Design, synthesis and biological evaluation of new bifunctional chemical degrader molecules (PROTACs) targeting hypoxia signalling pathway

Chiara Maniaci

Ph.D. Thesis

Supervisors: Professor Roberto Romeo

Professor Alessio Ciulli

Table of contents page

List of figures

List of schemes

List of abbreviations

ADME	absorption, distribution, metabolism, and excretion
AML	acute myeloid leukemia
ARNT	aryl hydrocarbon receptor nuclear translocator
BAIB	bis-acetoxy iodobenzene
BET	bromo-and extraterminal domain protein
bHLH/PAS	basic helix-loop-helix/Per-ARNT-Sim homology
BTB	bric-a-brac/tramtrack/broad complex
CBP	CREB binding protein
ccRCC	clear cell renal cell carcinoma
CDI	1,1'-carbonyldiimidazole
CRBN	cereblon
DCM	dichloromethane
DDB1	damage-specific DNA-binding protein 1
DIPEA	*N,N*-Diisopropyethylamine
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
EloBC	ElonginB-ElonginC complex
FIH	Factor Inhibiting HIF
GFP	green fluorescent protein
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate
HECT	homologous to E6AP carboxyl terminus
HIF	hypoxia inducible factor
HOAT	1-Hydroxy-7-azabenzotriazole
HRE	hypoxia-responsive element
Hyp	hydroxyproline
ITC	isothermal titration calorimetry
LHS	left-hand side
NHS	*N*-hydroxysuccinimide
PEG	polyethylene glycol

PPI	protein-protein interaction
PROTAC	proteolysis targeting chimera
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
RBR	RING-between-RING
RHS	right hand-side
RING	really interesting new gene
SAR	structure-activity relationship
SEC	size exclusion chromatography
SNIPER	specific and non-genetic inhibitor of apoptosis proteins (IAPs)-dependent protein eraser
TEA	triethylamine
TEMPO	2,2,6,6-Tetramethyl-1-piperidinyloxyl or (2,2,6,6-tetramethylpiperidin-1-yl)oxidanyl
TFA	trifluoroacetic acid
UPS	ubiquitin proteasome system
VHL	von Hippel-Lindau
PBS	phosphate buffer saline
BSA	Bovine serum albumin
TBS	Tris-buffered saline
ECL	enhanced chemiluminescence
TCEP	tris(2-carboxyethyl)phosphine
DMSO	dimethylsulfoxide
VBC	VHL-ElonginB-ElonginC
DTT	1,4-dithiothreitol

Declaration

I certify that this thesis is of my own writing and the work reported is based upon the results of my own work carried out under the supervision of Professor Alessio Ciulli, Professor Dario Alessi at University of Dundee and Roberto Romeo at University of Messina. Work other than my own has been specifically stated within the text by referencing relevant researchers and/or their relevant publications. No part of this thesis has previously been submitted for higher degree.

Abstract

Most synthetic small molecules function as inhibitors to modulate biological processes without affecting target level inside cells. Recent studies have demonstrated that inhibitors targeting more than a single protein i.e. pan-selective, can be turned into more selective and more effective chemical probes for individual target proteins, by using PROTACs (proteolysis targeting chimera). Using this approach, the ubiquitin proteasome pathway can be hijacked to induce degradation of the target protein as opposed to its simple blockade.

PROTACs are usually hetero-bifunctional compounds containing a ligand for an E3 Ubiquitin Ligase, and a ligand for the target protein, tethered by a linker. By bringing the two proteins into proximity, PROTACs can induce ubiquitination and subsequent degradation of the target protein through the proteasome. One E3 ligase successfully targeted by our laboratory is the von Hippel-Lindau protein (VHL), which as part of Cullin2 complex naturally functions in the oxygen signalling pathway. The work described in this thesis aims to target two of the major players involved in this pathway – VHL itself, and the prolyl hydroxylase enzymes (PHDs) that catalyze the key hydroxylation step that triggers the series of event that regulate the sensing of oxygen.

The crystal structures of PHD2 and VHL proteins with their respective inhibitor bound provided the start point for the PROTAC design, allowing to identify suitable positions for chemical derivatization. Two series of PROTACs against PHD (Hetero-PROTACs) and VHL (Homo-PROTACs) recruiting both VHL and a related E3 ligase (cereblon) were designed accordingly and synthetized. The biological evaluation of the compounds ability to degrade their target proteins was performed by in vitro treatments of different cancer cells lines and immunoblotting. Homo-PROTAC CM11 induced rapid and isoform specific degradation of VHL in the double-digit nanomolar concentrations. Degradation was shown to be dependent upon proteasomal activity, cullin neddylation and VHL binding. Formation of ternary complex was assessed by biophysical studies.

Together, this study provide first proof-of-concept for developing the PROTAC approach to trigger an E3 ligase to induce its own degradation, with wide implications for chemical biology and potentially drug discovery.

1 Introduction

1.1 Targeting protein degradation

Cellular proteins function in a dynamic state, being continuously synthesized and degraded[1]. This phenomenon, called proteostasis (a fusion of words "protein" and "homeostasis") is characterized by the integrated action of biological pathways that control the biogenesis, folding, trafficking and degradation of proteins present within and outside cell[2]. In particular, protein degradation plays a central role in many cellular functions. First and foremost, it helps to remove misfolded or damaged proteins from the cell to avoid toxicity. Secondly, it contributes to regulate protein concentration at the appropriate times and compartments in the life of a cell. Moreover, all proteins targeted for degradation, foreign or native, are digested into small peptides as part of the adaptive immune response[3].

Work over the past two decades has convincingly demonstrated that the main "system" responsible for the majority of degradation of soluble cellular proteins under most conditions is the ubiquitin-proteasome pathway (described in detail in section 1.1.1). Although lysosomes/vacuoles are important degradative organelles, particularly under stress conditions, most of the proteolysis of cytosolic proteins that occurs in lysosomes is relatively nonspecific[4]. Such proteins generally gain access to the hydrolytic enzymes within the lysosome by autophagic processes that envelop portions of the cytoplasm[5] There is some evidence for more specific uptake mechanisms, as well, but these likely account for only a small percentage of the specific proteolysis that occurs *in vivo*. On the other hand, many membrane proteins may be targeted to the vacuolar/lysosomal compartment by ubiquitin-dependent endocytic routes. Additional proteolytic enzymes are present within cells, some of which are attracting much attention e.g. the ICE-like proteases crucial for apoptosis and the calpains. However, the existing evidence suggests a more circumscribed role for this class of enzymes, limited either in the range of substrates attacked or in the

range of conditions under which they operate at rates that are physiologically significant[6].

Selective modulation of protein turnover frequently functions in concert with other control mechanisms[7]. It also offers several advantages over other types of genetic regulatory mechanisms. One advantage is from a kinetic standpoint. Affecting a protein at its post-translational level can be much faster than at the DNA or RNA level, because a new steady-state level can be reached quickly following the perturbation. A second advantage is irreversibility. Elimination of a protein's molecule removes any possibility of it being reactivated inappropriately. These amongst other features help to explain why protein degradation is almost always a component of regulatory mechanisms[8]. Examples of such mechanisms include cell cycle progression and signaling transduction pathways. Sensitive and rapid regulation of protein levels comes however at a cost – that is the relatively high energy required for resynthesizing the destroyed proteins. However, not all possible proteins encoded by the genome are expressed at any one time, and only a small fraction, primarily key regulatory proteins, undergo rapid and continuous turnover in the cell[6].

1.1.1 Overview of the ubiquitin proteasome system (UPS)

As previously pointed out, the ubiquitin-proteasome pathway is the primary degradation machinery of the cell[9]. The initial and often key factor that determines the post-translational faith of a particular cellular protein, notably its degradation by the proteasome, is the covalent modification of that protein by ubiquitin. Indeed, the discovery of ubiquitin-mediated protein degradation was recognized by the Nobel Prize in Chemistry in 2004, awarded to Aaron Ciechanover, Avram Hershko and Irwin Rose. Ubiquitin (Ub) is a 76-amino-acid-residue protein that is highly abundant and ubiquitously expressed[10]. Protein modification by ubiquitination requires the sequential action of three enzymes, E1, E2 and E3 (Figure 1). First, ubiquitin is chemically activated in an ATP-dependent manner by an E1-activating enzyme, forming a thioester bond

between the C-terminal glycine residue of ubiquitin and a conserved catalytic cysteine of the E1.

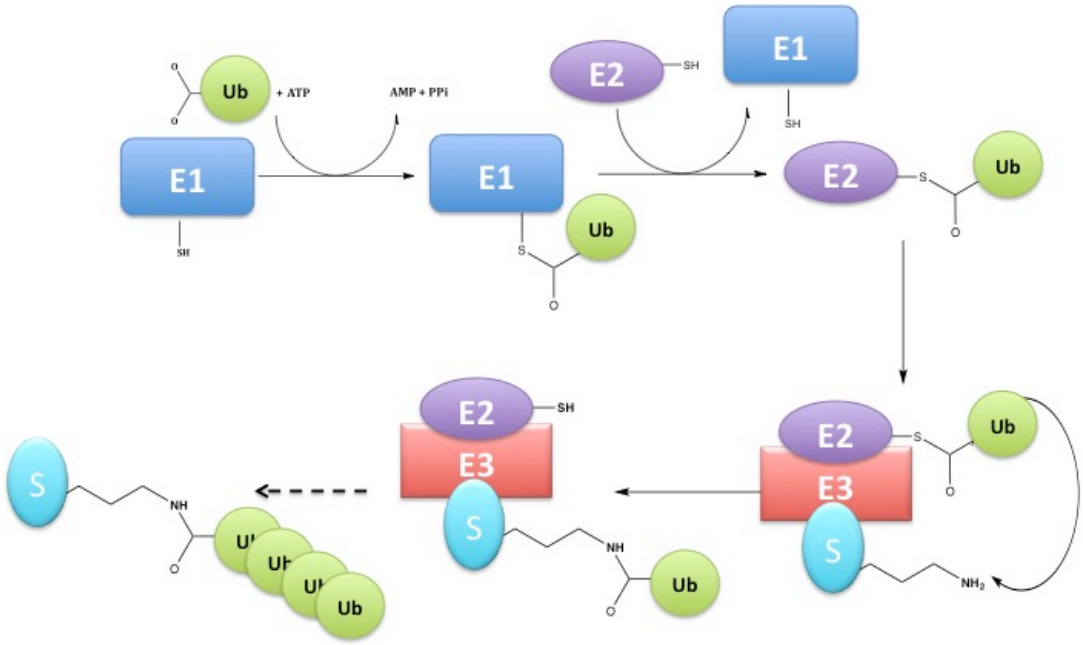


Figure 1. Schematic mechanism of the E1-E2-E3 protein ubiquitylation cascade.

Activated ubiquitin is next transferred on to the active site cysteine of an E2-conjugating enzyme, via a *trans*-thiolation reaction. Finally, an amide isopeptide bond between the ε-amino group of a lysine residue of a substrate protein and the C-terminal glycine residue of ubiquitin is formed via E3 ligase-mediated catalysis. Further Ub molecules are covalently linked to the previously attached Ub molecules in order to form poly-Ub chains. This process can occur at any of the several Lys residues present in the Ub molecule[11], however for protein degradation this predominantly occurs to Lys48, thus generating so-called K48-linked poly-Ub chains[12]. The consequence of this three-step conjugating cascade is that the resulting product, a poly-ubiquitinated substrate protein, can be recognized and ultimately degraded by the 26S proteasome[13]. Depending on the extent and chemical nature of the poly-Ub chain, a wide range of downstream cell signalling responses can be triggered in addition to the degradation of the substrate protein[14]. Non-degrading Ub signaling is beyond the scope of this thesis and will thus not be covered here.

1.1.2. E3 ubiquitin ligases

The human genome encodes two E1-activating enzymes (Uba1, also called Ube1, and Uba6), around 40 E2-conjugating enzymes and over 600 E3 ubiquitin ligases[11], that are all specific for ubiquitination. In the laboratory of Professor Ciulli we are particularly interested in understanding and targeting E3 ubiquitin ligases because many E3s are multi-subunit complexes that function upon multiple protein-protein interactions, and because they impart substrate specificity to the ubiquitin-proteasome pathway.

E3s are the most heterogeneous class of enzymes in the ubiquitination pathway as they mediate substrate specificity[15]. Currently, E3 ligases can be classified in three main classes depending on the presence of characteristic subunits or domains and on the detailed mechanism of ubiquitin transfer to the substrate protein: the homologous to E6AP carboxyl terminus (HECTs), the RING-between-RING (RBRs) and the really interesting new gene (RINGs) families[16].

1.1.2.1. HECT E3 ligases

The E3 ligases of the HECT domain family catalyze ubiquitin transfer to the substrate protein through a two-step reaction: ubiquitin is first transferred to a specific cysteine on the E3, and then from the E3 to the substrate[17]. The conserved HECT domain is located at the C terminus of the protein and is characterized by two lobes: the N-terminal lobe that interacts with the ubiquitin-charged E2, and the C-terminal lobe that contains the catalytic cysteine. A flexible hinge that modulates the relative orientation of the lobes during ubiquitin transfer tethers the two lobes. While the C-terminal HECT domain is involved in catalysis, substrate specificity is determined by the N-terminal domain. Based on their N-terminal extensions, human HECTs can be classified into three subfamilies: Nedd4, HERC (HECT and RCC1-like domain), and “other” HECTs that contain various domains. The catalytic activity of HECT E3s is often regulated by intramolecular interactions that keep the protein in an auto-inhibited state that is released in response to signals[18].

1.1.2.2. RBR E3 ligases

Similar to HECT E3s, RBR E3s catalyze ubiquitin transfer through a two-step reaction where ubiquitin is first transferred to a catalytic cysteine on the E3 and then to the substrate. The RBR name derives from the presence of two predicted RING domains (RING1 and RING2) separated by an in-between-RING (IBR) domain[19]. RING1 recruits the ubiquitin-charged E2, while RING2 domain possesses the catalytic cysteine; however, RING2 does not conform to the canonical RING E3 structure, and therefore it has been also called Rcat (required-for-catalysis) domain. The IBR domain adopts the same fold as the RING2 (or Rcat) domain while lacking the catalytic cysteine residue; therefore, it has been also called a BRcat (benign-catalytic) domain. RBR E3 ligases contain additional domains that are specific to each member. Several domains are involved in intramolecular interactions that keep the protein in an auto-inhibited state that can be released through various mechanisms[20].

1.1.2.3. RING E3 ligases

RING E3s are the most abundant types of ubiquitin ligases[21]. They are characterized by the presence of a zinc-binding domain called RING or by a U-box domain, which adopts the same RING fold but does not contain zinc metal ions in the domain structure. The RING and U-box domains are responsible for binding the ubiquitin-charged E2 and stimulating ubiquitin transfer.

RING E3s mediate a direct transfer of ubiquitin to the substrate, functioning as scaffolding enzymes to correctly orient the ubiquitin-charged E2 with respect to the substrate protein and juxtapose the reactive electrophilic and nucleophilic centers, respectively[12]. RING E3s can function as monomers, homodimers, or heterodimers. Homodimeric RINGs can normally bind two E2s (one per each monomer); this does not appear to be the case for heterodimeric RINGs. Similarly, U-box domains can function as monomers or homodimers[22].

Many RING E3s are composed by multiple subunits. This is the case for the cullin-RING ligases (CRLs), a highly diverse class of ubiquitin ligases characterized by several common features[23]. From a functional standpoint, catalytic activity of CRLs primarily leads to the poly-ubiquitination and ultimately proteasomal degradation of protein substrates[24]. CRLs are composed

of several subunits that are assembled around a cullin[25], that serves as a central scaffold subunit. The cullin subunit binds a RING-box protein (Rbx1 or Rbx2) at its C-terminus, while adaptor subunit(s) and a substrate receptor subunit are bound at the N-terminus of the cullin[23]. It is the latter end that is primarily responsible for substrate recognition and specificity. In contrast, the C-terminal end and the Rbx subunit are responsible for the docking of the charged E2-Ub subunit. Another important multi-subunit E3 is the anaphase-promoting complex/cyclosome (APC/C), a large assembly of 19 subunits that includes a RING subunit (Apc11) and a cullin-like subunit (Apc2)[26,27].
For the purpose of this thesis, we focus on CRLs as targets for small molecules.

1.1.3. Cullin RING Ligases (CRLs)

CRLs comprise over 200 members, making it the largest subfamily of all E3s. In some cell types, up to 20% of proteasome-dependent degradation of the proteome is mediated by CRL catalysis[28]. CRLs use modular subunit organization by utilizing interchangeable adaptors, substrate receptors, Cullin scaffolds and RING-box domain proteins to enable assembly of a large number of functionally diverse CRL complexes (Figure 2)[24].

1.1.3.1. The human cullin family

Human cells express seven different cullins that share similar structural architecture[25]. These are termed Cul1, 2, 3, 4A, 4B, 5 and 7. The classification of CRLs is based on the type of Cullin protein in the complex. For example, a cullin-RING ligase containing Cul5 is referred to as CRL5[29].
All cullin proteins contain two distinct domains: an N-terminal domain (NTD) and a C-terminal domain (CTD). The Cullin NTD is an elongated domain consisting of three repeats of a five-helix bundle – so-called “Cullin repeats” and often referred to as (CR)1–3. The NTD recruits different substrate receptors either directly or via one or more adaptor subunits. The interactions between the adaptor–receptor complex and the Cullin NTD tend to be very tight and constitute an important structural element for the stability of the CRL assembly[30]. The conserved globular CTD of Cullins serve instead as the docking

site for RING-box proteins, such as Rbx1 or Rbx2. These proteins recruit a cognate ubiquitin-loaded E2 enzyme and function to promote the discharge of ubiquitin from the E2 directly on to a substrate[31].

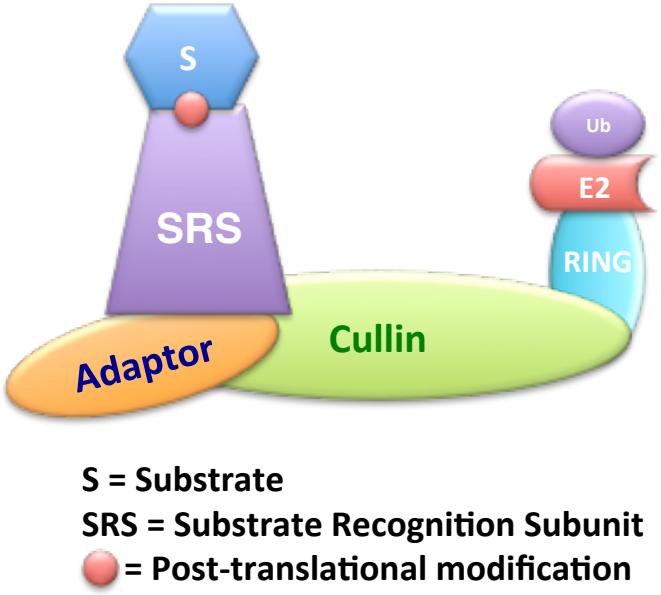


Figure 2. Schematic representation of the multisubunit assembly of cullin RING ligases.

1.1.3.2. Adaptor proteins

Most CRLs utilize adaptor subunits to bridge the central cullin to the substrate-recognition subunit[23,32]. Examples of adaptor proteins utilized in CRL assembly include Skp1, the binary complex between Elongin B and Elongin C (EloBC), bric-a-brac/tramtrack/broad complex (BTB) and DDB1 (damage-specific DNA-binding protein 1). Notably, Skp1, Elongin C and BTB display a high level of structural homology with each other and share a common fold that is often termed the Skp1/BTB/POZ fold[33]. In contrast, the DDB1 subunit differs from other adaptors in that it is considerably larger in size and consists of multiple distinctive β-propeller motifs[34].

1.1.3.3. Substrate specificity receptors

Specificity of CRL activity is determined by substrate-specific receptors. These subunits often recognize their cognate substrates upon recognition of specific post-translational modifications in the substrates, including phosphorylation and hydroxylation. Each receptor subunit can potentially bind multiple different

substrates, therefore expanding the functional range of CRL activities. Substrate receptors either consist of an individual subunit, which typically recruits the adaptor subunit(s) via a conserved short helical region called "box". Examples of such regions are the F-box, the VHL (von Hippel-Lindau)/SOCS-box or H-box domain[35]. Alternatively, substrate receptor can be merged with the adaptor in a single polypeptide chain, as in the case of the BTB proteins[36].

1.1.3.4. Mechanism and regulation of protein ubiquitination

Despite the importance of CRLs in cell signaling, detailed mechanisms underlining their catalytic activity and regulation remain elusive. The poly-ubiquitination catalytic mechanism of CRLS is complex, highly regulated and highly dynamic and it is still far from being completely understood. Indeed, it is a subject of intense investigation by biochemists and structural biologists alike[22,37-39].

Post-translational modifications and the diversity of CRL receptor subunits cooperate to specify which substrate is targeted for ubiquitination and when and where in the cell ubiquitylation occurs[28]. In addition, CRL activity is modulated by other regulatory control mechanisms, like neddylation and interaction with other partner proteins[39].

All cullins known to date can be modified by the covalent attachment of the ubiquitin-like protein NEDD8 to a conserved Lys residue in the cullin CTD that leads to a significant conformational change in the cullin CTD and RING-box subunit[31]. Like ubiquitination, neddylation is a reversible process. Deneddylation of CRLs is catalyzed by the COP9 signalosome (or CSN). After deneddylation, CAND1, a CRL exchange factor, binds to the de-neddylated CRL complexes to block the binding sites of adaptor proteins and Nedd8. This interaction keeps the CRL in a rigid off-state[39].

There is no direct experimental evidence to suggest that inside cells the transfer of the first Ub molecule is catalyzed by the CRL alone. Indeed, a recent study suggests that most CRL enzymes are poor at adding the first ubiquitin onto their substrate proteins[38]. Instead, a protein called ARIH, a member of the RBR E3s, is required with its own dedicated E2 to perform this job, before it hands over to

another E2 enzyme to catalyze with the CRL the subsequent poly-ubiquitination of the substrate[38].

1.1.3.5. Protein deubiquitination

Protein ubiquitination is a reversible process and the isopeptide bond can be hydrolyzed by protease enzymes called deubiquitinases (DUBs), of which over 80 have been identified in the human genome[40]. DUBs chop down poly-Ub chains on substrate proteins. When these are degradative chains, DUB activity opposes and counteracts E3s and proteasomal activity.

1.1.3.6. The 26S proteasome

The proteasome is a large cylindrical particle consisting of at least 33 subunits, with a total molecular weight of approximately 2.5 MDa[41]. There are several variants of the proteasome that perform slightly different functions. The 26S proteasome is the one responsible for the specific degradation of regulatory proteins as well as the removal of damaged proteins. It is composed of a central 20S core particle, capped by a 19S regulatory particle at one or both ends. The 20S core particle is a stack of four heptameric rings, that assemble to form a cylindrical structure[41]. The outer two rings are composed of α subunits, while the inner two rings are made of β subunits, which contain the proteolytic active sites in a central cavity. The degradation chamber can be reached through a channel that runs along a long axis at the center of the core particle. The entrance to the channel is narrow, therefore folded proteins must be at least partially unfolded before they can be threaded into the 20S core particle, cleaved and ultimately degraded. The 19S regulatory particle is composed of at least 19 subunits arranged into two sub-complexes – one called “the lid” and another called “the base”. The regulatory particle contains ATPase subunits, which catalyze ATP hydrolysis in order to energetically support its functions. These include gating the entrance to the degradation channel, mediating substrate recognition, deubiquitination, unfolding, and ultimately translocation into the 20S core particle[42].

Ubiquitinated proteins are delivered to the proteasome by various routes. Some substrates bind directly to the proteasome by interacting with specific subunits

in the 19S regulatory particle. Alternatively, ubiquitinated substrates can be brought to the proteasome by adaptors that bind both the proteasome and the ubiquitin chain on the substrate to deliver it for degradation[3]. Ubiquitin is usually recycled after targeting of the ubiquitinated protein to the proteasome, indicating that the proteasome contains or can recruit deubiquitinating enzymes. Short peptides are the products of proteasomal degradation. [6]

1.2 Monofunctional small molecules targeting the UPS

There is significant interest in targeting the ubiquitin proteasome system (UPS) with small molecules for the treatment of diseases such as cancer and neurodegeneration[43]. The UPS was initially targeted at the proteasome level using proteasome inhibitors of peptidic or peptide-like nature, such as the boronic acid bortezomib and the epoxyketone carfilzomib (described later). More recently, other molecules targeting key steps of the ubiquitination cascade upstream of the proteasome have been developed. An overview of the most representative compounds (Figure 3) is given next.

1.2.1. E1 inhibitor: MLN4924

MLN4924 is a selective inhibitor of NEDD8-activating enzyme (NAE), which is homologous to the E1 enzyme Ube1[44]. MLN4924 is currently in clinical trials for the treatment of solid tumors and hematological cancer e.g. relapsed/refractory multiple myeloma or lymphoma, and metastatic melanoma[45,46]. The anti-cancer therapeutic window of MLN4924 is believed to arise as a result of heavy reliance of cancer cells to the ubiquitin-proteasome pathway, in much the same way as in the case of proteasome inhibitors. MLN4924 creates a covalent NEDD8-MLN4924 adduct catalyzed by the enzyme[44]. The NEDD8-MLN4924 adduct mimics the structure of NEDD8-adenylate, the first intermediate in the NAE reaction cycle, but cannot be further utilized in the subsequent catalytic steps. The stability of the NEDD8-MLN4924 adduct within the NAE active site blocks enzyme activity, thereby accounting for the potent inhibition of the neddylation

pathway by MLN4924[47]. As a result, MLN4924 suppresses overall CRL activity in cells.

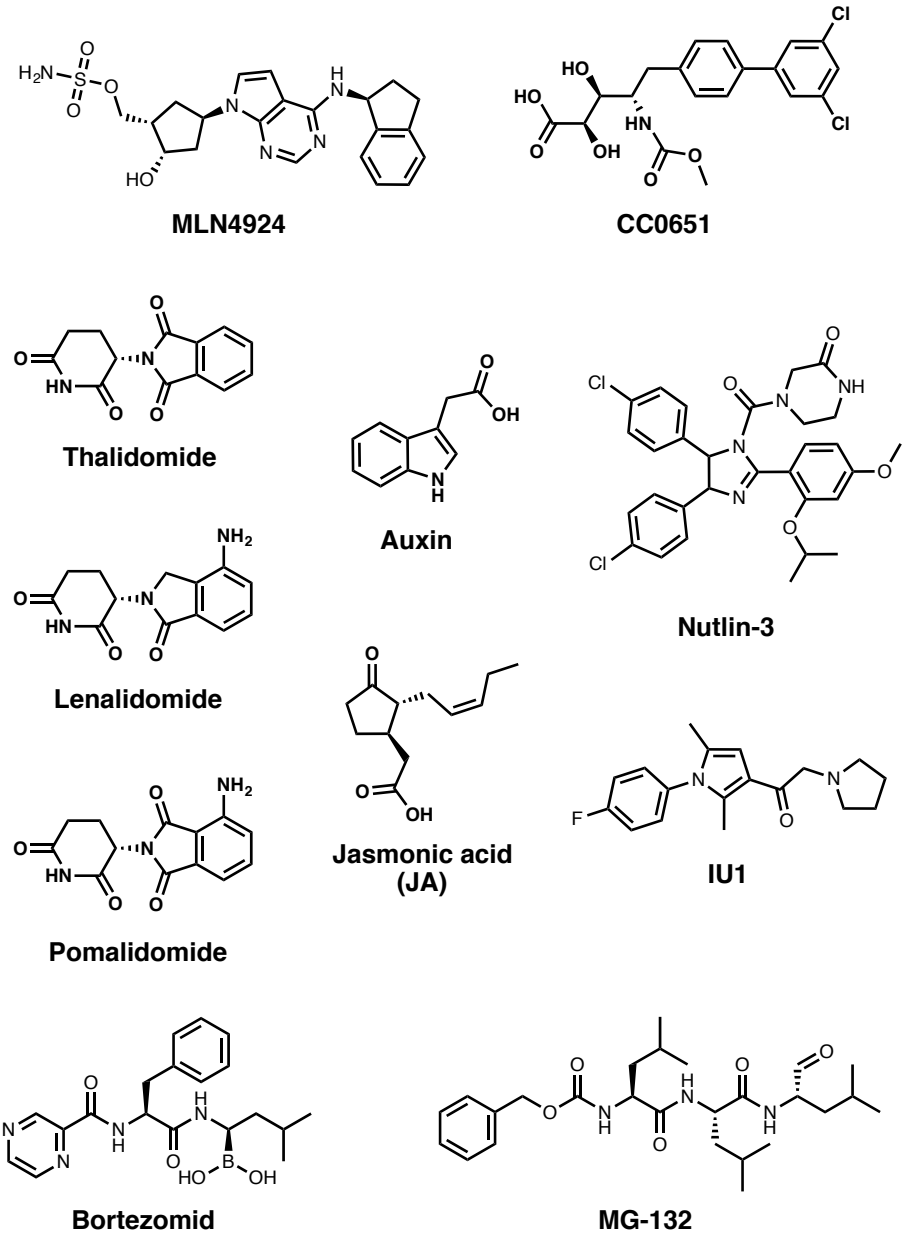


Figure 3. Chemical structures of monovalent molecules targeting the UPS.

1.2.2. E2 inhibitor: CC0651

The E2 enzyme hCdc34 (also known as UBE2R1) catalyzes the ubiquitination of hundreds of proteins in conjunction with the CRLs. Selective inhibitor CC0651 inserts into a cryptic binding pocket on the surface of hCdc34, distant from the catalytic site, causing the spatial rearrangement of E2 secondary structural

elements[48]. CC0651 does not affect hCdc34 interactions with E1 or E3 enzymes or the formation of the ubiquitin thioester but instead abrogates the initial transfer of ubiquitin to the substrate, thereby inhibiting the assembly of longer ubiquitin chains[49]. The discovery of CC0651 provides an example that E2 enzymes are susceptible to non-catalytic site inhibition[50]. CC0651 analogs inhibited proliferation of human cancer cell lines and caused accumulation of the key SCF^{Skp2} substrate p27Kip1.[48].

1.2.3. E3 inhibitor targeting the MDM2-p53 PPI: Nutlin

The E3 ligase MDM2 binds the tumor suppressor protein p53 with high-affinity and is responsible to the poly-ubiquitination and subsequent degradation of p53. Therefore MDM2 negatively regulates the transcriptional activity and stability of p53. Overexpression of MDM2, found in many human tumors, effectively impairs p53 function. Inhibition of the MDM2-p53 interaction can stabilize p53 and may offer a novel strategy for cancer therapy. The compound Nutlin, discovered by scientists at Roche, blocks the MDM2-p53 interaction, thereby preventing p53 from binding to MDM2 and in doing so it induces p53 stabilization. This inhibition mechanism led to anti-tumoral effects [51].

1.2.4. E3 inhibitors: Molecular "glues"

Compounds so-called "molecular glues" are a class of mono-functional molecules that bind to an E3 ligase and *de novo* "hijack" E3 activity toward a new substrate. By stabilizing protein-protein interactions between the E3 and its substrate, this mechanism can also lead to an increase in affinity of the E3 for the natural substrate that may have been weakened e.g. as a result of mutations. Two are the major examples known: plant hormones and the immunomodulatory drugs (IMiDs).

1.2.4.1. Plant hormones: auxin & jasmonate

Auxin is a gene-regulatory plant hormone that is crucial for $CRL1^{TIR1}$-mediated ubiquitination and degradation of substrate Aux (auxin)/IAA (indole-3-acetic

acid) transcriptional repressor proteins. Structural studies revealed that auxin binds in the substrate-binding pocket of TIR1 and promotes the productive recruitment of Aux/IAA[52]. Substrate Aux/IAA binds to the leucine-rich repeat domain of TIR1 in an auxin-mediated manner. This observation suggested that auxin analogues could potentially disrupt the TIR1-substrate interaction. Towards this aim, a series of alkylated derivatives of a natural auxin, IAA, was developed with the aim to modulate E3 ligase function and potentially prevent the substrate from ubiquitination[53].

A small-molecule-mediated interaction between a CRL and its substrate is not unique to the F-box protein TIR1. A similar mechanism was also identified for its close homologue COI1 (coronatine-insensitive protein 1) and the jasmonate class of plant signalling molecules[54]. JAZ1 (jasmonate/ZIM domain protein 1, where ZIM is zinc finger expressed in inflorescence meristem), is the substrate of $CRL1^{COI1}$. Crystal structures showed the plant hormone jasmonoyl-isoleucine (JA-Ile) to bind to the F-box protein COI1, promoting the interaction between COI1 and substrate JAZ1, as the mechanism leading to $CRL1^{COI1}$-dependent JAZ1 ubiquitination and proteasomal degradation[55]. No inhibitors of the COI1–JAZ1 interaction have been reported to date.

1.2.4.2. Immunomodulatory drugs: thalidomide and analogues

The immunomodulatory drug (IMiD) thalidomide was found to exert its infamous teratogenic effects primarily through targeting cereblon (CRBN)[56]. CRBN interacts with the adaptor DDB1 and Cul4A to form the $CRL4A^{CRBN}$ complex[34]. Later, thalidomide and its two close analogues, lenalidomide and pomalidomide were also found to display their anti-cancer effect by targeting $CRL4A^{CRBN}$ via direct interaction with CRBN[57-59]. Such interaction with the E3 ligase resulted in promoted degradative ubiquitination of several members of the Ikaros transcription factors family, in a manner similar to auxin-mediated degradation of Aux/IAA, described above. All three ligands bound in the same pocket of CRBN and demonstrated similar binding modes [34,60] (Figure 20) .

1.2.5. Ubiquitin chain targeting inhibitor: Ubistatin

The ubiquitin chain itself provides another potential opportunity for pharmacological intervention in this important pathway. Compounds called Ubistatins blocked the binding of ubiquitinated substrates to the proteasome by targeting the ubiquitin-ubiquitin interface of Lys48-linked chains [61].

1.2.6. Proteasome inhibitors: bortezomib and MG132

Drug discovery in the UPS has seen significant progress in the last decade in part due to the clinical success of the first-marketed proteasome inhibitors bortezomib (Velcade®, Millennium Pharmaceuticals) and more recently carfilzomib (Kyprolis®, Onyx Pharmaceuticals). The U.S. Food and Drug Administration (FDA) have approved these compounds in recent years for the treatment of relapsed and/or refractory multiple myeloma and mantle cell lymphoma[62]. It is believed that the anti-cancer activity of proteasome inhibitors is brought about in large part by the superior reliance on protein turnover of certain cancer cells relative to normal cells. Although proteasome inhibitors are reasonably selective against cancer cells to induce apoptosis, they suppress proteasome-mediated degradation of many proteins in the cell, leading to high risk of toxic side effects. In addition, clinical evidence for growing resistance against proteasome inhibitors is beginning to emerge. For a review of proteasome inhibitors see refs. [63,64].

1.2.7. Inhibitors of deubiquitinating enzymes

USP14 is one of the three distinct DUBs that are associated with proteasomes (RPN11, UCHL5 and USP14). USP14 disassembles proteasome-bound poly-Ub chains from their distal tips. USP14 activity leads to suppression of substrate degradation by disengaging the substrates from the proteasomes. A small molecule inhibitor of USP14, IU1, has been described which was found to enhance proteasomal function in cells[65]. This compound binds specifically to the activated, proteasome bound, form of USP14. Increasing proteasomal-mediated degradation by IU1 or similar inhibitors may be of therapeutic benefit in diseases such as Alzheimer's disease or Parkinson’s disease[66]. The situation

appears to be quite complex, however, since USP14 appears to be essential for the maintenance of synaptic ubiquitin levels and the development of neuromuscular junctions.

1.3 Physiological response to hypoxia: fundamental concepts

In mammals, the development of multiple complex physiological systems are required to ensure adequate O_2 delivery to all cells and tissues under normal conditions. It is well established that all nucleated cells in the human body sense O_2 concentration and respond to reduced O_2 availability (hypoxia) that is either acute or chronic in its duration[67,68]. Adaptive cellular responses to acute changes in O_2 concentration i.e. rapid changes lasting to no more than minutes, principally occur as a result of alterations of pre-existing proteins, for example, involving changes in phosphorylation or redox states. In contrast, chronic changes in O_2 concentration i.e. lasting from minutes to hours or longer, principally occur as a result of alterations in gene expression. The transcriptional regulator hypoxia-inducible factor 1 (HIF-1), first discovered in the mid-90s, is an essential mediator of O_2 homeostasis[69,70]. HIF-1 plays key physiological roles during cellular development and in fetal and postnatal life in both physiological and pathophysiological processes such as tumour growth, ischaemia and tissue repair[71]. Although there are other mechanisms of hypoxic adaptation, including those acting on a shorter time-scale than the HIF system, the extent and pleiotropic effects of the HIF system have led to this transcription factor being elected as a master regulator of the hypoxic response[72].

1.3.1 The transcriptional regulator hypoxia-inducible factor 1

Active HIF transcription factors are comprised of an oxygen- regulated α-subunit and a constitutively stable β-subunit. In humans there are three HIF-α paralogs, of which HIF-1α and HIF-2α are best characterized; together with HIF-β (also known as ARNT for aryl hydrocarbon receptor nuclear translocator) they form active transcription factors termed HIF-1 and HIF-2, respectively. HIF-1 and HIF-

2 are closely related, but regulate distinct (and sometimes overlapping) sets of genes in response to hypoxia. The α- and β-subunits of HIF belong to the bHLH/PAS (basic helix-loop-helix/Per-ARNT-Sim homology) family of transcription factors. The bHLH and PAS domains mediate α–β dimerisation and DNA binding, while N- and C-terminal transcriptional activation domains (NAD and CAD, respectively) recruit co-activator proteins to form active transcriptional complexes on DNA. Several crystal structures of large multi-domain regions of HIF-1α, HIF-2α and HIF-β complexes were recently solved[73]. HIF α-subunits also contain oxygen-dependent degradation (ODD) domains. Hydroxylation of the ODDs at conserved proline residues renders them labile to proteasomal degradation under oxygenated conditions[74,75].

1.3.2 Regulation of HIF-1 activity

The HIF-1 and HIF-2 α-subunits have short half-life in normoxia due to their rapid turnover by the ubiquitin-proteasome system. The $CRL2^{VHL}$ E3 ubiquitin ligase complex is composed of Cullin 2, Rbx1, Elongin C, Elongin B and the von Hippel-Lindau tumour suppressor protein (pVHL)[76]. This multisubunit complex catalyses the poly-ubiquitination of lysine residues that target HIF-α for degradation by the proteasome[77,78]. pVHL is the substrate recognition component of this complex and binds directly to HIF-α; this protein-protein interaction is substantially enhanced (>1,000 fold increase in binding affinity) by the hydroxylation of two proline residues in HIF-α (P402 and P564 in HIF-1α)[79,80]. These prolines are located within the N- and C-terminal oxygen-dependent degradation domains (NODD and CODD, respectively), each of which can interact independently with pVHL[81]. A single hydroxylation at either site (NODD or CODD) is sufficient for the $CRL2^{VHL}$ E3 ubiquitin ligase complex to target HIF-α for degradation. HIF prolyl hydroxylation is catalyzed by a set of non-haem iron- and 2-oxoglutarate (2OG)-dependent prolyl-4-hydroxylase enzymes PHD1, PHD2 and PHD3, which are encoded by the genes Egl nine homolog *EGLN2*, *EGLN1*, and *EGLN3*, respectively[82]. The catalytic activity of these enzymes is suppressed under conditions of sub-optimal oxygen availability,

leading to decreased hydroxylation of HIF-α. Because HIF-α molecules that lack prolyl hydroxylation are not recognised by pVHL, this form of HIF-α accumulates under hypoxia, translocates in to the nucleus where it dimerizes with HIF-β subunit and it binds to hypoxia responsive elements (HREs) of HIF-target genes, and as such it functions as a transcription factor (Figure 4).

$CRL2^{VHL}$ and PHD enzymes will be described in more details in Section 1.5.

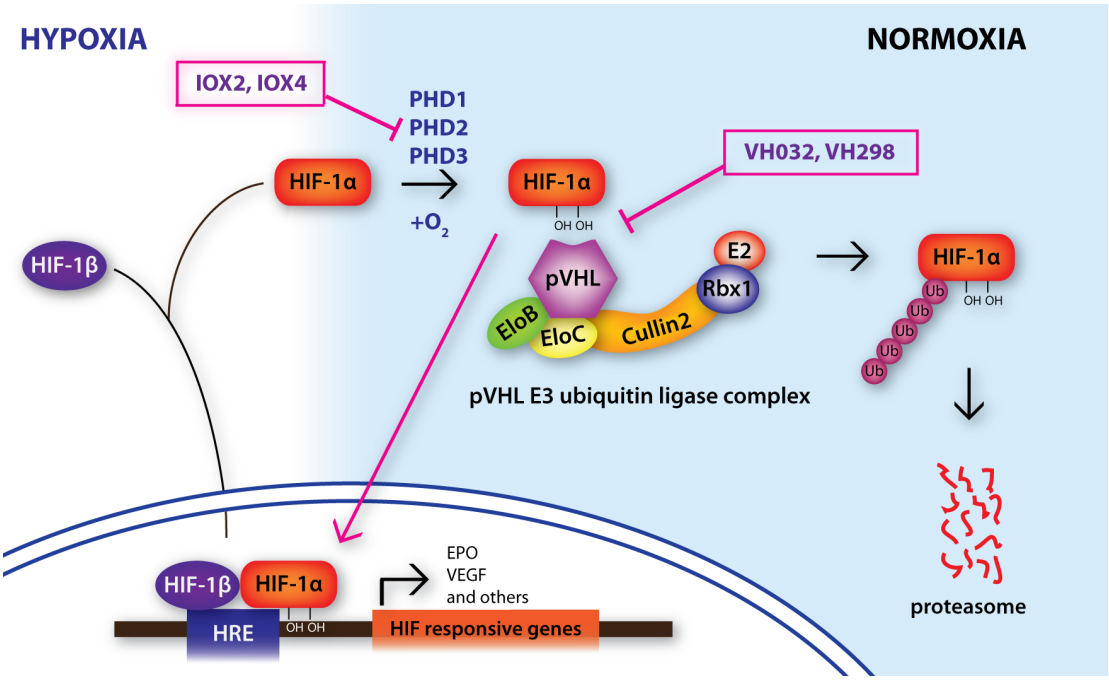


Figure 4. Schematic representation of the hypoxia-signalling pathway, showing the involvement of the key players HIFs, VHL, PHD enzymes and their key inhibitors.

A second mechanism of HIF regulation involves hydroxylation of a conserved asparagine residue (N803 in the HIF-1α CAD and N847 in HIF-2α) by the 2OG-dependent oxygenase Factor Inhibiting HIF (FIH)[83]. FIH is structurally and catalytically related to PHDs, as these enzymes are all members of the 2-oxoglutarate-dependent dioxygenase family[84]. Asparaginyl hydroxylation of the HIF-α CAD suppresses the interaction between HIF-α and the transcriptional co-activator proteins CREB binding protein (CBP) and p300, which are required for transcriptional activation of most HIF target genes. Like the PHDs, FIH is inactivated in hypoxia due to low levels of oxygen[82].

1.3.3 HIF-1 target genes

To activate transcription of target genes, during hypoxia conditions HIF-1α dimerizes with HIF-1β. The resulting-heterodimer binds to DNA at sites represented by the consensus sequence 5'-RCGTG-3' (the HRE). The HIF-1 binding site is present within a hypoxia response element, a *cis*-acting transcriptional regulatory sequence. It has been estimated that HIF activity regulates over 2% of the human genome, particularly genes related to oxygen sensing and the hypoxic response[85]. The target genes activated by HIF-1 include erythropoietin (*EPO*) and vascular endothelial growth factor (*VEGF*), which induce the production of red blood cells and blood vessels, respectively. Other HIF-target genes includes genes whose protein products are involved in energy metabolism, cell proliferation and viability, vascular remodeling, and vasomotor responses amongst other cellular processes[71,72].

1.3.4 HIF-1 is involved in the pathophysiology of many different diseases

Hypoxia plays an important role for tumor growth and invasion mechanisms[86,87]. Indeed, HIF-1α and HIF-2α are overexpressed in common human cancers[88]. This is due to the direct involvement of many HIF-1 and HIF-2 target genes in the growth of cancer mass[89]. Known HIF-1 target genes provide a molecular basis by which HIF-1 overexpression may promote key aspects of tumor progression. For example, glucose transporters e.g. GLUT1 and glycolytic enzymes promote metabolic adaptation to hypoxia; NOS2 and VEGF promote angiogenesis; IGF-2 promotes cancer cell survival and proliferation[89].

A key genetic lesion in cancer cells, i.e. hemangioblastoma and in clear cell renal cell carcinoma (ccRCC), is the functional inactivation of the *VHL* gene[90-92]. In renal carcinoma cell lines, VHL loss-of-function results in constitutive expression of HIF-1α and HIF-2α under non-hypoxic normal oxygen levels (also referred to as normoxic conditions) due to defective or lacking ubiquitination of HIF-1α.

In contrast to the roles of HIF in tumorigenesis, its pharmacological stabilization under normoxic conditions could provide therapeutic benefit for many conditions including anemia due to chronic kidney disease and anemia associated to cancer chemotherapy, ischemia and ischemic reperfusion injuries in the kidney, heart, brain, liver, or other organs, such as acute lung injuries and intestinal inflammation[93]. Recent studies have crucially identified VHL knockdown/knockout as protective during states of impaired mitochondrial function, supporting VHL and PHD inhibition as a potential treatment for human diseases associated with mitochondrial respiratory chain dysfunction[94].

1.4 Bifunctional molecules

Protein-protein interactions (PPIs) mediate almost all intracellular processes, so it is not surprising that modulation of these interactions using small molecules is one of the "holy grails" of pharmacology. However, because of the large, and often flat interfaces of many protein-protein complexes, it can be daunting to find or rationally design a small molecule that is able to bind to such PPI sites with the desired binding affinity and specificity, while retaining suitable pharmacological properties. As a result, there has been only limited success toward this goal, with some notable exceptions[95]. The opposite strategy, that of forming *de novo* or stabilizing PPIs with so-called interfacial inhibitors, has also proven challenging[96,97] On the other hand, a related approach that involves dimerization or oligomerization of proteins – that would not otherwise interact – mediated by small molecules has progressed further. This approach often relies on so-called chemical inducers of dimerization (CID)[98], that can be ascribed in general as bifunctional ligands[99].

A bifunctional molecule can be simply defined as a compound that interacts with two protein molecules simultaneously. Synthetic bifunctional molecules typically consist of two protein-binding moieties joined *via* a suitable linker of appropriate length[99]. If the two moieties are identical (homo-bifunctional), the molecule will dimerize two identical or highly homologous protein monomers.

By contrast, if the moieties differ (hetero-bifunctional), two different proteins can be recruited. The applications of these molecules can therefore be as broad as their design is diverse. For example, bringing two or more proteins together can trigger receptor signaling or activate transcription factors (as in the case of HIF). It can direct or cargo a protein in a different subcellular location, signal that protein for degradation, or induce the formation of a new binding surface recognized by other proteins e.g. via an allosteric mechanism.

1.4.1 Chemical inducers of dimerization

Molecules that function as CIDs are known in Nature. Amongst natural CIDs are complex natural products that have macrocyclic structures, including rapamycin and cyclosporine (Figure 5). Rapamycin simultaneously binds the FK506 binding protein FKBP12 and a specific domain of the protein mTOR (mammalian target of rapamycin). These interactions underline the immunosuppressive and cancer antiproliferative activities of rapamycin[100]. In a similar fashion, immunosuppressive small molecule drugs FK506 (also called tacrolimus) and cyclosporin induce the recruitment of the protein phosphatase calcineurin to form ternary complexes with the proteins FKBP12 and cyclophilin, respectively. The ultimate result of these induced dimerizations is that calcineurin activity is inhibited[101]. These discoveries were taken a step further by developing bifunctional synthetic ligands that can be used to control the intracellular oligomerization of specific proteins. One such ligand was designed as an artificial dimer of the ligand FK506, called FK1012, which promoted FKBP12 homodimerization[102]. By expressing each dimerizing FKBP12 protein as fusion constructs with other proteins of interest, these synthetic CIDs could be used to induce proximity or even interaction of proteins that would not normally interact, or that could not bind in a normal way any longer because they lacked key functional portions. As such, the approached provides small-molecule control or activation over numerous signaling pathways via proximity effects. Other variants of the CID approach include the so-called “yeast three-hybrid” systems for identifying protein targets of small molecules.

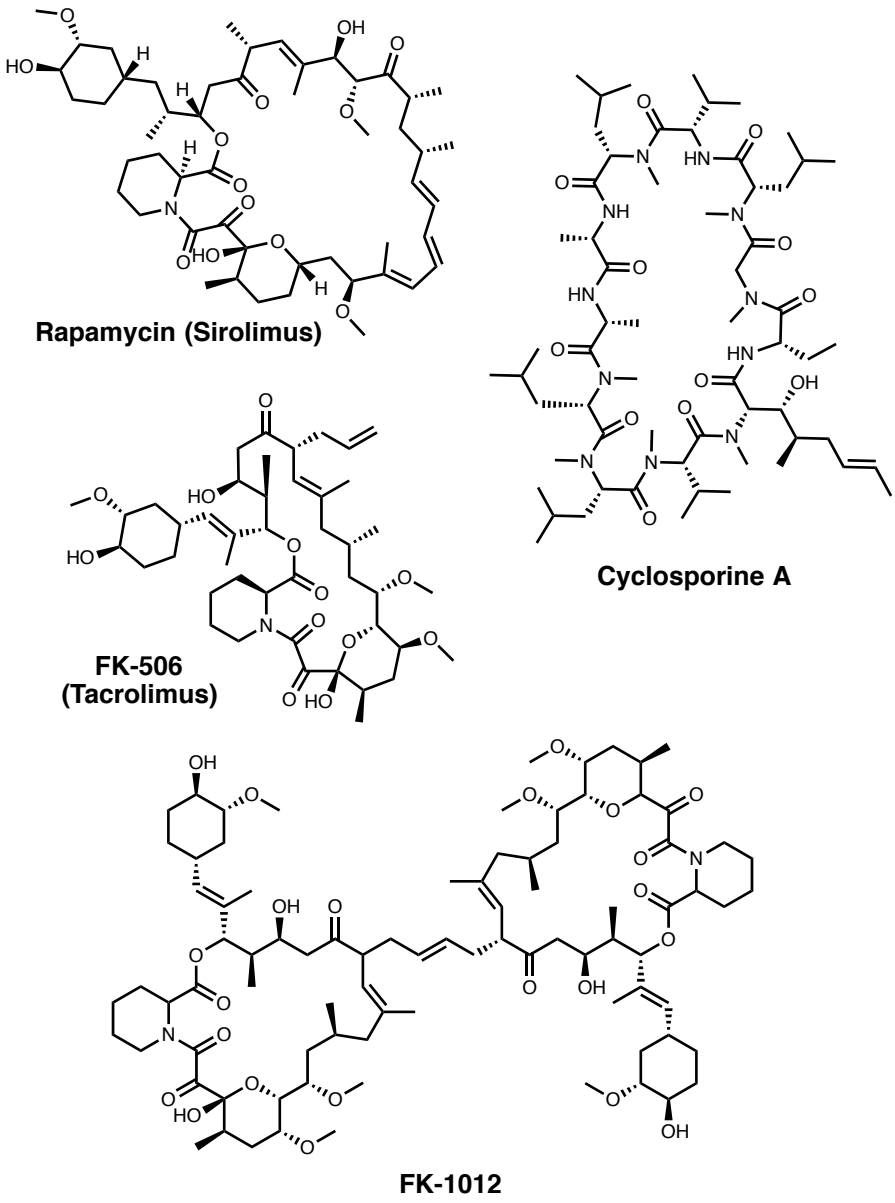


Figure 5. Chemical structures of natural and synthetic CIDs.

1.4.2 Bivalent inhibitors

In recent years, bi-functional synthetic ligands have been developed as bivalent inhibitors of their protein targets (Figure 6). Illendula and colleagues developed a bivalent protein-protein interaction inhibitor, termed AI-10-49, that binds to the transcription factor CBFβ-SMMHC (fusion of core binding factor b and smooth-muscle myosin heavy chain) and disrupts its binding to transcription factor RUNX1, which are key components that sustain acute myeloid leukemia (AML) growth. AI-10-49 was significantly more potent that the parent monovalent

compound, displayed favorable pharmacokinetics, and delayed leukemia progression in mice[103].

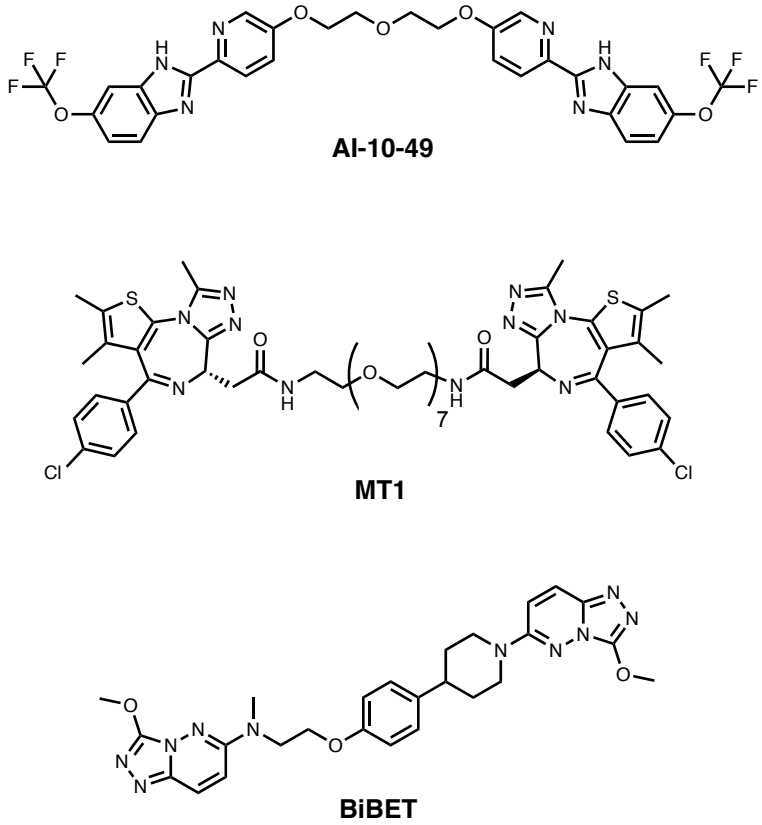


Figure 6. Chemical structures of representative bivalent inhibitors.

More recently, the laboratories of Bradner[104] and AstraZeneca[105] independently and at the same time reported bivalent inhibitors of the bromo and extraterminal domain (BET) proteins Brd2, Brd3 and Brd4. All BET proteins contain two separate bromodomains at their N-terminal region, termed BD1 and BD2. Because of their role in different forms of carcinomas and hematological malignancies, BET inhibitors have been developed with the aim of blocking BET protein function, prompting differentiation and specific anti-proliferative effects in BET-sensitive cell lines. Existing BET inhibitors, such as JQ1 (ref. [106]), bind to BET proteins monovalently. Given that a significant number of the 61 human bromodomain-containing proteins have more than one bromodomain within their overall structure (frequently appearing sequentially), a bi- or multi-valent approach to small-molecule inhibition stood as an attractive possibility for this

target class. Two different series of bivalent inhibitors of BET bromodomains were developed. Both MT1 (ref. [104]) and biBETs[105] were able to bind intramolecularly the two domains of a single BET protein. MT1 is a bifunctional compound in which two molecules of JQ1 are joined together by a PEG linker. MT1 was found to be 100-fold more potent than the corresponding monovalent ligand JQ1, at blocking Brd4 from binding to chromatin in cells. biBETs, initially developed to downregulate androgen receptor signaling, revealed also to be able to engage simultaneously to both BD1 and BD2 tandem domains of a BET protein. Cells treated with these bivalent compounds showed inhibition of cell growth in a manner consistent with sensitivity to Brd4 inhibition, and with a remarkable enhancement in potency, suggesting strong avidity effect.

1.4.3 Proteolysis-targeting chimeras (PROTACs)

Several strategies can be employed to knock down or decrease the expression level of disease-causing proteins. Conventional genetic knockdown techniques have been based on nucleic acids, for example antisense oligonucleotides, RNAi and siRNA, and more recently CRISPR-Cas9 (refs. [107,108]). All these approaches can abolish or decrease protein levels, however renowned issues of nucleic acids based in terms of their metabolic liabilities, poor bio-distribution, and frequent off-target effects have limited their therapeutic utility. A novel approach to address these challenges, which combines the favourable pharmacological "absorption, distribution, metabolism, and excretion" (ADME) properties of small molecules with the advantages of genetic knockdown techniques, is to control intracellular protein levels using small molecules. One approach to achieve this is by using so-called “chemical degrader” molecules to recruit target proteins to the UPS, thereby induce their selective degradation[43,109]. One prominent class of chemical degraders comprises hetero-bifunctional compounds that contain two recruiting ligands connected via a linker. One ligand is specific for a target protein of interest, while the other ligand specifically recruits an E3 ubiquitin ligase, for example a member of the CRL family of E3s (see Section 1.1.3). Because of their structural characteristics, PROTACs are able to bring in close proximity the target and the E3 ligase, thus facilitating substrate

poly-ubiquitination and subsequent proteasomal degradation. The resulting effect of PROTAC activity is the post-translational knock down of the targeted protein (for a schematic representation see Figure 13)[110].

1.4.3.1 First-generation peptidic PROTACs

First-generation PROTACs were based on a peptidic recognition motif of 7-amino acid from the HIF α-subunit ODD (ALA-Hyp-YIP) that binds to the VHL E3 ligase complex. These early compounds were used to target the estrogen receptor, the aryl hydrocarbon receptor and the FKBP12 protein[111]. Other all-peptidic PROTACs have since been reported. Amongst these are compounds that induce destruction of the X-protein of the hepatitis B virus (HBV) by fusing the ODD domain of HIF-1α with the X-protein oligomerization domain[112]. In all those cases, PROTACs required addition of a polyarginine cell-penetrating peptide in order to enter cells, increasing further molecular weight and limiting their synthetic accessibility and scope of chemical tools to probe biology.

Although the first-generation peptidic PROTACs had poor drug-like properties, they provided proof-of-concept, in that targeting E3 ligases could be exploited to modulate protein half-life, and as a viable strategy for potential drug development. The much-needed step forward in the development of PROTACs as a therapeutic strategy required conversion of the highly peptidic regions recruiting the E3 ligase into more drug-like compounds, thus allowing for the creation of all-small-molecule PROTACs. These would have several advantages over peptidic PROTACs, including better *in vivo* stability, improved bio-distribution, and potentially greater potency, leading to higher cellular activities.

1.4.3.2 Second-generation all-small-molecule PROTACs

The first reported all-small-molecule PROTACs were designed to target the E3 ligases cIAP1 – an approach termed SNIPER for "specific and non-genetic inhibitor of apoptosis proteins (IAPs)-dependent protein eraser" (ref. [113]) – and MDM2 (ref. [114]). These non-peptidic PROTACs were able to induce degradation of numerous proteins, however unfortunately they were only active in the double-

digit millimolar range of concentration. Therefore, there was a strong incentive to design better PROTACs in terms of potency and target selectivity.

After years of stepwise advances, the past couple of years have seen spectacular breakthroughs in the development of all-small-molecule PROTACs with high cellular efficacies and specificities. This has been made possible because of two low–molecular weight, specific, high-affinity E3 ligands, namely VHL inhibitor VH032 (ref. [115]) and cereblon-targeting ligand pomalidomide[34]. These new E3 ligands have enabled the assembly of potent and specific PROTACs able to degrade a range of targets[116-119]. Some of the recently reported PROTACs show more than 1,000-fold improvement in potency over earlier-generation compounds, with half-maximal effective concentration (EC_{50}) values ranging from <1 to 100 nM and half-lives for target degradation of only a few hours[120]. Most importantly, these new compounds are acting with a very high specificity.

