

# Development of novel chemical inducers of dimerization to regulate proteins with high spatial and temporal precision

## Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

Vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

Der Universität Basel

von

Ruben Cal

Aus Le Noirmont, Schweiz

Basel, 2015

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

**edoc.unibas.ch**



Dieses Werk ist unter dem Vertrag "Creative Commons Namensnennung-Keine kommerzielle Nutzung-Keine Bearbeitung 3.0 Schweiz" (CC BY-NC-ND 3.0 CH) lizenziert. Die vollständige

Lizenz kann unter

**[creativecommons.org/licenses/by-nc-nd/3.0/ch/](https://creativecommons.org/licenses/by-nc-nd/3.0/ch/)**

eingesehen werden.

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Matthias P. Wymann und Prof. Dr. Edwin Constable

Basel, den 21.4.2015

Dekan Prof. Dr. Jörg Schibler



## Namensnennung - Nicht-kommerziell - Keine Bearbeitung 3.0 Schweiz

(CC BY-NC-ND 3.0 CH)

**Sie dürfen:** Teilen - den Inhalt kopieren, verbreiten und zugänglich machen

**Unter folgenden Bedingungen:**



**Namensnennung** - Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.



**Keine kommerzielle Nutzung** - Sie dürfen das Material nicht für kommerzielle Zwecke nutzen.



**Keine Bearbeitungen erlaubt** - Sie dürfen diesen Inhalt nicht bearbeiten, abwandeln oder in anderer Weise verändern.

**Wobei gilt:**

- **Verzichtserklärung** - Jede der vorgenannten Bedingungen kann aufgehoben werden, sofern Sie die ausdrückliche Einwilligung des Rechteinhabers dazu erhalten.
- **Public Domain** (gemeinfreie oder nicht-schützbare Inhalte) - Soweit das Werk, der Inhalt oder irgendein Teil davon zur Public Domain der jeweiligen Rechtsordnung gehört, wird dieser Status von der Lizenz in keiner Weise berührt.
- **Sonstige Rechte** - Die Lizenz hat keinerlei Einfluss auf die folgenden Rechte:
  - Die Rechte, die jedermann wegen der Schranken des Urheberrechts oder aufgrund gesetzlicher Erlaubnisse zustehen (in einigen Ländern als grundsätzliche Doktrin des fair use bekannt);
  - Die **Persönlichkeitsrechte** des Urhebers;
  - Rechte anderer Personen, entweder am Lizenzgegenstand selber oder bezüglich seiner Verwendung, zum Beispiel für Werbung oder Privatsphärenschutz.
- **Hinweis** - Bei jeder Nutzung oder Verbreitung müssen Sie anderen alle Lizenzbedingungen mitteilen, die für diesen Inhalt gelten. Am einfachsten ist es, an entsprechender Stelle einen Link auf diese Seite einzubinden.

Quelle: <http://creativecommons.org/licenses/by-nc-nd/3.0/ch/>

Datum: 12.11.2013

## Acknowledgment

First of all, I would like to express my gratitude to Prof. Dr. Matthias P. Wymann for giving me the opportunity to work in his research group and his kind support throughout my Ph.D. studies. Matthias, you allowed me to explore the fascinating field of chemical inducers of dimerization and helped me to overcome the boundary between chemistry and biology, for that I am greatly thankful.

I would like also to thank Prof. Dr. Ed Constable for evaluating this thesis and his advice during the time of my PhD studies.

My special thanks goes to two co-workers, with whom I worked closely during all this time. I thank Dr. Florent Beaufls for his guidance in the Lab, his valuable teaching lessons, Wednesday sandwiches, and for "football total". I thank Mirjam Zimmermann for performing all the biological experiments and for many fruitful scientific discussions.

I thank Kaspar Zimmermann and PD. Dr. Daniel Häussinger for recording and analyzing the NMR spectra for the mechanistic study. I am thankful to Prof. Christian Bochet and his group member Elia Janett for measuring and evaluating the photophysical properties of the described molecules. Additionally, I would like to thank Prof. Dr. Marcel Mayor and his former group member Pascal Hess for technical support.

I would like to express my gratitude to Dr. Florent Beaufls and Dr. Kirsten Gillingwater for proofreading this thesis.

I thank all the past and present members of the Wymann group for an amenable working atmosphere, for their patience with the chemists and the great times we had together. I would like to express my special gratitude to the present and former chemists of the group, namely, Dr. Olivier Jacques, Dr. Eileen Jackson, Dr. Jean-Baptiste Langlois, Dr. Florent Beaufls and Denise Rageot for the nice spirit and fun we had in our lab during all these years.

Last but not least, I want to thank my family and friends. Without their constant support and love, none of the described work would have been possible.



## Table of contents

<b>Acknowledgment</b>	<b>i</b>
<b>Contents</b>	<b>iii</b>
<b>Abbreviations</b>	<b>iv</b>
<b>Abstract</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Targeting proteins with chemical tags and their application . . . . .	3
1.1.1 Self-labelling chemical tags . . . . .	3
1.1.2 Affinity-based chemical protein tags . . . . .	7
1.1.3 Enzyme-mediated chemical tags . . . . .	10
1.2 Labelling two proteins with a bifunctional molecule . . . . .	12
1.2.1 Protein cross-linkers . . . . .	13
1.2.2 Chemical inducers of dimerisation . . . . .	15
1.3 Photo-activatable tools to regulate the function of proteins . . . . .	24
1.3.1 Chemical tag based tools to modify proteins with light . . . . .	24
1.3.2 Optogenetic approaches to regulate protein-protein interactions . . . . .	28
<b>2 Scope of the thesis</b>	<b>30</b>
<b>3 Results and Discussion</b>	<b>31</b>
3.1 Efficient Synthesis of C16-carbamyl Rapalogs via Lewis Acid Decomplexation . .	31
3.2 Synthesis and evaluation of cell permeable and photocleavable HaXS derivatives .	38
3.2.1 Selection of the photolabile groups . . . . .	39
3.2.2 Synthesis of SNAP-tag and HaloTag substrate building blocks . . . . .	42
3.2.3 Synthesis and evaluation of nitrobenzyl derivative containing HaXS molecules	43
3.2.4 Synthesis and evaluation of coumarin derivative containing HaXS . . . . .	51
<b>4 Conclusion and outlook</b>	<b>64</b>
<b>5 Experimental Section</b>	<b>66</b>
5.1 General Information . . . . .	66
5.1.1 Reagents and Solvents . . . . .	66
5.1.2 Sensitive reactions . . . . .	66
5.2 Instrumentation . . . . .	66
5.2.1 Nuclear Magnetic Resonance (NMR) spectroscopy . . . . .	66
5.2.2 Two-dimensional nuclear magnetic resonance (2D NMR) spectroscopy . .	66
5.2.3 Mass Spectroscopy (MS) . . . . .	67
5.2.4 Thin-Layer Chromatography (TLC) and Preparative Thin-Layer Chromatography (PTLC) . . . . .	67
5.2.5 High-Performance Liquid Chromatography (HPLC) . . . . .	67
5.2.6 Ultra Performance Liquid Chromatography (UPLC) . . . . .	67
5.2.7 Column chromatography . . . . .	67
5.2.8 Ultraviolet/Visible (UV/Vis) absorption spectroscopy . . . . .	67
5.2.9 Photo reactors . . . . .	67
5.2.10 Microscopes . . . . .	68

5.3	Experimental procedures . . . . .	68
5.3.1	Synthetic procedures . . . . .	68
5.3.2	NMR experiments to determine the mechanism . . . . .	103
5.3.3	Determination of photophysical properties . . . . .	104
5.3.4	Biological procedures . . . . .	106
<b>6</b>	<b>Appendix</b>	<b>109</b>
6.1	Contributions . . . . .	109
6.2	Supplementary figures . . . . .	109
<b>7</b>	<b>References</b>	<b>112</b>

### Abbreviations

Å	Angström
A	Alanin
ABA	S-(+)-Absciscic acid
ACM	7-Aminocoumarinyl-4-methyl
AcOH	Acetic acid
ACP	Acyl carrier protein
AM	Acetoxymethyl
ATP	Adenosine triphosphate
azo	Azobenzene
BC	O <sup>2</sup> -benzylcytosine
BCys	Biscysteine
BG	O <sup>6</sup> -benzyl guanine
BL	Biotin ligases
Bn	Benzyl
Boc	<i>tert</i> -Butyloxycarbonyl
Boc <sub>2</sub> O	Di- <i>tert</i> -butyl dicarbonate
C	Cysteine
Calcd	Calculated
CALI	Chromophore-assisted light inactivation
CDI	1,1'-Carbodiimidazole
CFP	Cyan fluorescent protein
CIB1	Calcium and integrin-binding protein 1
CID	Chemical inducers of dimerization
CoA	Coenzyme A
COSY	Correlation spectroscopy
CRY	Cryptochrome
CsA	Cyclosporine A
Cys	Cysteine
d	Doublet
DBD	DNA-binding domain
DCC	N,N-Dicyclohexylcarbodiimide
DD	Destabilizing domain
DEX	Dexamethasone
DHFR	Dihydrofolate reductase
DIPEA	N,N-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	N,N'-Dimethylformamide
DMNB	Dimethoxy nitrobenzyl
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
E	Glutamic acid
<i>E. Coli</i>	<i>Escherichia coli</i>
EF-Tu	Elongation factor thermo unstable
e.g.	For example
eGFP	Enhanced green fluorescent protein

---

eYFP	Enhanced yellow fluorescent protein
EGFR	Epidermal growth factor receptor
eq	Equivalent
ESI	Electrospray ionization
Et	Ethyl
<i>et al.</i>	And others
Et <sub>2</sub> O	Diethyl ether
EtOAc	Ethyl acetate
EtOH	Ethanol
EWG	Electron-withdrawing group
F	Phenylalanine
FGE	Formylglycine-generating enzyme
FGly	Formylglycine
FITC	Fluorescein isothiocyanate
FKBP	FK506 binding protein
FRAP	Fluorescence Recovery after Photobleaching
FRB	FKBP12-rapamycin-binding
FRET	Fluorescence resonance energy transfer
FTase	Farnesyltransferase
G	Glycine
GA <sub>3</sub>	Gibberellin
GFP	Green fluorescent protein
GGTase	Geranylgeranyl transferase
GP	General procedure
GPCR	G protein-coupled receptors
GTP	Guanosine-5'-triphosphate
GyrB	Bacterial DNA gyrase
H	Histidine
hAGT	Human O <sup>6</sup> -alkylguanine-DNA alkyltransferase
HCM	7-Hydroxycoumarin-4-yl-methyl
HIFCS	Heat-inactivated fetal calf serum
HIV	Human immunodeficiency virus
HMBC	Heteronuclear Multiple-Bond Correlation
HMQC	Heteronuclear Multiple-Quantum Correlation
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectroscopy
HRP	Horseradish peroxidase
I	Isoleucine
K	Lysine
L	Leucine
LAP	Ligand acceptor protein
LID	Ligand induced degradation

## Abbreviations

---

LOV	Light oxygen voltage
LplA	Lipoic acid ligase
M	Molarity
m	Multiplet
MALDI	Matrix assisted laser desorption ionization
Me	Methyl
MeCN	Acetonitrile
MeNV	Methyl-6-nitroveratryl
MeOH	Methanol
MRI	Magnet Resonance Imaging
MS	Mass spectroscopy
mTORC1	Mammalian/Mechanistic target of rapamycin complex 1
MTX	Methotrexate
N	Asparagine
Nb	Nitrobenzyl
NEt <sub>3</sub>	Triethylamine
NHS	N-Hydroxysuccinimide
NLS	Nuclear localization sequence
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser effect spectroscopy
NPE	2-(4-Nitrophenyl)ethyl
NTA	Nitrilotriacetates
NVOC	2-Nitroveratryloxycarbonyl
P	Proline
<i>p</i> -TsOH	<i>p</i> -Toluenesulfonic acid
PBS	Phosphate-buffered saline
pc	Phenyl carbamate
PCB	Phycocyanobilin
PCP	Peptide carrier protein
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
Ph	Phenyl
pH	Negative logarithm of hydrogen ion concentration
PhyB	Phytochrome B
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
PIF	Phytochrome interacting factor
PKB	Protein kinase B
pNPP	<i>p</i> -Nitrophenyl phosphonate
POI	Protein of interest
PPI	Protein-protein interaction
ppm	Parts per million
PPTase	Phosphopantetheinyl transferases
PTLC	Preparative thin-layer chromatography
PYP	Photoactive yellow protein
Q	Glutamine

q	Quartet
r.t.	Room temperature
R	Arginine
r	Rate
Rac1	Ras-related C3 botulinum toxin substrate 1
Rap	Rapamycin
RFP	Red fluorescent protein
Rheb	Ras homolog enriched in brain
RNA	Ribonucleic acid
ROESY	Rotating frame nuclear Overhauser effect spectroscopy
ROS	Reactive oxygene species
S	Serine
s	Singlet
SEM	Standard error of the mean
SLF	Synthetic ligand of FKBP
SLL	Self-localizing ligands
T	Threonine
t	Triplet
TAD	Transcriptional activation domain
TC	Tetra cysteine
TD	Tetra aspartic acid
TFA	Trifluoroacetic acid
TFP	Teal fluorescent proteins
TGase	Transglutaminase
THF	Tetrahydrofuran
TIAM1	T-cell lymphoma invasion and metastasis-inducing protein 1
TLC	Thin-layer chromatography
TMP	Trimethoprim
TOF	Time of flight
TOSCY	Total Correlated Spectroscopy
tRNA	Transfer ribonucleic acid
TS	Tetra serine
UV/Vis	Ultraviolet/visible
UPLC	Ultra Performance Liquid Chromatography
V	Valine
VVD	Vivid
W	Tryptophane
Y	Tyrosine
YFP	Yellow fluorescent protein

### Abstract

The function of proteins is often regulated by the formation of multi-protein complexes or translocation events. Over the last few decades, several methods have been developed to force proteins to specific locations or induce protein-protein interactions. One of these approaches is based on small bi-functional molecules, so-called chemical inducers of dimerisation (CIDs), which bind to two protein tags. Therefore, two proteins of interest fused to the tag protein, are brought into close proximity in the presence of the CID. Since the pioneering report by Schreiber and Crabtree in 1996, a variety of novel CIDs were designed and used to regulate protein dimerisation and translocation. However, most of these systems exhibit relatively low temporal resolution, determined by the cell permeability of the small molecule. One of the few exceptions is **rapamycin**, a naturally occurring CID, which rapidly induces dimerisation between FKBP12 and FRB. Due to its immunosuppressant properties, several researchers used the "bump and whole" approach to design C16 rapamycin derivatives, which do not bind to wild-type FRB, but to a genetically engineered mutant. Although, various nucleophiles were introduced at the C16 position, no general method to introduce carbamates was published. Here we describe an efficient Lewis acid-mediated method to substitute **rapamycin** at the C16 position with various nucleophiles under mild conditions. Furthermore, we performed NMR experiments to elucidate the exact mechanism of the mentioned method. We observed that treatment of **rapamycin** with the Lewis acid  $\text{BF}_3\text{-Et}_2\text{O}$  in  $\text{CH}_2\text{Cl}_2$ , leads to the heterolytic cleavage of the C16 methoxy group and the formation of a  $\text{BF}_3$ -carbocation complex, which does not react with carbamates. The addition of THF further stabilises the carbocation, which results in a fast reaction with the nucleophile. To demonstrate the utility of this novel method, we introduced a series of representative carbamates at the C16 position of rapamycin. One of the major challenges, when synthesising rapalogs, is that small impurities of rapamycin or rapamycin by-products still inhibit mTOR. Therefore, we extensively purified the diastereomers of C16 phenyl carbamate (**pcRap**) with preparative HPLC. In a proof of concept, (*R*)-**pcRap** successfully induced dimerisation of FKBP and FRB fusion proteins without interfering with the mTOR pathway.

The application of light controllable tools is an alternative approach to regulate the localisation and the activity of proteins with high spatiotemporal precision. Here we report the synthesis of a series of photoactivatable HaXS dimerisers, which contain a photocleavable linker between the HaloTag-reactive chloroalkane ligand and the SNAP-tag-reactive  $\text{O}^6$ -benzylguanine. Chemical modification of the photocleavable linker afforded three powerful photocleavable and cell-permeable CIDs, which could be cleaved independently and orthogonally. Furthermore, in a proof of concept, one of these photocleavable HaXS molecules was successfully used to translocate proteins of interest to the plasma membrane, late endosomes, lysosomes, Golgi, mitochondria and the actin cytoskeleton. Subsequent irradiation of a specific location with either a FRAP-laser or a mercury lamp with a DAPI filter set, readily liberated the anchored proteins. Furthermore, we demonstrated the utility of photocleavable CIDs and explicitly **MeNV-HaXS** in kinetic studies of protein dynamics and the manipulation of subcellular enzyme activities. Therefore, **MeNV-HaXS** was used to anchor a nuclear probe to the Golgi. Subsequent irradiation of cells triggered the release and the relocation of the nuclear probe into the nucleus.

## 1 Introduction

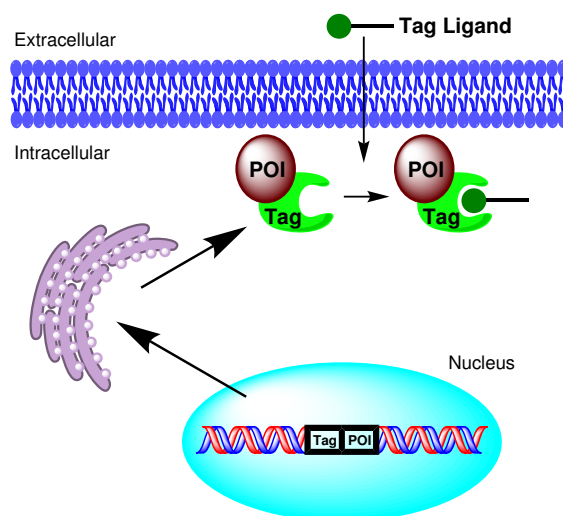
The function of a particular protein in signaling transduction is strongly dependent on its localisation; its concentration at a specific location; and its proximity and orientation towards other biomolecules[1]. Tools to manipulate these parameters *in vivo* are essential to understand the role of specific proteins in signaling pathways[2]. Genetic approaches to regulate protein functions, such as knockout, overexpression and mutations, allows the targeting of single proteins within a whole complex, but suffers from aspects such as low temporal resolution, low flexibility after the genetic change, possible cellular or molecular compensation and lastly, are often restricted to one specific application[3, 4]. In contrast, cell-permeable small molecules (e.g. inhibitors and activators) can regulate the function of proteins rapidly, reversibly and in a dose-dependent manner[4, 5]. However, the discovery of a specific chemical effector for the protein of interest (POI) is challenging. Moreover, most known synthetic small molecules have several targets, which may lead to side effects[4, 5]. Other chemical methods such as site-specific labeling or chemical ligation are more selective, but often require incorporation of specific amino acids into the protein[6, 7]. Additionally, protein modification is performed *in vitro* and microinjection of the protein into cells is needed for *in vivo* studies[6, 8].

In the last three decades, numerous chemical genetic techniques were designed to combine the benefits of these approaches[4, 9]. For example, the orthogonal chemical genetic approach (also referred to as "hole and bump" strategy) was developed to increase the selectivity of small molecules to one specific protein[4, 10]. In this approach, a known ligand for the POI is synthetically modified to the extent that it loses the ability to bind the target protein. In a second step, a cavity is genetically engineered into the protein to accept the modified ligand. By these means, Hwang and Miller altered the ligand specificity of the GTPase thermo unstable elongation factor (Ef-Tu)[11]. Mutation of a single amino acid residue (D138N) in the guanine ring binding motif (NKXD) led to mutants with low affinity for GTP, but high affinity for xanthine triphosphate (XTP). Interestingly, this approach was later extended to other members of the GTPase family, as the NKXD motif is conserved for all known GTPases[4]. Although, the orthogonal chemical genetic approach yields high protein specificity, it requires modification of a known ligand and mutation of the targeted protein.

Incorporation of unnatural amino acids is an alternative approach to regulate the function of proteins. A major breakthrough in this field was the development of an orthogonal aminoacyl-tRNA synthetases/tRNAs pair, to site-specifically introduce unnatural amino acids *in vivo*[12]. In this approach, a genetically engineered tRNA synthase specifically aminoacylates the corresponding foreign tRNA with the unnatural amino acids. The aminoacylated tRNA then recognises a selected codon (nonsense codon or stop codon) and site-specifically incorporates the unnatural amino acid into the protein. Until today, unnatural amino acids, containing spectroscopic probes, bio-orthogonal functionalities for chemo-selective reactions, photosensitive groups and other chemical moieties, were selectively incorporated into proteins in *Arabidopsis thaliana*, bacteria, *Caenorhabditis elegans*, *Drosophila melanogaster*, *E. coli*, mammalian cells and *Saccharomyces cerevisiae*[13]. However, development of new aminoacyl-tRNA synthase/tRNAs pair generally requires large screenings of approximately  $10^9$  synthase variants[14]. Furthermore, the resulting proteins are restricted to one specific application or depend on slow bio-orthogonal reactions, such as the condensation of ketons and aldehydes with amines, click chemistry based reactions, Diels-Alder cyclo-addition, Staudinger-ligation, and several others[6, 13, 15].

A more general approach to regulate the function of specific proteins is based on the genetic fusion of a gene of interest to a sequence encoding a well-studied ligand-binding domain (chemical tag)[16–18]. The function of the resulting chimeric protein (fusion protein) can be regulated by a cell-permeable ligand (figure 1).





**Figure 1:** Principle of chemical tags to target specific protein with small molecules *in vivo*. Cells expressing POI-Tag fusions are treated with a cell-permeable Tag ligand, which functionalises the tag.

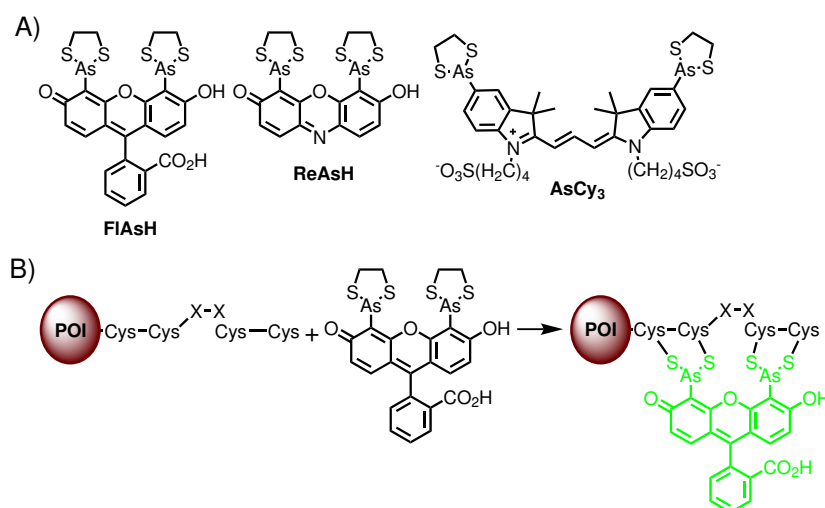
This approach combines the selectivity obtained with genetic encoding and the versatility provided by synthetic chemistry. Additionally, the use of small cell-permeable molecules allows the regulation of the POI in a time- and dose-dependent manner. Moreover, the same or slightly modified fusion proteins can be used for different applications, such as imaging of proteins, regulation of protein stability, control of protein aggregation, translocation of proteins to specific cellular compartments, manipulation of protein-protein interactions, regulation of gene expression, and generation of novel proteins *in vivo* by simply altering the properties of the small molecule[3, 19]. However, the synthetic molecule has to; (I) display fast reaction kinetics; (II) react quantitatively and bio-orthogonal; (III) be cell-permeable and water soluble; and display no cytotoxicity[16–18, 20]. Currently, several ligand-protein domain pairs have been developed, which fulfill all or most of these criteria.

In this introduction, we will mainly focus on general approaches to target fusion proteins and regulate their activity with small molecules. Effector molecules, which target only a specific protein or a protein family, will be excluded. Methods such as incorporation of unnatural amino acids and bio-conjugation will be briefly mentioned. For a more detailed summary of these methods, we refer to several excellent reviews in the field[6, 13, 15, 21].

## 1.1 Targeting proteins with chemical tags and their application

### 1.1.1 Self-labelling chemical tags

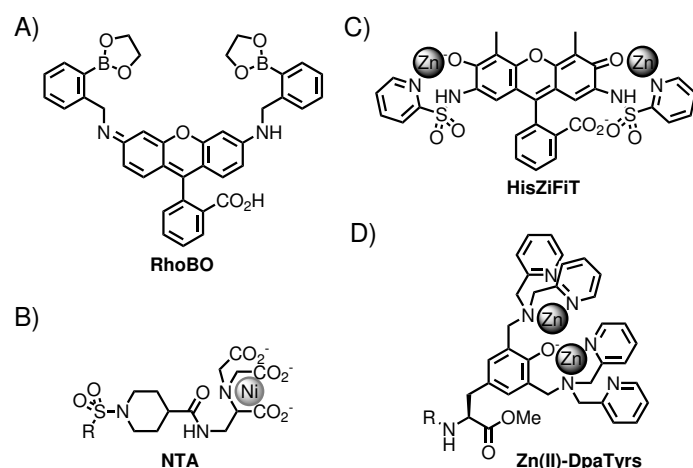
The first applied genetically encoded small polypeptide for specific protein labelling was the tetracysteine (TC) tag, a small polypeptide tag with a CCXXCC core. The cell permeable biarsenic fluorescein derivative (**FlAsH**) recognises and binds to the cysteines of the TC tag, forming a stable complex ( $K_d = 10$  pM) (figure 2, B)[22]. Up to date, several analogues of **FlAsH** have been published, including a red fluorescent version called **ReAsH**, and a bisarsenic cyanine substrate (**AsCy3**) with an orthogonal binding sequence (CCKEAACC) (figure 2, A)[23, 24]. One of the major advantages of **FlAsH** and **ReAsH** is that the fluorescence of the biarsenic compounds dramatically increases after binding to the tetracysteine motif, allowing to differentiate between bound and unbound substrate[22, 23]. **FlAsH** and its variants were used to image HIV-1 proteins in infected cells[25–27], G-protein-coupled receptors[26–29], the formation of amyloid[30, 31], ATP-gated P2X receptor[32], and  $\beta$ -tubulin[33] among many other proteins[34–40]. Despite their wide use, biarsenic compounds suffer from several drawbacks, such as unspecific binding to similar or related sequences and hydrophobic sites in cellular compartments, as well as from cytotoxicity of the biarsenic ligand and labelling conditions[41].



**Figure 2:** TC-tag technology A) Chemical structure of commonly used biarsenic TC ligands; greenfluorescent **FlAsH**, redfluorescent **ReAsH**, and orthogonal **AsCy3**. B) Labelling of TC-tag fused protein with **FlAsH**.

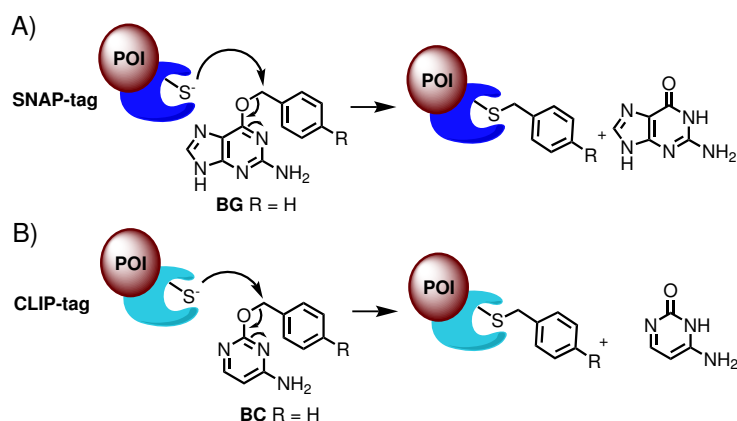
Recently, several groups reported conceptually similar metal binding peptide tags. The group of Schepartz demonstrated the *in vivo* labelling of a model protein fused to a tetraserine (TS) tag with a cell permeable bisboronic acid derivative of rhodamine (**RhoBO**) (scheme 1, A)[42]. Vogel and colleagues designed two  $\text{Ni}^{2+}$  chelating nitrilotriacetates (**NTA**), which recognise a polyhistidine motif (His-tag) (scheme 1, B)[43]. Interestingly, the paramagnetic nature of  $\text{Ni}^{2+}$  quenches nearby fluorophores in a distant dependent manner. The authors used this property to investigate the structure of the ionotropic 5HT3 serotonin receptor. Similarly, the research group of Tsien targeted the His-Tag with the  $\text{Zn}^{2+}$  chelating molecule **HisZiFiT** (scheme 1, C)[44]. In a proof of concept, the authors monitored the translocation of STIM1 from

the ER lumen to plasma membrane and demonstrated that its N-terminus was externalised. Ojida *et al.*, reported the high affinity labelling of a tetra aspartic acid (TD) sequences with multinuclear zinc(II) complexes (**Zn(II)-Dpa-Tyrs**) (scheme 1, D)[45]. In 2003, Imperiali and his group reported the *in vitro* labelling of a polypeptide tag with terbium(III)[46]. The major advantage of these approaches is the small tag size, which minimises the risk of interfering with the protein activity. However, this comes at the expense of lower specificity. To overcome this drawback, several groups designed covalent self-labelling protein tags.



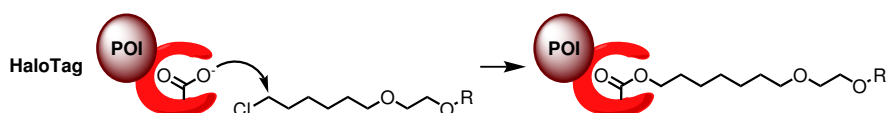
**Scheme 1:** Chemical structure of metal-binding peptide ligands. A) Tetraserine targeting **RhoBO** ligand. B) His-tag targeting **NTA**. C) His-tag targeting ligand **HisZiFiT**. D) Poly aspartic acid targeting **Zn(II)-DpaTyrs** ligand.

In this approach, the POI is fused to an enzyme or a receptor rather than a short peptide, which can subsequently be targeted with a corresponding cell permeable ligand. By these means, the group of Kai Johnsson developed the SNAP-tag, a mutant of the human DNA repair protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase (hAGT)[47, 48]. In its natural function, hAGT irreversibly transfers alkyl groups of O<sup>6</sup>-alkylated guanine to a resident cysteine. Mutations of residues close to the active-site led to versions with increased selectivity and activity for O<sup>6</sup>-benzyl guanine (**BG**) derivatives (figure 3, A)[48–50]. Later, the same group developed an hAGT mutant with orthogonal selectivity. The so-called CLIP-tag has a high affinity for O<sup>2</sup>-benzyl cytosine (**BC**) and very low reactivity towards **BG** (figure 3, B)[47]. Fluorescent SNAP-tag substrates were successfully used to study the localisation of centromere bound CENP-A, the trafficking of Na,K-ATPase and Smoothed, as well as the stoichiometry of class A and C GPCRs[51–54]. More recently, fluorescent **BG** derivatives were used in combination with super-resolution imaging techniques, to image cytoskeletal and membrane bound proteins[55], clathrine coated pits[56], histone proteins[57], and clustering of HIV-1 Gag protein[58].



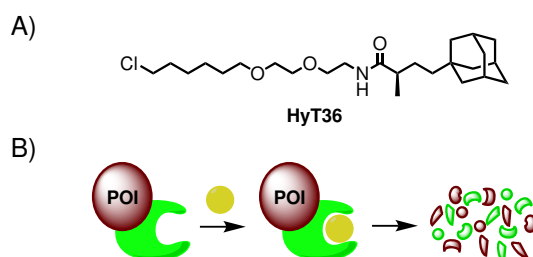
**Figure 3:** Schematic representation of covalent protein labelling based on hAGT. A) SNAP-tag; B) CLIP-tag technology.

A third covalent self-labelling chemical tag, called HaloTag, was introduced in 2008 by Promega[59]. The HaloTag is a modified bacterial haloalkane dehalogenase. Haloalkane dehalogenases naturally catalyse the two-step reaction of an alkylhalide to an alkyl alcohol. In the initial step, the alkyl group of the alkylhalide is transferred to the side chain of an aspartate located at the active-site. In the second step, the previously formed ester is hydrolysed by a water molecule, which is activated by a histidine. Mutation of this histidine in the HaloTag prevents the hydrolysis and the regeneration of aspartate. Thus, the alkyl group remains covalently attached to the enzyme (figure 4). With 33 kDa, the HaloTag is a rather large tag, which may limit its application. However, the lack of haloalkane dehalogenase in eukaryotes results in low background labelling. Recently, HaloTag technology was applied to monitor the distribution of



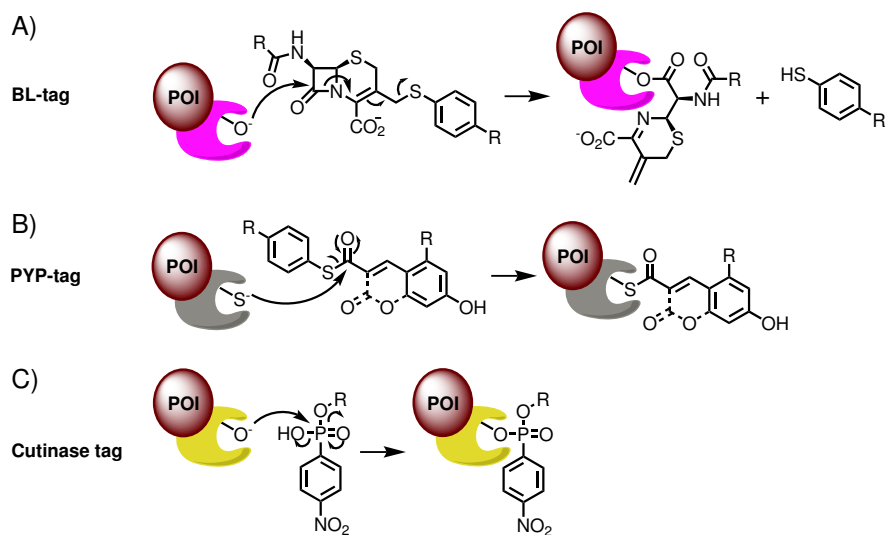
**Figure 4:** Schematic representation of covalent protein labelling with the Halo-tag technology.

MMP-2 between the mitochondria and the mitochondria-associated membrane in the heart[60], the localisation of outer mitochondrial membrane proteins and subunits of the respiratory chain complex II[61], and the Protocadherin-24 mediated translocation of  $\beta$ -catenin in a colon cancer cell line[62]. Mimicking the partially denatured state of proteins, the Crews group designed a hydrophobic HaloTag substrate (**HyT36**) to control proteolysis of proteins (figure 5, A). *In vivo* labelling of POI-HaloTag fusion with this hydrophobic small molecule readily induced proteolysis of the tagged protein (figure 5, B)[63]. The same group used this approach to regulate the protein levels of HRas1<sup>G12V</sup> in mice[64], Smad5 in zebrafish[64] and to study the unfolded protein response in the endoplasmic reticulum[65]. Recently, Chidely *et al.*, developed a bifunctional chimera tag, based on the HaloTag and the SNAP-tag, termed Covalin[66]. In a series of proof-of-principle experiments, the authors conjugated antibodies to fluorophores, enzymes and to magnetic beads.



**Figure 5:** HyT36 induced protein degradation A) Chemical structure of **HyT36**. B) Schematic representation of **HyT36** induced protein degradation.

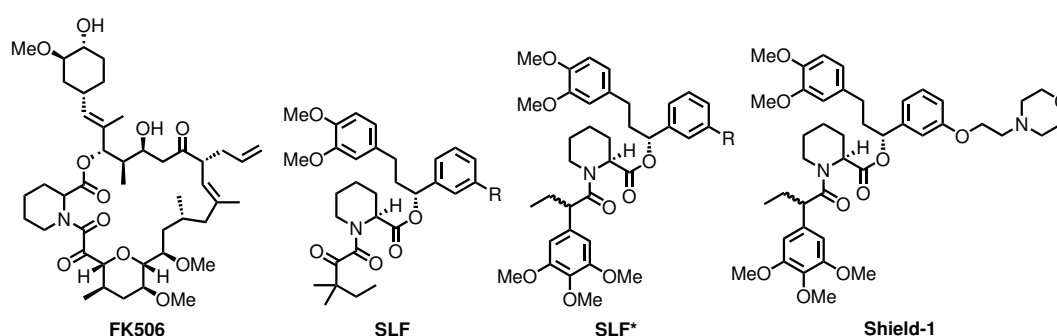
New self-labelling chemical tags continue to be introduced, but most of these are yet to be sufficiently tested. The Kikuchi lab genetically engineered a suicide protein based on the bacterial  $\beta$ -lactamase TEM-1 (BL-tag) (figure 6, A)[67]. The process whereby  $\beta$ -lactamases hydrolyse antibiotics containing  $\beta$ -lactam is conducted in two steps. In the first step, Ser70 reacts with the amide bond of the lactam, forming an ester, whilst in the second step, Glu166 catalyses the hydrolysis of this ester bond. Mutation of residue E166 to aspartic acid drastically reduced the deacylation reaction, leading to accumulation of protein-small molecule conjugate. Later, the same group developed an orthogonal tag-system based on the yellow protein (PYP-tag)[68–70]. PYP is a 125 amino acid small protein, which binds thioester derivatives of 4-hydroxycinnamic acid or 7-hydroxycoumarin through transthioesterification (figure 6, B). Importantly, PYP does not exist in mammalian cells and no endogenous target for its substrates is known. Bonasio *et al.* applied cutinase (a fungal serin esterase) variant that reacts with its suicide substrate *p*-nitrophenyl phosphonate (pNPP) as the tag (Cutinase tag) (figure 6, C). However, so far this system is limited to extracellular studies, due to the low cell permeability of the tag substrates.



**Figure 6:** Schematic representation of novel self-labelling tags based on proteins. A) BL-tag. B) PYP-tag. C) Cutinase-tag.

### 1.1.2 Affinity-based chemical protein tags

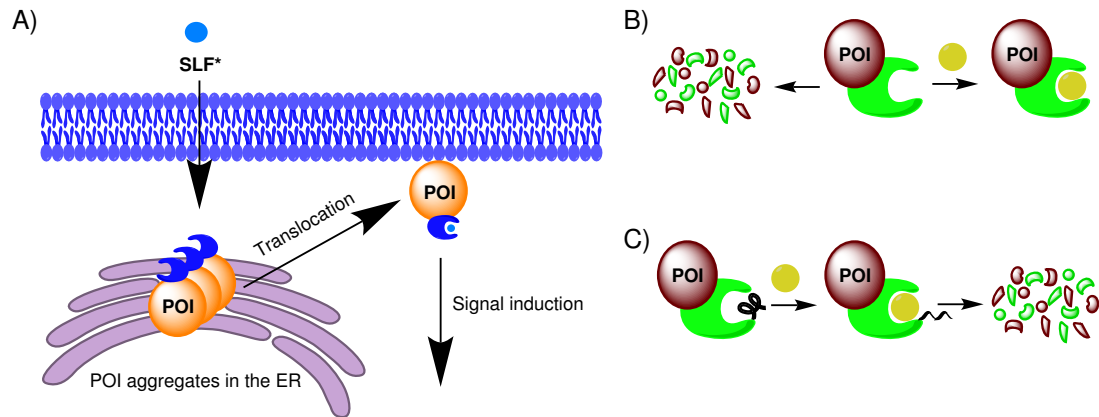
As the name suggests affinity-based chemical protein tags take advantage of high affinity interactions between a tag and their effectors. In contrast to most of the self-labelling tags discussed above, labelling of affinity-based tags can be reversed by addition of a competing tag substrate. One of the most established affinity tags is based on the immunosuppressant **FK506** (scheme 2), which binds tightly to the cytosolic Human FKBP12 protein and subsequently inhibits calcineurin. However, the ubiquitous and high expression of Human FKBP12 in mammalian cells limits this approach. To overcome this drawback, synthetic ligands for FKBP12 (**SLF**) lacking the immunosuppressant effect of **FK506**, were created (scheme 2)[71]. Applying the "bump and hole" approach, the research group of Clackson introduced a steric group into the synthetic ligand for FKBP12 (**SLF\***) to disrupt its binding to endogenous FKBP12 (scheme 2). Subsequent mutation of FKBP12(F36V), containing a corresponding "hole" re-established the binding of **SLF\***[72].



**Scheme 2:** Structure of the FKBP12 binding **FK506** and its none immunosuppressant analogs **SLF**, **SLF\*** and **Shield-1**

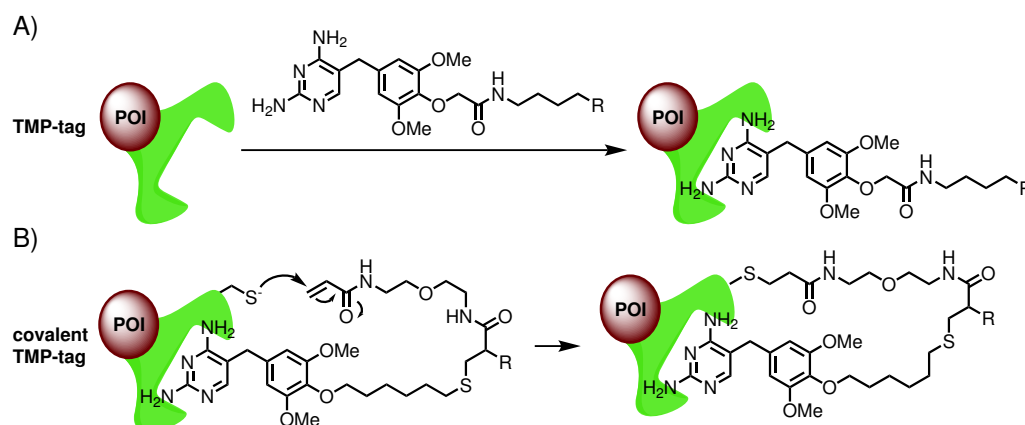
During the course of the FKBP remodeling study, Clackson and colleagues observed that the <sup>F36M</sup>FKBP mutant spontaneously formed dimers, which could be dissociated by addition of the **SLF\***[73]. Capitalising on this behavior, the group developed a "reverse dimerization" system to regulate protein secretion. In this system, POI fused to <sup>F36M</sup>FKBP form large aggregates, which are unable to exit the endoplasmic reticulum (figure 7, A). Incubation of cells with **SLF\*** dissolves the aggregates and ultimately restores the activity of POI. In this way, the authors regulated the secretion of insulin and growth hormones *in vivo*[74]. Extending the scope of applications of **FK506** based tags, the Wandless group genetically engineered an unstable FKBP mutant (DDFKBP) (figure 7, B)[5]. Proteins fused to this destabilising domain (DD), are readily and constitutively degraded when expressed in cells. In the presence of cell-permeable **Shield-1** (an **SLF** derivative with improved pharmacokinetic properties) (scheme 2), the protein fusion is rescued from proteolysis restoring their original function. To validate this approach, the authors regulated the levels of two kinases, two cell cycle regulatory proteins, and three small GTPases. Complementary to this, the same group developed a ligand-induced degradation system, fusing a 19 amino acid degradation sequence (degron) to the C-terminus of FKBP (LID)[75]. In the absence of **Shield-1**, the degron intra-molecularly binds to FKBP, thus preventing the degradation of the protein. Addition of a synthetic FKBP ligand displaces the degron and induces degradation of the POI-tag fusion protein (figure 7, C). In the proof of concept, the authors investigated the small molecule-dependent regulation of transcription factors involved in reprogramming differentiated cells. Additionally, the authors used **Shield-1** to simultaneously deplete

YFP-LID domain fusions and simultaneously rescue mCherry-DDFKBP from degradation.



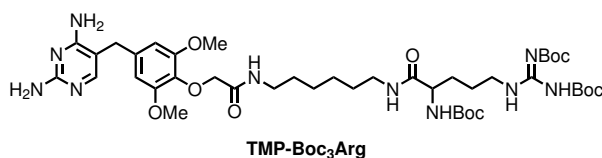
**Figure 7:** FK506-based approaches to regulate protein function. A) Self-assembling  $F^{36V}$ FKBPPOI fusion proteins accumulate in the ER. The membrane-permeable **SLF\*** binds to the FKBP mutant and disrupts the aggregate. The liberated chimeric proteins translocate to their active site, in this case the plasma membrane, where they regain their function. B) Proteins genetically fused to the destabilising FKBP mutant readily undergo proteolysis. Binding of **Shield-1**, stabilises the tag protein and rescues the fused protein from degradation. C) The protein of interest is fused to a FKBP mutant with C-terminal degron buried in the binding pocket. Addition of **Shield-1** displaces the degron, which induces proteolysis.

The group of Cornish capitalised on the high affinity of trimethoprim (**TMP**) to the protein dihydrofolate reductase from *E. coli* (eDHFR) to design an alternative high affinity tag (figure 8, A). **TMP** is easily modified at the para methoxy position and has low binding affinity for mammalian DHFR ( $K_d > 1 \mu\text{M}$ )[76]. The same group published a covalent version of the TMP-tag, using proximity-induced reactivity. Therefore, a lysine at the eDHFR surface was mutated to a cysteine. **TMP** binds to the mutated eDHFR and an acrylamide linked to the **TMP** reacts with the cysteine, forming a covalent bond (figure 8, B)[77].



**Figure 8:** Schematic representation of protein labelling with the TMP-tag technology. A) Reaction of **TMP** derivatives with eDHFR fusions B) **TMP**-directed covalent labelling of protein with a functionalised small molecule.

Analogues to their destabilising FKBP mutant, the Wandless group engineered a destabilising eDHFR domain (DDeDHFR), which can be prevented from inducing tag-POI fusion protein degradation by addition of **TMP**[78]. In the first applications, the authors regulated the protein levels of the transmembrane protein CD8 $\alpha$  and POIs in rat brain. Additionally, the group simultaneously and independently regulated the concentration of two different proteins in one cell, co-expressing DDeDHFR-POI and DDFKBPOI fusion proteins. Conceptually similar, the Hedstrom group designed a **TMP-Boc<sub>3</sub>Arg** derivative to post-translational control protein levels (scheme 3)[79]. This hydrophobic small molecule binds to its target protein and triggers proteasome-mediated protein degradation. Notably, the proteolysis of the fusion protein could be stopped with the addition of an excess of **TMP**.



**Scheme 3:** Chemical structure of **TMP-Boc<sub>3</sub>Arg**.

Recently Tsukiji and colleagues designed a family of self-localising ligands (SLLs) based on **TMP** and **SLF\*** to control the spatial distribution of POIs, *in vivo*[80]. In this approach the ligand for an endogenous protein or chemical tag is covalently linked to a specific small-molecule localisation motif. After cell incubation with this bi-functional small molecule, the POI is rapidly translocated to a targeted intracellular region. In the proof of concept, the authors translocated POIs to the inner leaflet of the plasma membrane, the nucleus and to the cytoskeleton.



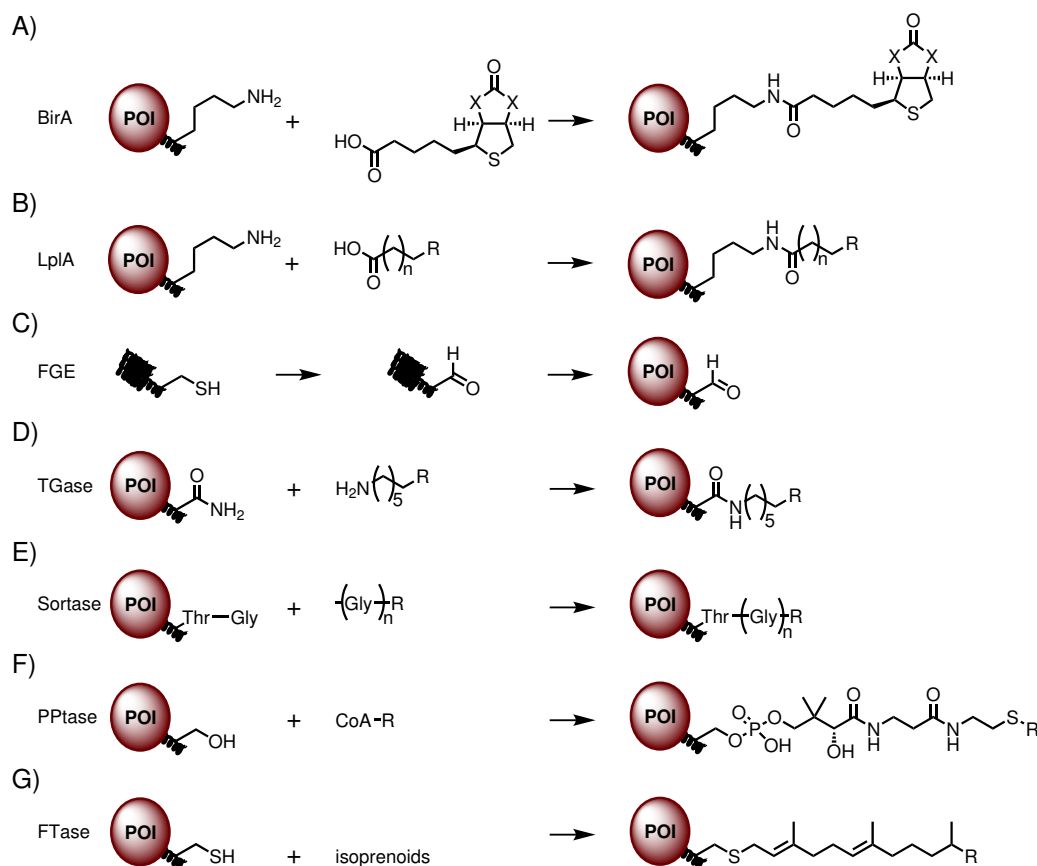
### 1.1.3 Enzyme-mediated chemical tags

Enzyme-mediated chemical tags were exploited to combine the advantages of metal-binding peptides and self-labelling chemical tags. In this approach, the POI is fused to a small peptide, which is modified by a small-molecule cofactor in the presence of the corresponding enzyme. **biotin** ligase from *E. coli* (BirA) catalyses the reaction of **biotin** with specific polypeptide, which is not recognised by **biotin** ligases from other species (figure 9, A). To reduce the size of the tag, the laboratory of Schatz identified a 14-residue peptide as a minimal requirement for BirA-catalysed biotinylation. Initially, the biotinylated proteins were further modified with streptavidin-conjugates[81, 82]. In 2005, the Ting lab discovered that BirA tolerates a ketone derivative of **biotin** as a substrate, which enabled the possibility to specifically functionalise the ketobiotin with hydrazide or aminooxy compounds in bioorthogonal reactions[83]. More recently, the same group applied biotin ligases from *yeast* (yBL) and *Pyrococcus horikoshii* (PhBL), to covalently bind biotin analogues with azide or alkyne groups to an acceptor protein[84]. However, more dramatic modifications of the **biotin** are not tolerated. Furthermore, the requirement of a second bioorthogonal reaction to introduce a specific function, limits this approach. Extending their studies, the same group investigated the covalent ligation of lipoic acid to a 22-amino acid long *E. coli* lipoic acid ligase (LplA) substrate (LplA-tag) (figure 9, B). Initially, LplA exhibited similar small molecule substrate tolerance as BirA and only small functional groups, such as azides or alkynes, were tolerated[85]. Mutation of the crucial residue W37 within LplA expanded the substrate specificity, allowing introduction of coumarins[86–88], aryl aldehydes[89], aryl hydrazines[89], aryl iodide[90], and *trans*-cyclooctene[91]. Despite these advances, most tolerated substrates depend on a second reaction to introduce functionalities for specific applications. Moreover, incorporation of different and more complex substrates requires an individual mutant for each substrate.