In the following section, two types of PROTACs are described in detail: 1) PROTACs recruiting $CRL2^{VHL}$ based on the VHL ligand VH032; and 2) PROTACs recruiting $CRL4^{CRBN}$ based on the structure of pomalidomide.

1.4.3.3 PROTACs recruiting the E3 ubiquitin ligase VHL

The first category of high-performance PROTACs recruiting $CRL2^{VHL}$ to induce degradation of a target protein was reported by the Ciulli laboratory and used VH032 as the VHL recruiting moiety[116]. VH032 is a VHL ligand of relatively small size (MW = 472 Da) but very high binding affinity (K_d = 180 nM)[115]. Structure and chemical-physical properties of this ligand will be described later (Sections 1.6.2 and 2.1.1). In this study, the pan-selective BET inhibitor JQ1 (ref. [106]) was linked to VH032 (refs. [115,121]) to afford PROTAC compounds MZ1 (Figure 7) and MZ2, with a linker length of 3 and 4 ethylene glycol units, respectively[116]. The aim was to develop a PROTAC compound able to induce degradation of BET proteins as opposed to inhibiting them, as BET proteins are known to be critical for cancer cell growth and survival. Both compounds potently and rapidly showed to induce selective, reversible, and long-lasting degradation of BET

proteins in cells. The authors also demonstrated the proteasome- and VHL-dependency of target degradation in cells. Notably, it was found that Brd4 could be preferentially degraded at lower compound concentrations over others BET paralogs Brd2 and Brd3 (ref. [116]). This finding was surprising, because it is known that JQ1 binds all BET bromodomains with very similar binding affinities[122]. As MZ1 also bound to all BET bromodomains with similar affinities[116], it was not expected to be able to discriminate between the BET proteins. This study demonstrates how targeted degradation by small molecules can add a layer of selectivity beyond simple binary target engagement by its constitutive ligand.[116]

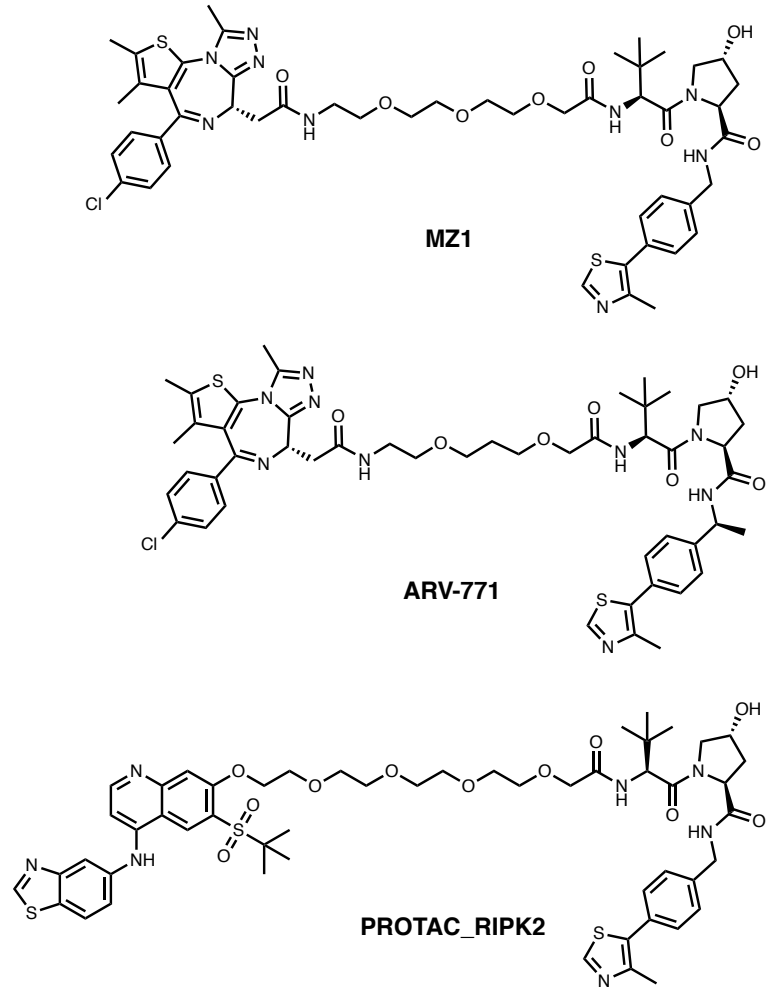


Figure 7. Chemical structures of bifunctional PROTACs designed to hijack VHL.

Another small-molecule PROTAC, reported later, that links together JQ1 to VH032 to target BET proteins is ARV-771 (Figure 7). ARV-771 differentiates from

the previously described MZ1 in minor structural elements, namely the use of a different linker and the addition of a methyl group in the VHL ligand. ARV-771 caused cell death in cultured prostate cancer cells by inducing degradation of BET proteins. This resulted in tumour growth inhibition or tumour regression in mouse models of castration-resistant prostate cancer, including models that express high levels of androgen receptor splice variant 7, the most important biomarker in this cancer type[123].

The same VHL ligand VH032 was linked to a thiazolidinedione-based selective ERRα ligand to form a PROTAC with the aim of target ERRα protein[119]. ERRα is an orphan nuclear hormone receptor that has been implicated as a master regulator of cellular energy homeostasis, regulating expression of genes involved in mitochondrial biogenesis, gluconeogenesis, oxidative phosphorylation and fatty acid metabolism. The VHL-based PROTAC was shown to induce 50% degradation of its target protein at 100 nM [119]. Importantly, this PROTAC proved to be active *in vivo*; treatment with the PROTAC reduced ERRα levels roughly by 40% in mouse heart, kidney, and xenografted tumors. Based on these promising results, another PROTAC was developed and published in the same study to target the serine/threonine kinase RIPK2, an important mediator of innate immune signaling. A RIPK2 inhibitor was connected via a 12-atom linker to VH032 to yield compound PROTAC_RIPK2 (Figure 7), which effectively mediated ubiquitination and subsequent proteasome degradation of RIPK2 with a maximal degradation ability (D_{max}) of 95% at concentrations as low as 10 nM. In addition, maximal RIPK2 knockdown with PROTAC_RIPK2 was achieved as early as 4 h, which compared favorably with the natural half-life of the protein, which is approximately 60 h.

VHL-based PROTACs were also developed for their ability to induce the degradation of intracellular green-fluorescent protein (GFP) fused to HaloTag7 (ref. [124]). Various linker lengths as well as two different linker attachment points were explored. The best molecule, HaloPROTAC3, was found to induce more than 90% GFP–Halotag7 degradation, with 50% degradation by 6 h. HaloPROTAC3

was found to induce complete knockdown of important kinases fused to HaloTag7 (HaloTag7–ERK1 and HaloTag7–MEK1) at 500 nM concentration.

1.4.3.4 PROTACs recruiting the E3 ubiquitin ligase CRBN

In recent years interest in thalidomide and its derivatives lenalidomide and pomalidomide has intensified as a result of their utility as potent immunomodulatory drugs (IMiDs). As mentioned previously (Section 1.2.4.2), recent studies have shown that lenalidomide binds to the E3 ligase cereblon, thereby inducing the degradation of IKZF1 and IKZF3, which are essential transcription factors in multiple myeloma[125,126]. This suggested that thalidomide and its derivatives could be used as starting ligands for new PROTACs targeting the E3 ligase CRBN.

Compound ARV825 (Figure 8) was developed by coupling BET inhibitor OTX015 (a close analogue of JQ1) to pomalidomide via a short alkyl linker, with the aim of induce degradation of BET proteins[118]. Unlike MZ1, ARV825 showed to induce pan-selective degradation of all BET paralogs Brd2, Brd3 and Brd4. The induced degradation was shown to be potent and complete within 6 h, and the effect was found to be long lasting (24 h). A similar PROTAC molecule was generated by coupling JQ1 to thalidomide[117]. The resulting PROTAC, called dBET1 (Figure 8), induced degradation of all BET proteins in the human AML cell line MV4-11 and functioned mechanistically as anticipated. The pharmacological properties of dBET1 were further evaluated in AML proliferation assays; interestingly, BET protein degradation suppressed growth in several lymphoma cells and was superior in efficacy to the parent inhibitor JQ1. In xenograft mouse model of human MV4-11 cells, it was shown that daily treatment with dBET1 led to tumor regression in 14 days[117].

The synthesis of PROTAC compounds that mediate the degradation of the protein kinases c-ABL and BCR-ABL by recruiting either VHL or CRBN was later reported[127]. This study showed how the two kinases could be degraded differentially by hijacking CRBN versus VHL, despite using the same kinase

inhibitor as scaffold. The results confirmed previous conclusions by Zengerle et al.[116], in that the ability of PROTACs to induce target degradation involves more than just target binding. The identity of the inhibitor warhead and the recruited E3 ligase, in addition to the nature and length of the linker, are all important factors in determining the degradation profiles of the compounds.

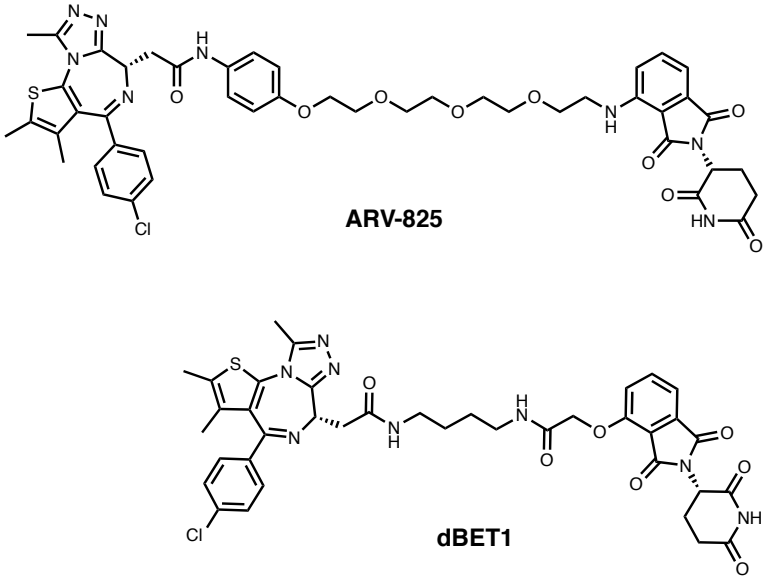


Figure 8. Chemical structures of bifunctional PROTACs designed to hijack CRBN.

1.5 PHDs and VHL as key HIF-alpha regulators: PPIs and mechanisms

As explained in Section 1.3, the HIF-α subunits are regulated by hydroxylation of their ODD domains catalysed by PHD enzymes under normal oxygen levels. These events lead to rapid turnover of HIF-α by $CRL2^{VHL}$ catalysed ubiquitination and subsequent proteasomal degradation. These observations provide solid foundation to understand how the PHD enzymes and VHL play a crucial role in the oxygen-signaling pathway. For these reasons, a closer look at the structure and functions of $CRL2^{VHL}$ and PHDs is given in the following sections.

1.5.1 VHL E3 ligase

1.5.1.1 Function of VHL and its different isoforms

As described above, pVHL is the substrate recognition component of an E3-ubiquitin ligase complex that assembles together with Elongins B and C, Cullin 2, and Rbx1 subunits. [79,128-130].

The gene encoding human pVHL protein is located on chromosome 3p25–26. The *VHL* gene includes three exons and it encodes two major isoforms of pVHL: a 213 amino-acid long, 30 kDa form (pVHL30) and a 160 amino-acid long, 19 kDa form (pVHL19). Although both isoforms are expressed in human cells, pVHL19 is the more prominent form in human tissues. pVHL19 lacks a 53 amino-acid long amino-terminal domain or N-terminal tail (pVHL-N), which is instead present in pVHL30. Interestingly VHL isoforms are also found located in different compartments of the cell. pVHL19 tends to be equally distributed in the nucleus and cytoplasm, whereas pVHL30 is found predominantly in the cytoplasm. This suggests that under certain circumstances the two isoforms can display distinct roles [131], and that the pVHL-N region could be responsible for the different subcellular localization. Regardless, functional studies have suggested that both isoforms have equivalent effects in different assays, and that both isoforms can have tumor suppressor activity *in vivo*[91,92].

The most extensively studied and validated substrate targets of pVHL-mediated ubiquitylation are the α subunits of HIF-1 and HIF-2. However other protein targets have been identified, including protein kinase C (PKC)[132]. Some clues for different HIF-1α-independent functions of pVHL have been proposed, but the molecular details of these potential roles remain elusive[133]. While the role of pVHL as part of the E3 ligase complex is the most widely studied and understood, several studies have suggested that pVHL may also assist in other proteasomal-independent and HIF-independent functions[134]. For example, pVHL has been involved with regulation of the extracellular matrix (ECM), by interacting with fibronectin, a glycoprotein that binds to integrin proteins to regulate the ECM[135].

Because of its role in hypoxia, HIF proteins play an important role in tumour proliferation, making them attractive targets in cancer therapeutics. Within certain contexts, pVHL plays a known role as tumour suppressor protein. Indeed,

VHL is frequently mutated or lost in various type of cancer, notably of the kidney, and give rise to the VHL disease[92]. VHL disease is an autosomal dominant inherited syndrome that is characterized by the predisposition and ultimately development of multiple tumors, both benign and malignant, particularly in the kidney, pancreas and central nervous system (CNS). The most common tumours are clear-cell renal cell carcinoma (RCC), renal cysts and phaeochromocytoma, and retinal and cerebellar haemangioblastomas[91,92]. Mutations in pVHL that give rise to these conditions typically weaken or disrupt the VHL:HIF-α interaction, or prevent the formation of a stable and functional $CRL2^{VHL}$ multi-protein complex. The result of these hampered interactions is ineffective HIF ubiquitylation activity, leading to constitutively stabilized HIF levels, and basal hypoxic and pro-angiogenic signalling. This promoted up-regulation of HIF activity correlates with tumour development and prognosis, providing further evidence for a role of the VHL-CRL as a tumour suppressor.

In contrast to the roles of VHL as tumor suppressor in kidney cancer and VHL disease tumorigenesis, recent studies have identified VHL inactivation by gene knockdown/knockout as protective or beneficial to lifespan in worm[136], and during states of mitochondrial dysfunction[94]. These evidences support VHL inhibition or induced degradation as a potential treatment for aging or human diseases associated with impaired mitochondrial function. There is also increasing evidence that VHL and HIF activity play an important role in the biology and signaling of T-cells[137], suggesting that modulating VHL function or its protein levels could be important for cancer immunotherapy.

1.5.1.2 Structure and interactions of VHL

The structure of the ternary complex between VHL, Elongin B and Elongin C (VBC complex) was first reported in 1999(ref. [129]). More recently the structure of VBC bound to the Cullin 2 NTD was also obtained [138], contributing to a more detailed structural elucidation of this machinery. However, full detailed picture of the VHL E3 ligase complex and understanding of its structure and functional mechanism still remain elusive (Figure 9). Indeed, only very few full-length

structures of cullins as part of a whole CRL complex have been solved since the landmark structure of the CRL1 complex was reported by Zheng et al[139].

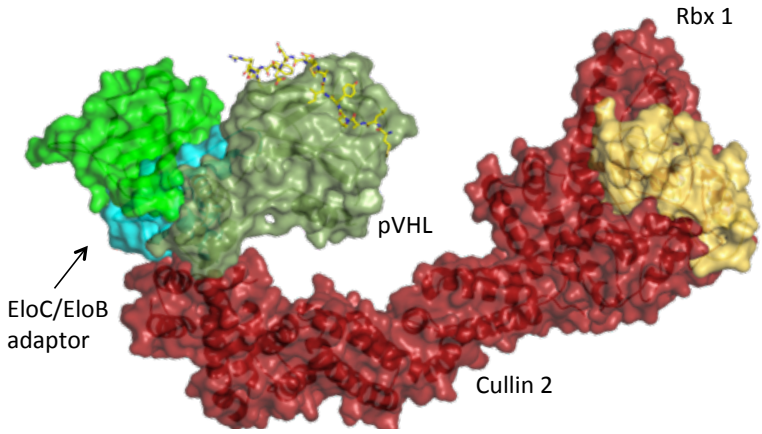


Figure 9. Structural model of the CRL2^{VHL} E3 ligase.

The crystallographic studies have revealed how VHL (residues 61–210 of human VHL) folds into two domains: 1) an N-terminal domain predominantly consisting of β-strands, and 2) a small α-helical domain in the C-terminal region. The α domain interacts directly with Elongin C, and to a lesser degree with Elongin B, and these interactions greatly stabilize the complex. Indeed, VHL cannot be recombinantly expressed alone in bacteria as it is unstable and unfolds into inclusion bodies[140]. These PPIs also mediate association of VHL with other components of the CRL complex, mainly Cul2. The β-domain (amino acids 63–154 and 192–210) instead binds specific E3 substrates, including HIF-α [141].

Two structures of VBC bound to a hydroxylated HIF-1α CODD peptide were published back-to-back in 2002 (ref. [79,80]). The CODD peptide adopts an extended conformation when bound to the N-terminal β-domain of pVHL (Figure 10). HIF-1α is held in place through extensive backbone and side-chain hydrogen bonds with pVHL, which limit its conformational flexibility. The C4 hydroxyproline residue in HIF-1α (Hyp564) is almost entirely buried and forms two key hydrogen bonds with the side chains of H115 and S111 in pVHL. These interactions contribute to the strict requirement for pVHL to bind a hydroxylated proline residue. However, the C4-*trans* stereochemistry at the carbon bearing

the hydroxyl group as well as the C4-*exo* conformation of the pyrrolidine ring both appear to be crucial determinants of pVHL binding[128]. The affinity of VHL for C4-*trans* hydroxylated CODD is over three orders of magnitude greater than for non-hydroxylated (or indeed C4-*cis* hydroxylated) CODD. This observation suggests that HIF-α prolyl hydroxylation has a switch-like regulatory effect on HIF signaling.

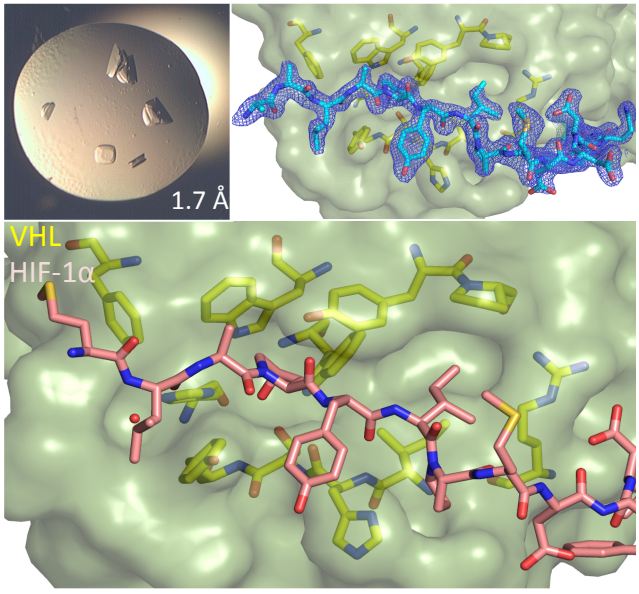


Figure 10. Crystal structure of a HIF-1α peptide bound to pVHL (PDB code 4AJY).

1.5.2 PHD enzymes

HIF prolyl hydroxylases (called PHD1-3 enzymes or EGLN 1-3 isoforms) are non-heam iron- and 2-oxoglutarate (2OG) dependent oxygenases[82]. They are responsible for HIF-α hydroxylation of two proline residues, P402 and P564, located within NODD and CODD, respectively[142,143]. Three PHD paralogs (PHD1, PHD2, and PHD3) have been identified, and their substrates are known to be quite diverse and isoform-specific. A brief account of each paralog of PHD enzymes follows.

1.5.2.1 Function of PHD enzymes

PHD1 is present in a wide variety of cell lines as a predominant species PHD1p43 and an additional species PHD1p40 (ref. [144]). PHD1 acts on several substrate proteins, including nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), and cell cycle regulator Cyclin D1, regulating their function in a hydroxylase activity-dependent manner. The large subunit of RNA polymerase II Rbp1 was also shown to be another substrate for PHD1, and prolyl hydroxylation is required for its VHL-dependent degradation[145]. PHD3 plays roles in neural development, immune system function, cell migration and apoptosis. Amongst the known protein substrates of PHD3 is non-muscle actin. Distinctly from its hydroxylase activity, PHD3 can act as a scaffolding protein and regulate various signaling pathways[146]. It is also interesting to observe that while PHD2 and PHD3 levels are strongly induced in hypoxia, expression of PHD1 (and PHD1p40 in particular) is suppressed by hypoxia, suggesting different roles of PHD enzymes in the hypoxic response.

PHD2 is the best-characterized and most extensively studied member of the HIF PHD enzymes, and it is thought to be the most important sensor of cellular oxygen level[82]. The enzyme utilizes mononuclear non-heme iron and 2-oxoglutarate (2OG, also known as α-ketoglutarate) co-substrates to form the desired product(s), succinate, and CO_2, incorporating one oxygen atom from O_2 into the hydroxylated product, and the other oxygen atom into succinate[147]. The enzyme has a high affinity for iron(II) and 2-oxoglutarate, and forms a long-lived complex with these molecules. It has been proposed that co-substrate and iron concentrations poise the HIF hydroxylases to respond to an appropriate "oxygen sensing window" for a particular cell type or tissue. Studies have revealed that PHD2 has a K_M for O_2 slightly above its atmospheric concentration[148,149].

1.5.2.2 Structure and interactions of PHD enzymes

PHD2 is a 46-kDa enzyme that consists of an N-terminal domain homologous to MYND zinc finger domains, and a C-terminal catalytic domain homologous to the 2-oxoglutarate dioxygenases. Crystal structures have elucidated the tertiary structure of the catalytic domain of PHD2, which contains a double stranded helix walls around the active site[150]. Compared to other 2OG dependent

oxygenases, PHD2 has a relatively deeper pocket with a narrower opening to the active site. The 2OG pocket of PHD2 consists of mainly hydrophobic residues. In contrast, the catalytic triad residues necessary for iron binding (His313-Asp315-His374) is highly conserved, and similar to other 2OG dependent oxygenases. Another residue important for cofactor binding and catalysis as reveal by site-directed mutagenesis is Arg383 (ref. [151]).

As with binding to pVHL, the conformation of the target proline residue is important for HIF binding to the PHDs, as shown by structural work with PHD2 (ref. [147]). The non-hydroxylated CODD proline adopts the C4-*endo* conformation when bound to PHD2; on the basis of crystallographic analysis, this conformation is proposed to be required for the productive binding and catalysis. NMR and other biophysical studies reveal that binding of the HIF-α CODDs to the PHDs involves substantial conformational changes, in particular involving a mobile loop region located between two β strands of PHD2 at the C-terminal region (see Figure 37 in section 3.1.1). The NODD is proposed to bind to the PHDs in a similar manner to CODD, however structural information for PHD-NODD complexes are still elusive[149]. Furthermore, no crystal structures of PHD3 or PHD1 are available to date in the Protein Data Bank (PDB). Other interacting partners that have been ascribed to PHD enzymes include the tumor suppressor protein LIMD1 (LIM domain containing protein 1), which is postulated to simultaneously bind PHDs and pVHL in a manner that promotes HIF-α degradation[152]. Many other interacting proteins have been reported to be PHD substrates on the basis of antibody and/or proteomic mass spectrometry analyses. The relevance of these interactions to the hypoxic response remains to be established.

1.6 PPI Inhibition of the VHL-HIF pathway

Targeting the VHL-HIF pathway using small molecules is a rapidly developing and highly dynamic field of drug discovery, with much therapeutic potential[149]. As explained before, stabilization of HIF, achieved by both genetic methods and pharmacologically by treatment with inhibitors, has been reported as a possible

therapeutic strategy for conditions including anemia and ischemic/vascular diseases. Moreover, blockade of HIF activity is proposed as an anti-cancer therapeutics. Inhibitors targeting the HIF pathway will be described next, focusing in particular on the ligands and inhibitors used for the studies described in this thesis (Figure 11).

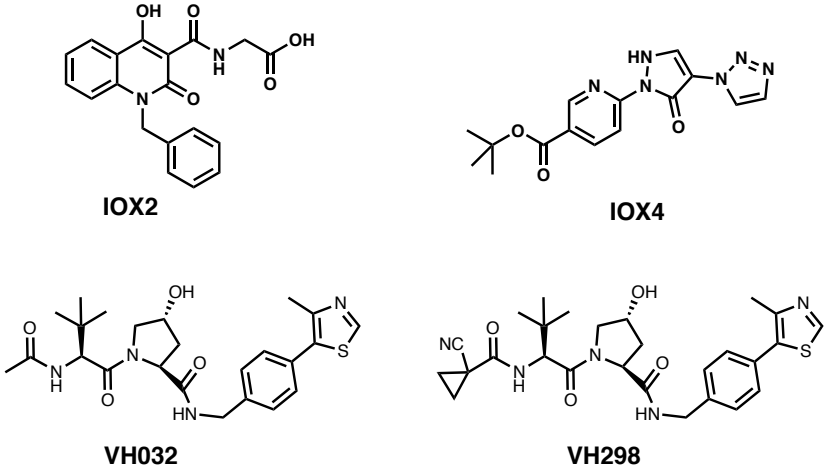


Figure 11. Chemical structures of PHD and VHL inhibitors relevant to this work.

1.6.1 PHD inhibitors

Targeting the PHDs via small-molecule inhibition has been a major focus of pharmaceutical efforts on the HIF system to date[153]. The vast majority of the reported PHD inhibitors, and all of those in clinical trials, have been developed to bind the active site iron and to compete with 2OG. Although there are subtle differences in binding modes, and compounds can be pan-selective for PHD enzymes over and above other 2OG oxygenases, virtually none of the PHD inhibitors described to date are selective for a single paralog over the other two proteins, as predicted given the similarity of the catalytic domain of the three human PHDs.

1.6.1.1 IOX2

Compound IOX2 (ref. [154]) was reported as a selective inhibitor of PHDs enzymes, with some selectivity for the PHD2 paralog. An analogue of IOX2 was shown to

bind into the relatively narrow pocket of the active site opening. The binding mode of this inhibitor will be described in more details in Section 3.1.1.

Biological evaluation of IOX2 and related inhibitors in both human renal carcinoma (RCC4) cell lines lacking VHL and VHL-competent cell lines, showed increased HIF-1α levels upon compound treatments, with EC_{50} of 0.033 μM. Related compound Roxadustat (FG-4592) is an oral hypoxia-inducible factor prolyl hydroxylase inhibitor that stimulates erythropoiesis and regulates iron metabolism. FG-4592 is currently in phase 2b study as an oral treatment for anemia associated with chronic kidney diseases (CKD) [93].

1.6.1.2 IOX4

Most of the previously reported PHD inhibitors, like IOX2, contain a carboxylic acid moiety to mimic the same group in the 2OG cofactor. However, carboxylic acid is often undesirable from a pharmacological perspective. Recently, compound IOX4 (ref. [155]) was developed as a PHD inhibitor which bears a completely different chemical scaffold compared with previously reported inhibitors. IOX4 is characterized by a tricyclic structure with linked pyridine-carboxylate, dihydropyrazole and triazole rings. The iron molecule is coordinated in a bidentate manner by the nitrogen atoms of pyridine and pyrazolone rings. The pyridine-based chain extends towards the entrance of the active site in a manner that differentiates this compound from the previous reported inhibitors. Studies in human cell lines support the efficacy of IOX4 as a potent and selective PHD inhibitor, with a maximal level of HIF-α induction being observed at low μM concentrations across all the cell lines tested. Although IOX4 was not tested against PHD1 and PHD3, the potent inhibition of NODD and CODD hydroxylation in cell lines containing all three PHD paralogs suggested that IOX4 also inhibits PHD1 and PHD3. The induction of HIF-α in mice treated with IOX4 demonstrated the utility for use *in vivo* of this inhibitor[155].

1.6.2 VHL inhibitors

A different approach to induce HIF-1α stabilization is by blocking its PPI with VHL. Professor Ciulli laboratory, initially in collaboration with scientists from

Yale University, developed the first series of inhibitors of the PPI between VHL and HIF-α (refs. [156-158]). The first generation of inhibitors was designed using a hydroxyproline fragment as a starting point (Figure 12). This choice was motivated by the knowledge that Hyp564 on HIF-1α makes key interactions with VHL and is crucial for HIF-1α binding, as previously described. The hydroxyproline fragment was grown initially guided by rational computer-aided design at the two ends of the molecules, called left-hand and right-hand sides (LHS and RHS, respectively)[156].

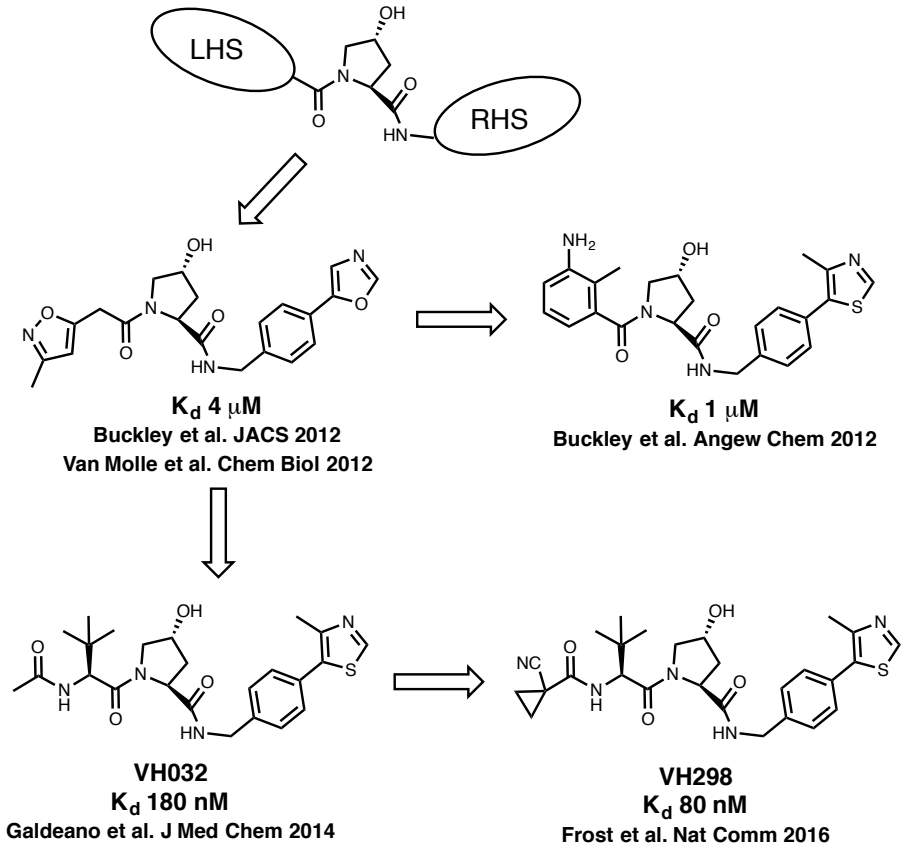


Figure 12. Progressive optimization of VHL ligands from a starting Hyp fragment to current most potent inhibitors.

An isoxazolylacetamide fragment moiety was introduced in order to interact with a water molecule previously identified as an important part of the hydrogen-bonding network between VHL and HIF-1α, and this elaborated compound was shown to constitute a minimal pharmacophore using NMR binding studies[156]. A benzyl group was positioned in order to interact with the side chain of Tyr98, and further elaborate to interact at the RHS pocket. The best

compound of this series achieved a K_d of 5 μM by ITC and FP, and the crystal structure of the compound bound to VHL was solved. This structure elucidated the key interactions formed, with the LHS of the inhibitor mimicking closely the recognition mode of the HIF peptide, while the RHS exploiting interactions deep in a VHL pocket that are not explored by the natural substrate[156].

With the aim to build further structure-activity relationships (SAR) around the chemical series, fragment-based group optimization studies were conducted around the starting hydroxyproline scaffold. These still achieved compounds with potencies that could not surpass single digit micromolar values of K_d and were still inactive in cells[157,158]. Nevertheless, group-based studies helped to identify a *tert*-butyl group as a beneficial group at the LHS[157]. Guided by extensive crystallographic analysis of co-crystal structures, and careful medicinal chemistry optimization, a new series of compounds was designed that achieved binding affinities in the nanomolar range and improved lipophilicities to aid cell permeability [115]. X-ray crystal structures and ITC studies reveal new interactions on the LHS of the pocket and elucidated the structural and thermodynamic determinants for the improvement in binding affinities. The best ligand of this series (VH032, compound 7 in the original paper[115], see Figure 12) was more potent than a model 10-mer HIF-1α peptide, providing an excellent lead compound targeting VHL[115]. Further optimization of the far LHS of the molecules by introducing a cyano-cyclopropyl moiety led to VH298 (Figure 12), a more potent (K_d of 80 nM by ITC and FP) and more cell-permeable compound than VH032, as measured using an artificial membrane permeability assay[121]. During cellular evaluation, VHL inhibitors showed to be highly specific for VHL as their only major cellular target, and to highly selectively stabilize the hydroxylated form of HIF-α, consistent with blockade of the HIF-VHL pathway downstream of HIF hydroxylation[121]. VH298 has been elected as novel chemical probe of the hypoxic signaling pathway and is available from the Chemical Probes portal (http://www.chemicalprobes.org/vh298). Despite the significantly higher potency and cellular efficacy, VH298 needed to be used at >10 μM concentration to drive a significant HIF response, underlining the challenges associated with blockade of the tight VHL-HIF interaction[121].

1.6.3 HIF inhibitors

Compounds that inhibit or block HIF transcriptional activity could provide novel useful anti-cancer therapies (recently reviewed in ref. [159]). However, targeting transcription factors is notoriously challenging, especially when structural information is lacking. Several strategies have been undertaken to achieve HIF inhibition pharmacologically. Amongst the most successful one is the use of small fragment-like molecule to "fill" a deep pocket in the PAS domain of HIF. Because the HIF-2α PAS domain was found structurally to be slightly larger in size than that of the analogous HIF-1α domain, such molecules were able to selectively target HIF-2α over HIF-1α, which is potentially attractive for cancer. More recently, such molecules were further developed into a potent clinical candidate (named PT2399, from Peloton Therapeutics) against renal cancer, validating HIF-2 as a target in ccRCC[160,161]. The recent crystal structure of large portions of the HIF-α and HIF-β proteins will no doubt facilitate future structure-based drug design efforts against the PAS and other domains, as well as targeting new PPIs[73].

Complementary approaches to block HIF activity have involved the use of peptidic inhibitors of the PPI between HIF-α and HIF-β. The Tavassoli group has developed cyclic hexapeptides that effectively block this PPI[162]. The best compound identified, termed cyclo-CLLFVY, revealed the capacity to disrupt HIF-1 hetero-dimerisation in cells by targeting selectively the HIF-1α PAS domain.

1.7 Inhibitors vs degraders

Amongst the examples of compounds introduced so far are conventional inhibitors, i.e. compounds that block a target protein, as well as inducers of target degradation. Several important conceptual and practical points are noteworthy when comparing and contrasting inhibitors with degraders.

Firstly, pharmacological activity of inhibitors is based on occupancy of the binding site of target protein. This underlines how potency is a critical factor when developing such class of compounds. The necessity of high concentration to achieve the desired effect in cells is another crucial point. This strict requirement could lead to off-target effects and toxicity, factors that need to be avoided when developing a chemical probe or a drug. Thirdly, in order to affect target function, inhibitors need to bind to a functional site on the protein, i.e. in such a way that they affect protein function either directly or allosterically. However, the concentration of the target protein will not decrease. Indeed, in many cases the concentration of the protein increases as feedback response of the inhibition, or because the potent ligand greatly stabilizes the target.

Degraders differ from inhibitors from structural and mechanistic points of view. Degraders act catalytically and independently from the functionality of the binding site on the target. Because of the coexistence of two binding moieties, PROTAC compounds are able to bring in close proximity the degrading machine and the target protein, inducing the degradation only of the latter. The degrading machinery and the compound will be not degraded and they will be available to act again, like in a catalytic cycle. Consequently, this will affect the final concentration of the target protein and will result in slower recovery from inhibition, because re-synthesis of the target is required, leading to prolonged efficacy. It has also been shown that degraders can be active in cells at concentrations significantly lower than the dissociation constants of their respective ligands, meaning the efficacy of the chemical intervention can be much more profound. Finally, removing an entire protein can have a very different effect from simply blocking a single binding site or a single activity, and it would be expected to pair well with genetic knock-down or knock-out of the target.

1.8 Aims and objectives

The oxygen signaling pathway has been described as one of the most important systems used by the cells to sense and react to lack of oxygen. Even if quite a lot has been discovered during the past few decades, still a lot of mechanisms behind this pathway are not well understood and many questions remain to be answered.

In the past, inhibitors have been developed to target the key enzymes and the key PPI involved in oxygen sensing, with the aim of induce HIF-1α stabilization for both therapeutics and research purposes. These inhibitors, such as IOX2, IOX4 for PHDs and VH032 and VH298 for VHL, are allowing researchers to address many biological questions. They also provide lead compounds for the development of potential new therapeutics targeting hypoxia signaling. However, even if active in cells, some of these compounds have been found to lack specificity for their target proteins, and they generally experience lower cellular potency relative to their in vitro affinity for the recombinant proteins. This means higher concentrations, typically in the micromolar range, are required for use in cells, leading to possible off target effect. On the other hand, previous studies have shown how it is possible to successfully turn inhibitors into more potent degraders. Moreover, recent studies from the Ciulli group have demonstrated that inhibitors targeting more than a single protein i.e. pan-selective, can be turned in to more selective and more effective degraders for individual target proteins, by using the PROTAC approach (Figure 13).

In light of these observations, the questions we wish to address in this project are:

- What would we learn by turning VHL and PHD inhibitors in bifunctional molecules?
- Can we induce degradation of the target proteins? if yes, could these molecules exhibit increased potency and cellular efficacy compared to the inhibitors? Could they exhibit an added layer of selectivity for some of

their targeted protein paralogs or isoforms? What would be the impact to the whole PHD-VHL signaling pathway?

The overall aim of this project is therefore to design, synthetize and evaluate from a biological point of view different series of PROTACs to target VHL and PHDs as two of the main players of the oxygen-signalling pathway.

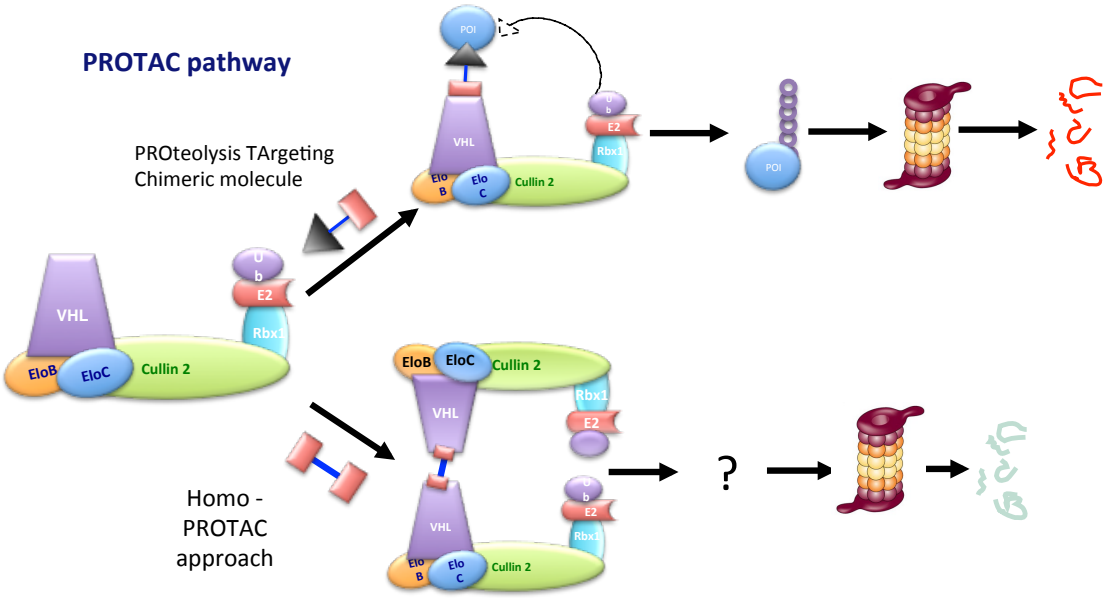


Figure 13. Comparison between conventional PROTAC approach (top pathway) and novel approach of Homo-PROTACs (bottom pathway).

2 VHL-targeting compounds

The design, synthesis and biological evaluation of VHL-targeting bifunctional molecules will be described in this chapter.

2.1 Design of VHL-targeting compounds

Two classes of bifunctional compounds were designed to target VHL – a first class of homo-bifunctional to recruit $CRL2^{VHL}$, and a second class of hetero-bifunctional compounds aiming to recruit $CRL4^{CRBN}$.

2.1.1 Homo-PROTACs recruiting $CRL2^{VHL}$

To address the question if VHL could degrade itself, the first idea was to develop a class of Homo-PROTACs compounds where a specific VHL ligand was used as lead scaffold. Structure of high-affinity binder VH032 (ref. [115]) was chosen as starting point for the design of these series of compounds because it was the most potent compound at the time of the start of the project.
Because the primary interest was to hijack $CRL2^{VHL}$ as degrading machinery to induce depletion in cellular levels of VHL itself, it was decided to link together two VH032 scaffolds in order to bring in close proximity two VHL protein molecules at the same time.
Design of bifunctional molecules required careful consideration of two major points: the position, type and length of the linking.

2.1.1.1 Linking position

In order to retain the strong binding affinity (K_d = 180 nM) that characterizes VH032, the crystal structure of VH032 bound to VHL (Figure 14) was analysed with the aim to identify solvent exposed regions from where the ligand could be derivatized without perturbing its binding mode.

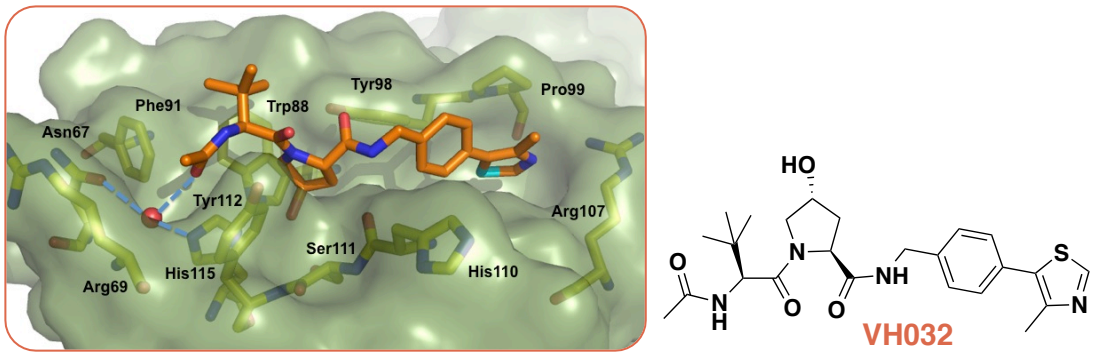


Figure 14: Crystal structure of VHL with bound VH032 ligand[115].

From previous studies (section 1.5.1.2), it is known that the hydroxyproline (Hyp) scaffold plays a key role in the interaction, and it is bound deeply in a pocket forming hydrogen bonds with the His115 and Ser111 side chains of VHL (Figure 14). Therefore, it was reasonably assumed that no modifications would be allowed on the Hyp core. It is also known that the methyl-thiazole ring binds in an energetically stable conformation, forming optimal interactions with right hand side (RHS) pocket present in the VHL structure, including a hydrogen bond with Arg107 and an aromatic C-H interaction with the backbone carbonyl of Pro99 (Figure 14). It was therefore deemed that the terminal thiazole ring should not be modified.

The *t*-butyl group as well was considered as not suitable for derivatization because pointing upward to make hydrophobic contacts with Phe91 and Trp88. In contrast, the methyl group of the terminal acetyl group is solvent exposed and points to an unfilled region of the VHL protein, where the natural HIF-1α peptide extends to interact with VHL[79,80]. This position therefore was thought to offer a suitable connecting point for a linker. This position was also used for the development of active PROTACs against the protein Brd4, later published by the Ciulli group, and it was found to have no detrimental impact on the binding mode of the VHL ligand[116].

Another position previously used for the developing of Halo-PROTACs by the Crews group and found to be suitable and available for the attachment of a linker is position 2 from the phenyl ring on the RHS of the ligand[124].

2.1.1.2 Type and length of the linker

Previous published papers on Brd4 degrading PROTACs suggest that the linker length between the two ligands moieties plays a major role in determining the efficacy of protein degradation[116]. It is also known that the chemical and physical properties of the linker could impact on the compound efficacy. Use of flexible linkers, such as polyethylene glycol (PEG) chains, or rigid rods, characterized by *p*-phenylacetylene chain (Figure 15), could have an impact on the ability of the compound to bring the E3 ligase productively in close proximity to the protein of interest for ubiquitination and subsequent proteasomal degradation.

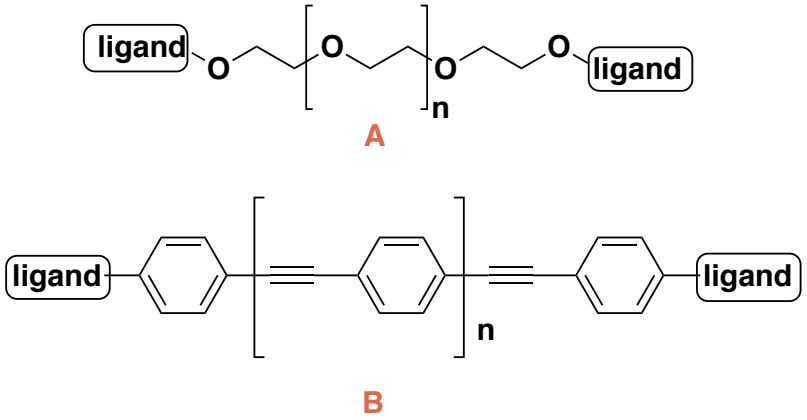


Figure 15: Two different linker types: PEG (A), and rigid rod (B).

A different VHL ligand later considered during the PROTAC design was VH298 (Figure 16). This compound was developed within the Ciulli group through a SAR optimization campaign of the VH032 scaffold, carried out in parallel during the course of our study. As explained before (section 1.6.2) the acetyl group on the left-hand side (LHS) of VH032 is substituted with a cyano-cyclopropyl moiety in VH298. This modification led to increase in binding affinity and cell permeability, resulting in a more active VHL inhibitor inside cells[121].

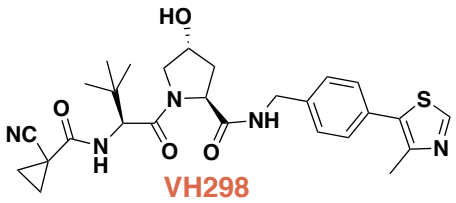


Figure 16: Structure of VH298.

To fully investigate the impact of the linkage position on the activity of the compounds inside cells, it was chosen to develop a first series of symmetric and asymmetric Homo-PROTACs (Figure 17) derivatizing VH032 from both the terminal acetyl groups and position 2 of the phenyl ring. In the case of VH298, it was chosen to derivatize only from position 2 of the phenyl ring. Linking from the LHS would require the possible impact of replacing the nitrile group of the cyclopropyl moiety to be evaluated, which was warranted for future work. Linkers with different lengths comprised of polyethylene glycol chains with either three, four or five ethylene glycol units were chosen to connect the two VHL ligands.

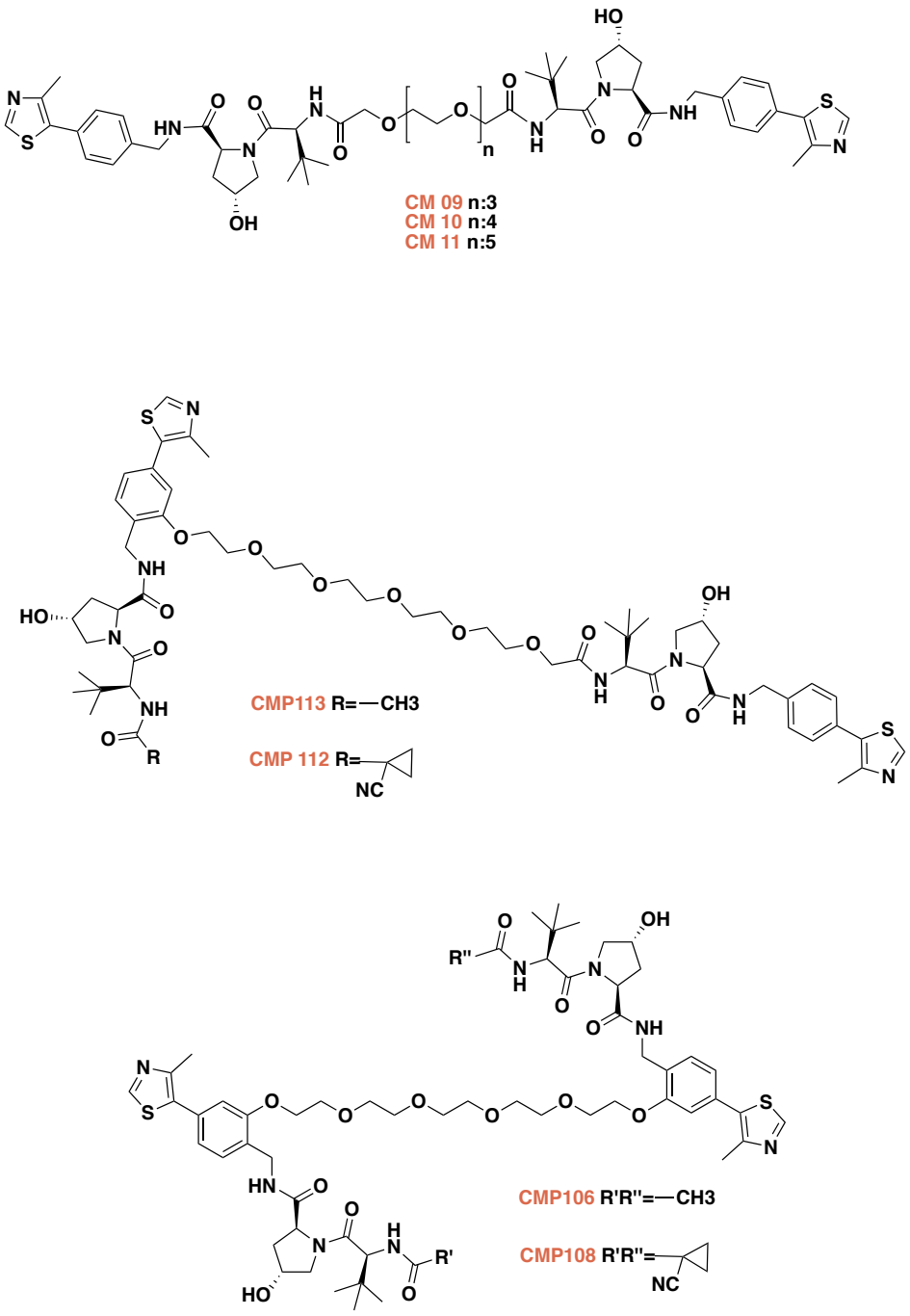


Figure 17: First series of symmetric and asymmetric Homo-PROTACs.

Compound DAT265 (a gift from Dr. Andrea Testa, Figure 18) was also taken in consideration due to the different nature of the linker. In fact, this compound is characterized by the presence of a rigid linker made by a *p*-phenylacetylene chain that confers to the final molecule a rigid structure. Due to its increased rigidity and in order to investigate the possibility of using more rigid linkers, DAT265 was deemed interesting to be added to the series.

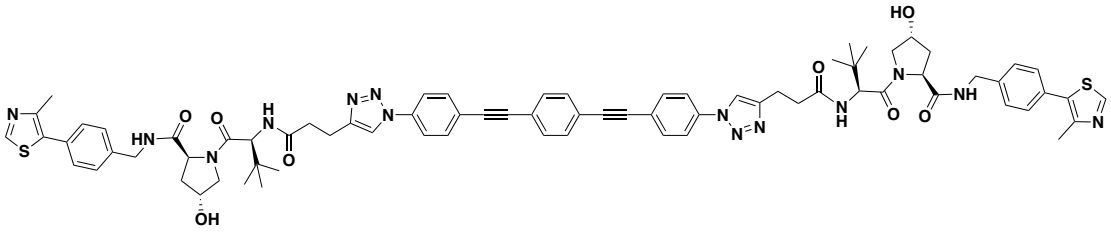


Figure 18: Structure of DAT265.

As negative controls, *cis-cis* and *cis-trans* equivalent of compound CM10 and CM11 respectively were designed. The *cis* conformation of the hydroxyl in the proline residue is known to disrupt the key interaction with His115 and Ser111 side chains of the VBC protein in the context of the VHL ligand alone[121]. We therefore designed two different epimer compounds (CMP98 and CMP99, see Figure 19). CMP98 is a *cis–cis* epimer, which would be expected to lose any binding with VHL and consequently be completely inactive. CMP99 is a *cis-trans* epimer compound, which would be expected to retain binding to a single pVHL molecule in a 1:1 fashion, thus act as inhibitor but not as degrader.

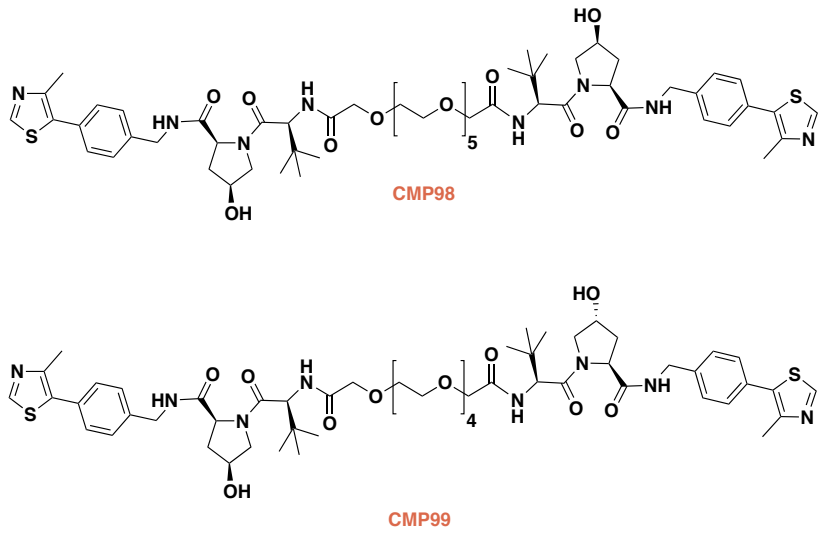


Figure 19: Negative controls CMP98 (*cis-cis*) and CMP99 (*cis-trans*).

2.1.2 Hetero PROTACs recruiting CRL4^{CRBN} to target VHL

In order to investigate the possibility of inducing the degradation of VHL using different degrading machinery, a series of so-called 'hetero' PROTACs recruiting

$CRL4^{CRBN}$ were designed. It was also deemed possible that this class of molecules could induce the degradation of CRBN as well as or instead of VHL. For the VHL recruiting moiety, it was chosen to use VH032 as ligand and the acetyl at the LHS of the molecule for the attachment of the linker.

Because of its greatest cellular stability, pomalidomide was chosen as starting point for design of the $CRL4^{CRBN}$ recruiting moiety (Figure 20). Analysis of the crystal structure of pomalidomide in complex with CRBN protein suggested that a 2-methylene chain linker appended to the amine at position 4 of the phthalimide anhydride would not affect the binding to CRBN and at the same time provide a convenient attachment of the linker from a synthetic point of view.

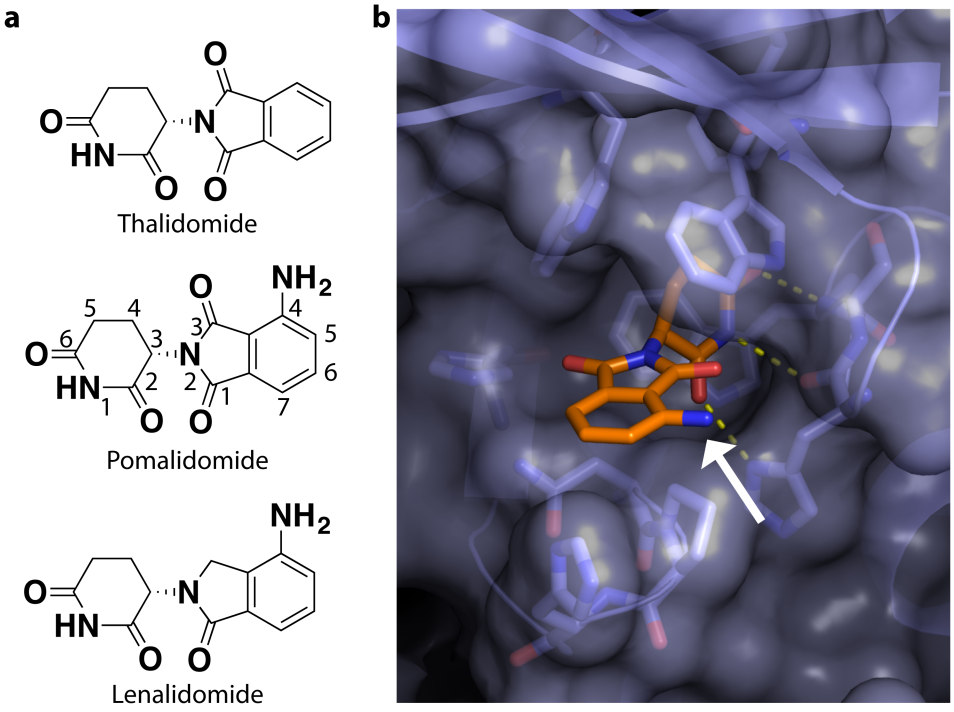
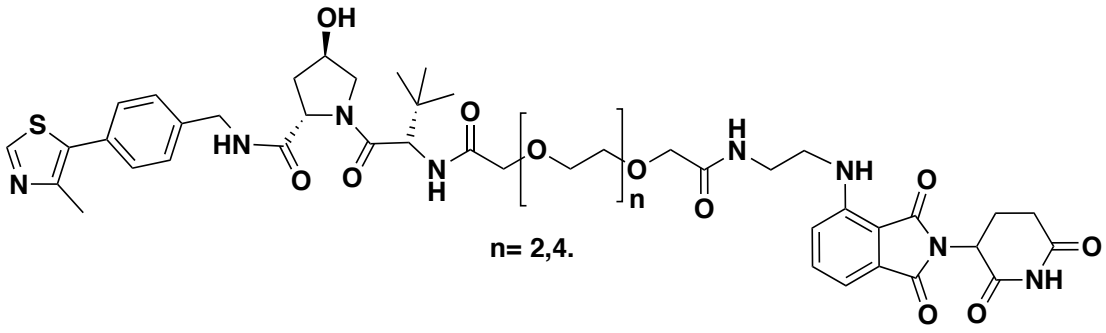


Figure 20: Immunomodulatory drugs targeting cereblon. (a) Cheimcal structures. (b) Crystal structure of pomalidomide bound to CRBN (PDB code 4CI3, ref. [34]).

During previous studies (Dr. Andrea Testa, unpublished observation) it was found that the amine in C4 does not react with carboxylic acid groups using common coupling agents. For this reason, a different approach was evaluated where a small 2-methylene chain with a terminal primary amine was attached at

C4. The free amine at the end of the chain was designed in order to facilitate peptide bond formation with the carboxylic acid at the end of the linker attached to the VHL ligand.

Structures of the final molecules are shown in Figure 21.



CMP85 n=2
CMP86 n=4

Figure 21: Structure of Hetero-PROTACs designed to recruit $CRL4^{CRBN}$ at one end and $CRL2^{VHL}$ at the other end.

2.2 Synthetic routes to VHL-targeting compounds

Next, the synthetic routes to achieve both homo- and hetero-bifunctional compounds targeting VHL are described.

2.2.1 Synthesis of Homo-PROTACS CM9, CM10, CM11, CMP98

The synthetic route to Homo-PROTACs CM9, CM10, CM11 and CMP98 (see general structures in Figure 17 and Figure 19) can be divided in three different parts: first, synthesis of compound **6**, that will constitute both the left and right sides of the final molecule; second, synthesis of linkers that will connect the two moieties together; and third, the assembly of the final compounds.

2.2.1.1 Synthesis of VHL ligand 6 and its *cis* analogue 13

Compound **6** was prepared as previously reported[115], with minor modifications that led to significant yield improvements. The first step, a palladium catalyzed direct arylation of 4-methylthiazole by 4-bromobenzonitrile in the presence of $Pd(OAc)_2$ as catalyst and KOAc as base, led to the formation of compound **1** with essentially quantitative yield (97%). Compound **1** was isolated by filtration after aqueous quenching and precipitation with NaOH.

The cyano derivative product **1** needs to be reduced to a primary amine. The conditions previously reported for the reduction involved the use of $NaBH_4$ and $CoCl_2$ in MeOH[115]. The reaction conducted in this way gave very low yield (~20%) and purification issues during the work up. The crude material was directly utilized in the next synthetic step, however it was found that a side product likely formed during this step could interfere during the further steps.
In order to optimize this reaction, different conditions were tested. First of all, hydrogenation reactions using different catalysts, such as Pd/C and Raney nickel, were taken in consideration. Initially, these reactions were performed in a H-Cube machine, a benchtop hydrogenation reactor where continuous flow of substances is combined with endogenous on-demand hydrogen, generated from electrolysis of water and with a disposable catalyst cartridge system. Unfortunately none of the conditions tested showed conversion of the starting material (Table 1). The same hydrogenation reaction was performed in THF solution under H_2 atmosphere in presence of Pd/C (10% w/v) as catalyst and triethylamine as base. After 18 h at 40 °C, no conversion of the starting material was observed.
Because the tested conditions of hydrogenation reactions proved ineffective, a different approach involving use of $LiAlH_4$ was evaluated[163]. A pilot reaction was conducted treating compound **1** with 1 M solution of $LiAlH_4$ in THF in a 1:2.5 molar ratio at reflux. After 2 h, it was possible to observe the conversion of the starting material in to the desired product (Table 1). Unfortunately, the formation of side products during the reaction led to the isolation of the desired amine in only low purity. Moreover, the formation of a big amount of aluminium hydroxide during the quenching with NaOH made the phase separation quite slow, promoting the formation of an emulsion. Different attempts were

performed in order to optimize the yield, improve the purity of the product and simplify the work up (Table 2).

Reaction	*Conditions*	*Result*
Hydrogenation using Pd/C	0.05 M solution of CM 01.1 H cube: 50 bar, 70 °C	No conversion of starting material
Hydrogenation using Raney Ni	0.05 M solution of CM 01.1 H cube: 50 bar, 70 °C	No conversion of starting material
Hydrogenation using Pd/C	THF, Pd/C 10%, formic acid Et_3N	No conversion of starting material
$LiAlH_4$ in THF (1 M)	$LiAlH_4$, THF, 0 °C →reflux, 2 h;	Conversion of the starting material. 25% yield.

Table 1: Attempts at reducing compound **1** in to compound **2**.

The large excess of $LiAlH_4$ used was identified as a possible cause of formation of the observed side products. For this reason the reduction was performed treating compound **1** with a 1 M solution of $LiAlH_4$ in THF in a 1:1.04 ratio at reflux. The work up was improved using a 10 times more concentrated solution of sodium hydroxide, reducing in this way the amount of water added to the mixture. Furthermore, chloroform was used instead of ethyl acetate, because of the greater solubility of the desired product in the chlorinated polar solvent. The precipitate of aluminium salt was filtered off under vacuum, facilitating the extraction step. The crude resulting after evaporation of the organic phase was purified by column chromatography over silica.

Reaction	**Conditions**	**Work up**	**Comments**
$LiAlH_4$ in THF (1 M)	CM 01.1/ $LiAlH_4$ 1:2.5 1. $LiAlH_4$ in dry THF, 0 °C, Ar. 2. reflux, O/N	• Addition of NaOH 1 M • Extraction with EtOAc • Extraction with HCl 1 M • rotaevaporation of water phase • Addition of NaOH 10 M till pH	• Formation of side products • Low purity

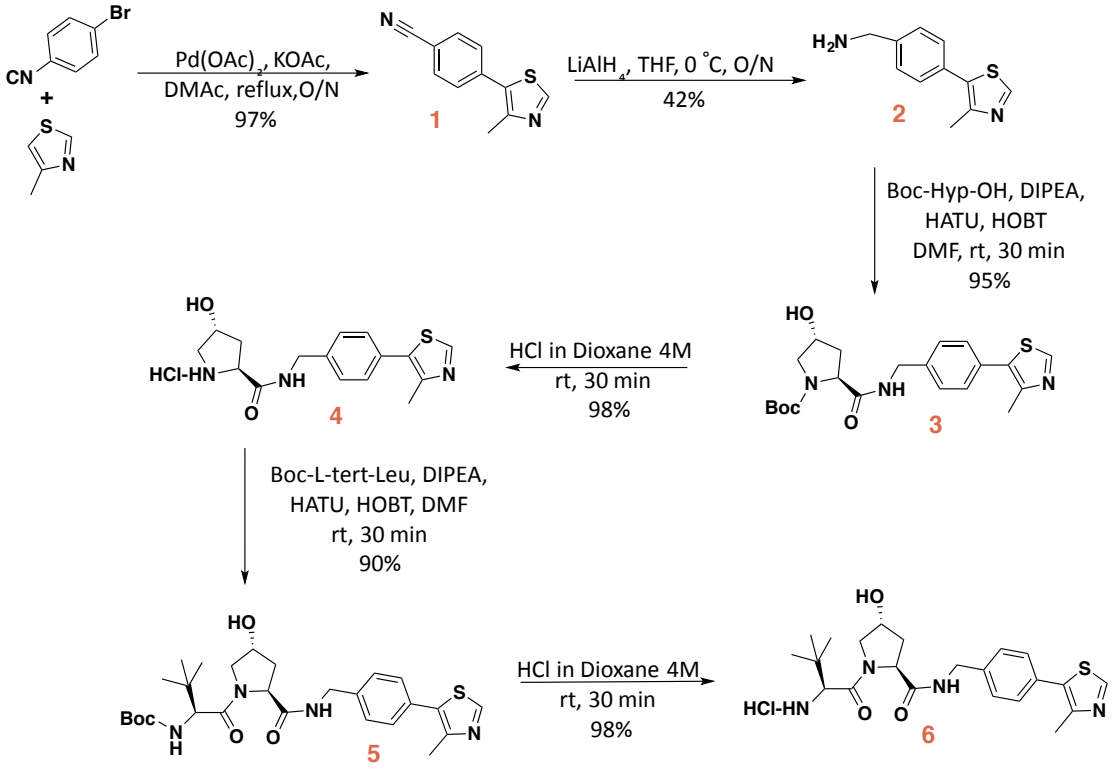
		8 • Extraction with EtOAc	
$LiAlH_4$ in THF (1 M)	CM 01.1/ $LiAlH_4$ 1:1.1 1.LiAlH4 in dry THF, 0°C, Ar. 2. reflux, O/N	• Addition of H_2O and NaOH 10 M till pH=8 • Extraction with $CHCl_3$ • Purification by column (MeOH from 0 to 10% in DCM)	• Less formation of side products • Yield increased to 42%
$LiAlH_4$ in THF (1 M)	CM 01.1/ $LiAlH_4$ 1:1.1 1. $LiAlH_4$ in dry THF, 0°C, Ar. 2. 50 °C, O/N	• For 15 mmol of $LIALH_4$: • add 600 µL of H_2O, 450 µL of 20% NaOH, 2 ml of H_2O • Filter of the aluminium salt and wash with $CHCl_3$ • Purification by column (MeOH from 0 to 10% in DCM)	• Pure product • Less formation of side products • Yield increased to 52%

Table 2: Optimization of nitrile reduction reaction using $LiAlH_4$

These changes significantly improved the work up and purification steps, although some side products and impurities were still detectable in the desired compound. For this reason, the temperature of the reaction was lowered to 50 °C and after 18 h the reaction was quenched by adding a very low amount of NaOH aqueous solution and water during the work up (Table 2). This led to the formation of a heavy precipitate of aluminium salt, which was much easier to filter off. Using these new conditions, the yield increased from the initial 20% to 52% and purity increased to >95% (Table 2). The compound obtained was ready to be used in the next step.

The amine **2** was coupled in the first instance with *trans*-*N*-(*tert*-butoxycarbonyl)-4-hydroxy-L-proline, in the presence of HATU and HOAT as coupling reagents, DIPEA as base and DMF as solvent. The resulting product **3,** obtained in 95% yield after column chromatography, was first deprotected using a 1:1 mixture of 4 M HCl in dioxane and DCM, giving compound **4**. The latter, obtained as a hydrochloride salt, was subsequently coupled to Boc-L-*tert*-

Leucine, in the same conditions used in the previous coupling step, to give compound **5**. Deprotection of **5** in the same acidic conditions described for the deprotection of **3** delivered compound **6** (Scheme 1).

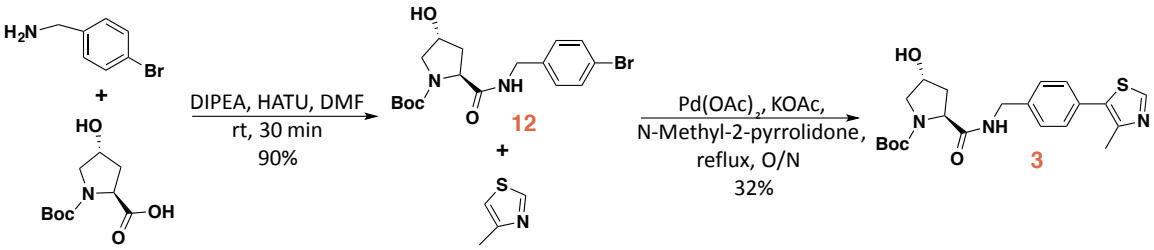


Scheme 1: Synthesis of VHL amine ligand, compound **6**.

Because of the difficulties connected with the reduction of the cyano-derivative, an alternative route was evaluated to synthetize compound **6**. The approach used by GSK (note [1]) to synthesize the same ligand was considered (Scheme 2). In this patent, bromo-benzylamine was coupled with *trans*-*N*-(*tert*-butoxycarbonyl)-4-hydroxy-L-proline, affording compound **12**. A palladium-catalyzed direct arylation of **12** with 4-methylthiazole was then employed. Using this procedure, it is possible to obtain the same ligand in two steps instead of the three required by our approach, also bypassing the challenging reduction step. For these reasons, this alternative approach was tested. The first step, the coupling of bromo-benzylamine with *trans*-*N*-(*tert*-butoxycarbonyl)-4-hydroxy-L-proline was successful and gave compound **12** with a yield of 98%.

[1] Patent WO 2014/108452

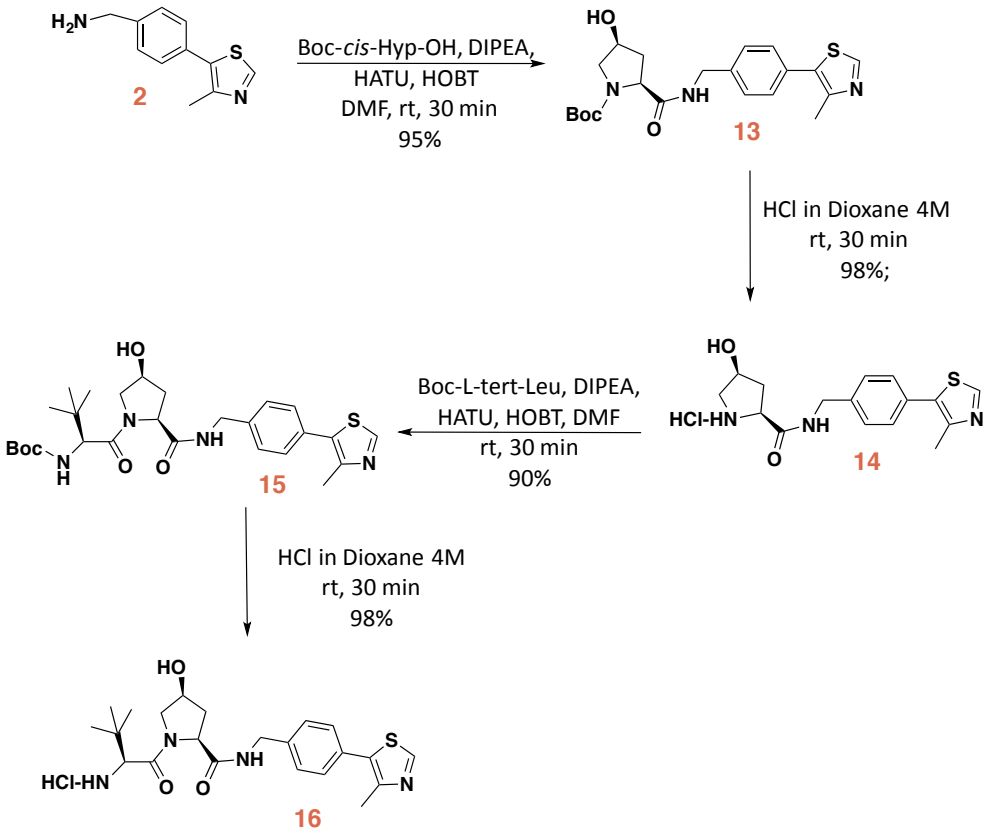
Unfortunately, the arylation step led to formation of the desired product **3** in only 32% yield (Scheme 2), which is lower compared to the 70% yield we obtained over two steps in the first route. Formation of side products leading to an impure final compound was also observed. For these reasons, it was decided not to pursue this new route further and maintain the route as previously developed and optimised.



Scheme 2: Alterative route for the synthesis of compound **3.**

As a negative control, **16** was prepared as the *cis* analogue of compound **6**. The corresponding PROTAC compound is expected to be inactive because the *cis*-epimer of hydroxyproline prevents binding to the VHL protein[121].

Compound **2** was coupled with *cis*-*N*-(*tert*-butoxycarbonyl)-4-hydroxy-L-proline unit instead of the *trans*-hydroxyproline, using the same conditions described before leading to the formation of compound **13** (Scheme 3). Boc deprotection and subsequent coupling of the latter with Boc-L-*tert*-Leucine afforded **15** in good yield. Deprotection of the latter in acidic conditions delivered compound **16** in quantitative yield.



Scheme 3: Synthesis of *cis* analogue **16**

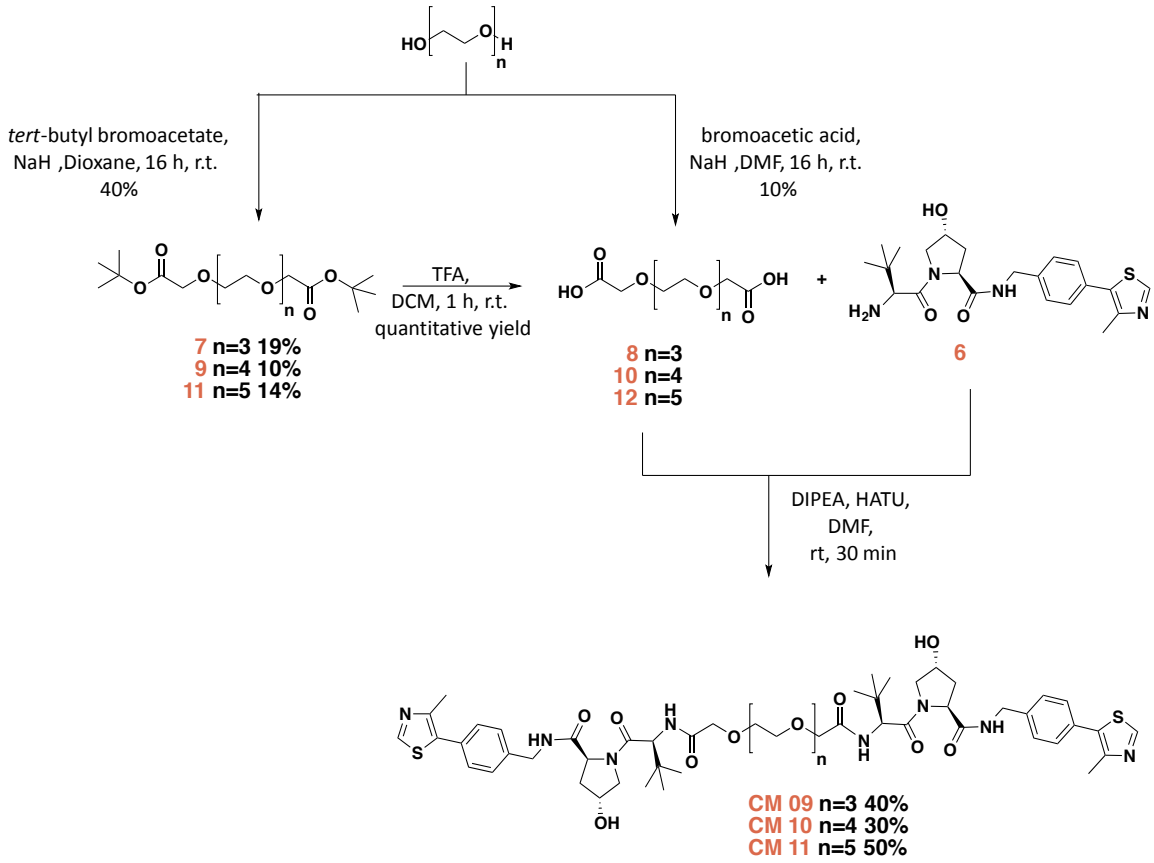
2.2.1.2 Synthesis of the linkers

For the synthesis of the linkers, two different strategies were pursued (Scheme 4). First, triethylene glycol was reacted with bromoacetic acid in a 1:2 ratio in the presence of NaH in dioxane. This approach led to the formation of a large amount of side products, probably due to the unprotected acid moiety of the bromoacetic acid reacting with itself in the presence of NaH. In addition, the product was really difficult to purify by column chromatography because of its high polarity due to the presence of free carboxylic acid moieties.

To overcome these challenges, a second strategy was devised (Scheme 4). In this new route, *tert*-butyl bromoacetate was reacted instead with triethylene glycol. Compound **7** was conveniently isolated after purification with moderate-to-low yield but high purity. Further deprotection of the *tert*-butyl protecting group using trifluoroacetic acid (TFA) afforded compound **8**. Compounds **10** and **12** were synthesized following the same route starting from tetraethylene glycol and pentaethylene glycol, respectively.

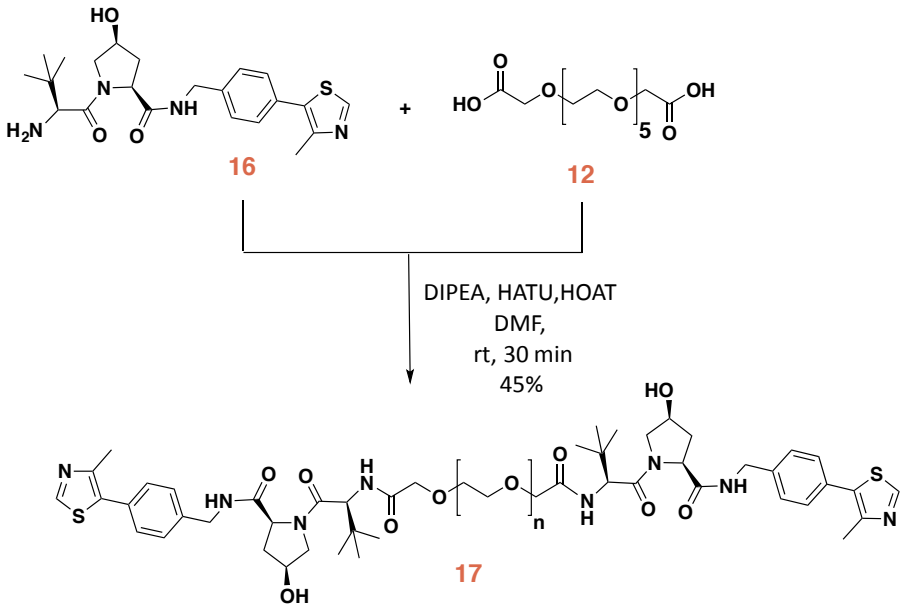
2.2.1.3 Assembly of CM9, CM10, CM11 and inactive compound CMP98

The final compounds **CM9**, **CM10** and **CM11** were obtained after coupling compound **6** with likers **8**, **10** and **12** respectively in the presence of HATU, HOAT, DIPEA and DMF (Scheme 4). All the final compounds were purified by preparative HPLC in order to obtain compounds of high purity required for biological evaluation.



Scheme 4: Synthesis of linkers **8**, **10**, **12** and of compound **CM09**, **CM10** and **CM11**

For the synthesis of the inactive PROTAC **17** (CMP98), **16** was coupled instead with linker **12** to afford the desired product (Scheme 5).

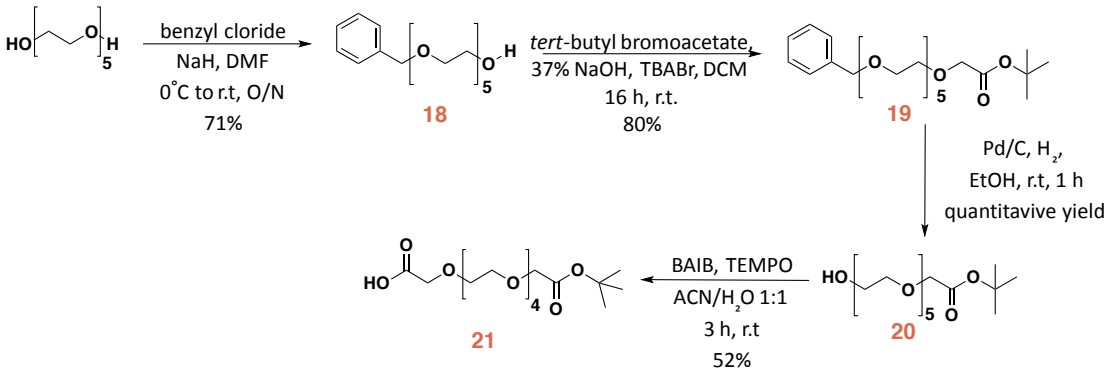


Scheme 5: Synthesis of compound **17** (CMP98).

2.2.2 Synthesis of asymmetric Homo-PROTAC CMP99

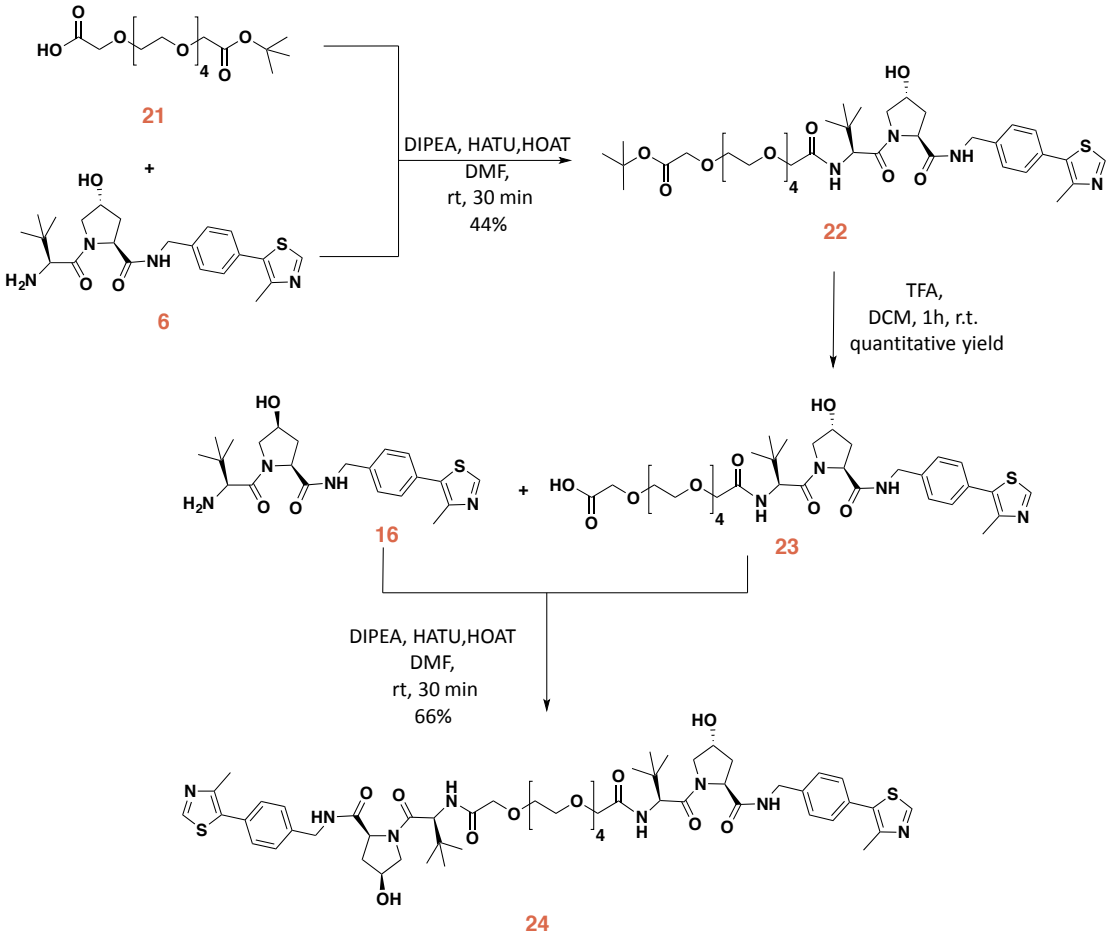
For the preparation of compound **24** (CMP99), a different route for the synthesis of the linker was established to allow subsequent controlled couplings first with **6**, followed by **16**. Pentaethylene glycol was converted in monobenzyl ether **18** in good yield (Scheme 6). For the conversion of the latter in to compound **19**, initially **18** was treated with *tert*-butyl bromoacetic acid in presence of NaH. This reaction led only to the formation of side products and no formation of the desired product was detected.

In a different approach, compound **19** was obtained in good yield by treatment of **18** with *tert*-butyl bromoacetic acid in DCM in the presence of 37% aqueous solution of NaOH and tetrabutyl ammonium bromide (TBABr). Formation of the carboxylic acid moiety was achieved by hydrogenation in the presence of Pd/C and subsequent oxidation of the free hydroxyl group by treatment with TEMPO and bis-acetoxy iodobenzene (BAIB), delivering compound **21** in 52% yield (Scheme 6).



Scheme 6: Synthesis of linker **21**.

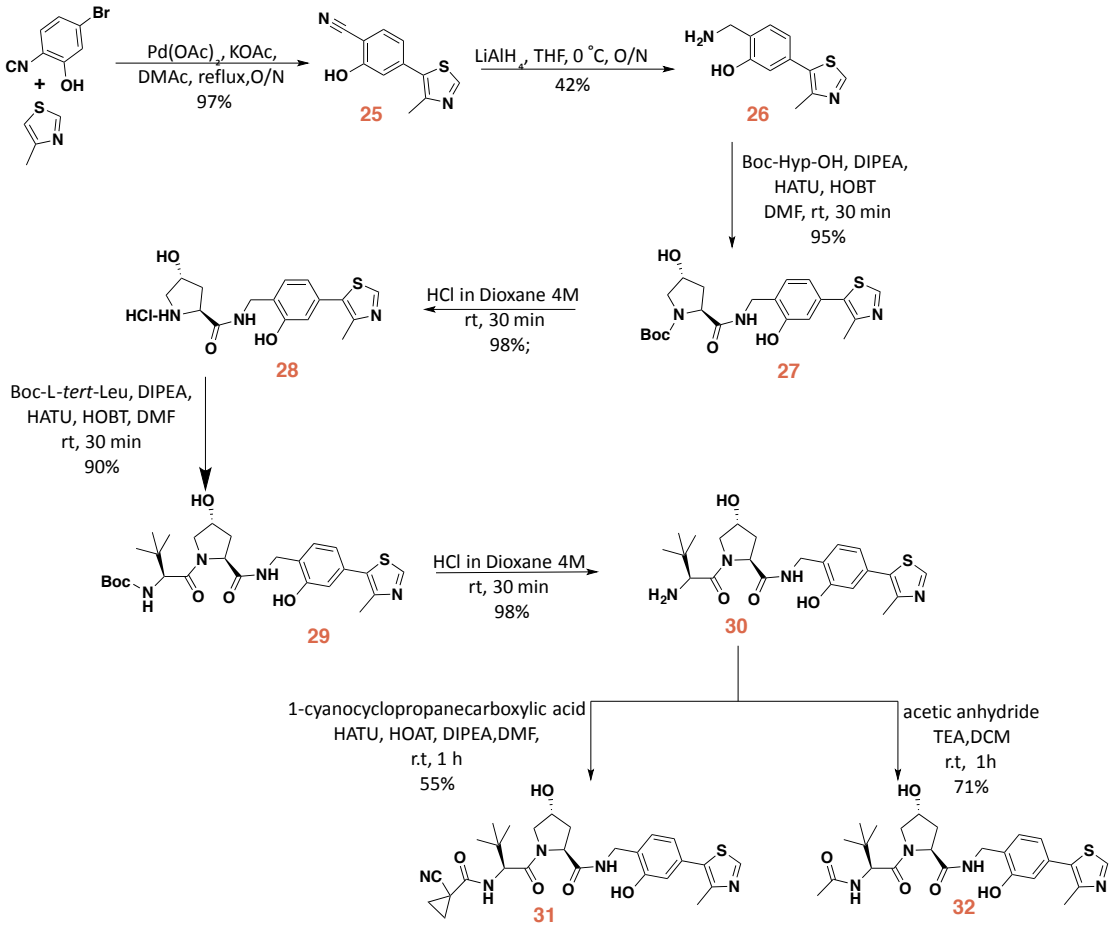
Compound **6** was coupled with linker **21** using the condition described above, affording compound **22**. The final compound **24** was then obtained after cleavage of the *tert*-butyl group using TFA and subsequently coupling with **16** (Scheme 7).



Scheme 7: Synthesis of compound **24** (CMP99).

2.2.3 Synthesis of Homo-PROTACs CMP106, CMP108, CMP112 and CMP113

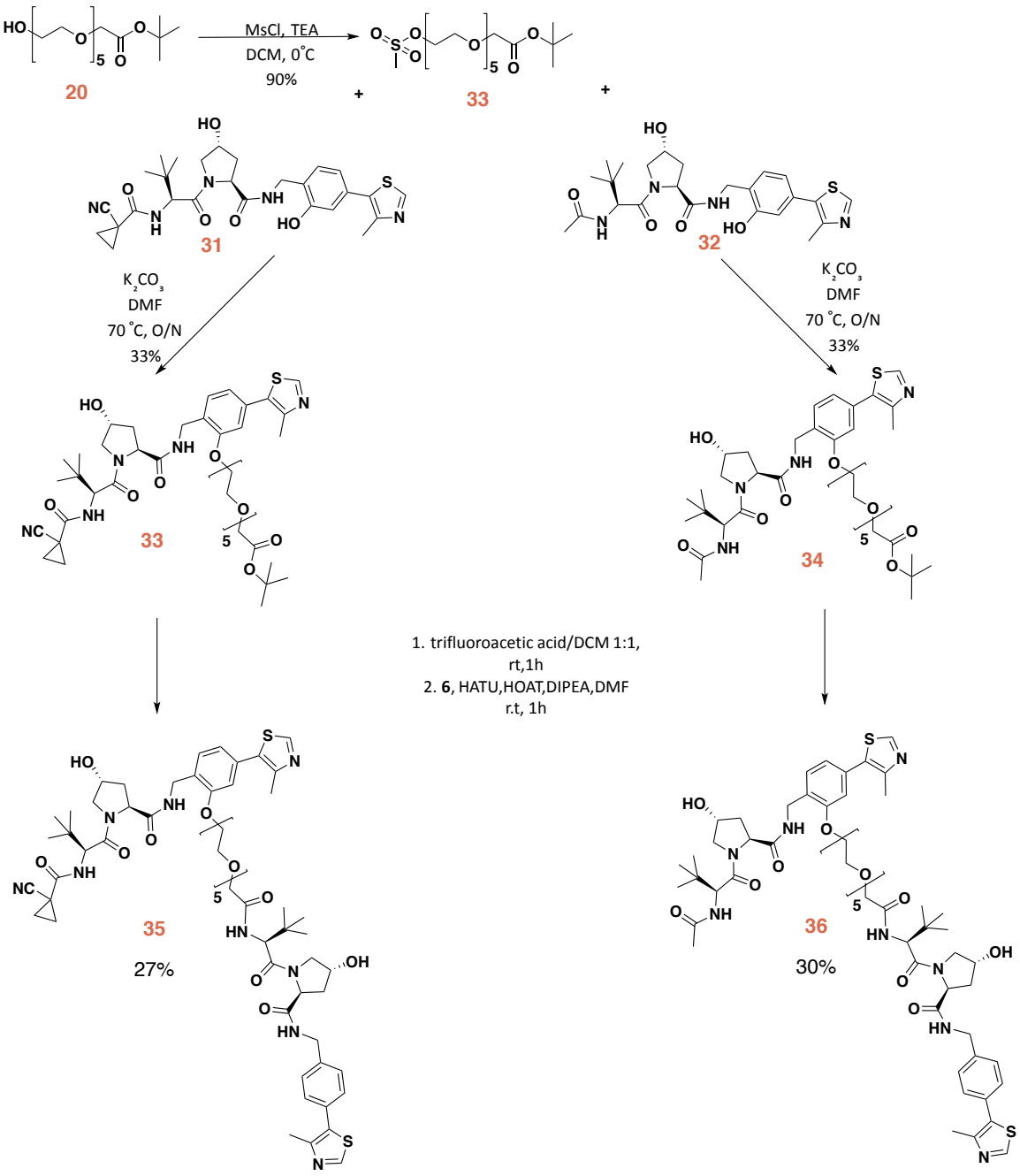
For the synthesis of these series of compounds, 4-bromo-2-hydroxybenzonitrile was used instead of 4-bromobenzonitrile as starting material. Compound **30** was synthesized following the procedure previously described (Scheme 8). The latter was treated with 1-cyanocyclopropanecarboxylic acid in the presence of coupling reagents HATU and HOAT to afford compound **31** in good yield. Instead, for the synthesis of **32,** compound **30** was acetylated using acetic anhydride and TEA in DCM at 0 °C.



Scheme 8: Synthesis of VHL ligands **31** and **32**.

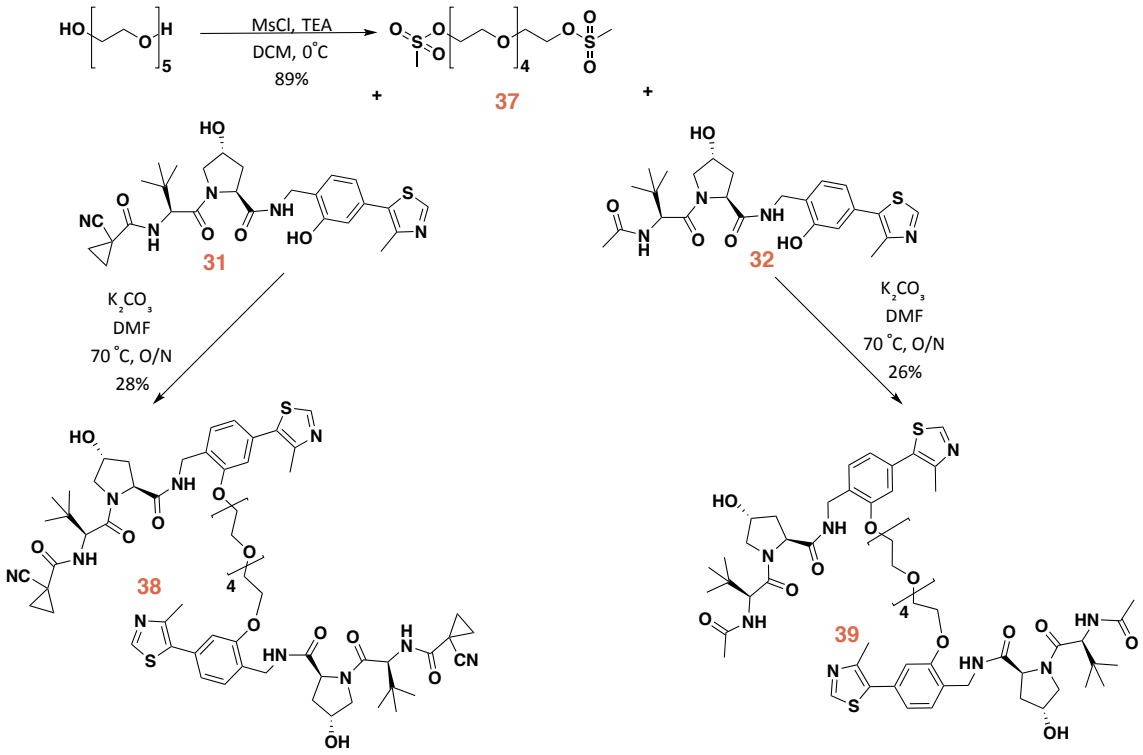
For the synthesis of the linker, **20** was converted in its mesylated derivative **33** to allow in the next step a nucleophilic attack by the phenoxy group present in

compounds **31** or **32.** These reactions led to the formation of compounds **33** and **34**, respectively (Scheme 9). Final compounds **36** (CMP112) and **35** (CMP113) were obtained after removal of the *tert*-butyl protecting group in acidic conditions and coupling of the free carboxylic acid moiety of the linker with the amine present in compound **6**.



Scheme 9: Synthesis of compounds **35** (CMP113) and **36** (CMP112).

In contrast, because of the necessity of couple the same VHL ligand, ether **31** or **32**, at both sides, for the synthesis of symmetric compounds CMP106 and CMP108, linker **37** was synthesised as a di-sulfonate derivative of pentaethylene glycol. Reaction of the **37** with compounds **31** or **32** in a 1:2 ratio led to the formation of **38** (CMP108) and **39** (CMP106), respectively (Scheme 10).

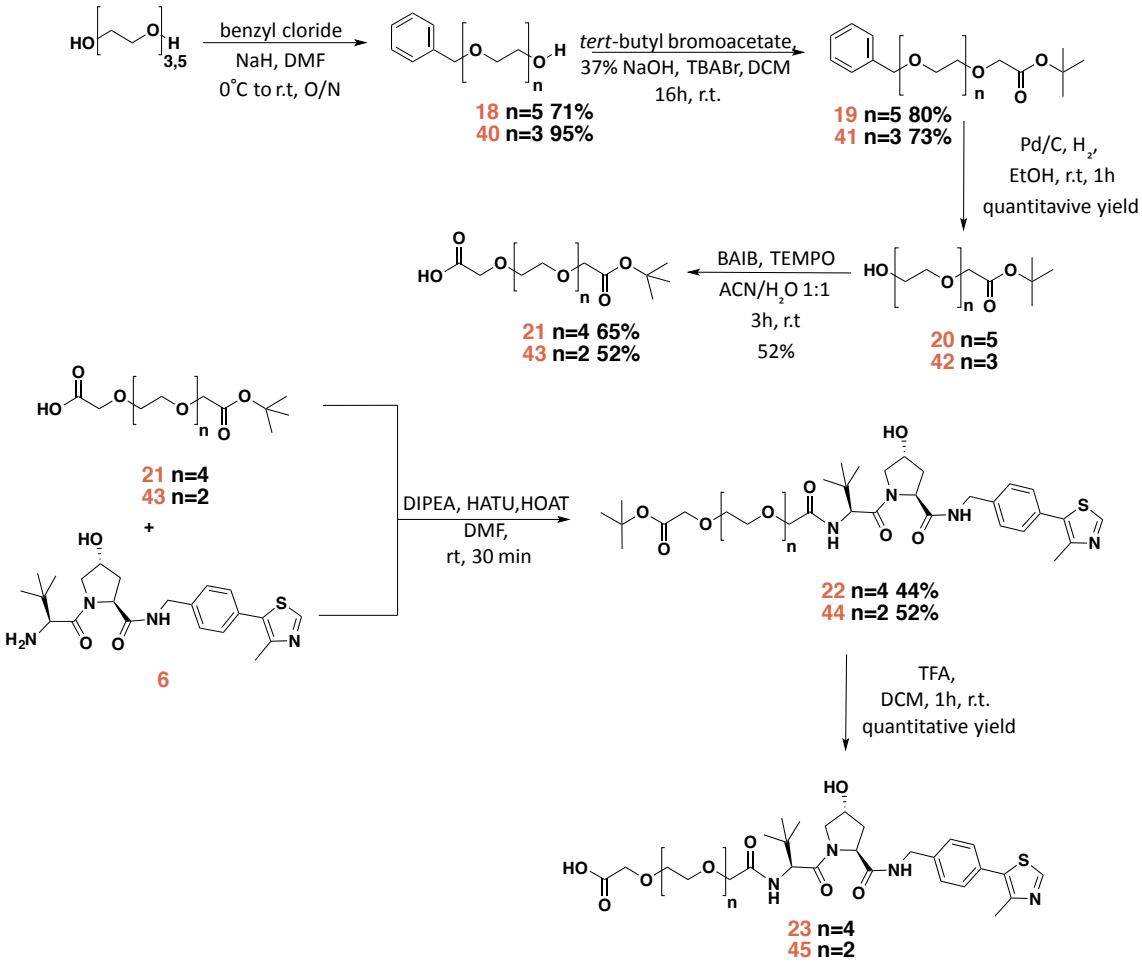


Scheme 10: Synthesis of compounds **38** (CMP108) and **39** (CMP106).

2.2.4 Synthesis of Hetero-PROTACs recruiting together CRL4^{CRBN} and CRL2^{VHL}.

For the synthesis of compounds CMP85 and CMP86 (structures shown in Scheme 12), the previously described linker **21** and its analogue with two PEG units **43** were synthesized adopting the same route used for **21** (Scheme 11). These linkers were then coupled to compound **6,** delivering compounds **22** and **44**, respectively. Subsequent deprotection of the *tert*-butyl group afforded

compound **23,** with a length of four PEG units, and **45,** with two PEG units instead (Scheme 11)**.**



Scheme 11: Synthesis of intermediates **23** and **45**.

Compound **48** (the desired thalidomide derivative, see Scheme 12) was synthesized as previously published by Lu et al[118]. In the first step, 3-fluorophthalic acid was dehydrated with acetic anhydride to obtain compound **40** in good yield. Reaction of compound **40** with L-glutamine and subsequent treatment with HCl 4 N solution led to the formation of compound **41**. Cyclization of **41** was performed at reflux in the presence of 1,1'-carbonyldiimidazole (CDI) and DMAP. The recommended time for this step was 5 h. After 2.5 h it was possible to observe the formation of a side product by LC-MS. For this reason, even if the reaction was not completed, the reaction was

cooled to r.t. and the resulting solid collected by filtration. During the purification step, performed by column chromatography over silica, compound **42** was isolated in good yield. The side product was isolated as well and analysed by NMR and identified to be compound **53** (Figure 22). Compound **53** is the product of an aromatic nucleophilic substitution at position 4 of the phthalic anhydride by the nitrogen lone pare of imidazole, which is itself a byproduct of the reaction between **47** and CDI.

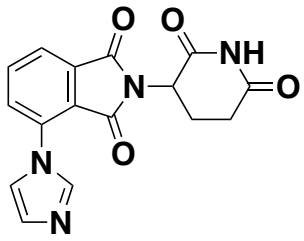
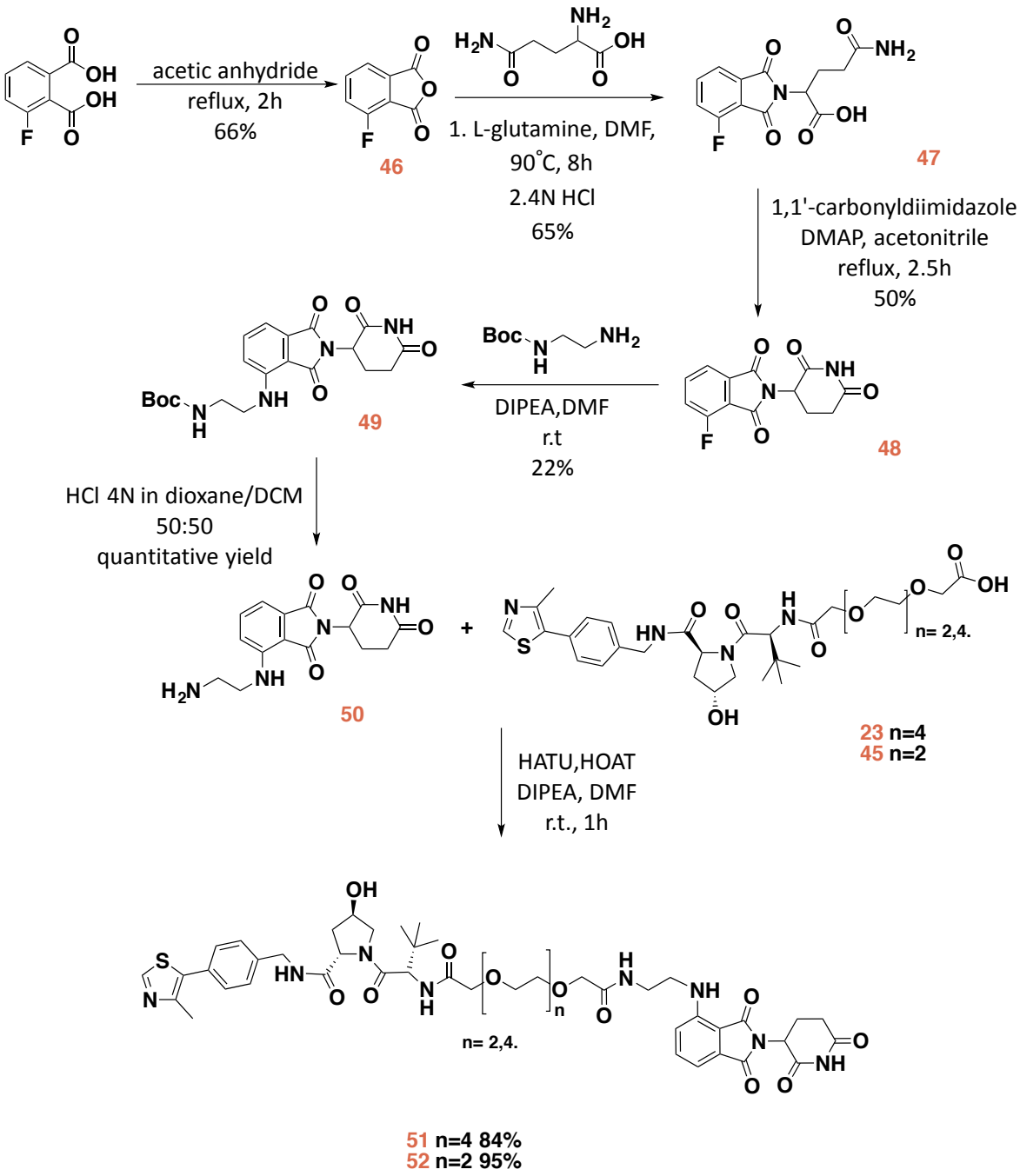


Figure 22: Side product **53** of cyclization reaction.

Compound **48** was converted in compound **50** in two steps (Scheme 12), by coupling with N-Boc-ethylenediamine and subsequent Boc deprotection in acidic conditions. Coupling of the latter with **23** or **45** afforded compounds **52** (CMP85) and **51** (CMP86) respectively in good yield.



Scheme 12: Synthesis of **52** (CMP85) and **51** (CMP86)

2.3 Biological evaluation of the VHL-targeting compounds

The following section outlines the results of the biological evaluation of both homo- and hetero-bifunctional compounds targeting VHL in cells.

2.3.1 Compounds screening: evaluation of cellular activity

In order to assess the activity of compounds inside cells, HeLa cells were treated with 1 mM of Homo-PROTACs CM09, CM10, CM11, DAT265 and the Hetero PROTACs recruiting $CRL4^{CRBN}$ to target VHL, i.e. CMP85 and CMP86. Dimethylsulfoxide (DMSO vehicle, 0.1% v/v), $CoCl_2$ (chemical inducer of HIF-1α), IOX2 and IOX4 (selective inhibitor of PHD2), VH032 (selective VHL inhibitor) were used as controls. The samples, obtained after 10 h of treatment and cell lysis, were resolved by SDS-PAGE followed by Western blot using the corresponding specific antibodies to probe for the following proteins (Figure 23):

- **VHL**: CM09, CM10 and CM11 demonstrated complete depletion of VHL levels, which featured as a preferential or selective degradation of the long isoform pVHL30. However, some degradation of the short isoform pVHL19 was also observed, albeit only around 20%. None of the other compounds were able to induce degradation of VHL.

- **Cullin2:** To assess if treatment with the series of compounds could have any effect on other subunits of the $CLR2^{VHL}$, protein levels of Cullin2 were evaluated. CM10 and CM11 showed to affect Cullin2 levels by inducing a reduction of approximately 20%.

- **CRBN:** No detectable effect was observed on CRBN levels upon treatment with CMP85 and CMP86.

- **HIF-1α and Hdy-HIF-1α:** To evaluate if VHL degraders could induce accumulation of HIF-1α, and specifically of its hydroxylated form (Hdy-HIF-1α), levels of these proteins were evaluated. It was observed during siRNA experiments (section 2.3.2) that VHL knockdown does not lead to HIF-1α stabilization. Indeed, even very low levels of VHL are capable of highly efficient catalysis on HIF-1α, leading to subsequent effective HIF-1α degradation. As expected, VHL depletion did not impact significantly on HIF-

1α level (compare the detected HIF-1α band with vehicle control). Nevertheless, a slight increase of HIF-1α level was induced by the active VHL degraders CM09, CM10 and CM11 (see HIF-1α band with longer exposure). This effect was even more pronounced on Hdy-HIF-1α, consistent with the stabilized HIF being in the hydroxylated form as expected from VHL knockdown.

- **PHD2 and PHD3:** to study potential hypoxic response of cells due to treatment with the compounds, levels of PHD2 and PHD3 were considered. No effect on the levels of these proteins was observed at this concentration.

The same experiments were performed in other cells lines to further assess the consistency of the cellular effects of our compounds, as different cell lines can have different expression levels of different proteins. For example, HEK293 are known to have higher expression levels of total VHL, which we confirmed by Western blot (Figure 23). The same activity profile in decreasing preferentially pVHL30 levels by CM09–11 was observed in HEK 293 (Figure 23). No major effects were observed on levels of the other proteins. Experiments conducted in U2OS cells showed the same results, confirming that the effect observed upon treatment with CM09, CM10 and CM11 is independent from cell type and it is consistent in all tested cell lines (Figure 23).

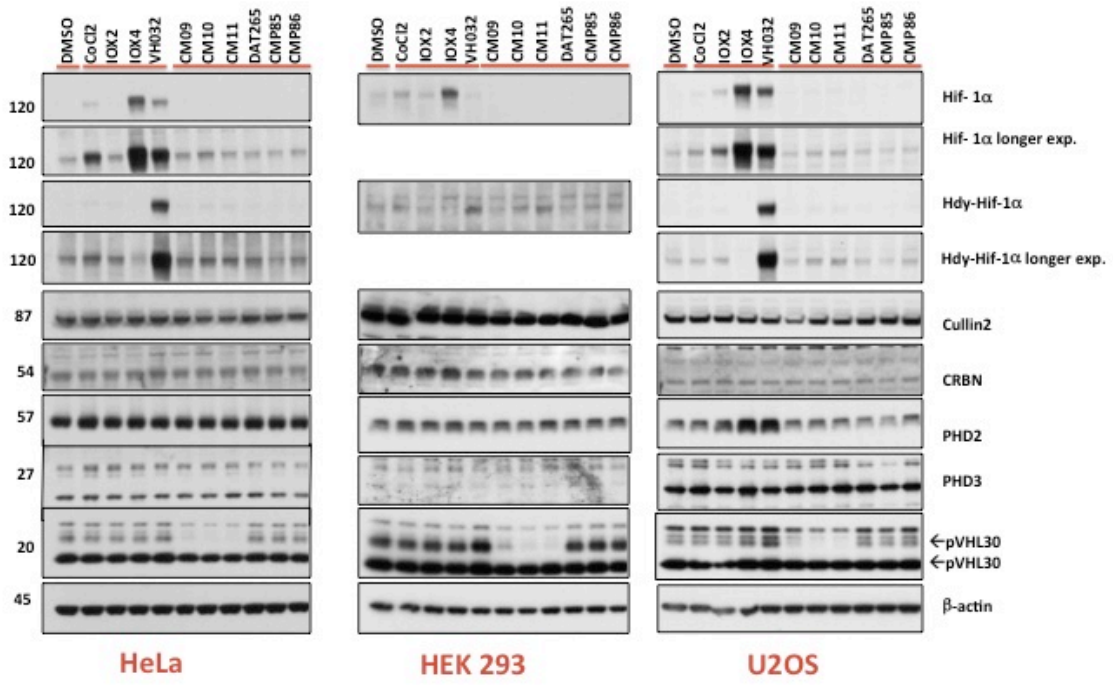


Figure 23: HeLa, Hek293 and U2OS cells were treated with 1 μM of CM09, CM10, CM11, DAT265, CMP85 or CMP86, 0.1% DMSO, $CoCl_2$ (100 μM), IOX2 (50 μM), IOX4 (50 μM) or VH032 (250 μM) for 10 h.

2.3.2 VHL siRNA experiment

To confirm the reliability of the VHL antibody, and that the bands observed decreasing in intensity are indeed for VHL, siRNA experiments was performed in HeLa, HEK 293 and U2OS cells. The silencing of VHL protein was achieved after 48 h of treatment with the siRNA. The samples obtained after cell lysis where evaluated by western immunoblotting, levels of VHL protein were monitored by incubation with the antibody used in the previous experiment and no bands relative to pVHL were found (**Figure 24**). This result demonstrates that the band we assumed to be related with pVHL was indeed the right one, and confirms that our compounds were working as expected by degrading VHL protein inside cells. Moreover, it was noticed that stabilization of Hdy-HIF-1α could be achieved by using CM11, but not with the siRNA. It is possible that this could be due to the longer treatment times (48 h) of the siRNA relative to the compound (10 h), as HIF levels are known to decrease significantly after 12-24 h even under

hypoxia[121]. Another intriguing possibility could be that another system in addition to VHL could be contributing to Hdy-HIF-1α degradation under these conditions.

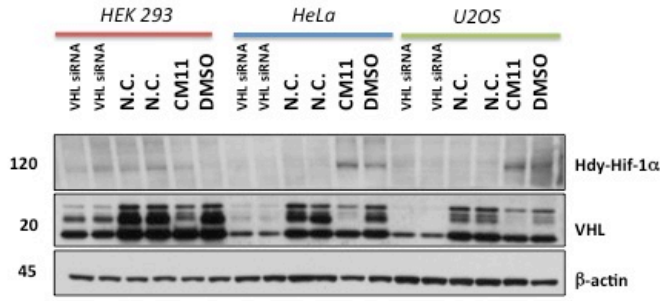


Figure 24: Different cells lines were treated with si-RNA targeting VHL proteins or negative control si-RNA (for 48 h), as well as with CM11 (1 μM) or 0.1% v/v DMSO for 10 h.

2.3.3 PROTACs-mediated degradation and influence of linker position on compound activity

To determine if the effect on VHL levels was dependent on PROTAC-mediated degradation and if the position of the linker could have an impact on the activity of the compounds, HeLa cancer cells were treated for 10 h with 1 μM of Homo-PROTACs CM09, CM10, CM11, negative control CMP98 (*cis-cis* homologue of CM11, expected to be a fully inactive epimer), CMP99 (*cis-trans* homologue of CM10, expected to act as mono-functional VHL ligand but not as VHL degrader) and CMP113, CMP112, CMP108, CMP106. DMSO (0.1%, vehicle) and VH032 (VHL inhibitor) were used as controls.

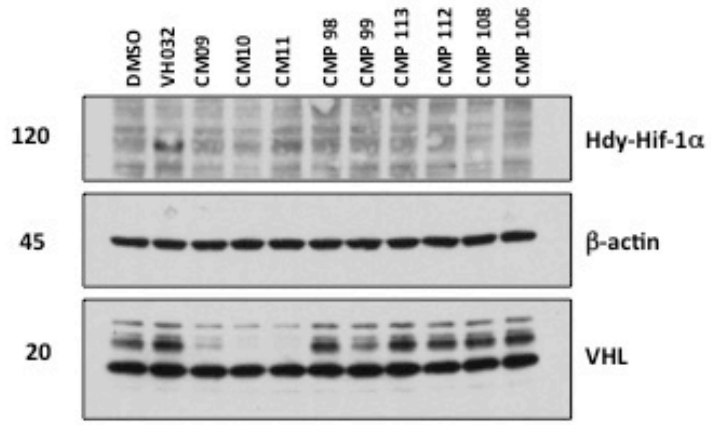


Figure 25: Treatment of 0.1 % DMSO, VH032 (150 μM) and 1 μM of the indicated compounds in HeLa cells for 10 h.

CMP98 and CMP99 were unable to induce degradation of any of the VHL isoforms, demonstrating that compound efficacy is dependent on productive recruitment of VHL. However, unexpectedly a small effect on the middle VHL band was observed with CMP99 (Figure 25). Compounds CMP113, CMP112, CMP108, CMP106 also were not able to induce target degradation (Figure 25). Together, these results demonstrate how important is the design step and highlight the impact of the position for the linkage of the two moieties on the compound cellular activity.

2.3.4 Concentration dependency experiment

To assess compounds dose- dependent activity, HeLa cells were treated with different concentrations of active VHL degraders CM09, CM10, CM11 at two different time points, 4 h and 24 h prior harvesting.

All three compounds showed to degrade preferentially one isoform of VHL in a concentration-dependent manner, with higher activity at higher concentrations (**Figure 26**). A 10–20% level depletion of the short isoform was still detectable, in particular for CM11. CM11 also proved to be the most active compound, with a total depletion of pVHL30 at 1 μM and more than 40% of pVHL20 removed at concentration below 1 μM.

The three compounds also showed to stabilize Hdy-HIF-1α in a concentration-dependent manner (**Figure 26**). HIF stabilization was particularly pronounced at the higher micromolar concentrations, at which compounds are expected to act preferentially as inhibitors over degrader due to the so-called "hook effect". This effect is often observed at high PROTAC concentration, whereby formation of binary complexes (PROTAC:target and PROTAC:ligase) competes with and eventually exceeds the formation of the productive ternary complex[119]. Indeed, stabilization of Hdy-HIF-1α upon treatment with all the three compounds at 100 μM was comparable with the effect obtained with inhibitor VH032 (**Figure 26**). However, CM11 showed to stabilize Hdy-HIF-1α level at even at nanomolar concentrations. Consistent with this effect, a slight decrease in Cullin2 concentration was observed only upon treatment with CM11.