The formylglycine-generating enzyme (FGE), another enzyme applied to site-selectively modify proteins, suffers from similar limitations. FGE oxidizes a cysteine within a six-amino acid motif (aldehyde tag) found in tyrosine sulfatases to formylglycine (FGly) (figure 9, C). The modification occurs prior to the protein folding and is required for the catalytic activity of sulfatases[92]. Bertozzi and colleagues genetically encoded the aldehyde tag to cytosolic, as well as membrane-associated proteins in *E. coli* and mammalian cells[92–94]. Co-expression of tagged POIs and FGE rapidly produced aldehyde-containing proteins, which could be further functionalised by condensation reactions with aminooxy or hydrazide probes[92–94]. Interestingly, endogenous proteins possessing the FGly residue (e.g. lysosomal and secreted sulfatases) did not undergo condensation reactions, outlining the high selectivity of the aldehyde tag[92]. Lin and Ting exploited the transglutaminase (TGase) catalysed amide formation between the central glutamine within a small polypeptide tag (Q-tag) and aliphatic amine derivatives (figure 9, D)[95]. In a proof of concept, the authors selectively labelled cell surface proteins fused to a Q-tag with functionalised cadaverine derivatives in the presence of TGase. Although, TGase tolerates a broad scope of amine substrates, its low selectivity for Q-tag and competing endogenous TGase substrates, limits this method to extracellular studies[95].

Sortase is another well-studied enzyme used to incorporate novel functionalities into proteins. One of the most commonly applied sortases is Sortase A from *Staphylococcus aureus*, which cleaves a 5 amino acid motif (LPXTG) between threonine and glycine and subsequently links the carboxyl group of threonine to the amino group of a N-terminal glycine in the presence of  $\text{Ca}^{2+}$  (figure 9, E)[96]. Popp *et al.*, successfully demonstrated labelling of POI-LPETG fusion proteins *ex vivo*[97]. Similarly, the group of Nagamune site-specifically labelled POIs expressed at the cell surface with a fluorophore- or biotin-containing peptide derivative[98]. More recently, the Ploegh group used the  $\text{Ca}^{2+}$ -independent sortase A from *Streptococcus pyogenes* to target

proteins in the cytosol and in the lumen of the endoplasmic reticulum[99].



**Figure 9:** Protein labelling with different enzyme-mediated chemical tags. A) BirA catalysed labelling of specific polypeptide with **biotin** (X = NH) or its ketone bearing analogue (X = CH<sub>2</sub>). B) LplA catalysed ligase of a lipoic acid derivative to the primary amine in the substrate sequence. C) FGE catalysed oxidation of the thiol to the corresponding aldehyde in the unfolded protein and subsequent protein folding. D) TGase catalysed amide formation between the primary amine and the central glutamine in the Q-tag. E) Sortase catalysed labelling of POI F) PPTase catalysed phosphorylation of a serine in the substrate G) FTase catalysed isoprenylation.

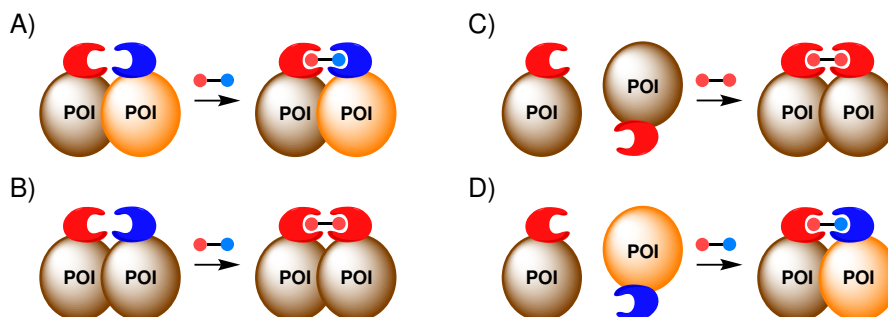
Phosphopantetheinyl transferases (PPTases) from *E. coli* (AcpS) and *B. subtilis* (Sfp) are arguably the most promising enzymes to modify proteins, as they tolerate a variety of functionalities and are orthogonal to mammalian PPTases[100]. Sfp and AcpS catalyse the transfer of a phosphopantetheinyl unit from coenzyme A (CoA) to a serine residue within peptide carrier protein (PCP) and acyl carrier protein (ACP), respectively (figure 9, F)[100, 101]. In a proof of concept, Walsh and co-workers genetically fused PCP to the N-terminus of  $\beta$ -galactosidase or luciferase. Sfp post-transcriptionally modified tagged proteins with biotinylated CoA, in cell lysates[101]. Simultaneously, the Johnsson group reported the AcpS-catalysed labelling of ACP-cell surface protein fusions[100]. However, the low cell permeability of CoA limited these approaches to extracellular studies. To overcome this drawback, Clarke *et al.* capitalised on the biosynthesis of CoA[102]. The authors incubated *E. coli* cells with a nonhydrolyzable and cell permeable coumarinyl derivative, which was biosynthetically converted to the corresponding

coumarinyl CoA analogue. Subsequently, Sfp recognised this analogue and linked it to a PCP fusion protein. To reduce the size of ACP and PCP tags, Walsh and colleagues screened a phage-displayed library[103]. The resulting 11-residue peptide YbbR13 was selectively recognised by Sfp. Extending their studies, the same group identified two new peptides (S6 and A1) with orthogonal reactivities for Sfp and AcpS[104].

Farnesyltransferase (FTase) and geranylgeranyl transferase-1 (GGTase-1) are enzymes, which attach 15 or 20-carbon long isoprenoids to a four amino acid sequence (CAAX) (Figure 9, G)[105]. Although, FTases and GGTases-1 tolerate a variety of isoprenoid derivatives, site-reactions with endogenous proteins containing the CAAX sequence, limit this approach to extracellular studies[105, 106]. However, the results in recent studies suggest that mutation of the acceptor peptide and enzymes from different species may lead to completely orthogonal enzyme/substrate pairs, which could be applied *in vivo*[107–109].

### 1.2 Labelling two proteins with a bifunctional molecule

In the last decades, *in vivo* studies revealed that the function of proteins is strongly regulated by the formation of multi-protein complexes. Each protein typically has several alternative interaction partners, and the selectivity of this protein-protein interaction (PPI) determines the specific function of the proteins. Techniques to identify interaction partners, such as Förster resonance energy transfer (FRET), *yeast* two hybrid systems, protein complementation or immunoprecipitation from cell lysates, are restricted to rather strong interactions[3]. In contrast, small bi-functional molecules can induce protein-protein interactions or increase their stability. Protein cross-linkers covalently link previously dimerised interacting proteins together (figure 10, A and B), whereas chemical inducers of dimerisation (CIDs) force the dimerisation of two proteins (figure 10, C and D). Two types of bi-functional molecules can be distinguished regarding their functional groups. Homobi-functional molecules contain identical functional groups at both reactive sites and heterobi-functional reagents contain two orthogonal functions[3, 110, 111].

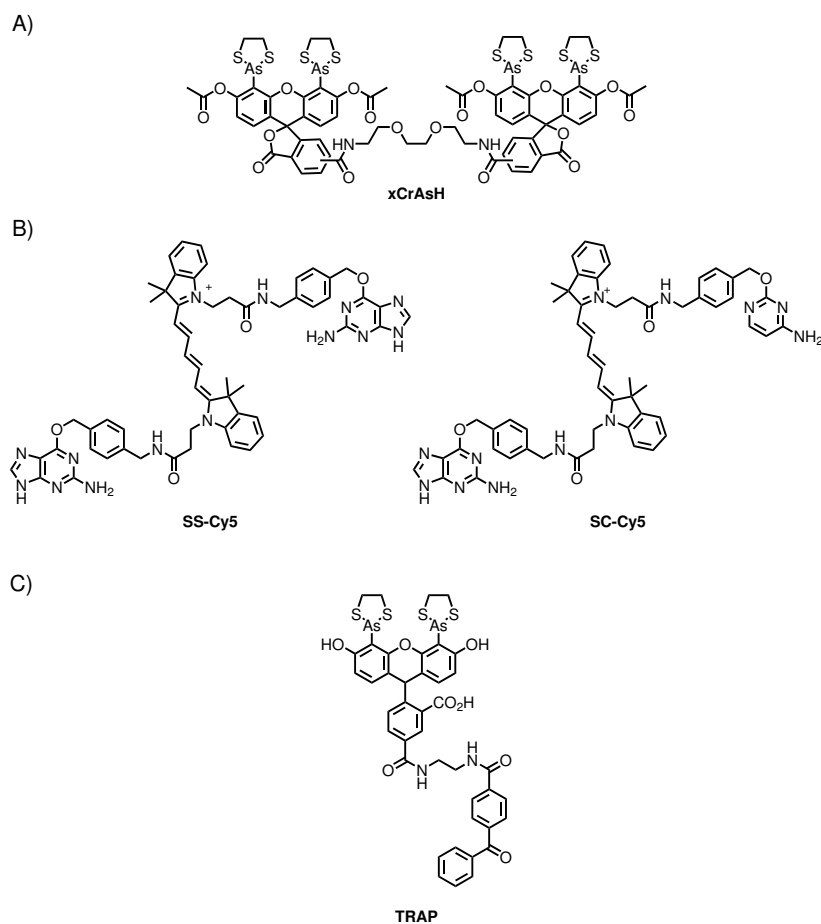


**Figure 10:** Concepts of chemical inducers of dimerisation and selective bi-functional cross-linkers. A) Cross-linking of two identical self-dimerising proteins. B) Cross-linking of heterodimerising proteins. C) Dimerisation of two identical proteins with a homodimeriser. D) Dimerisation of two different proteins with a heterodimeriser.

### 1.2.1 Protein cross-linkers

A wide variety of direct homobi-functional and heterobi-functional cross-linkers are commercially available[3]. Most of these bi-functional molecules contain NHS ester derivatives, maleimide and/or disulfides to label amines or thiolates on the protein surface. However, these cross-linkers are mainly limited to *in vitro* experiments, due to their poor selectivity. This drawback was partially solved for heterobi-functional cross-linkers by using a two-step reaction to link the two proteins of interest[3]. Therefore, a photoreactive moiety was introduced on one side of the cross-linker. Firstly, the photoinensitive functional group reacts with the protein and after irradiation, the photoreactive group reacts non-specifically with the nearest C-H or N-H bond[112]. Photoreactive groups used for this approach include aryl azides and diazirine rings, which form highly reactive nitrene, respectively carbene, intermediates. To further improve this approach, various groups introduced photoreactive amino acids at defined positions of the protein[113, 114]. However, their small size requires close proximity of the interacting proteins and incorporation of several unnatural amino acids into the protein. Alternatively, a chemical tag-based approach can be used to selectively cross-link interacting proteins. Interestingly, the group of Scherpartz split the TC tag into two bicysteine (BCys) motifs[115, 116]. The bipartied TC tag were fused apart from each other, either in two different proteins or in the same. Upon interaction or folding of the BCys containing proteins, the two dicysteine pairs were brought in close proximity and were recognised by **FlAsH** or **ReAsH**. However, the short distance between the two arsenic groups ( $= 4.65 \text{ \AA}$ ) limits this approach. To overcome this limitation, Schultz and co-workers synthesised a dimeric derivative of **FlAsH** (**xCrAsH**). Two bisarsenic derivatives of carboxyfluorescein were connected with a flexible spacer. **xCrAsH** successfully cross-linked FKBP fused to TC tag with FRB fused to TC Tag in the presence of the CID **rapamycin** (section 1.2.2.3), *in vivo* (scheme 4, A)[117]. Alternatively, the group of Johnsson designed a novel technique (S-CROSS) to detect and characterise hetero- and homo-dimeric protein-protein interactions in cell lysate[118]. S-CROSS is based on the specific cross-linking of SNAP-tag and/or CLIP-tag fusion proteins with a bi-functional small molecule consisting of fluorescent **BG** and/or **BC** derivatives (scheme 4, B). To demonstrate the utility of S-CROSS, the authors quantitatively and simultaneously studied a number of homotypic and heterotypic protein-protein interactions. Moreover, S-CROSS permitted to investigate the stability of protein complexes and to differentiate between strong and weak interactions. In collaboration with the Johnsson group, Gönczy and colleagues used fluorescent microscopy techniques in combination with S-CROSS to investigate interactions between centrosomal proteins[119]. Therefore, a set of 31 centrosomal proteins fused to SNAP-tag and CLIP-tag were pairwise incubated with a fluorescent BG-BC. Samples were subjected to SDS gel electrophoresis and analysed by in-gel fluorescence scanning, resulting in identification of three novel interacting partners. Recently, Sun *et al.*, reported the first *in vivo* application of S-CROSS[120]. In this study, S-CROSS was used to investigate the ligand-mediated homodimerisation of the epidermal growth factor receptor (EGFR) and subsequent activation of downstream transduction pathway. Thus, cells expressing recombinant EGFR-SNAP-tag fusion proteins were treated with a bi-functional BG cross-linker in the presence or absence of ligand stimulus. Although only small amounts of dimers were observed in unstimulated cells, these results support the hypothesis that EGFR monomers and dimers are in equilibrium. However, PPI interactions identified with bi-functional tag substrates require careful analysis as these cross-linkers may also induce dimerisation on their own[117, 121]. Alternatively, the Ting lab engineered an LplA mutant, which specifically labelled its acceptor peptide with a photoreactive aryl azide. To validate their method, cell lysate containing predimerised FRB and FKBPLAP fusion proteins and the LplA mutant were treated with the aryl azide derivative and subsequently irradiated[122, 123]. Mayer and co-workers designed a

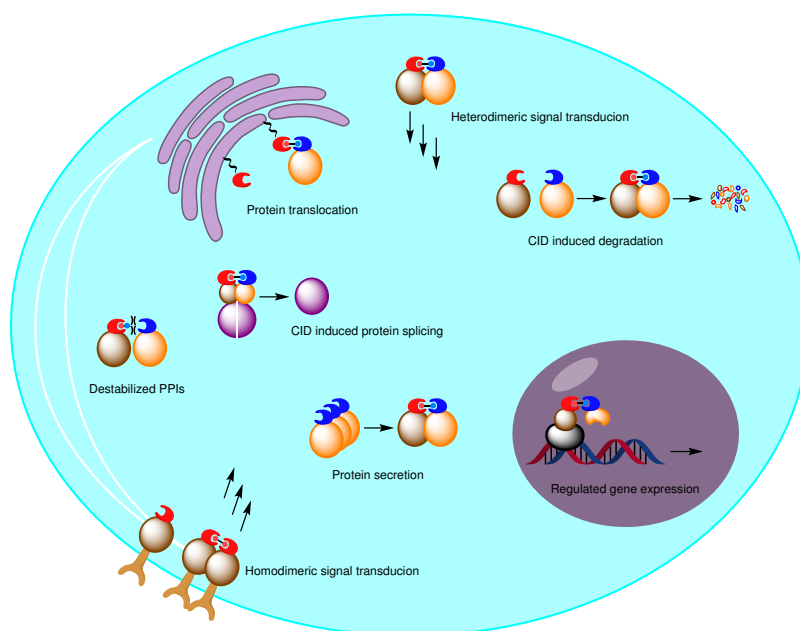
targeted and releasable affinity probe (**TRAP**), consisting of a **FlAsH** derivative covalently tethered to a photoreactive benzophenone (scheme 4, C)[124]. **TRAP** targets specific POIs fused to TC-tag and captures binding partners upon light activation. The authors successfully used **TRAP** to identify fibronectin as a binding partner of the abundant muscle membrane protein, phospholamban.



**Scheme 4:** Chemical structure of the self-labelling tag based cross-linkers. A) **xCrAsH** B) S-CROSS molecules: **SS-Cy5** and **SC-Cy5** C) **TRAP**

### 1.2.2 Chemical inducers of dimerisation

While cross-linkers are useful tools to detect and characterise naturally occurring PPIs, chemical inducers of dimerisation (CIDs) allow specific and independent regulation of protein activity. In principle, CIDs are small organic molecules, which trigger the assembly of two POI-Tag fusion proteins resulting in a trimeric complex. Homodimerisers promote the dimerisation of two identical proteins, whereas heterodimerisers force the dimerisation of two distinct proteins. The nature of the CID can vary from a simple natural bi-functional product that binds to proteins (e.g. **coumermycin**, **rapamycin**), molecules that induce a conformational change of the protein, thereby creating a binding site for a second protein (e.g. **S-(+)-abscisic acid**, **gibberellin**), to synthetic molecules, which are composed of two protein ligands connected with a flexible linker[125]. In general, natural CIDs display better biophysical properties, such as cell permeability and water solubility, but are harder to modify. In contrast, the linker of synthetic CIDs allows the relatively simple introduction of new functionalities or modification of the biophysical properties. Two general strategies are frequently used to artificially regulate protein activity (figure 11). In the first strategy, the CID interferes with the activity of the protein. In this way, the CIDs are used to dimerise two proteins to force a specific response[126, 127], translocate proteins to or away from their site of action to trigger or prevent signal transduction[128–132], induce protein secretion[133] and destabilise PPI by inducing unfavorable steric interactions[134, 135]. In the second strategy, the protein levels are controlled. In the most common approach, the CID triggers the dimerisation of a DNA-binding domain (DBD) fusion and a transcriptional activation domain (TAD) fusion, resulting in the transcription of a targeted gene[136, 137]. Other approaches regulate protein levels post-translationally through translocation POIs to the proteasom[138], dimerisation of the POI to the E3 ubiquitin ligase[139], reconstitution of split ubiquitin sequences[140], or rely on unstable protein domains[141, 142]. Finally, CIDs can be used to dimerise fusions of split-inteins, thereby generating new proteins with custom-made properties[143, 144].



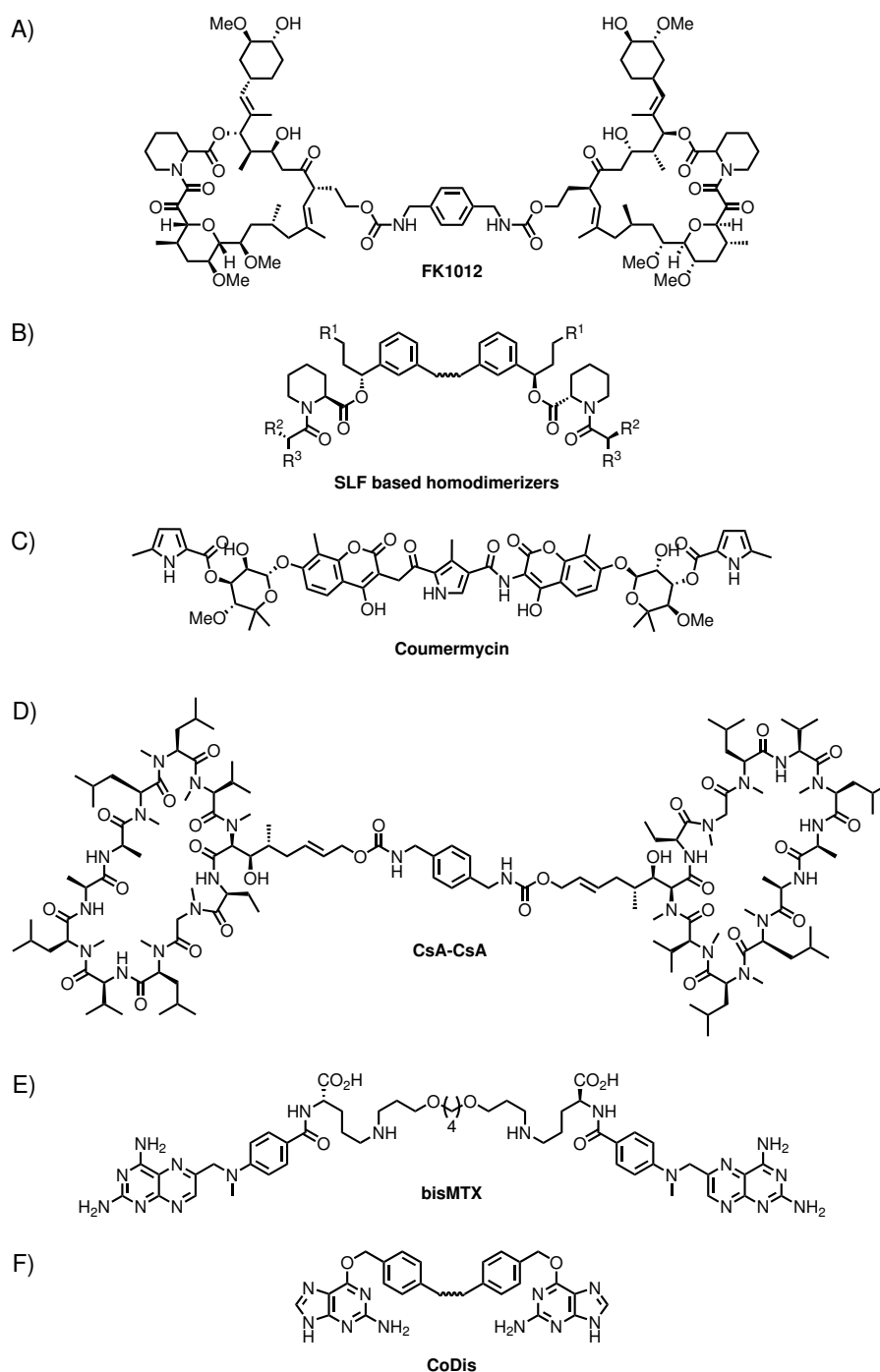
**Figure 11:** General approaches to regulate protein function with chemical inducers of dimerisation.

### 1.2.2.1 Homodimerisers

The groups of Schreiber and Crabtree pioneered in the field of CIDs in the early 90's[145]. Inspired by the naturally occurring heterodimeriser **FK506**, which binds to the cytoplasmic protein FKBP12 and calcineurin, they synthesised a symmetrical **FK506** dimer (**FK1012**) (scheme 5, A) able to bind to two FKBP12 proteins. To demonstrate the utility of this tool, the authors treated cells expressing FKBP12 domains fused to the  $\xi$ -subunit of the T-cell receptor with FK1012. Dimerisation of  $\xi$ -subunits led to rapid and dose-dependent activation of downstream targets. Although, FK1012 was used in several biological studies, binding to endogenous FKBP12, its relatively big size and intolerance for modifications limited its application. Since this first report, numerous synthetic FK1012 variants were designed with reduced affinity to endogenous FKBP12, higher specificity, reduced complexity and improved conformational orientation (scheme 5, B)[71, 72, 146, 147]. FK1012 and its improved analogues were frequently used to regulate numerous biological processes, including the activation of MEKK2 [148], the Fas-receptor and caspase mediated apoptosis[149–154], the activation of c-Abl Kinase[155], and protein synthesis in neurons[156], among many other biological processes.

A complementary homodimeriser system is based on the N-terminal domain (220 amino acids) of the B subunit of bacterial DNA gyrase (GyrB) and its natural bi-functional ligand, **coumermycin** (scheme 5, C)[157]. Importantly, the **coumermycin**-mediated dimerisation of GyrB can be readily reversed by addition of the monofunctional substrate, novobiocin. Moreover, neither **coumermycin** nor novobiocin has any high-affinity targets in eukaryotic cells. In a first application, Farrar *et al.* successfully activated MEK by homodimerising Raf-1-GyrB fusions in the cytosol[127]. Interestingly, downstream targets such as Erk1 and Erk2 were not affected by cytosolic Raf activation. In contrast, recruitment of Raf-GyrB fusion to the membrane, prior to activation with **coumermycin**, led to complete activation of downstream targets, underlining the importance of correct localisation to activate signaling pathways[158]. Apart from this study, the coumermycin/GyrB system was extensively used to investigate the Jak/Stat pathway[159–165], dimerise membrane-bound receptors[166–169], study intracellular signal transducers[170–172] and control gene expression[173].

Several different systems were developed to expand the palette of homodimerisers. In an early example, the groups of Crabtree and Schreiber used the orthogonal chemical genetic approach to create an artificial receptor-ligand pair based on cyclosporin A (**CsA**) and its target cyclophilin (CyP)[174]. Connecting two modified **CsA** molecules via a short linker, resulted in a cell-permeable non-toxic homodimeriser (**CsA-CsA**) (scheme 5, D)[175]. In a first application, this **CsA-CsA** dimeriser was used to induce Fas-mediated cell apoptosis. The group of Hu connected two **MTX** derivatives with a flexible linker (scheme 5, E). The resulting **bisMTX** molecule was able to homodimerise DHFR domains *in vitro*[176]. Interestingly, excessive addition of dimeriser did not lead to saturation and formation of dimeric bisMTX-DHFR complex, suggesting a cooperative binding model. Likewise, the group of Johnsson synthesised a series of small molecules containing two **BG** groups connected with different flexible linkers (**CoDis**) (scheme 5, F)[121]. In a proof of concept, the authors treated cells expressing fusions of POI-SNAP-tag, with **CoDis**, which resulted in rapid and proximity-dependent dimerisation.

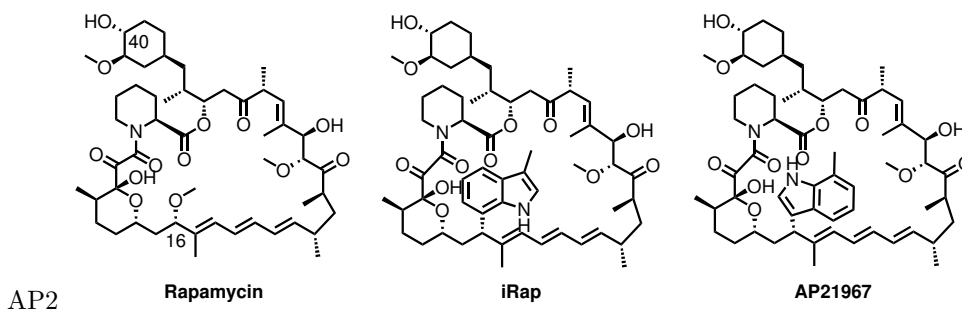


**Scheme 5:** Chemical structure of homobi-functional chemical inducers of dimerisation. A) FKBP12 dimerising **FK1012**. B) general structure of **SLF** based homodimerisers. C) GyrB dimerising **coumermycin**. D) Cyclophilin dimerising **CsA-CsA**. E) DHFR dimerising **bisMTX**. F) general structure of **CoDis**.



### 1.2.2.2 Heterodimeriser

The most widely used heterodimeriser system is based on the natural immunosuppressant **rapamycin** (Rap) or its analogues (Rapalogs) (scheme 6). **rapamycin** sequentially and rapidly binds to FKBP12 and the FRB (FKBP12-**rapamycin**-binding) domain of mTOR (mechanistic or mammalian target of **rapamycin**), thereby forming a tertiary FKBP-Rap-FRB complex[177]. The formation of this complex leads to efficient inhibition of mTOR complex 1 (mTORC1), a kinase strongly involved in cell growth and cell proliferation[178]. In its first application, Rivera *et al.*, regulated the gene expression of the secreted alkaline transferase in human fibrosarcoma cells[179]. Since this initial study, the **rapamycin** dimeriser system has frequently been used to study numerous biological processes[180]. However, the intrinsic activity of **rapamycin** dramatically limits its applications. To overcome this drawback, research groups subjected this system to the "bump and hole" approach, to generate a bioorthogonal ligand-protein pair. X-ray crystal structures of the FKBP12-**rapamycin**-FRB tertiary complex revealed that the C16-methoxy group of **rapamycin** directly points toward the FRB domain[178]. The group of Luengo demonstrated that substitution of this methoxy group can alter the FRB-binding affinity[181]. Inspired by this study, the Schreiber group synthesised a series of rapalogs, containing a bulky substituent at FRB-binding site and evaluated their toxicity (e.g. **iRap**, **AP21967**, scheme 6) [182]. Triple mutation in the  $\alpha$ -helix of FRB led to FRB mutants with a compensatory "hole" able to bind the identified non-toxic rapalogs. Up until now, numerous C16 rapalogs/FRB mutant pairs have been designed to improve the pharmacokinetics of the rapalogs[131, 141, 181].

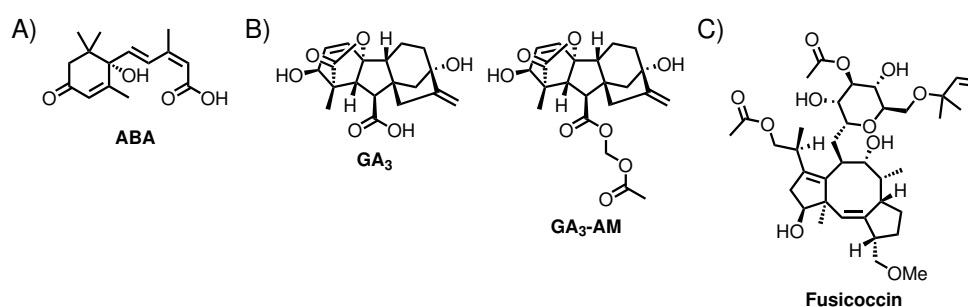


**Scheme 6:** Chemical structure of **rapamycin** and two mTORC1 safe rapaloges **iRAP** and **AP21967**.

Aside from the nucleophilic substitution at the C16 position, to disrupt endogenous FRB binding, modification of the C40-hydroxyl allows for introduction of specific functionalities without interfering with FKBP binding. Among many examples of C40 rapalogs, some of the most exciting ones are labelled with fluorescein for cell imaging[183, 184], attached to biotin derivatives to immobilise the rapalog on surfaces[183], or contain photocleavable groups to spatiotemporally control the dimerisation (section 1.3.1)[185, 186]. Since the initial report from Schreiber *et al.*, **rapamycin** and rapalogs, were commonly used to manipulate GTPase signaling pathways[131], intracellularly activate GPCRs[187, 188], and regulate phosphoinositide levels at specific cellular compartments[189–191]. Apart from inducing transduction pathways, the **rapamycin** system has been used to develop several techniques to translocate the POI to the proteasome, which induced complete degradation of the POI[138]. In contrast, Crabtree and co-workers designed a destabilised FRB mutant[142]. Analogues to DDFKBP and DDDHFR proteins, fused to this destabilised FRB mutant, are readily degraded, which is prevented by the addition of **rapamycin**. The group of Muir used the **rapamycin** system to post-translationally change the primary structure of a protein[143, 144]. Therefore, they combined the concepts of trans-splicing and CIDs. Splicing refers to a naturally occurring process, in which an internal domain of the

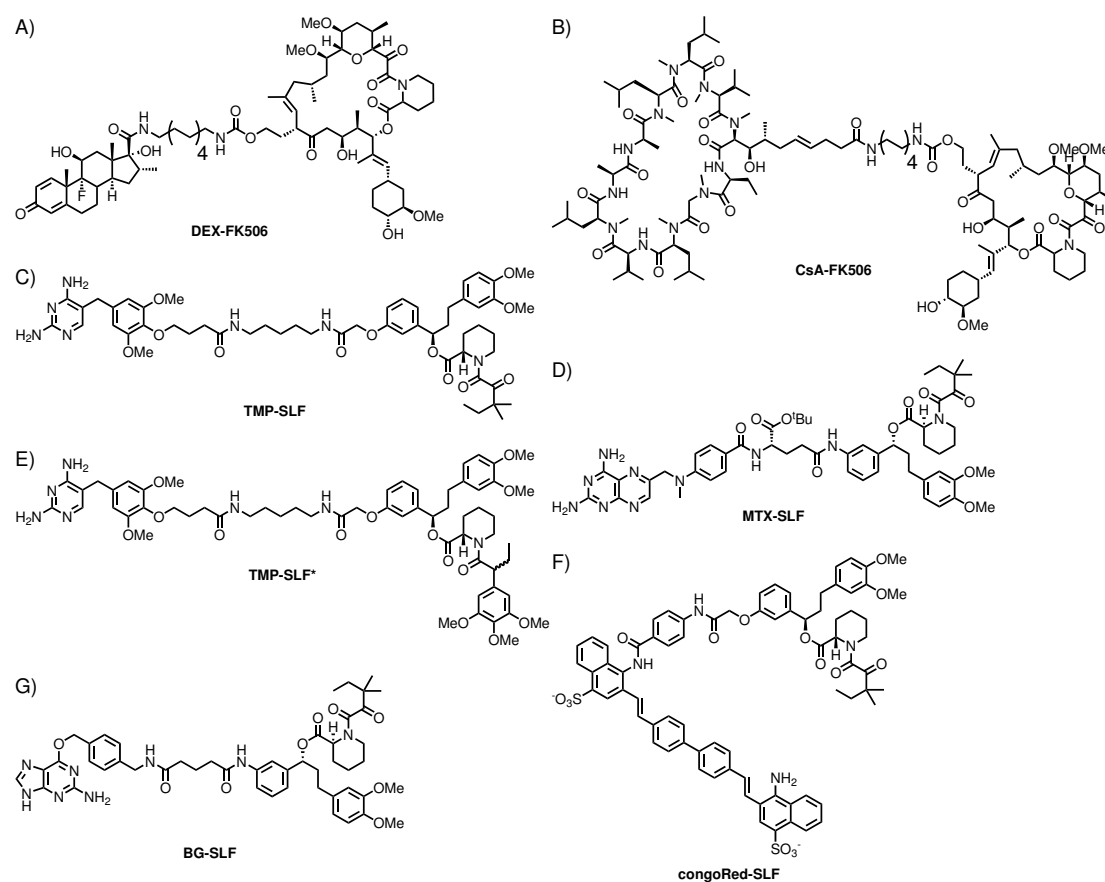
protein (intein) is excluded from the protein followed by ligation of the two adjacent sequences (exteins). In the presence of **rapamycin**, split-inteins fused to FKBP and FRB, were complemented and the exteins fused to the inteins, were ligated. Pratt *et al.* reported a conceptually similar approach to control protein levels[140]. They inserted a split-ubiquitin and FRB between a POI and a small sequence, that induces protein degradation (degron). Addition of **rapamycin** results in dimerisation with FKBP and the complementary split-ubiquitin. The reconstituted ubiquitin is subsequently recognised and cleaved by a protease, thus releasing and rescuing the POI[140]. Alternatively, the **rapamycin** method can be used to anchor POIs away from their site of action and thereby inhibiting the function of the protein. This approach was applied to arrest adaptor proteins to the mitochondrial membrane, and to trap nuclear proteins in the cytoplasm by dimerising these proteins to large ribosomal subunits[129, 130]. **Rapamycin** was also used to shuttle several proteins across the nuclear membrane[141, 192–195].

Recently, three other natural products were reported to chemically induce protein heterodimerisation. Binding of S-(+)-abscisic acid (**ABA**) (scheme 7, A) to a PYL1 protein domain, induces a conformational change in this domain, thereby creating an extensive binding surface for an ABI1 domain. **ABA** and its target domains fused to corresponding proteins were used to induce gene transcription, localise proteins to subcellular compartments, and regulate signaling transduction in mammalian cells. Moreover, the **ABA** system and the **rapamycin** system were used in parallel to translocate different proteins independently[126]. However, this novel heterodimeriser system displayed significant slower reaction kinetics compared to the **rapamycin** system. Analogues, the plant hormone gibberellin (**GA<sub>3</sub>**) (scheme 7, B) regulates the dimerisation of truncated GAI and GID1 fusion proteins. However, the negative charge of the carboxyl group reduces the cell permeability of this dimeriser dramatically. To overcome this drawback, Myamoto *et al.*, protected this side of the molecule with an acetoxymethyl group (**GA<sub>3</sub>-AM**) (scheme 7, B), which is readily hydrolysed by endogenous esterases[196]. The orthogonality of the **GA<sub>3</sub>** and **rapamycin** was demonstrated by recruiting CFP-FKBP to membrane bound FRB, and YFP-CID1 to mitochondria localised GAI. In contrast to **ABA**, **GA<sub>3</sub>** induced dimerisation of its target proteins with similar speed compared to **rapamycin**. Ottmann and his group used the natural product **Fusicoccin** (scheme 7, C), a diterpenoid glycoside to dimerize T14-3-3-cΔC and CT52 fusion proteins. In proof of concept, the authors successfully translocated POIs to the plasma membrane, into and out of the nucleus. Notably, the fusicoccin-induced dimerization of fusion proteins could be reversed by exchanging the medium. However both the dimerization and dissociation of T14-3-3-cΔC-CT52 fusion proteins proceeded at relatively slow rates[137].



**Scheme 7:** Chemical structure of protein dimerising plant hormones. A) **ABA**. B) **GA<sub>3</sub>** and the cell permeable analogue **GA<sub>3</sub>-AM** C) **Fusicoccin**.

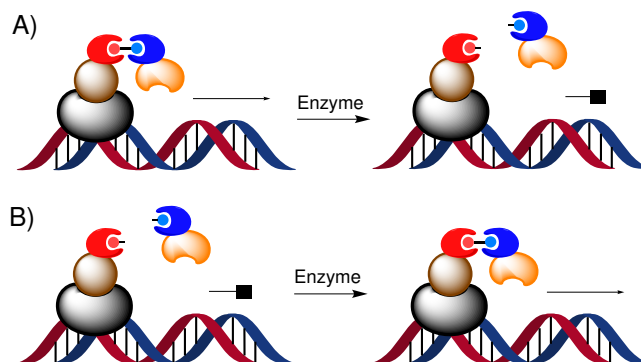
To increase the scope of application, several groups generated synthetic heterodimerisers connecting two tag ligands with a variety of linkers. An important technique developed on the basis of synthetic heterodimerisers, is the *yeast* three-hybrid system[198]. In contrast to standard gene expression assays, the synthetic CID consists of a known ligand for the DNA-binding domain fusion and a ligand for the unknown protein. Screening of activation domain fusion libraries allows the identification of protein targets for the ligand under investigation. In the course of developing this system, Licitra and Liu designed the first synthetic heterodimeriser, based on **FK506** and the glucocorticoid ligand (GR) dexamethasone (**DEX-FK506**) (scheme 8, A)[198]. Simultaneously, Belshaw *et al.*, reported a synthetic heterodimeriser, tethering **FK506** and **CsA** with an aliphatic linker (**FK506-CsA**) (scheme 8, B)[199]. First applications of this CID have included recruitment of a POI to a specific subcellular location and regulation of gene expression. In recent years, several conceptual similar CIDs were reported. The Cornish group for example covalently linked the tag substrates **TMP** and **SLF** (**TMP-SLF**) to promote the dimerisation of FKBP and eDHFR fusion proteins (scheme 8, C)[200]. In 2003, the group of Wandless designed a variation of this CID system. The authors introduced a very small linker between **MTX** and **SLF** (**MTX-SLF**) to destabilise protein-protein interactions (scheme 8, D)[134, 135]. The resulting steric collisions and electrostatic repulsions between the dimerised proteins were severe enough to favor reaction of the bi-functional molecule with only one tag, thus preventing dimerisation. In a model experiment, DHFR was selectively inhibited in cells lacking FKBP. In contrast, DHFR was resistant to treatment with the **MTX-SLF** dimeriser in cells expressing FKBP[135]. Gestwick *et al.*, used a conceptually similar approach to prevent aggregation of A $\beta$ -amyloid, a key process in the pathogenesis of Alzheimer[133]. Therefore the amyloid ligand, Congo Red, was chemically linked to **SLF** (**congoRed-SLF**) (scheme 8, E). The molecule-induced dimerisation of FKBP and A $\beta$ , which sterically prevented aggregation of the latter. In 2014, the Wu group regulated targeted gene expression with an **SLF\*-TMP** dimeriser (scheme 8, F)[137]. Notably, addition of monomeric **TMP** rapidly disrupted the heterodimeric complex. Similarly, Schultz and co-workers developed a reversible CID system based on the SNAP-tag and <sup>F36V</sup>FKBP[201]. Therefore, the authors synthesized **BG-SLF\*** (scheme 8, G). The addition of SLF\* readily led to the dissociation of BG-SLF\* -induced SNAP-tag and <sup>F36V</sup>FKBP tag fusions proteins. Feng *et al.* used this system to reversibly control PI3K activity in HeLa cells.



**Scheme 8:** Chemical structure of synthetic CIDs based on FKBP12-FK506 binding. A) **DEX-FK506**, B) **CsA-FK506**. C) **congoRed-SLF**. D) **MTX-SLF**. E) **TMP-SLF**. F) **BG-SLF**. G) **TMP-SLF\***.

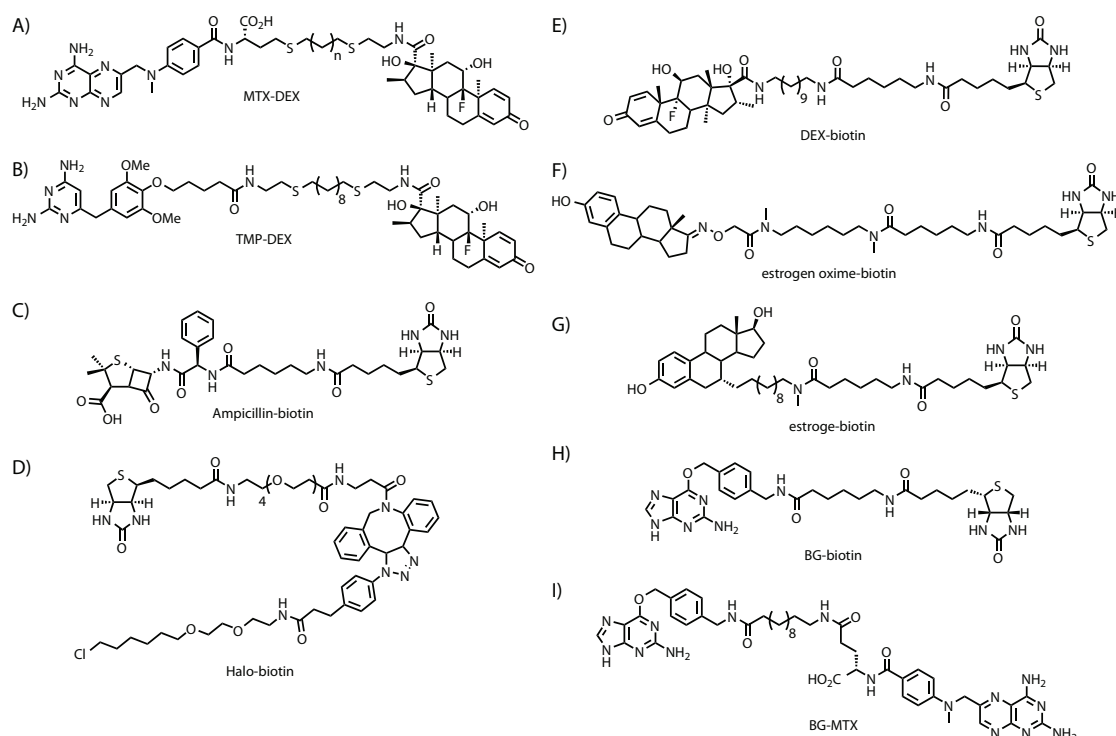
Given the broad utility of these heterofunctional CIDs, the Cornish lab prepared an alternative dimeriser system, containing ligand-receptor pairs, which avoid FKBP12-binding. Through this approach, DHFR targeting **MTX** and **DEX** were covalently connected with different linkers (scheme 9, A)[202, 203]. Based on these **MTX-DEX** dimerisers, the same group designed two three-hybrid systems to monitor the activity of enzymes[8]. In the first approach, the enzyme target is introduced as a linker between **DEX** and **MTX** as a linker. In the presence of active enzyme, the linker between the protein ligands is cleaved, leading to disruption of the reporter gene expression (figure 12, A)[8]. As a model reaction, the  $\beta$ -lactam hydrolase-mediated hydrolysis MTX-cephalosporin-DEX was investigated. This system was successfully applied to study the evolution of antibiotic resistant class C  $\beta$ -lactamase[204]. Complementary to this, the same group used split-CIDs to investigate enzyme-mediated complementation (figure 12, B). The authors used glycosynthases to bind **DEX** linked to cellobioside and **MTX** to lactosyl fluoride[205]. In more recent studies, the **MTX** moiety was replaced with the eDHFR selective inhibitor **TMP** to reduce potential reactions with endogenous DHFR (**TMP-DEX**) (scheme 9, B)[205]. Likewise, Athavankar *et al.*, controlled gene expression of POIs *in vivo* with unmodified **biotin** as a CID[206]. This system relies on the co-expression of BirA, streptavidin-POI fusion protein and a

biotin acceptor peptide, fused to a second POI. BirA catalyses the acylation of a lysine in the biotin acceptor peptide, and **biotin** yields a biotinylated fusion protein, which then binds with high affinity to streptavidin and completes the dimerisation. Later, the group of Fussenegger used this system to design a gene repression system[207]. Therefore, the authors co-expressed tetracycline-dependent transactivator (tTA)-AviTag, BirA and streptavidin-KRAB (krueppel-associated box protein of human *kox-1* gene) in mammalian cells. In the absence of **biotin**, tTA-AviTag binds and activates tetracycline-responsive promoters, whereas biotin-mediated dimerisation of tTA-AviTag fusion and streptavidin-KRAB fusion, represses gene expression.



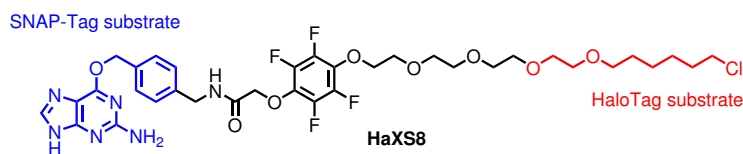
**Figure 12:** Principle of enzyme-mediated three-hybrid system. A) The heterodimeriser containing a substrate for the enzyme induces the reporter gene expression. In the presence of enzyme, the CID is cleaved and the expression stopped. B) The split CID is unable to induce reporter gene expression on its own. In the presence of enzyme, the CID is complemented and the reporter gene is expressed.

The group of Kikuchi, synthesised a heterodimeriser, biotin-6-aminohexanoylampicillin (**ampicillin-biotin**) (scheme 9, C) to dimerise  $E166N$ TEM-1 and streptavidin fusions[208]. In a first application, this group labelled  $POI-E166N$ TEM-1 with streptavidin containing super-paramagnetic iron oxide nanoparticles, which allowed detection of POI expression with  $^1H$ -MRI. The lab of Peterson synthesised a series of different heterodimerisers, fusing **biotin** to **DEX**, and the  $\beta$ -esteriol inhibitors **estrogen** or **estrone oxime** (scheme 9, E, F, G)[209, 210]. In a proof of concept, the authors successfully regulated the gene expression of a reported protein. Similarly, the group of Johnsson tethered **biotin** to **4-(Aminomethyl)-O<sup>6</sup>-benzyl guanine** (scheme 9, H)[211]. However, this CID was mainly used in phage displays to immobilise SNAP-tag fusion proteins. The same group also designed heterodimeriser systems based on the SNAP-tag and DHFR, connecting **4-(Aminomethyl)-O<sup>6</sup>-benzyl guanine** and **MTX** (scheme 9, I)[212]. This system rapidly induced dimerisation of DNA-bound LexA-SNAP-tag fusions and B42-DHFR fusions to trigger gene expression of a reporter protein. Etoc *et al.* developed a magnetogenetic approach to spatially induce signaling transduction of self-assembling proteins. In this approach, POI functionalised magnetic particles act as a hot spot for endogenous interacting partners[213]. To functionalise the magnetic particles, the authors used a heterodimeriser based on **biotin** and the HaloTag substrate (scheme 9, D). In proof-of-concept, the authors microinjected streptavidin containing magnetic particles into cells expressing TIAM1-HaloTag. Incubation of these cells with a biotinylated chlorohexane derivative recruited TIAM1-HaloTag to the magnetic particle. Subsequential translocation of the POI functionalised particles to the cell membrane with a magnetic force triggered TIAM1-Rac1 interaction, thus locally inducing a signaling response.



**Scheme 9:** Chemical structure of alternative heterodimerisers; A) **MTX-DEX**. B) **TMP-DEX**. C) **DEX-biotin**. D) **Halo-biotin**. E) **estrogen oxime-biotin**, F) **estrogen-biotin**. G) **BG-biotin**. H) **BG-MTX**.

Our group recently reported the first example of a completely covalent heterodimeriser, based on the SNAP-tag and HaloTag technologies[128]. Therefore, we covalently linked the SNAP-tag and the HaloTag-substrate with various linkers, giving rise to a series of CIDs named HaXS. To biologically validate these HaXS molecules, dimerisation of SNAP-GFP fusions and Halo-GFP fusions was performed in HeLa cells. The covalent nature of the dimerisation allowed simple SDS-page analysis. **HaXS8** (scheme 10), the best performing dimeriser in this study, bears a short tetrafluorophenyl moiety for improved cell permeability, as well as a small polyethylene glycol chain for better water solubility. In a proof of concept, we successfully translocated a protein of interest to different subcellular compartments. Further, **HaXS8**-mediated translocation of Halo-iSH2-GFP fusion protein to the plasma membrane, rapidly activated the PI3K/mTOR pathway. Although the described CIDs are powerful tools to regulate protein activity in living cells, they suffer from relative low temporal and spatial resolution based on their cell permeability and their diffusion inside the cells. To overcome such limitations, numerous light-activatable systems were developed to control protein functions *in vivo*.