An effect on PHD2 and PHD3 levels was also observed. In fact, the level of these proteins was busted up on treatment with the three compounds in a concentration-dependent manner, testifying a hypoxic response in the cells.

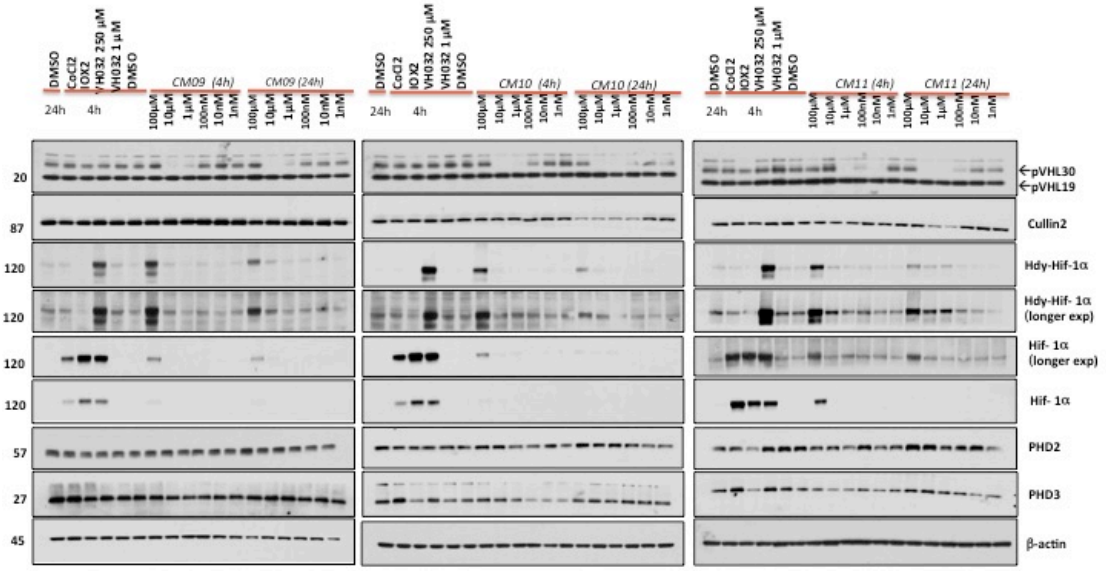


Figure 26: HeLa cells were treated with increasing concentration of indicated compound for 4 h or 24 h.

To confirm the observation made in the previous experiment, a similar experiment was performed treating U2OS cells for 10 h at the same

concentrations used previously, and a consistent activity profile was observed (Figure 27).

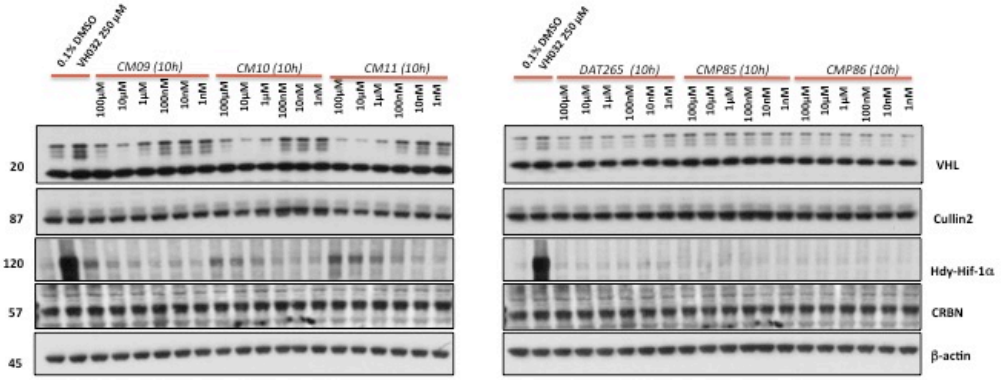


Figure 27: Concentration dependency experiment in U2OS (10 h treatment).

To assess if the non-detectable activity of compounds DAT265, CMP85 and CMP86 could potentially be due to the single concentration tested (1 μM, Figure 27), U2OS cells were treated with the three compounds for 10 h at concentrations ranging from 1 nM to 100 μM. No effect on the target protein level was observed, even at the highest concentrations used, consistent with these compounds being inactive inside cells, either as inhibitors or as degraders (Figure 27).

2.3.5 Time dependency intracellular activities

To further evaluate the activity of VHL degraders over time, HeLa cells where treated with CM09, CM10 and CM11 at 1 μM and levels of different proteins were monitored in a time course experiment. DMSO (0.1% v/v, vehicle), $CoCl_2$ (chemical inducer of HIF-1α), IOX2 and IOX4 (selective inhibitor of PHD2), VH032 (selective VHL inhibitor) and untreated cells were used as controls.

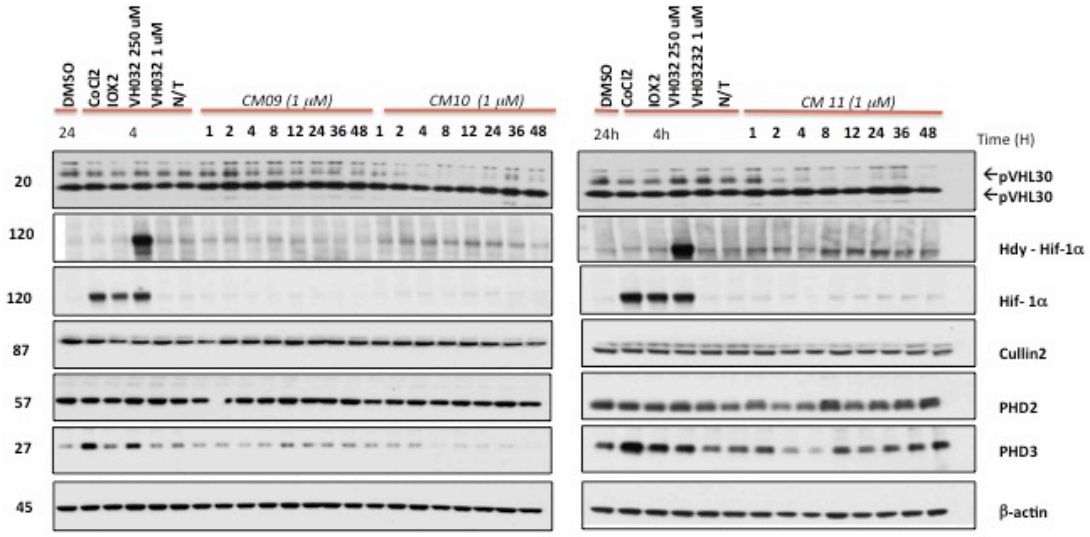


Figure 28: Time-course immunoblots of lysates from HeLa cells subjected to 0.1% DMSO, $CoCl_2$ (100 μM), IOX2 (150μM), VH032 (250 μM or 1 μM) or 1 μM of indicated compounds.

Progressive removal of VHL protein over time was observed and pVHL30 consistenly exibited the strongest and fastest reduction in protein level (Figure 28). No degradation was observed in the presence of any of the controls used. In particular, CM10 and CM11 were confirmed to be the most effective compounds, decreasing pVHL30 level by more that 90% already after 2 h of treatment. The depletion effect was retained up to 12 h, however interestingly protein level bounced back up after 24-36 h, to then decrease again after 48 h. We hypothesize that two effects synergistically account for this effect. First, at the very low level remaining after 12 h, VHL no longer efficiently catalyzes its own ubiquitination and degradation induced by the PROTAC. As a result, protein re-synthesis overcomes protein degradation. It is also possible that cells react to the induced depletion of pVHL protein by boosting its re-synthesis as feedback mechanism. Once VHL levels are restored close to normal, due to the catalytic mechanism of action of PROTAC compounds and because the compounds are still present inside cells, VHL levels are depleted again, and no pVHL can be detected at 48 h. Incomplete degradation of pVHL was observed upon treatment with CM09, even in the longer time points (Figure 28). Stabilization of Hdy-HIF-1α over time was

observed for all three compounds, most pronouncedly upon treatment with CM11.

To validate the observed effects in a different cell line, a second time-course experiment was performed in U2OS cells treated with CM09, CM10 and CM11 at 1 μM concentration at the desired time points. The results obtained were consistent with what observed in the previous experiment. However, in this cell line all the three compounds were able to induce complete degradation of pVHL over time (Figure 29). We hypothesize that this could be due to the lower expression level of VHL in U2OS, leading to faster cellular depletion compared to cell lines where VHL level is higher. CM09 and CM10 achieved complete degradation of the target protein after 2 h of treatment. CM11 again proved to be the most potent compound, achieving complete degradation of pVHL already after 1 h. Interestingly CM09 lost its cellular efficacy after 36 h. In constrast, both CM10 and CM11 retained their efficacy even at longer time points (Figure 29).

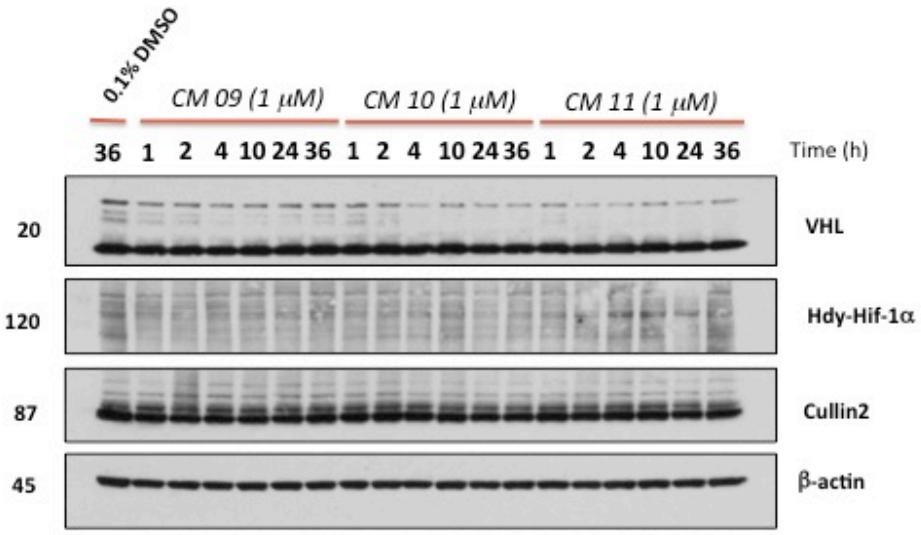


Figure 29: Time course experiments of lysate from U2OS.

2.3.6 VHL and proteasome dependency of PROTAC-mediated protein degradation – competition experiments.

To gain additional mechanistic insights, the $CRL2^{VHL}$ and proteasome dependency of Homo-PROTAC mediated degradation was examined. The

reliance of the PROTAC-induced protein degradation on $CRL2^{VHL}$ was assessed by inhibiting neddylation of Cullin2 using the NAE1 inhibitor MLN4924 (see Section 1.2.1 in the Introduction). Because neddylation activates Cullin RING E3 ligase enzymes (CRLs), the consequence of NAE1 inhibition will be the blockade of the function of all CRLs, including $CRL2^{VHL}$. Proteasome-dependency of the activity of these compounds was interrogated by treating cells with the proteasome inhibitor MG132 (see Section 1.2.6 in the Introduction).

To limit the known toxicity of MLN4924 and MG132 compounds in this experiment, HeLa cells where pre-treated with MLN4924 for 3h and MG132 before adding CM11 to the media. The plates were incubated for further 4h before harvesting. Single treatments with DMSO, MLN4924, MG132 and CM11 and combinations thereof were performed to disentangle the individual and combined effects of compounds. Degradation of pVHL induced by CM11 was completely abrogated when cells were pre-treated with MG132, establishing the expected proteasome-dependence of the chemical intervention (Figure 30). Degradation of the target protein by CM11 was overturn by pretreatment with MLN4924, confirming the dependency on the $CRL2^{VHL}$ degrading machinery (Figure 30). The same effect was observed when cells where co-treated with MLN4924, MG132 and CM11 (Figure 30).

To assess if the ability of CM11 to degrade the target protein could be blocked by pre-treatment with the VHL inhibitor VH032, a competition experiment was performed. HeLa cells were pre-treated with VH032 at 150 μM for 30 min before adding CM11 into the media. The plates were incubated for 4 h before harvesting. As expected, VHL inhibitor VH032 was able to block pVHL degradation (Figure 30), consistent with the hypothesis that VHL induces degradation of itself. In contrast, pre-treatment with IOX4, a PHD2 inhibitor, did not impact the cellular activity of CM11 at all (Figure 30).

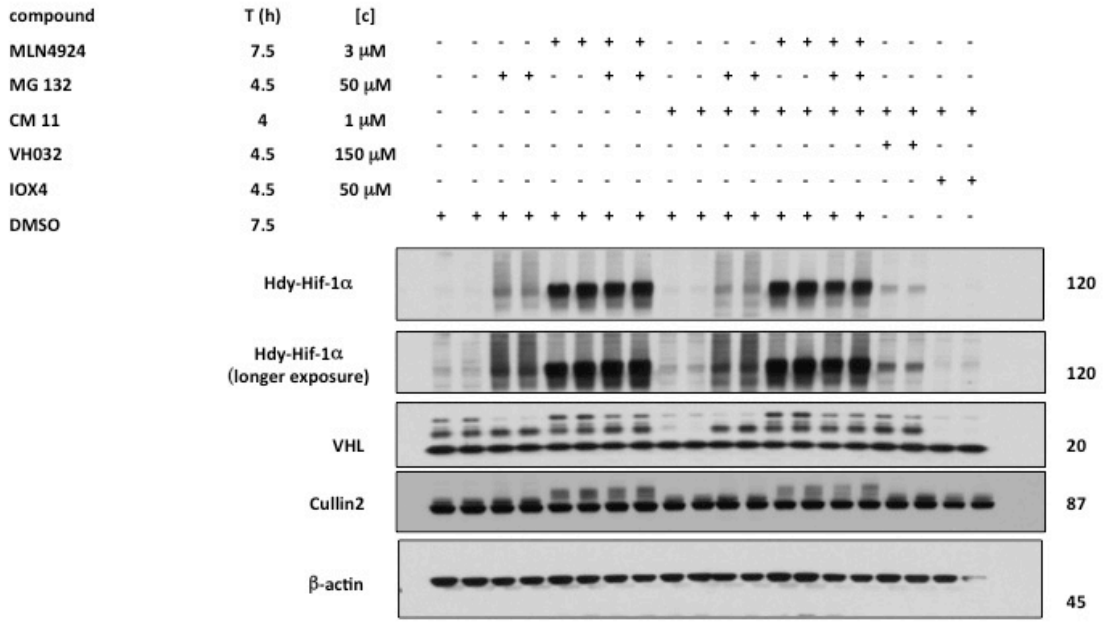


Figure 30: Hela cells treated with CM11 in the absence or presence of proteasome inhibitor MG132, MLN4924, VHL inhibitor VH032 or PHD2 inhibitor IOX4.

2.4 Biophysical assays

As explained previously (Section 1.4.3), PROTACs compounds induce protein degradation by bringing the target protein in close proximity to an E3 ubiquitin ligase, in our case VHL and $CRL2^{VHL}$ respectively, promoting the formation of a ternary complex VHL-PROTAC-$CRL2^{VHL}$. It is the actual formation of this complex that we believe is the key process in the functioning of this system.

To assess the formation of this ternary complex species from a biophysical point of view, two different techniques were employed: isothermal titration calorimetry (ITC) and size exclusion chromatography (SEC).

2.4.1 Isothermal titration calorimetry

To determine the thermodynamics of binding and complex formation, the interaction of VHL with PROTACs CM11, CMP98 and CMP99 was characterized using ITC. ITC is a calorimetric technique that measures the energetics of

biochemical reactions or molecular interactions, such as ligand-binding phenomena, at constant temperature. The heat released or absorbed by the stepwise addition of a ligand molecule to a solution containing the macromolecule under study is then monitored. Thermodynamic analysis of the observed heat effects permits quantitative characterization of the energetic process associated with the binding reaction. In ITC, the degree of saturation is defined in terms of the heat associated with the reaction. The goal is to generate a binding isotherm, i.e. a curve that represents the degree of saturation in terms of the ligand/protein molar ratio.

The pVHL protein, routinely expressed and purified from *E. coli* as complex with EloBC following a protocol established in Professor Ciulli's laboratory, was a gift from Dr. Scott Hughes. PROTACs compounds were titrated individually into a solution containing the VHL protein during different experiments. Data recorded was then fitted to a single-site binding model as detailed in the Methods (section 6.3.1) (Figure 31). Thermodynamic parameters measured during each experiment are summarised in Table 3. Previously published ITC data regarding VH032 (ref. [115]) were used as a reference to compare and contrast the results of our bifunctional compounds.

Compound	*n*	K_d (nM)	ΔG (kcal/mol)	ΔH (kcal/mol)	–TΔS (kcal/mol)
VH032 (ref. [115])	1.030 ± 0.001	188 ± 6	–9.30 ± 0.02	–5.53 ± 0.01	–3.65 ± 0.02
CM11	0.605 ± 0.002	8 ± 2	–11.0 ± 0.2	–11.6 ± 0.1	0.5 ± 0.2
CMP99	0.964 ± 0.005	146 ± 2	–9.33 ± 0.06	–6.23 ± 0.05	–3.1 ± 0.7

Table 3: Summary of biophysical data of PROTACs CM11 and CMP99 and comparison with inhibitor VH032 (from [115]).

In the ITC experiment in which Homo-PROTAC CM11 was titrated against VHL protein, the stoichiometry of binding (*n* value) was found to be equal to 0.6. This result is consistent with a scenario where CM11 binds to VHL in a 1:2 molar ratio, in contrast to the VHL inhibitor VH032 that binds to VHL in a 1:1 ratio.

The K_d value measured for CM11 was 8 nM. Closer examination of the titration curve revealed that only one point features during the inflection of the curve. Indeed, because we are using a protein concentration of 20 μM, the *c* value (defined as $[P]_{tot}/K_d$) calculated for this experiment is 2500, which is well above the upper limit of *c* (around 500-1000) that is a prerequisite for precise measurement of binding affinity[164]. Consequently, this analysis suggests that we maybe underestimating the binding affinity, i.e. we can conclude that K_d is ≤ 8nM.

The binding interaction between CM11 and VHL was driven by a large apparent binding enthalpy (ΔH = −11.6 kcal mol^{-1}), whereas the entropic term was unfavourable (−TΔS = 0.5 kcal mol^{-1}). This observation underlines how the thermodynamic signature of CM11 is also very different when compared with that of the inhibitor VH032, in which case both the enthalpic and entropic term contributed favourably to the ΔG of binding. By contrast, the thermodynamic values obtained for CMP99 were consistent with the ones of VH032 (Table 3). In particular, CMP99 bound to VHL in a 1:1 ratio, as expected due to the presence of the *cis*-Hyp in one of the two moieties, which is known to prevent the ligand to bind to the protein. CMP98, the inactive *cis*-*cis* epimer, did not show any binding, as expected.

Superposition of integrated heat curves of CM11, CMP98 and CMP99 is shown in Figure 32 and visually highlights the different behaviours of the compounds.

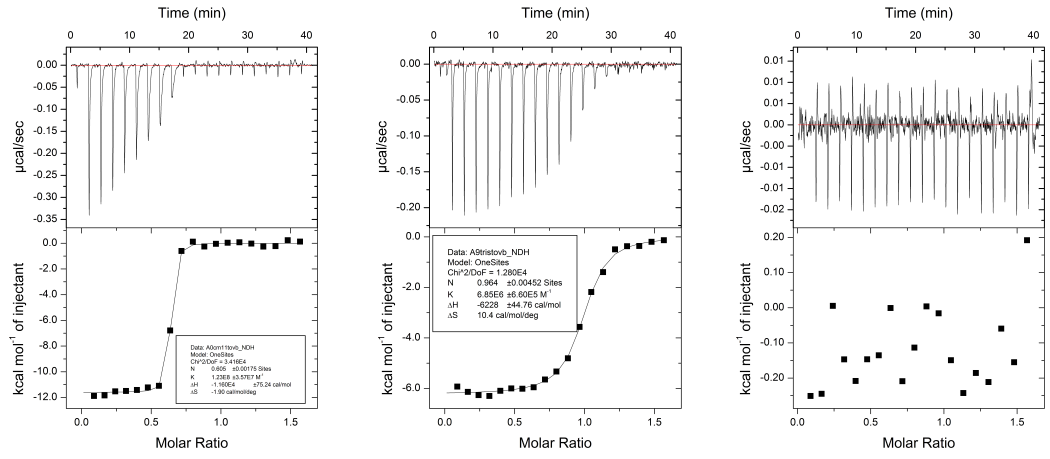


Figure 31: ITC curves obtained titrating CM11 (left), CMP99 (middle) or CMP98 (right) in to VHL protein.

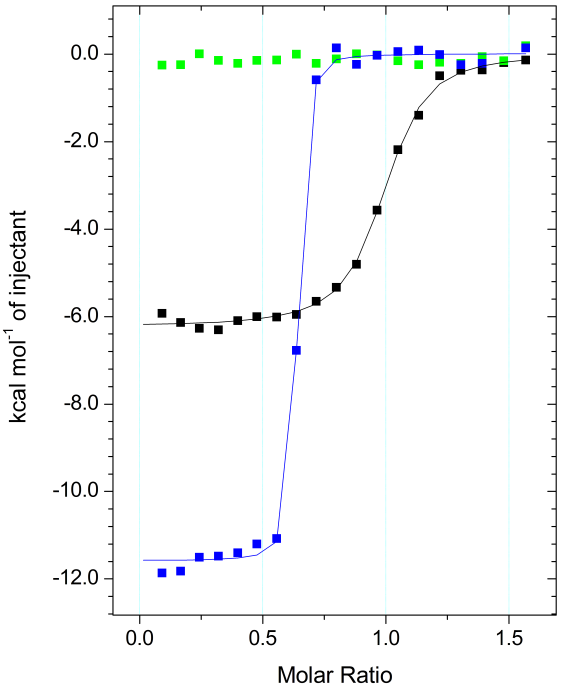


Figure 32: Superposition of the integrated heat curves of CM11 (blue), CMP99 (black) and CMP98 (green).

2.4.2 Size exclusion chromatography

SEC is a chromatographic method in which large protein molecules or their macromolecular complexes are separated by their size. A column made of adsorbent material (stationary phase) is used to separate molecules by trapping smaller molecules in the pores of the column. Instead, larger molecules simply pass through the pores because, due to their large dimensions, they cannot enter the pores. So larger molecules are not retained by the column, flow through more quickly than smaller molecules, and are thus characterized by shorter retention times. Instead smaller molecules are characterised by longer retention times.

In first instance, VBC protein mixed with 0.25% of DMSO was injected in the column in order to establish a reference point at which it should be expected to elute the protein alone. VBC eluted at a retention volume of 15.7 ml (black curve, Figure 33). Thereafter, to identify the correct concentration at which to perform

the experiment, VBC protein (50 μM) was mixed initially with CM11 at a concentration of 25 μM (Figure 33). A first peak eluted at 13.3 ml, which is around the expected elution volume for a complex of the size of the ternary complex VBC:CM11:VBC. A second peak eluted a 15.3 ml, possibly due to an excess of VBC that was not involved in the formation of the ternary complex (green curve, Figure 33). In a second attempt, by increasing CM11 concentration to 30 μM in was possible to achieve a single peak that eluted at the expected elution point, with no residual unbound VBC (red curve, Figure 33).

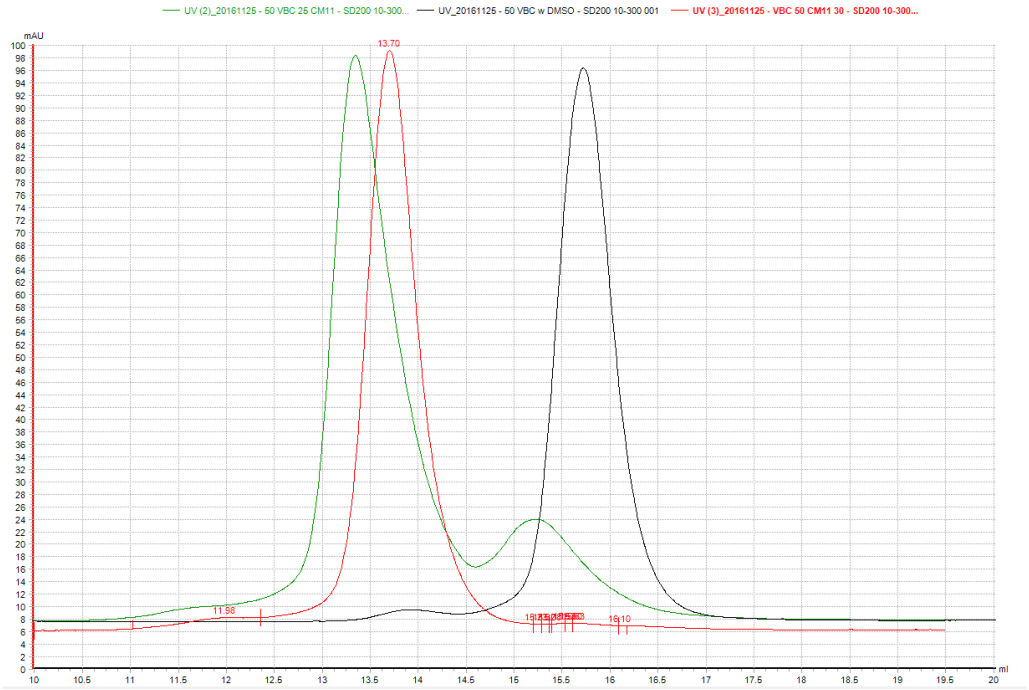


Figure 33: Superposition of curves obtained during SEC experiment: VBC+DMSO (black), VBC + 25 μM of CM11 (green), VBC + 30 μM of CM11 (red).

Once the optimal concentrations of components were identified, we compared the peaks obtained by mixing VBC protein with: CM11, the most active compound of the Homo-PROTACs series; CMP98, the *cis-cis* analogue of CM11 that showed to be totally inactive in cells; CMP99, the *cis-trans* analogue of CM10 that showed not to affect VHL levels in cells; and VH032, the VHL ligand used to develop this series of compounds (Figure 34).

The active compound CM11 shifted the VBC monomeric peak to a more quickly migrating dimeric peak when the protein was mixed with the compound in a 1:2 ratio. In contrast, the curves belonging to VBC incubated with inactive CMP98, CMP99 or ligand VH032 did not shift the monomeric peak. Only in the sample containing CMP99 a small peak eluted at 13.5 ml (purple curve, Figure 34). It is possible that such peak could be due to the formation of a small amount of ternary complex. This could explain the small VHL reduction observed during biological tests in cells (section 2.3.3, Figure 25).

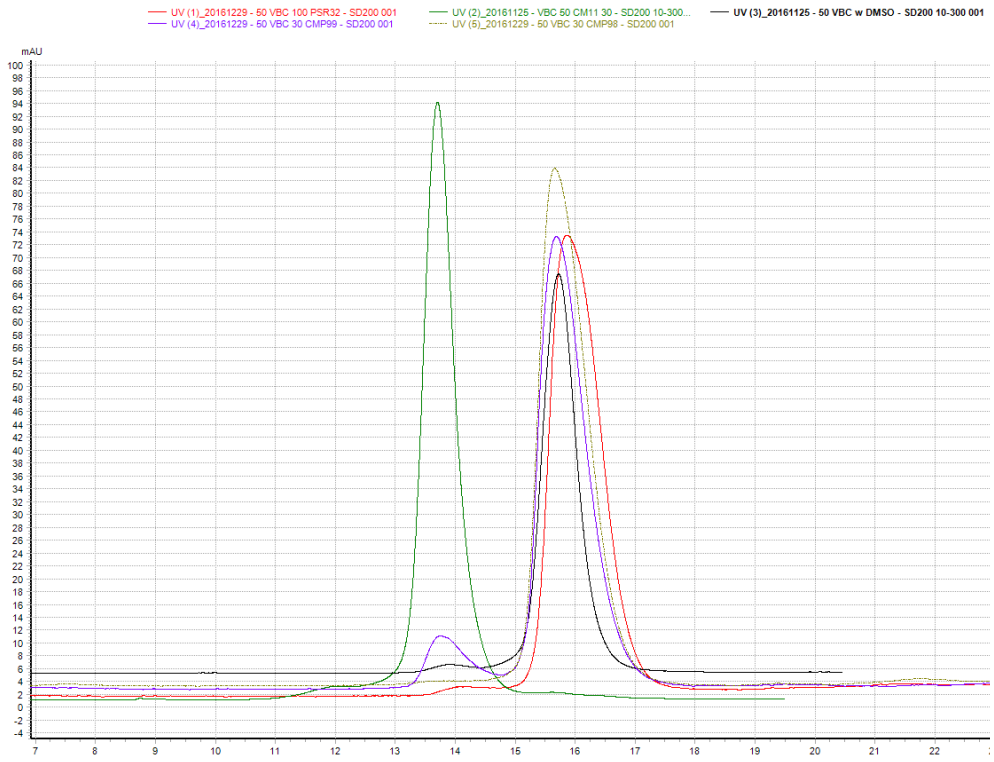


Figure 34: SEC assay of complex formation after incubation of CM11 (green), CMP98 (pale green), CMP99 (purple), VH032 (red) or DMSO (black) with VBC.

The optimized SEC assay was repeated in the presence of all remaining VHL-targeting PROTAC compounds. CM10 and CM09 showed the same ability to form a dimeric peak of identical retention time when compared to CM11 (Figure 35). A similar dimeric peak was also obtained by mixing VBC with the asymmetrically linked compound CMP113 (brown curve, Figure 35). This result suggests that this compound is also able to form the ternary complex, however interestingly the

compound was inactive in cells (section 2.3.3, Figure 25). We speculate that this complexation may not be sufficient to induce degradation of the target protein, probably suggesting that the relative orientation of the two proteins is also crucial to the effective ubiquitination mechanism.

DAT265, the Homo-PROTAC compound with the rigid linker, was not able to form a dimeric peak (purple curve, Figure 35). These data are consistent with the results obtained from biological studies, which showed the compound to be inactive in cells.

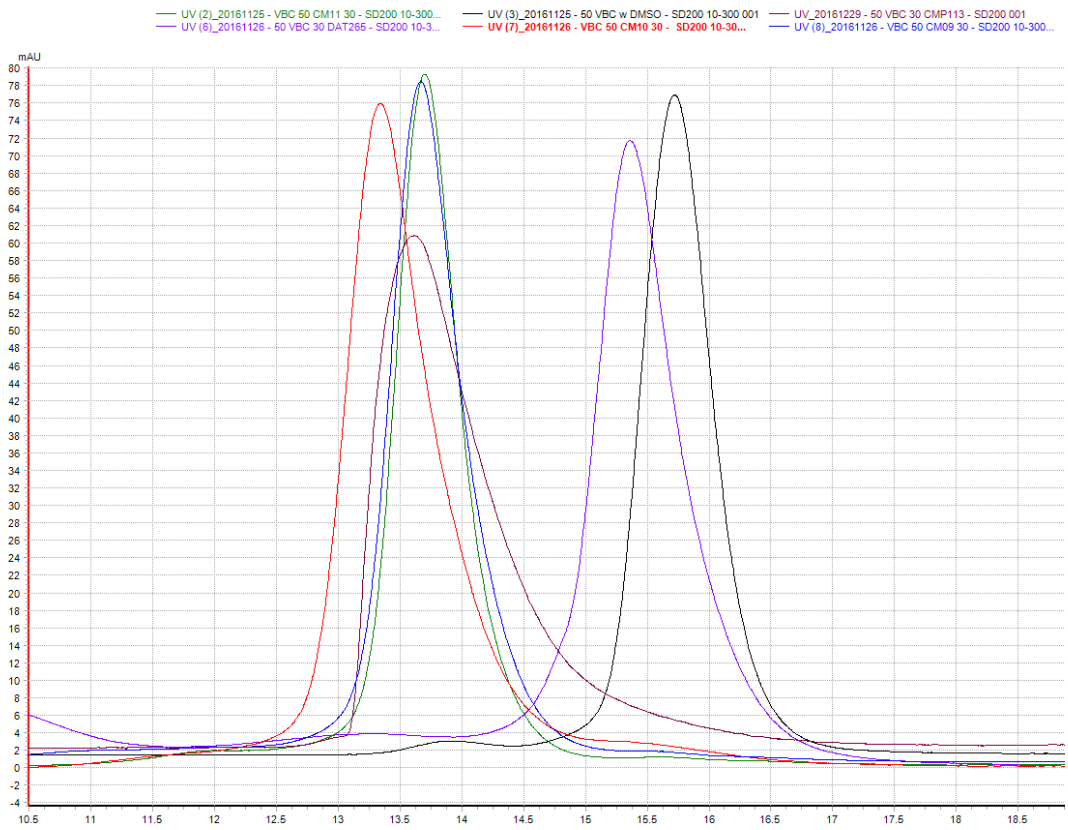


Figure 35: SEC assay of complex formation after incubation of CM11 (green), CM10 (red), CMP99 (blue), CMP113 (brown) or DAT265 (purple) with VBC. Protein mixed with DMSO is showed in black.

2.5 Summary

In summary, this section reports the discovery of novel homo-bifunctional PROTACs that effectively and selectively degrade the long isoform of VHL in cells.

The most active molecules also showed remarkable avidity for dimerizing VHL in biophysical assays using recombinant proteins. These features were very sensitive to the positioning of the linker, its chemical nature and its length, with most potent compound being symmetric homo-PROTAC CM11. Attempts to recruit VHL and CRBN to potentially induce degradation of either or both proteins proved unsuccessful.

3 PHD-targeting compounds

In this chapter, the design, synthesis and biological evaluation of PHD-targeting bifunctional molecules will be described.

3.1 Design of PHD targeting compounds

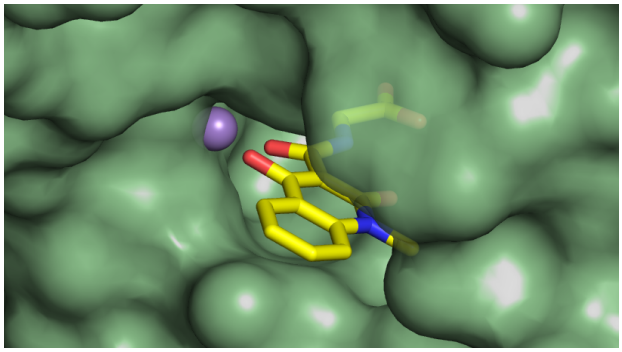
The PHD enzymes, PHD1, PHD2 and PHD3 are amongst the key players of the hypoxia signalling pathway (section 1.5.2). We therefore became interested in the opportunity to degrade these targets using a PROTAC approach. Such pharmacological intervention would be expected to achieve stabilization of HIF-1α and consequently induce a hypoxic response. As explained before, PHD enzymes are paralogs in a given organism e.g. humans, and the inhibitors commercially available, i.e. IOX2 and IOX4, are broadly pan-selective against the three paralogs, meaning they inhibit all three enzymes with comparable potencies. Previous work within the Ciulli group had shown how it is possible to turn a pan-selective inhibitor in more selective single-target degraders using the PROTAC approach[116]. Taking inspirations from these works, it was decided to apply this approach also on PHD inhibitors and thus develop two classes of compounds based on the structure of the two inhibitors IOX2 and IOX4.

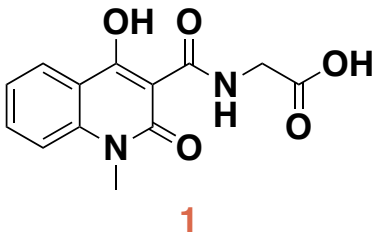
3.1.1 Hetero PROTACs based on IOX2 ligand

To rationally design this class of compounds, the chemical structure of IOX2 ligand was chosen as starting point. Careful consideration was required in choosing the position for the linking on the IOX2 scaffold in order not to disrupt its interaction with the target protein. For this reason, the crystal structure of an analogue of IOX2 ligand (**1**) bound to PHD2 (Figure 36)[154] was analysed to identify solvent-exposed regions from where the ligand could be derivatized without perturbing its binding mode.

The bicyclic heteroaromatic rings are set between the hydrophobic side chains of of Tyr310, Met299 and Trp389. The molecules also coordinate the key Fe^{2+} metal ion in a bi-dentate fashion through two oxygens atoms present in the ligand

structure (Figure 36). The side chain carboxylate is positioned to hydrogen bond with Arg383 and Tyr329.





1

Figure 36: Crystal structure of an analogue of the IOX2 derivative **1** bound to the PHD2 catalytic domain.

Based on this analysis, none of the previous mentioned positions were deemed suitable for derivatization due their involvement in the binding interactions of the molecule with the protein. Comparison of the co-crystal structures of the inhibitor and HIF-1α CODD (residues 556–574) each bound with PHD2 shows that inhibitor binding likely blocks productive binding of the conserved LXXLAP motif of HIF-1α (Figure 37)[147]. The aromatic ring of the molecule projects through the active site opening and likely makes a steric clash with the hydroxylated HIF-1α residue Pro564.

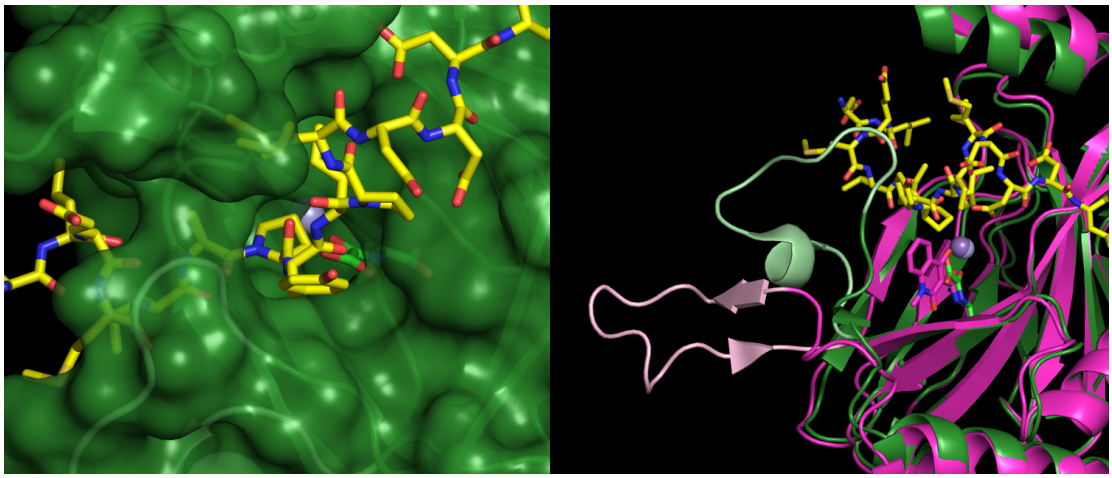


Figure 37: Crystal structures of PHD2 in complex with HIF-1α CODD (yellow carbons) and 2OG (green carbons, left), superposed with structure of inhibitor **1** bound (right).

These positions therefore were thought to offer suitable connection points for a linker. In particular, position 7 on the aromatic ring of the molecule seemed to be more exposed to the solvent compared with position 6 on the same ring.
As degrading machineries, it was chosen to investigate both $CRL2^{VHL}$ and $CRL4^{CRBN}$, developing two different series of compounds using VH032 ligand and pomalidomide to recruit VHL and CRBN, respectively. All the observation detailed previously (section 2.1) were also taken in consideration with regards to choosing the position where to attach the linker.

For VH032 it was chosen to derivatize from the acetyl group. Instead pomalidomide was derivatized from the amine in C-4, as done previously for the VHL targeting compounds. Flexible linkers made of PEG chains with either two, three, four, five or six ethylene glycol units were chosen to connect position 7 of PHD ligand **1** to the VHL ligand VH032. In contrast, linkers with a small size (two ethylene glycols units) and medium/long size (four ethylene glycols units) were chosen as representatives to develop a small series of compounds where the PHD IOX2-based ligand was linked to the pomalidomide scaffold. In both cases, the introduction of a carboxylic acid moiety was required in order to react with the free primary amine present in the linker. Structures of the final molecules are shown in Figure 38.

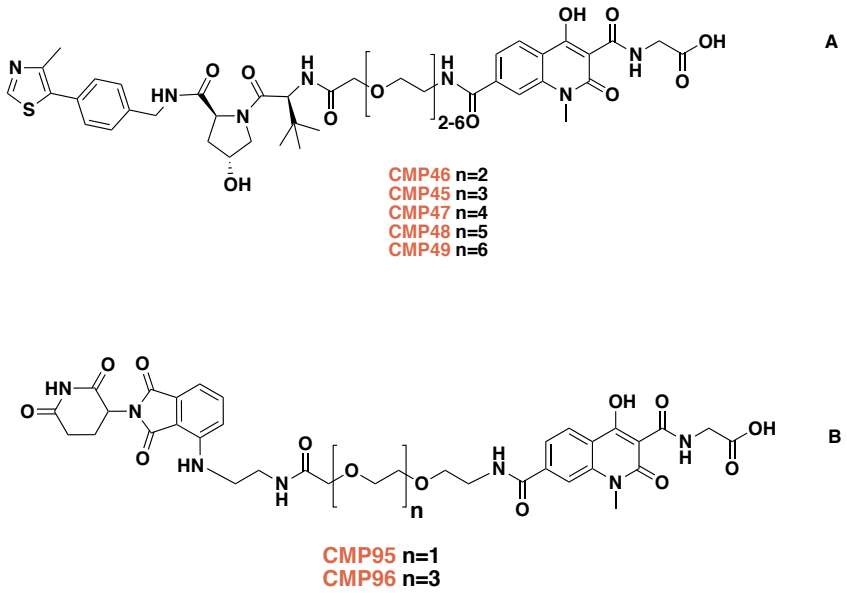


Figure 38: Structure of Hetero-PROTACs designed to recruit $CRL2^{VHL}$ (A) or $CRL4^{CRBN}$ (B) at one side and PHD enzymes at the other end.

To investigate if the free carboxylic acid on the RHS of IOX2 could impact on cell permeability of the final molecule, two compounds, CMP52 and CMP53, were designed substituting the carboxylic acid moiety with a methyl ester (Figure 39).

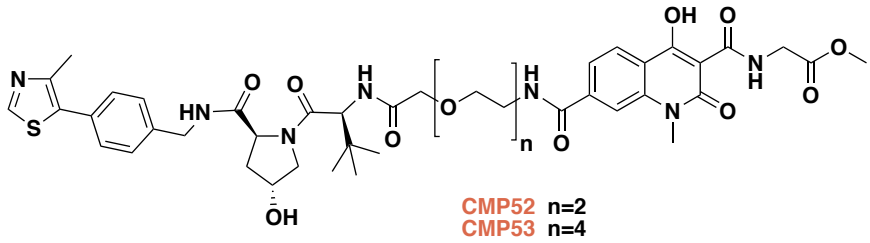


Figure 39: Structure of CMP52 and CMP53

The methyl group, because of its small size, should not interfere with the binding mode of the PHD ligand. It is supposed to be accommodating in a small pocket present in the catalytic domain (Figure 36). Moreover, the methyl ester could work in a pro-drug fashion, being potentially hydrolysed in cell by esterases, releasing the compound as its carboxylic acid derivative.

3.1.2 Hetero-PROTACs based on IOX4 ligand

The choice of IOX4 as a second ligand to recruit PHD enzymes was based on three main rationales:

1) Higher potency and selectivity: in comparison to previously identified PHD inhibitor IOX2, IOX4 was found to be 4-fold more potent *in vitro*, and 875-fold more selective for PHD enzymes over the other 2OG-dependent dioxygenases[155].

2) No presence of carboxylic acid group: to bind in the pocket occupied by the 2OG $(CH_2)_2COOH$ side chain, many of the previous PHD inhibitors contain in their structure a carboxylic acid moiety, which is often undesirable due to its possible impact on cell permeability. IOX4 showed a completely different structure when compared with its predecessor. IOX4 is characterized by a tricyclic structure with linked pyridine-carboxylate, dihydropyrazole and triazole rings (Figure 40). In this case, it is the triazole (an established carboxylate

mimic), rather than the pyridine-carboxylate (an established 2OG mimic) that occupies the 2OG-binding pocket (Figure 41).

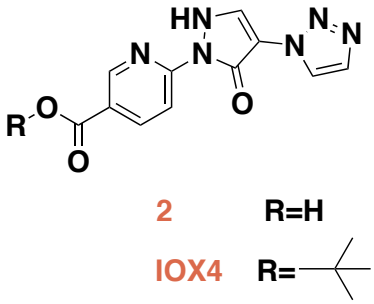


Figure 40: Structure of IOX4 and its free-acid derivative **2**.

3) More solvent exposed residue: the *tert*-butyl group present in the pyridine-based ring of IOX4 protrudes towards the entrance of the active site, preventing HIF-1α from binding because of a steric clash with this moiety (Figure 41).

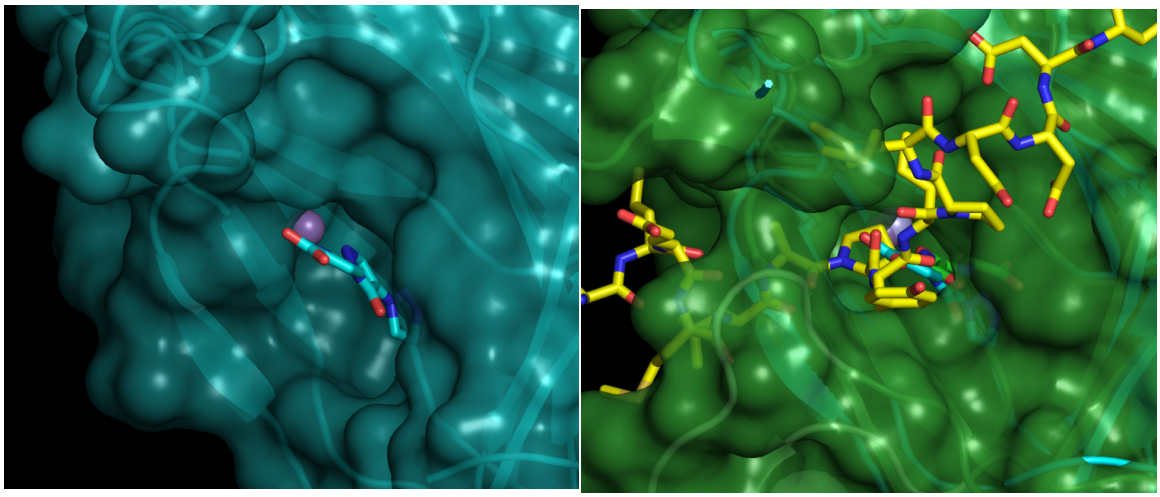


Figure 41: Left: crystal structure of IOX4 derivative **2** (cyan carbons) bound to the catalytic domain of PHD2. Right: superposition of HIF-1α (yellow), **2** and OG (green) bound to the PHD2 catalytic domain.

For all the reasons listed above, it was decided to develop a small series of compounds by derivatizing IOX4 analogue **2** (Figure 40) from the solvent-exposed carboxylic acid group. Indeed, this moiety could provide a convenient attachment for amide bond formation with the free primary amine already present in the previously designed and synthesized linkers. As degrading machinery, in this case also it was chosen to develop two series, one recruiting

CRL2^{VHL} and one recruiting CRL4^{CRBN}, using the same observations and consideration explained for the other compounds. With regards to linker length, PEG chains with two, three and four ethylene glycol units were chosen as representatives for the linkage with VHL ligand, and a medium size linker constituted of four ethylene glycol units for the linkage with CRBN ligand pomalidomide. Final structures of this series of compounds are presented in Figure 42.

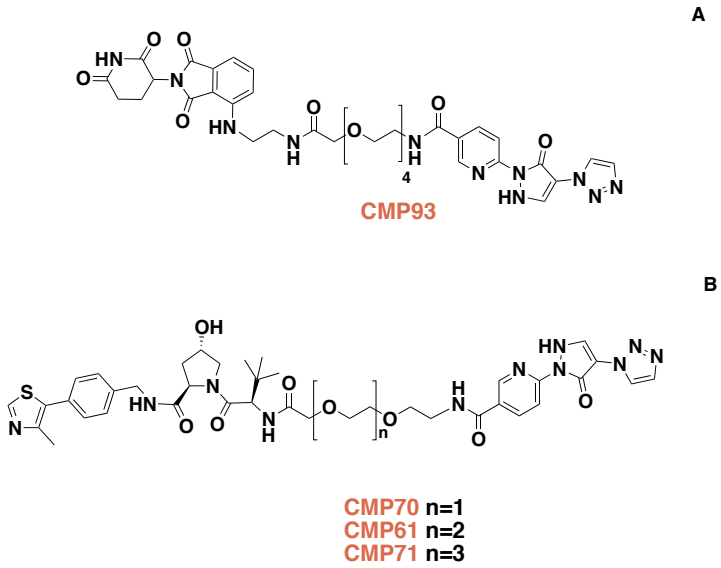


Figure 42: Structure of PHD targeting PROTAC recruiting CRL4^{CRBN} (A) and CRL2^{VHL} (B).

3.2 Synthetic routes to PHD-targeting compounds

Next, the synthetic routes to achieve both hetero-bifunctional compounds aimed to induce degradation of PHD enzymes are described.

3.2.1 Synthesis of Hetero-PROTACs based on IOX2 recruiting CRL2^{VHL} to target PHD enzymes

For the synthesis of a series of compounds targeting PHD proteins, it was decided to first assemble the IOX2-based PHD ligand. Commercially available 2-aminoterephthalic acid was converted to isatoic anhydride (Figure 43) derivative

54 in quantitative yield by reaction with triphosgene (Scheme 13). The latter was methylated both at its carboxylic acid moiety and on the nitrogen atom at position 1 by reaction with MeI in the presence of NaH via a S_N2 substitution reaction, leading to **55** in very good yield.

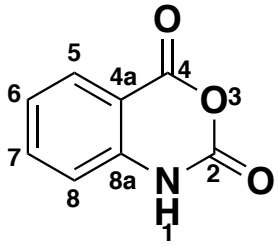
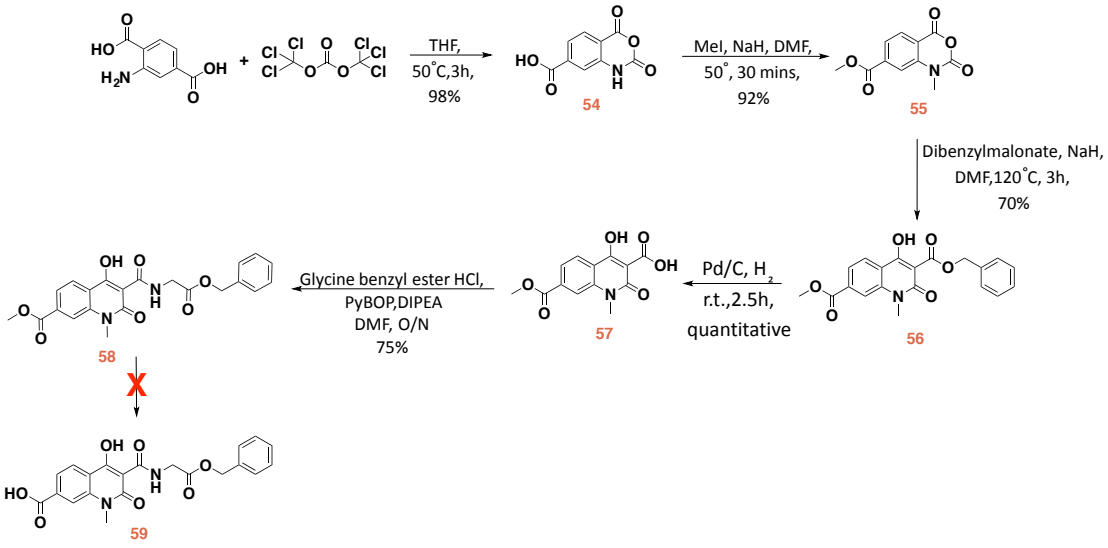


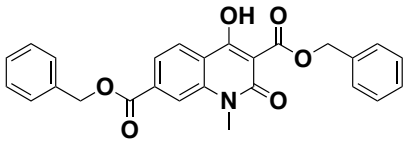
Figure 43: Chemical structure and numbering convention of isatoic anhydride.

For the synthesis of compound **56**, initially dibenzylmalonate was deprotonated by reaction with sodium hydride, resulting in the formation of an enolate anion. Upon addition of **55** to the reaction mixture, the enolate species previously formed performed a nucleophilic attack at position 4 of compound **55**, opening the isatoic anhydride ring**.** The resulting carbamate group is unstable and hydrolyses spontaneously into the amine derivative and CO_2. The resulting amine intramolecularly attacks the benzyl ester leading to the elimination of a benzyl alkoxide group, promptly neutralized with HCl, and ring cyclization to form product **56** that could be isolated in good yield (Scheme 13).



Scheme 13: Initial strategy for the synthesis of a linkable PHD ligand.

During this reaction, side product **60** was formed (Figure 44). This compound was isolated and analysed by NMR, and it showed to be a product of a side reaction, namely the nucleophilic attack of benzyl alkoxide species on the methyl ester group at position 7, forming the benzyl ester derivative **60**.



60

Figure 44: Chemical structure of side product **60**.

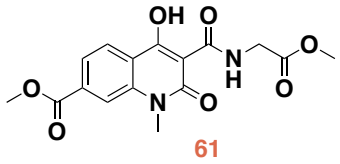
Cleavage of the benzyl group by hydrogenation afforded **57**. The carboxylic acid of **57** was coupled with glycine benzyl ester hydrochloride in the presence of PyBOP as coupling reagent, delivering **58** in 75% yield (Scheme 13).

In order to convert **58** to the free acid **59,** selective hydrolysis of the methyl ester at position 7 was required. In a first instance, compound **58** was treated with 1.1 equivalent of lithium hydroxide in a 1:1 solution of THF in H_2O (Table 4). After stirring for 2 h at room temperature and work up, unfortunately only a di-acid product was isolated. Such a product results from deprotection of both the methyl ester at position 7 and the benzyl ester at position 3. Traces of starting material were still detected.

A second attempt involved treating a solution of compound **58** in a 1:1 mixture of THF and water with 1.1 equivalents of lithium hydroxide at 0 °C (Table 4). The resulting mixture was stirred at 0 °C for 2 h and the reaction was monitored by LC-MS. As the starting material had not fully converted after the first 2 h of the reaction, the ice bath was removed and the reaction mixture was stirred O/N at room temperature. After quenching with a 1 M HCl solution and work up, only the di-acid product was isolated. The reaction had not gone to completion, since some starting material was still present. The same reaction was performed using 5% sodium hydroxide (1 eq.) in methanol at 0 °C, however no conversion of the starting material was observed (Table 4).

Another attempt was performed again using lithium hydroxide, however in this case in a 1:1 mixture of methanol in D_2O in order to monitor the reaction by NMR. No conversion of the starting material was observed (Table 4).

In a last attempt, compound **58** was treated with a 2 N solution of sodium hydroxide (1.1 eq.) in a mixture of methanol and DCM (Table 4). Upon preparative HPLC purification in acidic conditions, only the side product **61** coming from trans-esterification of the benzyl ester was isolated (Figure 45).



61

Figure 45: Chemical structure of side product **61**.

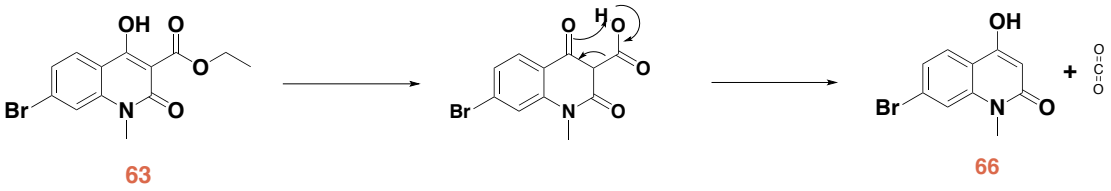
Conditions	Temp.	Reaction Time	Work up
LiOH H_2O 1.1 eq in THF/H_2O	r.t.	2 h	• Evaporated THF • Water extracted with EtOAc • Only di-acid product was isolated
LiOH H_2O 1.1 eq in THF/H_2O	1. 0 °C 2. r.t.	2 h O/N	• Quench with HCl • Wash with Brine • Extract water phase with $CHCl_3$/MeOH 10:1 • Combine organic phases • Only di-acid product was isolated
NaOH 5% (1 eq.) in MeOD	0 °C	48 h	None No conversion of s. m.
LiOH H_2O, 1.2 eq. in MeOD/D_2O	r.t.	48 h	None No conversion of s. m.
NaOH 2 N (1.1 eq.) in MeOH/CH_2Cl_2	r.t	10 min	HPLC (acidic) di methyl ester derivative was isolated

Table 4: Attempts at the selective deprotection of the methyl ester of **58**.

Due to the unsuccessful attempts to selectively hydrolyse the methyl ester at position 7 of **58**, a novel synthetic route was re-designed and developed from scratch. It was decided to start the synthesis of the inhibitor scaffold from commercially available 2-amino-4-bromobenzoic acid and to insert the required carboxylic acid moiety in a second time using a carbonylation reaction.

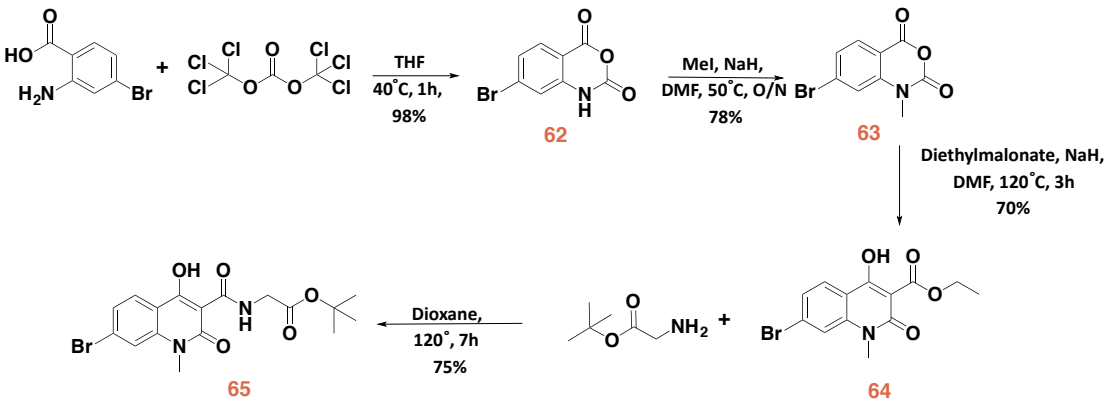
Starting material 2-amino-4-bromobenzoic acid was converted to **62** after reaction with triphosgene in 98% yield (Scheme 15). Upon methylation at position 1 using the same reaction conditions described above, compound **62** was reacted with the enolate derivative coming from a reaction between diethyl malonate and NaH. Diethyl malonate was chosen in this case instead of dibenzyl malonate because of the presence of the Br group that could be cleaved during the hydrogenation reaction needed for deprotection of the benzyl protecting group. Compound **63** was isolated in 70% yield.

To achieve the final desired compound **64** (Scheme 15)**,** in the first instance it was decided to deprotect the ethyl ester on **63.** Two attempts were conducted: initially, under acidic conditions using a 2.8 M solution of HCl in acetic anhydride. As this first attempt proved unsuccessful, the same reaction was repeated in basic conditions (1 N NaOH). Unfortunately in both cases no formation of the desired intermediate was observed. Instead a side product was the only product isolated. This side product was analyzed by NMR and shown to be **66**. Formation of this side product was due to the instability of the β-keto acid intermediate formed after hydrolysis of the ethyl ester, ending in a decarboxylation reaction leading to the formation of **66** (Scheme 14).



Scheme 14: Decarboxylation of **63** leading to the formation of **66**.