**Scheme 10:** Chemical structure of the covalent protein dimerizer **HaXS8**.

### 1.3 Photo-activatable tools to regulate the function of proteins

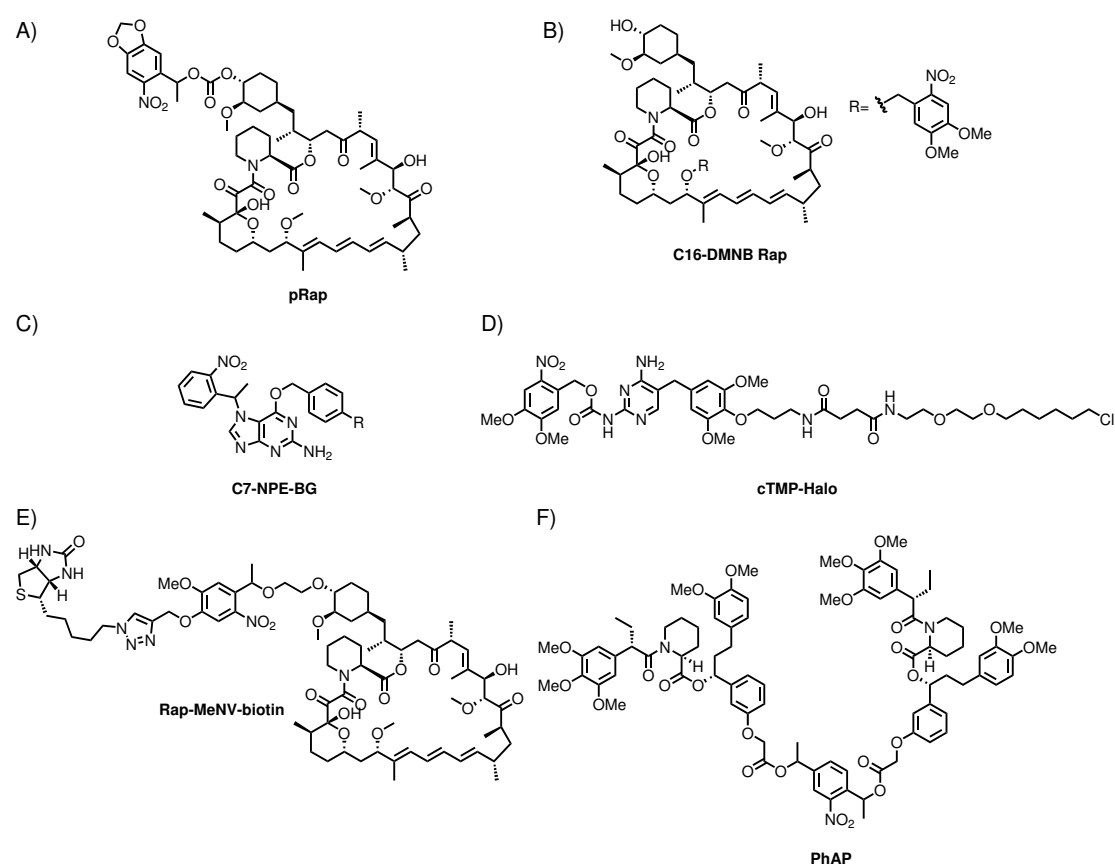
The concept of regulating biological processes with light was introduced into the field of chemical biology in the late 70's by the groups of Schlgler and Hoffman. In these early studies, biologically active compounds were deactivated or "caged" with photolabile protecting groups (PPGs). Since then, several approaches have been developed to manipulate protein function with light using protein tags and as a result, will be discussed in this section.

#### 1.3.1 Chemical tag based tools to modify proteins with light

Photosensitive protecting groups are most often applied to cage a bioactive molecule. Upon irradiation with light at an adequate wavelength, the cage is cleaved and the active compound is liberated. Until now, numerous examples of caged bioactive molecules have been reported, including important signaling molecules, several other small protein effector molecules and enzyme substrates[214]. Alternatively, protein activity can be regulated by directly caging the protein. Here, a caged unnatural amino acid is introduced at the active site of the protein, through unnatural amino acid site-directed mutagenesis and can then be expressed in cells using orthogonal synthase/tRNA pairs[21]. Cropp and co-workers for example, genetically encoded a caged cysteine into the TC-tag sequence, thereby dramatically decreasing the affinity of **FIAsH**. In fact, labelling of TC-tag fusion protein was not observed in *yeast* cells prior to irradiation[215]. The laboratory team of Hahn, introduced the photocleavable methyl-6-nitropiperonyloxycarbonyl group at the C40 position of rapamycin (**pRap**) (scheme 11, A)[186]. However, this photocaged rapalog was still able to induce dimerisation between FKBP and FRB. To overcome this limitation, the authors genetically engineered a mutant FKBP (iFKBP) to avoid interaction of tag with pRap prior to its uncaging. In contrast, Sadowski *et al.* caged the C16 position of **rapamycin** with a dimethyl nitrobenzyl group (**C16-DMNB Rap**) (scheme 11, B)[216]. However, cells were only incubated with **C16-DMNB Rap** after irradiation, hinting that **C16-DMNB rapamycin** might not be cell permeable or that photolysis in cells might be problematic. Similarly, Johnson and his group observed that the introduction of a photocleavable nitrophenyl group at the N7-position (**C7-NPE-BG**), but not at N9 of **BG**, disrupted SNAP-tag reactivity (scheme 11, C)[217]. Based on this observation, the authors successfully photoregulated fluorescent labelling, as well as homodimerisation of proteins, in cell lysate. Most recently, Ballister *et al.* developed a cell-permeable and photo-activatable CID system, based on a NVOC-caged TMP and the HaloTag substrate (**cTMP-Htag**) (scheme 11, D)[218]. In a proof of concept, the authors successfully recruited a POI protein to the centromere, kinetochores, centrosomes and mitochondria in HeLa cells, after the photocleavage of the protecting group.

In a complementary approach, the PPGs are incorporated into the core of molecules as photocleavable linkers. This approach was effectively used in several studies to selectively release bioactive molecules from nanoparticles, reversibly alter the cell permeability of molecules, degrade bioactive molecules, reversibly cross-link proteins, release proteins from specific locations, and deactivate proteins[214, 219]. Inoue and co-workers for example, introduced a photocleavable

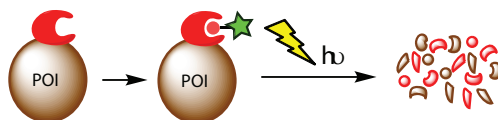
methyl-6-nitroveratryl group between a **biotin** and **He-rapamycin**, a C40 rapalog with similar properties as **rapamycin** (**Rap-MeNV-biotin**) (scheme 11, E)[185]. The high affinity of **biotin** towards avidin was used to arrest the designed rapalog extracellularly. Upon irradiation, **He-rapamycin** was released, penetrated the cell membrane and induced dimerisation inside the cell. Although, active **He-rapamycin** can be released with high spatial and temporal control, free diffusion of active **He-rapamycin** inside cells limits this approach. The lab of Williams synthesised a photocleavable homodimeriser based on the FKBP-**FK506** system[220]. The authors tethered two **SLF\*** with a photocleavable *o*-nitrobenzyl derivative (**PhAP**) (scheme 11, F) In proof of concept, they used **PhAP** in combination with a previously developed LC8<sub>TRAP</sub> protein to inhibit the interaction of LC8 with its endogenous ligands, ultimately leading to the dispersion of the endosome. Irradiation of cells with light at 365 nm for 10 minutes resulted in cleavage of the PhAP-LC8-LC8<sub>TRAP</sub> complex and gradual recovery of the endosomes to the perinuclear region. Although, these result demonstrate the high potential of cell-cleavable dimerisers, the relative long irradiation times required to induce photolysis may be critical.



**Scheme 11:** Photo-activatable CIDs and tag substrates. A) C40-caged rapalog **pRap**. B) C16-caged rapalog **C16-DMNB Rap**. C) C7-caged SNAP-tag substrate **C7-NPE-BG**. D) Caged **cTMP-Htag**. E) Photocleavable linker-containing rapalog **Rap-MeNV-biotin**. F) Photocleavable homodimerizer **PhAP**

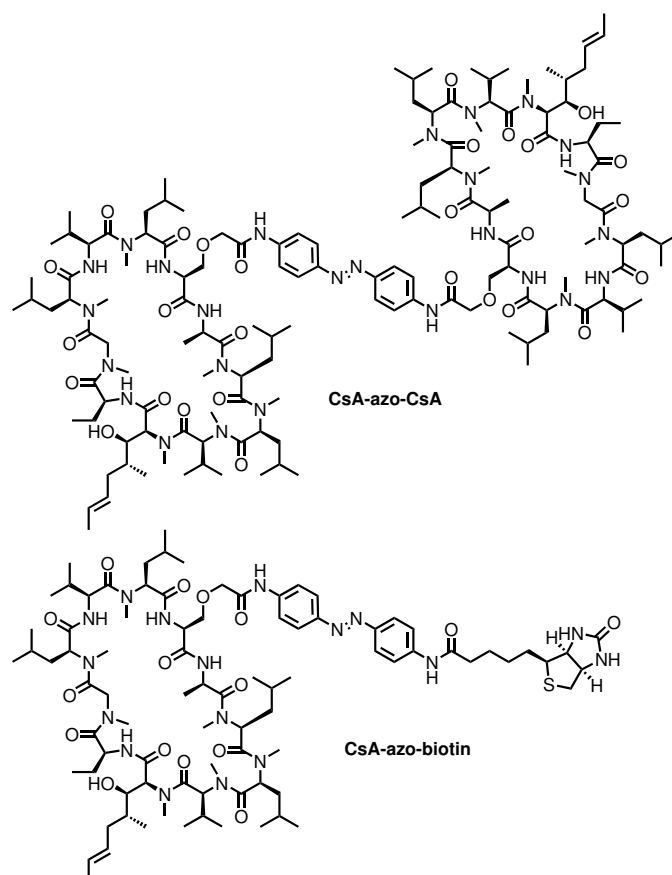


Incorporation of a photosensitive group into the backbone of a protein presents an attractive possibility to trigger proteolysis and thereby control the activity, lifetime and localisation of proteins. Therefore, several groups genetically or semi-synthetically introduced 2-nitrophenylglycine or 2-nitrophenylalanine into a protein[221–225]. This approach was successfully applied to regulate the activity of Shaker B  $K^+$  channel, nicotinic acetylcholine receptor, caspase-3, and Smad[221–223]. Chromophore-assisted light inactivation (CALI) is an alternative strategy to instantaneously down-regulate the activity of proteins with subcellular precision. In CALI, the reaction of an excited chromophore with oxygen is used to produce reactive oxygen species (ROS), which leads to the elimination of POIs (figure 13). To avoid cell damage and cell death caused by unspecific oxidation, ROS requires to be generated in close proximity to the POI. Originally, CALI used target-specific antibodies conjugated to the chromophore, Malachite Green, to selectively inactivate POIs[226]. Unfortunately, this system requires microinjection of the antibody-chromophore conjugate and the design of a unique antibody for each target. Additionally, Malachite Green is hydrophobic and tends to aggregate. Several years later, Surrey *et al.* used a fluorescein isothiocyanate (FITC) and GFP to address these limitations[227]. In a model experiment, a fluorescein labeled antibody inactivated  $\beta$ -galactosidase 50 times more effectively than the corresponding Malachite Green conjugate, whereas GFP only partially degraded the enzyme. In subsequent years, the most common chemical tags and corresponding fluorescent ligands were applied to target and inactivate POIs. **FIAsH** and **ReAsH** for example were used to regulate the polarity of cells, *in vivo*[228]. Similarly, Marks *et al.* inactivated  $\beta$ -galactosidase-FKBP12(F36V) fusion with light, in the presence of fluorescein labeled **SLF\***[229]. Although fluorescein **SLF\*** has no known endogenous targets, cells incubated with this ligand showed a considerable amount of off-target effects. To improve the labelling selectivity, the group of Ellenberg adopted the SNAP-tag system for CALI[230]. In a proof of concept,  $\alpha$ -tubulin and  $\gamma$ -tubulin, genetically fused to SNAP-tag, were expressed and labelled with **BG** diacetyl fluorescein in mitotic cells. Light-induced inactivation of  $\alpha$ -tubulin led to an arrest in metaphase, accompanied by a change in spindle morphology. In contrast, inactivation of centrosomal  $\gamma$ -tubulin disrupted nucleation of microtubules and impaired their growth. Takemoto *et al.*, demonstrated that eosin, another xanthene-based chromophore, exhibited a five-fold greater CALI efficiency and requires less intense irradiation compared to fluorescein[231]. Subsequently, an eosin haloalkane was administered to cells expressing HaloTag-POI fusion proteins. Regiospecific irradiation of cells containing the labelled fusion protein, rapidly led to POI inactivation. In this way, the authors regulated the translocation of PKC $\gamma$  to the cell membrane, cell survival and cell division. Although GFP does not generate ROS as effective as fluorescein, mutant variants with enhanced fluorescence (eGFP, eYFP etc.) are frequently used to image and inactivate proteins. Two additional protein tags, namely KillerRed and miniSOG, were applied in CALI experiments. In fact, KillerRed generates 1000-fold more ROS compared to eGFP and is the protein of choice for self-labelling CALI. The applications of CALI in combination with tags are versatile and are summarised in recent reviews[232–234].



**Figure 13:** Principle of chemical tag-based CALI. Tag-POI fusion proteins are labelled with a fluorophore containing tag substrate. Irradiation of the labelled tag rapidly leads to ROS induced fusion protein degradation.

In contrast to photoprotecting groups, photoswitchable groups allow reversible regulation of protein activity with small molecules. Most commonly used organic photoswitches, such as azobenzene, stilben, hemithioindigo, and spiropyran, undergo cis-trans isomerisation after irradiation. This geometrical alteration is accompanied with a more or less strong change in polarity, depending on the photoisomeriser. Ideally, these changes are severe enough to switch between an active and inactive state of the probe. However, none of the available organic photoswitches displays complete conversion from one isomer to the other. Nonetheless, photoswitchable compounds are becoming more and more popular in biology. Similar to photocleavable protecting groups, photoisomerisable groups can be used to reversibly alter the activity of a small effector molecule. In this manner, the activities of several protein classes were regulated, including enzymes, ligand-gated ion channels and G-protein-coupled receptors[235, 236]. Recently, Fischer and colleagues used a photoswitchable CsA derivative to reversibly regulate peptidyl prolyl cis-trans isomerase (PPIase)[237]. Interestingly, the authors introduced a second protein-binding moiety on the photochromic group to increase the activity difference of the two isoforms. Specifically, **biotin** or **CsA** were linked through an azobenzene to **CsA** (scheme 12). Light-mediated isomerisation from trans to cis of **CsA-azo-CsA** led to a seven-fold increase in inhibition of PPIase. Moreover, in cells transfected with streptavidin, the cis isomer of **CsA-azo-biotin** exhibited a twelve-fold higher suppressive activity than its trans isoform.



**Scheme 12:** Chemical structure of photoswitchable bi-functional small molecules **CsA-azo-CsA** and **CsA-azo-biotin**.

### 1.3.2 Optogenetic approaches to regulate protein-protein interactions

Optogenetics is a rapidly emerging method to control the activity of whole cells or specific proteins with light. In contrast to the techniques described above, optogenetics relies solely on genetically encoded photosensitive proteins. First applied in neuroscience, the concept was quickly adopted to regulate intracellular signaling[3, 214]. Therefore, several naturally occurring plant photoreceptors were re-designed to control PPIs in living cells. Upon irradiation, these optogenetic tags undergo chromophore-mediated conformational changes, which results in the binding of their effector domain. Some of the first studies in the field of optogenetic protein dimerisation used the phytochrome B (PhyB), the exogenous cofactor phycocyanobilin (PCB) and members of the phytochrome interacting factor (PIF) family. The light insensitive apo-PhyB autocatalytically ligates to the PCB chromophore, which is readily activated and rapidly binds to the PIF protein, after irradiation with light at 650 nm. This dimerisation can be reversed within seconds upon exposure to light at 750 nm. Quail and his group were the first to demonstrate the feasibility of this approach[238]. In a model experiment, light-mediated dimerisation of PhyB-DBD and PIF-TAD fusions reversibly induced the gene expression of a reporter protein. The research groups of Muir and Voigt adopted this approach to regulate protein splicing of POI[239] and to translocate the POI to membranes[240] with high spatial resolution. Light oxygen voltage (LOV) protein domain is another class of plant photoreceptor subdomain, frequently used to control PPIs. In response to blue light (440 - 473 nm), LOV domains bind to the endogenous chromophore, flavin, which either induces or disrupts the assembly of LOV, with its interacting partner. Kennedy *et al.* applied the LOV domain protein FKF1 and GIGANTEA, to control the activity of the small GTPase Rac1 and the transcription factor Gal4[241]. However, this system displays slow kinetics and the reaction is virtually irreversible. Similarly, the research groups of Yang and Gardner, used the homodimerising LOV domain, VVD, respectively EL222, to reversibly activate gene expression[242, 243]. In contrast, Wu *et al.*, fused the LOV2 domain to the carboxy-terminal of its binding partner, the helical extension ( $J\alpha$ ) and Rac1[244]. In the dark, LOV2 interacted with  $J\alpha$  adopting a sterically demanding closed conformation, which prevented the dimerisation of the Rac1 with its effectors. Irradiation of the fusion protein with blue light induced the formation of a covalent bond between LOV2 and flavin, thereby dissociating and unfolding the sterically inhibiting LOV2- $J\alpha$  complex and re-installing a Rac1-effector interaction. A third class of optogenetic system is based on Cryptochrome 2 (CRY2) and the N-terminal domain (CIBN) of its natural binding partner, CIB1 (Cryptochrome-interacting basic helix-loop-helix 1). Upon stimulation with blue light (405 - 488 nm), CRY2 binds to flavin; this in turn triggers the formation of a dimeric complex with CIBN, within a sub-second time scale. In the absence of light, previously excited CRY2 relaxes and the formed complex dissociates within a few minutes. In the proof of concept, Kennedy *et al.*, regulated the translocation of POI to the cell membrane, targeted gene expression and recombination of split-proteins[245]. Finally, the Lin group used a combination of two photoactivatable fluorescent Dronpa mutants (145K and 145N), to reversibly induce dimerisation. Notably, light-induced conformational change of the chromophore not only triggers dimerisation, but also alters the fluorescence of the proteins giving a direct read-out for successful dimerisation. In the monomeric state (dark state), Dronpa displays no green fluorescence. Illumination of this monomeric species with violet light (390 - 400 nm) results in the rapid formation of a green-fluorescent dimer, which can be reversed by irradiation with light at 490 nm. In an elegant application, the authors reversibly regulated the activity of intersectins and proteases with light. Therefore, Dronpa145K and Dronpa145N were fused to the N-terminus and C-terminus of POIs. Intramolecular dimerisation of Dronpa in cells resulted in reduced POI activity, which was completely restored after irradiation with light at 490 nm. Although all four approaches display excellent spatial and temporal control of protein

dimerisation, they suffer from several disadvantages, such as large protein tags, formation of unwanted homodimers and sensitivity to visible light[246].

## 2 Scope of the thesis

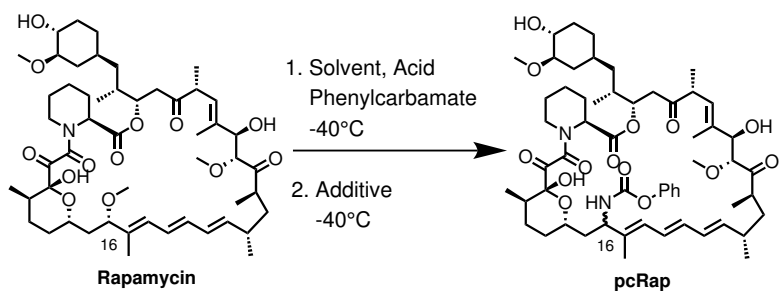
Chemical inducers of dimerisation are a powerful tool to specifically and independently regulate the activity of virtually any protein within a cellular context. However, the majority of the published CID systems, interfere with endogenous proteins or suffer from relatively low temporal resolution, requiring 10 or more minutes until significant amounts of dimerisers are formed. We therefore envisage to design novel CIDs, which enable the regulation of proteins with high spatial and temporal control, without interfering with endogenous proteins. Since its first application in 1996, CID **rapamycin** has become the most widely used heterodimeriser, due to its fast in vivo dimerisation kinetics, inducing dimerisation of FKBP and FRB fusion protein, on a timescale of seconds to minutes (manipulating signaling at will). However, the immunosuppressive activity of **rapamycin** may counteract the advantage of this system. Despite the efforts to synthesise rapamycin analogues with substituents at the FRB-binding site, small impurities of **rapamycin** or rapamycin byproducts in these C16 rapalogs, still interfere with mTORC1. Therefore, we plan to develop a simple and fast procedure to synthesise C16 rapalogs, which allows the reduction of the rapamycin content to a minimum, prior to the purification step. This procedure would enable the possibility to introduce a variety of substituents and proceed under mild conditions. Ultimately, to demonstrate the feasibility of this procedure, the synthesis of a C16 rapalog, which does not interfere with mTORC1 but induces dimerisation of FKBP, and an FRB mutant, was targeted.

In an alternative approach to regulate protein dimers or the location of proteins with high spatio-temporal control, the objective was the design and synthesis of cell permeable HaXS molecules, which rapidly trigger the formation of stable protein dimers and undergo photolysis upon irradiation. Such molecules would allow to regulate the dimerization and sequentially reverse the protein-protein interaction with two independent events. For cellular applications, these CIDs should possess rapid photolytic rates at wavelengths of  $> 300$  nm, without producing cytotoxic by-products. Therefore, the aim was to introduce different *o*-nitrobenzyl and C7 substituted coumarinyl-4-methyl derivatives into the core structure of the HaXS molecules, and to determine their photophysical properties. Ideally, the photophysical properties of two or more photocleavable HaXS CIDs are different enough, that they could be used simultaneously and cleaved independently. Additionally, we tried to synthesize a general photocleavable linker, which could be incorporated into different synthetic CIDs. Ultimately, the goal was to demonstrate the dimerization and subsequent cleavage of protein fusion proteins in a cellular context.

## 3 Results and Discussion

### 3.1 Efficient Synthesis of C16-carbamyl Rapalogs via Lewis Acid De-complexation

**Rapamycin** is a natural occurring CID, which binds sequentially to FKBP12 and only then to FRB, a subunit of mTORC1[125]. Thus, **rapamycin** can conditionally induce dimerisation of FKBP12-POI1 and FRB-POI2 fusion proteins[180]. Although **rapamycin** exhibits a very fast dimerisation rate, its interference with endogenous mTORC1 signaling, limits its application. Rapamycin analogues (so called rapalogs), which bear a substituent at the FRB-binding site, were reported to prevent cross-reaction with endogenous TORC1[181, 182], and allow these rapalogs only to interact with FRB domains with a compensatory small side chain mutation ("the hole"), to accommodate binding of the "bump" in rapamycin. The majority of these TORC1 safe rapalogs possess a bulky substituent at the C16 position of **rapamycin**. AP23102 is one of these non-immunosuppressive rapamycin derivatives, which bears a phenyl carbamate substituent at the C16 position. Although AP23102 was used in several biological studies to regulate the function of proteins[247, 248], no general method to introduce carbamates into **rapamycin** has been published to date. To develop a procedure to substitute the C16 methoxy group of rapamycin with carbamates, we decided to synthesise C16-phenyl carbamate rapamycin (**pcRap**), a C28 diastereomer of AP23102, as a model compound. Under acidic conditions, the C16-methoxy group of **rapamycin** undergoes a heterolytic cleavage, forming a relatively stable carbocation, which can be trapped with various nucleophiles[141, 181, 182]. Therefore, we treated solutions of **rapamycin** and phenyl carbamate (pc) with different acids. After basic work-up, we analysed the resulting crude mixtures with MALDI-TOF. Attempts to carry out the nucleophilic addition of pc to C16-rapamycin carbocation by using TFA[181] (table 1, entry 1) and *p*-toluenesulfonic acid[149] (table1, entry 2) in CH<sub>2</sub>Cl<sub>2</sub> at -40°C, gave a mixture of products containing only traces of **pcRap**. It is likely that **pcRap**, as well as pc itself, were not stable in the presence of protic acids, such as TFA and *p*-TsOH. We next applied the Lewis acid BF<sub>3</sub>-Et<sub>2</sub>O to generate the carbocation at the C 16 position[181]. Addition of the Lewis acid to a solution of Rap and pc in CH<sub>2</sub>Cl<sub>2</sub> at -40°C led to a new mixture of inseparable products containing starting material, the product of methoxy-elimination[181] and traces of **pcRap** (table 1, entry 3). These results and the observation of a dark-red colour change after addition of BF<sub>3</sub>-Et<sub>2</sub>O, suggest that the reactive carbocation intermediate is formed under these conditions, but the nucleophilic substitution is impaired. To increase the stability of the transient carbocation and increase the yield of the reaction, we used THF as the solvent instead of CH<sub>2</sub>Cl<sub>2</sub>. Surprisingly, no colour change was observed after the addition of BF<sub>3</sub>-Et<sub>2</sub>O and only rapamacin and pc were recovered, indicating that the transient carbocation complex is not formed in THF (table 1, entry 4). To generate the carbocation and subsequently stabilise it, we treated a solution of pc and rapamycin in CH<sub>2</sub>Cl<sub>2</sub> at -40°C with BF<sub>3</sub>-Et<sub>2</sub>O, followed five minutes later, by the addition of an excess of THF. MS- analysis of the crude mixture revealed formation of large amounts of **pcRap**, and starting materials (**pcRap**/**rapamycin** approximately 70:30). This mixture of compounds was re-dissolved in CH<sub>2</sub>Cl<sub>2</sub> and the resulting solution re-treated with BF<sub>3</sub>-Et<sub>2</sub>O and THF, to increase the formation of product and minimise the **rapamycin** content. After repeating this procedure an average of 3 to 4 times, the crude mixture contained mainly **pcRap** with small traces of rapamycin and elimination products. Column chromatography of the crude product afforded a diastereomeric mixture of **pcRap** with an 83% yield (table 1, entry 5). Similar results were obtained, when less or more polar ethers, such as Et<sub>2</sub>O or 1,4-dioxane were used as an additive (table 1, entries 6 and 7).

**Table 1:** Optimization of Lewis Acid-Mediated Synthesis of **pcRap**.


Entry	Solvent	Additive	Acid	Yield[%]
1	CH <sub>2</sub> Cl <sub>2</sub>	-	TFA <sup>[a]</sup>	Traces
2	CH <sub>2</sub> Cl <sub>2</sub>	-	<i>p</i> -TsOH <sup>[b]</sup>	Traces
3	CH <sub>2</sub> Cl <sub>2</sub>	-	BF <sub>3</sub> -Et <sub>2</sub> O <sup>[c]</sup>	Traces
4	THF	-	BF <sub>3</sub> -Et <sub>2</sub> O <sup>[c]</sup>	-
5	CH <sub>2</sub> Cl <sub>2</sub> <sup>[d]</sup>	THF	BF <sub>3</sub> -Et <sub>2</sub> O <sup>[c]</sup>	83
6	CH <sub>2</sub> Cl <sub>2</sub> <sup>[d]</sup>	Dioxane	BF <sub>3</sub> -Et <sub>2</sub> O <sup>[c]</sup>	76
7	CH <sub>2</sub> Cl <sub>2</sub> <sup>[d]</sup>	Et <sub>2</sub> O	BF <sub>3</sub> -Et <sub>2</sub> O <sup>[c]</sup>	79

**Table 1:** [a] pc (6 equiv.), TFA (10 equiv.); [b] pc (6 equiv.), *p*-TsOH (10 equiv.); [c] pc (6 equiv.), BF<sub>3</sub>-Et<sub>2</sub>O (4 equiv.); [d] 5 min.; [e] 20 min.

To validate the generality of this procedure, we extended the scope of the study by introducing a series of representative carbamates into rapamycin. Using benzyl carbamate, C16-benzylcarbamate-Rap (**bnRap**) was synthesised to provide an excellent yield (88%, table 2, entry 1). Reactions with the lower electron-donating and unhindered aliphatic *n*-butyl carbamate, gave C16-*n*-butylcarbamate-Rap (**buRap**) a yield of 73% (table 2, entry 2). Interestingly, comparable results were still achieved with the acid-sensitive, table 2, entry 3). We assumed that the compound **buRap** could be applied as an extended selective modification of the rapamycin core, via the C16-Boc-protected amine. Other nucleophiles, such as phenyl urea, phenyl thiourea, phenol or 3-methylindole, were also readily introduced into the C16-position under the described reaction conditions (table 2, entries 5-8). These products were only analysed by MALDI-TOF, since we were not able to separate the two diastereomers. Interestingly, 3-methylindole also reacted with rapamycin in the absence of THF additive, however to a much lesser extent. This result suggests that the nature of the nucleophile has an influence on the reaction outcome, which is in contrast to common unimolecular nucleophilic substitution reactions.

**Table 2:** BF<sub>3</sub>-Et<sub>2</sub>O/THF-mediated synthesis of C16 rapalogs.

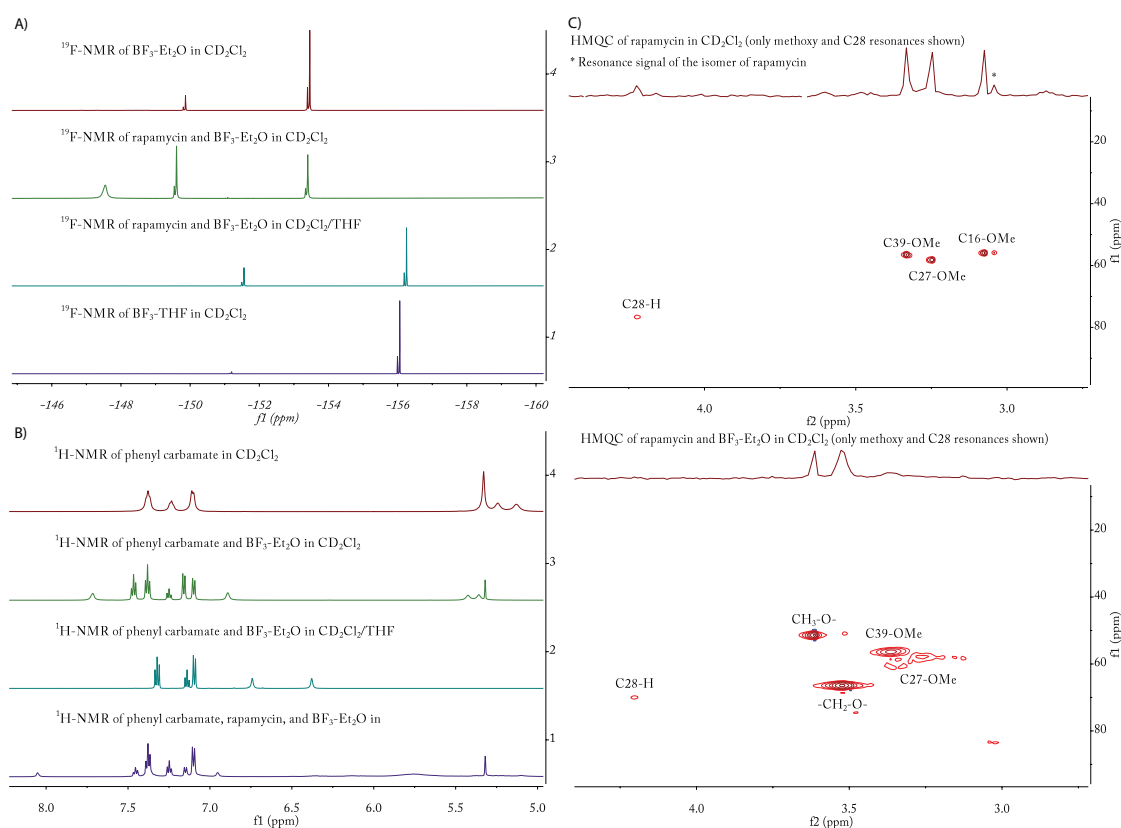
Entry	Nucleophile	Product	Yield[%] <sup>[a]</sup>	C16 ( <i>R</i> )/( <i>S</i> ) ratio
1	phenyl carbamate	<b>pcRap</b>	83	49:51
2	benzyl carbamate	<b>bnRap</b>	88	49:51
3	butyl carbamate	<b>buRap</b>	73	51:49
4	<i>ter</i> -butyl carbamate	<b><sup>t</sup>buRap</b>	78	50:50
5	phenylurea <sup>[b]</sup>	<b>puRap</b>	74	-
6	phenylthiourea <sup>[b]</sup>	<b>ptRap</b>	68	-
7	phenol <sup>[b]</sup>	<b>pRap</b>	81	-
8	3-methylindole <sup>[b]</sup>	<b>iRap</b>	83	-

**Table 2:** Indicated nucleophile (6 equiv.), CH<sub>2</sub>Cl<sub>2</sub>, -40°C, BF<sub>3</sub>-Et<sub>2</sub>O (4 equiv.), 5 min, THF, 20 min, -40°C. [a] Isolated yield of the diastereomeric mixture [b] products were only analyzed by MALDI-TOF

To elucidate the exact role of the additive and to determine the mechanism of the reaction of rapamycin with carbamates, we performed low temperature NMR studies. Therefore, we recorded and compared the <sup>1</sup>H-, <sup>11</sup>B- and <sup>19</sup>F-NMR spectrum of mixtures of BF<sub>3</sub>-Et<sub>2</sub>O, rapamycin, and phenyl carbamate, respectively, in CD<sub>2</sub>Cl<sub>2</sub> or CD<sub>2</sub>Cl<sub>2</sub>/THF with the corresponding homomolecular solutions (Figure 13, A and B). In the absence of phenyl carbamate, addition of BF<sub>3</sub>-Et<sub>2</sub>O to a solution of rapamycin in CD<sub>2</sub>Cl<sub>2</sub> resulted in the characteristic colour change of the solution. Comparing the <sup>19</sup>F-NMR spectra of this solution with the spectra of BF<sub>3</sub>-Et<sub>2</sub>O in the same solvent, exhibited a new broad signal, suggesting formation of a BF<sub>3</sub>-coordinated rapamycin intermediate (Scheme 13). This complex readily disassociated after the addition of an excess of THF, indicated by the disappearance of the broad signal and formation of BF<sub>3</sub>-THF. Since CH<sub>2</sub>Cl<sub>2</sub> and THF have similar dielectric constants, we assumed that the observed stabilisation of the carbocation is an effect of the coordinating ability of the additive rather than its polarity. This hypothesis is supported by the fact that the product **pcRap** is also formed when using less polar additives, such as diethyl ether. To determine the coordination site of BF<sub>3</sub> we next recorded the HMQC spectrum of rapamycin in the presence and absence of the Lewis acid (figure 14, C). As expected, the correlation peak of the C16 methoxy group disappeared after addition of BF<sub>3</sub>-Et<sub>2</sub>O, demonstrating that this position is the most favoured reaction site.

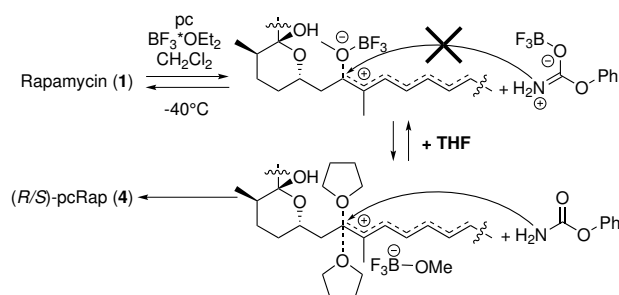
Although BF<sub>3</sub>-Et<sub>2</sub>O selectively reacts with the C16 methoxy group, irreversible formation of a free or solvent-separated carbocation is unlikely, as no nucleophilic substitution is observed under these conditions. More likely, a BF<sub>3</sub>-methanolate complex forms a tight and stable ion pair with the rapamycin ion. Addition of THF leads to the sufficient stabilisation of the carbocation, the subsequent formation of a solvent-separated ion pair and finally to a fast reaction with the nucleophile. In contrast, using THF as the solvent in the first step prevents the BF<sub>3</sub> from interacting with the methoxy group at C16, avoiding the formation of rapamycin carbocation (scheme 13). We next investigated the effect of the Lewis acid on phenyl carbamate. Analysis of the proton and fluorine NMR spectra of BF<sub>3</sub>-Et<sub>2</sub>O with phenyl carbamate in CD<sub>2</sub>Cl<sub>2</sub> revealed





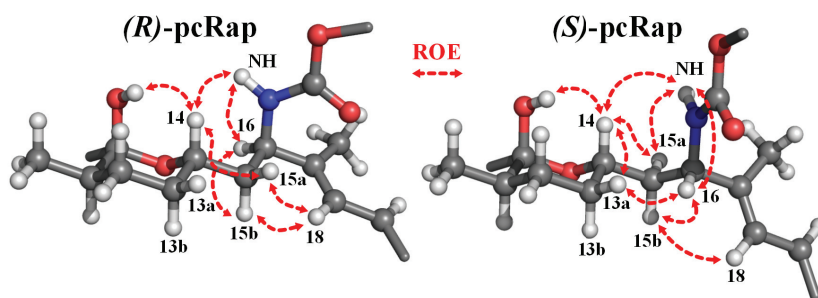
**Figure 14:** NMR study to elucidate the mechanism of the  $\text{BF}_3$ -mediated synthesis of pcRap. A)  $^{19}\text{F}$ -NMR spectra of the indicated mixtures were recorded at  $-30^\circ\text{C}$ ; 1.3 mg **rapamycin**, 0.74  $\mu\text{L}$   $\text{BF}_3\cdot\text{Et}_2\text{O}$  in 0.4  $\mu\text{L}$   $\text{CD}_2\text{Cl}_2$  respectively 0.4  $\mu\text{L}$   $\text{CD}_2\text{Cl}_2/\text{THF}$  (2:1) B)  $^1\text{H}$ -NMR spectra of indicated mixture at  $-30^\circ\text{C}$ ; 1.2 mg phenyl carbamate, 1.3 mg **rapamycin**, 0.74  $\mu\text{L}$   $\text{BF}_3\cdot\text{Et}_2\text{O}$  in 0.4  $\mu\text{L}$   $\text{CD}_2\text{Cl}_2$  respectively 0.4  $\mu\text{L}$   $\text{CD}_2\text{Cl}_2/\text{THF}$  (2:1). C) HMQC of 12 mg rapamycin in 0.4  $\mu\text{L}$   $\text{CD}_2\text{Cl}_2$  in the presence (down) or absence (up) of 0.74  $\mu\text{L}$   $\text{BF}_3\cdot\text{Et}_2\text{O}$  at  $-30^\circ\text{C}$ .

signals for two different carbamate species, as well as two different  $\text{BF}_3$  containing complexes. These results implicate that  $\text{BF}_3$  also coordinates to phenyl carbamate, presumably forming a boro-iminium complex. Moreover, NOESY experiments of phenyl carbamate and  $\text{BF}_3\cdot\text{Et}_2\text{O}$  revealed that the two phenyl carbamate species exist in a slow exchanging equilibrium. Similar to rapamycin- $\text{BF}_3$  complex, the addition of THF readily disrupted the boron-phenyl carbamate complex (scheme 13).



**Scheme 13:** Proposed mechanism for the Lewis acid-mediated formation of pcRap. After addition of  $\text{BF}_3 \cdot \text{Et}_2\text{O}$  to a mixture of rapamycin and phenyl carbamate, a  $\text{BF}_3$ -carbocation and a  $\text{BF}_3$ -iminiumion complex are formed, which are not able to further react. Addition of THF stabilizes the carbocation and the reaction with the liberated phenyl carbamate yield pcRap.

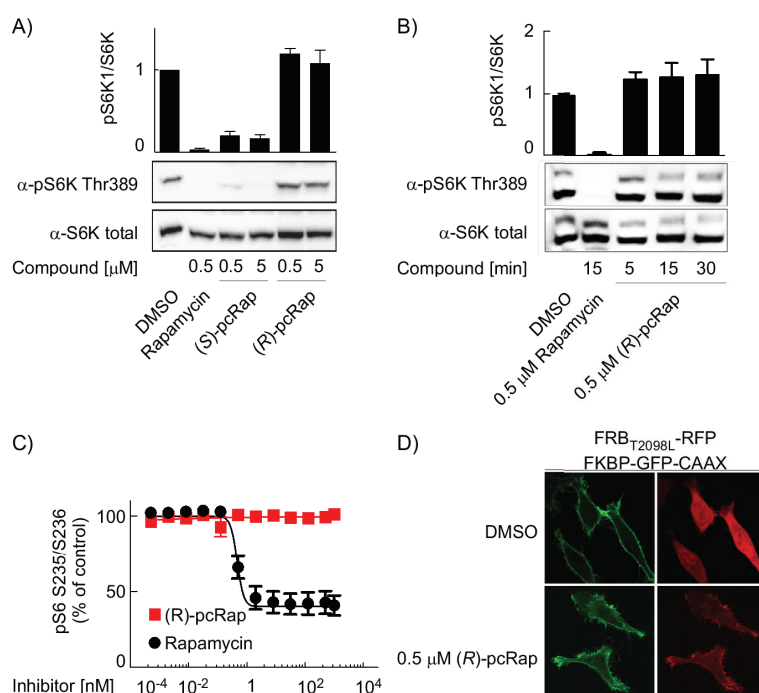
Previous studies on rapamycin derivatives demonstrated that the stereochemistry of the substitution at C16 can influence the binding affinity for wild-type and mutated FRB[141, 181]. Therefore, we expected that each **pcRap** diastereomer might possess a stereospecific profile towards the possible binding partners. To verify this hypothesis, we first separated the two diastereomers by column chromatography, assigned the stereochemistry of (*R*)-**pcRap** and (*S*)-**pcRap** via ROESY experiments and then validated their potential to interfere with endogenous TORC1. Correlations between the protons at the cyclic hemiacetal, at C15 and C16, as well as at the NH proton of the carbamate function, were used to determine the absolute configuration of the diastereoisomers. The most important ROE correlations, that were used to determine the absolute configuration, are shown in figure 15. Only the (*S*)-configuration at C16 could explain the observed correlation between H16 and H13a, and without violating any of the other correlations. In contrast, only the (*R*)-conformation could explain the observed correlation between the H15a and H18, and the correlation between the NH proton and H15a.



**Figure 15:** Key ROESY correlations to elucidate the absolute conformation of pcRap diastereomers. (*R*)-**pcRap** (d.r. 51%), left, and (*S*)-**pcRap** (d.r. 49%), right.

One of the major challenges of synthesising non-cytotoxic rapalogs is the purification phase, since minimal contamination of rapamycin or other FRB wild-type-binding rapamycin derivatives cause extensive TORC1 inhibition. To minimise the risk of such contamination, we further purified the two diastereomers using preparative HPLC (for an example of the UV-Vis spectra of HPLC separated reaction mixture see figure S1). Finally, we assayed the two **pcRap** isomers for their ability to inhibit TORC1 in cells (figure 16, A). Therefore, we treated HEK cells with either (*S*)-**pcRap**, (*R*)-**pcRap** or rapamycin for 15 minutes, lysed the cells and monitored the

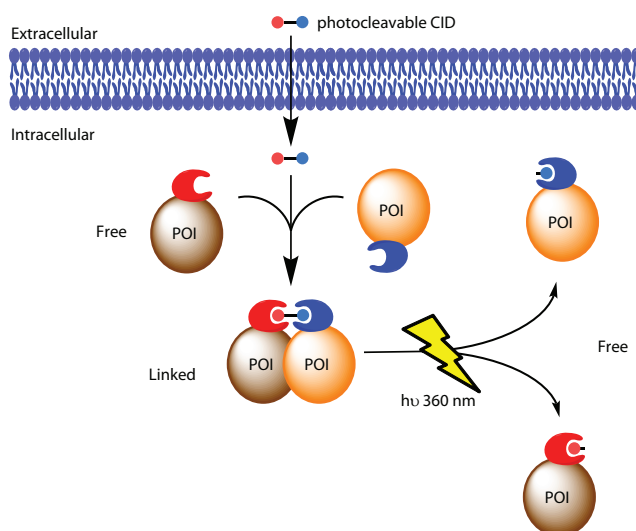
phosphorylation of Thr389 on the ribosomal protein S6 kinase (S6K1), a downstream target of TORC1. Incubation of HEK cells with rapamycin or **(S)-pcRap** resulted in a rapid inhibition of the TORC1, indicated by non-phosphorylated S6K1. However, it is presently not entirely clear if co-eluting contaminations of **rapamycin** and rapamycin by-products or **(S)-pcRap** impaired mTOR activity. In contrast, the phosphorylation of S6K1 was intact in cells treated with **(R)-pcRap**. Even after increasing the concentration of **(R)-pcRap** by a factor of ten, or increasing the incubation time to 30 minutes (figure 16, B), no inhibition of TORC1 was observed. Similarly, **(R)-pcRap** did not inhibit the phosphorylation of S6K1 in A2058 cells, after the cells were treated for 1 hour with the CID (figure 16, C). The elongated incubation time was necessary, as rapamycin did not down-regulate TORC1 after 15 minutes. However, under these conditions the IC<sub>50</sub> value of **(R)-pcRap** was measured to be above 1  $\mu$ M. In comparison, we measured an IC<sub>50</sub> value of 0.47 nM for rapamycin, which is comparable to values reported in the literature[249]. Although, we were able to synthesis pure and mTOR-safe **(R)-pcRap**, only small amounts (yield < 5%) were obtained after several rounds of purification by preparative HPLC (for an example of the UV-Vis spectra of purified **(R)-pcRap** see figure S2). . The majority of **(R)-pcRap** still contained small amounts of impurities, which inhibited mTOR and require further purification. To demonstrate the utility of the TORC1-safe **(R)-pcRap** isoform as a heterodimeriser, we investigated the CID-induced translocation of a cytosolic red fluorescent protein (monomeric RFP; TagRFP) to the cell plasma membrane (figure 16, D). Therefore, we co-transfected Hela cells with an FRB-RFP fusion protein and enabled the plasma membrane anchoring of an FKBP12-GFP-CAAX fusion. The addition of 0.5  $\mu$ M **(R)-pcRap** resulted in rapid translocation of the chimeric TagRFP-T<sup>2098L</sup>FRB protein to the cytosolic surface of the plasma membrane. In contrast, no translocation was observed in untreated transfected cells. These results demonstrate that **(R)-pcRap** is a powerful, non-toxic tool to chemically regulate the assembly and location of proteins of interest.



**Figure 16:** Biological evaluation of (*R*)-pcRap and (*S*)-pcRap. A) HEK cells grown in fetal calf serum supplemented media were exposed to DMSO, rapamycin, (*R*)-pcRap and (*S*)-pcRap for 15 min at 37°C, before cells were lysed and proteins were subjected to SDS PAGE, and immune-blotting using antibodies against total p70S6K (total S6K), and phosphorylated p70S6K (pS6K, Thr389). Data represent means SEM, n=3; difference from DMSO control; B) HEK cells grown in fetal calf serum supplemented media were exposed to DMSO, rapamycin, and (*R*)-pcRap for indicated time at 37°C. Samples were analysed as described in A); C) A2058 cells were treated with indicated amount of rapamycin or (*R*)-pcRap for 1 hour. Samples were analyzed as described in A); D) HeLa cells expressing the FKBP12-GFP-CAAX membrane anchor [green (-CAAX is the polybasic isoprenylation sequence from KRas-4B)] and FRB<sup>T2098L</sup>-RFP (red) fusion proteins were exposed to DMSO or 0.5 mM (*R*)-pcRap for 15 min. at 37°C. Translocation of FRB<sup>T2098L</sup>-RFP to the plasma membrane was imaged by confocal microscopy in fixed cells.

### 3.2 Synthesis and evaluation of cell permeable and photocleavable HaXS derivatives

Although common CIDs have been extensively used in numerous studies, many biological processes occur with spatiotemporal precision beyond the resolution of most common CIDs[3]. The concept of using light to alter the biological activity of CIDs, provides a non-invasive method to regulate biological processes in vivo with the required resolution to study fast occurring biological processes[219]. As described in the introduction, different caged CIDs were developed and used in proof of concept experiments. These caged CIDs do not induce dimerisation of the targeted proteins immediately, instead they require light-induced cleavage of the cage prior to triggering protein dimerisation[219]. In a complementary approach, we envisaged to design CIDs, which contain the photosensitive group in the core of the molecule, covalently linking the two tag substrates together. Photocleavable CIDs would allow to firstly trigger the dimerisation of POIs and then secondly, to irreversibly split the formed protein complex in a highly controlled way (figure 17). The possibility to regulate protein dimerisation by two independent events can be used for different applications. For example, POIs can be anchored away from their active site, followed by subsequent irradiation leading to liberation and relocation of the POI, thereby restoring its function. Other scenarios include the activation and controlled deactivation of metabolic pathways or *vice versa*, kinetic studies of protein trafficking, investigation of cell-compartment-associated signaling, and the simulation of cell-wide physiological and pathological signaling dynamics.

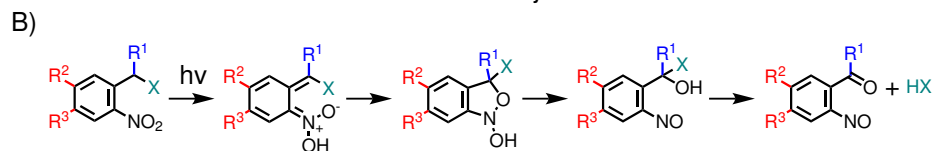
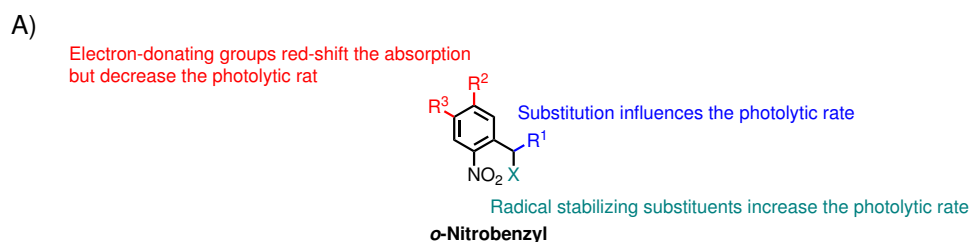


**Figure 17:** Principle of photocleavable CID. Treatment of cells with a photocleavable, and cell-permeable CID leads to dimerization of tag-POI fusion proteins. Illumination of CID cleaves the link between the POIs, and releases them from the covalent complex.