Because of the difficulties connected with the deprotection step, a different approach was adopted for the synthesis of **64**. Compound **63** was reacted with glycine *tert*-butyl ester HCl in a nucleophilic substitution at the carbonyl group of the ethyl ester. Because this reaction is reversible, in order to promote complete formation of the desired product, the ethanol formed was distilled by performing the reaction at 120 °C (Scheme 15). Compound **64** was obtained with 75% yield.



Scheme 15: Synthesis of compound **64**.

To introduce a carboxylic moiety at position 7 of compound **64** allowing attachment of the linker later in the synthesis, a carbonylation reaction was considered. In this specific case, it was chosen to prepare an *N*-hydroxysuccinamide (NHS) activated ester derivative of compound **64** using palladium-catalysed carbonylation (ref. [165]). To identify the best catalyst and solvent for this reaction, two combinations were evaluated (Table 5). In both attempts, the CO needed for this reaction was obtained from the decomposition of oxalyl chloride in an aqueous sodium hydroxide solution. The CO formed in this way filled a balloon that was subsequently connected to the flask where the carbonylation reaction took place. The only combination of catalyst and solvent that showed to lead to the formation of the desired product **10** was $Pd(dppf)Cl_2 \bullet$ DCM in dioxane. However, compound **10** was obtained in low yield, probably because the reaction was not performed in a properly sealed system, causing the loss of some CO.

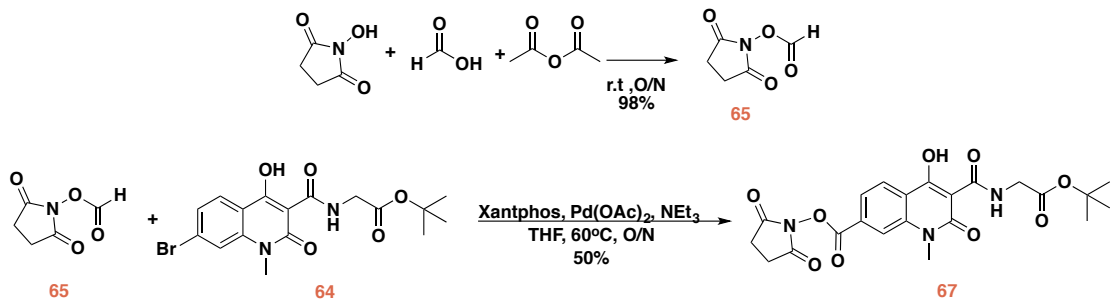
	Starting material	NHS	Cy_2NMe	Catalyst	Solvent, time, temp	Result
1	1 eq	1.4 eq	1.5 eq	$Pd(dppf)Cl_2$ • DCM	Toluene, 16 h, 95 °C	No product
2	1 eq	1.05 eq	1.5 eq	Palladium (*p*- cinnamyl) chloride dimer (3 mol%) + $HBF_4P(Tbu)_3$ (6 mol%)	Toluene, 16 h, 95 °C	No product
3	1 eq	1.4 eq	1.5 eq	$Pd(dppf)Cl_2$ • DCM	Dioxane, 16 h, 95 °C	Product, 27% yield

Table 5: Attempts to convert compound **64** in compound **67**.

A different approach, where *N*-hydroxysuccinimidyl formate **65** could be used as CO surrogate[166], was then considered for two main reasons: first, to have an internal source of CO gas needed for the carbonylation reaction and second, to work in an completely sealed reactor minimizing the possibility of leakage of CO that could reduce the efficacy of the reaction.

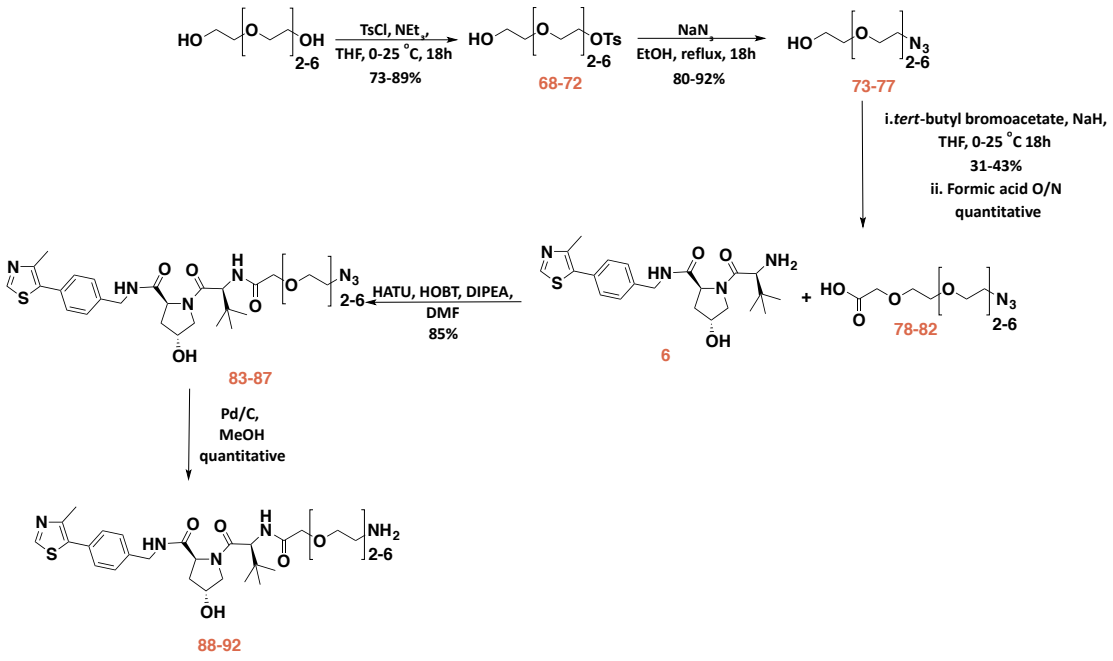
For the synthesis of **65,** NHS was reacted with acetic anhydride and formic acid delivering the desired product in quantitative yield (Scheme 16). Aryl halide **64** was then converted in the NHS activated ester **67** upon reaction with **65** in the presence of $Pd(OAc)_2$, XantPhos, and TEA. Gas evolution was observed upon addition of TEA to the mixture, due to reaction of the latter with **65**. In fact, the reaction between **65** and the tertiary amine leads to the formation of an unstable species that decomposes in NHS and releases in situ the CO needed for the carbonylation.

The desired product **67** was isolated after precipitation in good yield.



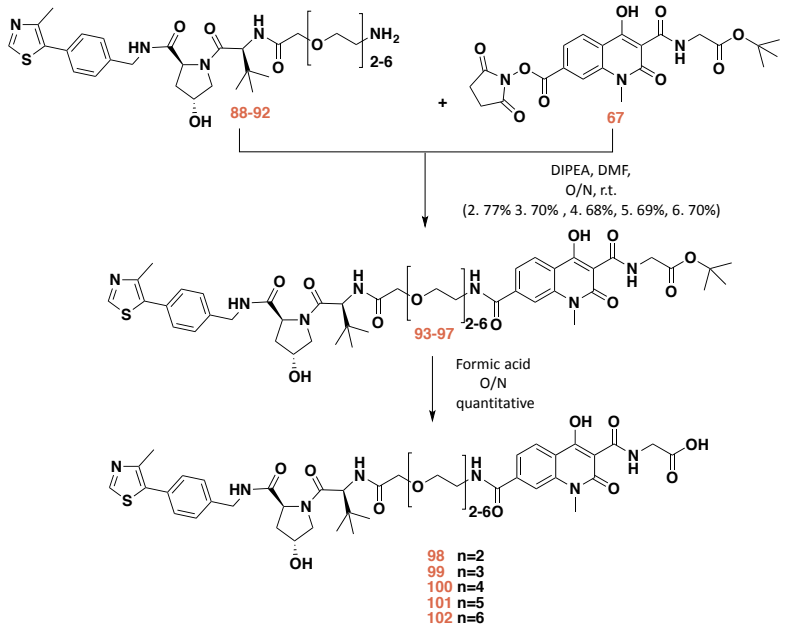
Scheme 16: Synthesis of activated ester **67**.

To connect the VHL ligand and **67** scaffolds together, PEG linkers with various lengths were synthesized (Scheme 17). Starting from PEG with two to six ethylene glycol units, in a first step one site of the PEG chains were tosylated, delivering compounds **68** to **72**. The introduced tosyl groups were substituted by an azide in a nucleophilic substitution reaction, forming compounds **73** to **77**. The remaining free hydroxyl group of the PEGs were then deprotonated with sodium hydride, to allow a nucleophilic attack to bromoacetic acid yielding PEG linkers **78** to **82** with an azide on one side and a reactive carboxylic acid on the other side. Compounds **78** to **82** were reacted with the free amine present in **6** in coupling conditions as described before, delivering compounds **83** to **87.** The terminal azide group of the latter was reduced to amine **88** to **92** using palladium on charcoal under an atmosphere of hydrogen (Scheme 17).



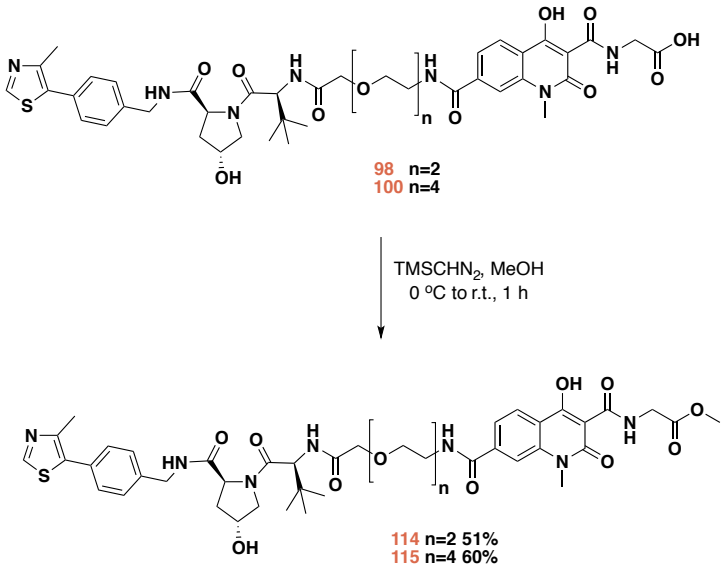
Scheme 17: Synthesis of the linkers and coupling with VHL ligand **6**.

The free amine of the linkers was then coupled in an amide bond formation reaction to **67,** affording compounds **93** to **97**. Hydrolysis of the *tert*-butyl ester in acidic conditions led to the formation of the desired final compounds **98** to **102** (Scheme 18).



Scheme 18: Synthesis of compound **98** (CMP46), **99** (CMP45), **100** (CMP47), **101** (CMP48), and **102** (CMP49).

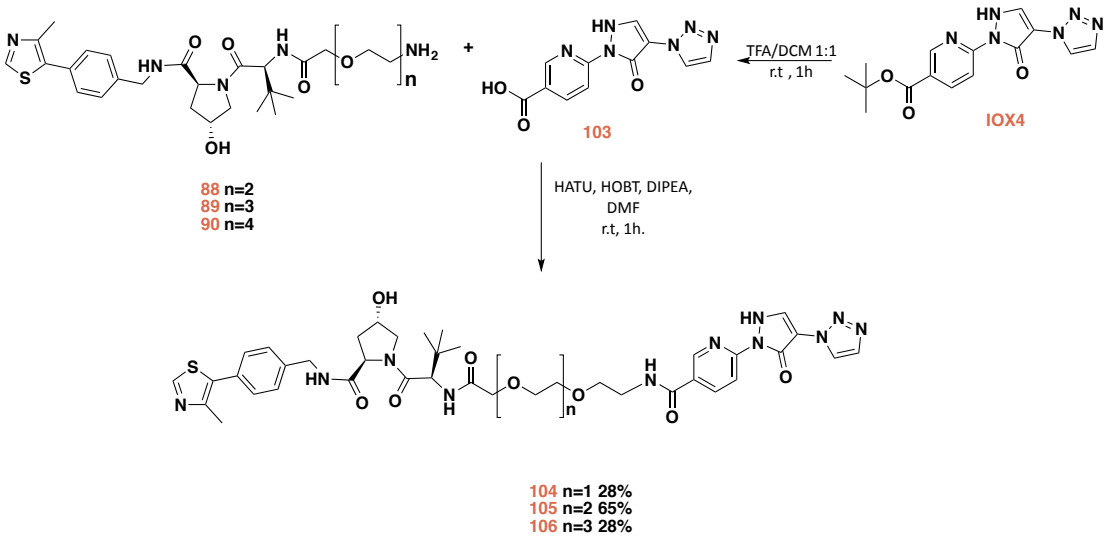
Methyl ester derivatives of compounds **98** and **100** where prepared titrating solutions of the two compounds in methanol with trimethylsilyldiazoethane ($TMSCHN_2$) (Scheme 19).



Scheme 19: Synthesis of compounds **114** (CMP52) and **115** (CMP53).

3.2.2 Synthesis of Hetero PROTACs based on IOX4 recruiting $CRL2^{VHL}$ to target PHD enzymes

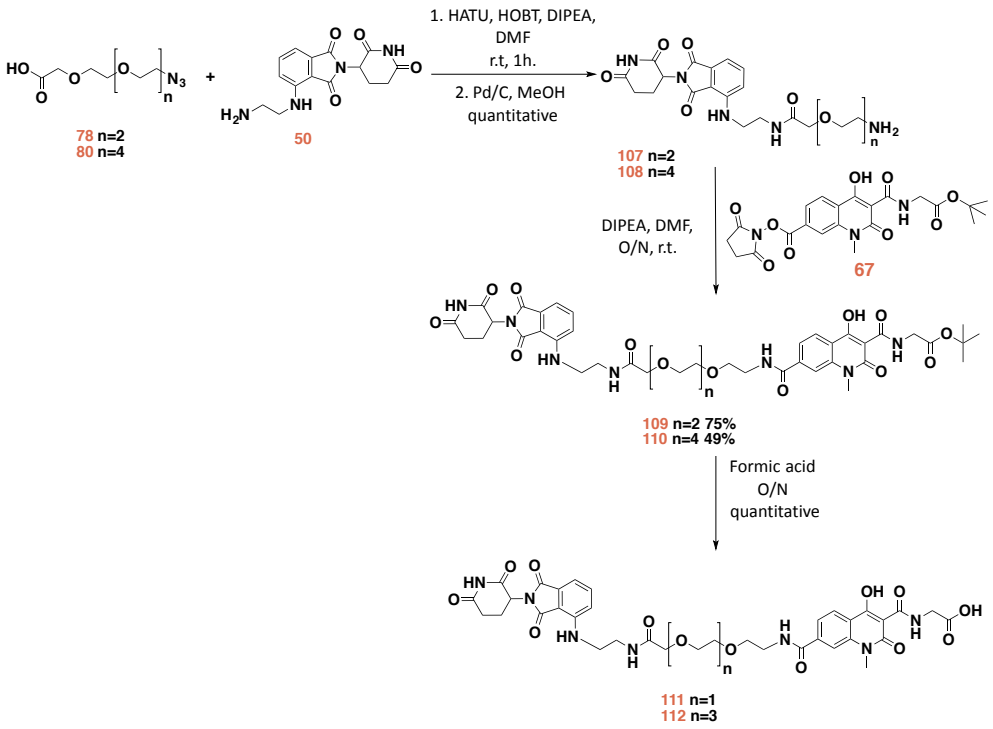
For the synthesis of this series of PROTACs, IOX4 compound was purchased from Cayman chemical. The *tert*-butyl ester group was hydrolysed using trifluoroacetic acid, delivering compound **103**. As previously described, **103** was connected to the previously synthesized VHL-linker amine compounds **88-90** by amide bond formation to obtain **104** (CMP70), **105** (CMP61), **106** (CMP71) (Scheme 20).



Scheme 20: Synthesis of **104** (CMP70), **105** (CMP61), and **106** (CMP71).

3.2.3 Synthesis of Hetero PROTACs based on IOX2 recruiting $CRL4^{CRBN}$ to target PHD enzymes

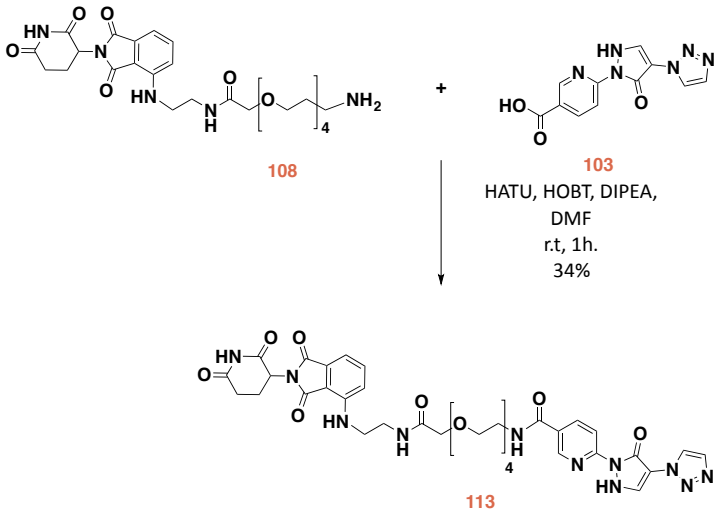
Compounds **107** and **108** were obtained after reaction of the free carboxylic acids **78** and **80** with the free amine **50** in coupling conditions as described before. Reduction of the azide led to the formation of compound **107** and **108**. The latter intermediates were coupled with the active ester **67** previously synthesized, affording compounds **109** and **110**. Hydrolysis of the *tert*-butyl ester in acidic conditions delivered compounds **111** (CMP95) and **112** (CMP96) (Scheme 21).



Scheme 21: Synthesis of compounds **111** (CMP95) and **112** (CMP96).

3.2.4 Synthesis of Hetero PROTACs based on IOX4 recruiting CRL4^{CRBN} to target PHD enzymes

For this series of compounds, compound **103** was connected to the previously synthesized **108** by amide bond formation, as previously described, to obtain **113** (Scheme 22).



Scheme 22: Synthesis of compound **113** (CMP93).

3.3 Biological evaluation of PHD-targeting compounds

The cellular evaluation of PHD-targeting compounds is described in the next section.

3.3.1 Compounds screening: evaluation of cellular activity

To assess the cellular activity of compounds, HeLa cells were treated with 1 μM of Hetero-PROTACs to target PHDs by recruiting $CRL2^{VHL}$ based on IOX2 structure (CMP46, CMP45, CMP47, CMP48, CMP49, CMP52, and CMP53); hetero-PROTACs recruiting $CRL2^{VHL}$ based on IOX4 structure (CMP61, CMP70, and CMP71); hetero-PROTAC CMP93 recruiting $CRL4^{CRBN}$ based on IOX4 structure; and hetero-PROTACs recruiting $CRL4^{CRBN}$ based on IOX2 structure (CMP95 and CMP96). DMSO (vehicle, 0.1% v/v), $CoCl_2$ (chemical inducer of HIF-1α), IOX2 and IOX4 (selective inhibitor of PHD2), and VH032 (selective VHL inhibitor) were used as controls. The samples, obtained after 10 h of treatment and cell lysis, were resolved by SDS- PAGE followed by Western blot using the corresponding specific antibodies to probe for the following proteins (Figure 46):

- **PHD2 and PHD3**: as explained previously, this series of compound were designed in order to target PHD enzymes, to induce their degradations. Unfortunately, no activity was detected against the target proteins when the bands belonging to the treated samples are compered with vehicle only. PHD1 was not evaluated because a specific antibody for this isoform was not commercially available.

- **HIF-1α:** To assess if treatment with this series of PROTACs could have any effect on stabilizing HIF-1α, protein levels of this protein were evaluated. No detectable effect was observed.

- **Hdy-HIF-1α:** Monitoring the level of hydroxylated HIF-1α could help to disentangle the effects of compounds on PHD enzymes from other hypoxia mimetic effects. We expect that if the compounds are able to bind the target

PHD proteins, and thus degrade them, or alternatively even block their activities, hydroxylation of HIF-1α should be inhibited and no HIF-1α in its hydroxylated form (Hdy-HIF-1α) should be detectable in cell lysates. Indeed, this would be consistent with what is observed when treating cells with either IOX2 or IOX4 inhibitors (Figure 46). Unfortunately, detectable levels of Hdy-HIF-1α where observed, suggesting that the compounds were not able to affect the level or activity of PHD enzymes.

- **VHL, Cullin2, CRBN**: because these series of compound were designed in order to recruit both $CRL2^{VHL}$ and $CRL4^{CRBN}$ as degrading machinery, levels of VHL, Cullin2 and CRBN proteins were evaluated as controls to monitor the level of these proteins too. As expected, the blots incubated with the specific antibody for the three proteins showed that treatment with Hetero PROTACs had no impact on the levels of these proteins.

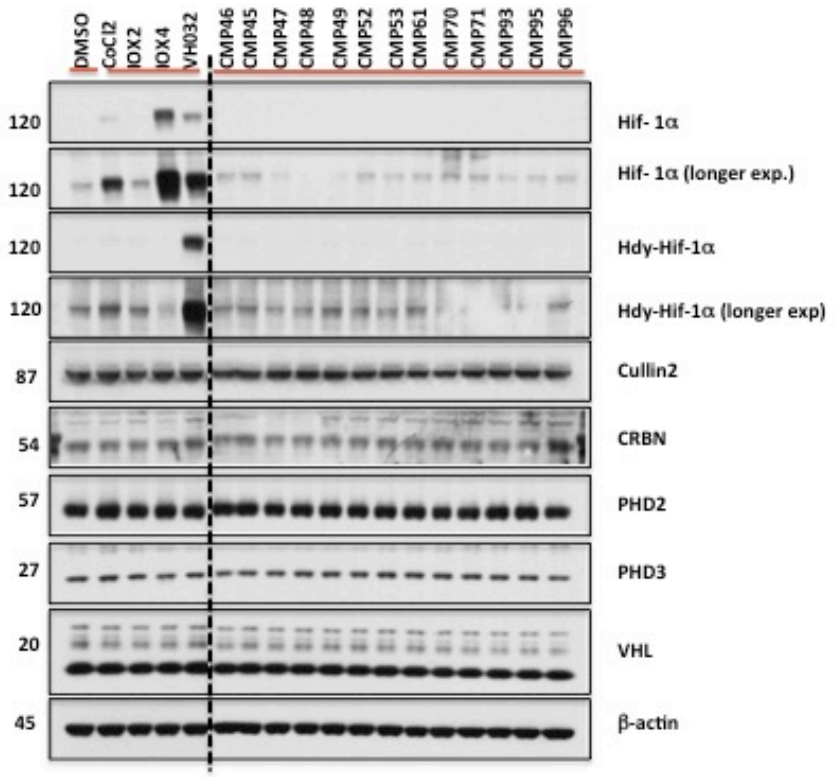


Figure 46: HeLa cells were treated with 1 μM of the indicated compounds and 0.1% DMSO, $CoCl_2$ (100 μM), IOX2 (50 μM), IOX4 (50 μM) or VH032 (250 μM) for 10 h.

The same experiment was performed in other two cell lines to assess if the not effectiveness of compounds activity could be due differences between the different cell lines. Experiments conducted in HEK293 and U2OS (Figure 47) confirmed the lack of activity of these series of PROTACs compounds.

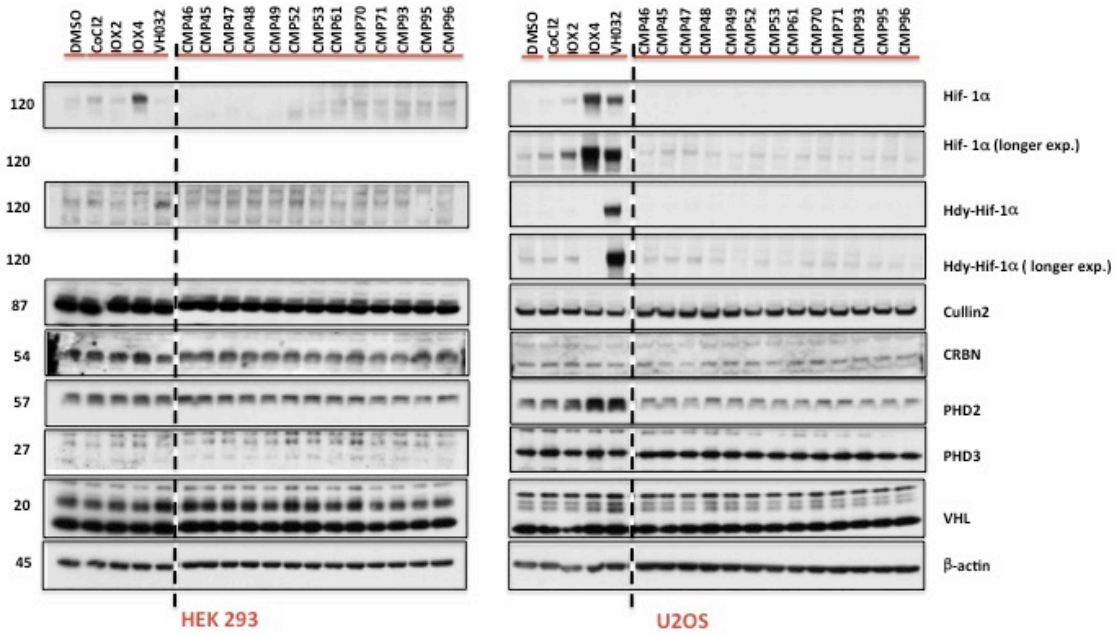


Figure 47: Treatment of 0.1 % DMSO, VH032 (250 μM), $CoCl_2$ (100 μM), IOX2 (50 μM), IOX4 (50 μM) and 1 μM of the indicated compounds in U2OS and HEK293 cells for 10 h.

3.3.2 PHD siRNA experiments

To confirm the reliability of the PHDs antibodies and to validate that the bands detected belong to the expected isoforms of PHD enzymes, siRNA experiments were performed in HeLa cells. The silencing of PHD2 and PHD1 was achieved after 48 h of treatment with specific siRNAs. For the silencing of PHD3, cells were pre-treated with a specific PHD3 siRNA for 24 h upon treatment with IOX2 inhibitor for further 24 h, in order to induce a hypoxic response and bust up the concentration of this otherwise low-abundance protein. The samples obtained after cell lysis were evaluated by immunoblotting, levels of target isoforms were monitored by incubation with the antibodies used in the previous experiment and no band relative to PHD2 or PHD3 were detected (Figure 48). These results demonstrated that the bands assumed to be related to both PHD2 and PHD3

were the right ones. During the same experiment, the same cells were treated with 1 μM of compounds CMP46, CMP45, CMP47, CMP48, CMP49 for 24h and results confirmed that these compounds were found to be inactive inside cells (Figure 48). Although you could comment that in this particular blot it can be observed a small but detectable effect on HIF-1α levels compared to DMSO control and similar to the one achieved using PHD1 siRNA, it could speculate that these effect is due to compounds inducing degradation of PHD1.

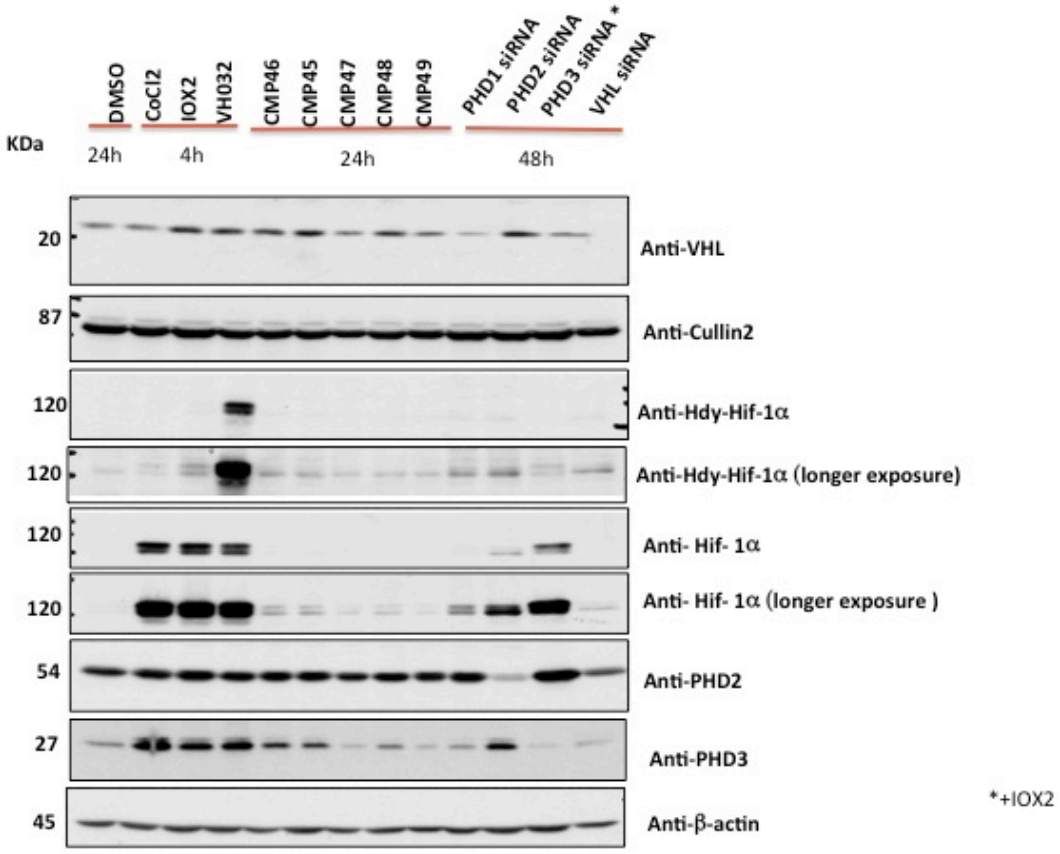


Figure 48: HeLa cells were treated with siRNA targeting the different isoforms of PHDs proteins or negative control VHL siRNA (for 48 h) as well as with the indicated compounds (24 h) and 0.1 % DMSO, $CoCl_2$ (100 μM), IOX2 (50 μM) or VH032 (250 μM).

3.3.3 Concentration dependency experiment

To assess if the lack of activity observed for this series of compound was due to the concentration at which the compounds were tested, a concentration dependency experiment was performed treating HeLa cells with different

concentrations of compounds CMP46, CMP45, CMP47, CMP48, CMP49 as representatives, at two different time points (4 h and 24 h) prior harvesting (Figure 49).

No detectable effect was observed on levels of both PHD2 and PHD3 at any of the concentration tested. Stabilization of HIF-1α in its hydroxylated form was however observed in particular at higher concentrations (Figure 49), suggesting that the VHL binding moiety is disrupting the protein-protein interaction between VHL and its substrate Hdy-HIF-1α. No detectable effect was observed on HIF-1α ever at high concentrations (Figure 49), suggesting that the compounds are not effective at blocking the HIF1α-PHDs interaction as the parent inhibitors IOX2 and IOX4 are.

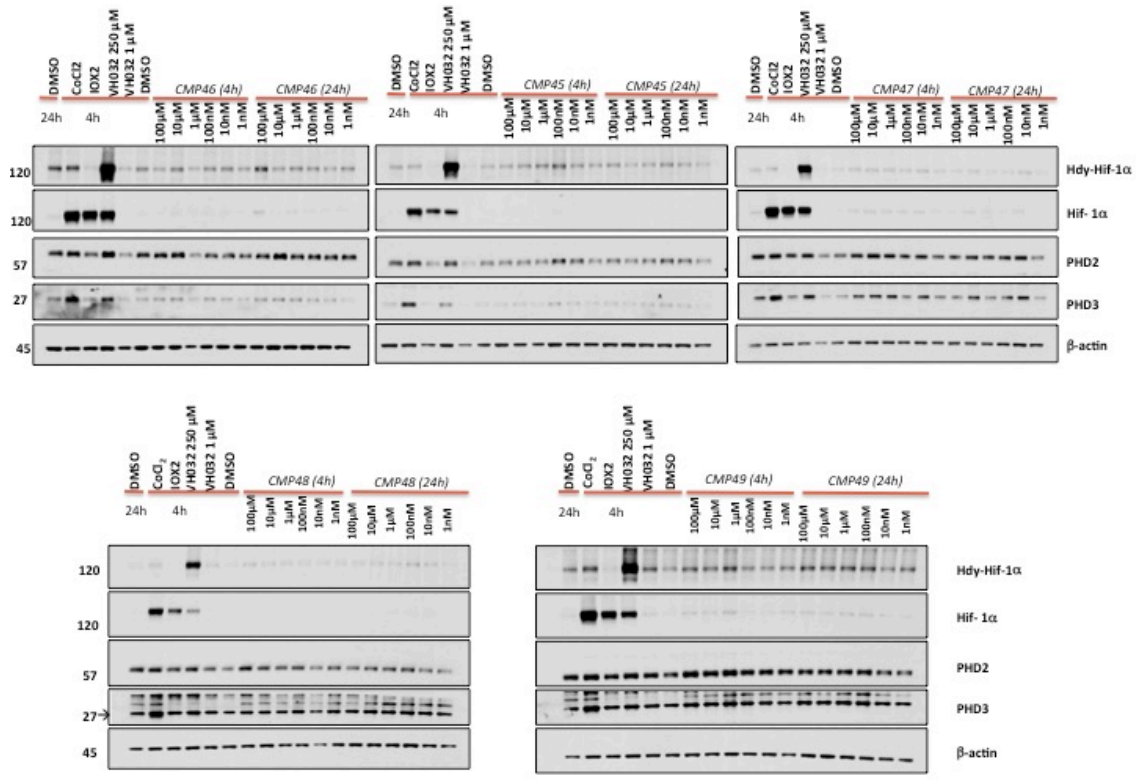


Figure 49: Hela cells were treated with increasing concentration of indicated compound for 4 h or 24 h and 0.1 % DMSO, $CoCl_2$ (100 µM), IOX2 (50 µM) or VH032 (250 µM).

3.3.4 Time dependent intracellular activities

To further confirm the inactivity of these compounds, a full time-dependency intracellular activity studies was performed treating HeLa cells with CMP46, CMP45, CMP47, CMP48, CMP49 at 1 μM concentration at different time points (Figure 50). No meaningful effect was observed on any of the protein evaluated.

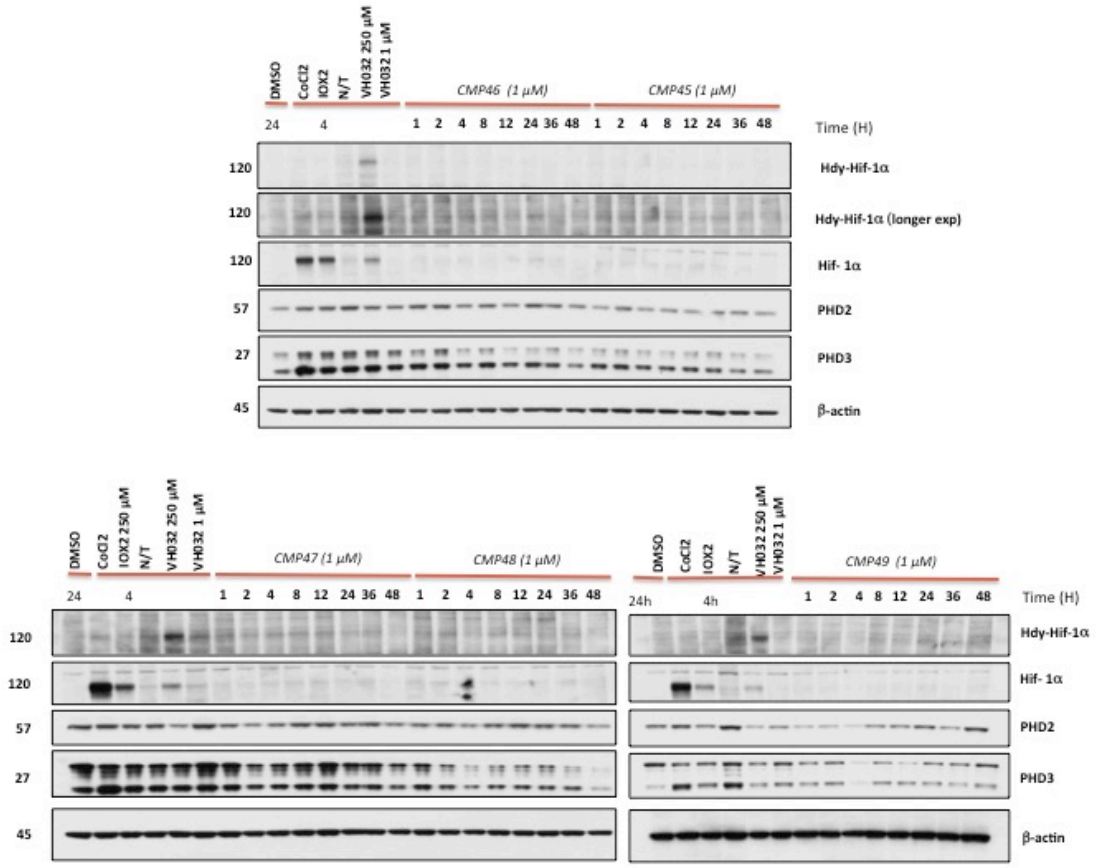


Figure 50: Time-course immunoblots of lysates from HeLa cells subjected to 0.1% DMSO, $CoCl_2$ (100 μM), IOX2 (150 μM), VH032 (250 μM or 1 μM) or 1 μM of indicated compounds.

3.4 Summary

In summary, this section reports the design, synthesis and biological evaluation of novel hetero-bifunctional compounds that aimed to induce the degradation of PHD enzymes by conjugating different PHD inhibitors to VHL and CRBN ligands. Several synthetic challenges were faced when derivatizing a PHD ligand (IOX2) to readily attach the linker group. In order to solve these challenges, new

synthetic routes and novel chemistry was developed to access carboxyl-derivatives of IOX2 primed for amide coupling with the linker. Unfortunately, cellular studies have been inconclusive and have failed to unambiguously detect depletion in cellular level of the monitored PHD2 and PHD3.

4 Discussion

To chemically interrogate the oxygen-signalling pathway in novel fashion using bifunctional molecules, it was strategically chosen to target for degradation two key molecular players of pathway – VHL and PHD enzymes – by hijacking both $CRL2^{VHL}$ and $CRL4^{CRBN}$. For this purpose, different series of all-small-molecule PROTACs were designed and synthesized by linking together the individual protein ligands to the CRLs binding moieties. To explore greater chemical space, the two molecules were linked from different positions, and by using linkers with different lengths and different chemical-physical characteristics. This combinatorial approach was deemed necessary in order to assess what impact the different permutations could have on final cellular activity. Most of the compounds achieved are hetero-bifunctional molecules, aimed to recruit two different proteins. Because VHL is considered for our purposes both as target protein and as CRL recognition subunit, homo-bifunctional molecules were obtained too, i.e. bivalent molecules composed of two instances of VHL ligands linked together. The design rationale for these molecules was to allow the recruitment of two VHL molecules in proximity with each other, in order to induce VHL-self degradation.

Overall, extensive cellular studies of all the molecules achieved synthetically in this thesis demonstrated that homo-bifunctional molecules proved the most effective degraders in cells. The best compounds were symmetric, with linkage from the terminal LHS acetyl group of VH032. In contrast, linkage at different positions proved ineffective, whether from VH032 or from the more potent ligand VH298. Indeed, no induced target degradation, and no complex formation too, were observed with these other Homo-PROTACs.

A remarkable discovery of this work is that, pleasingly, we observed a potent and selective degradation for the long isoforms of VHL, preferentially over the short isoform. The best compounds were CM09, CM10 and CM11, and they proved how length of the linker could affect potency in cells. Among these cell-active Homo-PROTACs, CM11 could induce complete depletion of levels of long

isoforms of pVHL after short treatments (4 h) already at 10 nM (**Figure 26**), suggesting rapid and pronounced activity. In the longer time treatments (>10 h) CM11 maintained potent and selective degradation for the long isoforms of VHL, with half-degrading concentration (DC_{50}) of approximately 100 nM. Interestingly, it was also observed that cellular levels of Cullin2 were also decreased to a measurable degree (approx. 20%) upon treatment with CM11. To my knowledge, this is the first demonstration that a protein different from the one directly targeted by a PROTAC can be degraded because it is part of the same complex.

Compared to its parent VHL inhibitor VH032, CM11 showed remarkable increase in cellular activity. Indeed, CM11 is active in the 10-100 nM range of concentrations, which compares very favorably with the inhibitor that is only active at 10-100 μM – an increase of almost 1000-fold. Treatment with CM11 showed detectable increases of HIF-1α levels at these low nanomolar concentrations. As the parent inhibitor is completely ineffective at these low concentrations, these downstream effects on HIF levels can only be due to the induced VHL degradation. As expected, it was the hydroxylated form of HIF-1α that was stabilized. Levels of HIF-1α remained fairly low e.g. compared to what could be achieved with the parent inhibitors when used at high micromolar concentrations. This is consistent with the results of the siRNA knockdown experiment, suggesting that VHL knockdown may need to be complete against all VHL isoforms (as it is only the case in $VHL^{-/-}$ cells) in order to drive a significant hypoxic response in cells. This observation also suggests that the un-degraded short isoform of VHL remaining in the cell may be sufficient to efficiently maintain low levels of HIF-1α. Future work to assess and quantify the extent of the hypoxic response induced by VHL degraders is warranted.

Provided in the thesis is extensive data that provide strong evidence that our cell-active Homo-PROTAC compounds work by forming a ternary complex VHL_2-PROTACS as the key intermediate species en-route to the ubiquitination and proteasomal degradation of VHL. Mechanistically, we provide strong evidence to support that CM11 works as we had initially hypothesized. Indeed, VHL

depletion was dependent on both proteasomal activity (as shown by co-treatment with MG132, Figure 30) and Cullin neddylation (as shown with MLN4924). It was also dependent on binding to VHL, as shown by competition experiment with excess VH032, and by lack of induced degradation activity of the *cis-trans* analogue. It was also shown that CM11 recruits two molecules of VHL simultaneously with a stoichiometry of 2:1 protein/PROTAC. Moreover, CM11 showed remarkable avidity (defined as the increase in binding affinity for multivalent events relative to the individual monovalent one), with an increase in binding affinity of more 20-fold compared to the parent inhibitor. It has been demonstrated how this effect is strictly dependent on the presence of two *trans*-Hyp moieties within the same molecule, supporting a model in which a highly cooperative ternary complex VHL-CM11-VHL function as the key species responsible for the induced degradation of VHL (Figure 51). We postulate that induced intermolecular interactions within this ternary complex may be accounting for increased avidity of complex formation by CM11, which will warrant future structural and mechanistic studies.

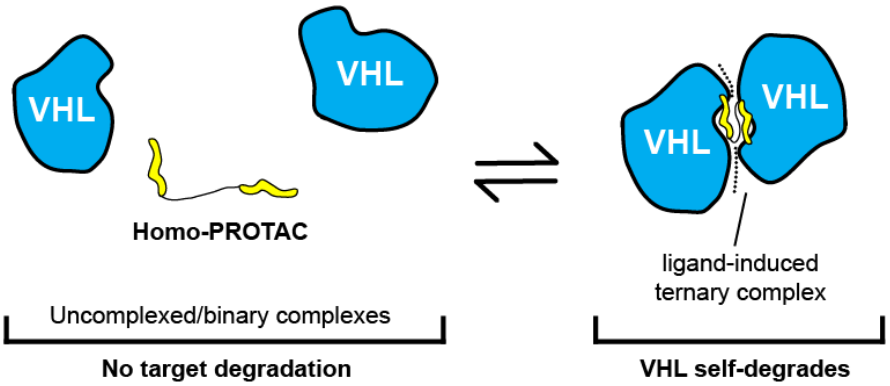


Figure 51. Proposed model for the mechanism of action of Homo-PROTAC CM11.

5 Conclusions and future work

The discoveries outlines in this thesis provide the foundation for future work evaluating the utility of CM11 as a chemical probe to study the biological consequences of inducing acute knockdown of VHL. In particular, the preferential degradation of the long relative to the short isoform of VHL will enable the possibility to dissect and further elucidate the relative functions and subtle difference between these two isoforms. These individual roles are still not fully understood, and current or chemical genetic tools to address these questions are inadequate. In this respect, this thesis provides a powerful demonstration of the development of innovative chemical tools that may enable in future researchers to probe biology in new ways, leading to many unforeseen discoveries.

In the short terms, our work will focus on complete the characterization of CM11 prior to disclosing it to the scientific community as a chemical probe. In addition, several outstanding questions remain as to the inactivity of the PHD-targeting compounds that would be interesting to address. Based on the results obtained in this thesis, some of the directions of our future studies are provided below:

1) Homo-PROTACs

- We believe that the enhanced avidity of CM11 results from the formation of a stable ternary complex, with CM11 being accommodated in a cavity formed between the two VHL proteins. To elucidate the structural basis of this bifunctional and highly cooperative recognition, our efforts are currently focused on obtaining a crystal structure of the ternary complex.

- To unambiguously identify the nature of the specific isoforms depleted selectively by CM11, isoform-specific siRNAs will be designed and assessed in cells. Furthermore, mass spectrometry-based proteomics will be employed to quantify the extent of the VHL depletion in cells and assess the impact of CM11 on the whole proteome in an unbiased manner.

- To assess the downstream impact of CM11 on HIF transcriptional activity, HRE-luciferase assays will be conducted and the mRNA levels of HIF responsive genes will be monitored. It could also be interesting to evaluate levels and activity of HIF-2α as well.

- It will be interesting to investigate whether the long linkers used to connect the two VHL-binding moieties could be the cause of the lack of activity of Homo-PROTACs connected via the phenyl rings. Indeed, considering the linking point and the necessity of bring in really close proximity the two proteins, it is possible that a long linker cannot be accommodated between the two proteins in this relative orientation, thus preventing the formation of a stable ternary complex. So, compounds with shorter linkers, like one or two PEG units, will be synthesized in future to test this hypothesis.

2) Hetero-PROTACs

- All PHDs targeting compounds were designed by linking the PHD recruiting moiety to the acetyl group present at the LHS of VH032. Based on evaluation of the crystal structures of both VHL and PHD2, the lack of activity shown by this series of compounds could be explained considering that VHL ligands bind on the surface of the VHL protein, whereas PHD ligands need to enter in the deep catalytic site of the PHD proteins. We hypothesize that in order to allow the PROTAC compound to productively bring the two proteins together so that their surfaces can interact to form a stable ternary complex, a different linkage e.g. from the phenyl group of the VHL ligand could be better tolerated. So different series of compounds, based on both IOX2 and IOX4 structure and linking VH032 and/or VH298 from a different linkage point, could be synthetized and tested.

6 Materials and Methods

6.1 Chemistry

6.1.1 Synthesis and characterization of PROTACs targeting VHL

General method A: To a solution of acid (1 eq.) in DMF was added HATU (1 eq.) and HAOT (1 eq.) and the solution was stirred at room temperature for 5 min. The amine (1 eq.) was added and the pH of the reaction mixture was adjusted to >9 by addition of DIPEA (3 eq.) and the mixture was stirred at room temperature until no presence of the starting materials was detected by LC-MS. Water was added and the mixture was extracted with ethyl acetate (×3). The combined organic phases were washed with brine (×2), dried over $MgSO_4$ and evaporated under reduced pressure to give the corresponding crude, which was purified by flash column chromatography using a gradient of 0 to 80% v/v acetone in heptane to yield the final compound.

General method B: The N-Boc-protected compound was dissolved in DCM. An equal volume of 4 M HCl in dioxane was added and the reaction mixture stirred at room temperature for 2 h. The solvent was removed under a stream of nitrogen and dried under reduce pressure.

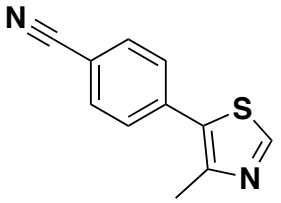
General method C: PEG was solubilised in dioxane anhydrous and NaH was added under stirring. The resulting mixture was stirred at r.t. for 3h. The mixture was cooled down to 0 ^{o}C using ice bath and *tert*-butylbromo acetate was added drop by drop. The resulting mixture was stirred at r.t O/N. The precipitate was filtered off and the organic phase evaporated to dryness. The resulting oil was taken up with ethyl acetate, washed with water, dried over $MgSO_4$ and evaporated to dryness. The resulting oil was purified by column chromatography using a gradient of ethyl acetate from 50% to 100% v/v in heptane.

General method D: product obtained following method C was dissolved in a solution of 50% v/v trifluroacetic acid in DCM. The resulting solution was stirred

for 1 h or until complete conversion of starting material. The solvent was removed under high vacuum.

General method E: To a solution of PEG linker (1 eq.) in 1 ml DMF was added HATU (17 mg, 0.045 mmol, 2.5 eq.) and the solution was stirred at room temperature for 5 min. Compound 6 (20 mg, 0,045 mmol, 2.5 eq.) was added and the pH of the reaction mixture was adjusted to >9 by addition of DIPEA (3 eq.). The mixture was stirred at room temperature until no presence of the starting materials was detected by LC-MS. Water was added and the mixture was extracted with ethyl acetate (×3). The combined organic phases were washed with brine (×2), dried over $MgSO_4$ and evaporated under reduced pressure to give the corresponding crude, which was purified by HPLC using a gradient of 20% to 95% v/v acetonitrile in 0.1% aqueous solution of ammonia to yield the final compound.

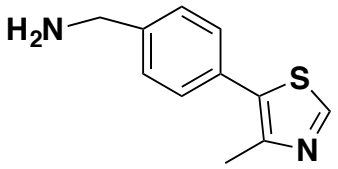
4-(4-methylthiazol-5-yl)benzonitrile (1)



Bromobenzonitrile (6 g, 33 mmol, 1 eq.) and Palladium (II) acetate (7.40 mg, 0.033 mmol, 0.001 eq.) were dissolved in DMAC (32 ml). Potassium acetate (6.47 g, 65.9 mmol, 2 eq.) and methylthiazole (6 ml, 65.9 mmol, 2 eq.) were added to the solution and the resulting mixture stirred ate 150 ^{o}C O/N. The mixture was cooled to r.t. and the reaction was quenched with water. The aqueous phase was extracted with DCM (×3). The combined organic phases where washed with water (×3), dried over $MgSO_4$ and evaporated to dryness. The residue was purified by column chromatography using a gradient of 0 to 60% v/v ethyl acetate in heptane. Yield: 5.34 g (80%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.79 (s, 1H), 7.77 - 7.73 (m, 2H), 7.61 - 7.57(m, 2H), 2.60 (s, 3H). **^{13}C NMR** (500 MHz, $CDCl_3$) δ

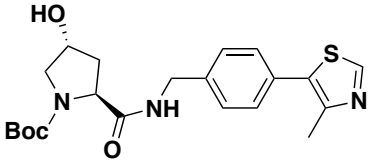
151.4, 149.8, 136.6, 132.6, 129.9, 129.5, 118.2, 111.2, 16.2. Analytical data matched those previously reported[115].

(4-(4-methylthiazol-5-yl)phenyl)methanamine (2)



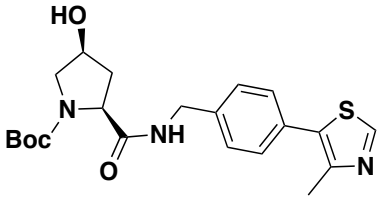
6.5 ml of $LiAlH_4$ 2M (1.1 eq., 13 mmol) solution in THF was chilled to 0 °C under N_2 atmosphere, to which a solution of **1** (2.50 g, 12.5 mmol, 1 eq.) in 25 ml of THF was added dropwise. The resulting mixture was stirred at 50 °C for 22 h. The mixture was then cooled to 4 °C and quenched sequentially with water (520 μl) then 20% w/v sodium hydroxide (390 μL) and after with water (1.73 ml). After standing for 15 min, the mixture was filtered through filter paper, then washed three times with chloroform. The organic layer was concentrated and then purified by column chromatography (10 to 30% v/v MeOH/DCM) to give a yellow oil that slowly solidifies upon storage. Yield: 1.620g (63%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.60 (1H, s), 7.37 - 7.30 (4H, m), 3.85 (2H, s), 2.46 (3H, s); **^{13}C NMR** (500 MHz, $CDCl_3$) δ 152.77, 149.07, 143.63, 133.42, 131.46, 130.49, 129.05, 46.23, 15.79. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{11}H_{12}N_2S$: 204.07; observed: 205.08.

***tert*-butyl (2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carboxylate (3)**



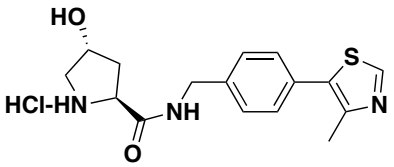
Following the general method A, from **2** (788 mg, 3.34 mmol, 1 eq.) and *trans*-*N*-(*tert*-Butoxycarbonyl)-4-hydroxy-L-proline (890 mg, 3.84 mmol, 1 eq.) compound **3** was obtained as a yellow solid (1.353 g, 95%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.53 (1H, s), 7.65 - 7.21 (4H, m), 5.07 (1H, s), 4.50 - 4.13 (4H, m), 3.66 - 3.30 (2H, m), 2.33 (3H, s), 2.26 - 1.89 (2H, m), 1.35 - 1.16 (9H, m); **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{21}H_{27}N_3O_4S$: 417.17; observed: 418.2. Analytical data matched those previously reported[115].

tert-butyl (2*S*,4*S*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carboxylate (13)



Following the general method A, from **2** (264 mg, 1.28 mmol, 1 eq.) and *cis*-*N*-(*tert*-Butoxycarbonyl)-4-hydroxy-L-proline (155 mg, 1.28 mmol, 1 eq.) compound **13** was obtained as a yellow solid (376 g, 70%). Analytical data matched those previously reported[116].

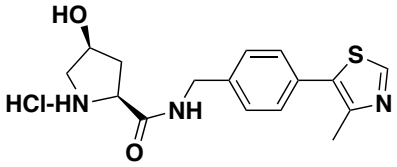
(2*S*,4*R*)-4-hydroxy-N-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide hydrochloride (4)



Following general method B, from **3** (1.271 g, 3.96 mmol, 1 eq.) and 10 ml of 4 M HCl in dioxane, compound **4** was obtained as light yellow foam after high vacuum (1.3 g, 3.68 mmol, 93%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.58 (1H, s), 7.54 - 6.45

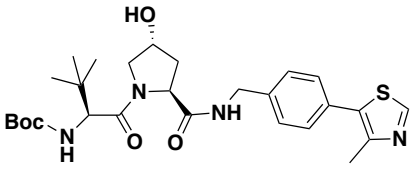
(5H, m), 4.46 - 4.15 (4H, m), 3.61 - 3.55 (1H, m), 3.42 (1H, s), 3.10 - 3.05 (2H, m), 2.37 - 2.05 (5H, m). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{16}H_{20}ClN_3O_2S$: 353.10; observed: 317.4 $(M-Cl)^+$. Analytical data matched those previously reported[116].

(2*S*,4*S*)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide hydrochloride (14)



Following general method B, from **13** (376 mg, 1.42 mmol, 1 eq.) and 4 M HCl in dioxane (3.56 ml, 14.2, 10 eq.), compound **14** was obtained as light yellow foam after high vacuum (305 mg, 0.862 mmol, 96%). Analytical data matched those previously reported[116].

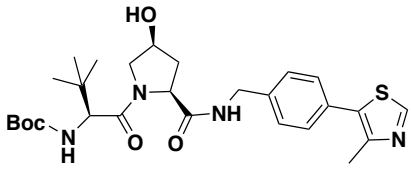
***tert*-butyl ((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamate (5)**



Following general method A, from **4** (763 mg, 2.40 mmol, 1 eq.) and Boc-L-tert-leucine (555 mg, 2.40 mmol, 1 eq.) compound **5** was obtained as a yellow solid (1.2 g, 94%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.62 (1H, s), 7.31 - 7.26 (4H, m), 5.07 (1H, d, *J*=10.0 Hz), 4.73 - 4.70 (1H, m), 4.54 - 4.46 (2H, m), 4.29 - 4.22 (1H, m), 4.07 (1H, t, *J*=8.9 Hz), 3.68 - 3.63 (4H, m), 3.50 - 3.47 (1H, m), 2.67 (1H, s), 2.58 -

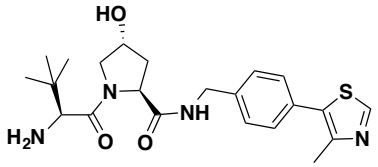
2.52 (1H, m), 2.46 (3H, s), 2.09 - 2.03 (1H, m), 1.35 (9H, s), 0.86 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.06, 171.22, 156.05, 150.45, 148.22,138.23, 131.68, 130.64, 129.36, 127.89, 80.08, 69.94, 58.81, 56.57, 55.44, 43.04, 36.48,35.43, 28.31, 26.28, 15.96. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{27}H_{38}N_4O_5S$: 530.26; observed: 531.27

tert-butyl ((_S_)-1-((2_S_,4_S_)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamate (15)



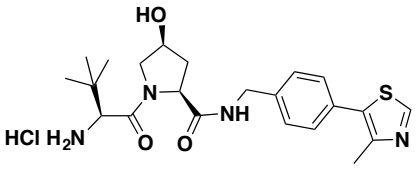
Following general method A, from **14** (353.46 mg, 0.863 mmol, 1 eq.) and Boc-L-tert-leucine (221 mg, 0.960 mmol, 1.1 eq.) compound **15** was obtained as a yellow solid (283 mg, 0.533 mmol, 61%). observed: 531.3. **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.69 (s, 1H), 7.64 (t, *J*=5.9 Hz, 1H), 7.37 – 7.28 (m, 4H), 5.34 (d, *J*=9.2 Hz, 1H), 4.72 (t, *J*=7.9 Hz, 1H), 4.53 (dd, *J*=14.8, 6.7 Hz, 2H), 4.33 (dd, *J*=15.1, 5.3 Hz, 1H), 4.23 (d, *J*=9.3 Hz, 1H), 3.94 (d, *J*=11.1 Hz, 1H), 3.71 – 3.63 (m, 1H), 2.56 – 2.45 (m, 3H), 2.39 (ddd, *J*=12.7, 7.8, 4.6 Hz, 1H), 2.17 – 2.05 (m, 1H), 1.41 (s, 9H), 0.94 (s, 9H). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{27}H_{38}N_4O_5S$: 530.26. Analytical data matched those previously reported[116].

(2_S_,4_R_)-1-((_S_)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-_N_-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (6)



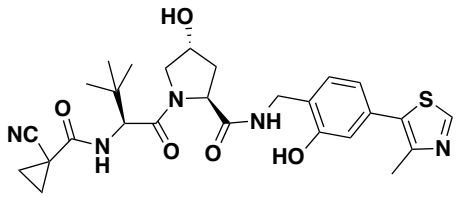
Following general method B, from **5** (1.2 g, 2.25 mmol, 1 eq.) in 10 ml of DCM and 10 ml of 4M HCl in dioxane, compound **6** was obtained as light yellow foam after high vacuum (1.3 g, 3.68 mmol, 93%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.61 (1H, s), 7.40 (1H, s), 7.30 - 7.25 (4H, m), 4.73 - 4.68 (1H, m), 4.48 - 4.41 (2H, m), 4.26 - 4.19 (1H, m), 3.69 - 3.52 (3H, m), 3.30 (1H, s), 2.74 (3H, s), 2.47 - 2.33 (3H, m), 2.09 - 2.02 (1H, m), 0.85 (9H, s); **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{22}H_{30}N_4O_3S$: 430.20; observed: 430.8. Analytical data matched those previously reported[115].