Based on the results of our previous study, the development of cell permeable HaloTag and SNAP-tag homodimerisers (HaXS)[128], we decided to introduce the photocleavable moiety into the core module, tethering the HaloTag-reactive chloroalkane ligand and the SNAP-tag-reactive **BG**. Synthetically, the modular synthesis enables a straight forward approach to build HaXS derivatives with different properties using a minimal number of steps. Biologically, HaXS molecules show excellent selectivity, good cell permeability and fast dimerisation. Additionally, the covalent and irreversible nature of these CIDs ensure that the formed protein complex is stable, prior to irradiation and thus allows simple analysis of the formed protein dimers to be carried out by Western blot.

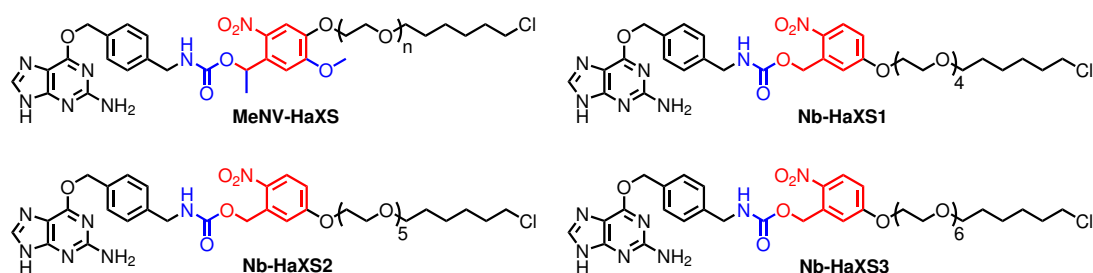
### 3.2.1 Selection of the photolabile groups

Photolabile protecting groups (PPGs) are widely used in chemistry as protecting groups or photolabile linkers in solid-phase syntheses. Due to their big impact in synthetic chemistry, numerous classes of photolabile groups have been developed to date. However only a handful display the properties required for biological applications. The PPG should (I) be stable towards hydrolysis, (II) absorb light at wavelengths higher than 300 nm to prevent cell damage, and (III) display high quantum yields for efficient photolysis. Additionally, the intermediates and products that are formed should not be cytotoxic or absorb the emitted light[214, 219, 250]. *o*-Nitrobenzyl, and coumarinyl-4-methyl derivatives are two of the classes of PPGs, which exhibit all or most of the necessary properties and were applied in several biological studies[219, 250]. Therefore, we decided to introduce these two PPGs into our HaXS molecules. The *o*-nitrobenzyl group and its many derivatives (scheme 14, A) are the most frequently used PPGs in biology[219]. Mechanistically, the absorption of a single electron leads to the formation of an excited nitro group ( $n$  to  $\pi^*$  electron transition), which abstracts a hydrogen from the *o*-alkyl substituent. Subsequently, the aci-nitro intermediate is formed, which is followed by intramolecular cyclisation and release of a nitrosoaldehyde and the leaving group (scheme 14, B)[250]. The photophysical properties of *o*-nitrobenzyl groups can easily be fine-tuned, introducing substituents at the benzylic position or at the aromatic ring, as well as changing the leaving group (Figure 14, A). For example, the *o*-nitrophenylethyl group, which bears a simple methyl group in  $\alpha$ -position displays increased quantum yields[251]. Introduction of a strong electron-withdrawing group, such as  $\text{CF}_3$ ,  $\text{CBr}_3$ , or  $\text{CN}$  at this position further improves the photolysis rate[250]. However, synthesis of these electron-withdrawing compounds remains challenging[250]. Modification of the aromatic region influences the absorption and consequently the quantum yield. Introduction of electron donating groups in meta and/or para position(s) to the nitro group shift the absorption maximum to longer wavelengths[250, 252, 253]. Despite having a bathochromic effect, these modifications also decrease the quantum yield of the photolytic reaction[250]. Beside benzylic or aromatic substituents, the leaving group also has an influence on the quantum yield. Bochet and co-workers demonstrated that the quantum yield of the photolysis correlates with the radical stabilisation properties of the leaving group (quantum yields decrease in the series alcohol > amine > N-carbamate > amide > O-carbamate > ester > carbonate [254]). Although alcohols and amines exhibit the highest quantum yields for the photolysis after irradiation with light at higher wavelength, they are generally protected as carbonates or carbamates, due to more convenient synthetic procedures.



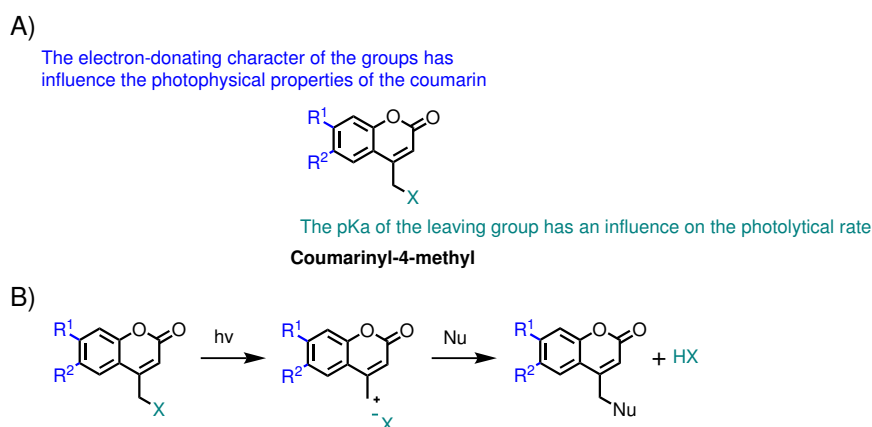
**Scheme 14:** Theoretical evaluation of *o*-nitrobenzyl based PPGs. A) Influence of the leaving group, aromatic and benzylic substitution on the photophysical properties of *o*-nitrobenzyl group. B) Mechanism of the photolytic decomposition of *o*-nitrobenzyl based PPGs.

Based on these considerations, we decided to incorporate the methyl-6-nitroveratryl (MeNV) moiety into the HaXS molecule via carbamate formation (**MeNV-HaXS**) (scheme 15). The methyl group of **MeNV-HaXS** should increase the photolytic rate, whereas the methoxy group should red-shift the absorption without decreasing the quantum yield too much. However, to evaluate the protein dimerisation properties of nitrobenzyl containing HaXS, we decided to first synthesise model compounds, which contain the commercially available 5-hydroxynitrobenzyl alcohol. Additionally, we envisaged to vary the number of PEG units, linking the HaloTag reactive chloroalkane and the PPG. Increasing the number of PEG units should improve the water solubility of the CID, reduce the chance of steric clashes between dimerised proteins, but may decrease the cell permeability. To find the optimal number of PEG units, we planned to synthesise three Nb-HaXS (scheme 15) molecules with either four, five or six PEG available in three steps, starting from the 5-hydroxynitrobenzyl alcohol, **BG-NH<sub>2</sub>** and **Halo-PEG<sub>n</sub>-Br** building block.



**Scheme 15:** Chemical structure of photocleavable HaXS molecules based on *o*-nitrobenzyl derivatives: **MeNV-HaXS**, **Nb-HaXS1**, **Nb-HaXS2**, and **Nb-HaXS3**.

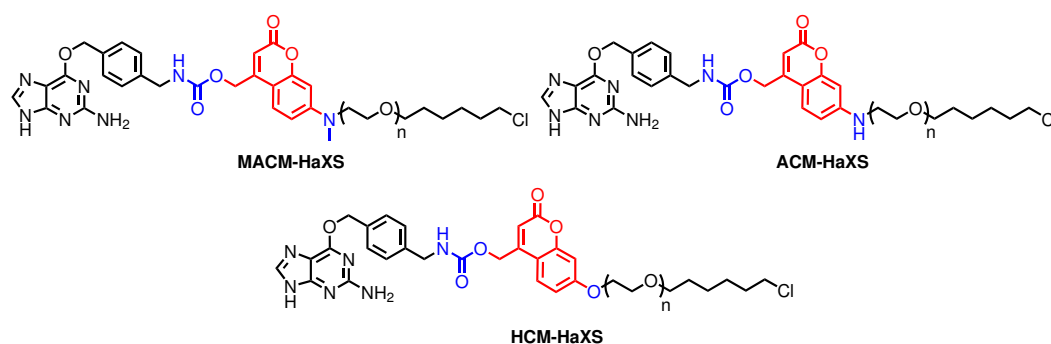
The second class of PPGs which we planned to incorporate into our HaXS, are based on coumarinyl-4-methyl group (scheme 16, A). These PPGs possess higher extinction coefficients, fast photolysis rates, fluorescent properties and are well suited for two-photon excitation[250]. After excitation with light, the electron from the  $\pi$  transitions to the  $\pi^*$  orbital and the 7-methoxycoumarinyl-4-methyl derivative relaxes via fluorescent emission, nonradiative processes or heterolytic C-X bond cleavage. The latter results in the formation of a tight ion pair. The coumarinylmethyl cation rapidly reacts with close-by nucleophiles or solvent molecules to generate a new stable product (scheme 16, B)[250]. Similar to *o*-nitrobenzyl group based PPGs, the photophysical properties of coumarinyl-4-methyl PPGs can be fine-tuned by substitution and selection of the leaving group (scheme 16, A). Introduction of electron-donating groups at the C7 and C6 position significantly improves the quantum yield of the photolysis, as well as shifts the absorption maximum to higher wavelengths[255]. The 7-methoxycoumarinyl-4-methyl derivatives for example (328 nm) displayed a  $\approx 20$  nm red-shifted compared to the parent molecule (310 nm)[250]. Amino substitution of C7 further shifts the absorption maxima to 350 - 400 nm. Additionally, these coumarinyl derivatives exhibited excellent quantum yields, ranging from 0.21 to 0.28[250]. The nature of the leaving group has a dramatic influence on the photolysis rate. For example, good leaving groups readily undergo photolysis and additionally minimise the possibility of ion pair recombination. In contrast, weak leaving groups such as alcohols, amines, phenols and thiols do not undergo photo-induced heterolysis. These groups are best protected as the corresponding carbonate, carbamate or thiocarbonate[250].



**Scheme 16:** Theoretical evaluation of coumarinyl-4-methyl based PPGs. A) Influence of the leaving group and substitutions on the photophysical properties of coumarinyl-4-methyl B) Mechanism of the photolytic decomposition of coumarinyl-4-methyl based PPGs.

Therefore, we planned to synthesise **ACM-HaXS** and **MACM-HaXS**, which contain a secondary or a tertiary amino substitution at C7 and a carbamate as the leaving group (scheme 17). However, to validate if coumarin containing HaXS still induce dimerisation of SNAP-tag and HaloTag fused proteins in vivo, we decided to synthesise the more readily accessible 7-alkoxycoumarinyl-4-methylhydroxyl derivative (**HCM-HaXS**) (scheme 17).

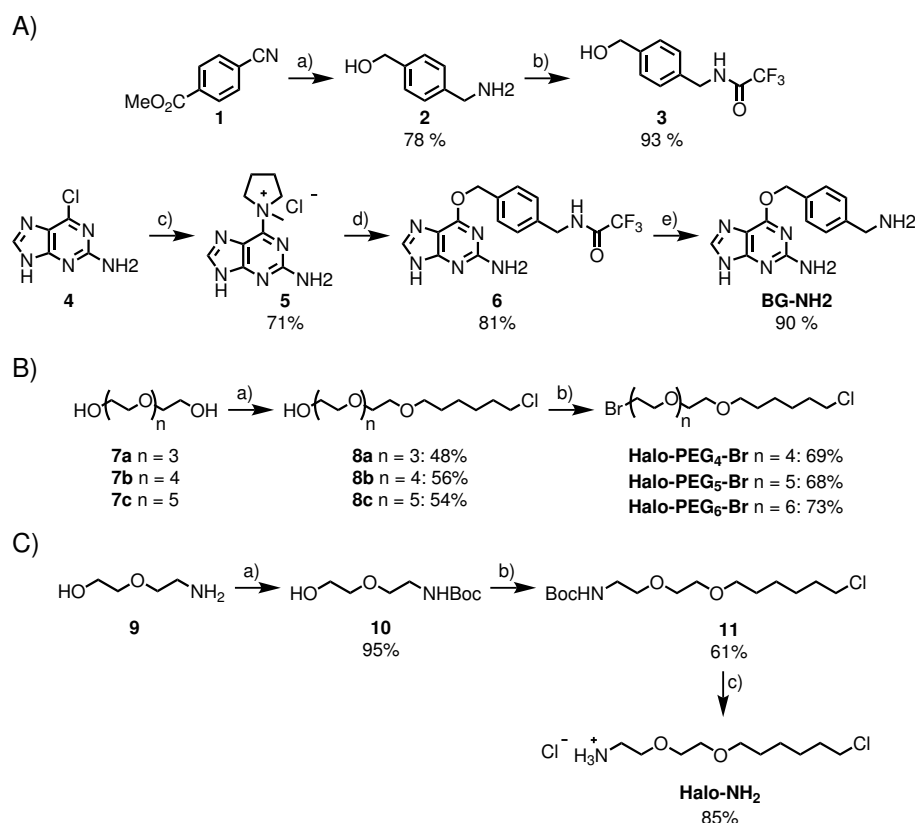




**Scheme 17:** Chemical structure of photocleavable HaXS molecules based on coumarinylmethyl derivatives: **MACM-HaXS**, **ACM-HaXS**, and **HCM-HaXS**.

### 3.2.2 Synthesis of SNAP-tag and HaloTag substrate building blocks

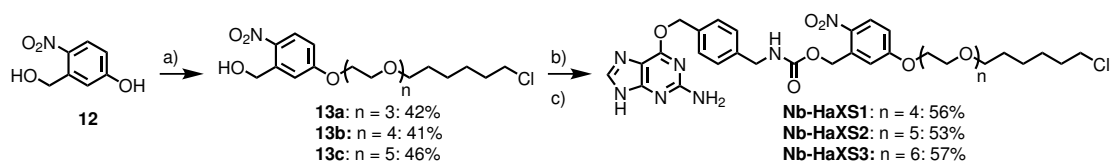
**BG-NH<sub>2</sub>** was prepared according to the procedure published by the group of Johnsson[48, 121] (scheme 18, A). The synthesis of **BG-NH<sub>2</sub>** was initiated with the reduction of methyl-4-cyanobenzoate (**1**) using an excess of LiAlH<sub>4</sub> giving 4-(aminomethyl)-benzyl alcohol (**2**) in acceptable yield. Subsequently, the amino group of the benzyl alcohol was protected with ethyl trifluoroacetate under mild basic conditions to yield **3**. The activated guanine derivative **5** was obtained after the reaction of 6-chloro-guanine with 1-methyl-pyrrolidine in DMF, which was further used with **3** under basic conditions to form the protected benzylguanine derivative **6** as a white solid. Finally, the deprotection of the trifluoroacetamide with methylamine gave **BG-NH<sub>2</sub>** in an overall yield of 55% over 5 steps. The **Halo-PEG<sub>n</sub>-Br** building blocks were synthesised as described by Erhart *et al.* (scheme 18, B)[128]. Thereafter, the polyethyleneglycols **7a**, **7b** and **7c** were each treated with 6-chloro-1-iodohexane, in the presence of sodium hydride for 16 h, to give a mixture of the HaloTag substrates **8a**, **8b**, and **8c** and their corresponding bis alkylated analogues in a ratio of approximately 2:1. The Apple reaction of **8a**, **8b**, and **8c** with triphenylphosphine and carbon tetrabromide readily produced the desired **Halo-PEG<sub>n</sub>-Br** building blocks in overall yields within the range of 30 to 42%, with slightly higher yields for the longer analogues. **Halo-NH<sub>2</sub>** was produced in a similar reaction sequence (scheme 18, C). However, an additional protection and deprotection step was required to prevent alkylation of the amine. Following the synthesis reported by So *et al.*[256] we protected the primary amine of 2-(2-aminoethoxy)ethanol (**9**) with a Boc protecting group using Boc<sub>2</sub>O. Reaction of the protected amine with 6-chloro-1-iodohexane in the presence of sodium hydride afforded **11** as a colorless oil. Cleavage of the Boc protecting group of **11** in a 4 M HCl solution (dioxane) and subsequent evaporation of the solvent, produced the HCl salt of **Halo-NH<sub>2</sub>** as a white solid.



**Scheme 18:** Synthesis of **BG – NH<sub>2</sub>**, **Halo – PEG<sub>n</sub>–Br** and **Halo – NH<sub>2</sub>** building blocks. A) Reaction conditions: a) LiAlH<sub>4</sub>, THF, 70°C, 3 h; b) Ethyl trifluoroacetate, DIPEA, MeOH, r.t., 2 h; c) 1-Methylpyrrolidinium, DMF r.t., 16 h; d) **3**, NaH, DMF, r.t., 5 h; e) Methylamine, MeOH, r.t., 16 h. B) Reaction conditions: a) NaH, THF, 6-chloro-1-iodohexane, r.t., 16 h; b) PPh<sub>3</sub>, CBr<sub>4</sub>, THF, r.t., 16 h. C) Reaction conditions: Boc<sub>2</sub>O, EtOH, r.t., 2 h; b) NaH, THF, 6-chloro-1-iodohexane, r.t., 16 h; c) 4 M HCl in dioxane, r.t., 3 h.

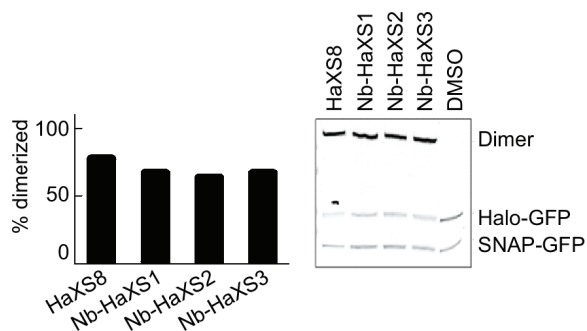
### 3.2.3 Synthesis and evaluation of nitrobenzyl derivative containing HaXS molecules

We initiated the synthesis of **Nb-HaXS1**, **Nb-HaXS2**, and **Nb-HaXS3** (scheme 19) with the alkylation of 5-hydroxynitrobenzyl alcohol (**12**) with the different **Halo – PEG<sub>n</sub>–Br** building blocks. The reaction proceeded under mild basic conditions and gave the desired products **13a**, **13b**, and **13c** in moderate yields as colorless oils. No major difference of reactivity was observed for the three electrophiles as the yields ranged from 41% to 46%. The carbamate formation between **BG – NH<sub>2</sub>** and **13a**, **13b**, respectively **13c** was achieved using various coupling procedures. However, the highest yields were obtained by activating the primary alcohol with bis(*p*-nitrophenyl) carbonate for 16 hours at room temperature, followed by the addition of **BG – NH<sub>2</sub>** and NEt<sub>3</sub>.



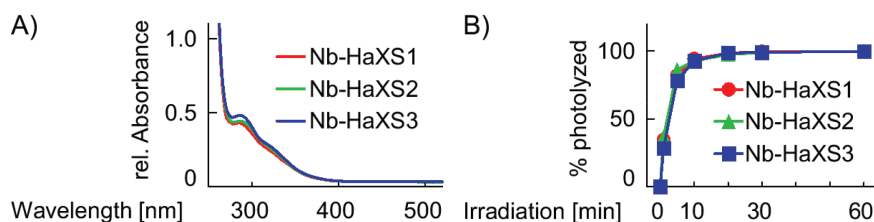
**Scheme 19:** Synthesis of **Nb-HaXS1**, **Nb-HaXS2**, and **Nb-HaXS3**. Reaction conditions: a) Corresponding **Halo-PEG<sub>n</sub>-Br**, K<sub>2</sub>CO<sub>3</sub>, DMF, 60°C, 16 h; b) bis(*p*-nitrophenyl) carbonate, triethylamine, DMF, r.t., 16 h; c) **BG-NH<sub>2</sub>**, 60°C, 6 h.

To validate the dimerisation efficacy of Nb-HaXS molecules in a cellular system, we treated HeLa cells, co-expressing Halo-GFP and SNAP-GFP fusions, with 0.5  $\mu\text{M}$  of **Nb-HaXS1**, **Nb-HaXS2**, **Nb-HaXS3** or **HaXS8** for 60 minutes (figure 18). After washing and lysing the cells, the tagged proteins were analysed by Western blot. Despite the difference in polarity, due to the increasing number of PEG units, no significant difference in the dimerisation activity between **Nb-HaXS1**, **Nb-HaXS2** and **Nb-HaXS3** was observed. All three compounds exhibited slightly lower dimerisation activity compared to **HaXS8**, as they induced approximately 65% of protein dimerisation compared to 78% dimerisation triggered by **HaXS8**.



**Figure 18:** Nb-HaXS induced *in vivo* dimerisation of Halo-GFP and SNAP-GFP fusion proteins. a) HeLa cells transfected with expression constructs for SNAP-tag-GFP (SNAP-GFP) and HaloTag-GFP (Halo-GFP) fusion proteins were exposed to 5  $\mu\text{M}$  **Nb-HaXS1**, **Nb-HaXS2**, **Nb-HaXS3** or **HaXS8** for 60 minutes in cell culture medium at 37°C. Subsequently, cells were lysed and proteins were subjected to SDS PAGE and immunoblotting. Tagged proteins were detected using anti- GFP (primary) and horseradish peroxidase labeled (secondary) antibodies, and chemiluminescence.

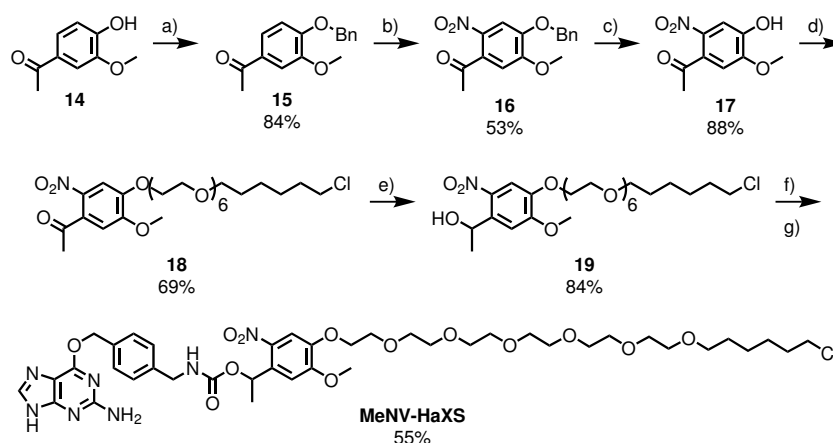
We next investigated the photophysical and photochemical properties of the Nb-HaXS molecules. As expected, the number of PEG units introduced had no influence on the photophysical properties of Nb-HaXS molecules. They all exhibited the highest absorption at 328 nm and displayed molar extinction coefficient of approximately  $1700 \text{ M}^{-1}\text{cm}^{-1}$  at 360 nm (figure 19, A). To evaluate the ability of the Nb-HaXS compounds to undergo photolysis, we irradiated 5  $\mu\text{M}$  solutions of corresponding CID in DMSO/water (10:90) in an LED based reactor with light at 360 nm for a given time (figure 19, B). The time courses for the photolysis reaction were analysed with a UPLC, monitoring the disappearance of the single ion signal of the starting material. Although, all compounds were successfully cleaved under the described conditions, 20 minutes of irradiation were required to achieve 95% of photolysis.



**Figure 19:** Photophysical properties of **Nb-HaXS1**, **Nb-HaXS2**, and **Nb-HaXS3**. A) The UV/Vis spectra of 0.5 mM Nb-HaXS molecules in DMSO were recorded at room temperature. B) 5 mM of solutions of **Nb-HaXS1**, **Nb-HaXS2**, and **Nb-HaXS3** in a DMSO/water (10:90) mixture were irradiated with four RPR-3500 Å lamps with light at 360 nm for the indicated time period. Subsequently, samples were subjected to a UPLC. Decrease of starting material was detected using a ESI-SQD mass spectrometer.

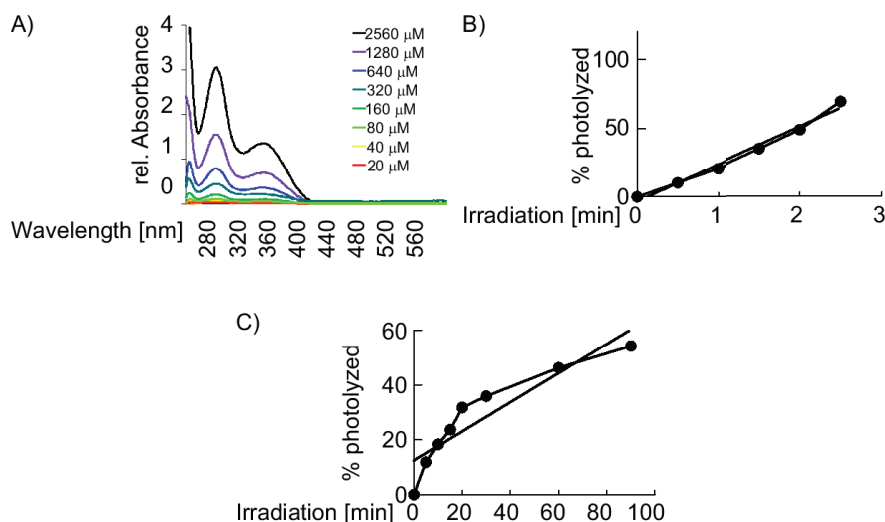
The results of the following part were published in 2014 at *Angewandte Chemie International Edition* Volume 53, Issue 18[257].

To improve the photophysical properties of nitrobenzyl based HaXS molecules, we carried out the synthesis of **MeNV-HaXS**. Incorporation of the methoxy group in the aromatic region should shift the absorption maximum to longer wavelengths[252] and thereby minimise possible cell damage, whereas the methyl group in an  $\alpha$ -position should increase the decomposition rate of the CID[251]. Due to the higher water solubility of **Nb-HaXS3** compared to its shorter analogues, we decided to use six PEG units to tether the chloroalkane with the MeNV moiety. Inspired by the synthesis of nitroveratril caged compounds reported by Tsai and Klinman[258], we initiated the synthesis of **MeNV-HaXS** with the protection of acetovanillone (**14**) with the benzyl protecting group (scheme 20). Under basic conditions, the phenol group was readily protected with benzylbromide to form **15** in an excellent yield as a white precipitate, which was used without further purification. Subsequent nitration of **15** in a mixture of nitric acid, acetic acid and acetic anhydride produced **16** as a 53% yield. After the cleavage of the benzyl protecting group in the presence of hydrogen bromide to give **17**, alkylation of the phenol group with **Halo-PEG<sub>6</sub>-Br** gave the intermediate **18** in a 60% yield over two steps. Reduction of the carbonyl group was successfully performed using NaBH<sub>4</sub> in a 1:1 mixture of MeOH/Dioxane to yield **19**. Finally, the coupling between **19** and **BG-NH<sub>2</sub>** with bis(4-nitrophenyl) carbonate produced **MeNV-HaXS** in an overall yield of 15% over 6 steps.



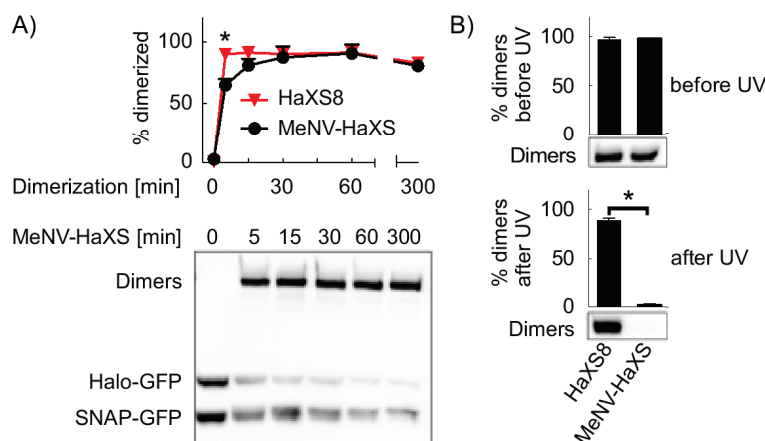
**Scheme 20:** Synthesis of **MeNV-HaXS**. Reaction conditions: a)  $\text{K}_2\text{CO}_3$ , benzyl bromide, DMF,  $80^\circ\text{C}$ , 16 h; b) Acetic acid, acetic anhydride,  $\text{HNO}_3$ , r.t., 16 h; c) Acetic acid, HBr,  $85^\circ\text{C}$ , 1.5 h; d)  $\text{K}_2\text{CO}_3$ , **Halo-PEG<sub>6</sub>-Br**, DMF,  $60^\circ\text{C}$ , 16 h; e)  $\text{NaBH}_4$ , MeOH/Dioxane, r.t., 2 h; f) bis(*p*-nitrophenyl) carbonate, triethylamine, DMF, r.t., 16 h; g) **BG-NH<sub>2</sub>**,  $60^\circ\text{C}$ , 6 h.

To investigate the photophysical properties of **MeNV-HaXS**, we recorded the UV-Vis spectra of different concentrated solutions of the photocleavable CID in DMSO (figure 20, A). As envisaged, the introduction of the methoxy group in a para position to the nitro group improved the absorption properties of the CID, red-shifting its absorption maximum by 35 nm to 355 nm and increasing the molar extinction coefficient at 360 nm ( $\epsilon = 4058 \text{ M}^{-1}\text{cm}^{-1}$ ) and 405 ( $\epsilon = 1238 \text{ M}^{-1}\text{cm}^{-1}$ ), compared to the parent Nb-HaXS series. To validate the influence of the methyl group in the  $\alpha$ -benzylic position on the reaction rate of the photolysis, we measured the decay rate of a  $0.5 \mu\text{M}$  solution of **MeNV-HaXS** in a DMSO/water (1:10) mixture, under the same conditions as described for the Nb-HaXS series. After 5 minutes of irradiation, over 95% of the substrate was successfully cleaved, increasing the photolytic activity by a factor of four compared to the first generation of photocleavable HaXS molecules (data not shown). However, this measurement strongly depends on several factors, such as the concentration, the power of the lamp, the distance of the lamp, and so on. In contrast, the quantum yield of a molecule only depends on the specific chemical reaction, the irradiation wavelength and the solvent. The quantum yield ( $\phi$ ) of a given molecule can be determined by dividing the photolytic rate by the number of photons entering the photoreactor per unit of time (quantum flow). Therefore, we used ferrioxalate actinometry to measure the quantum flow of the photoreactor. Based on these results, we calculated that **MeNV-HaXS** has an excellent quantum yield of 0.075 in a 10% solution of DMSO in water at 360 nm and 0.07 at 405 nm (figure 20, B and C).



**Figure 20:** Photophysical properties of **MeNV-HaXS**. A) The UV/Vis spectra of different **MeNV-HaXS** solutions in DMSO were recorded at room temperature. B) A 0.5 mM solution of **MeNV-HaXS** in DMSO/water mixture (10:90) was irradiated for the indicated time period with an LED lamp, emitting  $1.33 \cdot 10^{-8}$  E/s at 360 nm. Subsequently, samples were subjected to a UPLC. Decrease of starting material was detected using an ESI-SQD mass spectrometer. C) A 0.5 mM solution of **MeNV-HaXS** in DMSO/water mixture (10:90) was irradiated for the indicated time period with an LED lamp, emitting  $1.33 \cdot 10^{-8}$  E/s at 405 nm. Subsequently, samples were subjected to a UPLC. Decrease of starting material was detected using an ESI-SQD mass spectrometer.

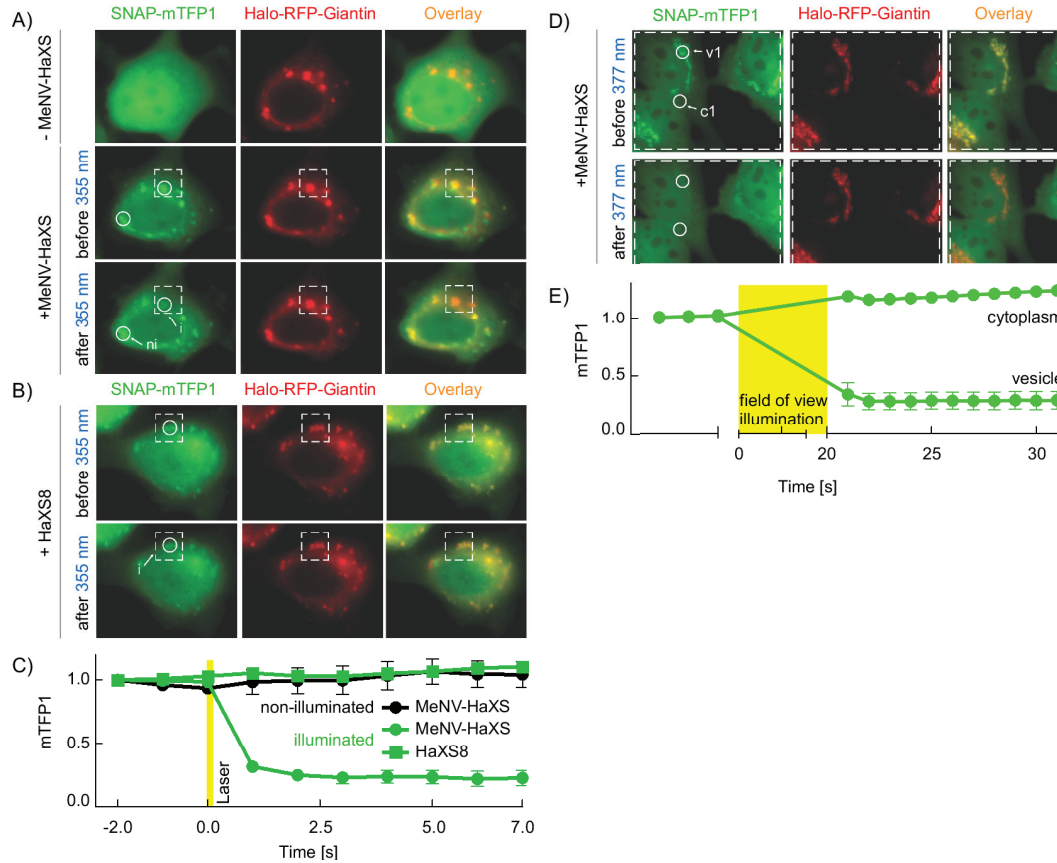
We next investigated the time-dependent dimerisation of Halo-GFP and SNAP-GFP fusion proteins expressed in HeLa cells, in response to the addition of **MeNV-HaXS** and **HaXS8** (figure 21, A). The two CIDs exhibited similar reaction rates and only slight differences were measured at earlier time points ( $t > 10$  min). The slightly inferior *in vivo* dimerisation rate of **MeNV-HaXS** can most probably be attributed to a reduced cell permeability, resulting from its increased molecular weight and higher polarity in respect to **HaXS8**. However, after 15 minutes of incubation the difference became insignificant and both, **MeNV-HaXS** and **HaXS8** induced over 90% fusion protein dimers, which were stable for more than 5 hours at ambient light. Thus, the noncleavable **HaXS8** can be used as control compound to analyze the efficiency of photocleavage, and to detect possible side effects arising from UV irradiation. In a proof of concept, we tested the intracellular photocleavage of a **MeNV-HaXS**-dimerised HaloTag-SNAP-tag complex. We incubated HeLa cells expressing Halo-GFP and SNAP-GFP fusion proteins with either **MeNV-HaXS** or **HaXS8** for 15 minutes. Irradiation of the cells with a 360 nm lamp (Blak-Ray, B-100A, UVP) for 10 minutes led to quantitative cleavage of **MeNV-HaXS**-induced dimers. In contrast, **HaXS8**-containing dimers remained intact (figure 21, B) during and after irradiation. **MeNV-HaXS** thus induces covalent dimerisation of Halo-POI and SNAP-POI fusion proteins, which can be readily reversed in bulk with high temporal resolution.



**Figure 21: MeNV-HaXS-induced intracellular dimerisation of HaloTag and SNAP-tag fusion proteins and subsequent cleavage upon UV illumination.** A) HeLa cells co-expressing SNAP-tag-GFP (SNAP-GFP) and HaloTag-GFP (Halo-GFP) fusion proteins were exposed to 5  $\mu$ M **MeNV-HaXS** or 5  $\mu$ M **HaXS8** for the indicated time periods in cell culture medium at 37°C. Subsequently, cells were lysed and proteins were subjected to SDS PAGE and immunoblotting. Tagged proteins were detected using anti- GFP (primary) and horseradish peroxidase labeled (secondary) antibodies, and chemiluminescence (means SEM, n = 3). B) HeLa cells co-expressing SNAP-GFP and Halo-GFP as in (A) were incubated with 5  $\mu$ M **MeNV-HaXS** or **HaXS8** for 15 minutes. Cells were then washed and submerged in phosphate-buffered saline (PBS) to remove unreacted compounds. Subsequently cells were illuminated with a high-intensity UV lamp (100 W, 5 cm distance) for 10 minutes. Analysis of dimerisation products was performed as in (A); values represent means SEM, n=3; \* indicates p < 0.05.

Many biological events occur at specific intracellular locations and to investigate these events manipulation techniques with high precision are required. To demonstrate the high spatial resolution of **MeNV-HaXS**, we anchored tagged proteins to Golgi (figure 22, A), and subsequently released the dimerised proteins. Therefore, we co-transfected HeLa cells with cytosolic teal fluorescent (cyan) SNAP-tag fusion protein (SNAP-mTFP1) and the Golgi-targeting, red fluorescent Halo-RFP-Giantin[259]. Incubation of these cells with 5  $\mu$ M of **MeNV-HaXS** or 5  $\mu$ M of **HaXS8** led to efficient translocation of SNAP-mTFP1 to the cytosolic surface of the Golgi membrane (figure 22, A). The irradiation of a specific region of the Golgi with 8x5 ms pulses at 355 nm from an XY scanning excitation FRAP laser, efficiently and selectively released SNAP-mTFP1 from the illuminated Golgi-derived vesicles, indicated by a reduced fluorescence intensity in the illuminated region. In contrast, no significant loss of fluorescence intensity was observed after the irradiation of **HaXS8**-induced SNAP-mTFP1-Halo-RFP-Giantin dimers (figure 22, B and C), confirming that the observed decrease in fluorescence at the Golgi, in **MeNV-HaXS** treated cells, results from the release of SNAP-mTFP1, and not from photobleaching. These results show that **MeNV-HaXS** can be used to regulate the localisation of a protein with excellent subcellular precision on a timescale of seconds. To underline the high potential of this approach, we tagged proteins to other intracellular compartments such as the plasma membrane, lysosomes, mitochondria, and the actin skeleton (see figure S3). Since a scanning FRAP laser is not part of the standard fluorescence microscopic equipment, we next explored the photocleavage of **MeNV-HaXS**-induced protein complexes using global field of view illumination with standard DAPI excitation filters (377 +/- 25 nm; standard mercury halide lamp). Irradiation of the cells for <20 s was sufficient to completely liberate SNAP-mTFP1 from Golgi membranes

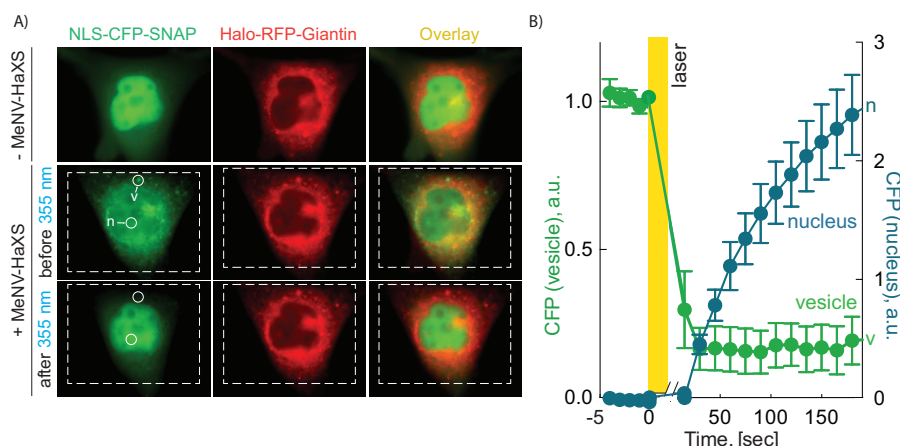
(figure 22, D and E), indicated by a decrease in the mTFP1 fluorescence at the Golgi. The corresponding increase of cyanic fluorescence in the cytoplasm demonstrates that the light-induced fluorescence decrease at vesicles, results from the release of SNAP-mTFP1 and not from global photobleaching. As expected the photocleavage of **MeNV-HaXS**-induced dimers under these conditions is slower, due to the limited excitation energy. However, the use of DAPI excitation filters on a conventional fluorescence microscope greatly expands the range of applications.



**Figure 22:** Translocation of cytosolic SNAP-mTFP1 proteins to the Golgi and their subsequent release after irradiation. A,B) HeLa cells expressing SNAP-mTFP1 and Halo-RFP-Giantin were treated with A) 5  $\mu$ M **MeNV-HaXS** or B) 5  $\mu$ M **HaXS8** in cell culture medium for 15 minutes at 37°C. The fluorescence of SNAP-mTFP1 was monitored in the indicated circular regions of interest by live cell microscopy, before and after illumination of a subcellular region within the cell (white, dotted square) with a scanning FRAP laser (8 areas 5 ms at 355 nm). C) Quantification of mTFP1 fluorescence intensity in selected regions of interest (labeled circles) for SNAP-mTFP1 after addition of **MeNV-HaXS** (circles) or **HaXS8** (squares). Quantifications of mTFP1 intensity in illuminated areas (green curves) and non-illuminated Golgi vesicles (black curve). D) HeLa cells as in (a) were exposed to 5  $\mu$ M **MeNV-HaXS** and illuminated for 20 s using a standard DAPI filter set on a conventional fluorescence microscope ( $t=20$  s,  $377 \pm 25$  nm). E) Quantification of mTFP1 fluorescence intensity in selected regions of interest (circles) at Golgi-derived vesicles, and in the cytoplasm before and after illumination as described in (d).



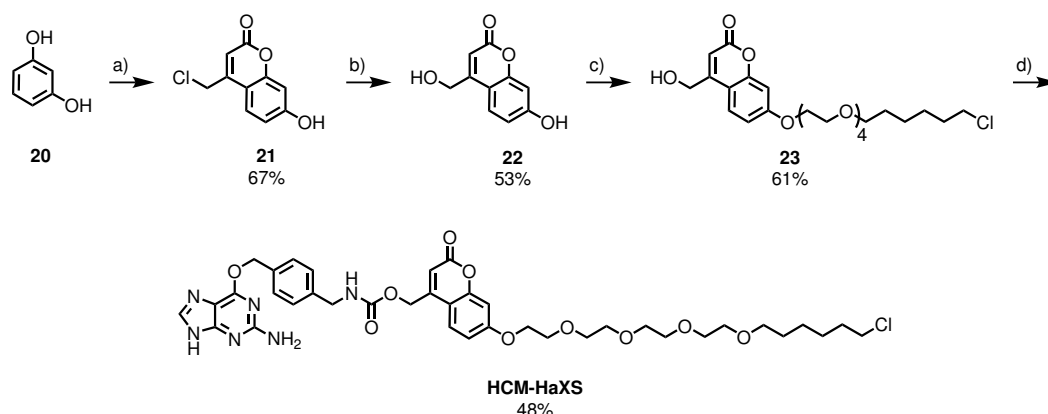
To further demonstrate the utility of **MeNV-HaXS**, we anchored a nuclear probe to the perinuclear sites of the Golgi. Therefore, we studied cells co-transfected with a fusion protein nucleus localising fusion protein consisting of the nuclear localisation sequence (NLS), cyan fluorescent protein and the SNAP-tag (NLS-CFP-SNAP) and a Golgi targeting Giantin-RFP-Halo fusion in the presence or absence of CID. In untreated cells, the NLS-CFP-SNAP fusion protein accumulated in the nucleus. However in the presence of **MeNV-HaXS**, NLS-CFP-SNAP is quickly trapped at perinuclear sites on the Golgi (figure 23, A). Upon irradiation, the multi protein complex was cleaved and NLS-CFP-SNAP relocated to the cytosol within seconds. The observed delay of the nuclear import of released NLS-CFP-SNAP (figure 23, B), is due to nuclear import processes[260] These results clearly demonstrate that **MeNV-HaXS** can be used to study translocation kinetics in real time in a simple experiment.



**Figure 23: MeNV-HaXS-controlled translocation of NLS-CFP-SNAP probe from the nucleus to the Golgi and light-induced nuclear import of NLS-CFP-SNAP.** A) HeLa cells expressing Halo-RFP-Giantin and NLS-CFP-SNAP were exposed to 5  $\mu$ M **MeNV-HaXS** in cell-culture medium for 15 minutes at 37°C and subsequently irradiated with a scanning FRAP laser (150 areas 5 ms at 355 nm). CFP fluorescence intensity was monitored in the indicated circular regions of interest by live-cell microscopy, before and after illumination of the cell (white, dashed rectangle). B) Quantification of CFP fluorescence intensity in selected regions of interest (labeled circles) monitoring vesicle-associated (green curve) and nuclear NLS-CFP-SNAP concentrations (blue curve) before and after illumination of the cells are shown; values are means SEM, n=7 cells, error bars not shown where smaller than symbols used.

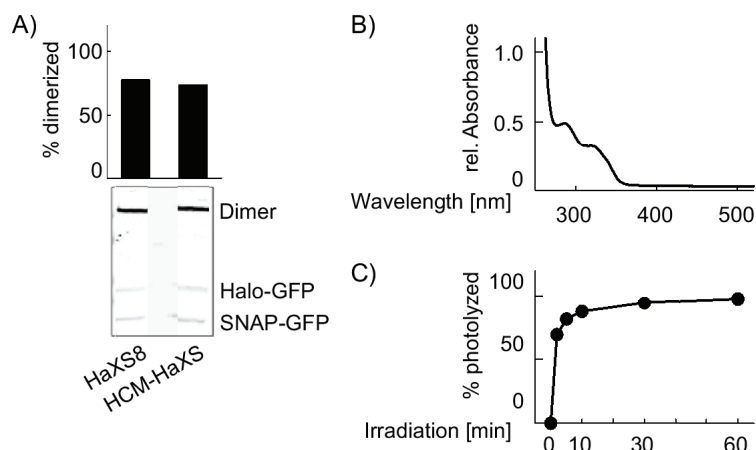
### 3.2.4 Synthesis and evaluation of coumarin derivative containing HaXS

Coumarinyl-4-methyl derivatives and especially 7-aminocoumarinyl-4-methyl (**ACM**) derivatives possess larger extinction coefficients, higher absorption maxima and faster photolysis rates compared to most common nitrobenzyl derivatives[250]. Additionally, coumarins are well suited for two-photon excitation, allowing cleavage at a very high wavelength, in respect to the single-photon photolysis. Therefore, incorporation of a coumarin-based PPG into the core of HaXS molecules would enable the possibility to cleave the formed protein complex more efficiently and under less toxic conditions. Moreover, the combination with **MeNV-HaXS** would enable the independent control of the formation and release of two different protein complexes, simultaneously. To validate the dimerisation activity of coumarin containing HaXS, we synthesised 7-alkoxycoumarinyl-4-methylhydroxyl HaXS derivatives (**HCM-HaXS**) in four steps, starting from the commercially available resorcinol (**20**) (scheme 21). In the presence of ethyl 4-chloroacetoacetate, resorcinol underwent an acid catalysed Pechmann condensation, giving a good yield of the desired 7-hydroxy coumarinyl-4-methyl chloride (**21**). Subsequent substitution of the chloro group with a hydroxyl group in a water/dioxane (1:1) mixture resulted in the formation of the coumarinyl-4-methyl alcohol **22**. Alkylation of the 7-hydroxy group with the **Halo-PEG<sub>4</sub>-Br** under mild basic conditions provided the HaloTag substrate **23**, as a yellow-tinted oil. Treatment of **23** with carbodiimidazole (CDI) furnished the corresponding activated carbamate, which was subsequently coupled to **BG-NH<sub>2</sub>** to provide **HCM-HaXS** in a moderate overall yield of 10 % over 4 steps.