(2*S*,4*S*)-1-((*S*)-2-amino-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (16)



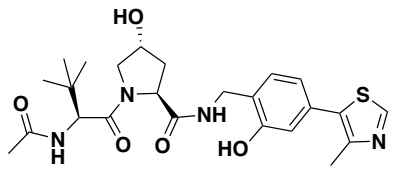
Following general method B, from **15** (283 mg, 0.533 mmol, 1 eq.) in 1 ml of DCM and 1 ml of 4 M HCl in dioxane, compound **16** was obtained as light yellow solid after high vacuum. Yield: 240 mg, 0.516 mmol (96%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.61 (1H, s), 7.40 (1H, s), 7.30 - 7.25 (4H, m), 4.73 - 4.68 (1H, m), 4.48 - 4.41 (2H, m), 4.26 - 4.19 (1H, m), 3.69 - 3.52 (3H, m), 3.30 (1H, s), 2.74 (3H, s), 2.47 - 2.33 (3H, m), 2.09 - 2.02 (1H, m), 0.85 (9H, s). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{22}H_{30}N_4O_3S$: 430.20; observed: 430.8. Analytical data matched those previously reported[115].

(2*S*,4*R*)-1-((*S*)-2-(1-cyanocyclopropane-1-carboxamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(2-hydroxy-4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (31)



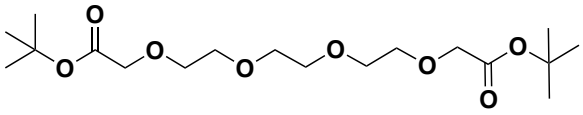
1-cyanocyclopropanecarboxylic acid (69 mg, 0.62 mmol, 1 eq.) was solubilized in 3 ml of DMF. HATU (235 mg, 0.62 mmol, 1 eq.) and HOAT (84.4 mg, 0.62 mmol, 1 eq.) were added and the resulting mixture was stirred at r.t. for 5 min. The amine precursor of **31** (gift from Dr. Andrea Testa) (300 mg, 0.62 mmol, 1 eq.) was added and the pH was adjusted to pH>9 using DIPEA (400 mg, 0.5 ml, 3.1 mmol, 5 eq.). The resulting mixture was stirred at r.t. until complete conversion of the starting material. Water was added, and the mixture was extracted with ethyl acetate (×3). The combined organic phases were washed with brine (×1), dried over MgSO4, and evaporated to afford the corresponding crude compound that was purified by flash column chromatography using a gradient of 10% to 70% acetone in heptane to yield the title compound as a white solid. Yield: 200 mg, 0.37 mmol (60 %) **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.25 (1H, s), 8.68 (1H, s), 8.06 - 8.02 (1H, m), 7.14 (1H, d, *J*=8.0 Hz), 7.02 - 6.98 (2H, m), 6.92 - 6.89 (1H, m), 4.80 - 4.74 (1H, m), 4.56 (1H, s), 4.48 - 4.41 (2H, m), 4.24 - 4.18 (1H, m), 3.90 (1H, d, *J*=11.9 Hz), 3.66 - 3.61 (1H, m), 2.53 (4H, s), 2.51 - 2.47 (1H, m), 2.14 - 2.07 (1H, m), 1.69 - 1.46 (4H, m), 0.91 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.8, 170.8, 165.8, 155.8, 150.5, 148.3, 133.3, 131.6, 131.2, 123.9, 120.9, 119.6, 118.2, 70.1, 58.6, 58.3, 56.7, 55.7, 40.0, 35.7, 26.2, 18.6, 17.9, 17.8, 17.2, 16.1, 13.8.. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{27}H_{33}N_5O_5S$: 539.22; observed: 540.3.

(2*S*,4*R*)-1-((*S*)-2-acetamido-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(2-hydroxy-4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (32)



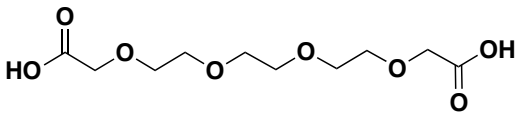
The amine precursor of **32**, a gift from Dr. Andrea Testa, (70 mg, 0.14 mmol, 1 eq.) was dissolved in 1 ml of DCM, and triethylamine (42.5 mg, 0.058 ml, 0.42 mmol, 3 eq.) was added to the solution. After stirring the mixture for 10 min at room temperature, the reaction mixture was cooled to 0 ^{o}C and acetic anhydride (17.35 mg, 0.016 ml, 0.17 mmol, 1.2 eq.) was added and the reaction was stirred at r.t for 1.5 h. The solvents were evaporated under reduced pressure to give the corresponding crude, which was purified by HPLC using a gradient of 5% to 70% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield the titled compound. Yield: 48.6 mg, 0.10 mmol (71%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.17 (1H, s), 8.61 (1H, s), 7.91 - 7.86 (1H, m), 7.06 (1H, d, *J*=7.9 Hz), 6.75 - 6.71 (2H, m), 6.64 (1H, d, *J*=9.3 Hz), 4.83 - 4.79 (1H, m), 4.50 - 4.33 (3H, m), 4.17 - 4.05 (2H, m), 3.52 - 3.47 (1H, m), 2.44 - 2.37 (4H, m), 2.16 - 2.08 (1H, m), 1.77 (3H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.8, 171.2, 155.9, 150.7, 148.1, 132.8, 131.7, 131.0, 124.2, 120.6, 117.1, 70.3, 58.1, 57.7, 57.1, 39.8, 35.5, 34.8, 26.3, 22.6, 16.0. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{24}H_{32}N_4O_5S$: 488.21; observed: 484.3.

di-*tert*-butyl 3,6,9,12-tetraoxatetradecanedioate (7)



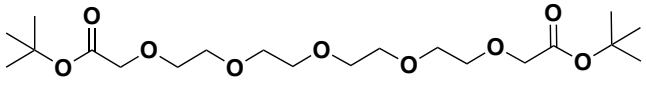
Following general method C, from triethylene glycol (1.125 g, 1 ml, 7.49 mmol, 1 eq.) in 10 ml of dioxane, NaH 60% in mineral oil (595.75 mg, 14.9 mmol, 2 eq.) and *tert*-Butyl bromoacetate (2.905 g, 2.19 ml, 14.9 mmol, 2 eq.), compound **7** was obtained as an oil after high vacuum. Yield: 538 mg, 1.42 mmol (19%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 3.81 (4H, s), 3.51 - 3.46 (12H, m), 1.26 (18H, s). Analytical data matched those previously reported [*J. Org. Chem.* **1998,** *63,* 5137-5143].

3,6,9,12-tetraoxatetradecanedioic acid (8)



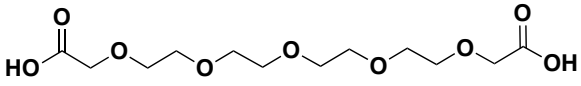
Following general method D, from compound **7** (538 mg, 1.42 mmol, 1 eq.) and trifluoroacetic acid (2 ml in 2 ml of DCM) compound **8** was obtained as an oil. Yield: 376 mg, (99%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.83 (2H, s), 4.26 - 4.08 (4H, m), 3.59 (12H, s). Analytical data matched those previously reported [*J. Org. Chem.* **1998,** *63,* 5137-5143].

di-*tert*-butyl 3,6,9,12,15-pentaoxaheptadecanedioate (9)



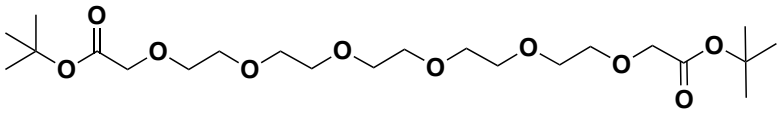
Following general method C, from tetrathylene glycol (1.125 g, 1 ml, 5.49 mmol, 1 eq.) in 10 ml of dioxane, NaH 60% in mineral oil (463 mg, 11.5 mmol, 2 eq.) and *tert*-Butyl bromoacetate (2.25 g, 1.7 ml, 11.5 mmol, 2 eq.), compound **9** was obtained as an oil after high vacuum. Yield: 500 mg, 1.18 mmol (10%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 3.86 (4H, s), 3.55 - 3.49 (16H, m), 1.31 (9H, s). Analytical data matched those previously reported [*J. Org. Chem.* **1998,** *63,* 5137-5143].

3,6,9,12,15-pentaoxaheptadecanedioic acid (10)



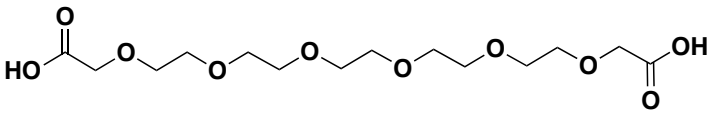
Following general method D, from compound **9** (500 mg, 1.18 mmol, 1 eq.) and trifluoroacetic acid (2 ml in 2 ml of DCM) compound **8** was obtained as an oil. Yield: 366 mg, (99%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 10.55 (2H, s), 4.17 (4H, s), 3.76 - 3.65 (16H, m). Analytical data matched those previously reported [*J. Org. Chem.* **1998,** *63,* 5137-5143].

di-*tert*-butyl 3,6,9,12,15,18-hexaoxaicosanedioate (11)



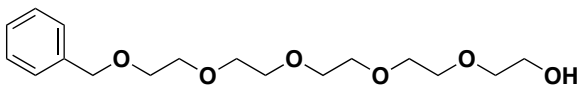
Following general method C, from pentathylene glycol (1.126 g, 1 ml, 4.72 mmol, 1 eq.) in 10 ml of dioxane, NaH 60% in mineral oil (377 mg, 9.45 mmol, 2 eq.) and *tert*-Butyl bromoacetate (1.872 g, 1.7 ml, 11.5 mmol, 2 eq.), compound **9** was obtained as an oil after high vacuum. Yield: 300 mg, 0,641 mmol (14%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 3.94 (4H, s), 3.66 - 3.56 (20H, m), 1.40 (18H, s). Analytical data matched those previously reported [*J. Org. Chem.* **1998,** *63,* 5137-5143].

3,6,9,12,15,18-hexaoxaicosanedioic acid (12)



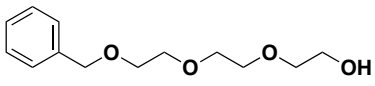
Following general method D, from compound **11** (500 mg, 1.18 mmol, 1 eq.) and trifluoroacetic acid (2 ml in 2 ml of DCM) compound **8** was obtained as an oil. Yield: 226 mg, (99%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 10.97 (2H, s), 4.19 (4H, s), 3.76 - 3.67 (20H, m). Analytical data matched those previously reported [*J. Org. Chem.* **1998,** *63,* 5137-5143].

1-phenyl-2,5,8,11,14-pentaoxahexadecan-16-ol (18)



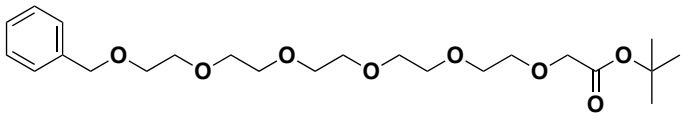
Pentaethylele glycol (9.53 g, 50 mmol, 5 eq.) was added dropwise to a suspension of NaH 60% in mineral oil (800 mg, 20 mmol, 2.5 eq.) in 20 ml of DMF at 0 °C. The resulting mixture was stirred at r.t for 1h. The reaction mixture was cooled to 0°C, benzyl chloride (1 ml, 1.1 g, 8.72 mmol, 1 eq.) was added. The resulting mixture was stirred O/N at r.t. The reaction was quenched with a saturated solution of NH_4Cl and the aqueous phase was extracted with ethyl acetate (×3). The combined organic phases were dried over $MgSO_4$ and evaporated to dryness. The resulting oil was purified by column chromatography (from 0 to 60% of ethyl acetate in heptane) to afforded the title compound as a oil. Yield: 2.055 g, 6.25 mmol (71%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.28 - 7.19 (5H, m), 4.50 (2H, s), 3.66 - 3.52 (20H, m), 2.50 (1H, s). Analytical data matched those previously reported [Patent US20160214972].

2-(2-(2-(benzyloxy)ethoxy)ethoxy)ethan-1-ol (40)



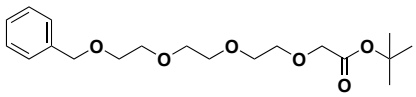
Compound **22** was synthetized starting from triethylene glycol (5 g, 33 mmol, 5 eq.) in presence of NaH in 60% mineral oil (670 mg, 16.6 mmol, 2.5 eq.) and benzyl chloride (1.1 g, 8.78 mmol, 1 eq.) in 14 ml of DMF as described for compound 17. The title compound was obtained as a oil. Yield 1.978 g, 8.23 mmol (95%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.27 - 7.19 (5H, m), 4.50 (2H, s), 3.66 - 3.52 (12H, m), 2.55 (1H, s). Analytical data matched those previously reported [Patent US20160214972].

***tert*-butyl 1-phenyl-2,5,8,11,14,17-hexaoxanonadecan-19-oate (19)**



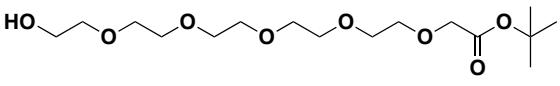
To a stirred solution of **18** (2.055 g, 6.25 mmol, 1 eq.) in 12.8 ml of DCM was added 37% solution of NaOH (12.8 ml), followed by *tert*-butylbromo acetate (4.882 g, 25 mmol, 4 eq.) and TBABr (2118 mg, 6.37 mmol, 1.02 eq.). the resulting solution was stired O/N at r.t. The reaction mixture was extracted with ethyl acetate (×3). The organic phases were combined and washed with brine (×1), dried over $MgSO_4$ and concentrate *in vacuo*. The resulting brow oil was purified by column chromatography (from 0 to 30% of ethyl acetate in petroleum) to afford the titled compound as colourless oil. Yield: 2.216 g, 5 mmol (80%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.28 - 7.20 (5H, m), 4.50 (2H, s), 3.95 (2H, s), 3.65 - 3.55 (20H, m), 1.40 (9H, s). **^{13}C NMR** (126 MHz, $CDCl_3$) δ169.7, 128.4, 127.7, 127.6, 81.5, 73.2, 70.7, 70.7, 70.6, 70.6, 69.4, 69.1, 28.1 **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{23}H_{38}O_8$: 442.26; observed: 387.2.

tert-butyl 1-phenyl-2,5,8,11-tetraoxatridecan-13-oate (41)



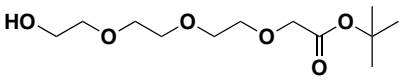
Compound **41** was synthetized starting **40** (513 mg, 2.134 mmol, 1 eq.) in presence of tertbutylbromo acetate (1.665 g, 1.228 ml, 8.54 mmol, 4 eq.), TBABr (723 mg, 2.17 mmol, 1.02 eq.), 37% NaOH aqueous solution (4.4 ml) in 4.4 ml of DCM as described for compound **19**. The title compound was obtained as a oil. Yield 551 mg, 1.56 mmol (73%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.28 - 7.26 (5H, m), 4.50 (2H, s), 3.95 (2H, s), 3.66 - 3.54 (12H, m), 1.40 (9H, s). Analytical data matched those previously reported [Patent US20160214972].

tert-butyl 17-hydroxy-3,6,9,12,15-pentaoxaheptadecanoate(20)



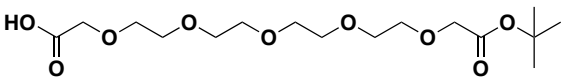
19 (2.216 g, 5 mmol, 1 eq.) was dissolved in 75 ml of ethanol, Pd/C (10 wt %) was added and the resulting mixture was place under hydrogen and stirred at r.t. until complete conversion of the starting material. The reaction mixture was filtered through celite, the celite pad was washed few times using ethanol. The filtrate was concentrated in vacuum to give an oil. Yield: 1764 g, 5 mmol (quantitative). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 3.95 (2H, s), 3.68 - 3.53 (20H, m), 1.41 (9H, s). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 165.0, 76.8, 67.8, 66.0, 65.8, 65.8, 65.8, 65.8, 65.8, 65.6, 64.3, 57.0, 23.4**.**

tert-butyl 2-(2-(2-(2-hydroxyethoxy)ethoxy)ethoxy)acetate (42)



Compound **42** was synthesized starting from **41** (551 mg, 1.56 mmol, 1 eq.) in presence of Pd/C (10 wt %) in 10 ml of ethanol as described for compound **19**. The title compound was obtained as an oil. Yield 381 g, 1.46 mmol (93%).
^{1}H NMR (500 MHz, $CDCl_3$) δ 3.96 (2H, s), 3.68 - 3.54 (12H, m), 1.41 (9H, s). Analytical data matched those previously reported [Patent US20160214972].

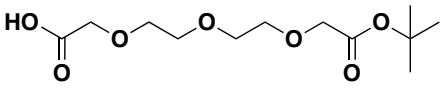
19,19-dimethyl-17-oxo-3,6,9,12,15,18-hexaoxaicosanoic acid (21)



BAIB (3.546 g, 11.01 mmol, 2.2 eq.) and TEMPO (171.87 mg, 1.10 mmol, 0.22 eq.) were added to a solution of ACN/H_2O 1:1 containing **20** (1.764 g, 5 mmol, 1 eq.). The resulting mixture was stirred at r.t until complete conversion of the starting material. The crude was purified using ISOLUTE® PE-AX anion exchange column. The column was equilibrate with methanol, the reaction mixture poured in the column and let it adsorbed in the pad. The column was washed with methanol (×3) to elute all the unbound material. Then, the titled product was

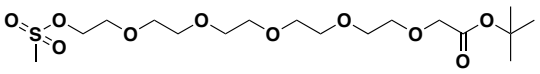
eluted using a 50% solution of formic acid in methanol. The organic phase was evaporated to dryness to afford the title compound as oil. Yield: 1.200 g, 3.27 mmol (65%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 4.10 (2H, s), 3.95 (2H, s), 3.72 - 3.58 (16H, m), 1.41 (9H, s).

13,13-dimethyl-11-oxo-3,6,9,12-tetraoxatetradecanoic acid (43)



Compound **43** was synthetized starting from **42** (380 mg, 1.44 mmol, 1 eq.) in presence of BAIB (1.018 g, 3.16 mmol, 2.2 eq.) and TEMPO (49 mg, 0.316 mmol, 0.22 eq.) in 3 ml of a solution of ACN/H_2O 1:1 as described for compound 20. The title compound was obtained as an oil. Yield 210 mg, 0.75 mmol (52%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.09 (2H, s), 3.95 (2H, s), 3.72 - 3.64 (8H, m), 1.41 (9H, s).

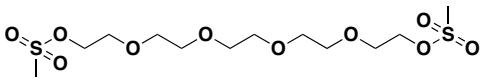
***tert*-butyl 17-((methylsulfonyl)oxy)-3,6,9,12,15-pentaoxaheptadecanoate (33)**



33 (251 mg, 0.712 mmol, 1 eq.) was dissolved in 1.4 ml of dry DCM. The temperature of the resulting mixture was cooled down to 0°C and methanesulfonyl chloride (122.3 mg, 0.082 ml, 1.068 mmol, 1.5 eq.) was added followed by triethylamine (216.14 mg, 0.3 ml, 2.136 ml, 3 eq.). The resulting mixture was stirred at 0 °C for 4 h. A 10% aqueous solution of $NaHSO_4$ was added till pH=3. The aqueous phase was extracted with DCM (x5). The organic phases were combined, dried over $MgSO_4$ and concentrated in vacuum to afford the title compound as a orange oil. Yield: 276 mg, 0.641 mmol (90%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.32 - 4.30 (2H, m), 3.95 (2H, s), 3.71 - 3.57 (18H, m), 3.02

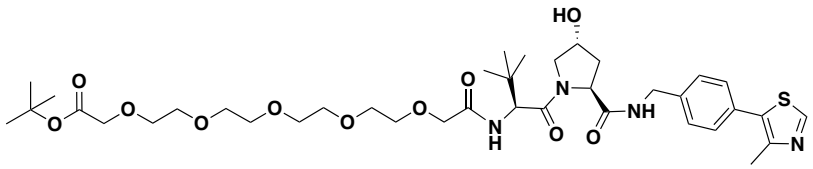
(3H, s), 1.41 (9H, s). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 169.7, 81.5, 70.72, 70.65, 70.61, 70.58, 70.5, 69.3, 69.0, 37.7, 28.1.

3,6,9,12-tetraoxatetradecane-1,14-diyl dimethanesulfonate (37)



Pentaethylene glycole (476.56 mg, 0.423 ml, 2 mmol, 1 eq.) was dissolved in 4 ml of dry DCM. The temperature of the resulting mixture was cooled down to 0°C and methanesulfonyl chloride (687.3mg, 0.464 ml, 16 mmol, 3 eq.) was added followed by triethylamine (1011.9 g, 1.39 ml, 10 mmol, 5 eq.). The resulting mixture was stirred at 0 °C for 4 h. A 10% aqueous solution of $NaHSO_4$ was added till pH=3. The aqueous phase was extracted with DCM (x5). The organic phases were combined, dried over $MgSO_4$ and concentrated in vacuum to afford the title compound as a orange oil. Yield: 701 mg, 1.77 mmol (89%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.33 - 4.30 (4H, m), 3.72 - 3.69 (4H, m), 3.62 - 3.56 (12H, m), 3.02 (6H, s). Analytical data matched those previously reported [1. Kimura et al. *J. Polym. Sci. Part A: Polym. Chem.* **54,** (2016).]

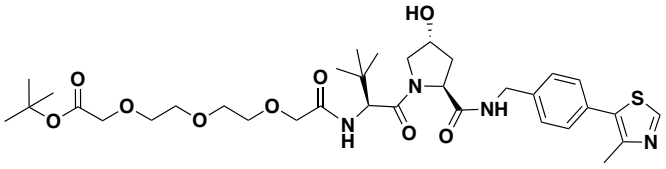
***tert*-butyl (*S*)-19-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-20,20-dimethyl-17-oxo-3,6,9,12,15-pentaoxa-18-azahenicosanoate (22)**



Following general method E, from compound **6** (100 mg, 0.215 mmol, 1 eq.), compound **21** (78.8 mg , 0.215 mmol, 1 eq.), HATU (81.74 mg, 0.215 mmol, 1 eq.), HOAT (29.26 mg, 0.215 mmol, 1 eq.), DIPEA (80.13 mg, 0.106 ml, 0.645 mmol, 3 eq.) in 1.5 ml of DMF, **22** was obtained as a white solid. Yield: 75.6 mg,

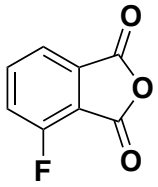
0.094 mmol (44%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.94 (1H, s), 7.42 - 7.36 (1H, m), 7.32 - 7.20 (5H, m), 4.67 - 4.63 (1H, m), 4.53 - 4.42 (3H, m), 4.31 - 4.25 (1H, m), 4.01 - 3.87 (5H, m), 3.64 - 3.55 (18H, m), 2.47 - 2.34 (4H, m), 2.11 - 2.04 (1H, m), 1.40 (9H, s), 0.91 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.3, 171.1, 170.5, 170.0, 151.7, 139.1, 129.4, 128.3, 82.0, 71.1, 70.6, 70.4, 70.4, 70.3, 70.3, 70.2, 70.2, 68.9, 58.7, 57.3, 56.8, 43.1, 36.3, 35.1, 28.1, 26.4, 15.1. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{38}H_{58}N_4O_{11}S_2$: 778.38; observed: 779.4.

tert-butyl (_S_)-16-((2_S_,4_R_)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-17,17-dimethyl-14-oxo-3,6,9,12-tetraoxa-15-azaoctadecanoate (44)



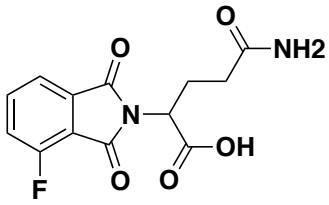
Following general method E, from compound **6** (100 mg, 0.215 mmol, 1 eq.), compound **43** (59.83 mg, 0.215 mmol, 1 eq.), HATU (81.74 mg, 0.215 mmol, 1 eq.), HOAT (29.26 mg, 0.215 mmol, 1 eq.), DIPEA (80.13 mg, 0.106 ml, 0.645 mmol, 3 eq.) in 1.5 ml of DMF, **44** was obtained as a white solid. Yield: 72.8 mg, 0.109 mmol (51%).%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.94 (1H, s), 7.42 - 7.36 (1H, m), 7.32 - 7.20 (5H, m), 4.67 - 4.63 (1H, m), 4.53 - 4.42 (3H, m), 4.31 - 4.25 (1H, m), 4.01 - 3.87 (5H, m), 3.64 - 3.55 (18H, m), 2.47 - 2.34 (4H, m), 2.11 - 2.04 (1H, m), 1.40 (9H, s), 0.91 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.3, 171.1, 170.5, 170.0, 151.7, 139.1, 129.4, 128.3, 82.0, 71.1, 70.6, 70.4, 70.4, 70.3, 70.3, 70.2, 70.2, 68.9, 58.7, 57.3, 56.8, 43.1, 36.3, 35.1, 28.1, 26.4, 15.1. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{34}H_{50}N_4O_9S_2$: 690.33; observed: 691.35.

4-fluoroisobenzofuran-1,3-dione (46)



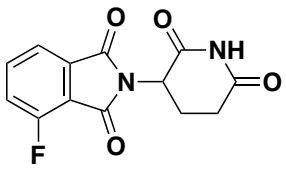
A mixture of 3-fluorophtalic acid (2.5 g, 13.58 mmol, 1 eq.) and acetic anhydride (20 ml) was refluxed for 2 h. The volaties were removed in vacuum and the residue was crystallized in acetic anhydride to afford the title product as a white solid. Yield: 1.490 g, 8.97 mmol (66%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_6H_3FO_3$: 166.01; observed: 167.1. Analytical data matched those previously reported[118].

5-amino-2-(4-fluoro-1,3-dioxoisoindolin-2-yl)-5-oxopentanoic acid (47)



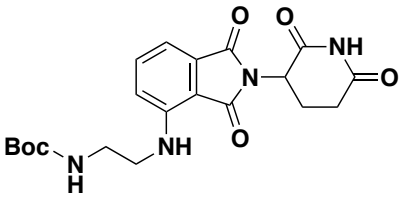
A mixture of **46** (1.49 g, 8.97 mmol, 1 eq.) and L-glutamine (1.309 g, 8.97 mmol, 1 eq.) in 7.5 ml of dry DMF was stirred at 90 ^oC for 8 h. The solvent was removed under reduced pressure and re-dissolved in 7.5 ml of HCl 4 N and stirred for additional 8h. The resulting precipitate was collected by filtration, washed with water and dried under reduced pressure. Yield: 1.77 g, 6 mmol (65%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{13}H_{11}FN_2O_5$: 294.07; observed: 295.1.
Analytical data matched those previously reported [118].

2-(2,6-dioxopiperidin-3-yl)-4-fluoroisoindoline-1,3-dione (48)



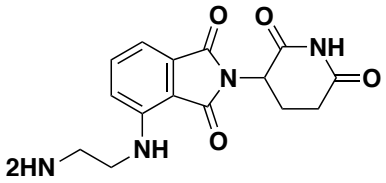
A mixture of **47** (1.77 g, 6.01 mmol, 1 eq.), 1,1'-carbonyldiimidazole (CDI) (1.169 g, 7.21 mmol, 1.2 eq.) and 4-dimethylaminopyridine (DMAP) (74.52 mg, 0.6 mmol, 0.1 eq.) in acetonitrile (3.8 mL) was refluxed for 2.5 h. The reaction mixture was cooled to r. t. and the resulting solid was collected by filtration, and washed with acetonitrile to afford the crude product, which was purified by column chromatography (using 1-10% of MeOH in DCM) to afford **30** as a light yellow solid. Yield: 827 mg, 3 mmol (50%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{13}H_9FN_2O_4$: 276.05; observed: 277.1. Analytical data matched those previously reported[118].

***tert*-butyl (2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethyl)carbamate (49)**



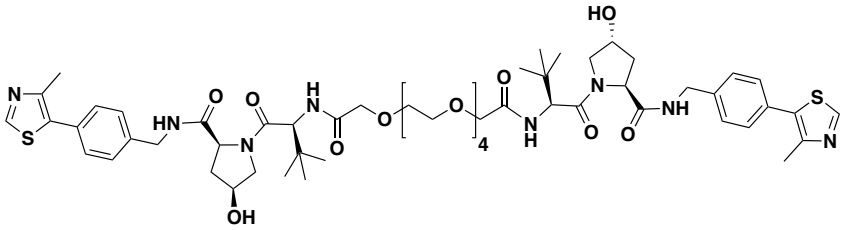
A mixture **30** (200 mg, 0.724 mmol, 1 eq.), *N*-Boc-ethylenediamine (127.6 mg, 0.796 mmol, 1.1 eq.), DIPEA (134 mg, 0.177 ml, 1.086 mmol, 1.5 eq.) in 5 ml of DMF was stirred at r.t. until complete conversion of starting material. The solvent was removed under reduce pressure. The crude product was purified by HPLC using a gradient of 5% to 95% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield the titled compound. Yield: 67 mg, 0.160 mmol (22%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{20}H_{24}N_4O_6$: 416.17; observed: 317.1.

4-((2-aminoethyl)amino)-2-(2,6-dioxopiperidin-3-yl)isoindoline-1,3-dione (50)



Following general method B, from **31** (67 mg, 0.160 mmol, 1 eq.) in 2 ml of a 1:1 mixture of 4 M HCl in dioxane and DCM, compound **32** was obtained as yellow solid after high vacuum (60 mg, 0.189 mmol, quantitative). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{15}H_{16}N_4O_4$: 316.12.; observed: 317.1.

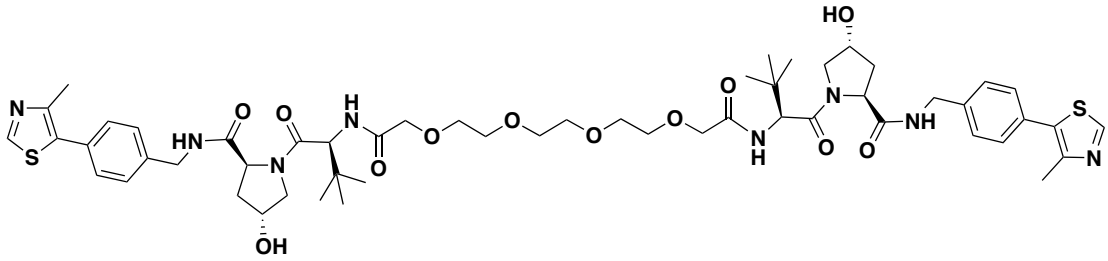
***N*^{1}-((*R*)-1-((2*R*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-*N*^{17}-((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-3,6,9,12,15-pentaoxaheptadecanediamide (CMP99)**



Following general method E, from compound **16** (2.77 mg, 0.006 mmol, 1 eq.), compound **23** (5.5 mg, 0.006 mmol, 1 eq.), HATU (2.28 mg, 0.0.006 mmol, 1 eq.), HOAT (1 mg, 0.0.006 mmol, 1 eq.), DIPEA (2.23 mg, 0.002 ml, 0.018 mmol, 3 eq.), **CMP99** was obtained as a white solid. Yield: 4.5 mg, 0.004 mmol (66%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.68 - 8.65 (2H, m), 7.30 - 7.27 (9H, m), 7.11 (1H, d, *J*=8.3 Hz), 4.69 - 4.57 (3H, m), 4.53 - 4.38 (5H, m), 4.30 - 4.19 (2H, m), 3.98 - 3.52 (27H, m), 2.45 (6H, s), 2.25 - 2.04 (4H, m), 0.90 (9H, s), 0.88 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 173.1, 172.4, 170.7, 170.3, 153.3, 153.2, 144.5, 140.0, 134.0, 129.2,

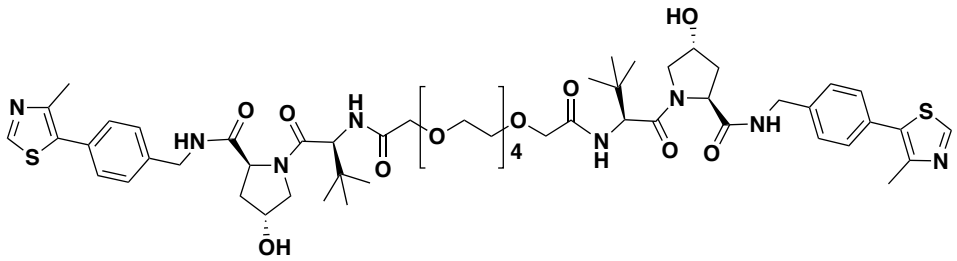
129.0, 128.4, 128.3, 127.8, 70.9, 70.5, 70.2, 70.1, 69.7, 68.2, 67.7, 59.7, 59.4, 56.8, 56.7, 54.9, 42.9, 42.3, 39.9, 37.6, 36.3, 35.7, 34.7, 25.6, 25.5, 13.1. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{56}H_{78}N_8O_{13}S_2$: 1134.51; observed: 1135.58.

N^1,*N*^{14}-bis((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-3,6,9,12-tetraoxatetradecanediamide (CM09)



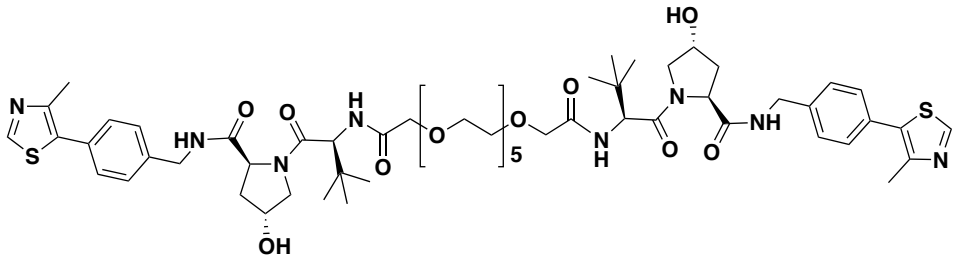
Following general method E, from compound **8** (5.43 mg, 0.018 mmol, 1 eq.), compound CM09 was obtained as a white solid. Yield: 8 mg, 0.007 mmol (40%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.61 (2H, s), 7.48 - 7.45 (2H, m), 7.31 - 7.27 (8H, m), 7.23 (2H, d, *J*=10.2 Hz), 4.64 - 4.59 (2H, m), 4.52 - 4.46 (4H, m), 4.41 - 4.38 (2H, m), 4.31 - 4.25 (2H, m), 4.01 - 3.94 (4H, m), 3.82 (2H, d, *J*=15.7 Hz), 3.62 - 3.52 (12H, m), 2.45 (6H, s), 2.42 - 2.34 (2H, m), 2.12 - 2.06 (2H, m), 1.19 (2H, s), 0.89 (18H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ. 170.2, 169.9, 169.6, 149.3, 147.5, 137.3, 130.6, 129.9, 128.4, 127.1, 69.9, 69.5, 69.3, 69.1, 57.6, 56.1, 55.9, 42.2, 35.5, 34.6, 25.4, 15.1. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{54}H_{74}N_8O_{12}S_2$: 1090.49; observed: 1091.4.

N^1,*N*^{17}-bis((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-3,6,9,12,15-pentaoxaheptadecanediamide (CM10)



Following general method E, from compound **10** (5.58 mg, 0.018 mmol, 1 eq.), compound CM09 was obtained as a white solid. Yield: 6 mg, 0.005 mmol (30%). **^{1}H NMR** (400 MHz, MeOD) δ 9.28 (2H, s), 7.43 - 7.36 (10H, m), 5.39 (2H, s), 4.77 (10H, s), 4.59 (2H, s), 4.50 - 4.43 (4H, m), 4.42 - 4.38 (2H, m), 4.26 (2H, d, *J*=17.2 Hz), 3.96 - 3.92 (4H, m), 3.77 (2H, d, *J*=11.1 Hz), 3.73 - 3.68 (2H, m), 3.56 (16H, m), 3.22 - 3.20 (10H, m), 2.43 (6H, s), 2.16 - 2.14 (2H, m), 2.13 (2H, m), 2.02 - 1.95 (2H, m); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 173.1, 172.4, 170.7, 170.3, 153.3, 153.2, 144.5, 140.0, 134.0, 129.2, 129.0, 128.4, 128.3, 127.8, 70.9, 70.5, 70.2, 70.1, 69.7, 68.2, 67.7, 59.7, 59.4, 56.8, 56.7, 54.9, 42.9, 42.3, 39.9, 37.6, 36.3, 35.7, 34.7, 25.6, 25.5, 13.1. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{56}H_{78}N_8O_{13}S_2$: 1134.51; observed: 1135.55.

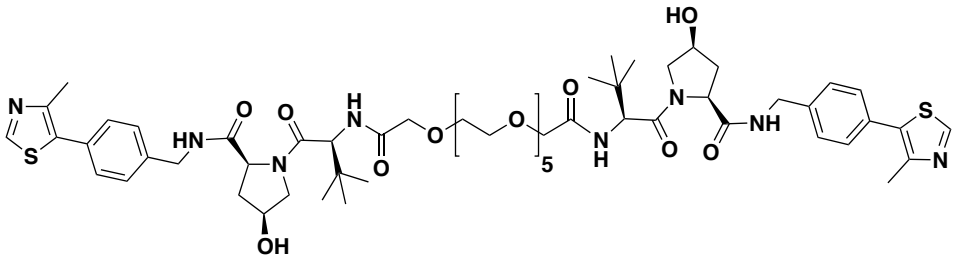
N^1,N^{20}-bis((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-3,6,9,12,15,18-hexaoxaicosanediamide (CM11)



Following general method E, from compound **12** (5.58 mg, 0.018 mmol, 1 eq.), compound CM09 was obtained as a white solid. Yield: 11.74 mg, 0.0099 mmol (55%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.61 (2H, s), 7.41 - 7.38 (2H, m), 7.29 (10H, t, *J*=7.6 Hz), 4.66 - 4.61 (2H, m), 4.49 - 4.41 (6H, m), 4.35 - 4.29 (2H, m), 3.98 - 3.91

(6H, m), 3.62 - 3.50 (24H, m), 2.45 (6H, s), 2.42 - 2.35 (2H, m), 2.11 - 2.06 (2H, m), 0.88 (18H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.2, 170.9, 170.4, 150.3, 148.5, 138.3, 131.6, 130.9, 129.5, 128.1, 71.2, 70.61, 70.59, 70.5, 70.4, 70.3, 58.6, 57.0, 43.2, 36.5, 35.6, 26.4, 16.1. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{58}H_{82}N_8O_{14}S_2$: 1178.54; observed: 1179.60.

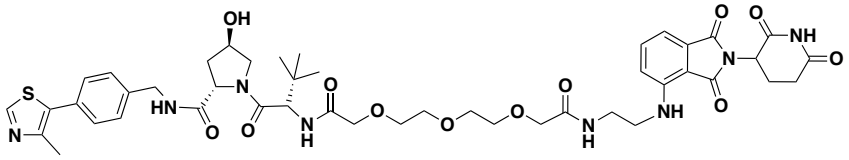
N^1,N^{20}-bis((*S*)-1-((2*S*,4*S*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-3,6,9,12,15,18-hexaoxaicosanediamide (CMP98)



Following general method E, from compound **12** (7.12 mg, 0.028 mmol, 1 eq.), compound **16** (18.06, 0.040 mmol, 2.1 eq.), HATU (15.2 mg, 0.040 mmol, 2 eq.), HOAT (5.44 mg, 0.040 mmol, 2 eq.), DIPEA (7.45 mg, 0.0010 ml, 3 eq.), compound CMP98 was obtained as a white solid. Yield: 10.58 mg, 0.0089 mmol (45%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 9.09 (2H, s), 8.02 (2H, s), 7.31 (4H, d, *J*=8.5 Hz), 7.22 (4H, d, *J*=8.0 Hz), 7.16 (2H, d, *J*=9.2 Hz), 4.75 - 4.64 (4H, m), 4.51 (2H, d, *J*=8.9 Hz), 4.41 - 4.37 (2H, m), 4.24 - 4.17 (2H, m), 3.94 (4H, d, *J*=3.2 Hz), 3.84 - 3.81 (4H, m), 3.62 - 3.54 (20H, m), 2.49 - 2.47 (2H, m), 2.44 (6H, s), 2.26 - 2.17 (4H, m), 0.93 (18H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 173.2, 171.5, 169.7, 151.8, 138.8, 132.9, 129.5, 129.2, 128.3, 71.2, 71.1, 70.6, 70.48, 70.45, 70.4, 70.3, 59.9, 58.5, 56.5, 43.2, 35.6, 35.2, 26.4, 15.0. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{58}H_{82}N_8O_{14}S_2$: 1178.54; observed: 1179.60.

(2*S*,4*R*)-1-((2*R*)-2-(*tert*-butyl)-17-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)-4,14-dioxo-6,9,12-trioxa-3,15-

diazaheptadecanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (CMP85)

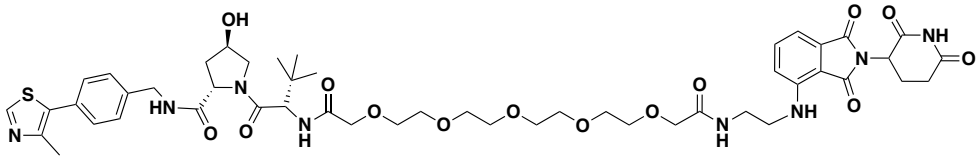


Following general method D, from compound 44 (72.3 mg, 0.105 mmol, 1 eq.) and trifluoroacetic acid (1 ml in 1 ml of DCM) carboxylic acid derivative was obtained as an oil. The compound was used for the next step without further purification. Yield: 99.3 mg, 0.156 mmol (quantitative). HRMS (ESI) m/z: $[M+H]^+$ calculated for: $C_{30}H_{42}N_4O_9S$: 634.27; observed: 635.3.

To a solution of the just obtained carboxylic acid (21.16 mg, 0.028 mmol,1 eq.) in 0.5 ml DMF was added HATU (10.64 mg, 0.028 mmol, 1 eq.) and HOAT (3.81 mg, 0.028 mmol, 1 eq.). The resulting mixture was stirred at room temperature for 5 min. Compound **50** (10 mg, 00.028 mmol, 1 eq.) was added and the pH of the reaction mixture was adjusted to >9 by addition of DIPEA (3 eq.). The mixture was stirred at room temperature until no presence of the starting materials was detected by LC-MS. The solvent was evaporated under reduced pressure to give the corresponding crude, which was purified by HPLC using a gradient of 5% to 95% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield the final compound as a yellow solid. Yield: 25 mg, 0.026 mmol (95%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.61 (1H, d, *J*=2.6 Hz), 7.60 - 7.54 (1H, m), 7.48 - 7.38 (2H, m), 7.30 - 7.25 (4H, m), 7.01 (1H, d, *J*=7.4 Hz), 6.90 (1H, d, *J*=8.4 Hz), 6.41 - 6.34 (1H, m), 4.83 - 4.78 (1H, m), 4.59 - 4.50 (2H, m), 4.48 - 4.44 (1H, m), 4.26 - 4.21 (1H, m), 3.97 - 3.81 (5H, m), 3.63 - 3.40 (12H, m), 2.67 (3H, t, *J*=46.8 Hz), 2.44 (3H, d, *J*=1.8 Hz), 2.35 - 2.29 (1H, m), 2.12 - 2.05 (1H, m), 2.03 - 1.97 (2H, m), 0.89 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ172.2, 171.4, 171.3, 171.3, 170.9, 170.8, 170.3, 170.2, 169.4, 169.2, 169.1, 167.6, 150.4, 150.4, 148.4, 148.4, 146.8, 146.8, 138.5, 138.5, 136.2, 132.5, 132.5, 131.7, 131.7, 130.7, 130.7, 129.4, 129.3, 128.1, 116.7, 111.8, 111.8, 110.3, 110.2, 71.1, 70.9, 70.8, 70.3, 70.2, 70.2, 70.1, 60.4, 59.0, 59.0, 57.0, 56.9, 56.8, 50.7, 48.9, 48.9, 43.1, 41.7, 38.8, 38.6, 36.6, 35.6, 35.5, 31.5, 26.5, 26.3,

22.7, 22.7, 16.0, 14.2. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{45}H_{56}N_8O_{12}S$: 932.37; observed: 933.3 .

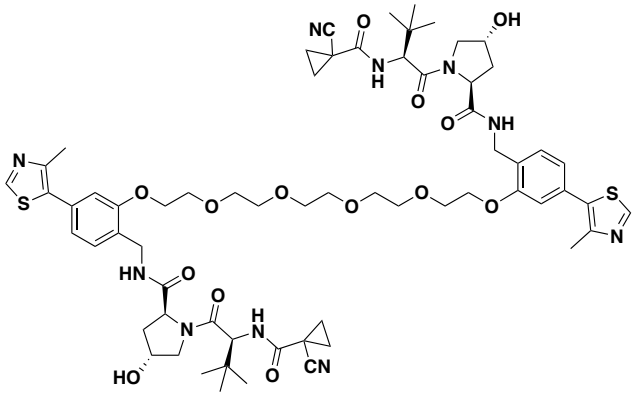
N^1-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethyl)-N^{17}-((*R*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)-3,6,9,12,15-pentaoxaheptadecanediamide (CMP86)



Following general method D, from compound **22** (75.6 mg, 0.094 mmol, 1 eq.) and trifluoroacetic acid (1 ml in 1 ml of DCM), the carboxylic acid derivative was obtained as an oil. The compound was used for the next step without further purification. Yield: 86.8 mg, 0.120mmol (quantitative). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{34}H_{50}N_4O_{11}S_2$: 722.32; observed: 723.3.

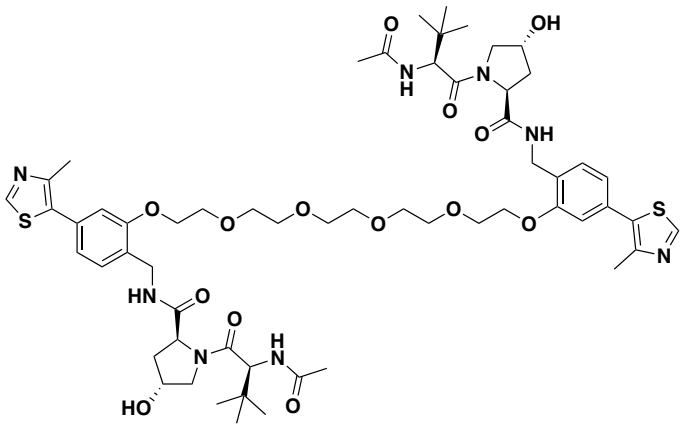
Following the method used for the synthesis of CMP85, from the carboxylic acid (23.88 mg, 0.028 mmol, 1 eq.), **50** (10 mg, 0.028 mmol, 1 eq.), HATU (10.64 mg, 0.028 mmol, 1 eq.) and HOAT (3.81 mg, 0.028 mmol, 1 eq.) in 0.5 ml of DMF, the title compound was obtained as a white solid. Yield: 24 mg, 0.023 mmol (84%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.70 (1H, d, *J*=8.2 Hz), 7.64 - 7.58 (1H, m), 7.55 - 7.49 (1H, m), 7.43 - 7.40 (1H, m), 7.32 - 7.27 (5H, m), 7.02 - 6.95 (2H, m), 4.84 - 4.77 (1H, m), 4.66 - 4.59 (1H, m), 4.56 - 4.48 (2H, m), 4.28 - 4.23 (1H, m), 3.95 - 3.82 (4H, m), 3.63 - 3.36 (22H, m), 2.46 - 2.44 (4H, m), 2.41 - 2.34 (1H, m), 2.14 - 2.08 (1H, m), 0.90 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ172.0, 171.3, 171.25, 171.21, 171.17, 171.1, 170.5, 170.3, 169.4, 168.9, 167.6, 150.7, 150.6, 147.9, 146.8, 146.8, 138.7, 138.6, 136.20, 136.16, 132.5, 132.0, 130.45, 130.40, 129.4, 129.3, 128.2, 117.0, 111.7, 110.1, 70.9, 70.55, 70.45, 70.42, 70.39, 70.3, 70.2, 70.1, 70.0, 59.0, 58.8, 57.0, 56.9, 48.9, 43.2, 41.9, 38.4, 38.3, 36.5, 35.4, 35.3, 31.5, 26.4, 22.7, 15.82, 15.79. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{49}H_{64}N_8O_{14}S$: 1020.43; observed: 1021.6.

(2*S*,2'*S*,4*R*,4'*R*)-*N*,*N*'-((((3,6,9,12-tetraoxatetradecane-1,14-diyl)bis(oxy))bis(4-(4-methylthiazol-5-yl)-2,1-phenylene))bis(methylene))bis(1-((*S*)-2-(1-cyanocyclopropane-1-carboxamido)-3,3-dimethylbutanoyl)-4-hydroxypyrrolidine-2-carboxamide) (CMP108)



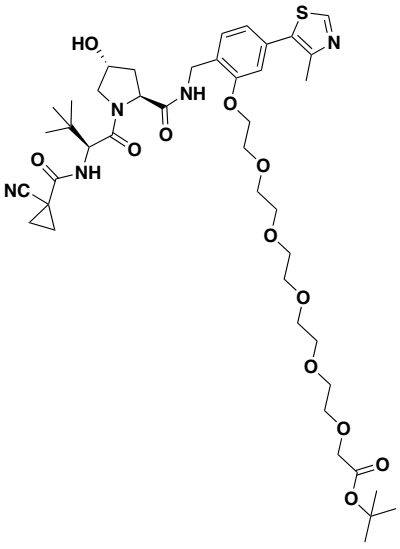
A mixture of **31** (27 mg, 0.05 mmol, 2 eq.), **37** (11.83 mg, 0.03 mmol, 1.2 eq.) and K_2CO_3 (41.46 mg, 0.3 mmol, 6 eq.) was stirred O/N at 70 ^{o}C. The reaction mixture was filtered off to afford the crude product, which was purified by HPLC using a gradient of 5% to 95% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield the final compound as a white solid. Yield: 9.1 mg, 0.007 mmol (28%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.61 (2H, s), 7.41 - 7.38 (2H, m), 7.26 (2H, d, *J*=8.1 Hz), 7.00 (2H, d, *J*=8.1 Hz), 6.91 - 6.88 (2H, m), 6.85 - 6.81 (2H, m), 4.57 - 4.52 (2H, m), 4.44 - 4.36 (8H, m), 4.19 - 4.08 (4H, m), 3.89 - 3.53 (22H, m), 2.45 (6H, s), 2.24 - 2.17 (2H, m), 2.08 - 2.02 (2H, m), 1.61 - 1.37 (8H, m), 0.88 (18H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 170.9, 170.5, 170.2, 170.0, 165.4, 156.9, 150.4, 148.5, 132.3, 131.7, 130.0, 126.9, 122.0, 119.6, 112.9, 71.3, 70.7, 70.6, 70.5, 70.4, 70.4, 70.3, 70.2, 69.6, 67.9, 58.9, 58.7, 58.4, 56.6, 42.8, 39.2, 37.0, 36.6, 36.0, 35.7, 26.3, 17.9, 17.7, 16.2, 13.7. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{64}H_{84}N_{10}O_{14}S_2$: 1280.56; observed: 1281.66.

(2*S*,2'*S*,4*R*,4'*R*)-*N*,*N*'-((((3,6,9,12-tetraoxatetradecane-1,14-diyl)bis(oxy))bis(4-(4-methylthiazol-5-yl)-2,1-phenylene))bis(methylene))bis(1-((*S*)-2-acetamido-3,3-dimethylbutanoyl)-4-hydroxypyrrolidine-2-carboxamide) (CMP106)



Following the same procedure used for the synthesis of **CMP108,** from **32** (24.3 mg, 0.05 mmol, 2 eq.), **37** (11.83 mg, 0.03 mmol, 1.2 eq.) and and K_2CO_3 (41.46 mg, 0.3 mmol, 6 eq.), the title compound was obtained as a white solid. Yield: 7.8 mg, 0.006 mmol (26%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.60 (2H, s), 7.39 - 7.35 (2H, m), 7.26 (2H, d, *J*=7.6 Hz), 6.91 - 6.88 (2H, m), 6.83 - 6.80 (2H, m), 6.36 - 6.13 (2H, m), 4.60 - 4.32 (10H, m), 4.18 - 4.05 (4H, m), 3.97 - 3.79 (6H, m), 3.71 - 3.54 (18H, m), 2.44 (6H, s), 2.17 - 1.86 (8H, m), 0.87 (18H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.3, 171.1, 171.0, 170.7, 170.5, 156.8, 156.8, 150.3, 148.5, 132.2, 131.7, 130.0, 129.8, 127.1, 126.9, 122.1, 122.0, 112.8, 112.8, 71.3, 70.7, 70.6, 70.5, 70.5, 70.5, 70.4 , 70.2, 70.1, 69.7, 67.9, 58.9, 58.6, 57.6, 57.5, 56.9, 56.7, 42.7, 39.1, 39.0, 37.1, 36.4, 35.4, 35.1, 26.4, 26.4, 23.2, 23.1, 16.2. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{58}H_{82}N_8O_{14}S_2$: 1178.54; observed: 1281.66.

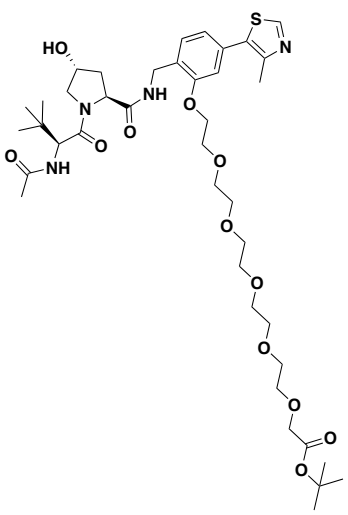
***tert*-butyl (14-(2-(((2*S*,4*R*)-1-((*S*)-2-(1-cyanocyclopropane-1-carboxamido)-3,3-dimethylbutanoyl)-4-hydroxypyrrolidine-2-carboxamido)methyl)-5-(4-methylthiazol-5-yl)phenoxy)-3,6,9,12-tetraoxatetradecyl) carbonate (33)**



Following the same procedure used for the synthesis of **CMP108,** from **31**(27 mg, 0.05 mmol, 1 eq.), **33** (26 mg, 0.06 mmol, 1.2 eq.) and and K_2CO_3 (20.73 mg, 0.15 mmol, 3 eq.), the title compound was obtained as a white solid. Yield: 17 mg, 0.02 mmol (33%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.61 (1H, s), 7.33 - 7.25 (2H, m), 6.97 (1H, d, *J*=9.1 Hz), 6.92 - 6.89 (1H, m), 6.84 (1H, d, *J*=1.5 Hz), 4.59 - 4.55 (1H, m), 4.45 - 4.38 (4H, m), 4.22 -

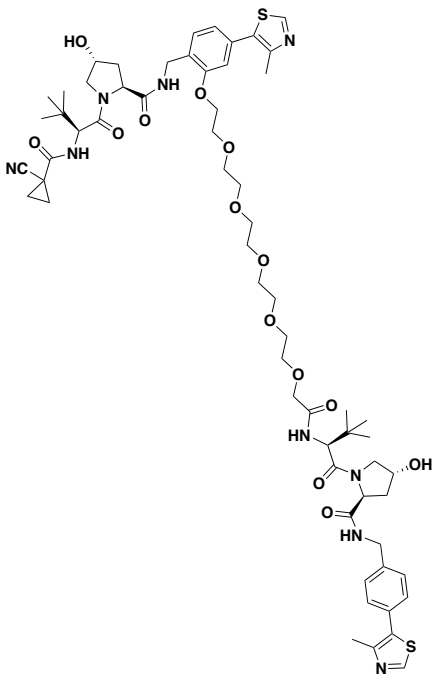
4.10 (2H, m), 3.93 - 3.54 (24H, m), 2.46 (3H, s), 2.32 - 2.24 (1H, m), 2.10 - 2.04 (1H, m), 1.63 - 1.52 (2H, m), 1.45 - 1.39 (12H, m), 0.87 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 170.6, 170.1, 169.7, 165.4, 156.9, 150.3, 148.5, 132.3, 131.7, 130.0, 126.9, 122.0, 119.7, 112.9, 81.7, 70.72, 70.66, 70.5, 70.4, 70.3, 69.6, 69.0, 68.0, 58.8, 58.4, 56.6, 39.3, 36.7, 35.8, 28.1, 26.3, 17.8, 16.2, 13.7. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{43}H_{63}N_5O_{12}S$: 873.42; observed: 874.49.

tert-butyl 17-(2-(((2*S*,4*R*)-1-((*S*)-2-acetamido-3,3-dimethylbutanoyl)-4-hydroxypyrrolidine-2-carboxamido)methyl)-5-(4-methylthiazol-5-yl)phenoxy)-3,6,9,12,15-pentaoxaheptadecanoate (34)



Following the same procedure used for the synthesis of **CMP108,** from **32** (24.3 mg, 0.05 mmol, 1 eq.), **33** (26 mg, 0.06 mmol, 1.2 eq.) and K_2CO_3 (20.73 mg, 0.15 mmol, 3 eq.), the title compound was obtained as a white solid. Yield: 17 mg, 0.021 mmol (33%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.61 (1H, s), 7.27 (2H, d, *J*=7.3 Hz), 6.91 - 6.88 (1H, m), 6.82 (1H, d, *J*=1.8 Hz), 4.59 - 4.55 (1H, m), 4.47 - 4.38 (2H, m), 4.33 - 4.30 (1H, m), 4.16 - 4.10 (2H, m), 3.94 (2H, d, *J*=6.5 Hz), 3.64 - 3.57 (20H, m), 3.02 (2H, s), 2.45 (3H, s), 2.32 - 2.24 (1H, m), 2.10 - 2.04 (1H, m), 1.92 (3H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.2, 170.8, 170.4, 169.7, 156.8, 150.3, 148.5, 132.2, 131.7, 129.8, 126.9, 122.0, 112.8, 81.6, 70.8, 70.71, 70.69, 70.60, 70.57, 70.55, 70.52, 70.49, 70.47, 70.1, 69.6, 69.3, 69.02, 68.98, 67.9, 58.6, 57.5, 56.7, 39.0, 37.7, 36.5, 35.2, 28.1, 26.4, 23.2, 16.1. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{40}H_{62}N_4O_{12}S$: 822.41; observed: 823.5.

(2*S*,4*R*)-1-((*S*)-2-(*tert*-butyl)-20-(2-(((2*S*,4*R*)-1-((*S*)-2-(1-cyanocyclopropane-1-carboxamido)-3,3-dimethylbutanoyl)-4-hydroxypyrrolidine-2-carboxamido)methyl)-5-(4-methylthiazol-5-yl)phenoxy)-4-oxo-6,9,12,15,18-pentaoxa-3-azaicosanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (CMP113)

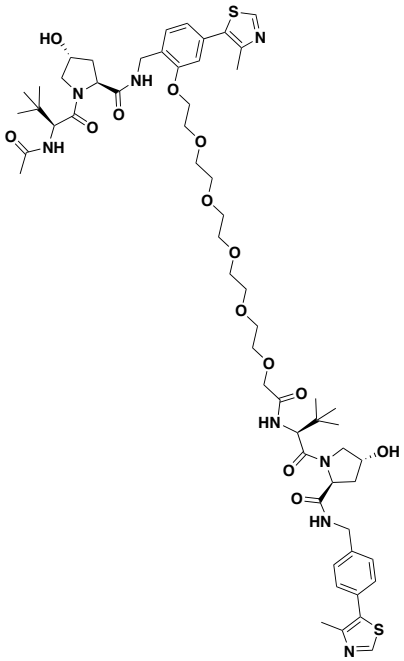


Following general method D, from compound **37** (17 mg, 0.02 mmol, 1 eq.) and trifluoroacetic acid (0.5 ml in 0.5 ml of DCM), the carboxyli acid derivative was obtained as an oil. Yield: 15.7 mg, 0.019 mmol (quantitative). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{39}H_{55}N_5O_{12}S$: 817.36; observed: 818.4.