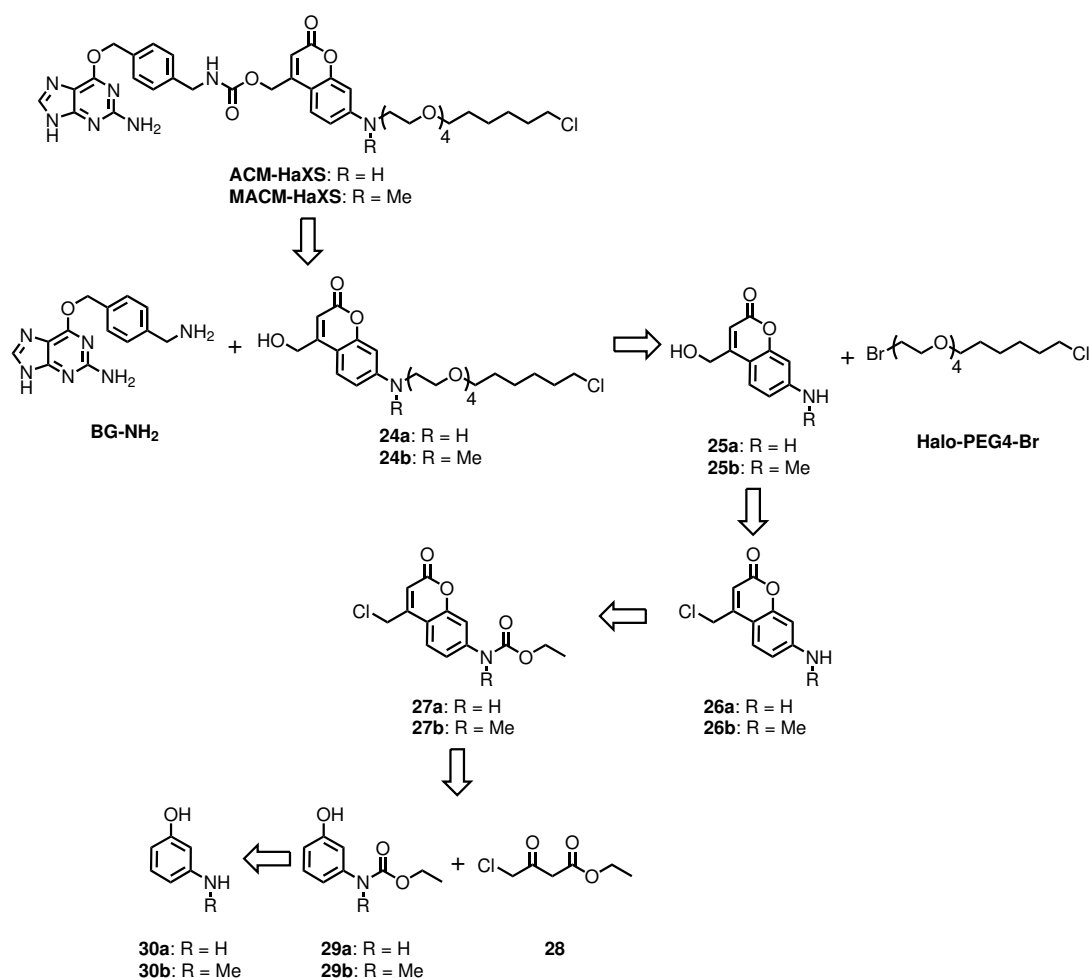
**Scheme 21:** Synthesis of **HCM-HaXS**. Reaction conditions: a) Ethyl 4-chloroacetoacetate, MeHSO<sub>4</sub>, 60°C, 16 h; b) H<sub>2</sub>O/Dioxane, 100°C, 16 h; c) K<sub>2</sub>CO<sub>3</sub>, **Halo-PEG<sub>4</sub>-Br**, DMF, rt, 16 h; d) CDI, DMF, rt. 16 h; e) **BG-NH<sub>2</sub>**, DIPEA, DMF, 50°C, 5 h.

To validate the dimerisation activity of **HCM-HaXS** *in vivo*, we incubated HeLa cells, co-expressing Halo-GFP and SNAP-GFP fusions, with 5 μM of **HCM-HaXS** or **HaXS8** for varying 60 minutes. Western blot analysis revealed that **HCM-HaXS** and **HaXS8** both induced over 70% of SNAP-tag and HaloTag fusion protein dimerisation (figure 24, A). We next measured the UV-Vis absorption spectra and the photolytic rate of **HCM-HaXS** (figure 24, B and C). As expected from reported values[201], the **HCM-HaXS** exhibited a relatively low absorption maxima (322 nm) and molar extinction coefficients at 360 nm (990 M<sup>-1</sup>cm<sup>-1</sup>) and 405 nm (678 M<sup>-1</sup>cm<sup>-1</sup>).



**Figure 24:** Evaluation of the photophysical and biological properties of **HCM-HaXS**. A) Absorption spectra of a 0.50 mM solution of HCM-HaXS in a DMSO/water (10:90) mixture. B) The UV/Vis spectra of 0.5  $\mu$ M **HCM-HaXS** molecules in DMSO were recorded at room temperature. C) A 5  $\mu$ M solution of **HCM-HaXS** in DMSO/water (10:90) was irradiated for the indicated amount of time with four RPR-3500 Å lamps emitting light at 360 nm. CID in a DMSO/water (10:90) mixture was irradiated with four RPR-3500 Å lamps with light at 360 nm for the indicated amount of time. Decrease of starting material was detected using a UPLC equipped with an ESI-SQD mass spectrometer.)

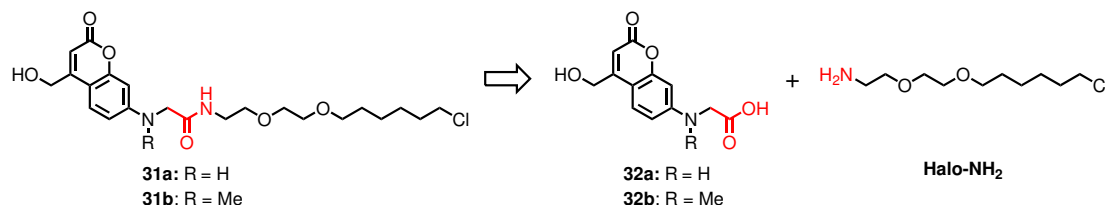
The photolytic rate of **HCM-HaXS** was far inferior to the photolytic rate measured for **MeNV-HaXS** at 360 nm, reflecting the low extinction coefficient. As previously mentioned, the introduction of electron-donating groups at either the C7 or C6 position improves the spectroscopic and photochemical properties of coumarin derivatives[255]. Among these coumarinylmethyl derivatives, 7-amino substituted exhibit the highest quantum yields (0.21-0.28) and absorption maxima between 350-400 nm[250]. Therefore, we further targeted the synthesis of 7-aminocoumarin-4-methylhydroxyl and 7-aminomethylcoumarin-4-methylhydroxyl containing HaXS molecules (**ACM-HaXS** and **MACM-HaXS**) (scheme 22). Similar to **HCM-HaXS**, we built the core coumarin structure (**25a** and **25b**) by Pechmann condensation of either aminophenol (**26a**) or methylaminophenol (**26b**) with ethyl 4-chloroacetoacetate. Alkylation of the amino group of either the methylchlorocoumarin or the methylhydroxylcoumarin with the **Halo-PEG<sub>4</sub>-Br** building block under basic conditions should result in the corresponding HaloTag substrate **24a** and **24b**. Finally, we envisaged forming the carbamate between **24a**, or **24b** and **BG-NH<sub>2</sub>** using common coupling reagents, such as DCC, CID or PyBOP.



**Scheme 22:** Retrosynthetic analysis of the synthesis of **ACM-HaXS** and **MACM-HaXS**.

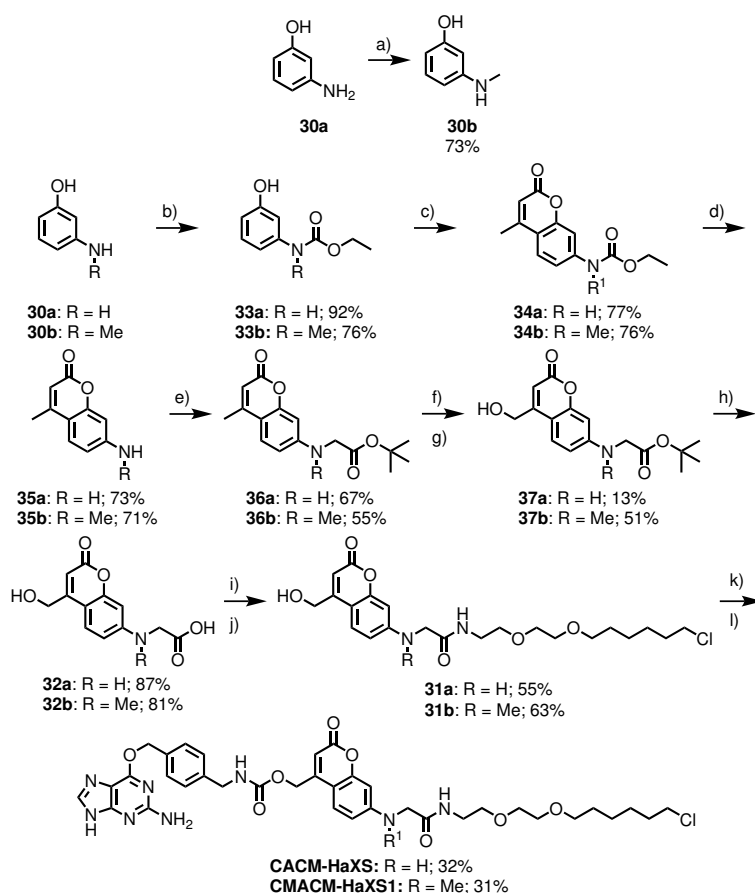
The reaction sequence to synthesise **MACM-HaXS** started with the alkylation of the commercially available 3-aminophenol (**30a**) with iodomethane under basic conditions. In our initial attempts, a one to one mixture of mono- and bismethylated product was observed and the yield of the desired 3-methylaminophenol (**30b**) was increased to 73% by using a slight excess of the starting material, as well as a dropwise addition of iodomethane in a diluted solution of 3-aminophenol in DMF. In contrast to the synthesis of **HCM-HaXS**, the electron donating character of the amino group prevents the direct cyclisation of aminophenol derivatives. Therefore, we protected **30a** and **30b** with ethyl chloroformate, enabling the deactivated aminophenols **29a** and **29b** to occur with good yields. Treatment of these protected aminophenols with ethyl 4-chloroacetoacetate in methanesulfonic acid at 60°C and subsequently quenching with an excess of ice cold water, created 7-aminocoumarin-4-methyl chloride (**27a**) and 7-methylaminocoumarin-4-methyl chloride (**27b**) as an orange precipitate. The hydrolysis of the carbamate to de-protect the amino group was carried out in a mixture of H<sub>2</sub>SO<sub>4</sub> and acetic acid at 100°C and gave **24a** and **24b** as a dark red precipitate. Subsequent substitution of the chlorides with a hydroxyl group in water proceeded with moderate yields. Unfortunately, all attempts to alkylate the coumarin

compounds **25a** or **25b** with **Halo – PEG<sub>4</sub>-Br** failed. In an alternative approach to synthesise **24a** and **24b**, we attempted to introduce a Halo – PEG<sub>4</sub> moiety prior to the cyclisation reaction. The first two steps, the alkylation of aminophenol with the **Halo – PEG<sub>4</sub>-Br** and the subsequent protection with ethyl chloroformate proceeded with no difficulties under basic conditions. However, the Pechmann condensation under acidic conditions led to coumarinyl-4-methyl chloride products missing the chloroalkane. Therefore, we modified the synthesis of **ACM-HaXS** and **MACM-HaXS**, incorporating an acetamide moiety between the HaloTag substrate and the coumarinyl group, which could be formed by coupling the corresponding amine and carboxylic acid (scheme 23).



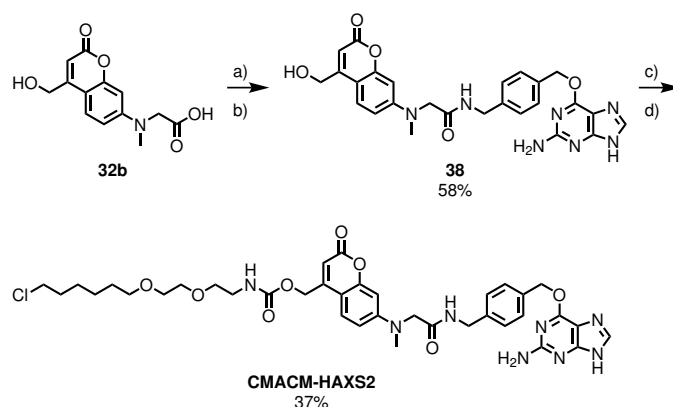
**Scheme 23:** Key steps of the synthesis of novel designed coumarin-4-methyl containing HaXS molecules.

Analogues to the coumarinylmethyl ester synthesis previously described by Hagen *et al.*[261], we attempted to introduce a *tert*-butyl protected carboxylic acid after the coumarin formation. In order to prevent possible side reactions with the chlorine or the hydroxyl group, we investigated the alkylation of the 7-aminocoumarine-4-methyl and 7-methylaminocoumarine-4-methyl derivatives. Compounds **33a** and **33b** were obtained as described above (scheme 24). Pechmann condensation of **33a** and **33b** with ethyl acetoacetate in MeHSO<sub>3</sub> provided coumarinyl derivatives **34a** and **34b** as a red precipitate, which was used without further purification. Subsequently, de-protection of the amino group under acidic conditions yielded the amines **35a** and **35b**. Alkylation of **35a** and **35b** with an excess of *tert*-butyl bromoacetate at 100°C led to the formation of **36a** and **36b** in acceptable yields. The allylic oxidation of **36b** with freshly sublimated SeO<sub>2</sub> gave the corresponding aldehyde, which was subsequently reduced to the alcohol **37b** in moderate yields, using NaBH<sub>4</sub>. In contrast, the same oxidation and reduction sequence of the non-methylated analogue **36a** provided 10 % of the desired product **37a** and increased amounts of degraded product. Cleavage of the *tert*-butyl group with TFA yielded two general photocleavable building blocks **37a** and **37b**, which could be used to attach a variety of functionalities in two synthetic steps. The carboxylic acids of **37a** and **37b** were activated using DCC and NHS, followed by amide formation with **Halo – NH<sub>2</sub>** to give compounds **32a** and **32b**. Carbamate formation between **32a** and **32b** and **BG – NH<sub>2</sub>** under the same conditions described for **HCM-HaXS** yielded **CACM-HaXS** and **CMACM-HaXS1** in **32**, respectively 39%, yields.



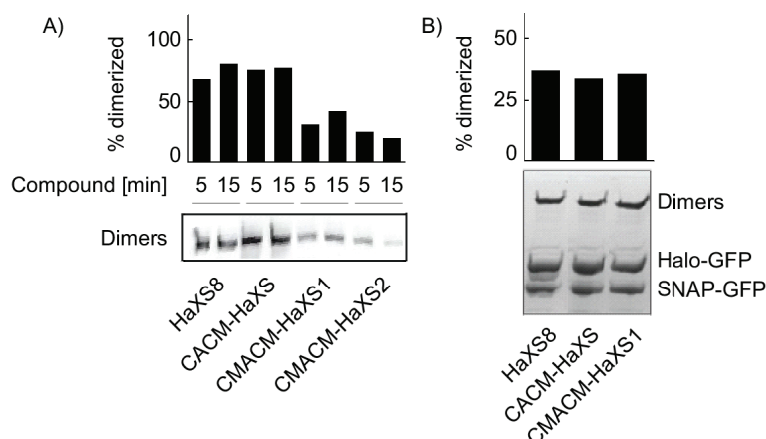
**Scheme 24:** Synthesis of **CACM-HaXS**, **CMACM-HaXS1** Reaction conditions: a)  $K_2CO_3$ , methyl iodide DMF, r.t., 16 h; b)  $K_2CO_3$ , Ethyl chloroformate, DMF, r.t., 16 h; c) Ethyl acetoacetate,  $MeHSO_3$ ,  $60^\circ C$ , 16 h; d) Conc. sulfuric acid, acetic acid  $100^\circ C$ , 16 h; e)  $K_2CO_3$ , *tert*-butyl bromoacetate DMF,  $100^\circ C$ , 16 h; f)  $SeO_2$ , *p*-xylene,  $140^\circ C$ , 16 h; g) Sodium borohydride, THF/MeOH rt, 4 h; h) TFA,  $CH_2Cl_2$ , r.t., 3 h; i) DCC, NHS, DMF, rt, 2 h; j) **Halo** -  $NH_2$ , TEA, r.t., 16 h; k) bis(*p*-nitrophenyl) carbonate  $CH_2Cl_2$  r.t. 3 h; l) **BG** -  $NH_2$ , DMF, TEA,  $60^\circ C$ , 3 h.

Starting from the general building block **37b**, we additionally synthesised **CMACM-HaXS2**, an analogue of **CMACM-HaXS1**, with reversed orientation of the tag substrates (Scheme 25). Treatment of **37b** with DCC and NHS followed by the addition of **BG** -  $NH_2$  under basic conditions gave compound **38** in moderate yields. Activation of the alcohol **38** with *p*-nitrophenyl carbonate for 16 hours at room temperature and a subsequent reaction with **Halo** -  $NH_2$ , produced **CMACM-HaXS2** in comparable yields to **CMACM-HaXS1**.



**Scheme 25:** Synthesis of **CMACM-HaXS2**. Reaction conditions: a) DCC, NHS, DMF, r.t., 2 h; b) **BG-NH<sub>2</sub>**, TEA, r.t., 16 h; c) bis(*p*-nitrophenyl) carbonate, CH<sub>2</sub>Cl<sub>2</sub>, rt, 3 h; d) **Halo-NH<sub>2</sub>**, DMF, TEA, 60, 3 h

To validate the *in vivo* dimerisation activity of these new 7-aminocoumarin-4-methyl HaXS molecules, we treated Hela cells co-expressing SNAP-GFP and Halo-GFP fusion proteins with 5  $\mu$ M **CACM-HaXS**, **CMACM-HaXS1**, **CMACM-HaXS2** and **HaXS8** (figure 25, A). The results of these *in vivo* assays clearly demonstrate that **CACM-HaXS** and **HaXS8** induced dimerisation of Halo-GFP and SNAP-GFP fusion proteins at similar rates. Both compounds triggered the formation of 80% fusion protein dimers after 15 minutes. In contrast to the previously discussed **MeNV-HaXS**, no difference was observed at early time points ( $t = 5$  minutes). Thus, **CACM-HaXS** may be preferred in biological experiments where rapid dimerisation is required. Surprisingly, the methylated compounds **CMACM-HaXS1** and **CMACM-HaXS2** induced dimerisation approximately two times less effectively with respect to **CACM-HaXS**. The orientation of the tag substrate has only minor influence on the dimerisation activity of the CID, as both **CMACM-HaXS1** and **CMACM-HaXS2** trigger dimerisation of fusion protein at similar rates. To investigate reasons for different *in vivo* reactivities of methylated and non-methylated coumarin containing HaXS molecules, we next treated recombinant SNAP-GFP with **HaXS8**, **CACM-HaXS** or **CMACM-HaXS1**, followed by the addition of recombinant Halo-GFP. Interestingly, all three compounds induced similar amounts of fusion protein dimers *in vitro* (figure 25, B). These results suggest that the lower *in vivo* dimerisation activity of the **CMACM-HaXS** derivatives may be due to a reduced bioavailability.

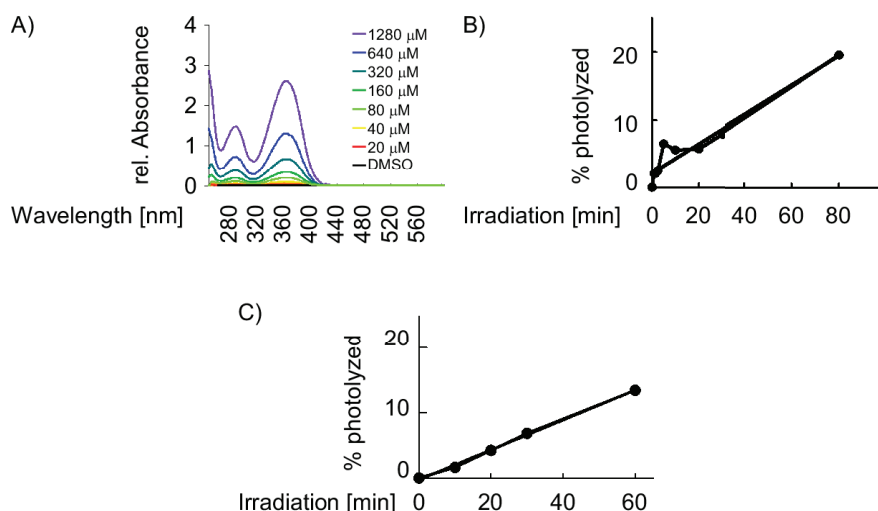


**Figure 25:** *In vivo* and *in vitro* dimerisation efficacy of **CACM-HaXS**, **CMACM-HaXS1** and **CMACM-HaXS2**. A) HeLa cells co-expressing SNAP-GFP and Halo-GFP fusion proteins were incubated with 5  $\mu$ M **HaXS8**, **CACM-HaXS**, **CMACM-HaXS1** and **CMACM-HaXS2** for 5 or 15 minutes in cell culture medium at 37°C before cells were lysed and proteins were subjected to SDS PAGE, immunoblotting and chemiluminescence. B) 5  $\mu$ M of recombinant SNAP-GFP fusion protein was incubated with 5  $\mu$ M of **HaXS8**, **CACM-HaXS** or **CMACM-HaXS1** for 30 minutes, prior to the addition of 5  $\mu$ M Halo-GFP. Thereafter, proteins were denatured, separated by SDS PAGE and detected by Coomassie-Blue staining.

To determine the molar extinction coefficient of **CACM-HaXS** at 360 nm and 405 nm, we measured the UV-Vis absorption spectra of different concentrated solutions in DMSO. The absorption maxima of **CACM-HaXS** (365 nm) is only slightly higher than the absorption maxima of **MeNV-HaXS** (355 nm) and significantly lower compared to other amino-substituted 7-aminocoumarinyl-4-methyl derivatives, which exhibit absorption maxima typically in the range of 380 - 400 nm (measured in different aqueous solutions, methanol or ethanol)[250]. Most likely the relatively low absorption maxima **CACM-HaXS** can be attributed to a solvent effect. In fact, more polar solvents cause a red shift of the absorption bands of  $\pi$  to  $\pi^*$  (coumarinyl derivatives) electron transitions[262, 263]. In contrast, the absorption bands of  $n$  to  $\pi^*$  (*o*-nitobenzyl derivatives) electron transitions experience a blue shift[262, 263]. Thus, we assume that the absorption maxima of **CACM-HaXS** in water or in a DMSO/water mixture should be red shifted with respect to the measured values, whereas the absorption maxima of **MeNV-HaXS** should be prominently blue shifted. However, **CACM-HaXS** still exhibits four-times higher molar extinction coefficients at 360 nm ( $23120 \text{ M}^{-1}\text{cm}^{-1}$ ) and two-times higher molar extinction coefficient at 405 nm ( $2622 \text{ M}^{-1}\text{cm}^{-1}$ ) compared to **MeNV-HaXS** ( $\epsilon_{360} = 4058 \text{ M}^{-1}\text{cm}^{-1}$  respectively  $\epsilon_{405} = 1238 \text{ M}^{-1}\text{cm}^{-1}$ ). To determine the quantum yield of **CACM-HaXS** at 360 nm and 405 nm we irradiated 3 mL of 0.5 mM solutions of the CID in a Lumos 43A photoreactor for indicated time. Aliquots of the samples were transferred to a HPLC and the disappearance of the starting material was monitored. **CACM-HaXS** exhibited significant lower quantum yields (0.006 at 360 nm and 0.016 nm at 405 nm), compared to the quantum yields reported for other 7-aminocoumarinyl-4-methyl carbamates (0.04 - 0.11)[264, 265]. Interestingly, samples irradiated at 360 nm displayed a higher photolytic rate at early time points (5 times higher at  $t < 5$  minutes). This may indicate that the formed by-products absorb at 360 nm and therefore interfere with the photolysis reaction[266]. Based on these results it is difficult to compare the photophysical properties of **CACM-HaXS** and **MeNV-HaXS** and draw conclusions on their



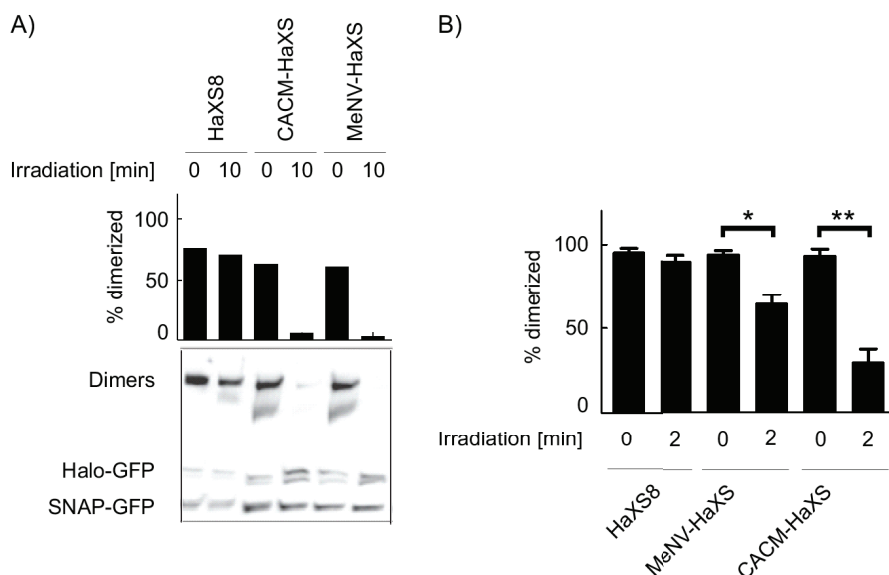
behaviour in cells. To obtain results, which better reflect the *in vivo* measurements, the UV-Vis spectra of **CACM-HaXS** and **MeNV-HaXS** should be recorded in the same medium used in the biological experiments. Additionally, the quantum yield experiments should be repeated using the same time scales for both compounds and lower concentrations. However, the product of the quantum yield and the molar extinction coefficient, which is proportional to the efficiency to release the leaving group at a given excitation wavelength[250], of **CACM-HaXS** ( $139 \text{ M}^{-1}\text{cm}^{-1}$  at 360 nm and  $42 \text{ M}^{-1}\text{cm}^{-1}$  at 405 nm) is only slightly lower in respect to the values measured for **MeNV-HaXS** ( $304 \text{ M}^{-1}\text{cm}^{-1}$  at 360 nm and  $87 \text{ M}^{-1}\text{cm}^{-1}$  at 405 nm). Thus, we assume that irradiation of **CACM-HaXS**-induced protein dimers at 360 nm or 405 nm should lead to efficient cleavage. Experiments to determine the two-photon cross section of **CACM-HaXS** and **MeNV-HaXS** at wavelength  $>700 \text{ nm}$  are ongoing.



**Figure 26:** Photophysical properties of **CACM-HaXS**. A) The UV/Vis spectra of different **CACM-HaXS** solutions in DMSO were recorded at room temperature. B) A 0.5 mM **CACM-HaXS** solution of in DMSO/water mixture (10:90) was irradiated for the indicated amount of time with an LED lamp, emitting  $5.44 \cdot 10^{-7} \text{ E/min}$  at 360 nm. Subsequently, samples were subjected to a UPLC. Decrease of starting material was detected using a ESI-SQD mass spectrometer and plotted against time. C) A 0.5 mM **CACM-HaXS** solution of in DMSO/water mixture (10:90) was irradiated for the indicated amount of time with an LED lamp, emitting  $5.44 \cdot 10^{-7} \text{ E/min}$  at 405 nm. Samples were analysed as analogues to B).

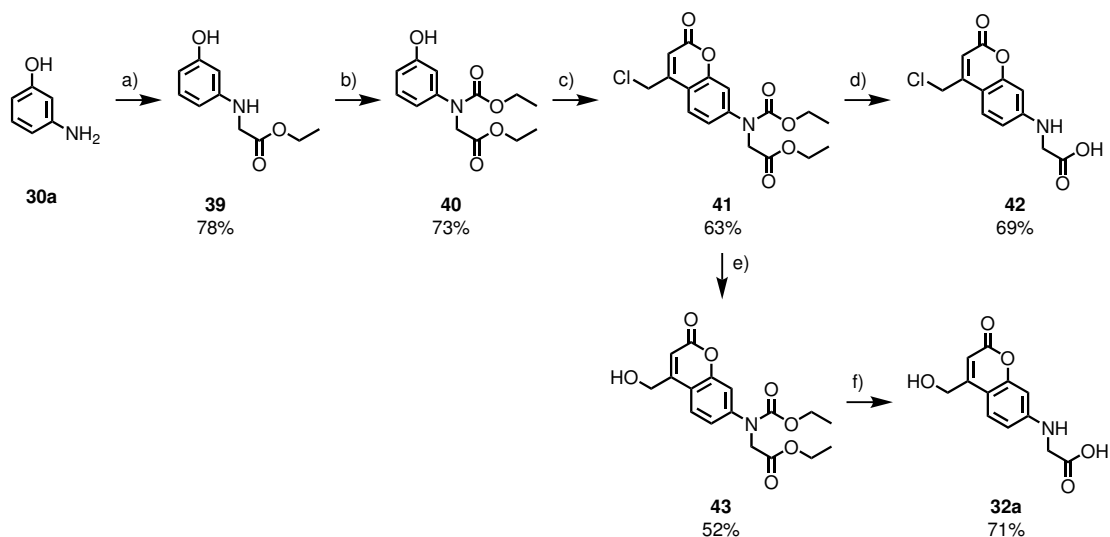
We next compared the *in vivo* photolysis reaction of **CACM-HaXS** and **MeNV-HaXS** dimerised SNAP-GFP and Halo-GFP fusion proteins at 360 nm (figure 27). Therefore, we incubated HeLa cells co-expressing Halo-GFP and SNAP-GFP with 5 μM of either **CACM-HaXS**, **MeNV-HaXS** or the non-cleavable **HaXS8** for 15 minutes, prior to irradiation of the cells with a black ray B-100A 360 nm long wave lamp. After 10 minutes of irradiation, approximately 90% of **CACM-HaXS**- and **MeNV-HaXS**-induced protein dimers were cleaved (figure 27, A). In contrast, dimers were not cleaved in cells treated with the photostable **HaXS8**. Interestingly, **CACM-HaXS**-induced dimers were approximately two times more efficiently cleaved, compared to **MeNV-HaXS**-induced dimers, when the irradiation time was reduced to 2 minutes (figure 27, B). This may be due to the increased absorption of coumarinyl compounds and reduced absorption of *o*-nitrobenzyl compounds in polar solvents[267, 268], and the higher quantum yield of **CACM-HaXS** displayed at early time points. Thus, **CACM-HaXS** can be used to covalently

dimerise Halo-POI and SNAP-POI fusion proteins, which can be cleaved with high temporal precision under mild conditions. Experiments comparing the photolysis of **CACM-HaXS**- and **MeNV-HaXS**-induced dimers are currently under investigation.



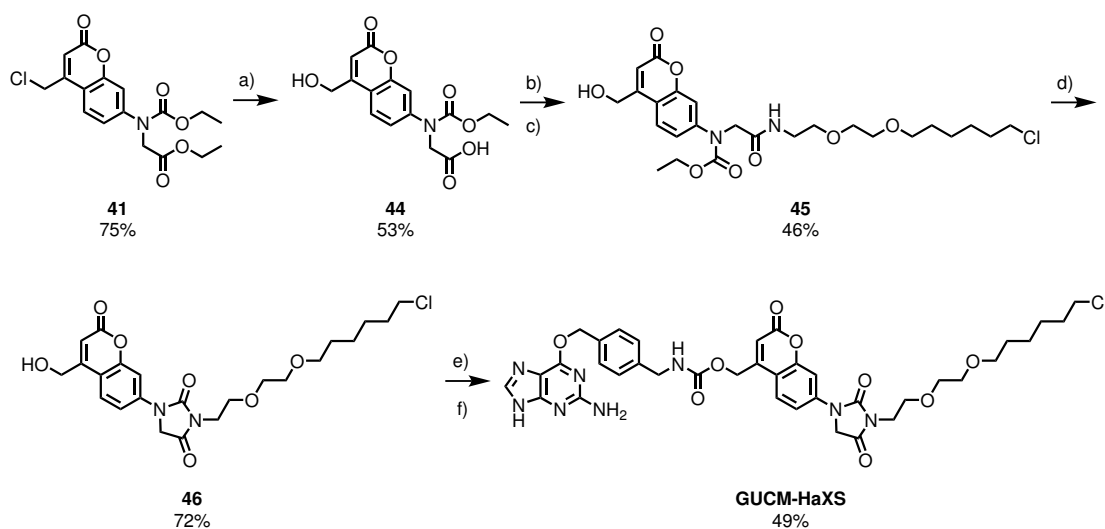
**Figure 27:** Comparison of the photocleavage of **MeNV-HaXS** and **CACM-HaXS** at 360 nm. A) HeLa cells expressing SNAP-GFP and Halo-GFP fusion proteins were treated with 5  $\mu$ M of **HaXS8**, **MeNV-HaXS** or **CACM-HaXS** for 15 minutes. Subsequently, cells were washed with PBS to remove unreacted compounds and illuminated with a high-intensity UV lamp (100 W, 5 cm distance) emitting light at 360 nm for 10 minutes A) or 2 minutes B) in 600  $\mu$ L PBS. Analysis of dimerisation products was performed as described previously. Values in B) represent means  $\pm$  SEM, n=3.

Given the high potential of **CACM-HaXS**, we attempted to increase the overall yield of the general building block **32a**. To avoid the allylic oxidation with  $\text{SeO}_2$ , we designed a similar route as for the planned synthesis of **ACM-HaXS**. However, this time we decided to substitute the amino group prior to the Pechman condensation (scheme 26). Therefore, we started the synthesis of **32a** with the alkylation 3-aminophenol (**30a**) with ethyl-2-iodoacetate under basic conditions, which gave **39** in good yields. We then deactivated the amino group with ethyl chloroformate under the same conditions as those previously described for **33a**. The acid catalysed cyclisation of **39** with ethyl 4-chloroacetoacetate generated the coumarine **40** with a good yield. In contrast to the Pechmann condensation procedure discussed above, chromatographic purification of **40** was necessary, as the product did not precipitate. Hydrolysis of the ethyl carbamate and the ethyl ester in a mixture of  $\text{H}_2\text{SO}_4$  and acetic acid at  $100^\circ\text{C}$  yielded **41** as a red solid. Unfortunately, the substitution of the chloride with a hydroxy group was not achieved using the same procedure as that described for the synthesis of **22**. Attempts to use basic conditions to substitute the chloride with the hydroxyl group, were not made, but rather we substituted the chloride of **40** with the hydroxyl group to give **42**. The saponification and cleavage of the protecting group with  $\text{H}_2\text{SO}_4$  and acetic acid, followed by basification to pH 5.3, led to the precipitation of **32a**, which was purified by recrystallisation in isopropanol. With this synthetic strategy, we were able to increase the overall yield of **32a** from 3.5% to 21%.



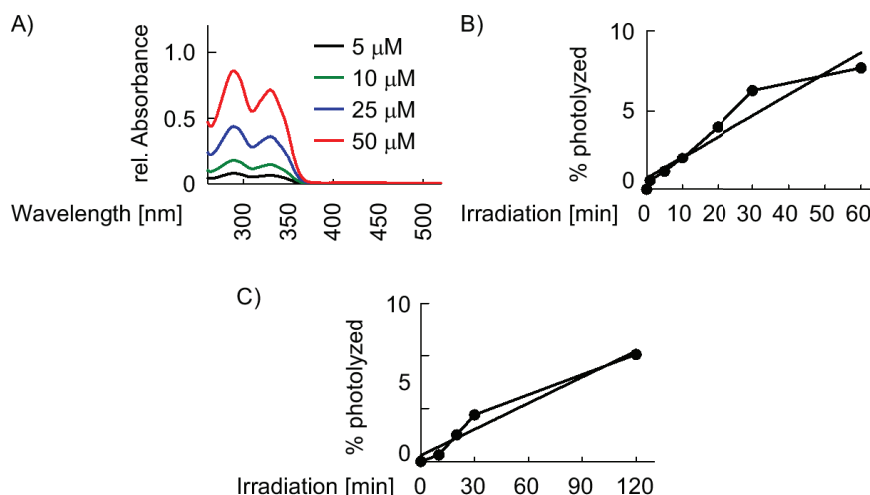
**Scheme 26:** Alternative synthetic route to the synthesis of **32a**. Reaction conditions: a) Ethyl-2-iodoacetate,  $\text{K}_2\text{CO}_3$ , DMF, r.t., 16 h; b) Ethyl chloroformate,  $\text{K}_2\text{CO}_3$ , DMF, r.t., 16 h; c) Ethyl 4-chloroacetoacetate,  $\text{MeHSO}_4$ ,  $60^\circ\text{C}$ , 16 h; d) Conc. sulfuric acid, acetic acid  $100^\circ\text{C}$ , 16 h; e)  $\text{H}_2\text{O}$ /Dioxane,  $100^\circ\text{C}$ , 16 h; f) Conc. sulfuric acid, acetic acid,  $100^\circ\text{C}$ , 16 h.

This route opened the path to the synthesis of an alternative coumarin containing an HaXS molecule (scheme 27). Therefore, we refluxed compound **41** for 3 days in water, rather than 16 h (see above), which resulted in the formation of compound **44**. Coupling of the **Halo** – **NH<sub>2</sub>** building block to compound **44** proceeded under various coupling conditions to create compound **45**. However, the best yields were obtained using DCC and NHS to activate the carboxylic acid, followed by coupling under mild basic conditions with **Halo** – **NH<sub>2</sub>**. Treatment of **45** under harsh basic conditions resulted in the formation of the glycolylurea derivative **46** with high yields. Attempts to carry out the reaction under milder reactive conditions failed. Finally, carbamate formation between **46** and **BG** – **NH<sub>2</sub>** using *p*-nitrophenyl carbonate generated **GUCM-HaXS** with good yields.



**Scheme 27:** Synthesis of **GUCM-HaXS**. Reaction conditions: a) H<sub>2</sub>O/Dioxane, 100°C, 3 days; b) DCC, NHS, DMF, r.t., 2 h; c) **Halo** – **NH<sub>2</sub>**, TEA, DMF, r.t., 16 h; d) NaOH, H<sub>2</sub>O, 16 h, 60°C; e) bis(*p*-nitrophenyl) carbonate, triethylamine, DMF, r.t., 16 h; f) **BG** – **NH<sub>2</sub>**, 60°C, 6 h.

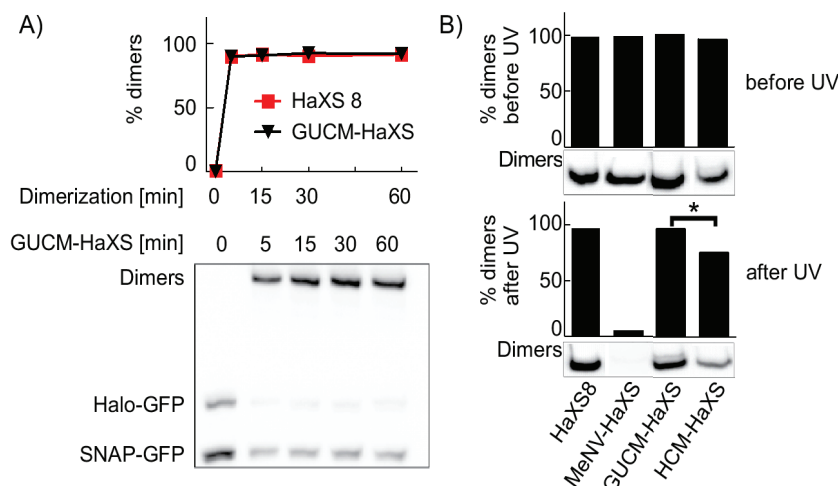
To investigate the absorption properties of **GUCM-HaXS**, we measured the UV-Vis spectra of different concentrated solutions in DMSO (figure 28, A). Not surprisingly, the electron withdrawing glycourea substituent at the C7 position of the coumarin core, blue shifted the absorption band to 330 nm and reduced the molar extinction coefficient at 360 (2346 M<sup>-1</sup> cm<sup>-1</sup>) and 405 nm (1313 M<sup>-1</sup> cm<sup>-1</sup>), compared to **CACM-HaXS**. The quantum yields at 360 nm (figure 28, B) and 405 nm (figure 28, C) were determined as described above. **GUCM-HaXS** exhibited quantum yields of 0.007 at 360 nm and 0.009 at 405 nm. However, the relative low absorption of **GUCM-HaXS** at these wavelength suggest that efficient cleavage of dimers may be difficult.



**Figure 28:** Photophysical properties of **GUCM-HaXS**. A) The UV/Vis spectra of different **GUCM-HaXS** solutions in DMSO were recorded at room temperature. B) A 0.5 mM **GUCM-HaXS** solution of in DMSO/water mixture (10:90) was irradiated for the indicated amount of time with an LED lamp, emitting  $5.44 \cdot 10^{-7}$  E/min at 360 nm. Subsequently, samples were subjected to a UPLC. Decrease of starting material was detected using a ESI-SQD mass spectrometer and plotted against time. C) A 0.5 mM **GUCM-HaXS** solution of in DMSO/water mixture (10:90) was irradiated for the indicated amount of time with an LED lamp, emitting  $5.44 \cdot 10^{-7}$  E/min at 405 nm. Samples were analysed as analogues to B).

We then investigated the time-dependent **GUCM-HaXS**-mediated dimerisation of SNAP-GFP and Halo-GFP. We treated HeLa cells, co-expressing SNAP-GFP and Halo-GFP, with 0.5 μM of the **GUCM-HaXS**, and **HaXS8**, for the indicated time period (figure 29). Like most of the photocleavable HaXS molecules described above, the dimerisation activity of **GUCM-HaXS** was comparable to **HaXS8** (figure 29, A). Both CIDs induced over 85% of fusion protein dimers after only 5 minutes of treatment, which were stable at ambient light, for a minimum of 60 minutes. To validate the *in vivo* photocleavage of **GUCM-HaXS**-induced fusion protein dimers, we irradiated cells containing the dimerised fusion proteins, with a black ray B-100A 360 nm long wave lamp and compared the results to the photolysis rate of **HaXS8**-, **MeNV-HaXS**-, and **HCM-HaXS**-induced dimers (figure 29, B). Surprisingly, fusion protein complexes which formed with **GUCM-HaXS** did not undergo photolysis after irradiation for 10 minutes. In contrast, dimers formed by **HCM-HaXS** were slightly cleaved, despite having an approximately 2-times lower molar extinction coefficient at 365 nm compared to **GUCM-HaXS**. No irregularities were observed in the control experiments, as **HaXS8**-induced protein complexes were photo insensitive and **MeNV-HaXS**-induced dimers were quantitatively cleaved. Based on the results of the *in*

*vivo* and *in vitro* photolysis experiments, we concluded that **GUCM-HaXS** can be used to induce the formation of a Halo-POI-SNAP-POI complex, which can be cleaved orthogonally to dimers formed by **MeNV-HaXS**.



**Figure 29:** *In vivo* dimerisation and cleavage efficacy of **GUCM-HaXS**. A) HeLa cells co-expressing SNAP-GFP and Halo-GFP fusion proteins were incubated with 5  $\mu$ M **HaXS8**, and **GUCM-HaXS** in cell culture medium at 37°C for indicated time, before cells were lysed and proteins were subjected to SDS PAGE, immunoblotting and chemiluminescence. B) HeLa cells expressing SNAP-GFP and Halo-GFP fusion proteins were treated with 5  $\mu$ M of **HaXS8**, **GUCM-HaXS**, **MeNV-HaXS** or **HCM-HaXS1** for 15 minutes. Subsequently, cells were washed with PBS to remove unreacted compounds and illuminated with a high-intensity UV lamp (100 W, 5 cm distance) emitting light at 360 nm for 10 minutes. Samples were analysed as described in A)

## 4 Conclusion and outlook

In the first part of this project, we successfully developed a novel Lewis acid-mediated method to synthesis C-16 modified rapalogs. The relatively mild reaction conditions allow the introduction of acid sensitive groups, such as different carbamates, ureas and thioureas. Notably, we were able to introduce a *tert*-butyl carbamate at the C16 position. We assumed that this rapalog could be applied in an extended selective modification of the rapamycin core after de-protection of the amine. The mechanism of the addition of carbamates onto the C16-carbocation appears to involve the prior establishment of a tight rapamycin-BF<sub>3</sub> ion pair, which exhibits reduced reactivity towards nucleophiles. Additionally, BF<sub>3</sub> also forms transient complexes with carbamate, reducing the nucleophilicity of the carbamate. The addition of THF subsequently leads to the irreversible formation of the desired rapalogs in high yields. The resulting C16 carbamyl diastereomers were readily separated by column chromatography and the absolute configuration of the C16 stereocenter was determined through NMR spectroscopic analysis. Moreover, we demonstrated that phenyl carbamate substitution can abolish potential interference with endogenous mTOR. However, extensive purification was required to remove traces of mTOR inhibiting byproducts and a method to quantitatively purify the product will be the focus of future studies. Nonetheless, **(R)-pcRap** is a powerful, non-toxic tool to rapidly induce protein-protein interactions. Future work will focus on the scale up of **(R)-pcRap** purification, as well as on its biological applications. The combination of **(R)-pcRap** with other orthogonal CID systems (e.g. HaXS molecules) would be an attractive method to simultaneously regulate multiple proteins, without off-target effects.

In the second part of this project, we synthesised a series of cell-permeable HaXS, which rapidly induced the formation of covalently linked and photocleavable protein dimers. Most of the synthesised photocleavable HaXS molecules induced significant amounts of dimers after a short period of time, highlighting the feasibility to modify the properties of HaXS molecules without decreasing the dimerisation activity. The covalent link formed between the photocleavable HaXS molecules and its target proteins, simplifies the analysis of protein complexes and prevents degradation of the complex prior to irradiation. The fast photolytic rates of the photocleavable HaXS molecules provides a choice between irradiation with a high-intensity UV lamp, a standard fluorescence microscope equipped with a mercury halide lamp or with a FRAP laser. Photocleavable HaXS molecules are therefore a novel asset to the CID toolbox, which combines intracellular protein dimerisation with photocleavage. Photocleavable HaXS molecules enable the possibility to sequester any protein of interest away from its functional compartment. Sequential photo-induced cleavage of the CID can release the anchored proteins and restore their function. Alternatively, photocleavable HaXS molecules can be used to reversibly switch the activity of proteins on or off. The choice between global and local illumination triggering dissociation of the CID-mediated complex permits the investigation of cell-compartment-associated signaling, and the simulation of cell-wide physiological and pathological signaling dynamics. Given the influence of the solvent on the photophysical properties of photocleavable groups, it might be interesting to investigate whether **MeNV-HaXS** and **CACM-HaXS** behave differently at various subcellular localisations. Combinations of cleavable HaXS molecules simultaneously with other orthogonal CIDs systems, would be an interesting approach to controlling multiple proteins. The combination of **MeNV-HaXS** or **CACM-HaXS** with other photo-activatable CIDs, such as **cTMP-Htag** would certainly be of specific interest. This would allow, either sequentially (in the case of **CACM-HaXS**) or simultaneously (in the case of **MeNV-HaXS**) to photo-induce the cleavage of a protein complex and independently trigger the formation of another protein-protein interaction. A combination of orthogonal photocleavable HaXS molecules could also be used to regulate multiple protein networks; however such experiments would require an elaborate

design. One possibility would be to control the expression of the fusion proteins and to treat the cells sequentially with **MeNV-HaXS** and **CACM-HaXS**. Alternatively, the intermediate **32a** presents a valuable precursor to tether different tag substrates, thus enabling the possibility to develop a series of photocleavable CIDs with orthogonal reactivities in two synthetic steps. Finally, the protection of the guanylyl nitrogen in **MeNV-HaXS** with a re-shifted coumarin-based photocage would give access to a highly powerful tool, which would allow the formation and the dissociation of a protein complex with light, to be controlled. Similar CIDs could be designed, using the aminocoumarinyl-4-methyl protected **TMP** or **BG-NH<sub>2</sub>**, linked via a nitrobenzyl linker to a second tag substrate.



## 5 Experimental Section

### 5.1 General Information

#### 5.1.1 Reagents and Solvents

Materials and reagents were of the highest commercially available grade and used as received from ABCR GmbH Co. KG (Karlsruhe, Germany), Acros Organics<sup>TM</sup> (Basel, Switzerland), Betapharma Co., Ltd (Shanghai, China), Fluka AG (Buchs, Switzerland), Fluorochem Ltd. (Hadfield, United Kingdom), or Sigma-Aldrich<sup>®</sup> (Buchs, Switzerland) unless stated otherwise. Dry solvents were purchased as anhydrous grade from Acros Organics or Fluka. Technical-grade solvents used for extractions and chromatographies were obtained from Brenntag AG (Basel, Switzerland) and were used without purification. Deuterated solvents for NMR spectroscopy were purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA, USA).

#### 5.1.2 Sensitive reactions

All reactions were performed at ambient light as none of the photocleavable molecules showed any sensitivity towards ambient light. Air- and water-sensitive reactions were performed under an inert atmosphere using nitrogen from PanGas AG (Darmstadt, Switzerland). Low temperature reaction were performed with an external cooling bath, containing isopropanol and dry ice.

### 5.2 Instrumentation

#### 5.2.1 Nuclear Magnetic Resonance (NMR) spectroscopy

All NMR spectra were recorded on a Bruker Avance III NMR spectrometer operating at 600 MHz (for <sup>1</sup>H-NMR), 151 MHz (<sup>13</sup>C-NMR), 565 MHz (<sup>19</sup>F-NMR) or 193 MHz (<sup>11</sup>B-NMR) equipped with a z-axis pulsed field gradient broadband direct detection probe-head, on a Bruker Ultra Shield Avance III with a BBO probe-head spectrometer operating at 500 MHz (for <sup>1</sup>H-NMR), 126 MHz (<sup>13</sup>C-NMR) or a Bruker Avance III 400 with a BFO<sup>+</sup> probe-head operating at 400 MHz (<sup>1</sup>H-NMR), 101 MHz (<sup>13</sup>C-NMR), 376 MHz (<sup>19</sup>F-NMR). Chemical shifts (δ) are reported in parts per million (ppm) relative to residual solvent peaks or trimethylsilane (TMS). Coupling constants (J) are reported in Hertz (Hz). NMR experiments to determine the structure of compounds were performed at 298 K. Experiments to determine the mechanism of the Lewis acid-mediated C16 substitution of rapamycin were performed at 243 K. For low temperature measurements the spectrometer was calibrated using a methanol sample. The multiplicities of the signals are described as: s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, h = hexet, hept = heptet and m = multiplet. For multiplets, the range of the chemical shift (ppm) of the signal is reported, for the other signals the centre of the signal is reported.