To a solution of obtained carboxylic acid (15.7 mg, 0.019 mmol, 1 eq.) in 0.5 ml DMF was added HATU (7.22 mg, 0.019 mmol, 1 eq.) and HOAT (2.58 mg, 0.019 mmol, 1 eq.). The resulting mixture was stirred at room temperature for 5 min. Compound **6** (8.73 mg, 0.019 mmol, 1 eq.) was added and the pH of the reaction mixture was adjusted to >9 by addition of DIPEA (3 eq.). The mixture was stirred at room temperature until no presence of the starting materials was detected by LC-MS. The solvent was evaporated under reduced pressure to give the corresponding crude, which was purified by HPLC using a gradient of 5% to 95% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield the final compound as awhite solid. Yield: 6.3 mg, 0.005 mmol (27%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.61 (2H, s), 7.58 - 7.54 (1H, m), 7.31 - 7.25 (5H, m), 7.02 (1H, d, *J*=9.7 Hz), 6.88 - 6.85 (1H, m), 6.78 (1H, d, *J*=1.5 Hz), 4.59 - 4.56 (2H, m), 4.47 - 4.25 (6H, m), 4.13 - 3.52 (20H, m), 2.47 - 2.42 (6H, m), 2.34 - 2.27 (1H, m), 2.19 - 2.03 (5H, m), 1.63 - 1.52 (2H, m), 1.48 - 1.37 (2H, m), 0.90 (18H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.2, 171.1, 170.7, 170.3, 165.4, 156.6, 150.3, 148.4, 148.3, 138.3, 132.0, 131.8, 131.7, 130.7, 129.5, 129.4, 127.9, 126.9, 122.0, 119.7, 112.6, 71.0, 70.7, 70.5, 70.4, 70.32, 70.28, 70.25, 69.6, 67.9, 59.1, 58.8, 58.5, 57.3,

57.1, 56.7, 43.1, 39.0, 37.3, 36.8, 36.2, 35.4, 26.4, 26.3, 17.9, 17.7, 16.1, 16.0, 13.7. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{61}H_{83}N_9O_{14}S_2$: 1229.55; observed: 1230.66.

(2*S*,4*R*)-1-((*S*)-2-acetamido-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(2-(((*S*)-19-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-20,20-dimethyl-17-oxo-3,6,9,12,15-pentaoxa-18-azahenicosyl)oxy)-4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (CMP112)



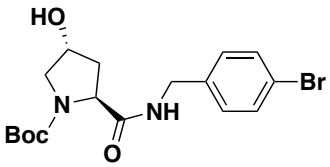
Following general method D, from compound **38** (17 mg, 0.021 mmol, 1 eq.) and trifluoroacetic acid (0.5 ml in 0.5 ml of DCM), the carboxylic acid derivative or 38 was obtained as an oil. Yield: 13 mg, 0.017 mmol (quantitative). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{36}H_{54}N_4O_{12}S$: 766.35; observed: 767.4.

To a solution of carboxylic acid (13 mg, 0.017 mmol, 1 eq.) in 0.5 ml DMF was added HATU (6.49 mg, 0.017 mmol, 1 eq.) and HOAT (2.31 mg, 0.017 mmol, 1 eq.). The resulting mixture was stirred at room temperature for 5 min. Compound **6** (7.90 mg, 0.017 mmol, 1 eq.) was added and the pH of the reaction mixture was adjusted to >9 by addition of DIPEA (3 eq.). The mixture was stirred at room temperature until no presence of the starting materials was detected by LC-MS. The solvent was evaporated under reduced pressure to give the corresponding crude, which was purified by HPLC using a gradient of 5% to 95% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield the final compound as awhite solid. Yield: 6 mg, 0.005 mmol (30%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.61 (2H, s), 7.49 - 7.45 (1H, m), 7.32 - 7.24 (6H, m), 6.90 - 6.87 (1H, m), 6.79 (1H, d, *J*=2.4 Hz), 6.24 (1H, d, *J*=8.9 Hz), 4.61 - 4.29 (10H, m), 4.11 - 3.52 (27H, m), 2.44 (6H, s), 2.30 (1H, t, *J*=13.3 Hz), 2.18 - 2.03 (3H, m), 0.87 (9H, s); **^{13}C**

NMR (126 MHz, $CDCl_3$) δ 171.2, 171.1, 170.7, 170.4, 156.7, 150.3, 148.4, 138.3, 132.2, 130.9, 129.7, 129.4, 128.0, 127.0, 122.0, 112.8, 70.9, 70.6, 70.5, 70.4, 70.3, 70.2, 69.6, 67.9, 59.0, 58.8, 57.7, 57.1, 43.1, 39.0, 37.1, 36.8, 35.6, 35.5, 26.42, 26.38, 23.0, 16.13, 16.06. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{58}H_{82}N_8O_{14}S_2$: 1178.54; observed: 1179.6.

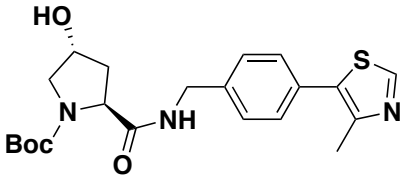
Alternative method for the synthesis of VHL ligand:

***tert*-butyl (2*S*,4*R*)-2-((4-bromobenzyl)carbamoyl)-4-hydroxypyrrolidine-1-carboxylate (12)**



An ice-cooled mixture of *trans*-*N*-(*tert*-Butoxycarbonyl)-4-hydroxy-L-proline (182.46 mg, 0.79 mmol, 1 eq.) and Bromobenzonitrile (147 mg , 0.79 mmol, 1 eq.) in 20 ml DMF was treated with DIPEA (0.4 ml, 2.37 mmol, 3 eq.) and with HATU (380.23 mg, 1 mmol, 1.2 eq.) and the mixture was stirred at r.t. for 30 min. The reaction mixture was quenched with water and extracted with ethyl acetate (×2). The combined organic phases layers were washed with a saturated solution of $NaHCO_3$ (×2), water (×1), brine (×1), dried over $MgSO_4$ and evaporated to dryness. The product was purified by column chromatography using a gradient of 0 to 10% v/v methanol in DCM to yield the final compound. Yield: 265 mg (90%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{17}H_{23}BrN_2O_4$: 398.08; observed: 399.37. Analytical data matched those previously reported [Patent WO 2014/108452].

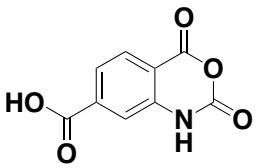
***tert*-butyl ((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)carbamate (3)**



A mixture of compound **13** (265 mg, 0.66 mmol**,** 1 eq.) 4-methylthiazole (0.120 ml, 1.32 mmol, 2 eq.), $Pd(OAc)_2$ (14.81 mg, 0.066 mmol, 0.1 eq.), KOAc (129.5 mg, 1.32 mmol, 2 eq.) in 5 ml *N*-Methyl-2-pyrrolidone was stirred at 120 °C under N_2 for 18h. The reaction was quenched with water and the aqueous phase extracted with ethyl acetate (×4). The combined organic phase was washed with brine (×5), dried over $MgSO_4$ and evaporated to dryness. The crude was purified by column chromatography using a gradient of 0 to 10% methanol in DCM. Yield: 98.64 mg (32%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for: $C_{21}H_{27}N_3O_4S$: 417.17; observed: 418.19. Analytical data matched those previously reported [Patent WO 2014/108452].

6.1.2 Synthesis and characterization of Hetero PROTACs targeting PHD enzymes

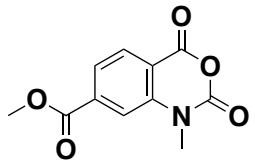
2,4-dioxo-1,4-dihydro-2*H*-benzo[*d*][1,3]oxazine-7-carboxylic acid (54)



A mixture of 2-aminoterephtalic acid (3 g, 166 mmol, 3 eq.) and triphosgene (1.64 g, 5.60 mmol, 1 eq.) in 80 ml of dry THF was heated at 50 °C for 3 h. The solution was concentrate and heptane was added. The precipitate was filtered and washed with methanol. The resulting solid was dried at 50 °C to afford title compound. Yield: 3.389 g (98%). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.65 (1H, *s*),

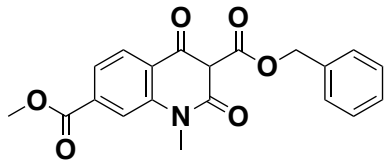
11.88 (1H, s), 8.01 (1H, d, *J*=7.8 Hz), 7.73 - 7.67 (2H, m); **^{13}C NMR** (126 MHz, DMSO-d_6) δ 166.3, 159.9, 147.4, 141.9, 138.3, 129.8, 123.8, 116.5, 114.0. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_9H_5NO_5$: 207.02; observed: 169.2.

methyl 1-methyl-2,4-dioxo-1,4-dihydro-2*H*-benzo[*d*][1,3]oxazine-7-carboxylate (55)



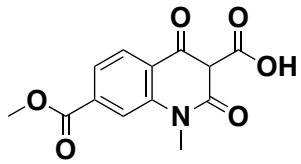
NaH 60% in mineral oil (425 mg, 10.13 mmol, 2.1 eq.) was added to 20 ml of DMF at 0 °C under nitrogen. A solution of **54** (1 g, 4.82 mmol, 1 eq.) in 3 ml of DMF was added and the resulting mixture was stirred at 0 °C until hydrogen gas evolution ceased. A yellow suspension resulted. MeI (1,334 g, 0.6 ml, 9.5 mmol, 1.95 eq.) was added, the resulting mixture was stirred at 50 °C for 30 min. The reaction was quenched at 0 °C using water. The aqueous solution was extracted using DCM (×3). The combined organic phases were washed with a saturated solution of $NaHCO_3$, water and brine, dried over $MgSO_4$, filtered and concentrate to afford the title compound as a yellow solid. Yield: 1.050 mg (92%). **^{1}H NMR** (500 MHz, acetone-d_6) δ 8.06 (1H, d, *J*=7.9 Hz), 7.80 - 7.75 (2H, m), 3.84 (3H, s), 3.52 (3H, s); **^{13}C NMR** (126 MHz, acetone-d_6) δ 167.9, 158.3, 147.6, 142.7, 137.5, 130.2, 123.5, 115.2, 114.3, 22.0, 13.5. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{11}H_9NO_5$: 235.05; observed: 238.2.

3-benzyl 7-methyl 1-methyl-2,4-dioxo-1,2,3,4-tetrahydroquinoline-3,7-dicarboxylate (56)



To a solution of dibenzyl malonate (3.497 g, 3.06 ml, 12.3 mmol, 5 eq.) in 13 ml of DMF was added NaH 60% in mineral oil (118 mg, 2.95 mmol, 1.2 eq.) in portions. The resulting mixture was stirred at r.t for 10 minutes before addition of a solution of **55** (580 mg, 2.46 mmol, 1 eq.) in 13 ml of DMF. The resulting mixture was stirred at 60 °C for 1.5 h. after this time, the reaction was cooled to r.t and 2N HCl was added. The aqueous phase was extracted 3 times with ethyl acetate. The combined organic phases were washed with water, brine, dried over $MgSO_4$, filtered and concentrate to give a yellow solid. The crude was purified by column chromatography (from 0 to 50% EtOAc in Heptane) to afford the title compound as a yellow solid. Yield: 434 mg (48%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.10 - 8.04 (1H, m), 7.85 - 7.82 (1H, m), 7.74 - 7.69 (1H, m), 7.46 (2H, d, *J*=7.6 Hz), 7.32 - 7.21 (3H, m), 5.39 (2H, s), 3.88 (3H, s), 3.57 (3H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.1, 170.8, 165.9, 159.1, 141.0, 135.2, 134.9, 128.6, 128.5, 128.3, 128.2, 127.8, 125.9, 122.1, 117.7, 115.5, 99.2, 67.6, 52.7, 29.4. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{20}H_{17}NO_6$: 367.11; observed: 368.2.

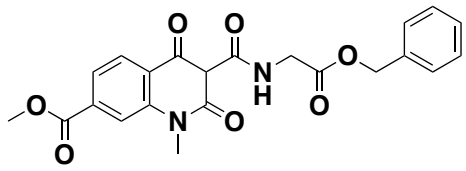
7-(methoxycarbonyl)-1-methyl-2,4-dioxo-1,2,3,4-tetrahydroquinoline-3-carboxylic acid (57)



Compound **56** (530 mg, 1.57 mmol, 1 eq.) was solubilized in 20 ml of ethyl acetate, Pd/C was added and the resulting mixture was place under hydrogen and stirred at r.t for 2.5 h. The reaction mixture was filtered through celite, the celite pad was washed few times using a 1:1 mixture of DCM/EtOAc. The filtrate

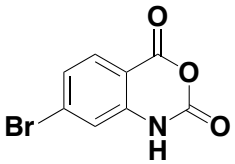
was concentrated in vacuum to give the title compound as an off-white powder. Yield: 320 mg (73%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 15.35 (1H, s), 14.54 (1H, s), 8.26 (1H, d, *J*=7.4 Hz), 8.10 (1H, s), 7.96 (1H, d, *J*=8.3 Hz), 3.94 (3H, s), 3.75 (3H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 173.0, 170.6367, 165.6, 139.5, 135.7, 126.3, 123.9, 118.6, 116.4, 96.4, 53.0, 31.9, 22.7. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{13}H_{11}NO_6$: 277.06; observed: 260.1.

methyl 3-((2-(benzyloxy)-2-oxoethyl)carbamoyl)-1-methyl-2,4-dioxo-1,2,3,4-tetrahydroquinoline-7-carboxylate (58)



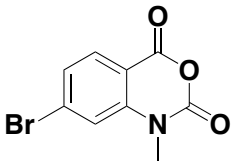
Compound **57** (410 mg, 1.47 mmol, 1 eq.) and glycin benzyl ester HCl (594.86 mg, 2.95 mmol, 1.5 eq.) were dissolved in 10 ml of DMF, DIPEA (1.525 mg, 2 ml, 11.8 mmol, 6 eq.) and pyBOP (3.075 g, 5.91 mmol, 3 eq.) were added and the resulting mixture was stirred at r.t O/N. The reaction mixture was diluted with water and extracted with DCM. The organic phases was washed with water, dried over $MgSO_4$ and evaporated to dryness. The crude was purified by column chromatography (from 0 to 60 % ethyl acetate in heptane) to give the title compound as an white solid. Yield: 471 mg (75%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.16 (1H, d, *J*=9.0 Hz), 7.96 (1H, s), 7.83 - 7.79 (1H, m), 7.34 - 7.25 (5H, m), 5.16 (2H, s), 4.19 (2H, d, *J*=5.5 Hz), 3.91 (3H, s), 3.65 (3H, s); ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.0, 170.9, 168.9, 166.0, 162.4, 139.7, 135.2, 134.5, 128.6, 128.5, 128.4, 125.8, 122.6, 119.0, 115.7, 98.0, 67.2, 67.0, 52.7, 42.2, 41.2, 29.4. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{22}H_{20}N_2O_7$: 424.13; observed: 425.1.

7-bromo-2*H*-benzo[*d*][1,3]oxazine-2,4(*1H*)-dione (61)



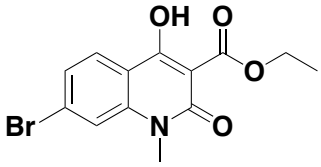
To a solution of 2-amino-4-bromobenzoic acid (629 mg, 2.91 mmol, 3 eq.) in 10 ml of THF was added triphosgene (294 mg, 0.99 mmol, 1 eq.). The resulting mixture was stirred at 40 °C for 1h. The solid was filtered off and washed with diethyl ether and dried under vacuum to afford the title compound as white crystals. Yield: 512 mg (73%). **^{1}H NMR** (400 MHz, DMSO-*d*$_6$) δ 11.91 (1H, s), 7.83 (1H, d, *J*=8.5 Hz), 7.43 (1H, dd, *J*=1.9, 8.5 Hz), 7.35 (1H, d, *J*=1.9 Hz). **^{13}C NMR** (126 MHz, DMSO-*d*$_6$) δ 159.7, 147.3, 142.9, 131.2, 130.7, 126.9, 118.1, 110.2. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_8H_4BrNO_3$: 240.94; observed: 169.2.

7-bromo-1-methyl-2*H*-benzo[*d*][1,3]oxazine-2,4(1*H*)-dione(62)



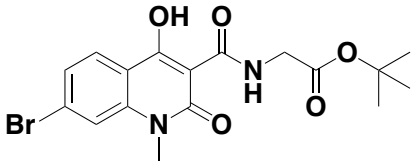
NaH 60% in mineral oil (285 mg, 7.14 mmol, 1.2 eq.) was added to 10 ml of DMF under vigorous stirring. The resulting mixture was cooled to 0 °C and a solution of compound **61** (1440 mg, 5.95 mmol, 1 eq.) in 10 ml of DMF was added. The resulting mixture was stirred at r.t for 1h. MeI (0.407 ml, 6.545 mmol, 1.1 eq.) was added and the mixture was stirred O/N at r.t. `The reaction was quenched with H_2O and the solid collected by filtration, washed with H_2O and diethyl ether, dried in oven at 50 °C to afford the title compound. Yield: 1.145g (76%). **^{1}H NMR** (500 MHz, DMSO-*d*$_6$) δ 7.91 (1H, d, *J*=8.1 Hz), 7.71 (1H, d, *J*=1.7 Hz), 7.53 (1H, dd, *J*=1.7, 8.4 Hz), 3.47 (3H, s); **^{13}C NMR** (126 MHz, DMSO-*d*$_6$) δ 158.9, 148.0, 143.7, 131.5, 131.3, 127.0, 118.1, 111.4, 32.3. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_9H_6BrNO_3$: 254.95; observed: 256.95

ethyl 7-bromo-4-hydroxy-1-methyl-2-oxo-1,2-dihydroquinoline-3-carboxylate (63)



To a solution of diethyl malonate (3.42 ml, 22.55 mmol, 5 eq.) in 15 ml of DMF was added NaH 60 in mineral oil (216 mg, 5.41 mmol, 1.2 eq.) in portions. The resulting mixture was stirred at r.t for 10 min or till a clear yellow solution resulted. A solution of compound **62** (1145 mg, 4.51 mmol, 1 eq.) in 15 ml of DMF was added and the mixture was stirred at 60 ^{o}C for 2.5h. the reaction mixture was cooled to r.t , the precipitate was collected by filtration, dissolved in water and HCl conc. was added to the aqueous solution. The precipitate crystals were collected by filtration and dried under high vacuum. Yield: 1.027 g (70%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 7.95 (1H, d, *J*=8.6 Hz), 7.74 (1H, d, *J*=1.8 Hz), 7.49 (1H, dd, *J*=1.8, 8.6 Hz), 4.35 - 4.30 (2H, m), 3.54 (3H, s), 3.32 (1H, s); **^{13}C NMR** (126 MHz, DMSO-d_6) δ 172.4, 171.2, 159.3, 141.9, 129.4, 127.0, 125.2, 117.1, 113.6, 98.078, 62.4, 29.5, 14.1. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{13}H_{12}BrNO_4$: 324.99; observed: 326.0.

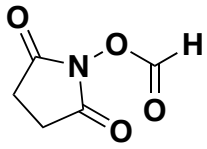
***tert*-butyl (7-bromo-4-hydroxy-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycinate (64)**



A solution of compound **63** (618 mg, 1.89 mmol, 1 eq.) and glycine *tert*-butyl ester hydrochloride (349 mg, 2.08 mmol, 1.1 eq.) in 120 ml of dioxane was

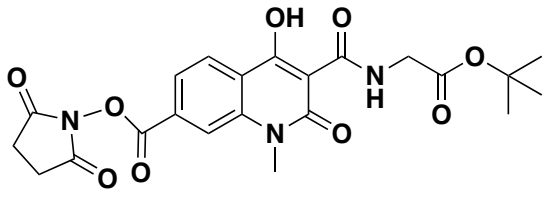
stirred at 120 °C under distillation until complete conversion of the starting material. The crude obtained after evaporation of the organic solvent was taken up with ethyl acetate and filtered. The organic phase was evaporated and the solid residue was purified by column chromatography (from 0 to 60% of ethyl acetate in heptane) to afford the desired product. Yield : 244 mg (31%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 10.56 (1H, s), 7.95 (1H, d, *J*=8.6 Hz), 7.44 (1H, s), 7.32 (1H, d, *J*=8.3 Hz), 4.05 (2H, d, *J*=5.3 Hz), 3.57 (3H, s), 1.44 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.3, 170.8, 168.1, 162.3, 140.7, 128.7, 126.8, 125.6, 117.3, 114.8, 96.9, 82.4, 41.8, 29.7, 29.2, 28.0, 22.6. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{17}H_{19}BrN_2O_5$: 410.05; observed: 354.9.

2,5-dioxopyrrolidin-1-yl formate (65)



Acetic anhydride (7.7 ml, 0.08 mol, 8 eq.) was cooled to 0°C. Formic acid (3.8 ml, 0.1 mol, 10 eq.) was added drop wise and the resulting solution was stirred at r.t for 2h. N-hydroxysuccinimide (1.170 g, 0.01 mol, 1 eq.) was added and the resulting mixture was stirred O/N at r.t.. The organic solvent was removed to dryness and the resulting solid dried under vacuum to afford the desired product. Yield: 1.43 g (98%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.11 (1H, s), 2.81 (2H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 155.0, 25.5, 25.3564, 17.57. Mass not recorded due to fragmentation of the compound.

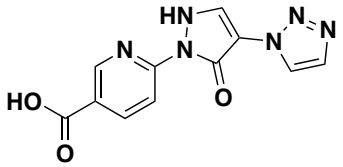
2,5-dioxopyrrolidin-1-yl 3-((2-(*tert*-butoxy)-2-oxoethyl)carbamoyl)-4-hydroxy-1-methyl-2-oxo-1,2-dihydroquinoline-7-carboxylate (67)



Method 1: in a microwave vial were placed 30 ml of a NaOH 2N solution. The vial was capped, evacuated and filled with nitrogen. A deflated balloon was inserted through the septum of the cap. Oxalyl chloride (510 μL) was added through a syringe in to the stirred solution of NaOH. The consequent formation of CO gas flatted the balloon. In an other microwave vial, compound **64** (100 mg, 0.243 mmol, 1 eq.), N-hydroxysuccinimide (39.3 mmg, 0.341 mmol, 1.4 eq.), [1,1′-Bis(diphenylphosphino)ferrocene]dichloropalladium(II), complex with dichloromethane (9.92 mmg, 0.012, 0.05 eq.) were added and the vial was capped and evacuated. Dioxane (2 ml) was added followed by *N,N*-Dicyclohexylmethylamine (70 μL, 0.364 mmol, 1.5 eq.). The system was filled with nitrogen and evacuated two times. The CO balloon was insert through the septum of the cap. The mixture was stirred vigorously at 95 °C O/N. The residue was taken up with methanol and the precipitate was filtered and dried under vacuum to afforded a grey powder. Yield: 32 mg (27%)

Method 2:into a 20 ml vial were placed compound **64**,comppund **65**, xantphos and $Pd(OAc)_2$. The vial was ealed with a septum cap, evacuated and then purged three times with argon. THF dry was added and the resulting mixture was stirred at 60 °C for few minutes before removing the source of argon and adding quickly TEA. The resulting mixture was stirred O/N at 60 °C. The reaction mixture was diluted with methanol and the precipitate was filtered off. The solid grey solid was dried under vacuum to afford the titled compound. Yield: 66.7mg (58%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.35 (1H, d, *J*=7.9 Hz), 8.13 (1H, s), 8.03 - 7.99 (1H, m), 4.14 (2H, d, *J*=5.3 Hz), 3.75 (3H, s), 2.88 (2H, s), 1.51 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 170.7, 170.6, 168.9, 168.0, 162.3, 161.2, 139.7, 129.2, 126.4, 123.1, 120.6, 116.6, 98.6, 82.5, 41.9, 29.4, 28.0, 25.7. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{22}H_{23}N_3O_9$: 473.14; observed: 418.10.

6-(5-oxo-4-(1*H*-1,2,3-triazol-1-yl)-2,5-dihydro-1*H*-pyrazol-1-yl)nicotinic acid (103)



IOX4 (purchased from Cayman Chemical CAS № 1154097-71-8) was in a solution of 50% trifluoroacetic acid in DCM and stirred at r.t until complete conversion of the starting material. Solvent was removed under high vacuum and final product was used for the next step without any further purification. Due to solubility issue it was not possible to analyse this compound by NMR. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{11}H_8N_6O_3$: 272.07; observed: 273.1.

Method A: PEG (3 eq.) was added to a solution of triethyl amine in THF. At 0°C a solution of tosyl chloride in 5 ml of THF was added dropwise over 45 mins. The resulting mixture was stirred at r.t O/N. After this time, solvent was removed under vacuum and the residue dissolved in ethyl acetate. The resulting organic phase was washed with brine, dried over $MgSO_4$ and concentrate in vacuum. The resulting oil was purified by column chromatography (from 20% to 60% acetone in heptane).

Method B: Tosylate was dissolved in ethanol, sodium azide was added and the resulting solution refluxed O/N. Solvent was removed and the residue was dissolved in DCM. The organic phase was washed with water and $NaHCO_3$ solution. The organic phase was dried over $MgSO_4$ and evaporated under vacuum.

Method C: NaH was suspended in dry THF, product from method B was added slowly to the mixture at 0 °C. Ice bath was removed and the resulting mixture was stirred at r.t. for 1h. Bromo acetic acid, dissolved in THF was added to the mixture at 0 °C and stirred O/N at r.t.. Solvent was removed under vacuum and

the resulting residue was acidified to pH 2 with HCl 1N. The aqueous phase was extracted with DCM (×3). Combined organic phases were dried over $MgSO_4$ and dried under vacuum. The crude was purified by column chromatography (from 1% to 7% Methanol in DCM).

Method D: To a solution of acid (1 eq.) in 1 ml DMF was added HATU (1 eq.), HOAT (1 eq.) and the solution was stirred at room temperature for 5 min. Amine (1 eq.) was added and the pH of the reaction mixture was adjusted to >9 by addition of DIPEA (3 eq.). The mixture was stirred at room temperature until no presence of the starting materials was detected by LC-MS. Water was added and the mixture was extracted with ethyl acetate (×3). The combined organic phases were washed with brine (×2), dried over $MgSO_4$ and evaporated under reduced pressure to give the corresponding crude, which was purified by HPLC using a gradient of 20% to 95% v/v acetonitrile in 0.1% aqueous solution of ammonia to yield the final compound.

Method E: azides **83-87** (1 eq.) were solubilized in 2 ml methanol, Pd/C (10 wt %) was added and the resulting mixture was place under hydrogen and stirred at r.t. until complete conversion of the starting material. The reaction mixture was filtered through celite pad and washed few times using methanol. The filtrate was concentrated in vacuum to give an oil.

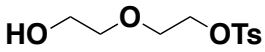
General methods F : To a solution of compound **67** or **103** (1 eq.) in 0.5 ml DMF amine (1 eq.) was added and the pH of the reaction mixture was adjusted to >9 by addition of DIPEA (3 eq.). The mixture was stirred at room temperature until no presence of the starting materials was detected by LC-MS. The solvent was evaporated under reduced pressure to give the corresponding crude, which was purified by HPLC using a gradient of 5% to 95% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield the desired products.

Method G: *tert*-butyl ester compounds were dissolved in a 50% trifluoroacetic acid in DCM and stirred at r.t until complete conversion of the starting material. Solvent was removed under high vacuum and the resulting crude was purified by

HPLC using a gradient of 5% to 95% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield desired products.

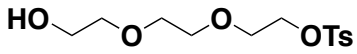
Method H: compounds CMP46 or CMP47 were dissolved in 1 ml of methanol anhydrous and the temperature was cooled down to 0°C. Trimethylsilyldiazomethane was added drop wise till the solution turned yellow. The resulting mixture was stirred at r.t. for 1h. the reaction mixture was quenched with acetic acid glacial and the organic solvent evaporated to dryness. The crude was purified by HPLC using a gradient of 5% to 95% v/v acetonitrile in 0.1% aqueous solution of formic acid to yield final products CMP52-CMP53.

2-(2-hydroxyethoxy)ethyl 4-methylbenzenesulfonate (68)



Following general method A, from diethylene glycol (7.96 g, 7.11 ml, 75 mmol, 3 eq.), and tosyl chloride (4.77 g, 25 mmol, 1 eq.) and TEA (6.93 ml, 5.06 g, 50 mmol, 2 eq.), **68** was obtained as a white solid. Yield: 4.77 g, 18.3 mmol (73%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.80 (2H, d, *J*=6.7 Hz), 7.36 (2H, d, *J*=7.7 Hz), 4.21 - 4.16 (2H, m), 3.71 - 3.53 (6H, m), 2.45 (3H, s); Analytical data matched those previously reported[116].

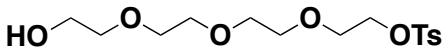
2-(2-(2-hydroxyethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (69)



Following general method A, from compound triethylene glycol (13.2g, 12 ml, 87.9 mmol, 3 eq.), and tosyl chloride (5.59 g, 29.3 mmol, 1 eq.) and TEA (8.12 ml, 5.93 g, 58.6 mmol, 2 eq.) **69** was obtained as a white solid. Yield: 5.77 g, 19 mmol (65%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.82 (2H, d, *J*=8.5 Hz), 7.36 (2H, d, *J*=8.5 Hz),

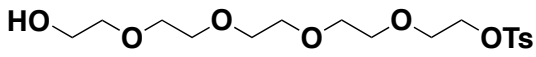
4.20 - 4.18 (2H, m), 3.74 - 3.71 (4H, m), 3.63 (4H, s), 3.61 - 3.58 (2H, m); Analytical data matched those previously reported[116].

2-(2-(2-(2-hydroxyethoxy)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (70)



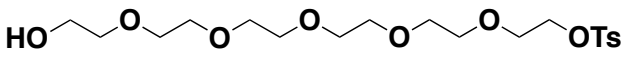
Following general method A, from compound tretraethylene glycol (14,57g, 13 ml, 75 mmol, 3 eq.), and tosyl chloride (4.77g, 25 mmol, 1 eq.) and TEA (6.93 ml, 5.06 g, 58.6 mmol, 2 eq.), **70** was obtained as a white solid. Yield: 7.91 g, 22.4mmol (89%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.82 (2H, d, *J*=7.6 Hz), 7.36 (2H, d, *J*=8.3 Hz), 4.20 - 4.16 (2H, m), 3.75 - 3.60 (14H, m), 2.47 (3H, s); Analytical data matched those previously reported[116].

14-hydroxy-3,6,9,12-tetraoxatetradecyl 4-methylbenzenesulfonate (71)



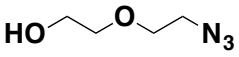
Following general method A, from compound penta glycol (4.05 g, 17 mmol, 3 eq.), and tosyl chloride (1.08 g, 5.67 mmol, 1 eq.) and TEA (1.57 ml, 1.15 g, 11.3 mmol, 2 eq.) in 20 ml of THF, **71** was obtained as a white solid. Yield: 1.6 g, 4 mmol (72%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.81 (2H, d, *J*=7.8 Hz), 7.36 (2H, d, *J*=7.8 Hz), 4.19 - 4.15 (2H, m), 3.74 - 3.59 (18H, m), 2.51 (1H, s), 2.46 (3H, s); Analytical data matched those previously reported[116].

17-hydroxy-3,6,9,12,15-pentaoxaheptadecyl 4-methylbenzenesulfonate (72)



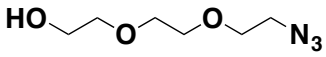
Following general method A, from compound exaethylene glycol (4.80 g, 17 mmol, 3 eq.), and tosyl chloride (1.08 g, 5.67 mmol, 1 eq.) and TEA (1.57 ml, 1.15 g, 11.3 mmol, 2 eq.) in 20 ml of THF, **72** was obtained as a white solid. Yield: 1.20 g, 2.74 mmol (48%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.81 (2H, d, *J*=8.0 Hz), 7.35 (2H, d, *J*=8.0 Hz), 4.19 - 4.15 (2H, m), 3.76 - 3.59 (22H, m), 2.45 (3, s); Analytical data matched those previously reported [116].

2-(2-azidoethoxy)ethan-1-ol (73)



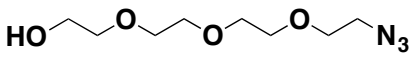
Following general method B, from compound **68** (4.69 g, 18 mmol, 1 eq.), and NaN_3 (5.85 g, 90 mmol, 5 eq.) in 50 ml of ethanol, **73** was obtained as a white solid. Yield: 1.90 g, 14.5 mmol (80%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 3.72 - 3.53 (6H, m), 3.38 - 3.33 (2H, m); Analytical data matched those previously reported[116].

2-(2-(2-azidoethoxy)ethoxy)ethan-1-ol (74)



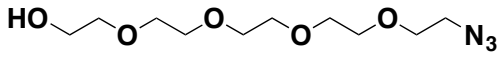
Following general method B, from compound **69** (3.04 g, 10 mmol, 1 eq.), and NaN_3 (3.25 g, 50 mmol, 5 eq.) in 50 ml of ethanol, **74** was obtained as a white solid. Yield: 1.61 g, 9.19 mmol (92%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 3.78 - 3.74 (2H, m), 3.72 - 3.69 (6H, m), 3.65 - 3.62 (2H, m), 3.44 - 3.40 (2H, m); Analytical data matched those previously reported[116].

2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethan-1-ol (75)



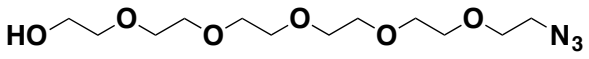
Following general method B, from compound **70** (3.48 g, 10 mmol, 1 eq.), and NaN_3 (3.25 g, 50 mmol, 5 eq.) in 50 ml of ethanol, **75** was obtained as a white solid. Yield: 1.93 g, 8.80 mmol (88%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 3.77 - 3.73 (2H, m), 3.71 - 3.69 (10H, m), 3.65 - 3.62 (2H, m), 3.43 - 3.40 (2H, m). Analytical data matched those previously reported[116].

14-azido-3,6,9,12-tetraoxatetradecan-1-ol (76)



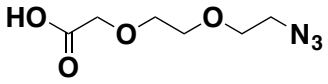
Following general method B, from compound **71** (1.60 g, 4.08 mmol, 1 eq.), and NaN_3 (1.33 g, 20.4 mmol, 5 eq.) in 25 ml of ethanol, **76** was obtained as a white solid. Yield: 1.07 g, 0.004 mmol (99%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 3.75 - 3.59 (18H, m), 3.40 (2H, t, *J*=4.9 Hz); Analytical data matched those previously reported [116].

17-azido-3,6,9,12,15-pentaoxaheptadecan-1-ol (77)



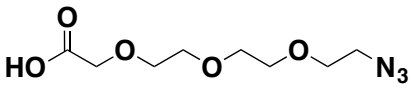
Following general method B, from compound **72** (1.20g, 2.75 mmol, 1 eq.), and NaN_3 (894 mg, 13.7 mmol, 5 eq.) in 25 ml of ethanol, **77** was obtained as a white solid. Yield: 845 mg, 2.74 mmol (99%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 3.76 - 3.60 (22H, m), 3.42 - 3.37 (2H, m); Analytical data matched those previously reported [116].

2-(2-(2-azidoethoxy)ethoxy)acetic acid (78)



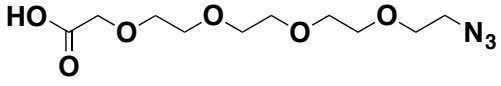
Following general method C, from compound **73** (1.90g, 14.5 mmol, 1 eq.), and NaH (1.16 g, 29 mmol, 2 eq.), bromoacetic acid (2.01 g, 14.5 mmol, 1 eq.) in 50 ml of ethanol, **78** was obtained as a oil. Yield: 1.18 g, 6.24 mmol (43%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 4.21 (2H, s), 3.82 - 3.71 (6H, m), 3.47 - 3.42 (2H, m). Analytical data matched those previously reported [116].

2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)acetic acid (79)



Following general method C, from compound **74** (1.61g, 9.19 mmol, 1 eq.), NaH (765 mg, 18.4 mmol, 2 eq.), bromoacetic acid (1.28g, 9.19 mmol,1eq.) in 20 ml of ethanol, **79** was obtained as a oil. Yield: 857,3 mg, 3.67 mmol (40%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 4.10 (2H, s), 3.72 - 3.60 (10H, m), 3.33 (2H, t, *J*=4.9 Hz); Analytical data matched those previously reported [116].

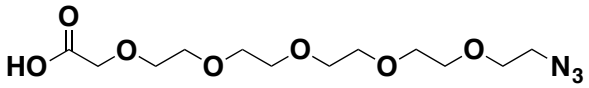
14-azido-3,6,9,12-tetraoxatetradecanoic acid (80)



Following general method C, from compound **75** (1.93g, 8.80 mmol, 1 eq.), NaH (704 mg, 17.6 mmol, 2 eq.), bromoacetic acid (1.22g, 8.80mmol,1eq.) in 20 ml of ethanol, **80** was obtained as a oil. Yield: 732 mg, 2.64 mmol (38%). **^{1}H NMR**

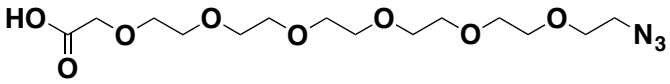
(500 MHz, $CDCl_3$) δ 4.19 (2H, s), 3.80 - 3.69 (10H, m), 3.44 - 3.41 (2H, m); Analytical data matched those previously reported [116].

17-azido-3,6,9,12,15-pentaoxaheptadecanoic acid (81)



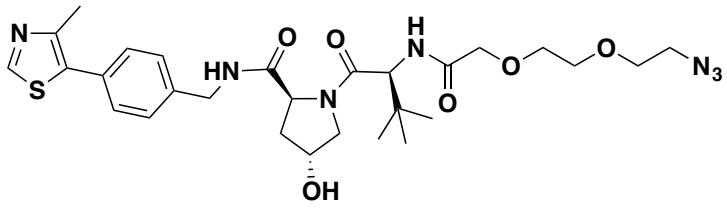
Following general method C, from compound **76** (1.07 g, 4.08 mmol, 1 eq.), NaH (196 mg, 4.90 mmol, 1.2 eq.), bromoacetic acid (690 mg, 4.90 mmol,1.2 eq.) 20 ml of ethanol, **81** was obtained as a white solid. Yield: 800 mg, 2.49 mmol (61%). Analytical data matched those previously reported [116].

20-azido-3,6,9,12,15,18-hexaoxaicosanoic acid (82)



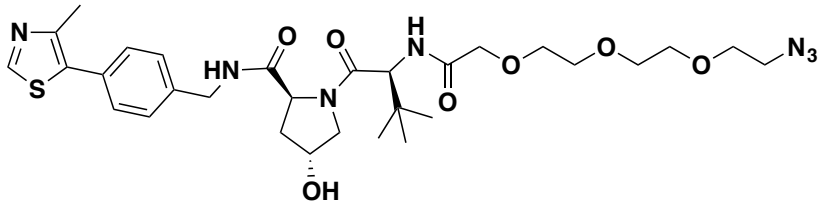
Following general method C, from compound **77** (845 g, 2.75 mmol, 1 eq.), NaH (132 mg, 3.30 mmol, 1.2 eq.), bromoacetic acid (464,7 mg, 3.30mmol,1.2 eq.) 20 ml of ethanol, **82** was obtained as a white solid. Yield: 450 mg, 1.23 mmol (45%). Analytical data matched those previously reported [116].

(2*S*,4*R*)-1-((*S*)-2-(2-(2-(2-azidoethoxy)ethoxy)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (83)



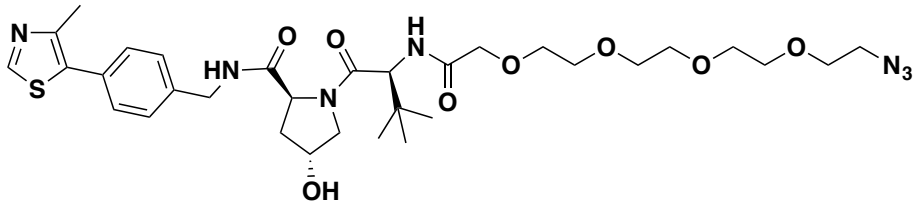
Following general method D, from compound **6** (200 mg, 0.430 mmol, 1 eq.), **78** (81.34 mg, 0.430 mmol, 1 eq.), HATU (163.49 mg, 0.430 mmol, 1 eq.), HOAT (58.52 mg, 0.430 mmol, 1 eq.), DIPEA (160 mg, 0.212 ml, 1.290 mmol, 3 eq.), 3 ml of DMF the title product was obtained as a white solid. Yield: 140 mg, 0.230 mmol (54%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{28}H_{38}N_7O_6S$: 601.27; observed: 602.29. Analytical data matched those previously reported [116].

(2*S*,4*R*)-1-((*S*)-14-azido-2-(*tert*-butyl)-4-oxo-6,9,12-trioxa-3-azatetradecanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (84)



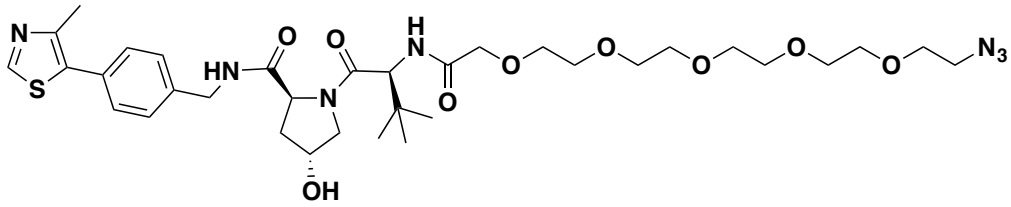
Following general method D, from compound **6** (100 mg, 0.232 mmol, 1 eq.), **79** (77.08 mg, 0.278 mmol, 1.2 eq.), HATU (132.32 mg, 0.348 mmol, 1.5 eq.), DIPEA (119.94 mg, 0.157 ml, 0.928 mmol, 4 eq.), 3 ml of DMF the title product was obtained as a white solid. Yield: 130 mg, 0.201 mmol (72%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{30}H_{43}N_7O_7S$: 645.29; observed: 646.32. Analytical data matched those previously reported [116].

(2*S*,4*R*)-1-((*S*)-17-azido-2-(*tert*-butyl)-4-oxo-6,9,12,15-tetraoxa-3-azaheptadecanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (85)



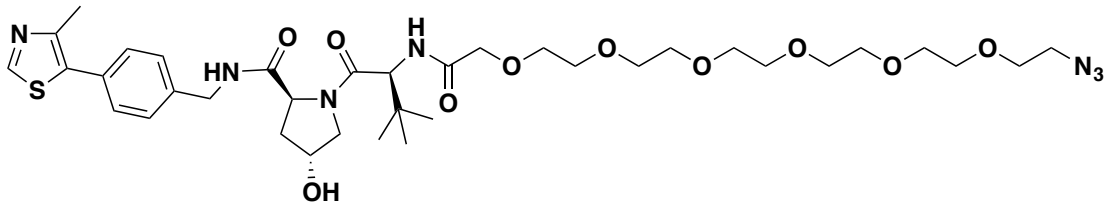
Following general method D, from compound **6** (100 mg, 0.232 mmol, 1 eq.), **80** (89.32 mg, 0.278 mmol, 1.2 eq.), HATU (132.32 mg, 0.348 mmol, 1.5 eq.), DIPEA (119.94 mg, 0.157 ml, 0.928 mmol, 4 eq.), 3 ml of DMF the title product was obtained as a white solid. Yield: 120 mg, 0.173 mmol (62%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{32}H_{47}N_7O_8S$: 689.82; observed: 690.3. Analytical data matched those previously reported [116].

(2*S*,4*R*)-1-((*S*)-20-azido-2-(*tert*-butyl)-4-oxo-6,9,12,15,18-pentaoxa-3-azaicosanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (86)



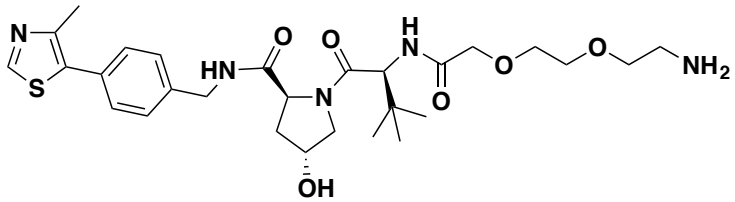
Following general method D, from compound **6** (100 mg, 0.232 mmol, 1 eq.), 81 (101.57 mg, 0.278 mmol, 1.2 eq.), HATU (132.32 mg, 0.348 mmol, 1.5 eq.), DIPEA (119.94 mg, 0.157 ml, 0.928 mmol, 4 eq.), 3 ml of DMF the title product was obtained as a white solid. Yield: 130 mg, 0.177 mmol (63%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{34}H_{51}N_7O_9S$: 733.35; observed: 734.3. Analytical data matched those previously reported [116].

(2*S*,4*R*)-1-((*S*)-23-azido-2-(*tert*-butyl)-4-oxo-6,9,12,15,18,21-hexaoxa-3-azatricosanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (87)



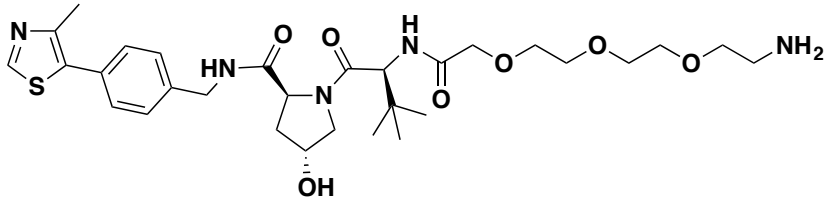
Following general method D, from compound **6** (100 mg, 0.232 mmol, 1 eq.), **82** (113.82 mg, 0.278 mmol, 1.2 eq.), HATU (132.32 mg, 0.348 mmol, 1.5 eq.), DIPEA (119.94 mg, 0.157 ml, 0.928 mmol, 4 eq.), 3 ml of DMF the title product was obtained as a white solid. Yield: 141 mg, 0.181 mmol (65%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{36}H_{55}N_7O_{10}S$: 777.37; observed: 778.4. Analytical data matched those previously reported [116].

(2*S*,4*R*)-1-((*S*)-2-(2-(2-(2-aminoethoxy)ethoxy)acetamido)-3,3-dimethylbutanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (88)



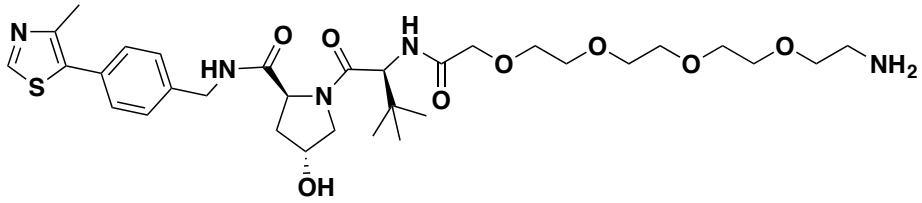
Following general method E, from **83** (32.6 mg, 0.05 mmol, 1 eq.) the title product was obtained as a white solid. Yield: 25 mg, 0.043 mmol (80%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{28}H_{41}N_5O_6S$: 575.28; observed: 576.30. Analytical data matched those previously reported [116].

(2*S*,4*R*)-1-((*S*)-14-amino-2-(*tert*-butyl)-4-oxo-6,9,12-trioxa-3-azatetradecanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide(89)



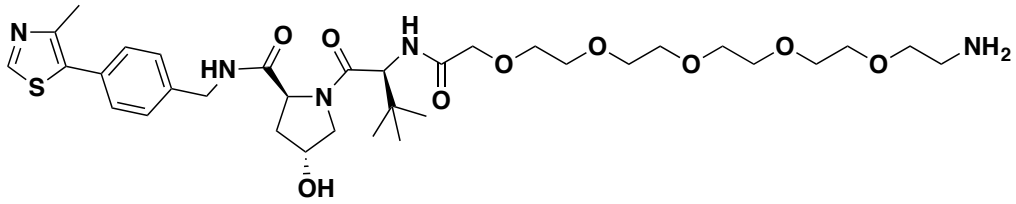
Following general method E, from **84** (34.5 mg, 0.05 mmol, 1 eq.) the title product was obtained as a white solid. Yield: 24.5 mg, 0.039 mmol (79%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{30}H_{45}N_5O_6S$: 619.32; observed: 620.33. Analytical data matched those previously reported[116].

(2*S*,4*R*)-1-((*S*)-17-amino-2-(*tert*-butyl)-4-oxo-6,9,12,15-tetraoxa-3-azaheptadecanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (90)



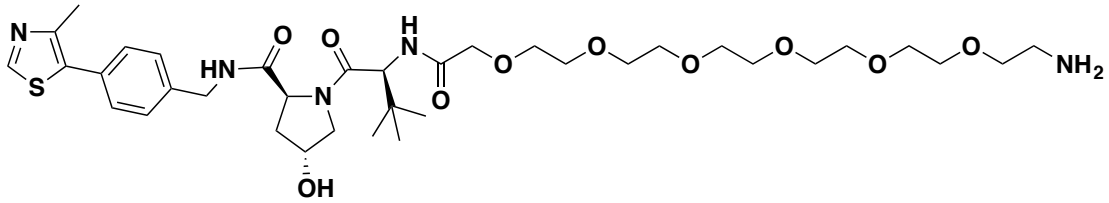
Following general method E, from **85** (34.5 mg, 0.05 mmol, 1 eq.) the title product was obtained as a white solid. Yield: 27.4 mg, 0.041 mmol (78 %). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{32}H_{49}N_5O_8S$: 663.33; observed: 664.4. Analytical data matched those previously reported[116].

(2*S*,4*R*)-1-((*S*)-20-amino-2-(*tert*-butyl)-4-oxo-6,9,12,15,18-pentaoxa-3-azaicosanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (91)



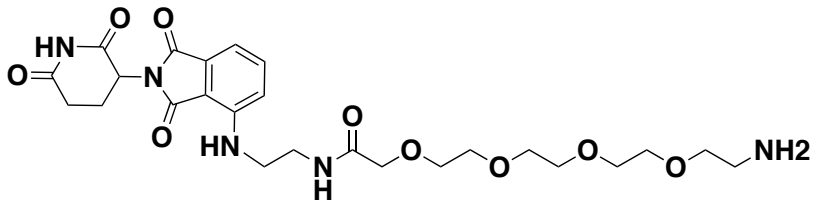
Following general method E, from **86** (37.5 mg, 0.05 mmol, 1 eq.) the title product was obtained as a white solid. Yield: 27.2 mg, 0.038 mmol (75 %). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{34}H_{53}N_5O_9S$: 707.36; observed: 708.38. Analytical data matched those previously reported[116].

(2*S*,4*R*)-1-((*S*)-23-amino-2-(*tert*-butyl)-4-oxo-6,9,12,15,18,21-hexaoxa-3-azatricosanoyl)-4-hydroxy-*N*-(4-(4-methylthiazol-5-yl)benzyl)pyrrolidine-2-carboxamide (92)



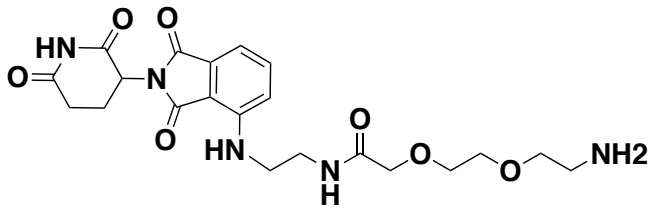
Following general method E, from **87** (39 mg, 0.05 mmol, 1 eq.) the title product was obtained as a white solid. Yield: 25.2 mg, 0.033 mmol (66 %). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{36}H_{57}N_5O_{10}S$: 751.38; observed: 752.40.
Analytical data matched those previously reported[116].

14-amino-*N*-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethyl)-3,6,9,12-tetraoxatetradecanamide



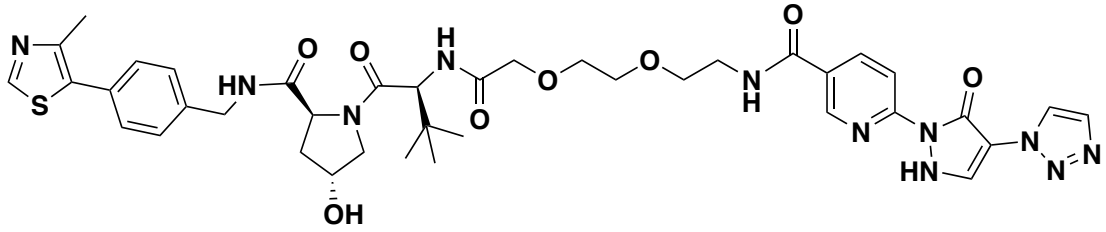
Following general method D, from compound **50** (20 mg, 0.056 mmol, 1 eq.), **80** (15.52 mg, 0.056 mmol, 1 eq.), HATU (21.29 mg, 0.056 mmol, 1 eq.), HOAT (7.61 mg, 0.056 mmol, 1 eq.), DIPEA (23 mg, 0.027 ml, 0.162 mmol, 3 eq.), 0.5 ml of DMF the azide derivative was obtained as a white solid. Yield: 22 mg, 0.038 mmol (68%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{25}H_{33}N_7O_9$: 575.23; observed: 576.3. Following general method E, from azide-compound (22 mg, 0.038 mmol, 1 eq.) the title product was obtained as a white solid. Yield: 21.2 mg, mmol (quantitative). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{25}H_{35}N_5O_9$: 549.24; observed: 550.3.

2-(2-(2-aminoethoxy)ethoxy)-*N*-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethyl)acetamide



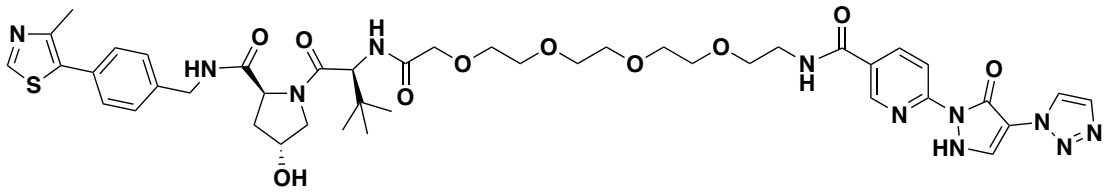
Following general method D, from compound **50** (20 mg, 0.056 mmol, 1 eq.), **78** (10.59 mg, 0.056 mmol, 1 eq.), HATU (21.29 mg, 0.056 mmol, 1 eq.), HOAT (7.61 mg, 0.056 mmol, 1 eq.), DIPEA (23 mg, 0.027 ml, 0.162 mmol, 3 eq.), 0.5 ml of DMF the azide derivative was obtained as a white solid. Yield: 20 mg, 0.041 mmol (73%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{21}H_{25}N_7O_7$: 487.18; observed: 488.2. Following general method E, from azide-compound (20 mg, 0.041 mmol, 1 eq.) the title product was obtained as a white solid. Yield: 21.2 mg, 0.055 mmol (quantitative) **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{21}H_{27}N_5O_7$: 461.19; observed: 462.3 .

***N*-(2-(2-(2-(((*S*)-1-((2*S*,4*R*)-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)-6-(5-oxo-4-(1*H*-1,2,3-triazol-1-yl)-2,5-dihydro-1*H*-pyrazol-1-yl)nicotinamide (CMP70)**



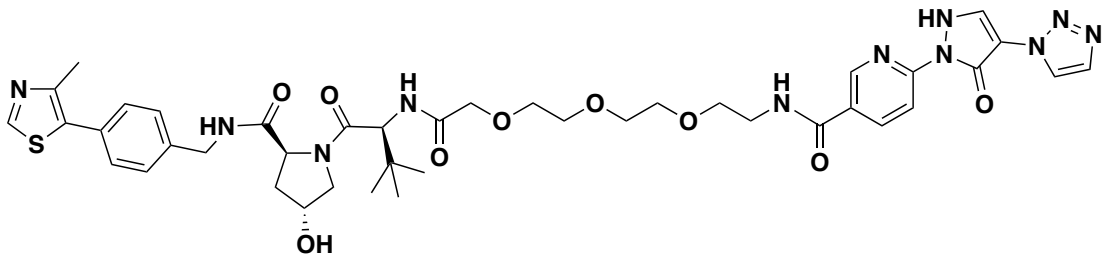
Following general method D, from compound **88** (21 mg, 0.036 mmol, 1 eq.), **103** (10 mg, 0.036 mmol, 1 eq.), HATU (14 mg, 0.036 mmol, 1 eq.), HOAT (5 mg, 0.036 mmol, 1 eq.) DIPEA (13 mg, 0.017 ml, 0.108 mmol, 3 eq.) 0.5 ml of DMF, **CMP70** was obtained as a yiellow solid. Yield: 8.5 mg, 0.010 mmol (28%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.81 (1H, s), 8.59 (1H, s), 8.18 (1H, d, *J*=9.9 Hz), 7.99 (1H, d, *J*=8.6 Hz), 7.89 (1H, s), 7.76 (1H, s), 7.49 (1H, d, *J*=9.3 Hz), 7.23 (2H, d, *J*=8.3 Hz), 7.17 (2H, d, *J*=7.8 Hz), 4.62 (1H, d, *J*=11.3 Hz), 4.54 - 4.50 (1H, m), 4.39 - 4.32 (1H, m), 4.25 - 4.19 (1H, m), 3.93 (1H, s), 3.68 - 3.53 (8H, m), 3.39 - 3.34 (1H, m), 2.83 - 2.78 (1H, m), 2.42 (3H, s), 2.20 - 2.10 (2H, m), 1.13 - 1.06 (6H, m), 0.94 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.0, 170.3, 165.2, 154.0, 150.2, 148.4, 146.4, 138.7, 138.1, 136.1, 133.2, 131.5, 130.8, 129.3, 127.9, 127.0, 123.6, 112.8, 103.1, 71.2, 70.1, 70.0, 59.0, 57.3, 56.9, 53.2, 43.0, 41.5, 39.9, 37.3, 35.9, 26.5, 17.9, 16.0, 11.5. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{39}H_{47}N_{11}O_8S$: 829.33; observed: 830.37.

***N*-((*S*)-16-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-17,17-dimethyl-14-oxo-3,6,9,12-tetraoxa-15-azaoctadecyl)-6-(5-oxo-4-(1*H*-1,2,3-triazol-1-yl)-2,5-dihydro-1*H*-pyrazol-1-yl)nicotinamide (CMP71)**



Following general method D, from compound **90** (24 mg, 0.036 mmol, 1 eq.), **103** (10 mg, 0.036 mmol, 1 eq.), HATU (14 mg, 0.036 mmol, 1 eq.), HOAT (5 mg, 0.036 mmol, 1 eq.) DIPEA (13 mg, 0.017 ml, 0.108 mmol, 3 eq.) 0.5 ml of DMF, **CMP71** was obtained as a white solid. Yield: 8.9 mg, 0.010 mmol (28%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.81 (1H, s), 8.59 (1H, s), 8.18 (1H, d, *J*=9.9 Hz), 7.99 (1H, d, *J*=8.6 Hz), 7.89 (1H, s), 7.76 (1H, s), 7.49 (1H, d, *J*=9.3 Hz), 7.23 (2H, d, *J*=8.3 Hz), 7.17 (2H, d, *J*=7.8 Hz), 4.62 (1H, d, *J*=11.3 Hz), 4.54 - 4.50 (1H, m), 4.39 - 4.32 (1H, m), 4.25 - 4.19 (1H, m), 3.93 (1H, s), 3.68 - 3.53 (16H, m), 3.39 - 3.34 (1H, m), 2.83 - 2.78 (1H, m), 2.42 (3H, s), 2.20 - 2.10 (2H, m), 1.13 - 1.06 (6H, m), 0.94 (9H, s); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 171.2, 171.1, 170.3, 165.2, 153.9, 150.2, 148.4, 146.4, 138.3, 135.9, 133.2, 131.6, 130.7, 129.4, 127.9, 126.7, 123.4, 112.4, 70.8, 70.4, 70.3, 70.2, 70.1, 70.0, 69.8, 58.9, 57.1, 57.0, 53.1, 43.1, 41.4, 40.0, 36.8, 35.4, 26.4, 18.1, 16.1, 11.7. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{43}H_{55}N_{11}O_{10}S$: 917.39; observed: 918.43.