#### 5.2.2 Two-dimensional nuclear magnetic resonance (2D NMR) spectroscopy

<sup>1</sup>H-<sup>1</sup>H COSY, <sup>1</sup>H-<sup>1</sup>H NOESY, <sup>1</sup>H-<sup>1</sup>H ROESY, <sup>1</sup>H-<sup>1</sup>H TOSCY, <sup>1</sup>H-<sup>13</sup>C HMQC, and <sup>1</sup>H-<sup>13</sup>C HMBC experiments, required to assign <sup>1</sup>H- and <sup>13</sup>C-signal, determine the absolute configuration of rapalogs or for the mechanistic study, were performed on Bruker Avance III NMR spectrometer equipped with a z-axis pulsed field gradient broadband direct detection probe-head at 600 and 151 MHz, respectively, or on a Bruker Ultra Shield Avance III with a BBO probe head spectrometer at 500 and 126 MHz, respectively. Phase-sensitive ROESY experiments as well as TOCSY experiments were performed with 2048 time points in F2 and 1024 time increments in the indirect dimension F1, which corresponds to acquisition times of 155 ms in F2 and 77 ms in F1. The

Spinlock pulse for the ROESY experiment was set to 350 ms and the experiment times were between 3 and 18 hours. The mixing time for the TOCSY experiment was set to 200 ms and the experiment times were between 3 and 10 hours.

### 5.2.3 Mass Spectroscopy (MS)

Mass spectra were recorded with a fast atom bombardment (FAB) VG70-250 spectrometer, a electron ionization (EI) Finnigan MAT MS 312 spectrometer, a electrospray ionization (ESI) Finnigan MAT LSQ spectrometer, or a matrix assisted laser desorption ionization time of flight (MALDI-TOF) Bruker microflex<sup>TM</sup> spectrometer. High resolution mass spectra (**HRMS**) were recorded with a Thermo Fisher Scientific LTQ Orbitrap XL, nanoelectrospray ion source.

### 5.2.4 Thin-Layer Chromatography (TLC) and Preparative Thin-Layer Chromatography (PTLC)

TLC was performed using Merck silica gel 60 F<sub>254</sub> plates with a thickness of 0.25 mm. Compounds were visualised by UV lamps (253 or 366 nm), and with ceric ammonium molybdate (CAM), KMnO<sub>4</sub>, ninhydrin, or vanillin staining. For PTLC silica gel glass plates with a thickness of 2.0 mm from Merck KGaA were used.

### 5.2.5 High-Performance Liquid Chromatography (HPLC)

Normal-phase analytical HPLC measurements were performed on a Shimadzu LC-20AT instrument using a Reprospher 100 Si, 5 µm column (100 x 10 mm) from Maisch GmbH. Compounds were detected with a SPD-M20A diode array detector. Preparative normal-phase HPLC separations were performed on a Dionex P680 HPLC instrument using a Reprosil 100 Si, 5 µm column (250 x 40 mm) from Maisch GmbH and compounds were detected with a UV-Vis spectrometer.

### 5.2.6 Ultra Performance Liquid Chromatography (UPLC)

Reverse-phase UPLC measurements were performed on a Acquity H-Class UPLC system equipped with an ESI-SQD mass spectrometer, using a Acquity UPLC BEH column (2.1 x 50 mm) with a pore size of 1.7 µm.

### 5.2.7 Column chromatography

Column chromatography were performed using Silica gel 60 (40-63 µm) from Sigma Aldrich or SilicaFlash P60 (40-63 µm) from Silicycle®.

### 5.2.8 Ultraviolet/Visible (UV/Vis) absorption spectroscopy

UV/Vis spectra were recorded on a Perkin Elmer Lambda 40 UV/Vis spectrometer, or a NanoDrop 2000/2000c spectrophotometer, using a quartz standard absorption cells. All measurements were performed at room temperature and the wavelength of the absorption band ( $\lambda$ ) are reported in nm.

### 5.2.9 Photo reactors

Photolysis reactions were performed in a Lumos 43A equipped with a LED lamp or in a Transilluminator reactor equipped with 4 RPR-3500 Å, which emitted light at 360 nm or 405 nm.

### 5.2.10 Microscopes

Images were acquired on a Leica Live Imaging Microscope equipped with a HCX Plan Fluotar 63x/1.4 oil objective, a Photometrics CCD Camera CoolSnap HQ2 with Metamorph 7.1 software (Molecular devices) and a 50 mW FRAP scanning laser (UV-Diode Laser, 355 nm, 30 Hz) or a Leica EL600 microscope equipped with a mercury metal halide lamp (filter cube: ex: 377±25 nm, em: 447±30 nm, dichroic mirror: 409 nm; 63x objective: 63xHCX Plan-Fluotar, NA: 1.25 [oil]).

## 5.3 Experimental procedures

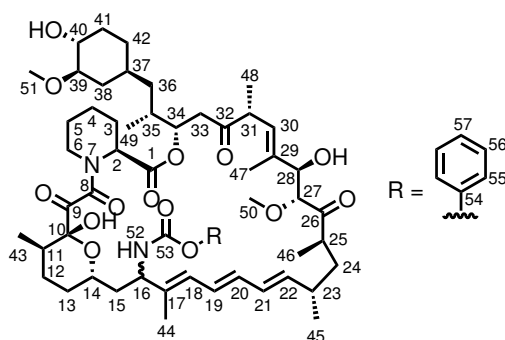
### 5.3.1 Synthetic procedures

#### 5.3.1.1 Synthesis of C16-carbamyl rapalog

**General Procedure 1 (GP1): Lewis acid-mediated substitution of the C16 methoxy group of rapamycin**

**Rapamycin** (0.10 mmol) and the indicated carbamate (0.60 mmol) were dissolved in CH<sub>2</sub>Cl<sub>2</sub> (20 mL). The reaction mixture was cooled to -40°C prior to the addition of BF<sub>3</sub>·Et<sub>2</sub>O (0.40 mmol). This solution was stirred for 5 min prior to the addition of THF (10 mL). After stirring the mixture for 30 min stirring at -40°C, the reaction was quenched with a saturated aqueous solution of NaHCO<sub>3</sub>. The aqueous layer was extracted three times with EtOAc, the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. To determine the amount of remaining rapamycin, an aliquot (1 mg) of the resulting white solid was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (1 mL) and subjected to HPLC analysis (injection volume 50 µL). The sample was eluted using a 60-min linear gradient of 100:0 to 96:4 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH) and a flow rate of 2 mL·min<sup>-1</sup>. The crude mixture was redissolved in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and cooled down to -40°C. BF<sub>3</sub>·Et<sub>2</sub>O (0.40 mmol) were added and the mixture was stirred for 5 min, prior to THF addition (10 mL). The mixture was stirred for 30 min at -40°C. The reaction was quenched with a saturated aqueous solution of NaHCO<sub>3</sub> and the aqueous layer were extracted three times with EtOAc. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. This reaction procedure was repeated 3-4 times until complete consumption of rapamycin (HPLC monitoring). The product was purified and the diastereomers were separated by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH).

#### Preparation of C16-phenyl carbamyl rapamycin (pcRap)



According to GP1, starting with **rapamycin** (100 mg, 0.11 mmol, 1 eq), phenyl carbamate

(90 mg, 0.66 mmol, 6 eq),  $\text{BF}_3\text{-Et}_2\text{O}$  (57  $\mu\text{L}$ , 0.44 mmol, 4 eq), column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  40:1 to 20:1) afforded 92 mg of **pcRap** as a mixture of two diastereomers (83% yield; 52:48 ratio). The product obtained after column chromatography was further purified by preparative HPLC using a 90-min linear gradient of 100:0 to 94:6 ( $\text{CH}_2\text{Cl}_2$ :methanol) and a flow rate of 30  $\text{mL}\cdot\text{min}^{-1}$ . Purity of the resulting fractions was determined using analytical HPLC. Therefore, 50  $\mu\text{L}$  of a 100  $\mu\text{M}$  solution of product in  $\text{CH}_2\text{Cl}_2$  were injected and eluted using a 60-min linear gradient of 100:0 to 96:4 ( $\text{CH}_2\text{Cl}_2$ :methanol) and a flow rate of 2  $\text{mL}\cdot\text{min}^{-1}$ .

Mixture of diastereomers:

**ESI-MS**  $m/z$ : 1041.57, 999.56, 864.53, 685.44, 607.39, 551.33, 412.17, 351.25, 279.02, 256.00, 189.96.

**HRMS** for  $\text{C}_{57}\text{H}_{82}\text{O}_{14}\text{N}_2$   $[\text{M}-\text{Na}]^+$ : calcd.: 1041.5658, found 1041.5664.

### (*R*)-pcRap:

Rt analytical: 40.3 min

Rt preparative: 37.9 min

**$^1\text{H-NMR}$**  (600MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 7.87 (d,  $J = 7.8$  Hz, 1H, H52), 7.37 (dd,  $J = 6.8$  Hz, 8.0 Hz, 2H, H56), 7.19 (t,  $J = 6.8$  Hz, 1H, H57), 7.09 (d,  $J = 8.0$  Hz, 2H, H55), 6.40 (dd,  $J = 10.8$  Hz, 11.4 Hz, 1H, H19), 6.39 (s, 1H, OH10), 6.21 (dd,  $J = 11.4$ , 12.4 Hz, 1H, H20), 6.13 (d,  $J = 10.6$  Hz, 12.4 Hz, 1H, H21), 6.08 (d,  $J = 10.8$  Hz, 1H, H18), 5.52 (dd,  $J = 9.4$  Hz, 12.4 Hz, 1H, H22), 5.24 (d,  $J = 4.4$  Hz, 1H, OH28), 5.22 (d,  $J = 9.8$  Hz, 1H, H30), 5.10 (m, 1H, H34), 4.96 (d,  $J = 5.3$  Hz, 1H, H2), 4.59 (d,  $J = 4.3$  Hz, 1H, OH40), 4.07 (m, 1H, H28), 4.02 (m, 1H, H16), 4.01 (m, 1H, H27), 3.95 (m, 1H, H14), 3.57 (m, 1H, H6), 3.31 (m, 3H, H51), 3.27 (m, 1H, H31), 3.18 (m, 3H, H50), 3.16 (m, 1H, H40), 3.06 (m, 1H, H6), 2.81 (m, 1H, H39), 2.81 (m, 1H, H33), 2.44 (m, 1H, H33), 2.39 (m, 1H, H25), 2.25 (m, 1H, H23), 2.16 (m, 1H, H3), 2.04 (m, 1H, H11), 1.91 (m, 1H, H38), 1.84 (m, 1H, H15), 1.75 (m, 1H, H41), 1.74 (m, 3H, H47), 1.72 (m, 3H, H44), 1.70 (m, 1H, H35), 1.67 (m, 1H, H4), 1.64 (m, 1H, H13), 1.61 (m, 1H, H12), 1.60 (m, 1H, H3), 1.55 (m, 1H, H12), 1.52 (m, 3H, H5, H15, H42), 1.40 (m, 1H, H24), 1.36 (m, 2H, H4, H5), 1.31 (m, 1H, H13), 1.27 (m, 1H, H37), 1.17 (m, 1H, H41), 1.10 (m, 1H, H36), 1.04 (m, 1H, H24), 0.99 (m, 3H, H45), 0.96 (m, 1H, H36), 0.91 (m, 3H, H48), 0.85 (m, 1H, H42), 0.81 (m, 3H, H46), 0.80 (m, 3H, H49), 0.76 (m, 3H, H43), 0.58 (m, 1H, H38).

**$^{13}\text{C-NMR}$**  (126 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 210.2 (C26), 208.0 (C32), 199.0 (C9), 169.4 (C1), 167.0 (C8), 153.8 (C53), 151.1 (C54), 140.0 (C17), 138.4 (C22), 136.9 (C29), 131.8 (C20), 130.4 (C21), 129.2 (C56), 127.6 (C19), 124.8 (C57), 124.6 (C30), 123.6 (C18), 121.6 (C55), 99.0 (C10), 85.5 (C27), 83.8 (C39), 75.7 (C28), 73.4 (C34), 73.2 (C40), 66.6 (C14), 56.9 (C50), 56.8 (C51), 53.7 (C16), 51.2 (C2), 44.8 (C31), 43.9 (C6), 40.4 (C33), 39.9 (C25), 39.6 (C24), 39.4 (C15), 38.3 (C36), 35.2 (C38), 34.7 (C11), 34.4 (C23), 33.5 (C35), 32.9 (C41), 32.6 (C37), 31.3 (C13, C42), 26.6 (C3), 26.4 (C12), 24.4 (C5), 21.5 (C45), 20.5 (C4), 16.0 (C48), 15.6 (C43), 15.2 (C44), 15.0 (C49), 13.8 (C47), 12.9 (C46).

### (*S*)-pcRap:

Rt analytical: 37.6 min

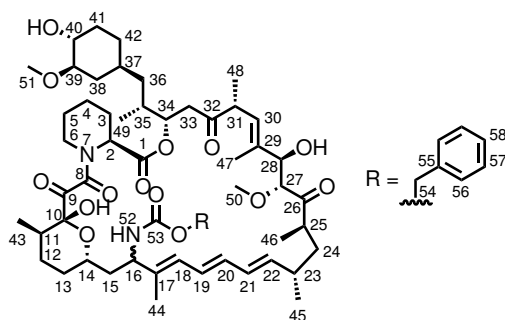
Rt preparative: 35.5 min

**$^1\text{H-NMR}$**  (600MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 8.07 (d,  $J = 8.8$  Hz, 1H, H52), 7.37 (dd,  $J = 7.38$  Hz, 7.68 Hz, 2H, H56), 7.19 (t,  $J = 7.38$  Hz, 1H, H57), 7.07 (d,  $J = 7.68$  Hz, 2H, H55), 6.51 (s, 1H, OH10), 6.36 (dd,  $J = 11.3$  Hz, 13.1 Hz, 1H, H19), 6.20 (dd,  $J = 10.8$  Hz, 13.1 Hz, 1H, H20), 6.12 (dd,  $J = 10.8$  Hz, 13.1 Hz, 1H, H21), 6.05 (d,  $J = 10.9$  Hz, 1H, H18), 5.48 (dd,  $J = 9.2$  Hz, 13.1 Hz, 1H, H22), 5.27 (d,  $J = 4.4$  Hz, 1H, OH28), 5.10 (d,  $J = 10.0$  Hz, 1H, H30), 4.99 (m, 1H, H34), 4.96 (d,  $J = 6.3$  Hz, 1H, H2), 4.60 (m, 1H, OH40), 4.13 (m, 1H, H16), 4.01 (m, 1H, H28), 3.94

(m, 1H, H14), 3.87 (d,  $J = 5.0$  Hz, 1H, H27), 3.45 (m, 1H, H6), 3.33 (m, 1H, H31), 3.32 (m, 3H, H51), 3.25 (m, 1H, H6), 3.17 (m, 1H, H40), 3.16 (m, 3H, H50), 2.83 (m, 1H, H39), 2.76 (m, 1H, H33), 2.49 (m, 1H, H25), 2.44 (m, 1H, H33), 2.22 (m, 1H, H23), 2.08 (m, 1H, H3), 2.05 (m, 1H, H11), 1.96 (m, 1H, H15), 1.92 (m, 1H, H38), 1.82 (m, 1H, H13), 1.75 (m, 6H, H41, H42, H44), 1.74 (m, 3H, H47), 1.70 (m, 1H, H35), 1.67 (m, 1H, H4), 1.63 (m, 1H, H3), 1.58 (m, 1H, H5), 1.54 (m, 1H, H12), 1.52 (m, 1H, H12), 1.43 (m, 1H, H15), 1.41 (m, 2H, H4, H24), 1.31 (m, 1H, H5), 1.28 (m, 1H, H37), 1.25 (m, 1H, H13), 1.18 (m, 1H, H41), 1.07 (m, 1H, H36), 1.06 (m, 1H, H24), 0.98 (m, 3H, H45), 0.97 (m, 1H, H36), 0.89 (m, 3H, H48), 0.86 (m, 1H, H42), 0.85 (m, 3H, H46), 0.79 (m, 3H, H49), 0.76 (m, 3H, H43), 0.60 (m, 1H, H38).

**$^{13}\text{C}$ -NMR** (126 MHz, DMSO –  $d_6$ )  $\delta$ : 210.9 (C26), 207.7 (C32), 198.9 (C9), 169.2 (C1), 166.8 (C8), 154.2 (C53), 151.1 (C54), 138.9 (C22), 138.2 (C17), 132.0 (C20), 130.4 (C21), 129.3 (C56), 127.3 (C19), 125.5 (C18), 124.9 (C57), 125.3 (C30), 121.7 (C55), 99.0 (C10), 85.9 (C27), 83.8 (C39), 75.8 (C28), 73.7 (C34), 73.2 (C40), 66.8 (C14), 57.1 (C50), 56.7 (C51), 54.9 (C16), 50.6 (C2), 45.3 (C31), 43.4 (C6), 39.7 (C33), 39.6 (C24), 39.2 (C25), 38.4 (C15, C36), 37.4 (C29), 35.4 (C38), 34.9 (C11, C23), 33.3 (C35), 32.9 (C41), 32.6 (C37), 31.2 (C42), 29.6 (C13), 26.4 (C3), 26.2 (C12), 24.4 (C5), 21.6 (C45), 20.4 (C4), 15.6 (C43), 15.5 (C48), 14.7 (C49), 13.7 (C46), 13.2 (C47), 11.9 (C44).

**Preparation of C16-benzyl carbamyl rapamycin (bnRap)** According to GP1, starting



with **rapamycin** (100 mg, 0.11 mmol, 1 eq), benzyl carbamate (100 mg, 0.66 mmol, 6 eq),  $\text{BF}_3\text{-Et}_2\text{O}$  (57  $\mu\text{L}$ , 0.44 mmol, 4 eq), column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  40:1 to 20:1) afforded 99 mg of X as a mixture of two diastereomers (88% yield; 51:49 ratio).

Mixture of diastereomers:

**ESI-MS**  $m/z$ : 1055.58, 1033.60, 1015.59, 882.54, 864.52, 846.51, 814.49, 793.56, 663.45, 607.39, 496.27, 411.17, 351.25, 309.2.

**HRMS** for  $\text{C}_{57}\text{H}_{82}\text{O}_{14}\text{N}_2$   $[M]$ : calcd: 1033.5995, found 1033.5992.

**(R)-benzyl carbamyl rapamycin:**

**$^1\text{H}$ -NMR** (600MHz, DMSO –  $d_6$ )  $\delta$ : 7.45 (d,  $J = 7.7$  Hz, 1H, H52), 7.38 (m, 2H, H57), 7.35 (m, 1H, H58), 7.08 (m, 2H, H56), 6.37 (dd,  $J = 10.8$  Hz, 12.4 Hz, 1H, H19), 6.34 (s, 1H, OH10), 6.17 (dd,  $J = 10.8$  Hz, 12.4 Hz, 1H, H20), 6.13 (dd,  $J = 10.8$  Hz, 12.0 Hz, 1H, H21), 6.02 (d,  $J = 10.8$  Hz, 1H, H18), 5.50 (dd,  $J = 8.8$  Hz, 12.0 Hz, 1H, H22), 5.23 (m, 1H, OH28), 5.21 (m, 1H, H30), 5.08 (m, 1H, H34), 5.01 (m, 2H, H54), 4.95 (d,  $J = 5.3$  Hz, 1H, H2), 4.59 (d,  $J = 4.65$  Hz, 1H, OH40), 4.07 (m, 1H, H28), 3.95 (m, 1H, H16), 3.88 (m, 1H, H27), 3.85 (m, 1H, H14), 3.57 (m, 1H, H6), 3.30 (m, 3H, H51), 3.27 (m, 1H, H31), 3.18 (m, 3H, H50), 3.15 (m, 1H, H40), 3.03

(m, 1H, H6), 2.82 (m, 1H, H39), 2.79 (m, 1H, H33), 2.50 (m, 2H, H33, H25), 2.42 (m, 1H, H23), 2.15 (m, 1H, H3), 2.02 (m, 1H, H11), 1.90 (m, 1H, H38), 1.76 (m, 1H, H41), 1.73 (m, 3H, H47), 1.72 (m, 1H, H15), 1.69 (m, 1H, H35), 1.68 (m, 1H, H4), 1.65 (m, 3H, H44), 1.59 (m, 1H, H3), 1.58 (m, 1H, H12), 1.56 (m, 2H, H13, H42), 1.52 (m, 1H, H12), 1.51 (m, 1H, H5), 1.40 (m, 2H, H15, H24), 1.39 (m, 1H, H5), 1.38 (m, 1H, H4), 1.29 (m, 1H, H13), 1.27 (m, 1H, H41), 1.26 (m, 1H, H37), 1.09 (m, 1H, H36), 1.03 (m, 1H, H24), 0.99 (m, 3H, H45), 0.95 (m, 1H, H36), 0.90 (m, 3H, H48), 0.86 (m, 3H, H46), 0.85 (m, 1H, H42), 0.79 (m, 3H, H49), 0.73 (m, 3H, H43), 0.57 (m, 1H, H38).

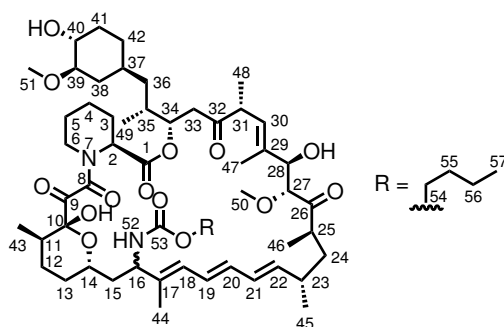
**<sup>13</sup>C-NMR** (126 MHz, DMSO – d<sub>6</sub>) δ: 210.2 (C26), 207.9 (C32), 199.1 (C9), 169.4 (C1), 167.0 (C8), 156.3 (C55), 155.7 (C53), 137.2 (C17), 138.3 (C22), 137.2 (C29), 131.5 (C20), 130.4 (C21), 128.3 (C57), 127.7 (C58), 127.6 (C19), 127.4 (C56), 124.4 (C30), 123.3 (C18), 99.0 (C10), 85.9 (C27), 83.8 (C39), 75.8 (C28), 73.4 (C34), 73.2 (C40), 67.0 (C14), 65.2 (C54), 56.9 (C50), 56.8 (C51), 53.7 (C16), 51.3 (C2), 44.9 (C31), 43.8 (C6), 40.4 (C33), 39.9 (C25), 39.6 (C24), 39.5 (C15), 38.4 (C36), 35.4 (C38), 34.7 (C11), 34.4 (C23), 33.5 (C35), 33.0 (C41), 32.6 (C37), 31.8 (C13), 31.2 (C42), 26.4 (C3, C12), 24.5 (C5), 21.5 (C45), 20.5 (C4), 16.0 (C48), 15.6 (C43), 15.0 (C49), 14.1 (C46), 14.0 (C47), 12.0 (C44).

**(S)-benzyl carbamyl rapamycin:**

**<sup>1</sup>H-NMR** (600MHz, DMSO – d<sub>6</sub>) δ: 7.58 (d, J = 9.0 Hz, H52, 1H), 7.34 (m, 2H, H57), 7.32 (m, 1H, H58) 7.07 (m, 2H, H56), 6.45 (s, 1H, OH10), 6.33 (dd, J = 10.8 Hz, 12.4 Hz, 1H, H19), 6.19 (dd, J = 10.2 Hz, 12.4 Hz, 1H, H20), 6.12 (dd, J = 10.2 Hz, 12.8 Hz, 1H, H21), 6.01 (d, J = 10.8 Hz, 1H, H18), 5.49 (dd, J = 8.8 Hz, 12.8 Hz, 1H, H22), 5.28 (m, 1H, OH28), 5.10 (m, 1H, H30), 5.03 (m, 1H, H54), 4.98 (m, 1H, H34), 4.95 (d, J = 4.9 Hz, 1H, H2), 4.59 (d, J = 4.38 Hz, 1H, OH40), 4.11 (m, 1H, H16), 4.01 (m, 1H, H28), 3.87 (m, 1H, H27), 3.86 (m, 1H, H14), 3.42 (m, 1H, H6), 3.32 (m, 4H, H31, H51), 3.21 (m, 1H, H6), 3.18 (m, 1H, H40), 3.16 (m, 3H, H50), 2.83 (m, 1H, H39), 2.76 (m, 1H, H33), 2.50 (m, 2H, H33, H25), 2.23 (m, 1H, H23), 2.08 (m, 1H, H3), 2.02 (m, 1H, H11), 1.92 (m, 1H, H38), 1.88 (m, 1H, H15), 1.76 (m, 1H, H41), 1.74 (m, 3H, H47), 1.69 (m, 1H, H35), 1.68 (m, 3H, H44), 1.65 (m, 1H, H4), 1.62 (m, 1H, H3), 1.54 (m, 3H, H5, H12, H13), 1.53 (m, 1H, H42), 1.45 (m, 1H, H12), 1.39 (m, 1H, H4), 1.37 (m, 1H, H15), 1.36 (m, 1H, H24), 1.31 (m, 1H, H5), 1.30 (m, 1H, H37), 1.26 (m, 2H, H13, H41), 1.12 (m, 1H, H36), 1.08 (m, 1H, H24), 0.99 (m, 3H, H45), 0.96 (m, 1H, H36), 0.89 (m, 3H, H48), 0.86 (m, 1H, H42), 0.85 (m, 3H, H46), 0.79 (m, 3H, H49), 0.75 (m, 3H, H43), 0.60 (m, 1H, H38).

**<sup>13</sup>C-NMR** (126 MHz, DMSO – d<sub>6</sub>) δ: 210.8 (C26), 207.7 (C32), 198.8 (C9), 169.2 (C1), 166.9 (C8), 156.4 (C55), 155.6 (C53), 139.1 (C17), 138.7 (C22), 137.2 (C29), 131.8 (C20), 130.5 (C21), 128.3 (C57), 127.7 (C58), 127.6 (C19), 127.4 (C56), 125.1 (C30), 125.0 (C18), 98.8 (C10), 85.9 (C27), 83.9 (C39), 75.9 (C28), 73.8 (C34), 73.3 (C40), 67.0 (C14), 65.2 (C54), 57.2 (C50), 56.8 (C51), 53.4 (C16), 50.8 (C2), 45.4 (C31), 43.6 (C6), 39.7 (C24), 39.6 (C33), 39.3 (C25), 38.7 (C15), 38.3 (C36), 35.5 (C38), 34.9 (C11, C23), 33.4 (C35), 33.0 (C41), 32.8 (C37), 31.3 (C13, C42), 26.4 (C3, C12), 24.5 (C5), 21.7 (C45), 20.5 (C4), 15.7 (C43), 15.5 (C48), 14.9 (C49), 13.7 (C46), 13.4 (C47), 12.1 (C44).

**Preparation of C16-butyl carbamyl rapamycin (buRap)** According to GP1, starting with



**rapamycin** (100 mg, 0.11 mmol, 1 eq), butyl carbamate (77 mg, 0.66 mmol, 6 eq),  $\text{BF}_3\text{-Et}_2\text{O}$  (57  $\mu\text{L}$ , 0.44 mmol, 4 eq), column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  40:1 to 20:1) afforded 79 mg of X as a mixture of two diastereomers (73% yield; 51:49 ratio).

Mixture of diastereomers:

**ESI-MS**  $m/z$ : 1141.83, 1097.80, 1021.59, 979.59, 699.38, 619.44, 351.25, 279.02, 208.13.

**HRMS** for  $\text{C}_{55}\text{H}_{86}\text{O}_{14}\text{N}_2$   $[\text{M}-\text{Na}]^+$ : calcd.: 1021.5971, found 1021.5945.

**(R)-butyl carbamyl rapamycin:**

**$^1\text{H-NMR}$**  (600MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 7.43 (d,  $J = 7.8$  Hz, 1H, H52), 6.36 (dd,  $J = 10.7$  Hz, 12.3 Hz, 1H, H19), 6.34 (s, 1H, OH10), 6.17 (dd,  $J = 11.0$  Hz, 12.3 Hz, 1H, H20), 6.13 (dd,  $J = 11.0$  Hz, 12.3 Hz, 1H, H21), 6.01 (d,  $J = 10.7$  Hz, 1H, H18), 5.50 (dd,  $J = 9.8$  Hz, 12.3 Hz, 1H, H22), 5.23 (d,  $J = 4.1$  Hz, 1H, OH28), 5.22 (d,  $J = 11.1$  Hz, 1H, H30), 5.09 (m, 1H, H34), 4.95 (d,  $J = 5.1$  Hz, 1H, H2), 4.59 (d,  $J = 4.53$  Hz, 1H, OH40), 4.07 (m, 1H, H28), 4.03 (m, 1H, H27), 3.95 (m, 1H, H16), 3.93 (m, 2H, H54), 3.86 (m, 1H, H14), 3.57 (m, 1H, H6), 3.30 (m, 3H, H51), 3.26 (m, 1H, H31), 3.18 (m, 3H, H50), 3.16 (m, 1H, H40), 3.03 (m, 1H, H6), 2.81 (m, 1H, H39), 2.78 (m, 1H, H33), 2.50 (m, 1H, H33), 2.35 (m, 1H, H25), 2.25 (m, 1H, H23), 2.16 (m, 1H, H3), 2.02 (m, 1H, H11), 1.91 (m, 1H, H38), 1.76 (m, 1H, H41), 1.73 (m, 3H, H47), 1.72 (m, 3H, H15, H55), 1.69 (m, 1H, H35), 1.68 (m, 1H, H4), 1.67 (m, 3H, H44), 1.61 (m, 1H, H3), 1.57 (m, 1H, H13), 1.53 (m, 3H, H5, H12), 1.52 (m, 3H, H42, H55), 1.45 (m, 1H, H15), 1.39 (m, 1H, H24), 1.38 (m, 1H, H4), 1.37 (m, 1H, H5), 1.29 (m, 2H, H56), 1.26 (m, 1H, H37), 1.25 (m, 1H, H13), 1.17 (m, 1H, H41), 1.09 (m, 1H, H36), 1.03 (m, 1H, H24), 0.99 (m, 3H, H45), 0.95 (m, 1H, H36), 0.90 (m, 3H, H48), 0.85 (m, 4H, H42, H57), 0.81 (m, 3H, H46), 0.80 (m, 3H, H49), 0.74 (m, 3H, H43), 0.57 (m, 1H, H38).

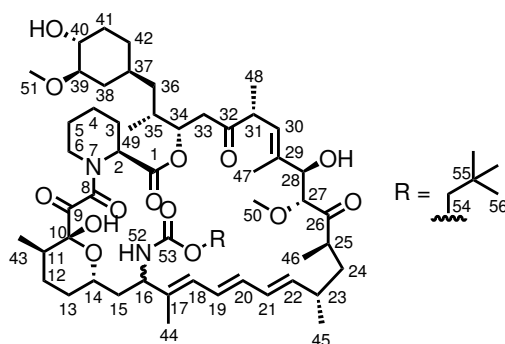
**$^{13}\text{C-NMR}$**  (126 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 209.9 (C26), 207.8 (C32), 198.8 (C9), 169.3 (C1), 167.0 (C8), 155.6 (C53), 140.4 (C17), 138.0 (C22), 136.7 (C29), 130.2 (C21), 127.4 (C19), 124.1 (C30), 123.0 (C18), 99.0 (C10), 85.2 (C27), 83.6 (C39), 75.5 (C28), 73.2 (C34), 73.0 (C40), 66.8 (C14), 65.4 (C54), 56.6 (C50), 56.5 (C51), 53.5 (C16), 50.8 (C2), 44.6 (C31), 43.6 (C6), 39.8 (C25), 39.6 (C33), 39.3 (C15, C24), 38.1 (C36), 35.0 (C38), 34.4 (C11), 34.2 (C23), 33.2 (C35), 32.8 (C41), 32.7 (C55), 32.4 (C37), 31.0 (C13, C42), 26.2 (C3, C12), 24.2 (C5), 22.3 (C56), 21.3 (C45), 20.3 (C4), 15.8 (C48), 15.4 (C43), 15.0 (C44), 14.8 (C49), 14.7 (C57), 13.7 (C47), 12.6 (C46).

**(S)-butyl carbamyl rapamycin:**

**<sup>1</sup>H-NMR** (600MHz, DMSO – d<sub>6</sub>) δ: 7.25 (d, J = 8.8 Hz, 1H, H52), 6.44 (s, 1H, OH10), 6.33 (dd, J = 10.9 Hz, 12.4, 1H, H19), 6.19 (dd, J = 10.4 Hz, 12.4, 1H, H20), 6.12 (dd, J = 10.4 Hz, 11.3 Hz, 1H, H21), 6.02 (d, J = 11.3 Hz, 1H, H18), 5.48 (dd, J = 9.5 Hz, 12.4 Hz, 1H, H22), 5.10 (s, 1H, OH28), 5.09 (m, 1H, H30), 4.99 (m, 1H, H34), 4.96 (m, 1H, H2), 4.59 (d, J = 4.37 Hz, 1H, OH40), 4.10 (m, 1H, H16), 4.01 (d, J = 5.2 Hz, 1H, H28), 3.93 (m, 2H, H54), 3.88 (d, J = 5.2Hz, 1H, H27), 3.83 (M, 1H, H14), 3.44 (m, 1H, H6), 3.32 (m, 3H, H51), 3.31 (m, 1H, H31), 3.23 (m, 1H, H6), 3.17 (m, 1H, H40), 3.14 (s, 3H, H50), 2.81 (m, 1H, H39), 2.75 (m, 1H, H33), 2.49 (m, 1H, H25), 2.41 (m, 1H, H33), 2.22 (m, 1H, H23), 2.09 (m, 1H, H3), 2.02 (m, 1H, H11), 1.91 (m, 1H, H38), 1.88 (m, 1H, H15), 1.79 (m, 1H, H13), 1.76 (m, 2H, H4, H41), 1.74 (m, 4H, H35, H47), 1.73 (m, 3H, H44), 1.55 (m, 1H, H5), 1.53 (m, 2H, H42, H3), 1.52 (m, 2H, H55), 1.47 (m, 1H, H12), 1.44 (m, 1H, H12), 1.41 (m, 1H, H24), 1.34 (m, 1H, H15), 1.30 (m, 2H, H4, H5), 1.28 (m, 1H, H37), 1.26 (m, 2H, H56), 1.24 (m, 1H, H13), 1.16 (m, 1H, H41), 1.09 (m, 1H, H36), 1.06 (m, 1H, H24), 0.99 (m, 3H, H45), 0.96 (m, 1H, H36), 0.89 (m, 3H, H48), 0.86 (m, 3H, H57), 0.85 (m, 1H, H42), 0.84 (m, 3H, H46), 0.79 (m, 3H, H49), 0.75 (m, 3H, H43), 0.60 (m, 1H, H38).

**<sup>13</sup>C-NMR** (126 MHz, DMSO – d<sub>6</sub>) δ: 210.8 (C26), 207.9 (C32), 198.8 (C9), 169.1 (C1), 166.4 (C8), 155.5 (C53), 138.5 (C22), 137.5 (C17), 131.7 (C20), 130.2 (C21), 127.1 (C19), 125.0 (C30), 124.9 (C18), 99.0 (C10), 85.5 (C27), 83.6 (C39), 75.5 (C28), 73.5 (C34), 72.9 (C40), 67.0 (C14), 65.4 (C54), 57.3 (C50), 56.8 (C16), 56.5 (C51), 50.4 (C2), 45.0 (C31), 43.2 (C6), 39.5 (C25, C33), 39.3 (C24), 38.2 (C15, C36), 35.1 (C38), 34.6 (C11, C23), 32.7 (C55), 32.6 (C35, C41), 31.2 (C42), 30.9 (C37), 29.6 (C5), 28.7 (C13), 26.0 (C3, C55), 25.9 (C12), 22.3 (C56), 21.4 (C45), 20.3 (C4), 15.4 (C43), 15.3 (C48), 14.7 (C57), 14.5 (C49), 13.5 (C46), 13.0 (C44), 12.9 (C47).

**Preparation of C16-*tert*-butyl carbamyl rapamycin (<sup>t</sup>buRap)** According to GP1,



starting with **rapamycin** (100 mg, 0.11 mmol, 1 eq), *tert*-butyl carbamate (77 mg, 0.66 mmol, 6 eq), BF<sub>3</sub>·Et<sub>2</sub>O (57 μL, 0.44 mmol, 4 eq), column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 40:1 to 20:1) afforded 85 mg of X as a mixture of two diastereomers (78% yield; 51:49 ratio).

Mixture of diastereomers:

**ESI-MS** m/z: 1021.60, 999.62, 897.55, 879.54, 846.51, 412.17, 351.25.

**HRMS** for C<sub>55</sub>H<sub>86</sub>O<sub>14</sub>N<sub>2</sub> [M-Na]<sup>+</sup>: calcd.: 1021.5971, found 1021.5865.

**(R)-C16-*tert*-butyl carbamyl rapamycin:**

**<sup>1</sup>H-NMR** (600MHz, DMSO – d<sub>6</sub>) δ: 6.98 (d, J = 9.0 Hz, 1H, H52). 6.42 (s, 1H, OH10), 6.33



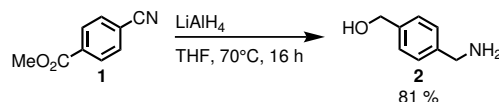
(dd,  $J = 11.1$  Hz, 12.4 Hz, 1H, H19), 6.18 (dd,  $J = 11.9$  Hz, 12.4 Hz, 1H, H20), 6.11 (dd,  $J = 11.9$  Hz, 12.4 Hz, 1H, H21), 5.97 (d,  $J = 11.1$  Hz, 1H, H18), 5.49 (dd,  $J = 9.1$  Hz, 12.4 Hz, 1H, H22), 5.11 (m, 2H, OH28, H30), 5.08 (m, 1H, H34), 4.95 (d,  $J = 4.9$  Hz, 1H, H2), 4.52 (m, 1H, OH40), 4.08 (m, 1H, H16), 4.07 (m, 1H, H28), 4.05 (m, 1H, H27), 3.83 (m, 1H, H14), 3.55 (m, 1H, H6), 3.31 (m, 3H, H51), 3.26 (m, 1H, H31), 3.18 (m, 3H, H50), 3.17 (m, 1H, H40), 3.00 (m, 1H, H6), 2.81 (m, 1H, H39), 2.77 (m, 1H, H33), 2.50 (m, 1H, H25), 2.43 (m, 1H, H33), 2.23 (m, 1H, H23), 2.11 (m, 1H, H3), 1.92 (m, 1H, H11), 1.91 (m, 1H, H38), 1.75 (m, 1H, H41), 1.74 (m, 3H, H47), 1.70 (m, 1H, H15), 1.69 (m, 1H, H35), 1.67 (m, 1H, H4), 1.66 (m, 3H, H44), 1.60 (m, 1H, H3), 1.57 (m, 1H, H42), 1.52 (m, 1H, H12), 1.50 (m, 2H, H13, H5), 1.42 (m, 1H, H15), 1.40 (m, 2H, H24, H12), 1.37 (m, 5H, H5, H4, H55), 1.26 (m, 2H, H37, H13), 1.25 (m, 1H, H41), 1.09 (m, 1H, H36), 1.03 (m, 1H, H24), 0.98 (m, 3H, H45), 0.95 (m, 1H, H36), 0.90 (m, 3H, H48), 0.86 (m, 4H, H46, H42), 0.79 (m, 3H, H49), 0.73 (m, 3H, H43), 0.59 (m, 1H, H38).

**<sup>13</sup>C-NMR** (126 MHz, DMSO – d<sub>6</sub>)  $\delta$ : 210.1 (C26), 207.7 (C32), 198.8 (C9), 169.2 (C1), 166.9 (C8), 155.0 (C53), 139.1 (C17), 138.7 (C22), 137.2 (C29), 131.8 (C20), 130.5 (C21), 127.6 (C19), 125.1 (C30), 125.0 (C18), 98.8 (C10), 85.4 (C27), 83.9 (C39), 77.6 (C54), 75.9 (C28), 73.8 (C34), 73.3 (C40), 67.0 (C14), 57.2 (C50), 56.8 (C51), 53.4 (C16), 50.8 (C2), 45.4 (C31), 43.6 (C6), 39.7 (C24), 39.6 (C33), 39.3 (C25), 38.7 (C15), 38.4 (C36), 35.5 (C38), 34.9 (C11, C23), 33.4 (C35), 33.0 (C41), 32.8 (C37), 31.3 (C13, C42), 28.3 (C55), 26.4 (C3, C12), 24.5 (C5), 21.7 (C45), 20.5 (C4), 15.7 (C43), 15.5 (C48), 14.9 (C49), 13.7 (C46), 13.4 (C47), 12.1 (C44).

**(S)-C16-*tert*.-butyl carbamyl rapamycin:**

**<sup>1</sup>H-NMR** (600MHz, DMSO – d<sub>6</sub>)  $\delta$ : 7.07 (d,  $J = 9.0$  Hz, 1H, H52). 6.39 (s, 1H, OH10) 6.36 (m, 1H, H19), 6.15 (m, 1H, H20), 6.12 (m, 12.4 Hz, 1H, H21), 5.99 (m, 1H, H18), 5.51 (dd,  $J = 9.1$  Hz, 12.4 Hz, 1H, H22), 5.22 (s, 1H, OH28), 5.20 (m, 1H, H30), 4.98 (m, 1H, H34), 4.95 (d,  $J = 5.60$  Hz, 1H, H2), 4.04 (m, 1H, H16), 4.02 (m, 1H, H27), 3.89 (m, 1H, H28), 3.84 (m, 1H, H14), 3.44 (m, 1H, H6), 3.32 (m, 3H, H51), 3.30 (m, 1H, H31), 3.20 (m, 1H, H6), 3.17 (m, 1H, H40), 3.15 (m, 3H, H50), 2.83 (m, 1H, H39), 2.75 (m, 1H, H33), 2.50 (m, 1H, H25), 2.43 (m, 1H, H33), 2.23 (m, 1H, H23), 2.11 (m, 1H, H3), 2.01 (m, 1H, H11), 1.91 (m, 1H, H38), 1.83 (m, 1H, H15), 1.76 (m, 1H, H35), 1.74 (m, 1H, H41), 1.73 (m, 3H, H47), 1.67 (m, 1H, H4), 1.66 (m, 3H, H44), 1.55 (m, 2H, H5, H13), 1.54 (m, 1H, H12), 1.53 (m, 1H, H42), 1.52 (m, 2H, H12), 1.51 (m, 1H, H3), 1.40 (m, 1H, H24), 1.39 (m, 1H, H4), 1.37 (m, 9H, H55), 1.32 (m, 2H, H5, H15), 1.31 (m, 1H, H13), 1.28 (m, 1H, H37), 1.18 (m, 1H, H41), 1.17 (m, 1H, H36), 1.06 (m, 1H, H24), 0.97 (m, 3H, H45), 0.89 (m, 3H, H48), 0.86 (m, 1H, H36), 0.85 (m, 1H, H42), 0.83 (m, 3H, H46), 0.78 (m, 3H, H49), 0.74 (m, 3H, H43), 0.60 (m, 1H, H38).

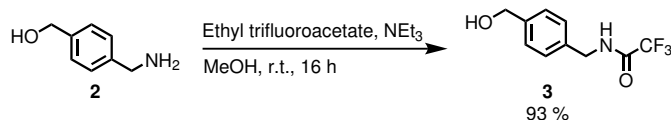
**<sup>13</sup>C-NMR** (126 MHz, DMSO – d<sub>6</sub>)  $\delta$ : 209.9 (C26), 207.9 (C32), 198.8 (C9), 169.5 (C1), 167.0 (C8), 155.2 (C53), 138.2 (C22), 137.8 (C29), 136.8 (C17), 131.4 (C20), 130.5 (C21), 127.8 (C19), 124.4 (C30), 123.1 (C18), 99.0 (C10), 85.4 (C27), 83.9 (C39), 77.7 (C54), 75.8 (C28), 73.4 (C34), 73.2 (C40), 67.3 (C14), 56.9 (C50), 56.8 (C51), 53.7 (C16), 51.3 (C2), 44.9 (C31), 43.7 (C6), 40.0 (C25), 39.8 (C15), 39.6 (C33), 39.5 (C24), 38.7 (C36), 35.5 (C38), 35.4 (C11), 34.5 (C23), 33.5 (C35), 33.0 (C41), 31.8 (C37), 31.3 (C42), 31.2 (C13), 28.3 (C55), 26.5 (C3, C12), 24.5 (C5), 21.6 (C45), 20.6 (C4), 16.0 (C48), 15.6 (C43), 15.1 (C49), 14.1 (C47), 14.0 (C46), 12.1 (C44).

5.3.1.2 Synthesis of BG-NH<sub>2</sub>[48]Preparation of (4-(aminomethyl)phenyl)methanol (**2**)

LiAlH<sub>4</sub> (10.4 g, 273 mmol, 3.0 eq) was added portionwise to a solution of methyl-4-cyanobenzoate **1** (14.7 g, 91 mmol, 1.0 eq) in THF (500 mL) over a period of 2 h. The mixture was stirred at 70°C for 16 h. After allowing the reaction to cool to room temperature, the remaining excess of LiAlH<sub>4</sub> was quenched by carefully adding H<sub>2</sub>O (25 mL), followed by addition of an aqueous 1M NaOH solution (25 mL) and H<sub>2</sub>O (25 mL). This mixture was vigorously stirred for 1h at room temperature. The resulting white precipitate was filtrated over a short pat of celite. The filtrate was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduce pressure to afford compound **2** (11.1 g, 81%), which was used in the next step without purification.

<sup>1</sup>H-NMR (400MHz, DMSO – d<sub>6</sub>) δ: 7.29 (d, J = 8.0 Hz, 2H), 7.22 (d, J = 8.0 Hz, 2H), 4.60 (s, 2H), 3.77 (s, 2H), 2.37 (s, 3H).

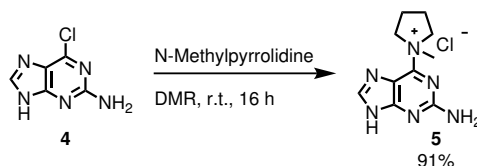
The <sup>1</sup>H-NMR spectrum is in agreement with the literature[48].

Preparation of 2,2,2-trifluoro-N-(4-(hydroxymethyl)benzyl)acetamide (**3**)

Compound **2** (10.9 g, 79.5 mmol, 1.0 eq) was dissolved in MeOH (100 mL), prior to the addition of NEt<sub>3</sub> (11.1 μL, 79.5 mmol, 1.0 eq) and ethyl trifluoroacetate (12.5 mL, 103.4 mmol, 1.3 eq). The reaction mixture was stirred at room temperature for 16 h. The reaction was quenched with H<sub>2</sub>O and the aqueous solution was extracted three times with EtOAc. The organic layer was washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude product was purified by column chromatography (cylcohexane/EtOAc 3:1) to afford **3** (16.5 g, 93%).

<sup>1</sup>H-NMR (400MHz, DMSO – d<sub>6</sub>) δ: 9.98 (s, 1H), 7.30 (d, J = 8.1 Hz, 2H), 7.23 (d, J = 8.1 Hz, 2H), 5.17 (t, J = 5.7 Hz, 1H), 4.48 (d, J = 5.7 Hz, 2H) 4.37 (d, J = 6.0 Hz, 2H).

The <sup>1</sup>H-NMR spectrum is in agreement with the literature[48].

Preparation of 1-(2-amino-9H-purin-6-yl)-1-methylpyrrolidin-1-ium chloride (**5**)

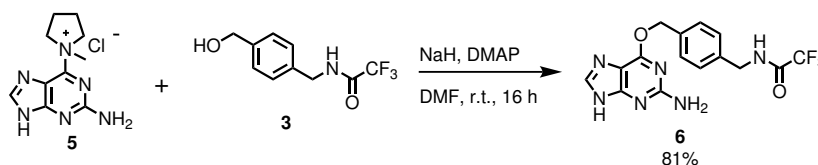
Chloroguanine **4** (15 g, 88 mmol, 1.0 eq) was dissolved in DMF (250 mL), N-methylpyrrolidine (28 mL, 264 mmol, 3.0 eq) was added and the mixture was stirred at room temperature for 16 h. After the addition of acetone (175 mL), the formed precipitate was filtered, washed with diethyl ether, and dried under reduced pressure to yield **5** (20.3 g, 91%).

## Experimental Section

**<sup>1</sup>H-NMR** (400MHz, DMSO – d<sub>6</sub>) δ: 13.45 (s, 1H), 8.34 (s, 1H), 7.11 (s, 2H), 4.59 (dt, J = 12.3, 6.6 Hz, 2H), 3.97 (dt, J = 13.5 Hz, 7.51 Hz, 2H), 3.65 (s, 3H), 2.26-2.22 (m, 2H), 2.07-2.03 (m, 2H).

The **<sup>1</sup>H-NMR** spectrum is in agreement with the literature[48].

### Preparation of N-(4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)-2,2,2-trifluoroacetamide (**6**)

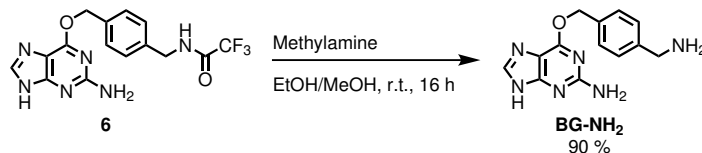


Compound **3** (2.704 g, 11.60 mmol, 1.0 eq) was dissolved in DMF (20 mL). NaH (60%, 1.392 g, 34.8 mmol, 3.0 eq) was added portionwise and the mixture was stirred for 30 min, prior to the addition of **5** (2.955 g, 11.60 mmol, 1.0 eq) and DMAP (0.142 g, 1.16 mmol, 0.1 eq). The reaction mixture was further stirred at room temperature for 16 h. The reaction was carefully quenched with a saturated aqueous solution of NH<sub>4</sub>Cl. The product was extracted three times with EtOAc, dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 20:1 to 10:1) to afford **6** (3.6 g, 85%).

**<sup>19</sup>F-NMR** (400MHz, DMSO – d<sub>6</sub>) δ: -74.25.

The **<sup>19</sup>F-NMR** spectrum is in agreement with the literature[48].

### Preparation of 6-((4-(aminomethyl)benzyl)oxy)-9H-purin-2-amine (BG-NH<sub>2</sub>)



A 33% solution of methylamine in ethanol (80 mL) was added to a mixture of **6** (3.0 g, 8.1 mmol) in dry MeOH (20 mL). The reaction mixture was stirred at room temperature for 16 h. The resulting precipitate was filtered, washed with cold EtOH and dried under reduced pressure to yield BG-NH<sub>2</sub> (2.0 g, 90%).

**<sup>1</sup>H-NMR** (400MHz, DMSO – d<sub>6</sub>) δ: 7.81 (s, 1H), 7.43 (d, J = 8.1 Hz, 2H), 7.34 (d, J = 8.1 Hz, 2H), 6.26 (s, 2H), 5.44 (s, 2H), 3.71 (s, 2H).

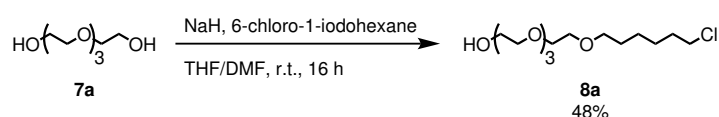
The **<sup>1</sup>H-NMR** spectrum is in agreement with the literature[48].

### 5.3.1.3 Synthesis of Halo-PEG<sub>n</sub>-Br building blocks<sup>[128]</sup>

#### General procedure 2 (GP2): Alkylation of 1-chloro-6-iodohexane with polyethylene glycols

NaH (60% in mineral oil, 1.3 mmol) was added portionwise to a solution of poly-ethyleneglycol (1 mmol) in a THF/DMF (vol. 3:1) mixture at 0°C. After stirring the reaction mixture for 30 min, 6-chloro-1-iodohexane (1 mmol) was added at 0°C. The mixture was further stirred at room temperature for 16 h. The excess of NaH was carefully quenched with water and the crude mixture was poured into water and extracted three times with EtOAc. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, concentrated under reduce pressure and the crude oil was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH).

#### Preparation of 18-chloro-3,6,9,12-tetraoxaoctadecan-1-ol (**8a**)

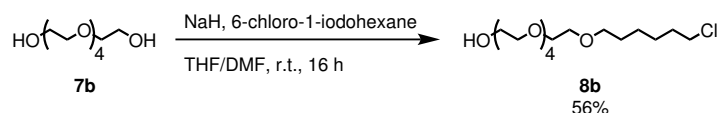


According to GP2, starting from NaH (60% in mineral oil, 1.23 g, 30 mmol, 1.3 eq), tetraethyleneglycol (**7a**) (5 g, 25.7 mmol, 1 eq) and 6-chloro-1-iodohexane (4.3 mL, 28 mmol, 1 eq), column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 30:1) afforded **8a** as colorless oil (2.9 g, 48%).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ: 3.73-3.70 (m, 2H), 3.67-3.57 (m, 14H), 3.53 (t, J = 6.7 Hz, 2H), 3.38 (t, J = 6.7 Hz, 2H), 1.81-1.74 (m, 2H), 1.63-1.56 (m, 2H), 1.49-1.35 (m, 4H).

The <sup>1</sup>H-NMR spectrum is in agreement with the literature<sup>[128]</sup>.

#### Preparation of 21-chloro-3,6,9,12,15-pentaoxahenicosan-1-ol (**8b**)

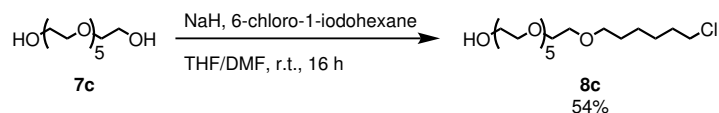


According to GP2, starting from NaH (60% in mineral oil, 1.72 g, 46 mmol, 1.3 eq), pentaethyleneglycol (**7b**) (10 g, 42 mmol, 1 eq) and 6-chloro-1-iodohexane (6.4 mL, 42 mmol, 1 eq), column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 30:1) afforded **8b** (5.3 g, 56%).

<sup>1</sup>H-NMR (400 MHz, CD<sub>3</sub>OD) δ: 3.67-3.61 (m, 16H), 3.58-3.54 (m, 6H), 3.47 (t, J = 6.5 Hz, 2H), 1.80-1.73 (m, 2H), 1.62-1.55 (m, 2H), 1.49-1.37 (m, 4H).

<sup>13</sup>C-NMR (100.6 MHz, CD<sub>3</sub>OD) δ: 73.6, 72.1, 71.6, 71.5, 71.5, 71.4, 71.1, 62.2, 45.7, 33.7, 30.5, 27.7, 26.5.

#### Preparation of 24-chloro-3,6,9,12,15,18-hexaoxatetracosan-1-ol (**8c**)



According to GP2, starting from NaH (60% in mineral oil, 370 mg, 9.2 mmol, 1.3 eq), hexaethyleneglycol (**7c**) (2.1 mL, 8.4 mmol, 1 eq) and 6-chloro-1-iodohexane (1.3 mL, 8.4 mmol, 1 eq), column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 20:1) afforded **8c** (2.0 g, 54%).

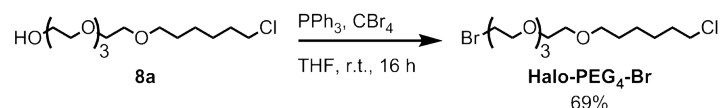
**<sup>1</sup>H-NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$ : 3.67-3.61 (m, 20H), 3.58-3.54 (m, 6H), 3.47 (t,  $J$  = 6.51 Hz, 2H), 1.80-1.73 (m, 2H), 1.62-1.55 (m, 2H), 1.50-1.37 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, CD<sub>3</sub>OD)  $\delta$ : 73.6, 72.1, 71.6, 71.5, 71.5, 71.5, 71.1, 62.2, 45.7, 33.7, 30.5, 27.7, 26.5.

**General procedure 3 (GP3): Apple reaction of the HaloTag substrates **8a**, **8b**, and **8c** with CBr<sub>4</sub>**

Triphenylphosphine (1.15 mmol) and carbon tetrabromide (1.15 mmol) were added portionwise to a solution of corresponding chloroalkane (1 mmol) in THF (5 mL) at 0°C. The resulting mixture was stirred at room temperature for 16 h. The solvent was evaporated under reduced pressure and the crude oil was purified by column chromatography (cylcohexane/EtOAc) to yield the desired product as a colourless oil.

**Preparation of 1-bromo-18-chloro-3,6,9,12-tetraoxaoctadecane (Halo-PEG<sub>4</sub>-Br)**

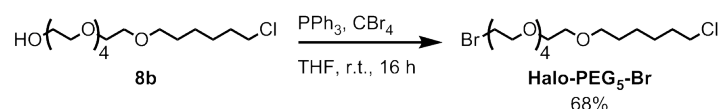


According to GP3, starting from triphenylphosphine (1.45g, 5.53 mmol, 1.15 eq), carbon tetrabromide (1.83g, 5.53 mmol, 1.15 eq) and compound **8a** (1.5 g, 4.8 mmol, 1 eq), column chromatography (cylcohexane/EtOAc 3:1) yielded **Halo-PEG<sub>4</sub>-Br** (1.23 g, 69%).

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$ : 3.81 (t,  $J$  = 6.3 Hz, 2H), 3.68-3.63 (m, 10H), 3.59-3.66 (m, 2H), 3.53 (t,  $J$  = 6.7 Hz, 2H), 3.49-3.44 (m, 4H), 1.81-1.74 (m, 2H), 1.63-1.56 (m, 2H), 1.49-1.35 (m, 4H).

The **<sup>1</sup>H-NMR** spectrum is in agreement with the literature[128].

**Preparation of 1-bromo-21-chloro-3,6,9,12,15-pentaoxahenicosane (Halo-PEG<sub>5</sub>-Br)**

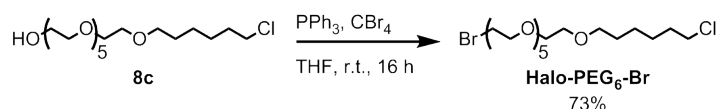


According to GP3, starting from triphenylphosphine (1.69 g, 6.44 mmol, 1.15 eq), carbon tetrabromide (2.13g, 6.44 mmol, 1.15 eq) and compound **8b** (2 g, 5.6 mmol, 1 eq), column chromatography (cylcohexane/EtOAc 3:1) yielded **Halo-PEG<sub>5</sub>-Br** (1.6 g, 68%).

**<sup>1</sup>H-NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$ : 3.80 (t,  $J$  = 6.0 Hz, 2H), 3.66-3.61 (m, 14H), 3.58-3.54 (m, 4H), 3.51 (t,  $J$  = 6.0 Hz, 2H), 3.78 (t,  $J$  = 6.5 Hz, 2H), 1.77 (m, 2H), 1.58 (m, 2H), 1.36-1.52 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, CD<sub>3</sub>OD)  $\delta$ : 72.3, 72.1, 71.6, 71.5, 71.5, 71.4, 71.1, 45.7, 33.7, 31.4, 30.5, 27.7, 26.4.

### Preparation of 1-bromo-24-chloro-3,6,9,12,15,18-hexaoxatetracosane (Halo-PEG<sub>6</sub>-Br)



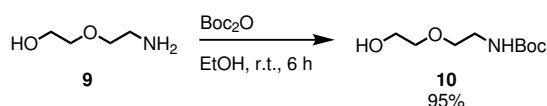
According to GP3, starting from triphenylphosphine (721 mg, 2.76 mmol, 1.15 eq), carbon tetrabromide (900 mg, 2.76 mmol, 1.15 eq) and compound **8c** (950 mg, 2.37 mmol, 1 eq) column chromatography (cyclohexane/EtOAc 3:1) yielded **Halo-PEG<sub>6</sub>-Br** (800 mg, 73%).

**<sup>1</sup>H-NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$ : 3.80 (t, *J* = 6.0 Hz, 2H), 3.66-3.61 (m, 18H), 3.58-3.54 (m, 4H), 3.51 (t, *J* = 6.0 Hz, 2H), 3.48 (t, *J* = 6.6 Hz, 2H), 1.80-1.73 (m, 2H), 1.62-1.55 (m, 2H), 1.49-1.37 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, CD<sub>3</sub>OD)  $\delta$ : 72.3, 72.1, 71.6, 71.4, 71.2, 45.7, 33.8, 31.4, 30.6, 27.7, 26.5.

#### 5.3.1.4 Synthesis of Halo-NH<sub>2</sub> building block[256]

##### Preparation of *tert*.-butyl (2-(2-hydroxyethoxy)ethyl)carbamate (**10**)

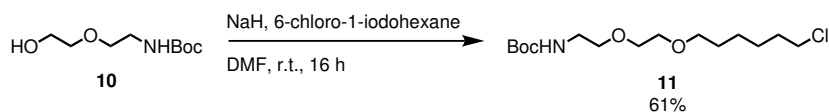


Boc<sub>2</sub>O (17 g, 78 mmol, 1.0 eq) was slowly added to a solution of 2-(2-aminoethoxy)ethanol (**9**) (8 g, 78 mmol, 1.0 eq) in anhydrous EtOH (150 mL) at 0°C. The mixture was stirred at room temperature for 6 h. The reaction mixture was diluted with EtOAc and washed with brine. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure to give **10** (15.2 g, 95%) as a colorless oil, which was used without further purification.

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$ : 5.11 (s, 1H), 3.75-3.73 (m, 2H), 3.59-3.54 (m, 4H), 3.35-3.31 (m, 2H), 2.64 (s, 1H), 1.45 (s, 9H).

The **<sup>1</sup>H-NMR** spectrum is in agreement with the literature[256].

##### Preparation of *tert*.-butyl (2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamate (**11**)

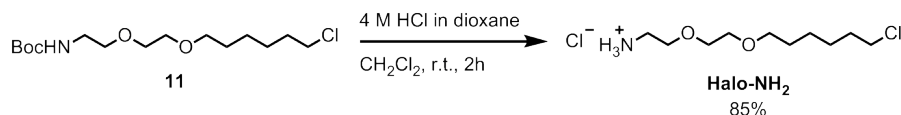


NaH (60% in mineral oil, 2.7 g, 67 mmol, 1.3 eq) was portionwise added to a solution of **10** (10.7 g, 52 mmol, 1.0 eq) in DMF (20 mL) at 0°C. The mixture was stirred for 30 min at 0°C prior to the addition of 1-chloro-6-iodohexane **18** (12.8 g, 51.8 mmol, 1.0 eq). The resulting mixture was stirred at room temperature for 16 h. The reaction was quenched with a saturated aqueous solution of NH<sub>4</sub>Cl and the product was extracted three times with EtOAc. The combined organic layer was washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude product was purified by column chromatography (cyclohexane/EtOAc 3:1 to 1:1) to afford **11** (7.5 g, 46%).

**<sup>1</sup>H-NMR** (400 MHz, DMSO – d<sub>6</sub>) δ: 6.7 (s, 1H), 3.61 (t, J = 6.60 Hz, 2H), 3.49-3.44 (m, 4H), 3.38-3.35 (m, 4H), 3.06 (q, J = 6.0 Hz, 2H), 1.74-1.67 (m, 2H), 1.52-1.45 (m, 2H), 1.41-1.28 (m, 13H).

The **<sup>1</sup>H-NMR** spectrum is in agreement with the literature[256].

#### Preparation of 2-((6-chlorohexyl)oxy)ethoxyethan-1-amine (Halo-NH<sub>2</sub>)



Compound **11** (5 g, 16 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (10 mL). 10 mL of a 4M HCl solution in dioxane were added and the reaction mixture was stirred at 80°C for 2 h. After cooling to room temperature the solvent was evaporated under reduced pressure to afford **Halo-NH<sub>2</sub>** as its HCl salt (3.5 g, 85%).

**<sup>1</sup>H-NMR** (400MHz, DMSO – d<sub>6</sub>) δ: 8.13 (s, 3H), 3.64-3.60 (m, 4H), 3.57-3.48 (m, 4H), 3.38 (t, J = 6.6 Hz, 4H), 2.93 (t, J = 5.2 Hz, 2H), 1.74-1.67 (m, 2H), 1.53-1.46 (m, 2H), 1.42-1.28 (s, 13H).

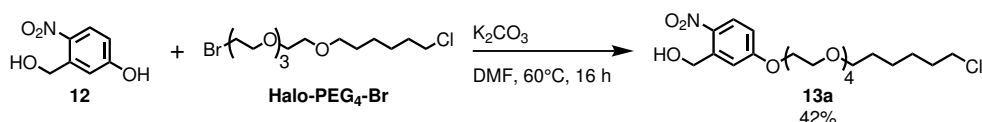
The **<sup>1</sup>H-NMR** spectrum is in agreement with the literature[256].

#### 5.3.1.5 Synthesis of Nb-HaXS molecules

##### General procedure 4 (GP4): Alkylation of 3-(hydroxymethyl)-4-nitrophenol with Halo-PEG<sub>n</sub>-Br

**Halo-PEG<sub>n</sub>-Br** (1 mmol) and 5-hydroxynitrobenzyl alcohol (1 mmol) were dissolved in DMF (5 mL). K<sub>2</sub>CO<sub>3</sub> (1 mmol) was added and the mixture was stirred at 60°C for 16 h. The reaction was quenched with H<sub>2</sub>O (75 mL) and extracted three times with EtOAc. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduce pressure. The crude oil was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH) to yield the desired product as a yellowish oil.

##### Preparation of (5-((18-chloro-3,6,9,12-tetraoxaoctadecyl)oxy)-2-nitrophenyl)methanol (**13a**)

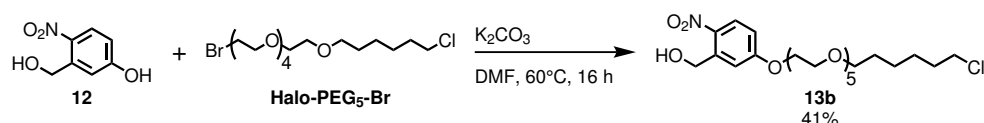


According to GP4, starting from K<sub>2</sub>CO<sub>3</sub> (106.7 mg, 0.77 mmol, 1 eq), 5-hydroxynitrobenzyl alcohol (**12**) (130.6 mg, 0.77 mmol, 1 eq) and compound **Halo-PEG<sub>4</sub>-Br** (290 mg, 0.77 mmol, 1 eq), column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 30:1) afforded **13a** (150 mg, 42%).

**<sup>1</sup>H-NMR** (400 MHz, CD<sub>3</sub>OD) δ: 8.10 (d, J = 9.1 Hz, 1H), 7.38 (d, J = 2.8 Hz, 1H), 6.95 (dd, J = 9.1, 2.8 Hz, 1H), 4.95 (s, 2H), 4.79 (s, 1H), 4.25-4.23 (m, 2H), 3.88-3.85 (m, 2H), 3.69-3.51 (m, 14H), 3.43 (t, J = 6.5 Hz, 2H), 1.76-1.69 (m, 2H), 1.57-1.50 (m, 2H), 1.44-1.32 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, CD<sub>3</sub>OD) δ: 164.7, 143.00, 141.00, 128.5, 114.3, 113.9, 72.1, 71.7, 71.5, 71.5, 71.4, 71.1, 70.4, 69.3, 62.2, 45.7, 33.7, 30.5, 27.7, 26.4.

**Preparation of (5-((21-chloro-3,6,9,12,15-pentaoxahenicosyl)oxy)-2-nitrophenyl)methanol (**13b**)**

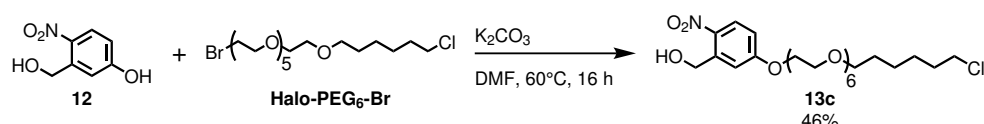


According to GP4, starting from  $K_2CO_3$  (235 mg, 2.14 mmol, 1 eq), 5-hydroxynitrobenzyl alcohol (**12**) (362 mg, 2.14 mmol, 1 eq) and compound **Halo-PEG<sub>5</sub>-Br** (900 mg, 2.14 mmol, 1 eq), column chromatography ( $CH_2Cl_2/MeOH$  30:1) afforded **13b** (450 mg, 41%).

**<sup>1</sup>H-NMR** (400 MHz,  $CD_3OD$ )  $\delta$ : 8.12 (d,  $J = 9.1$  Hz, 1H), 7.39 (d,  $J = 2.8$  Hz, 1H), 6.97 (dd,  $J = 9.1, 2.8$  Hz, 1H), 4.95 (s, 2H), 4.79 (s, 1H), 4.26-4.24 (m, 2H), 3.88-3.86 (m, 2H), 3.69-3.51 (m, 18H), 3.43 (t,  $J = 6.6$  Hz, 2H), 1.77-1.70 (m, 2H), 1.58-1.51 (m, 2H), 1.46-1.33 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz,  $CD_3OD$ )  $\delta$ : 164.8, 143.0, 141.1, 128.5, 114.4, 114.0, 72.1, 71.8, 71.6, 71.5, 71.1, 70.5, 69.3, 62.2, 45.7, 33.7, 30.5, 27.7, 26.4.

**Preparation of (5-((24-chloro-3,6,9,12,15,18-hexaoxatetracosyl)oxy)-2-nitrophenyl)methanol (**13c**)**



According to GP4, starting from  $K_2CO_3$  (295 mg, 2.14 mmol), 5-hydroxynitrobenzyl alcohol (**12**) (362 mg, 2.14 mmol) and compound **Halo-PEG<sub>6</sub>-Br** (1.09 g, 2.14 mmol), column chromatography ( $CH_2Cl_2/MeOH$  30:1) afforded **13c** (550 mg, 46%).

**<sup>1</sup>H-NMR** (400 MHz,  $CDCl_3$ )  $\delta$ : 8.03 (d,  $J = 9.1$  Hz, 1H), 7.30 (d,  $J = 2.8$  Hz, 1H), 6.78 (dd,  $J = 9.1, 2.8$  Hz, 1H), 4.92 (s, 2H), 4.17-4.15 (m, 2H), 3.81-3.78 (m, 2H), 3.63-3.42 (m, 22H), 3.36 (t,  $J = 6.6$  Hz, 2H), 1.71-1.64 (m, 2H), 1.53-1.46 (m, 2H), 1.39-1.26 (m, 4H).

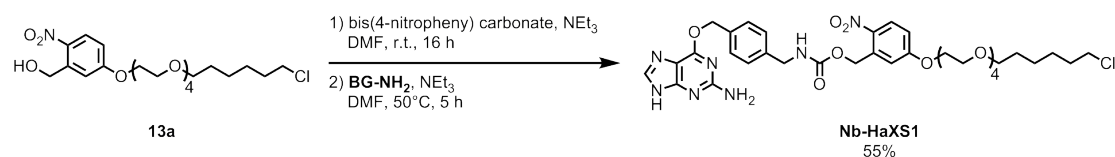
**<sup>13</sup>C-NMR** (100.6 MHz,  $CDCl_3$ )  $\delta$ : 163.5, 141.4, 140.0, 132.1, 132.0, 128.6, 128.5, 127.7, 113.8, 113.5, 71.2, 70.9, 70.6, 70.6, 70.6, 70.5, 70.1, 69.5, 68.2, 62.3, 45.1, 32.5, 29.4, 26.7, 25.4.

**General procedure 5 (GP5): Carbamate formation between BG-NH<sub>2</sub> and compounds **13a**, **13b** and **13c****

Compounds **13a**, **13b** or **13c** (1 mmol) were slowly added to a solution of bis(4-nitrophenyl) carbonate (1 mmol) in DMF (5 mL) at 0°C.  $NEt_3$  (1.4 mmol) was added and the reaction mixture was stirred at room temperature for 16 h. **BG-NH<sub>2</sub>** (1 mmol) and  $NEt_3$  (1.4 mmol) were added to the solution and the mixture was stirred at 50°C for 5 h. The mixture was poured onto  $H_2O$  and the product was extracted with  $EtOAc$ . The combined organic layer was dried over  $Na_2SO_4$  and concentrated under reduce pressure. The crude oil was purified by column chromatography ( $CH_2Cl_2/MeOH$ ) to yield the desired product.



**Preparation of 5-((18-chloro-3,6,9,12-tetraoxaoctadecyl)oxy)-2-nitrobenzyl (4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)carbamate (Nb-HaXS1)**

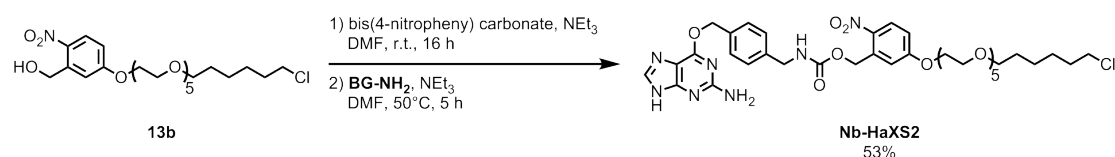


According to GP5, starting from bis(4-nitrophenyl) carbonate (35 mg, 0.22 mmol, 1 eq), compound **13a** (100 mg, 0.22 mmol), NEt<sub>3</sub> (43  $\mu$ L, 0.38 mmol, 1.4 eq), **BG-NH<sub>2</sub>** (58 mg, 0.22 mmol, 1 eq) and NEt<sub>3</sub> (43  $\mu$ L, 0.38 mmol, 1.4 eq), column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 30:1) yielded **Nb-HaXS1** (90 mg, 55%).

**<sup>1</sup>H-NMR** (400 MHz, DMSO - d<sub>6</sub>)  $\delta$ : 12.4 (s, 1H), 8.18 (d, J = 9.1 Hz, 1H), 8.08 (s, 1H), 7.79 (s, 1H), 7.46 (d; J = 7.9 Hz, 2H), 7.30 (d, J = 7.9 Hz, 2H), 7.17 (d, J = 2.8 Hz, 1H), 7.12 (dd, J = 9.1, 2.8 Hz, 1H), 6.27 (s, 2H), 5.45 (s, 2H), 5.40 (s, 2H), 4.24-4.21 (m, 2H), 3.77-3.75 (m, 2H), 3.61-3.32 (m, 18H), 1.70-1.63 (m, 2H), 1.47-1.41 (m, 2H), 1.36-1.23 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, CD<sub>3</sub>OD)  $\delta$ : 164.7, 161.6, 158.3, 131.2, 140.5, 138.1, 136.9, 129.7, 128.8, 128.5, 114.5, 72.1, 71.8, 71.5, 71.5, 77.1, 70.5, 69.4, 68.7, 64.5, 45.7, 45.3, 33.7, 30.5, 27.7, 26.5. Photophysical properties:  $\lambda_{max}$ : 329 nm,  $\epsilon_{360}$ : 1690 M<sup>-1</sup>cm<sup>-1</sup>,  $\epsilon_{405}$ : 682 M<sup>-1</sup>cm<sup>-1</sup>.

**Preparation of 5-((21-chloro-3,6,9,12,15-pentaoxahenicosyl)oxy)-2-nitrobenzyl (4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)carbamate (Nb-HaXS2)**



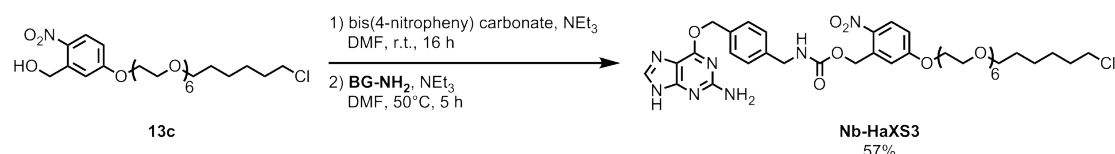
According to GP5, starting from bis(4-nitrophenyl) carbonate (259 mg, 0.85 mmol, 1 eq), compound **13b** (430 mg, 0.85 mmol, 1.4 eq), NEt<sub>3</sub> (134  $\mu$ L, 1.19 mmol, 1 eq), **BG-NH<sub>2</sub>** (229 mg, 0.85 mmol, 1 eq) and NEt<sub>3</sub> (134  $\mu$ L, 1.19 mmol, 1 eq) column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 30:1) afforded **Nb-HaXS2** (353 mg, 53%).

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$ : 8.02 (d, J = 9.1 Hz, 1H), 7.27 (d, J = 8.0 Hz, 2H), 7.2 (d, J = 8.0 Hz, 2H), 7.03 (s, 1H), 6.8 (d, J 9.1 Hz, 1H), 5.46-5.29 (m, 6H), 4.33-4.31 (m, 2H), 4.06-4.05 (m, 2H), 3.74-3.73 (m, 2H), 3.47-3.50 (m, 14H), 3.46 (t, J = 6.7 Hz, 2H), 3.38 (t, J = 6.7 Hz, 2H), 1.72-1.65 (m, 2H), 1.55-1.48 (m, 2H), 1.40-1.26 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, CDCl<sub>3</sub>)  $\delta$ : 163.2, 159.4, 156.4, 139.7, 138.9, 136.8, 135.1, 129.0, 127.6, 113.6, 113.1, 71.2, 70.7, 70.5, 70.4, 69.3, 68.0, 63.6, 45.1, 32.5, 29.4, 26.6, 25.4.

Photophysical properties:  $\lambda_{max}$ : 329 nm,  $\epsilon_{360}$ : 1730 M<sup>-1</sup>cm<sup>-1</sup>,  $\epsilon_{405}$ : 674 M<sup>-1</sup>cm<sup>-1</sup>.

**Preparation of 5-((24-chloro-3,6,9,12,15,18-hexaoxatetracosyl)oxy)-2-nitrobenzyl 4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)carbamate (Nb-HaXS3)**



According to GP5, starting from bis(4-nitrophenyl) carbonate (180 mg, 0.59 mmol, 1 eq), compound **13c** (324 mg, 0.59 mmol, 1 eq),  $\text{NEt}_3$  (92  $\mu\text{L}$ , 0.82 mmol, 1.4 eq), **BG-NH<sub>2</sub>** (158 mg, 0.59 mmol, 1 eq) and  $\text{NEt}_3$  (92  $\mu\text{L}$ , 0.82 mmol, 1.4 eq), column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  30:1) afforded **Nb-HaXS3** (285 mg, 57%).

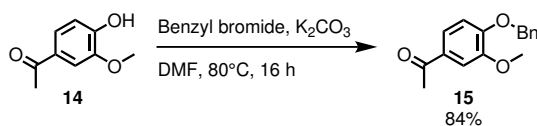
**<sup>1</sup>H-NMR** (400 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 12.43 (s, 1H), 8.18 (d,  $J = 9.1$  Hz, 1H), 8.10 (s, 1H), 7.83 (s, 1H), 7.47 (d,  $J = 8.0$  Hz, 2H), 7.31 (d,  $J = 8.0$  Hz, 2H), 7.18 (d,  $J = 2.7$  Hz, 1H), 7.12 (dd,  $J = 9.1, 2.7$  Hz, 1H), 6.26 (s, 2H), 5.46 (s, 2H), 5.41 (s, 2H), 4.26-4.21 (m, 2H), 3.77-3.75 (m, 2H), 3.61-3.43 (m, 20H), 3.34 (t,  $J = 6.5$  Hz, 2H), 1.71-1.64 (m, 2H), 1.50-1.43 (m, 2H), 1.39-1.26 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$ : 163.2, 159.5, 156.4, 139.9, 139.0, 136.9, 135.3, 129.1, 127.8, 113.6, 113.2, 71.2, 70.8, 70.6, 70.5, 70.1, 69.3, 68.1, 63.7, 45.1, 32.6, 29.5, 26.7, 25.4.

Photophysical properties:  $\lambda_{\text{max}}$ : 329 nm,  $\epsilon_{360}$ :  $1796 \text{ M}^{-1}\text{cm}^{-1}$ ,  $\epsilon_{405}$ :  $677 \text{ M}^{-1}\text{cm}^{-1}$ .

### 5.3.1.6 Synthesis of MeNV-HaXS

**Preparation of 1-(4-(benzyloxy)-3-methoxyphenyl)ethan-1-one (15)**



Acetovanillone (**14**) (5 g, 30.1 mmol, 1 eq) and benzyl bromide (5.15 g, 30.1 mmol, 1 eq) were dissolved in DMF (5 mL).  $\text{K}_2\text{CO}_3$  (4.15 g, 30.1 mmol, 1 eq) was added and the mixture was stirred at  $80^\circ\text{C}$  for 16 h. The reaction was quenched with a saturated aqueous solution of  $\text{NH}_4\text{Cl}$  and extracted three times with  $\text{EtOAc}$ . The combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure to give **15** (6.5 g, 84%), which was used in the next step without further purification.

**<sup>1</sup>H-NMR** (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.48 (d,  $J = 2.01$  Hz, 1H), 7.43 (dd,  $J = 8.26$  Hz, 2.01 Hz, 1H), 7.36 (d,  $J = 7.23$  Hz, 2H), 7.31 (td,  $J = 7.23$  Hz, 1.54 Hz, 2H), 7.25 (tt,  $J = 7.12$  Hz, 1.54 Hz, 2H), 6.82 (d,  $J = 8.26$  Hz, 1H), 5.16 (s, 2H), 3.88 (s, 3H), 2.47 (s, 3H).

**<sup>13</sup>C-NMR** (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$ : 196.9, 152.5, 149.6, 136.4, 130.8, 128.7, 128.1, 127.2, 123.1, 112.2, 110.5, 70.8, 56.1, 26.2.

**HRMS**  $\text{C}_{16}\text{H}_{17}\text{O}_3$   $[\text{M}-\text{H}]^+$  calcd: 257.1172, found: 257.1166

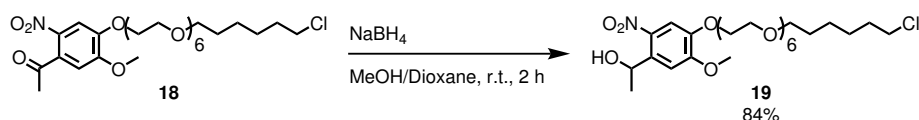


**<sup>1</sup>H-NMR** (400 MHz, DMSO – d<sub>6</sub>) δ: 7.69 (s, 1H), 6.76 (s, 1H), 4.28 (m, 2H), 3.96 (s, 3H), 3.92 (m, 2H), 3.74 (m, 2H), 3.69-3.62 (m, 18H), 3.56 (m, 2H), 3.53 (t, J = 6.8 Hz, 2H), 3.45 (t, J = 6.8 Hz, 2H), 2.50 (s, 3H), 1.77 (m, 2H), 1.59 (m, 2H), 1.49-1.34 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, DMSO – d<sub>6</sub>) δ: 200.1, 154.4, 149.0, 138.3, 133.0, 108.9, 108.7, 71.2, 71.0, 70.7, 70.6, 70.1, 69.4, 69.3, 56.7, 53.6, 45.1, 32.6, 30.4, 29.5, 26.7, 25.5.

**HRMS** C<sub>27</sub>H<sub>45</sub>O<sub>11</sub>ClN [M-H]<sup>+</sup> calcd: 594.2676, found: 594.2664.

**Preparation of 1-(4-((24-chloro-3,6,9,12,15,18-hexaoxatetracosyl)oxy)-5-methoxy-2-nitrophenyl)ethanol (19)**



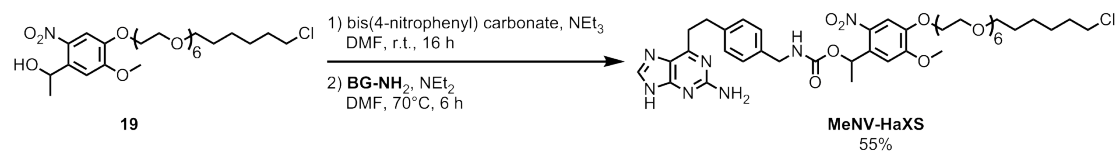
Compound **18** (675 mg, 1.1 mmol, 1 eq) was dissolved in MeOH/dioxane mixture (7 mL each). NaBH<sub>4</sub> (51 mg, 2.0 mmol, 1.8 eq) was added portionwise at 0°C. The mixture was stirred at room temperature for 2 h. The mixture was poured onto water, neutralised with a 1 M solution of HCl and extracted three times with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1) to yield **19** (568 mg, 84%) as yellow oil.

**<sup>1</sup>H-NMR** (400 MHz, DMSO – d<sub>6</sub>) δ: 7.61 (s, 1H), 7.29 (s, 1H), 5.53 (q, J = 6.3 Hz, 1H), 4.21 (m, 2H), 3.95 (s, 3H), 3.88 (m, 2H), 3.70 (m, 2H), 3.59-3.66 (m, 18H), 3.55 (m, 2H), 3.51 (m, 2H), 3.43 (m, 2H), 2.50 (s, 3H), 1.75 (m, 2H), 1.57 (m, 2H), 1.52 (d, J = 6.3 Hz, 3H), 1.32-1.46 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, DMSO – d<sub>6</sub>) δ: 153.4, 146.3, 138.9, 138.1, 109.1, 108.6, 70.2, 70.0, 69.8, 69.5, 68.8, 68.4, 66.4, 56.0, 54.9, 45.3, 32.0, 29.1, 26.1, 25.2, 25.0.

**HRMS** C<sub>27</sub>H<sub>47</sub>O<sub>11</sub>ClNNa [M-Na]<sup>+</sup> calcd: 619.2730, found: 619.2724.

**Preparation of 1-(4-((24-chloro-3,6,9,12,15,18-hexaoxatetracosyl)oxy)-5-methoxy-2-nitrophenyl)ethyl (4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)carbamate (MeNV-HaXS)**



Compound **19** (829 mg, 1.4 mmol, 1 eq) was slowly added to a solution of bis(4-nitrophenyl) carbonate (423 mg, 1.4 mmol, 1 eq) in DMF (10 mL) at 0°C. NEt<sub>3</sub> (320 μL, 2.0 mmol, 1.4 eq) was added and the mixture was stirred at room temperature for 16 h. **BG-NH<sub>2</sub>** (379 mg, 1.4 mmol, 1 eq) and NEt<sub>3</sub> (320 μL, 2.0 mmol, 1.4 eq) were added and the reaction was further stirred at 70°C for 6 h. The reaction was quenched with H<sub>2</sub>O and extracted three times with EtOAc. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1) to yield **MeNV-HaXS** (687 mg, 55%).

**<sup>1</sup>H-NMR** (400 MHz, MeOD) δ: 7.66 (s, 1H), 7.43 (d, J = 7.9, 2H), 7.21 (d, J = 7.9, 2H), 7.14 (s, 1H), 6.28 (q, J = 6.5 Hz, 1H), 5.51 (m, 2H), 4.14-4.31 (m, 5H), 3.80-3.86 (m, 6H), 3.68-3.70 (m, 2H), 3.50-3.64 (m, 24H), 3.43 (t, J = 6.6 Hz, 2H), 1.69-1.74 (m, 2H), 1.59 (d, J = 6.5 Hz, 3H), 1.52-1.59 (m, 2H), 1.30-1.46 (m, 4H).

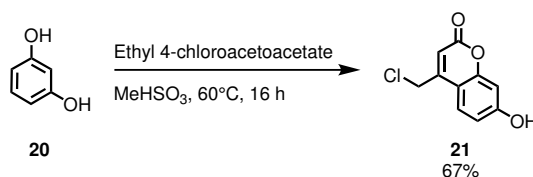
**$^{13}\text{C}$ -NMR** (100.6 MHz, MeOD)  $\delta$ : 148.7, 141.0, 140.4, 129.7, 128.4, 110.8, 109.5, 72.2, 71.8, 71.6, 71.5, 71.2, 70.7, 70.3, 70.1, 68.8, 56.9, 45.8, 45.1, 33.8, 30.6, 27.7, 26.5, 22.4.

**HRMS**  $\text{C}_{41}\text{H}_{59}\text{O}_{13}\text{ClN}_7$   $[\text{M}-\text{H}]^+$  calcd: 892.3854, found: 892.3842.

Photophysical properties:  $\lambda_{\text{max}}$ : 355 nm,  $\epsilon_{360}$ : 4058  $\text{M}^{-1}\text{cm}^{-1}$ ,  $\epsilon_{405}$ : 1248  $\text{M}^{-1}\text{cm}^{-1}$ ,  $\phi_{360}$ : 0.075,  $\phi_{405}$ : 0.072.

### 5.3.1.7 Synthesis of HCM-HaXS

#### Preparation of 4-(chloromethyl)-7-hydroxy-2H-chromen-2-one (21)

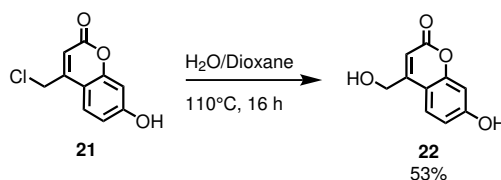


Resorcinol **20** (5.0 g, 45.5 mmol, 1 eq) was dissolved in  $\text{MeHSO}_3$  (10 mL). Ethyl 4-chloroacetoacetate (7.5 g, 45.5 mmol, 1 eq) was added and the mixture was stirred at  $60^\circ\text{C}$  for 16 h. The reaction mixture was poured onto ice water and stirred for 30 min. The precipitate was filtered, washed with water and cold EtOAc and dried in vacuum to give the desired compound **21** (6.4 g, 67% yield), which was used in the next step without further purification.

**$^1\text{H}$ -NMR** (400 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 7.66 (d,  $J = 8.8$  Hz, 1H), 6.82 (dd,  $J = 8.8, 2.4$  Hz, 1H), 6.74 (d,  $J = 2.4$  Hz, 1H), 6.39 (s, 2H), 4.92 (s, 2H).

**$^{13}\text{C}$ -NMR** (100.6 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 161.5, 160.2, 155.3, 150.9, 126.5, 113.1, 111.1, 109.4, 102.6, 41.40.

#### Preparation of 7-hydroxy-4-(hydroxymethyl)-2H-chromen-2-one (22)

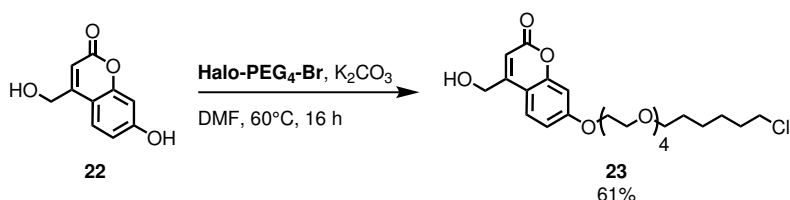


Compound **21** (4 g, 19.1 mmol) was dissolved in a mixture of water and dioxane (10 mL each). The reaction mixture was stirred at  $110^\circ\text{C}$  for 16 h and diluted with water. The product was extracted with three times EtOAc, the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent evaporated under reduced pressure. The product was purified by column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  20:1) to give the desired compound **22** (1.94 g, 53%).

**$^1\text{H}$ -NMR** (400 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 7.48 (d,  $J = 8.7$  Hz, 1H), 6.75 (dd,  $J = 8.7, 2.3$  Hz, 1H), 6.71 (d,  $J = 2.3$  Hz, 1H), 6.23 (s, 2H), 4.69 (s, 2H).

**$^{13}\text{C}$ -NMR** (100.6 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 161.1, 160.8, 156.8, 155.0, 125.5, 112.9, 109.7, 106.7, 102.4, 59.2.

**Preparation of 7-((18-chloro-3,6,9,12-tetraoxaoctadecyl)oxy)-4-(hydroxymethyl)-2H-chromen-2-one (**23**)**

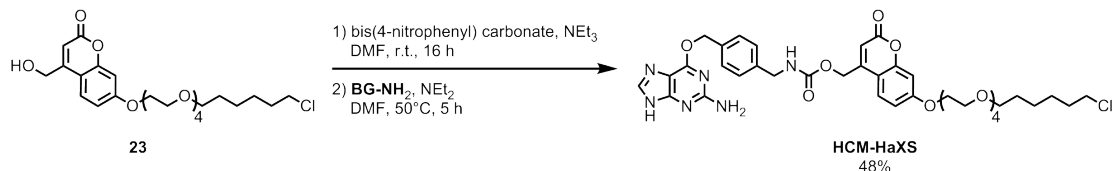


**Halo-PEG<sub>4</sub>-Br** (460 mg, 1 mmol, 1 eq) was added to a solution of **22** (192 mg, 1 mmol, 1 eq) and  $K_2CO_3$  (135 mg, 1 mmol, 1 eq) in DMF (2 mL). The mixture was stirred at 60°C for 16 h. The reaction mixture was poured onto a saturated aqueous solution of  $NH_4Cl$  and extracted three times with EtOAc. The combined organic layer was dried over  $Na_2SO_4$ , the solvent was evaporated under reduced pressure and the crude product was purified by column chromatography ( $CH_2Cl_2/MeOH$  40:1) to afford **23** (300 mg, 61%).

**<sup>1</sup>H-NMR** (400 MHz,  $CD_3OD$ )  $\delta$ : 7.51 (d,  $J$  = 8.5 Hz, 1H), 6.89 (dd,  $J$  = 8.8, 2.5 Hz, 1H), 6.86 (d,  $J$  = 2.5 Hz, 1H), 6.37 (s, 2H), 4.79 (d,  $J$  = 6.8 Hz, 2H), 4.20-4.17 (m, 2H), 3.87-3.84 (m, 2H), 3.70-3.50 (m, 14H), 3.42 (t,  $J$  = 6.5 Hz, 2H), 1.75-1.68 (m, 2H), 1.57-1.50 (m, 2H), 1.45-1.31 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz,  $CD_3OD$ )  $\delta$ : 163.6, 163.4, 158.0, 156.4, 126.0, 113.9, 112.3, 108.7, 102.5, 72.1, 71.8, 71.6, 71.5, 71.1, 70.5, 69.3, 60.8, 45.7, 33.7, 30.5, 27.7, 26.5.

**Preparation of (7-((18-chloro-3,6,9,12-tetraoxaoctadecyl)oxy)-2-oxo-2H-chromen-4-yl)methyl 4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)carbamate (HCM-HaXS)**



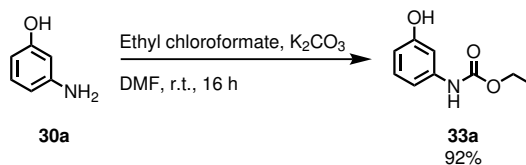
Compound **23** (200 mg, 0.41 mmol) and CDI (67 mg, 0.41 mmol) were dissolved in DMF (2 mL). The mixture was stirred at room temperature for 16 h. **BG-NH<sub>2</sub>** (107 mg, 0.41 mmol) and DIPEA (70  $\mu$ L, 0.41 mmol) were added to the solution. The mixture was stirred at 50°C for 5 h. The mixture was poured onto  $H_2O$  and extracted three times with EtOAc. The combined organic layer was dried over  $Na_2SO_4$  and concentrated under reduced pressure. The crude product was purified by flash column chromatography ( $CH_2Cl_2/MeOH$  10:1) to yield **HCM-HaXS** (154 mg, 48%).

**<sup>1</sup>H-NMR** (400 MHz,  $CDCl_3$ )  $\delta$ : 7.26-7.08 (m, 6H), 6.71 (d,  $J$  = 8.3 Hz, 1H), 6.66 (s, 1H), 6.32 (s, 1H), 5.58 (s, 2H), 5.17 (s, 4H), 4.35 (s, 2H), 3.79 (s, 2H), 3.66-3.53 (m, 12H), 3.47 (t,  $J$  = 6.7 Hz, 2H), 3.40 (t,  $J$  = 6.6 Hz, 2H), 1.74-1.67 (m, 2H), 1.57-1.50 (m, 2H), 1.40-1.24 (m, 2H).

**<sup>13</sup>C-NMR** (100.6 MHz,  $CDCl_3$ )  $\delta$ : 161.9, 161.5, 159.8, 159.6, 156.4, 156.1, 155.1, 150.7, 138.6, 136.9, 135.2, 129.0, 128.9, 127.7, 124.4, 112.9, 110.6, 109.0, 101.7, 71.3, 70.8, 70.6, 70.5, 70.1, 69.4, 68.0, 61.6, 45.1, 32.6, 29.5, 26.7, 25.4. Photophysical properties:  $\lambda_{max}$ : 320 nm,  $\epsilon_{360}$ : 990  $M^{-1}cm^{-1}$ ,  $\epsilon_{405}$ : 678  $M^{-1}cm^{-1}$ .

### 5.3.1.8 Synthesis of CACM-HaXS

#### Preparation of ethyl (3-hydroxyphenyl)carbamate (**33a**)

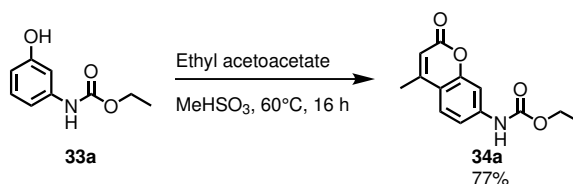


3-Aminophenol (**30a**) (20.0 g, 183.2 mmol, 1 eq) and  $K_2CO_3$  (37.8 g, 273.2 mmol, 1.5 eq), were dissolved in DMF (100 mL). Ethyl chloroformate (23.9 g, 219.8 mmol, 1.2 eq) was added and the mixture was stirred at room temperature for 16 h. The reaction was quenched with a saturated aqueous solution of  $NH_4Cl$ . The product was extracted three times with EtOAc, the combined organic layer was dried over  $Na_2SO_4$  and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (cyclohexane/EtOAc, 3:1) to give the desired compound **33a** (30.5 g, 92%).

**$^1H$ -NMR** (400 MHz,  $CDCl_3$ )  $\delta$ : 7.38 (s, 1H), 7.13 (t,  $J$  = 8.1 Hz, 1H), 6.73 (s, 1H), 6.63 (dd,  $J$  = 18.1 Hz, 2.2 Hz, 1H), 6.58 (dd,  $J$  = 8.1 Hz, 2.2 Hz, 1H), 4.23 (q,  $J$  = 7.1 Hz, 2H), 1.31 (t,  $J$  = 7.1 Hz, 3H).

**$^{13}C$ -NMR** (100.6 MHz,  $CDCl_3$ )  $\delta$ : 157.1, 154.2, 139.0, 130.1, 110.8, 110.5, 106.0, 61.8, 14.6.

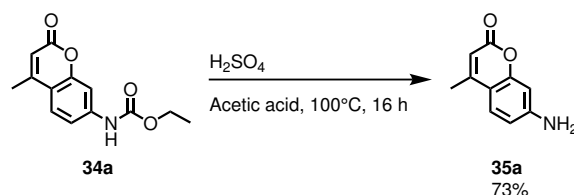
#### Preparation of ethyl (4-methyl-2-oxo-2H-chromen-7-yl)carbamate (**34a**)



Compound **33a** (20.0 g, 110.4 mmol, 1 eq) was dissolved in  $MeHSO_3$  (30 mL). Ethyl acetoacetate (17.2 g, 132.4 mmol, 1.2 eq) was added and the mixture was stirred at 60°C for 16 h. The reaction mixture was poured onto ice water and stirred for 30 min. The precipitate was filtered, washed with EtOAc and dried in vacuum to give the desired compound **34a** (21.0 g, 77% yield), which was used in the next step without further purification.

**$^1H$ -NMR** (400 MHz,  $DMSO-d_6$ )  $\delta$ : 10.1 (s, 1H), 7.65 (d,  $J$  = 8.7 Hz, 1H), 7.53 (d,  $J$  = 2.1 Hz, 1H), 7.38 (dd,  $J$  = 8.7 Hz, 2.1 Hz, 1H), 6.20 (d,  $J$  = 1.3 Hz, 1H), 4.15 (q,  $J$  = 7.1 Hz, 2H), 2.36 (d,  $J$  = 1.3 Hz, 3H), 1.25 (t,  $J$  = 7.1 Hz, 3H).

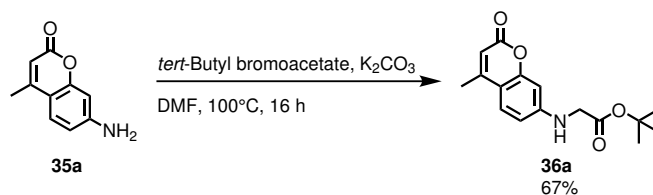
**$^{13}C$ -NMR** (100.6 MHz,  $DMSO-d_6$ )  $\delta$ : 160.2, 153.9, 153.3, 143.0, 126.0, 114.3, 111.9, 104.4, 60.8, 18.1, 14.5.

Preparation of 4-methyl-7-(methyamino)-2H-chromen-2-one (**35a**)

Compound **34a** (10.0 g, 38.3 mmol, 1 eq) was dissolved in a mixture of conc. sulfuric acid (5 mL) and acetic acid (5 mL). The reaction mixture was stirred at 100°C for 16 h. The reaction was allowed to cool to room temperature. The resulting precipitated was filtered, washed with water and dried under reduced pressure to give compound **35a** (5.3 g, 73%), which was used in the next step without further purification.