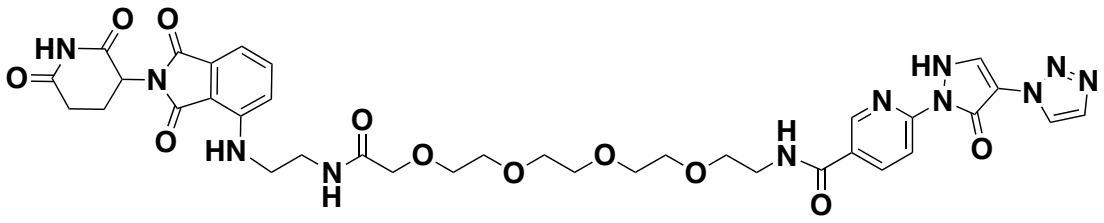
***N*-((*S*)-13-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-14,14-dimethyl-11-oxo-3,6,9-trioxa-12-azapentadecyl)-6-(5-oxo-4-(1*H*-1,2,3-triazol-1-yl)-2,5-dihydro-1*H*-pyrazol-1-yl)nicotinamide(CMP61)**



Following general method D, from compound **89** (19 mg, 0.030 mmol, 1 eq.), **103** (8 mg, 0.030 mmol, 1 eq.), HATU (11 mg, 0.030 mmol, 1 eq.), HOAT (4 mg, 0.036 mmol, 1 eq.) DIPEA (12 mg, 0.015 ml, 0.091 mmol, 3 eq.) 0.5 ml of DMF was obtained as a white solid. Yield: 17 mg, 0.019 mmol (65%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.81 (1H, s), 8.59 (1H, s), 8.18 (1H, d, *J*=9.9 Hz), 7.99 (1H, d, *J*=8.6 Hz), 7.89 (1H, s), 7.76 (1H, s), 7.49 (1H, d, *J*=9.3 Hz), 7.23 (2H, d, *J*=8.3 Hz), 7.17 (2H, d, *J*=7.8 Hz), 4.62 (1H, d, *J*=11.3 Hz), 4.54 - 4.50 (1H, m), 4.39 - 4.32 (1H, m), 4.25 - 4.19 (1H, m), 3.93 (1H, s), 3.68 - 3.53 (12H, m), 3.39 - 3.34 (1H, m), 2.83 - 2.78 (1H, m), 2.42 (3H, s), 2.20 - 2.10 (2H, m), 1.13 - 1.06 (6H, m), 0.94 (9H, s);); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 172.2, 169.6, 169.1, 164.5, 151.9, 148.2, 147.5, 139.9, 138.7, 133.6, 131.6, 130.2, 129.3, 129.1, 128.6, 127.9, 127.4, 124.8, 111.8, 105.3, 70.9, 70.2, 70.1, 70.1, 69.3, 59.2, 57.0, 56.2, 42.1, 40.6, 40.4, 40.2, 36.2, 26.6, 16.4.

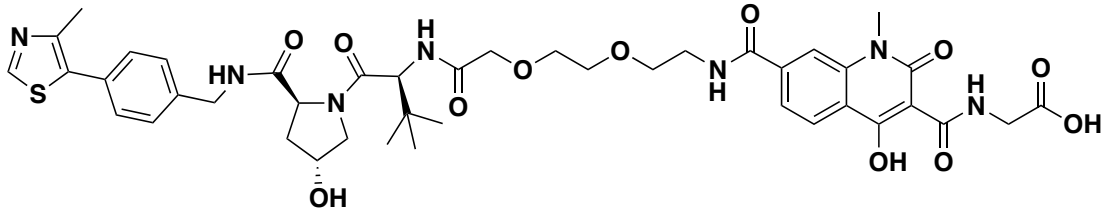
HRMS (ESI) m/z: $[M+H]^+$ calculated for $C_{41}H_{51}N_{11}O_9S$: 873.36; observed: 874.3.

***N*-(1-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)-4-oxo-6,9,12,15-tetraoxa-3-azaheptadecan-17-yl)-6-(5-oxo-4-(1*H*-1,2,3-triazol-1-yl)-2,5-dihydro-1*H*-pyrazol-1-yl)nicotinamide (CMP93)**



Following general method D, from compound **108** (10 mg, 0.018 mmol, 1 eq.), **103** (5 mg, 0.018 mmol, 1 eq.), HATU (6.84 mg, 0.018 mmol, 1 eq.), HOAT (2.44 mg, 0.018 mmol, 1 eq.), DIPEA (0.008 ml, 0.054 mmol, 3 eq.) 0.5 ml of DMF, the title compound was obtained as a yellow solid. Yield: 5 mg, 0.006 mmol (34%). NMR spectra was not recorded due to solubility issues. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{36}H_{41}N_{11}O_{11}$: 803.3; observed: 804.3.

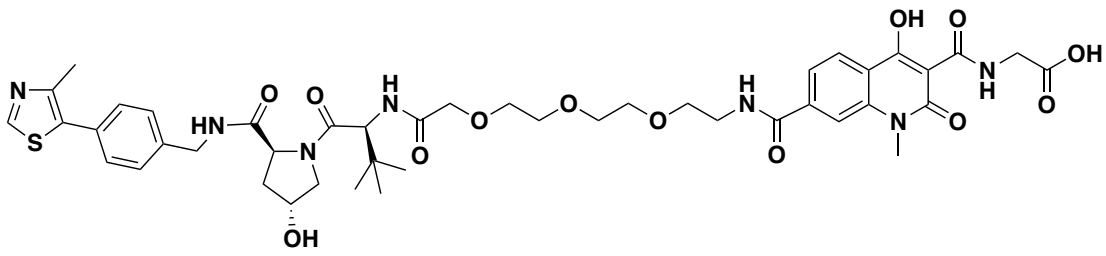
(4-hydroxy-7-((2-(2-(2-(((*S*)-1-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidin-1-yl)-3,3-dimethyl-1-oxobutan-2-yl)amino)-2-oxoethoxy)ethoxy)ethyl)carbamoyl)-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycine (CMP46)



Following general method F, from compound **88** (19 mg, 0.033 mmol, 1 eq.), **67** (16 mg, 0.033 mmol, 1.2 eq.), DIPEA (12.79 mg, 0.017 ml, 0.099 mmol, 3 eq.) 0.5 ml of DMF, **93** was obtained as a white solid. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{46}H_{59}N_7O_{12}S$: 933.39; observed: 934.45.

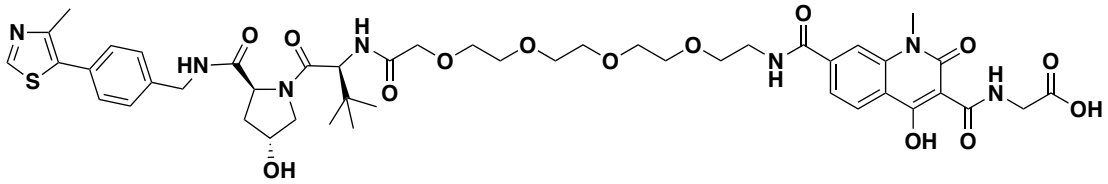
Following general method G, from compound **93** in 0.5 ml of 1:1 solution of TFA in DCM, final product CMP46 was obtained as a white solid. Yield: 25.8 mg, 0.030 mmol (89%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 10.48 - 10.43 (1H, m), 8.90 - 8.81 (2H, m), 8.54 - 8.49 (1H, m), 8.05 (1H, d, *J*=8.4 Hz), 7.89 (1H, s), 7.69 (1H, d, *J*=9.3 Hz), 7.38 (2H, d, *J*=9.8 Hz), 4.49 (1H, d, *J*=9.8 Hz), 4.39 - 3.84 (10H, m), 3.59 - 3.41 (14H, m), 2.42 - 2.34 (6H, m), 2.01 - 1.95 (1H, m), 1.85 - 1.79 (1H, m), 1.14 (1H, s), 0.88 - 0.83 (10H, m). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{42}H_{51}N_7O_{12}S$: 877.33; observed: 878.35.

(4-hydroxy-7-(((*S*)-13-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-14,14-dimethyl-11-oxo-3,6,9-trioxa-12-azapentadecyl)carbamoyl)-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycine(CMP45)



Following general method F, from compound **89** (21 mg, 0.033 mmol, 1 eq.),**67** (16 mg, 0.033 mmol, 1.2 eq.), DIPEA (12.79 mg, 0.017 ml, 0.099 mmol, 3 eq.) 0.5 ml of DMF, **94** was obtained as a white solid. Yield: 20 mg, 0.020 mmol (60%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{48}H_{63}N_7O_{13}S$: 977.42; observed: 978.47. Following general method G, from compound **94** (20 mg, 0.020 mmol, 1 eq.) in 0.5 ml of 1:1 solution of TFA in DCM, final product **CMP45** was obtained as a white solid. Yield: 18.7 mg, 0.019 mmol (98%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 10.48 - 10.43 (1H, m), 8.90 - 8.81 (2H, m), 8.54 - 8.49 (1H, m), 8.05 (1H, d, *J*=8.4 Hz), 7.89 (1H, s), 7.69 (1H, d, *J*=9.3 Hz), 7.38 (2H, d, *J*=9.8 Hz), 4.49 (1H, d, *J*=9.8 Hz), 4.39 - 3.84 (10H, m), 3.59 - 3.41 (18H, m), 2.42 - 2.34 (6H, m), 2.01 - 1.95 (1H, m), 1.85 - 1.79 (1H, m), 1.14 (1H, s), 0.88 - 0.83 (10H, m). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{44}H_{55}N_7O_{13}S$: 922.36; observed: 923.38.

(4-hydroxy-7-(((*S*)-16-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-17,17-dimethyl-14-oxo-3,6,9,12-tetraoxa-15-azaoctadecyl)carbamoyl)-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycine(CMP47)

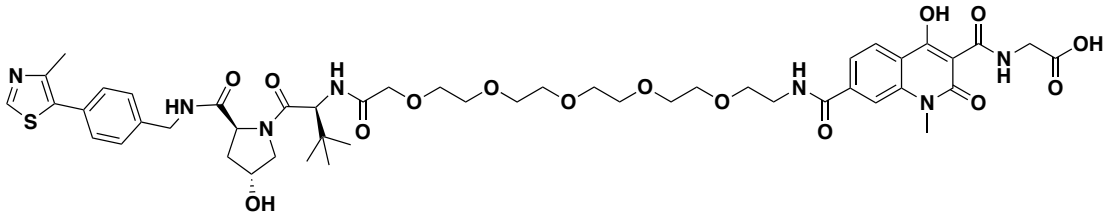


Following general method F, from compound **90** (22 mg, 0.033 mmol, 1 eq.), **67** (16 mg, 0.033 mmol, 1.2 eq.), DIPEA (12.79 mg, 0.017 ml, 0.099 mmol, 3 eq.) 0.5 ml of DMF, **95** was obtained as a white solid. Yield: 23.4 mg, 0.023 mmol (68%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{50}H_{67}N_7O_{14}S$: 1021.45; observed: 1022.47.

Following general method G, from compound **95** (23.4 mg, 0.023 mmol, 1 eq.) in 0.5 ml of 1:1 solution of TFA in DCM, final product **CMP47** was obtained as a white solid. Yield: 21.4 mg, 0.020 mmol (98%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 10.48 - 10.43 (1H, m), 8.90 - 8.81 (2H, m), 8.54 - 8.49 (1H, m), 8.05 (1H, d, *J*=8.4 Hz), 7.89 (1H, s), 7.69 (1H, d, *J*=9.3 Hz), 7.38 (2H, d, *J*=9.8 Hz), 4.49 (1H, d, *J*=9.8

Hz), 4.39 - 3.84 (10H, m), 3.59 - 3.41 (22H, m), 2.42 - 2.34 (6H, m), 2.01 - 1.95 (1H, m), 1.85 - 1.79 (1H, m), 1.14 (1H, s), 0.88 - 0.83 (10H, m). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{46}H_{59}N_7O_{14}S$: 965.38; observed: 966.40.

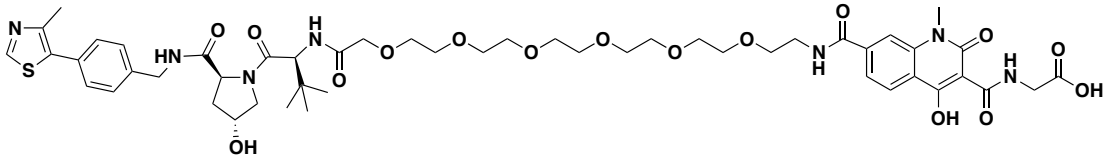
(4-hydroxy-7-(((*S*)-19-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-20,20-dimethyl-17-oxo-3,6,9,12,15-pentaoxa-18-azahenicosyl)carbamoyl)-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycine(CMP48)



Following general method F, from compound **91**(23 mg, 0.033 mmol, 1 eq.), **67** (16 mg, 0.033 mmol, 1.2 eq.), DIPEA (12.79 mg, 0.017 ml, 0.099 mmol, 3 eq.) 0.5 ml of DMF, **96** was obtained as a white solid. Yield: 23.4 mg, 0.023 mmol (68%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{50}H_{67}N_7O_{14}S$: 1021.45; observed: 1067.6.

Following general method G, from compound **96** (23.4 mg, 0.023 mmol, 1 eq.) in 0.5 ml of 1:1 solution of TFA in DCM, final product CMP48 was obtained as a white solid. Yield: 21.6 mg, 0.021 mmol (96%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 10.48 - 10.43 (1H, m), 8.90 - 8.81 (2H, m), 8.54 - 8.49 (1H, m), 8.05 (1H, d, *J*=8.4 Hz), 7.89 (1H, s), 7.69 (1H, d, *J*=9.3 Hz), 7.38 (2H, d, *J*=9.8 Hz), 4.49 (1H, d, *J*=9.8 Hz), 4.39 - 3.84 (10H, m), 3.59 - 3.41 (26H, m), 2.42 - 2.34 (6H, m), 2.01 - 1.95 (1H, m), 1.85 - 1.79 (1H, m), 1.14 (1H, s), 0.88 - 0.83 (10H, m). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{48}H_{59}N_7O_{14}S$: 1009.41; observed: 1010.46.

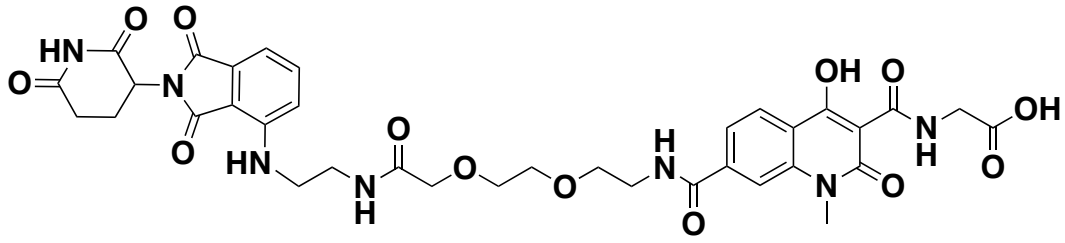
(4-hydroxy-7-(((*S*)-22-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-23,23-dimethyl-20-oxo-3,6,9,12,15,18-hexaoxa-21-azatetracosyl)carbamoyl)-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycine(CMP49)



Following general method F, from compound **92** (25 mg, 0.033 mmol, 1 eq.), **67** (16 mg, 0.033 mmol, 1.2 eq.), DIPEA (12.79 mg, 0.017 ml, 0.099 mmol, 3 eq.) 0.5 ml of DMF, **97** was obtained as a white solid. Yield: 20.5 mg, 0.018 mmol (56%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{54}H_{75}N_7O_{16}S$: 1109.50; observed: 1110.8.

Following general method G, from compound **97** (20.5 mg, 0.018 mmol, 1 eq.) in 0.5 ml of 1:1 solution of TFA in DCM, final product X was obtained as a white solid. Yield: 17 mg, 0.016 mmol (90%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 10.48 - 10.43 (1H, m), 8.90 - 8.81 (2H, m), 8.54 - 8.49 (1H, m), 8.05 (1H, d, *J*=8.4 Hz), 7.89 (1H, s), 7.69 (1H, d, *J*=9.3 Hz), 7.38 (2H, d, *J*=9.8 Hz), 4.49 (1H, d, *J*=9.8 Hz), 4.39 - 3.84 (10H, m), 3.59 - 3.41 (32H, m), 2.42 - 2.34 (6H, m), 2.01 - 1.95 (1H, m), 1.85 - 1.79 (1H, m), 1.14 (1H, s), 0.88 - 0.83 (10H, m). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{50}H_{67}N_7O_{16}S$: 1053.44; observed: 1054.49.

(7-((2-(2-(2-((2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)ethyl)amino)-2-oxoethoxy)ethoxy)ethyl)carbamoyl)-4-hydroxy-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycine (CMP95)

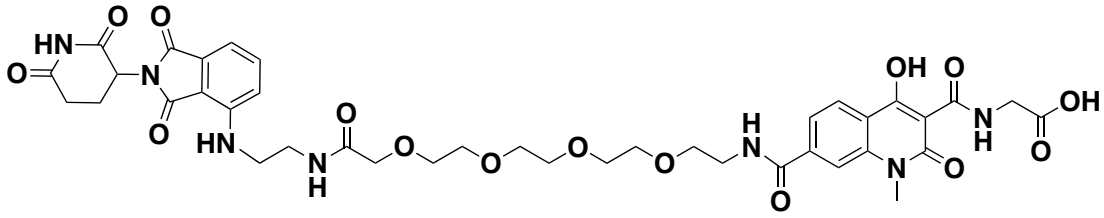


Following general method D, from compound **107** (7.5 mg, 0.016 mmol, 1 eq.), **67** (7.68 mg, 0.016 mmol, 1 eq.), DIPEA (0.009 ml, 6.97 mg, 0.054 mmol, 3 eq.) 0.5 ml of DMF, **110** was obtained as a yellow solid. Yield: 9.8 mg, 0.012 mmol (75%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{43}H_{53}N_7O_{15}$: 819.31; observed:

820.3.

Following general method G, from compound 110 (9.8 mg, 0.012 mmol, 1 eq.) in 0.5 ml of 1:1 solution of TFA in DCM, final product CMP95 was obtained as a white solid. Yield: 9 mg, 0.011 mmol (98%) NMR experiments were not recorded due to solubility issues. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{35}H_{37}N_7O_{13}$: 763.24; observed: 764.3.

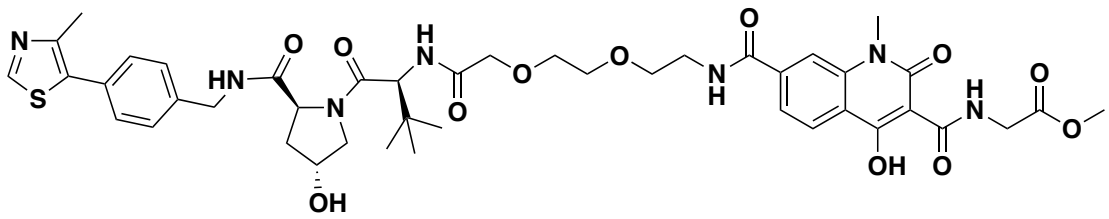
(7-((1-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)-4-oxo-6,9,12,15-tetraoxa-3-azaheptadecan-17-yl)carbamoyl)-4-hydroxy-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycine (CMP96)



Following general method D, from compound **108** (10 mg, 0.018 mmol, 1 eq.), **67** (8.51 mg, 0.018 mmol, 1 eq.), DIPEA (0.009 ml, 6.97 mg, 0.054 mmol, 3 eq.) 0.5 ml of DMF, **109** obtained as a yellow solid. Yield: 8 mg, 0.009 mmol (49%). **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{43}H_{53}N_7O_{15}$: 907.36; observed: 908.4.

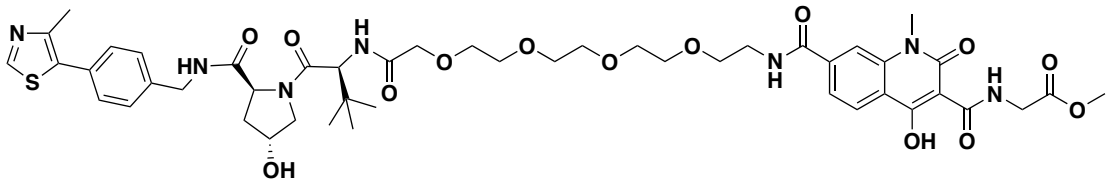
Following general method G, from compound 109 (8 mg, 0.009 mmol, 1 eq.) in 0.5 ml of 1:1 solution of TFA in DCM, final product CMP96 was obtained as a white solid. Yield: 5 mg, 0.005 mmol (65%). NMR experiments were not recorded due to solubility issues. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{39}H_{45}N_7O_{15}$: 851.3; observed: 852.3.

methyl (4-hydroxy-7-((2-(2-(2-(((*S*)-1-((2*S*,4*R*)-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)-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycinate (CMP52)



Following general method H, from compound **CMP46** (8.77 mg, 0.010 mmol, 1 eq.) in 0.5 ml of 1:1 solution of TFA in DCM, final product CMP52 was obtained as a white solid. Yield: 4.5 mg, 0.005 mmol (51%). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.96 (1H, s), 8.87 (1H, s), 8.60 - 8.56 (1H, m), 8.15 (1H, d, *J*=9.1 Hz), 7.93 (1H, s), 7.73 (1H, d, *J*=9.8 Hz), 7.45 (2H, d, *J*=11.3 Hz), 4.58 (1H, d, *J*=10.6 Hz), 4.46 - 4.35 (4H, m), 4.28 - 4.19 (4H, m), 3.99 (3H, s), 3.69 - 3.60 (20H, m), 3.51 (4H, d, *J*=6.8 Hz), 2.43 (4H, s); ^{13}C NMR was not recorded because not enough material was available. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{43}H_{53}N_7O_{12}S$: 891.35; observed: 892.3.

methyl (4-hydroxy-7-(((*S*)-16-((2*S*,4*R*)-4-hydroxy-2-((4-(4-methylthiazol-5-yl)benzyl)carbamoyl)pyrrolidine-1-carbonyl)-17,17-dimethyl-14-oxo-3,6,9,12-tetraoxa-15-azaoctadecyl)carbamoyl)-1-methyl-2-oxo-1,2-dihydroquinoline-3-carbonyl)glycinate (CMP53)



Following general method H, from compound **CMP47** (9.66 mg, 0.010 mmol, 1 eq.) in 0.5 ml of 1:1 solution of TFA in DCM, final product CMP53 was obtained as a white solid. Yield: 5.6 mg, 0.006 mmol (60%). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.96 (1H, s), 8.87 (1H, s), 8.60 - 8.56 (1H, m), 8.15 (1H, d, *J*=9.1 Hz), 7.93 (1H, s), 7.73 (1H, d, *J*=9.8 Hz), 7.45 (2H, d, *J*=11.3 Hz), 4.58 (1H, d, *J*=10.6 Hz), 4.46 - 4.35 (4H, m), 4.28 - 4.19 (4H, m), 3.99 (3H, s), 3.69 - 3.60 (28H, m), 3.51 (4H, d, *J*=6.8 Hz), 2.43 (4H, s); ^{13}C NMR was not recorded because not enough material was

available. **HRMS (ESI)** m/z: $[M+H]^+$ calculated for $C_{47}H_{61}N_7O_{14}S$: 979.40; observed: 980.40.

6.2 Biology

6.2.1 Cell culture

Human cell lines HeLa, U2OS and HEK 293 purchased from ATCC were propagated in DMEM supplemented with 10% fetal bovine serum (FBS), L-glutamine, 100 μg ml^{-1} of penicillin/streptomycin at 37 °C and 5% CO_2. Cells were maintained for no more than 30 passages. All cell lines were routinely tested for mycoplasma contamination using MycoAlert kit from Lonza.

6.2.2 Cell treatment

6.2.2.1 Small interfering RNA

- For siRNA inhibition studies, 3×10^5 cells were seeded into each well of a 6-well plate in order to achieve 70% of confluence on the day of transfection.
- siRNA (SMARTpool: ON-TARGETplus VHL siRNA L-003936-00-0005) was prepared as a 20 μM solution in RNase-free 1× siRNA buffer. Negative control siRNA (siRNA from Life Technologies, cat. # 4390843) was used as negative control.
- On the day of transfection, old medium was replaced with fresh one.
- siRNA solution (5 μL) of both VHL targeting siRNA and negative control were added to 250 μL of Opti-mem in 1.5 mL tube. This solution was prepared in duplicate. The content in each tube was mixed by pipetting.
- Lipofectamine RNAiMax (5 μL) was added to 250 μL of Opti-mem in another 1.5 mL tube. The solution was prepared in duplicate. The content in each tube was mixed by pipetting.
- The solution from step 2 was added to the tube in step 3. The solution was mixed by brief vortex ad incubated at r.t. for 20 min.

- The tubes were centrifuged briefly. The whole volume of transfection mix was added to the 6-well plate. Plate was swirled gently back and forth to mix the content.
- Plates were incubated at 37 ^{o}C and 5% CO_2 for 48 h before harvesting.

6.2.2.2 Single point treatment

For single time point treatment experiments, cells were transferred in 6-well plates with 5 × 10^5 cells per well in 2 ml media in order to achieve 80% confluence the following day.

Stock concentrations of compounds were prepared by solubilizing the powder in 100% v/v DMSO to the final desired stock concentration.

On the day of treatment, all compound samples were prepared as 100-fold concentrated compound solution using DMEM just before treatment.

The experiment samples (20 μL) were added to the 6 well plate containing 2 ml of media. The final DMSO concentration was 0.1% v/v. Cells were incubated at 37 ^{o}C and 5% CO_2 for the desired time before harvesting.

6.2.2.3 Time course experiments

For time dependent treatment, cells were transferred in 6-well plates with 3 × 10^5 cells per well in 2 ml media. Samples were prepared as detailed above or the single time point experiments. Treatment was conducted at given time points prior to harvest.

6.2.2.4 ML4924 and MG132 treatment

For this experiment, cells were transferred in 6-well plates with 5 × 10^5 cells per well in 2 ml media in order to achieve 80% confluence the day after. At t=0, MLN4924 was added into the desired wells at 3 μM final concentration and 0.1% v/v of DMSO. DMSO (0.1% v/v final conc.) was added to the remaining wells in order to match identical conc. of vehicle in all wells. At t=3h, MG 132 was added into the desired wells at 50 μM final conc. and 0.1% v/v of DMSO. DMSO (0.1%

v/v final conc.) was added to the remaining wells in order to achieve the same conc. of vehicle in all the wells. At t=3.5 h, the desired wells were treated with 1 μM of CM11 in 0.1% v/v DMSO final concentration. DMSO (0.1% v/v final conc.) was added to the remaining wells in order to obtain the same conc. of vehicle in all the wells. The total final concentration of DMSO was therefore 0.3% v/v. Plates were incubated for 4 h at 37 °C and 5% CO_2 before harvesting.

6.2.2.5 Competition experiments with VH032

For these experiments, cells were transferred in 6-well plates with 5×10^5 cells per well in 2 ml media in order to achieve 80% confluence the day after. On the day of experiment, cells were treated with VH032 at the final conc. of 150 μM for 30 min prior to treatment with CM11 at 1 μM final concentration for 4 h. Plates were incubated for the desired time at 37 °C and 5% CO_2 before harvesting.

6.2.2.6 Co-treatment with IOX4 and CM11 to investigate upstream effect experiment

For this experiments, cells were transferred in 6-well plates with 5×10^5 cells per well in 2 ml media in order to achieve 80% confluence the day after. On the day of experiment, cells were treated with IOX4 at the final concentration of 50 μM for 30 min prior to treatment with CM11 at 1 μM final concentration for 4 h. Plates were incubated for the desired time at 37 °C and 5% CO_2 before harvesting.

6.2.3 Immunoblotting

Cells were lysed in lysis buffer (20 mM Tris pH 8, 150 mM NaCl, 1% Triton x100) and a protease inhibitor cocktail (Roche) per 10 ml buffer.
For protein extracts, the dishes were placed on ice. The media was aspirated and the tissue layer washed twice with ice-cold phosphate buffer saline (PBS). Lysis buffer (120 μl) was added and the cells detached from the surface with a cell scraper. After removal of the insoluble fraction by centrifugation, the protein

concentration of the supernatant was determined by Pierce™ Coomassie (Bradford) Protein Assay Kit. Protein extracts were fractionated by SDS-PAGE on 4–12% Tris-Acetate NuPage® Novex® (Life Technologies) polyacrylamide gels and transferred to a nitrocellulose membrane using wet transfer. The membrane was then blocked with 5 % w/v Bovine serum albumin (BSA) in Tris-buffered saline (TBS) with 0.1 % w/v Tween-20. For detecting proteins the following primary antibodies in the given concentrations were used: anti-β-Actin (Cell Signaling Technology, 4970S, 13E5) 1:2000, anti-VHL (Cell Signaling Technology, #68547) 1:1000, anti-Hif-1α (BD Biosciences, 610959, clone 54) 1:1000, anti-hydroxy-HIF-1α (Hyp564) (Cell Signaling Techonology; #3434) 1:1000, anti-PHD2 (Bethyl Laboratories; A300-322A) 1:1000, anti-PHD3 (Bethyl Laboratories; A300-327A) 1:1000, anti-CRBN (Proteintech; 11435-1-AP) 1:1000. Following incubation with a horseradish peroxidase-conjugated secondary antibody (Cell Signaling Techonology), the signal was developed using enhanced chemiluminescence (ECL) Western Blotting Detection Kit (Amersham) on Amersham Hyperfilm ECL film (Amersham).

6.3 Biophysical studies

6.3.1 Isothermal titration calorimetry (ITC)

Titrations were performed on an ITC200 micro-calorimeter (GE Healthcare). PROTACs (CM11, CMP98 or CMP99) were diluted from a 100 mM DMSO stock solution to 150 μM in a buffer containing 20 mM Bis-tris propane, 150 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), pH 7.4. The final DMSO concentration was 0.15% v/v.

VBC protein experiments were carried out in a buffer containing 20 mM Bis-tris propane, 150 mM NaCl, 1 mM TCEP, 0.15% v/v DMSO, pH 7.4.

The titrations consisted of 19 injections of 2 μL compounds solution (150 μM, in the syringe) at a rate of 2 s/μL at 120 s time intervals into the VBC protein solution (20 μM, in the cell). An initial injection of compound solution (0.4 μL)

was made and discarded during data analysis. All experiments were performed at 25 °C, whilst stirring the syringe at 600 rpm.
The data were fitted to a single binding site model to obtain the stoichiometry n, the dissociation constant K_d and the enthalpy of binding ΔH using the Microcal LLC ITC200 Origin software provided by the manufacturer.

6.3.2 Size exclusion chromatography (SEC)

SEC experiments were carried out in a ÄKTA pure system (GE Healthcare) at room temperature. The oligomeric state of the VBC complex in solution was analyzed by gel filtration in a buffer containing 20 mM Bis-Tris propane (pH 7), 150 mM NaCl and 1 mM 1,4-dithiothreitol (DTT) using a Superdex 200 10/300GL column (GE Healthcare) calibrated with globular proteins of known molecular weight (GE Healthcare, 28-4038-41/42). VBC protein (50 μM) and ligands CM11 (30 μM), CMP98 (30 μM), CMP99 (30 μM), VH032 (30 μM) or DMSO were mixed and incubated at 20 °C for 20 min before injection. Total volume for each injection was 200 μL. Eluting peaks were monitored using ultraviolet absorbance at wavelength of 280 nm with a flow rate of 0.5 mL/min.

7 References

1. Heilbrunn, L. V. The Dynamic State of Body Constituents. Rudolf Schoenheimer. *Physiological Zoology* **16,** 134–134 (1943).
2. Hutt, D. & Balch, W. E. Cell Biology. The proteome in balance. *Science* **329,** 766–767 (2010).
3. Schrader, E. K., Harstad, K. G. & Matouschek, A. Targeting proteins for degradation. *Nat Chem Biol* **5,** 815–822 (2009).
4. Knop, M., Schiffer, H. H., Rupp, S. & Wolf, D. H. Vacuolar/lysosomal proteolysis: proteases, substrates, mechanisms. *Curr. Opin. Cell Biol.* **5,** 990–996 (1993).
5. Kroemer, G. & Jäättelä, M. Lysosomes and autophagy in cell death control. *Nature Reviews Cancer* **5,** 886–897 (2005).
6. Hochstrasser, M. Ubiquitin-dependent protein degradation. *Annual review of genetics* **30,** 405–439 (1996).
7. Ravid, T. & Hochstrasser, M. Diversity of degradation signals in the ubiquitin-proteasome system. *Nat. Rev. Mol. Cell Biol.* **9,** 679–690 (2008).
8. Goldberg, A. L. Protein degradation and protection against misfolded or damaged proteins. *Nature* **426,** 895–899 (2003).
9. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67,** 425–479 (1998).
10. Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70,** 503–533 (2001).
11. Komander, D. The emerging complexity of protein ubiquitination. *Biochem. Soc. Trans.* **37,** 937–953 (2009).
12. Petroski, M. D. & Deshaies, R. J. Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34. *Cell* **123,** 1107–1120 (2005).
13. Rubin, D. M. & Finley, D. Proteolysis. The proteasome: a protein-degrading organelle? *Curr. Biol.* **5,** 854–858 (1995).
14. Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta* **1695,** 55–72 (2004).
15. Buetow, L. & Huang, D. T. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **17,** 626–642 (2016).
16. Morreale, F. E. & Walden, H. Types of Ubiquitin Ligases. *Cell* **165,** 248–248.e1 (2016).
17. Bernassola, F., Karin, M., Ciechanover, A. & Melino, G. The HECT family of E3 ubiquitin ligases: multiple players in cancer development. *Cancer Cell* **14,** 10–21 (2008).
18. Rotin, D. & Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **10,** 398–409 (2009).
19. Spratt, D. E., Walden, H. & Shaw, G. S. RBR E3 ubiquitin ligases: new structures, new insights, new questions. *Biochem. J.* **458,** 421–437 (2014).
20. Duda, D. M. *et al.* Structure of HHARI, a RING-IBR-RING ubiquitin ligase: autoinhibition of an Ariadne-family E3 and insights into ligation mechanism. *Structure* **21,** 1030–1041 (2013).
21. Deshaies, R. J. & Joazeiro, C. A. P. RING domain E3 ubiquitin ligases. *Annu.*

Rev. Biochem. **78,** 399–434 (2009).
22. Berndsen, C. E. & Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nat. Struct. Mol. Biol.* **21,** 301–307 (2014).
23. Bulatov, E. & Ciulli, A. Targeting Cullin-RING E3 ubiquitin ligases for drug discovery: structure, assembly and small-molecule modulation. *Biochem. J.* **467,** 365–386 (2015).
24. Petroski, M. D. & Deshaies, R. J. Function and regulation of cullin-RING ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **6,** 9–20 (2005).
25. Sarikas, A., Hartmann, T. & Pan, Z.-Q. The cullin protein family. *Genome Biol.* **12,** 220 (2011).
26. Chang, L., Zhang, Z., Yang, J., McLaughlin, S. H. & Barford, D. Atomic structure of the APC/C and its mechanism of protein ubiquitination. *Nature* **522,** 450–454 (2015).
27. Brown, N. G. *et al.* Dual RING E3 Architectures Regulate Multiubiquitination and Ubiquitin Chain Elongation by APC/C. *Cell* **165,** 1440–1453 (2016).
28. Bennett, E. J., Rush, J., Gygi, S. P. & Harper, J. W. Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. *Cell* **143,** 951–965 (2010).
29. Lamsoul, I., Uttenweiler-Joseph, S., Moog-Lutz, C. & Lutz, P. G. Cullin 5-RING E3 ubiquitin ligases, new therapeutic targets? *Biochimie* **122,** 339–347 (2016).
30. Bulatov, E. *et al.* Biophysical studies on interactions and assembly of full-size E3 ubiquitin ligase: suppressor of cytokine signaling 2 (SOCS2)-elongin BC-cullin 5-ring box protein 2 (RBX2). *J. Biol. Chem.* **290,** 4178–4191 (2015).
31. Duda, D. M. *et al.* Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* **134,** 995–1006 (2008).
32. Zimmerman, E. S., Schulman, B. A. & Zheng, N. Structural assembly of cullin-RING ubiquitin ligase complexes. *Curr. Opin. Struct. Biol.* **20,** 714–721 (2010).
33. Schulman, B. A. *et al.* Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* **408,** 381–386 (2000).
34. Fischer, E. S. *et al.* Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* **512,** 49–53 (2014).
35. Kile, B. T. *et al.* The SOCS box: a tale of destruction and degradation. *Trends in Biochemical Sciences* **27,** 235–241 (2002).
36. Xu, L. *et al.* BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* **425,** 316–321 (2003).
37. Duda, D. M. *et al.* Structural regulation of cullin-RING ubiquitin ligase complexes. *Curr. Opin. Struct. Biol.* **21,** 257–264 (2011).
38. Scott, D. C. *et al.* Two Distinct Types of E3 Ligases Work in Unison to Regulate Substrate Ubiquitylation. *Cell* **166,** 1198–1214.e24 (2016).
39. Lydeard, J. R., Schulman, B. A. & Harper, J. W. Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO reports* **14,** 1050–1061 (2013).
40. D'Arcy, P., Wang, X. & Linder, S. Deubiquitinase inhibition as a cancer therapeutic strategy. *Pharmacol. Ther.* **147,** 32–54 (2015).
41. Coux, O., Tanaka, K. & Goldberg, A. L. Structure and functions of the 20S

and 26S proteasomes. *Annu. Rev. Biochem.* **65,** 801–847 (1996).
42. Finley, D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* **78,** 477–513 (2009).
43. Huang, X. & Dixit, V. M. Drugging the undruggables: exploring the ubiquitin system for drug development. *Cell Res.* **26,** 484–498 (2016).
44. Soucy, T. A. *et al.* An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458,** 732–736 (2009).
45. Sarantopoulos, J. *et al.* Phase I Study of the Investigational NEDD8-Activating Enzyme Inhibitor Pevonedistat (TAK-924/MLN4924) in Patients with Advanced Solid Tumors. *Clin. Cancer Res.* **22,** 847–857 (2016).
46. Shah, J. J. *et al.* Phase I Study of the Novel Investigational NEDD8-Activating Enzyme Inhibitor Pevonedistat (MLN4924) in Patients with Relapsed/Refractory Multiple Myeloma or Lymphoma. *Clin. Cancer Res.* **22,** 34–43 (2016).
47. Brownell, J. E. *et al.* Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ. *Mol. Cell* **37,** 102–111 (2010).
48. Ceccarelli, D. F. *et al.* An allosteric inhibitor of the human Cdc34 ubiquitin-conjugating enzyme. *Cell* **145,** 1075–1087 (2011).
49. Huang, H. *et al.* E2 enzyme inhibition by stabilization of a low-affinity interface with ubiquitin. *Nat Chem Biol* **10,** 156–163 (2014).
50. Harper, J. W. & King, R. W. Stuck in the middle: drugging the ubiquitin system at the e2 step. *Cell* **145,** 1007–1009 (2011).
51. Vassilev, L. T. *et al.* In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303,** 844–848 (2004).
52. Tan, X. *et al.* Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446,** 640–645 (2007).
53. Hayashi, K.-I. *et al.* Small-molecule agonists and antagonists of F-box protein-substrate interactions in auxin perception and signaling. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 5632–5637 (2008).
54. Thines, B. *et al.* JAZ repressor proteins are targets of the SCF(COI1) complex during jasmonate signalling. *Nature* **448,** 661–665 (2007).
55. Sheard, L. B. *et al.* Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature* **468,** 400–405 (2010).
56. Ito, T. *et al.* Identification of a primary target of thalidomide teratogenicity. *Science* **327,** 1345–1350 (2010).
57. Zhu, Y. X. *et al.* Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide. *Blood* **118,** 4771–4779 (2011).
58. Lopez-Girona, A. *et al.* Cereblon is a direct protein target for immunomodulatory and antiproliferative activities of lenalidomide and pomalidomide. *Leukemia* **26,** 2445–2445 (2012).
59. Gandhi, A. K. *et al.* Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4(CRBN.). *Br. J. Haematol.* **164,** 811–821 (2014).
60. Chamberlain, P. P. *et al.* Structure of the human Cereblon-DDB1-lenalidomide complex reveals basis for responsiveness to thalidomide

analogs. *Nat. Struct. Mol. Biol.* **21,** 803–809 (2014).
61. Verma, R. *et al.* Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain. *Science* **306,** 117–120 (2004).
62. Berenson, J. R. *et al.* Replacement of bortezomib with carfilzomib for multiple myeloma patients progressing from bortezomib combination therapy. *Leukemia* **28,** 1529–1536 (2014).
63. Lee, D. H. & Goldberg, A. L. Proteasome inhibitors: valuable new tools for cell biologists. *Trends Cell Biol.* **8,** 397–403 (1998).
64. Rentsch, A. *et al.* Synthesis and pharmacology of proteasome inhibitors. *Angew. Chem. Int. Ed. Engl.* **52,** 5450–5488 (2013).
65. Lee, B.-H. *et al.* Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* **467,** 179–184 (2010).
66. Reinhart, P. *et al.* Developing Usp14 inhibitors as disease-modifying therapeutics for protein aggregation diseases. *Alzheimer's & Dementia* **9,** P512–P513 (2013).
67. Semenza, G. L. Oxygen Sensing, Homeostasis, and Disease. *New England Journal of Medicine* **365,** 537–547 (2011).
68. Kaelin, W. G. & Ratcliffe, P. J. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Mol. Cell* **30,** 393–402 (2008).
69. Semenza, G. L. & Wang, G. L. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. *Mol. Cell. Biol.* **12,** 5447–5454 (1992).
70. Wang, G. L., Jiang, B. H., Rue, E. A. & Semenza, G. L. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 5510–5514 (1995).
71. Semenza, G. L. HIF-1: mediator of physiological and pathophysiological responses to hypoxia. *Journal of Applied Physiology* **88,** 1474–1480 (2000).
72. Semenza, G. L. HIF-1 and human disease: one highly involved factor. *Genes Dev.* **14,** 1983–1991 (2000).
73. Wu, D., Potluri, N., Lu, J., Kim, Y. & Rastinejad, F. Structural integration in hypoxia-inducible factors. *Nature* **524,** 303–308 (2015).
74. Ivan, M. *et al.* HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science* **292,** 464–468 (2001).
75. Jaakkola, P. *et al.* Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* **292,** 468–472 (2001).
76. Kamura, T. *et al.* Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase. *Science* **284,** 657–661 (1999).
77. Maxwell, P. H. *et al.* The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399,** 271–275 (1999).
78. Ohh, M. *et al.* Ubiquitination of hypoxia-inducible factor requires direct binding to the beta-domain of the von Hippel-Lindau protein. *Nat. Cell Biol.* **2,** 423–427 (2000).
79. Hon, W.-C. *et al.* Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature* **417,** 975–978 (2002).

80. Min, J.-H. *et al.* Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling. *Science* **296,** 1886–1889 (2002).
81. Masson, N., Willam, C., Maxwell, P. H., Pugh, C. W. & Ratcliffe, P. J. Independent function of two destruction domains in hypoxia-inducible factor-alpha chains activated by prolyl hydroxylation. *EMBO J.* **20,** 5197–5206 (2001).
82. Schofield, C. J. & Ratcliffe, P. J. Oxygen sensing by HIF hydroxylases. *Nat. Rev. Mol. Cell Biol.* **5,** 343–354 (2004).
83. Hewitson, K. S. *et al.* Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural family. *J. Biol. Chem.* **277,** 26351–26355 (2002).
84. Dann, C. E., Bruick, R. K. & Deisenhofer, J. Structure of factor-inhibiting hypoxia-inducible factor 1: An asparaginyl hydroxylase involved in the hypoxic response pathway. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 15351–15356 (2002).
85. Semenza, G. L. Life with oxygen. *Science* **318,** 62–64 (2007).
86. Graeber, T. G. *et al.* Hypoxia-mediated selection of cells with diminished apoptotic potential in solid tumours. *Nature* **379,** 88–91 (1996).
87. Rankin, E. B. & Giaccia, A. J. The role of hypoxia-inducible factors in tumorigenesis. *Cell Death Differ.* **15,** 678–685 (2008).
88. Zhong, H. *et al.* Overexpression of hypoxia-inducible factor 1alpha in common human cancers and their metastases. *Cancer Res.* **59,** 5830–5835 (1999).
89. Semenza, G. L. Targeting HIF-1 for cancer therapy. *Nature Reviews Cancer* **3,** 721–732 (2003).
90. Kaelin, W. G. The von Hippel-Lindau gene, kidney cancer, and oxygen sensing. *J. Am. Soc. Nephrol.* **14,** 2703–2711 (2003).
91. Kaelin, W. G. The von Hippel-Lindau tumour suppressor protein: O2 sensing and cancer. *Nature Reviews Cancer* **8,** 865–873 (2008).
92. Gossage, L., Eisen, T. & Maher, E. R. VHL, the story of a tumour suppressor gene. *Nature Reviews Cancer* **15,** 55–64 (2015).
93. Eltzschig, H. K., Bratton, D. L. & Colgan, S. P. Targeting hypoxia signalling for the treatment of ischaemic and inflammatory diseases. *Nat Rev Drug Discov* **13,** 852–869 (2014).
94. Jain, I. H. *et al.* Hypoxia as a therapy for mitochondrial disease. *Science* **352,** 54–61 (2016).
95. Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450,** 1001–1009 (2007).
96. Pommier, Y. & Marchand, C. Interfacial inhibitors: targeting macromolecular complexes. *Nat Rev Drug Discov* **11,** 25–36 (2011).
97. Thiel, P., Kaiser, M. & Ottmann, C. Small-molecule stabilization of protein-protein interactions: an underestimated concept in drug discovery? *Angew. Chem. Int. Ed. Engl.* **51,** 2012–2018 (2012).
98. Pollock, R. & Clackson, T. Dimerizer-regulated gene expression. *Curr. Opin. Biotechnol.* **13,** 459–467 (2002).
99. Corson, T. W., Aberle, N. & Crews, C. M. Design and Applications of Bifunctional Small Molecules: Why Two Heads Are Better Than One. *ACS Chem. Biol.* **3,** 677–692 (2008).
100. Brown, E. J. *et al.* A mammalian protein targeted by G1-arresting

rapamycin-receptor complex. *Nature* **369,** 756–758 (1994).
101. Liu, J. *et al.* Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66,** 807–815 (1991).
102. Spencer, D. M., Wandless, T. J., Schreiber, S. L. & Crabtree, G. R. Controlling signal transduction with synthetic ligands. *Science* **262,** 1019–1024 (1993).
103. Illendula, A. *et al.* Chemical biology. A small-molecule inhibitor of the aberrant transcription factor CBFβ-SMMHC delays leukemia in mice. *Science* **347,** 779–784 (2015).
104. Tanaka, M. *et al.* Design and characterization of bivalent BET inhibitors. *Nat Chem Biol* **12,** 1089–1096 (2016).
105. Waring, M. J. *et al.* Potent and selective bivalent inhibitors of BET bromodomains. *Nat Chem Biol* **12,** 1097–1104 (2016).
106. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468,** 1067–1073 (2010).
107. Tiscornia, G., Singer, O., Ikawa, M. & Verma, I. M. A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 1844–1848 (2003).
108. Gaj, T., Gersbach, C. A. & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31,** 397–405 (2013).
109. Lai, A. C. & Crews, C. M. Induced protein degradation: an emerging drug discovery paradigm. *Nat Rev Drug Discov* (2016). doi:10.1038/nrd.2016.211
110. Toure, M. & Crews, C. M. Small-Molecule PROTACS: New Approaches to Protein Degradation. *Angew. Chem. Int. Ed. Engl.* **55,** 1966–1973 (2016).
111. Schneekloth, J. S. *et al.* Chemical genetic control of protein levels: selective in vivo targeted degradation. *J. Am. Chem. Soc.* **126,** 3748–3754 (2004).
112. Montrose, K. & Krissansen, G. W. Design of a PROTAC that antagonizes and destroys the cancer-forming X-protein of the hepatitis B virus. *Biochem. Biophys. Res. Commun.* **453,** 735–740 (2014).
113. 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.* **132,** 5820–5826 (2010).
114. Schneekloth, A. R., Pucheault, M., Tae, H. S. & Crews, C. M. Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics. *Bioorg. Med. Chem. Lett.* **18,** 5904–5908 (2008).
115. Galdeano, C. *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.* **57,** 8657–8663 (2014).
116. Zengerle, M., Chan, K.-H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **10,** 1770–1777 (2015).
117. Winter, G. E. *et al.* DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348,** 1376–1381

(2015).
118. Lu, J. *et al.* Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem. Biol.* **22,** 755–763 (2015).
119. Bondeson, D. P. *et al.* Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat Chem Biol* **11,** 611–617 (2015).
120. Deshaies, R. J. Protein degradation: Prime time for PROTACs. *Nat Chem Biol* **11,** 634–635 (2015).
121. Frost, J. *et al.* Potent and selective chemical probe of hypoxic signalling downstream of HIF-α hydroxylation via VHL inhibition. *Nat Commun* **7,** 13312 (2016).
122. Baud, M. G. J. *et al.* Chemical biology. A bump-and-hole approach to engineer controlled selectivity of BET bromodomain chemical probes. *Science* **346,** 638–641 (2014).
123. Raina, K. *et al.* PROTAC-induced BET protein degradation as a therapy for castration-resistant prostate cancer. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 7124–7129 (2016).
124. Buckley, D. L. *et al.* HaloPROTACS: Use of Small Molecule PROTACs to Induce Degradation of HaloTag Fusion Proteins. *ACS Chem. Biol.* **10,** 1831–1837 (2015).
125. Krönke, J. *et al.* Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* **343,** 301–305 (2014).
126. Lu, G. *et al.* The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **343,** 305–309 (2014).
127. Lai, A. C. *et al.* Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL. *Angew. Chem. Int. Ed. Engl.* **55,** 807–810 (2016).
128. Loenarz, C. *et al.* Evidence for a stereoelectronic effect in human oxygen sensing. *Angew. Chem. Int. Ed. Engl.* **48,** 1784–1787 (2009).
129. Stebbins, C. E., Kaelin, W. G. & Pavletich, N. P. Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function. *Science* **284,** 455–461 (1999).
130. Illingworth, C. J. R., Loenarz, C., Schofield, C. J. & Domene, C. Chemical basis for the selectivity of the von Hippel Lindau tumor suppressor pVHL for prolyl-hydroxylated HIF-1alpha. *Biochemistry* **49,** 6936–6944 (2010).
131. Robinson, C. M. & Ohh, M. The multifaceted von Hippel-Lindau tumour suppressor protein. *FEBS Lett.* **588,** 2704–2711 (2014).
132. Okuda, H. *et al.* The von Hippel-Lindau tumor suppressor protein mediates ubiquitination of activated atypical protein kinase C. *J. Biol. Chem.* **276,** 43611–43617 (2001).
133. Minervini, G. *et al.* Isoform-specific interactions of the von Hippel-Lindau tumor suppressor protein. *Sci Rep* **5,** 12605 (2015).
134. Li, M. & Kim, W. Y. Two sides to every story: the HIF-dependent and HIF-independent functions of pVHL. *J. Cell. Mol. Med.* **15,** 187–195 (2011).
135. Ohh, M. *et al.* The von Hippel-Lindau tumor suppressor protein is required for proper assembly of an extracellular fibronectin matrix. *Mol. Cell* **1,** 959–968 (1998).
136. Mehta, R. *et al.* Proteasomal regulation of the hypoxic response modulates aging in C. elegans. *Science* **324,** 1196–1198 (2009).
137. Ramsay, G. & Cantrell, D. Environmental and metabolic sensors that control T cell biology. *Front Immunol* **6,** 99 (2015).

138. Nguyen, H. C., Yang, H., Fribourgh, J. L., Wolfe, L. S. & Xiong, Y. Insights into Cullin-RING E3 ubiquitin ligase recruitment: structure of the VHL-EloBC-Cul2 complex. *Structure* **23,** 441–449 (2015).
139. Zheng, N. *et al.* Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416,** 703–709 (2002).
140. Takagi, Y., Pause, A., Conaway, R. C. & Conaway, J. W. Identification of elongin C sequences required for interaction with the von Hippel-Lindau tumor suppressor protein. *J. Biol. Chem.* **272,** 27444–27449 (1997).
141. Czyzyk-Krzeska, M. F. & Meller, J. von Hippel–Lindau tumor suppressor: not only HIF's executioner. *Trends Mol Med* **10,** 146–149 (2004).
142. Epstein, A. C. *et al.* C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107,** 43–54 (2001).
143. Bruick, R. K. & McKnight, S. L. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science* **294,** 1337–1340 (2001).
144. Tian, Y.-M., Mole, D. R., Ratcliffe, P. J. & Gleadle, J. M. Characterization of different isoforms of the HIF prolyl hydroxylase PHD1 generated by alternative initiation. *Biochem. J.* **397,** 179–186 (2006).
145. Mikhaylova, O. *et al.* The von Hippel-Lindau tumor suppressor protein and Egl-9-Type proline hydroxylases regulate the large subunit of RNA polymerase II in response to oxidative stress. *Mol. Cell. Biol.* **28,** 2701–2717 (2008).
146. Kenneth, N. S. & Rocha, S. Regulation of gene expression by hypoxia. *Biochem. J.* **414,** 19–29 (2008).
147. Chowdhury, R. *et al.* Structural basis for binding of hypoxia-inducible factor to the oxygen-sensing prolyl hydroxylases. *Structure* **17,** 981–989 (2009).
148. Chowdhury, R. *et al.* Structural basis for oxygen degradation domain selectivity of the HIF prolyl hydroxylases. *Nat Commun* **7,** 12673 (2016).
149. Wilkins, S. E., Abboud, M. I., Hancock, R. L. & Schofield, C. J. Targeting Protein-Protein Interactions in the HIF System. *ChemMedChem* **11,** 773–786 (2016).
150. McDonough, M. A. *et al.* Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc. Natl. Acad. Sci. U.S.A.* **103,** 9814–9819 (2006).
151. Kim, S. Y. & Yang, E. G. Recent Advances in Developing Inhibitors for Hypoxia-Inducible Factor Prolyl Hydroxylases and Their Therapeutic Implications. *Molecules* **20,** 20551–20568 (2015).
152. Foxler, D. E. *et al.* The LIMD1 protein bridges an association between the prolyl hydroxylases and VHL to repress HIF-1 activity. *Nat. Cell Biol.* **14,** 201–208 (2012).
153. Rabinowitz, M. H. Inhibition of hypoxia-inducible factor prolyl hydroxylase domain oxygen sensors: tricking the body into mounting orchestrated survival and repair responses. *J. Med. Chem.* **56,** 9369–9402 (2013).
154. Chowdhury, R. *et al.* Selective small molecule probes for the hypoxia inducible factor (HIF) prolyl hydroxylases. *ACS Chem. Biol.* **8,** 1488–1496 (2013).
155. Chan, M. C. *et al.* Potent and Selective Triazole-Based Inhibitors of the

Hypoxia-Inducible Factor Prolyl-Hydroxylases with Activity in the Murine Brain. *PLoS ONE* **10,** e0132004 (2015).
156. Buckley, D. L. *et al.* Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1α interaction. *J. Am. Chem. Soc.* **134,** 4465–4468 (2012).
157. Van Molle, I. *et al.* Dissecting fragment-based lead discovery at the von Hippel-Lindau protein:hypoxia inducible factor 1α protein-protein interface. *Chem. Biol.* **19,** 1300–1312 (2012).
158. Buckley, D. L. *et al.* Small-molecule inhibitors of the interaction between the E3 ligase VHL and HIF1α. *Angew. Chem. Int. Ed. Engl.* **51,** 11463–11467 (2012).
159. Martin, A. R., Ronco, C., Demange, L. & Benhida, R. Hypoxia inducible factor down-regulation, cancer and cancer stem cells (CSCs): ongoing success stories. *Med. Chem. Commun.* (2017). doi:10.1039/C6MD00432F
160. Cho, H. *et al.* On-target efficacy of a HIF-2α antagonist in preclinical kidney cancer models. *Nature* **539,** 107–111 (2016).
161. Chen, W. *et al.* Targeting renal cell carcinoma with a HIF-2 antagonist. *Nature* **539,** 112–117 (2016).
162. Miranda, E. *et al.* A Cyclic Peptide Inhibitor of HIF-1 Heterodimerization That Inhibits Hypoxia Signaling in Cancer Cells. *J. Am. Chem. Soc.* **135,** 10418–10425 (2013).
163. Vassiliou, S. *et al.* A synthetic method for diversification of the P1‘ substituent in phosphinic dipeptides as a tool for exploration of the specificity of the S1’ binding pockets of leucine aminopeptidases. *Bioorg. Med. Chem.* **15,** 3187–3200 (2007).
164. Turnbull, W. B. & Daranas, A. H. On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* **125,** 14859–14866 (2003).
165. de Almeida, A. M., Andersen, T. L., Lindhardt, A. T., de Almeida, M. V. & Skrydstrup, T. General method for the preparation of active esters by palladium-catalyzed alkoxycarbonylation of aryl bromides. *J. Org. Chem.* **80,** 1920–1928 (2015).
166. Barré, A. *et al.* Palladium-Catalyzed Carbonylation of (Hetero)Aryl, Alkenyl and Allyl Halides by Means of N-Hydroxysuccinimidyl Formate as CO Surrogate. *J. Org. Chem.* **80,** 6537–6544 (2015).

Acknowledgements

I would like to thank:

My mentor and supervisor Professor Alessio Ciulli, to welcome me in his group as a visiting PhD student for more than 2 years now; to continuously probe me with ever deeper and more challenging questions; to guide me during this beautiful but complicated experience, and for always finding an answer to my millions doubts.

Professor Dario Alessi, for giving me the possibility to join his group and deepen my knowledge of biology and for the very interesting discussion.

Professor Sonia Rocha, for always finding the time for our scientific discussions, for sharing with me information, and for helping me to understand the complicated oxygen signalling pathway.

Professor Giovanni Romeo, to understand my thirst for knowledge and to let me go my own way, for always being there for me and for offering his support.

Professor Roberto Romeo, for accepting to step in as supervisor when required.

Dr. Andrea Testa, for all the time spent together, for all the help and support in the lab, for the chance to learn the beauty of organic chemistry from such a young but so talented scientist.

Dr. Scott Hughes, for introducing me to the new world of structural biology with patience and consistency, but also for all the fun during the short period of time we spent together.

Hannah Tovell, first of all for teaching me, as "a poor medicinal chemist", how to venture into biology; but also for the patience, for the support, for all the laughs, for the all the hugs at the right time, for the help… to give me hope when I lost it. Nothing would be the same without her.

My lovely friend Elda, for her support, for all the chats… and to forgive me all the time I let her down.

To all the other scientists and colleagues that I met during this long journey, they all left me something that I will never forget.

To my parents and my aunt, even if they don't always appreciate it, I know that you are always there for me, and it is the best that I could desire.

To my syster Miriam, because she never left me alone.

To my lover, to inspire me everyday, for being my best friend, my only love… my everything.

Last but not least, I would like to thank me… for never giving up.