**<sup>1</sup>H-NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ: 7.39 (d, *J* = 8.6 Hz, 1H), 6.65 (dd, *J* = 8.6 Hz, 2.0 Hz, 1H), 6.41 (d, *J* = 2.0 Hz, 1H), 6.09 (s, 2H), 5.89 (s, 1H), 2.29 (s, 3H).

**<sup>13</sup>C-NMR** (100.6 MHz, DMSO-*d*<sub>6</sub>) δ: 160.7, 155.4, 153.7, 153.1, 126.2, 111.2, 108.8, 107.5, 98.5, 18.0.

Preparation of *tert*.-butyl 2-(methyl(4-methyl-2-oxo-2H-chromen-7-yl)amino)acetate (**36a**)

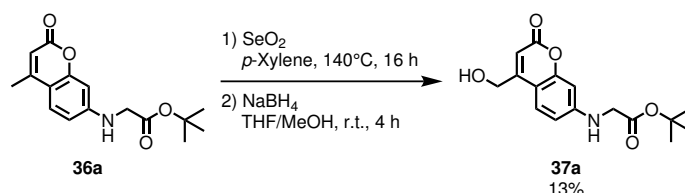
Compound **35a** (10 g, 57.1 mmol, 1 eq) and K<sub>2</sub>CO<sub>3</sub> (39.2 g, 284.8 mmol, 5 eq) were dissolved in DMF (100 mL). *tert*.-butyl bromoacetate (33.2 g, 170.9 mmol, 3 eq) were added and the mixture was stirred at 100°C for 16 h. The reaction mixture was allowed to cool to room temperature and quenched with a saturated aqueous solution of NH<sub>4</sub>Cl. The mixture was extracted three times with EtOAc. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (cylcohexane/EtOAc, 3:1) to yield compound **36a** (11.2 g, 67%).

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>) δ: 7.37 (d, *J* = 8.7 Hz, 1H), 6.64 (dd, *J* = 8.7 Hz, 2.4 Hz, 1H), 6.39 (d, *J* = 2.4 Hz, 1H), 6.00 (d, *J* = 1.3 Hz, 1H), 4.86 (q, *J* = 5.14 Hz, 1H), 3.84 (d, *J* = 4.75 Hz, 2H), 2.34 (d, *J* = 5.1 Hz, 3H), 1.51 (s, 9H).

**<sup>13</sup>C-NMR** (100.6 MHz, CDCl<sub>3</sub>) δ: 169.4, 161.9, 156.0, 153.0, 150.4, 125.7, 111.2, 110.7, 110.0, 98.5, 82.8, 45.8, 28.2, 18.7.



Preparation of *tert*.-butyl 2-((4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)amino)acetate (**37a**)

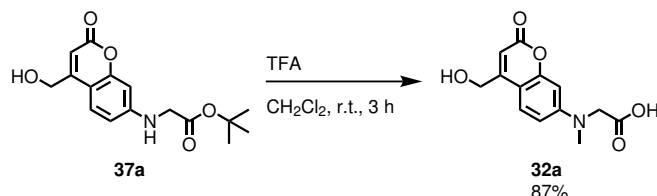


Compound **36a** (11.0 g, 38.0 mmol, 1 eq) was dissolved in *p*-xylene (50 mL). Freshly sublimated SeO<sub>2</sub> (5.1 g 45.6 mmol, 1.2 eq) was added and the mixture was vigorous stirring at 140°C for 16 h. The reaction was allowed to cool to room temperature, the mixture was filtered and concentrated under reduced pressure. The resulting oil was dissolved in a mixture of THF (20 mL) and methanol (5 mL). NaBH<sub>4</sub> (1.4 g, 38.0 mmol, 1 eq) was added and the solution was stirred for 4 h at room temperature. The suspension was neutralised with 1 M aqueous HCl, diluted with water, and partially concentrated under reduced pressure to remove methanol. The mixture was extracted three times with CH<sub>2</sub>Cl<sub>2</sub>, the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was removed under reduced pressure. The crude product was purified by column chromatography (cylcohexane/EtOAc, 3:1) to give compound **37a** (1.5 g, 13%).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ: 7.29 (d, *J* = 8.7 Hz, 1H), 6.51 (dd, *J* = 8.7 Hz, 2.3 Hz, 1H), 6.38 (d, *J* = 2.3 Hz, 1H), 6.32 (s, 1H), 4.82 (m, 3H), 3.83 (d, *J* = 3.6 Hz, 2H), 1.51 (s, 9H).

<sup>13</sup>C-NMR (100.6 MHz, CDCl<sub>3</sub>) δ: 169.5, 162.5, 156.0, 155.2, 150.4, 124.6, 110.9, 108.4, 106.6, 98.6, 82.9, 60.9, 45.7, 28.2.

Preparation of 2-((4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)amino)acetic acid (**32a**)

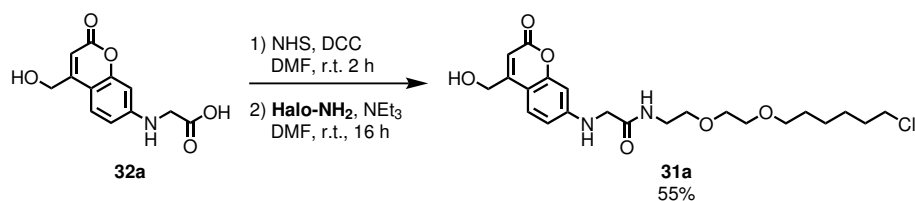


Compound **37a** (1 g, 3.3 mmol, 1 eq) was dissolved in a mixture of CH<sub>2</sub>Cl<sub>2</sub> (5 mL) and TFA (5 mL). The mixture was stirred at room temperature for 3 h. The solvent was evaporated under reduced pressure and the crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 10:1, 0.1% acetic acid) to give compound **32a** (715 mg, 87%).

<sup>1</sup>H-NMR (400 MHz, DMSO - d<sub>6</sub>) δ: 7.36 (d, *J* = 8.80 Hz, 1H); 6.72 (s, 1H), 6.62 (dd, *J* = 8.8 Hz, 1.9 Hz, 1H), 6.40 (d, *J* = 1.9 Hz, 1H), 4.65 (s, 2H), 3.80 (s, 2H).

<sup>13</sup>C-NMR (100.6 MHz, DMSO - d<sub>6</sub>) δ: 171.8, 160.6, 155.9, 152.3, 151.2, 125.8, 110.5, 108.4, 106.6, 97.2, 44.2, 41.5.

**Preparation of N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-((4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)amino)acetamide (31a)**

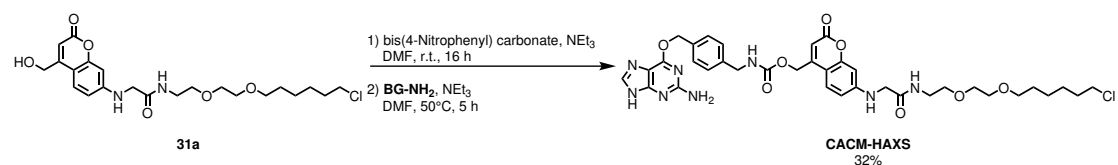


Compound **32a** (500 mg, 2.0 mmol, 1 eq) was dissolved in DMF. DCC (414 mg, 2.0 mmol, 1 eq) and NHS (230 mg, 2.0 mmol, 1 eq) were added and the mixture was stirred at room temperature for 2 h. **Halo-NH<sub>2</sub>** (5.18 mg, 2.0 mmol, 1 eq) and NEt<sub>3</sub> (560  $\mu$ L, 4.0 mmol, 2 eq) were added and the solution was further stirred at room temperature for 16 h. The reaction was quenched with a saturated aqueous solution of NH<sub>4</sub>Cl, extracted three times with EtOAc and the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated under reduced pressure and the crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1) to give compound **31a** (500 mg, 55%).

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$ : 7.31 (d, *J* = 8.7 Hz, 1H), 6.81 (t, *J* = 5.3 Hz, 1H), 6.52 (dd, *J* = 8.7 Hz, 2.4 Hz, 1H), 6.41 (d, *J* = 2.4 Hz, 1H), 6.33 (t, *J* = 1.2 Hz, 1H), 5.08 (t, *J* = 5.3 Hz, 1H), 4.81 (t, *J* = 2.4 Hz, 2H), 3.83 (d, *J* = 5.2 Hz, 2H), 3.61 - 3.41 (m, 12 H), 1.79 - 1.74 (m, 2H), 1.62-1.54 (m, 2H), 1.47 - 1.32 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, DMSO - d<sub>6</sub>)  $\delta$ : 169.3, 161.1, 157.0, 155.4, 151.9, 124.8, 110.7, 106.8, 104.2, 97.0, 70.2, 69.6, 69.4, 69.1, 69.0, 59.1, 46.0, 45.4, 32.0, 29.1, 26.1, 24.9.

**Preparation of (7-((2-((2-((5-chloropentyl)oxy)ethoxy)ethyl)amino)-2-oxoethyl)amino)-2-oxo-2H-chromen-4-yl)methyl 4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzylcarbamate (CACM-HaXS)**



Compound **31a** (300 mg, 0.66 mmol, 1 eq) was added slowly to a solution of bis(4-nitrophenyl) carbonate (201 mg, 0.66 mmol, 1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) at 0°C. NEt<sub>3</sub> (138  $\mu$ L, 0.99 mmol, 1.5 eq) was added and the mixture was stirred at room temperature for 3 h. The reaction was quenched with a saturated aqueous solution of NH<sub>4</sub>Cl. The product was extracted three times with EtOAc. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. Filtration over a short pad of silica gave the carbonate of **31a**, which was directly used for the carbamate formation. Therefore, the carbonate was dissolved in DMF (3 mL), **BG-NH<sub>2</sub>** (178 mg, 0.66 mmol, 1 eq) and NEt<sub>3</sub> (138  $\mu$ L, 0.99 mmol, 1.5 eq) were added and the solution was stirred for 3 h at 60°C. The reaction was quenched with water and the product was extracted three times with EtOAc. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 10:1) to give compound **CACM-HaXS** (158 mg, 32%).

**<sup>1</sup>H-NMR** (400 MHz, DMSO - d<sub>6</sub>)  $\delta$ : 12.42, (s, 1H), 8.08 (t, *J* = 6.0 Hz, 1H), 8.05 (t, *J* = 6.0 Hz, 1H), 7.78 (s, 1H), 7.44 (d, *J* = 7.9 Hz, 2H), 7.40 (d, *J* = 8.8 Hz, 1H), 7.28 (d, *J* = 7.9 Hz,

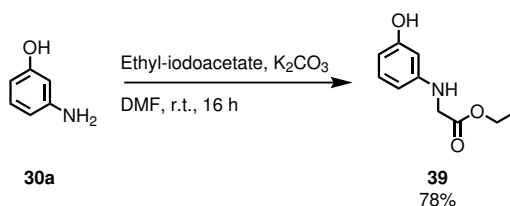
2H), 7.00 (t,  $J = 6.0$  Hz, 1H), 6.61 (dd,  $J = 8.8$  Hz, 2.3 Hz, 1H), 6.38 (d,  $J = 2.3$  Hz, 1H), 5.98 (s, 1H), 5.44 (s, 2H), 5.21 (s, 2H), 4.23 (d,  $J = 6.1$  Hz, 2H), 3.74 (d,  $J = 5.9$  Hz, 2H), 3.57 (t,  $J = 6.6$  Hz, 2H), 3.45-3.23 (m, 12H), 1.71-1.64 (m, 2H), 1.49-1.42 (m, 2H), 1.39-1.25 (m, 4H).

**$^{13}\text{C}$ -NMR** (100.6 MHz, DMSO- $d_6$ )  $\delta$ : 169.2, 160.6, 159.9, 159.7, 155.7, 155.5, 155.2, 152.2, 152.1, 139.3, 137.8, 135.5, 128.6, 127.1, 125.0, 113.5, 110.6, 106.3, 104.8, 97.0, 70.2, 69.6, 69.4, 69.0, 66.5, 61.1, 55.8, 45.9, 45.3, 43.7, 32.0, 29.0, 26.1, 24.9.

Photophysical properties:  $\lambda_{max}$ : 365 nm,  $\epsilon_{360}$ : 23120 M $^{-1}\text{cm}^{-1}$ ,  $\epsilon_{405}$ : 2622 M $^{-1}\text{cm}^{-1}$ ,  $\phi_{360}$ : 0.006,  $\phi_{405}$ : 0.016.

### 5.3.1.9 Alternative synthesis of **32a**

#### Preparation of ethyl 2-((3-hydroxyphenyl)amino)acetate (**39**)

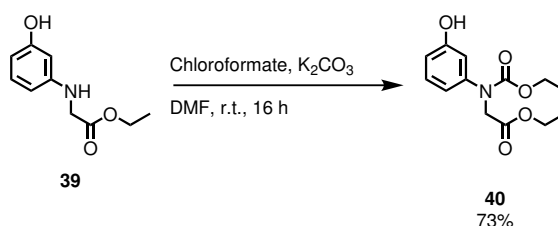


3-Aminophenol (**30a**) (10 g, 91.7 mmol, 1.2 eq) and  $\text{K}_2\text{CO}_3$  (12.67 g, 91.7 mmol, 1.2 eq) were dissolved in DMF (150 mL). Ethyl-2-iodoacetate (9.0 mL, 76.4 mmol, 1 eq) was added and the mixture was stirred at room temperature for 16 h. The reaction was quenched with a saturated aqueous solution of  $\text{NH}_4\text{Cl}$ . The product was extracted three times with EtOAc, the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was evaporated under reduced pressure. The resulting oil was purified by column chromatography (cylcohexane/EtOAc, 3:1) to give the desired compound **39** (11.6 g, 78%).

**$^1\text{H}$ -NMR** (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 6.98 (t,  $J = 8.1$  Hz, 1H), 6.64 (dd  $J = 8.10, 1.94$  Hz, 1H), 6.14 (dd,  $J = 8.1, 1.94$  Hz, 1H) 6.06 (t,  $J = 1.94$ , 1H), 4.20 (q,  $J = 7.12$  Hz, 2H), 3.81 (s, 2H), 1.25 (t,  $J = 7.12$  Hz, 3H).

**$^{13}\text{C}$ -NMR** (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$ : 171.9, 157.0, 148.5, 130.3, 105.7, 105.6, 100.5, 61.6, 45.8, 14.1.

#### Preparation of ethyl 2-((ethoxycarbonyl)(3-hydroxyphenyl)amino)acetate (**40**)

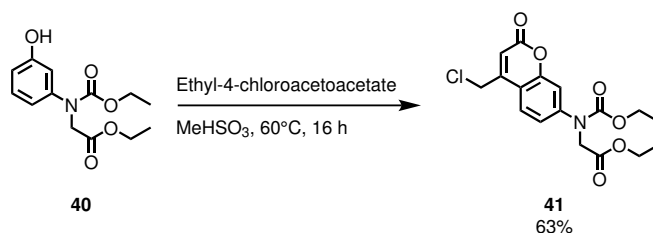


Compound **39** (10 g, 51.2 mmol, 1.2 eq) and  $\text{K}_2\text{CO}_3$  (7.0 g, 51.2 mmol, 1.2 eq) were dissolved in DMF (100 mL). Chloroformate was added and the mixture was stirred at room temperature for 16 h. The reaction was quenched with a saturated aqueous solution of  $\text{NH}_4\text{Cl}$ . The product was extracted three times with EtOAc, the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (cylcohexane/EtOAc, 3:1) to give the desired compound **40** (9.9 g, 73%).

**<sup>1</sup>H-NMR** (400 MHz, DMSO – d<sub>6</sub>) δ: 9.52 (s, 1H), 7.13 (t, J = 8.1 Hz, 1H), 6.71-6.69 (m, 2H), 6.14 (dd, J = 8.11, 1.94 Hz, 1H), 4.29 (s, 2H), 4.14 (q, J = 7.01 Hz, 2H), 4.06 (q, J = 7.01 Hz, 2H), 1.20 (t, J = 7.01 Hz, 3H), 1.13 (t, J = 7.01 Hz, 3H).

**<sup>13</sup>C-NMR** (100.6 MHz, DMSO – d<sub>6</sub>) δ: 169.5, 157.5, 154.5, 143.2, 129.6, 116.7, 115.9, 113.5, 61.4, 60.7, 51.9, 14.4, 14.04.

**Preparation of ethyl 2-((4-(chloromethyl)-2-oxo-2H-chromen-7-yl)(ethoxycarbonyl)amino)acetate (41)**

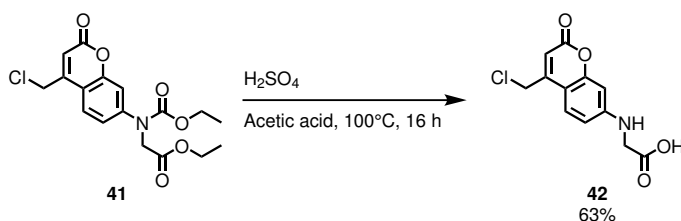


Compound **40** (5 g, 18.7 mmol, 1 eq) was dissolved in MeHSO<sub>3</sub> (8 mL). Ethyl 4-chloroacetoacetate (3.7 g, 22.4 mmol, 1.2 eq) was added and the mixture was stirred at 60°C for 16 h. The reaction was quenched with water. The product was extracted three times with EtOAc, the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (cylcohexane/EtOAc, 3:1) to give the desired compound **41** (4.3 g, 63%).

**<sup>1</sup>H-NMR** (400 MHz, CDCl<sub>3</sub>) δ: 7.65 (d, J = 8.59 Hz, 1H), 7.33 (m, 2H), 6.54 (t, J = 0.66 Hz, 1H), 5.30 (s, 1H), 4.65 (d, J = 0.66 Hz, 2H), 4.40 (s, 2H), 4.27-4.22 (m, 4H), 1.32-1.25 (m, 6H).

**<sup>13</sup>C-NMR** (100.6 MHz, CDCl<sub>3</sub>) δ: 169.8, 169.3, 160.3, 154.9, 154.2, 149.2, 145.9, 129.9, 124.6, 122.4, 115.5, 63.0, 61.8, 51.9, 41.3, 14.5, 14.3.

**Preparation of ethyl 2-((4-(chloromethyl)-2-oxo-2H-chromen-7-yl)(ethoxycarbonyl)amino)acetate (42)**

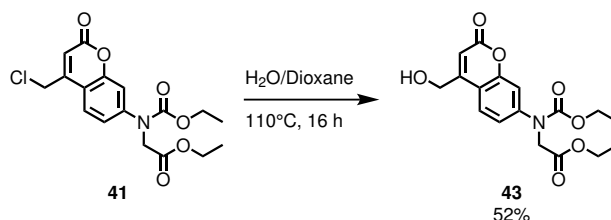


Compound **41** (2 g, 2.7 mmol) was dissolved in a mixture of conc. sulfuric acid (2 mL) and acetic acid (2 mL). The reaction mixture was stirred at 100°C for 16 h. The reaction was allowed to cool to room temperature and set to pH 5.3. The resulting precipitated was filtered, washed with ice cold EtOH and dried under reduced pressure to give compound **42** (499 mg, 69%), which was used in the next step without further purification.

**<sup>1</sup>H-NMR** (400 MHz, DMSO – d<sub>6</sub>) δ: 7.53 (d, J = 8.8 Hz, 1H), 6.68 (dd, J = 8.8, 2.2 Hz, 1H), 6.45 (d, J = 2.2 Hz, 1H), 6.22 (s, 1H), 4.89 (s, 2H), 3.95 (s, 2H).

**<sup>13</sup>C-NMR** (100.6 MHz, CDCl<sub>3</sub>) δ: 171.8, 160.6, 155.9, 152.3, 151.3, 138.0, 125.8, 110.5, 108.4, 106.6, 97.2, 44.1, 41.5.

Preparation of ethyl 2-((4-(chloromethyl)-2-oxo-2H-chromen-7-yl)(ethoxycarbonyl)amino)acetate (**43**)

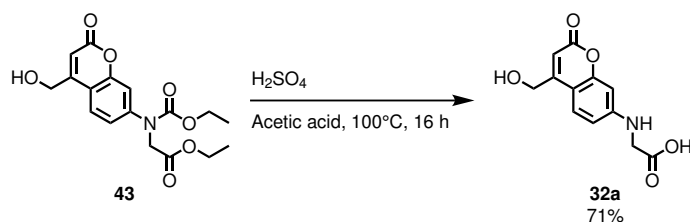


Compound **41** (2 g, 2.7 mmol) was dissolved in a mixture of water and dioxane (10 mL each). The reaction mixture was stirred at 110°C for 16 h and diluted with water. The product was extracted with three times EtOAc, the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent evaporated under reduced pressure. The product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 40:1) to give the desired compound **22** (490 mg, 52%).

<sup>1</sup>H-NMR (400 MHz, DMSO - d<sub>6</sub>) δ: 7.68 (d, J = 8.6 Hz, 1H), 7.36 (dd, J = 2.2 Hz, 1H), 7.31 (dd, J = 8.6, 2.2 Hz, 1H), 6.43 (s, 1H), 5.67 (t, J = 5.5 Hz, 1H), 4.75 (d, J = 5.5 Hz, 2H), 4.47 (s, 2H), 4.18-4.10 (m, 4H), 1.22-1.14 (m, 6H).

<sup>13</sup>C-NMR (100.6 MHz, CDCl<sub>3</sub>) δ: 169.2, 160.1, 154.8, 154.2, 149.2, 145.9, 124.5, 122.3, 115.5, 115.3, 114.4, 67.2, 62.9, 61.7, 51.9, 14.5, 14.3.

Preparation of ethyl 2-((4-(chloromethyl)-2-oxo-2H-chromen-7-yl)(ethoxycarbonyl)amino)acetate (**32a**)

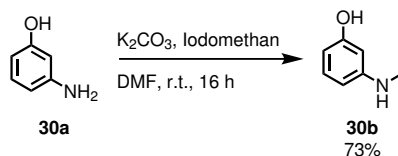


Compound **43** (200 mg, 0.6 mmol) was dissolved in a mixture of conc. sulfuric acid (1 mL) and acetic acid (1 mL). The reaction mixture was stirred at 100°C for 16 h. The reaction was allowed to cool to room temperature and set to pH 5.3. The resulting precipitated was filtered, washed with ice cold EtOH and dried under reduced pressure to give compound **32a** (101 mg, 71%), which was used in the next step without further purification.

<sup>1</sup>H-NMR (400 MHz, DMSO - d<sub>6</sub>) δ: 7.36 (d, J = 8.80 Hz, 1H); 6.72 (s, 1H), 6.62 (dd, J = 8.8 Hz, 1.9 Hz, 1H), 6.40 (d, J = 1.9 Hz, 1H), 4.65 (s, 2H), 3.80 (s, 2H).

<sup>13</sup>C-NMR (100.6 MHz, DMSO - d<sub>6</sub>) δ: 171.8, 160.6, 155.9, 152.3, 151.2, 125.8, 110.5, 108.4, 106.6, 97.2, 44.2, 41.5.

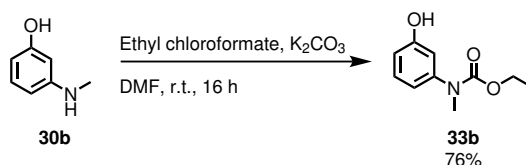
## 5.3.1.10 Synthesis of CMACM-HaXS1 and CMACM-HaXS2

Preparation of 3-methyl aminophenol (**30b**)

3-Aminophenol (**30a**) (20 g, 183.2 mmol, 1.2 eq) and  $\text{K}_2\text{CO}_3$  (25.2 g, 183.2 mmol, 1.2 eq), were dissolved in DMF (100 mL). Iodomethan (21.6 g, 152.8 mmol, 1 eq) was added dropwise and the mixture was stirred at room temperature for 16 h. The reaction was quenched with a saturated aqueous solution of  $\text{NH}_4\text{Cl}$ . The product was extracted three times with EtOAc, the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (cylcohexane/EtOAc, 3:1) to give the desired compound **30b** (18.0 g, 73%).

**$^1\text{H-NMR}$**  (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.02 (t,  $J = 8.0$ , 1H), 6.20 (m, 2H), 6.08 (t,  $J = 2.3$  Hz, 1H), 4.37 (m, 2H), 2.76 (s, 3H).

**$^{13}\text{C-NMR}$**  (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$ : 156.9, 150.9, 130.3, 105.9, 104.9, 99.9, 31.0.

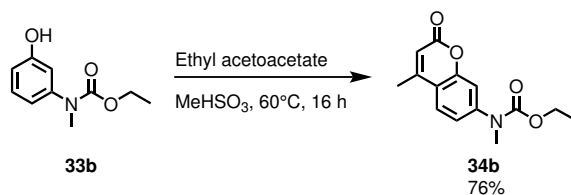
Preparation of ethyl (3-hydroxyphenyl)(methyl)carbamate (**33b**)

Compound **30b** (13.0 g, 105.5 mmol, 1 eq) and  $\text{K}_2\text{CO}_3$  (21.9 g, 158.3 mmol, 1.5 eq), were dissolved in DMF (100 mL). Ethyl chloroformate (17.2 g, 158.3 mmol, 1.5 eq) was added and the mixture was stirred for 16 h at room temperature. The reaction was quenched with a saturated aqueous solution of  $\text{NH}_4\text{Cl}$ . The product was extracted three times with EtOAc, the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the crude product was purified by column chromatography (cylcohexane/EtOAc, 3:1) to give the desired compound **33b** (15.6 g, 76%).

**$^1\text{H-NMR}$**  (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.15 (t,  $J = 8.0$  Hz, 1H), 6.71 (m, 3H), 4.18 (q,  $J = 7.1$  Hz, 2H), 3.25 (s, 3H), 1.25 (t,  $J = 7.1$  Hz, 3H).

**$^{13}\text{C-NMR}$**  (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$ : 156.8, 156.3, 144.0, 129.6, 117.3, 113.7, 113.3, 62.2, 37.8, 14.6.7.

Preparation of ethyl methyl(4-methyl-2-oxo-2H-chromen-7-yl)carbamate (**34b**)

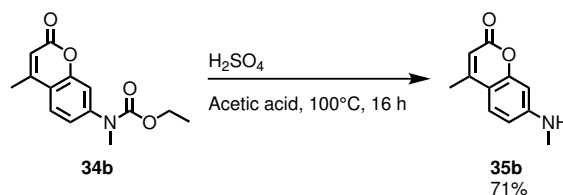


Compound **33b** (10.0 g, 51.2 mmol, 1 eq) was dissolved in MeHSO<sub>3</sub> (20 mL). Ethyl acetoacetate (8.0 g, 61.5 mmol, 1.2 eq) was added and the mixture was stirred for 16 h at 60°C. The reaction mixture was poured onto ice water and stirred for 30 min. The precipitate was filtered and washed with EtOAc. The precipitate was dried in vacuum to give the desired compound **34b** (10.4 g, 78% yield), which was used in the next step without further purification.

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ: 7.71 (d, J = 8.5 Hz, 1H), 7.37 (m, 2H), 6.33 (s, 1H), 4.12 (q, J = 7.1 Hz, 2H), 3.29 (s, 3H), 2.41 (s, 3H), 1.20 (t, J = 7.1 Hz, 3H).

<sup>13</sup>C-NMR (100.6 MHz, CDCl<sub>3</sub>) δ: 159.8, 154.3, 153.0, 152.9, 146.1, 125.3, 120.8, 116.5, 113.4, 112.0, 61.6, 36.7, 18.0, 14.3.

Preparation of 4-methyl-7-(methylamino)-2H-chromen-2-one (**35b**)

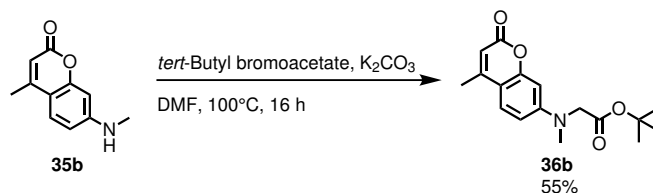


Compound **34b** (10.0 g, 38.3 mmol, 1 eq) was dissolved in a mixture of conc. sulfuric acid (5 mL) and acetic acid (5 mL). The reaction mixture was stirred at 100°C for 16 h. The reaction was allowed to cool to room temperature. The resulting precipitate was filtered, washed with water and dried under reduced pressure to give compound **35b** (5.3 g, 73%).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ: 7.35 (d, J = 8.6 Hz, 1H), 6.52 (dd, J = 8.6 Hz, 2.3 Hz, 1H), 6.44 (d, J = 2.3 Hz, 1H), 5.97 (d, J = 1.0 Hz, 1H), 4.39 (s, 1H), 2.90 (s, 3H), 2.34 (d, J = 1.0 Hz, 3H).

<sup>13</sup>C-NMR (100.6 MHz, CDCl<sub>3</sub>) δ: 162.1, 156.0, 153.1, 152.5, 125.4, 110.4, 110.1, 109.3, 97.7, 30.2, 18.6.

Preparation of *tert*.-butyl 2-(methyl(4-methyl-2-oxo-2H-chromen-7-yl)amino)acetate (**36b**)



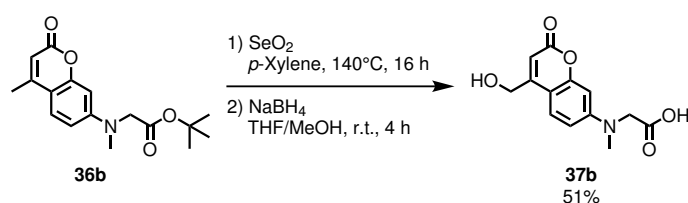
Compound **35b** (5 g, 26.5 mmol, 1 eq) and K<sub>2</sub>CO<sub>3</sub> (18.2 g, 132.2 mmol, 5 eq) were dissolved in DMF (50 mL). *tert*.-butyl bromoacetate (15.4 g, 79.3 mmol, 3 eq) were added and the mixture

was stirred at 100°C for 16 h. The reaction mixture was allowed to cool to room temperature and quenched with a saturated aqueous solution of  $\text{NH}_4\text{Cl}$ . The mixture was extracted three times with EtOAc. The combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (cylcohexane/EtOAc, 3:1) to give compound **36b** (4.4 g, 55%).

**$^1\text{H-NMR}$**  (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.41 (d,  $J = 8.9$  Hz, 1H), 6.59 (dd,  $J = 8.9$  Hz, 2.6 Hz, 1H), 6.50 (d,  $J = 2.6$  Hz, 1H), 5.99 (d,  $J = 1.1$  Hz, 1H), 4.01 (s, 2H), 3.13 (s, 3H), 2.35 (d,  $J = 1.1$  Hz, 3H), 1.44 (s, 9H).

**$^{13}\text{C-NMR}$**  (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$ : 169.1, 162.1, 155.7, 152.9, 152.1, 125.6, 110.6, 110.0, 108.9, 98.8, 82.4, 55.0, 39.9, 28.2, 18.6.

#### Preparation of *tert.*-butyl 2-((4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)amino)acetate (**37b**)

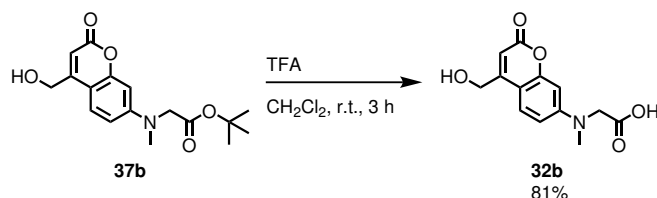


Compound **36b** (4.1 g, 13.5 mmol, 1 eq) was dissolved in *p*-xylene (50 mL). Freshly sublimated  $\text{SeO}_2$  (1.8 g 16.2 mmol, 1.2 eq) was added and the mixture was vigorous stirring at 120°C for 16 h. After the reaction cooled down to room temperature the mixture was filtered and concentrated under reduced pressure. The resulting oil was dissolved in a mixture of THF (20 mL) and methanol (5 mL). Sodium borohydride (500 mg, 13.5 mmol, 1 eq) was added and the solution was stirred for 4 h at room temperature. The suspension was neutralised with 1 M aqueous HCl, diluted with water, and partially concentrated under reduced pressure to remove methanol. The mixture was extracted three times with  $\text{CH}_2\text{Cl}_2$ , the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$ , and the solvent was removed under reduced pressure. The crude product was purified by column chromatography (cylcohexane/EtOAc, 3:1) to yield compound **37b** (2.2 g, 51%).

**$^1\text{H-NMR}$**  (400 MHz,  $\text{CDCl}_3$ )  $\delta$ : 7.28 (d,  $J = 8.9$  Hz, 1H), 6.53 (dd,  $J = 8.9$  Hz, 2.6 Hz, 1H), 6.46 (d,  $J = 2.6$  Hz, 1H), 6.30 (t,  $J = 1.37$  Hz, 1H), 4.76 (s, 2 H), 4.00 (s, 2H), 3.21 (s, 1H), 3.10 (s, 3H), 1.45 (s, 9H).

**$^{13}\text{C-NMR}$**  (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$ : 169.3, 162.7, 155.6, 155.3, 151.9, 124.3, 109.1, 107.8, 106.4, 98.8, 82.6, 60.7, 54.9, 39.8, 28.2.

#### Preparation of 2-((4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)(methylamino)acetic acid (**32b**)





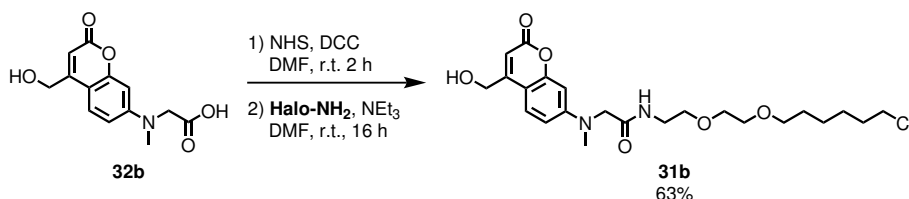
## Experimental Section

Compound **37b** (2 g, 6.3 mmol, 1 eq) was dissolved in a mixture of CH<sub>2</sub>Cl<sub>2</sub> (5 mL) and TFA (5 mL). The mixture was stirred at room temperature for 3 h. The solvent was evaporated under reduced pressure and the crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 10:1, 0.1% acetic acid) to give compound **32b** (1.3 g, 81%).

<sup>1</sup>H-NMR (400 MHz, DMSO – d<sub>6</sub>) δ: 7.36 (d, J = 8.9 Hz, 1H), 6.60 (d, J = 8.9 Hz, 1H), 6.44 (s, 1H), 6.06 (s, 1H), 4.64 (s, 2H), 3.87 (s, 2H), 3.02 (s, 3H).

<sup>13</sup>C-NMR (100.6 MHz, DMSO – d<sub>6</sub>) δ: 161.2, 157.0, 155.2, 152.6, 124.4, 109.0, 105.9, 104.0, 97.4, 59.0, 55.8, 39.7.

### Preparation of N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-((4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)(methyl)amino)acetamideacid (**31b**)

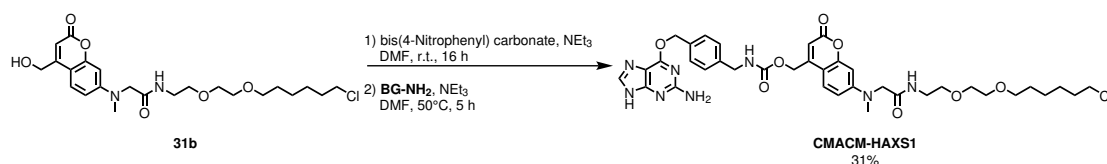


Compound **32b** (500 mg, 1.9 mmol, 1 eq) was dissolved in DMF (3 mL). DCC (392 mg, 1.9 mmol, 1 eq) and NHS (219 mg, 1.9 mmol, 1 eq) were added and the mixture was stirred at room temperature for 2 h. Halo-NH<sub>2</sub> (492 mg, 1.9 mmol, 1 eq) and NEt<sub>3</sub> (530 μL, 3.8 mmol, 2 eq) were added and the mixture was further stirred at room temperature for 16 h. The reaction was quenched with a saturated aqueous solution of NH<sub>4</sub>Cl, extracted three times with EtOAc and the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated under reduced pressure and the crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1) to give compound **31b** (553 mg, 63%).

<sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>) δ: 7.23 (d, J = 8.8 Hz, 1H), 6.94 (t, J = 5.4 Hz, 1H), 6.53 (dd, J = 8.8 Hz, 2.6 Hz, 1H), 6.49 (d, J = 2.6 Hz, 1H), 6.24 (t, J = 1.4 Hz, 1H); 4.69 (d, J = 1.4 Hz, 2H), 3.97 (s, 2H), 3.55-3.36 (m, 12 H), 3.10 (s, 3H), 1.76-1.69 (m, 2H), 1.56-1.49 (m, 2H), 1.44-1.28 (m, 4H).

<sup>13</sup>C-NMR (100.6 MHz, CDCl<sub>3</sub>) δ: 169.3, 162.1, 155.5, 154.9, 151.9, 124.4, 109.5, 108.3, 107.0, 99.4, 71.4, 70.3, 70.00, 69.9, 60.6, 57.3, 45.2, 39.8, 39.3, 32.6, 29.5, 26.8, 25.5.

### Preparation of (7-((2-((2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)amino)-2-oxoethyl)(methyl)amino)-2-oxo-2H-chromen-4-yl)methyl (4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)carbamate (CMACM-HaXS1)



Compound **31b** (400 mg, 0.85 mmol, 1 eq) was added slowly to a solution of bis(4-nitrophenyl) carbonate (260 mg, 0.85 mmol, 1 eq) in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) at 0°C. NEt<sub>3</sub> (178 μL, 1.28 mmol, 1.5 eq) was added and the mixture was stirred for 3 h at rt. The reaction was quenched with a saturated aqueous solution of NH<sub>4</sub>Cl. The product was extracted three times with EtOAc the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced

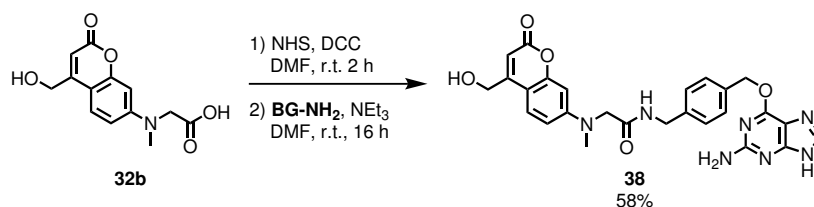
pressure. Filtration over a short pad of silica gave the carbonate of **31b**, which was directly used for the carbamate formation. Therefore, the carbonate was dissolved in DMF (5 mL), **BG-NH<sub>2</sub>** (300 mg, 0.66 mmol, 1 eq) and NEt<sub>3</sub> (178  $\mu$ L, 1.28 mmol, 1.5 eq) were added and the solution was stirred for 3 h at 60°C. The reaction was quenched with water and the product was extracted three times with EtOAc. The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 10:1) to give compound **CMACM-HaXS1** (350 mg, 46%).

**<sup>1</sup>H-NMR** (400 MHz, DMSO - d<sub>6</sub>)  $\delta$ : 12.4 (s, 1H), 8.13-8.09 (m, 2H), 7.80 (s, 1H), 7.49-7.45 (m, 3H), 7.30 (d, *J* = 8.0 Hz, 2H), 6.65 (dd, *J* = 9.0, 2.4 Hz, 1H), 6.53 (d, *J* = 2.4 Hz, 1H), 6.29 (s, 2H), 6.04 (s, 1H), 5.46 (s, 2H), 5.25 (s, 2H), 4.24 (d, *J* = 6.0 Hz, 2H), 4.06 (s, 2H), 3.60 (t, *J* = 6.6 Hz, 2H), 3.54-3.20 (m, 10H), 3.07 (s, 3H), 1.71-1.65 (m, 2H), 1.48-1.43 (m, 2H), 1.37-1.25 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, DMSO - d<sub>6</sub>)  $\delta$ : 168.7, 160.6, 159.8, 159.6, 155.7, 155.2, 152.3, 151.9, 139.3, 137.8, 135.4, 129.6, 128.5, 127.1, 125.0, 113.5, 109.2, 106.2, 105.3, 97.8, 70.2, 69.6, 69.4, 69.0, 66.5, 61.1, 55.8, 44.4, 43.8, 40.0, 38.5, 32.0, 29.0, 26.1, 24.9.

$\lambda_{max}$ : 370 nm,  $\epsilon_{360}$ : 11230 M<sup>-1</sup>cm<sup>-1</sup>,  $\epsilon_{405}$ : 2411 M<sup>-1</sup>cm<sup>-1</sup>.

#### Preparation of N-(4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)-2-((4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)(methyl)amino)acetamide (**38**)

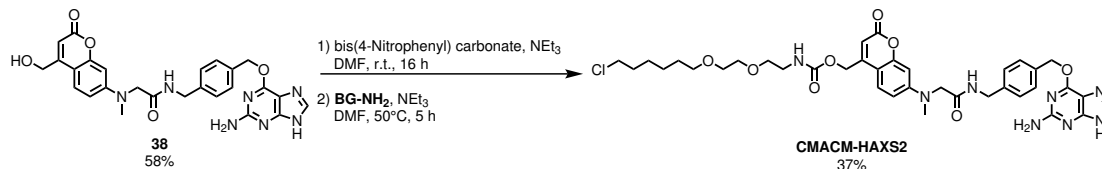


Compound **32b** (300 mg, 1.1 mmol, 1 eq) was dissolved in DMF (5 mL). DCC (227 mg, 1.1 mmol, 1 eq) and NHS (127 mg, 1.1 mmol, 1 eq) were added and the mixture was stirred at room temperature for 2 h. **BG-NH<sub>2</sub>** (286 mg, 1.1 mmol, 1 eq) and NEt<sub>3</sub> (153  $\mu$ L, 1.1 mmol, 1 eq) were added and the mixture was stirred at room temperature for 16 h. The reaction was quenched with a saturated aqueous solution of NH<sub>4</sub>Cl, extracted three times with EtOAc and the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated under reduced pressure and the crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 10:1) to give compound **38** (329 mg, 58%).

**<sup>1</sup>H-NMR** (400 MHz, DMSO - d<sub>6</sub>)  $\delta$ : 12.44 (s, 1H), 8.56 (t, *J* = 6.1 Hz, 1H), 7.81 (s, 1H), 7.46 (d, *J* = 9.0 Hz, 1H), 7.43 (d, *J* = 8.2 Hz, 2H), 7.26 (d, *J* = 8.2 Hz, 2H), 6.65 (dd, *J* = 9.0 Hz, 2.5 Hz, 1H), 6.55 (d, *J* = 2.5 Hz, 1H), 6.29 (s, 2H), 6.12 (t, *J* = 1.4 Hz, 1H), 5.57 (s, 2H), 4.69 (dd, *J* = 5.7 Hz, 1.5 Hz, 2H), 4.29 (d, *J* = 5.9 Hz, 2H), 4.12 (s, 2H), 3.09 (s, 3H).

**<sup>13</sup>C-NMR** (100.6 MHz, DMSO - d<sub>6</sub>)  $\delta$ : 168.8, 161.0, 159.6, 156.9, 155.2, 155.1, 152.1, 139.2, 137.8, 135.3, 134.5, 128.5, 127.2, 124.8, 118.4, 109.0, 106.7, 104.7, 97.8, 66.5, 59.0, 55.0, 41.9.

**Preparation of (7-((2-((4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)amino)-2-oxoethyl)(methyl)amino)-2-oxo-2H-chromen-4-yl)methyl (2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)carbamate (CMACM-HaXS2)**



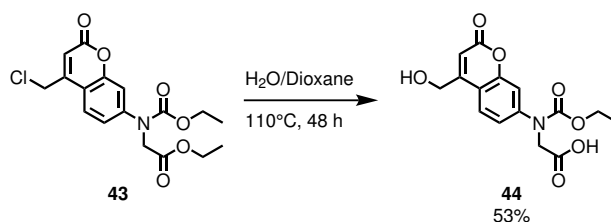
Compound **38** (200 mg, 0.39 mmol, 1 eq) was added slowly to a solution of bis(4-nitrophenyl) carbonate (120 mg, 0.39 mmol, 1 eq) in  $\text{CH}_2\text{Cl}_2$  (5 mL) at 0°C.  $\text{NEt}_3$  (83  $\mu\text{L}$ , 0.59 mmol, 1.5 eq) was added and the mixture was stirred for 3 h at rt. The reaction was quenched with a saturated aqueous solution of  $\text{NH}_4\text{Cl}$ . The product was extracted three times with EtOAc the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was evaporated under reduced pressure. Filtration over a short pad of silica gave the carbonate of **38**, which was directly used for the carbonate formation. Therefore, the carbonate was dissolved in DMF (3 mL), **Halo-NH<sub>2</sub>** (101 mg, 0.39 mmol, 1 eq) and  $\text{NEt}_3$  (108  $\mu\text{L}$ , 0.78 mmol, 2 eq) were added. The solution was stirred for 3 h at 60°C. The reaction was quenched with water and the product was extracted three times with EtOAc. The combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ , 10:1) to give compound **CMACM-HaXS2** (120 mg, 41%).

**<sup>1</sup>H-NMR** (400 MHz, DMSO -  $d_6$ )  $\delta$ : 12.44 (s, 1H), 8.58 (s, 1H), 7.80 (s, 1H), 7.57 (s, 1H), 7.48 (d,  $J$  = 9.0 Hz, 1H), 7.44 (d,  $J$  = 8.0 Hz, 2H), 7.27 (d,  $J$  = 8.0 Hz, 2H), 6.69 (dd,  $J$  = 9.0, 2.5 Hz, 1H), 6.57 (d,  $J$  = 2.5 Hz, 1H), 6.29 (s, 1H), 6.02 (s, 1H), 5.44 (s, 2H), 5.24 (s, 1H), 4.29 (d,  $J$  = 5.9 Hz, 1H), 4.14 (s, 1H), 3.60 (t,  $J$  = 6.6 Hz, 2H), 3.51-3.23 (m, 10H), 3.10 (s, 3H), 1.70-1.65 (m, 2H), 1.50-1.44 (m, 2H), 1.32-1.26 (m, 4H).

**<sup>13</sup>C-NMR** (100.6 MHz, DMSO -  $d_6$ )  $\delta$ : 168.7, 160.6, 159.6, 159.8, 155.5, 155.2, 152.4, 152.0, 147.9, 137.8, 135.3, 129.6, 128.5, 127.2, 125.0, 113.5, 109.3, 106.2, 105.2, 97.9, 70.2, 69.6, 69.4, 69.0, 66.5, 60.9, 55.0, 45.4, 41.9, 32.0, 29.0, 26.1, 24.9. Photophysical data:  $\lambda_{\text{max}}$ : 370 nm,  $\epsilon_{360}$ : 11598  $\text{M}^{-1}\text{cm}^{-1}$ ,  $\epsilon_{405}$ : 2390  $\text{M}^{-1}\text{cm}^{-1}$ .

### 5.3.1.11 Synthesis of GUCM-HaXS

**Preparation of ethyl 2-((3-hydroxyphenyl)amino)acetate (44)**

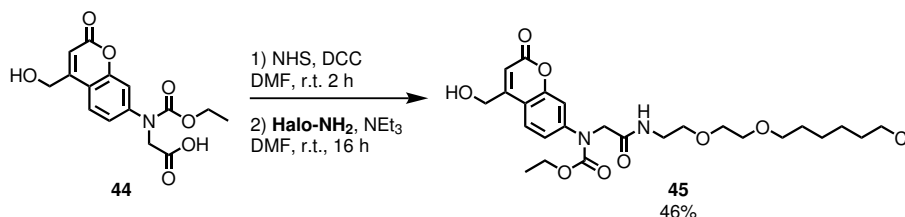


Compound **41** (4.0 g, 10.8 mmol) was dissolved in a mixture of water and dioxane. The reaction mixture was stirred at 110°C for 48 h. The reaction mixture was diluted with water and acidified with an aqueous 1M HCl solution. The product was extracted three times with EtOAc, the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent evaporated under reduced pressure. The product was purified by column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  20:1 with 0.1% acetic acid) to give the desired compound **44** (1.85 g, 53%).

**$^1\text{H}$  NMR** (400 MHz, DMSO- $d_6$ )  $\delta$ : 7.67 (d,  $J$  = 8.63 Hz, 1H), 7.35 (d,  $J$  = 2.26 Hz, 1H), 7.31 (dd,  $J$  = 8.63, 2.26 Hz, 1H), 6.43 (t,  $J$  = 1.53 Hz, 1H), 4.76 (d,  $J$  = 1.53 Hz, 2H), 4.38 (s, 2H), 4.12 (q,  $J$  = 7.02 Hz, 2H), 1.17 (t,  $J$  = 7.02 Hz, 3H).

**$^{13}\text{C}$ -NMR** (100.6 MHz, DMSO- $d_6$ )  $\delta$ : 170.8, 160.1, 156.1, 154.2, 153.0, 145.1, 124.3, 121.7, 114.8, 113.0, 110.0, 61.9, 59.0, 51.3, 14.3.

**Preparation of ethyl (2-((2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)amino)-2-oxoethyl)(4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)carbamate (45)**

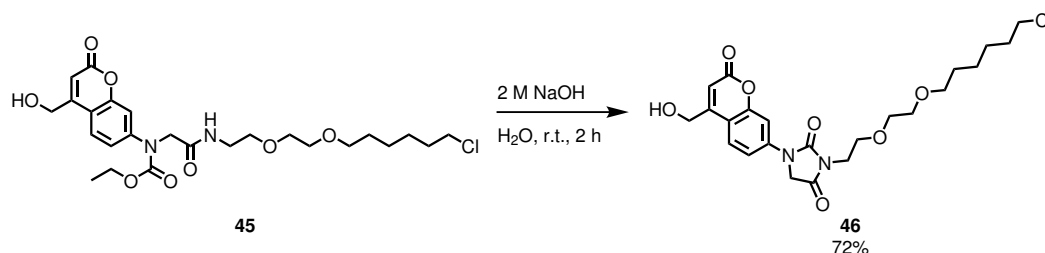


Compound **44** (1 g, 3.1 mmol, 1 eq) was dissolved in DMF (5 mL). DCC (642 mg, 3.1 mmol, 1 eq) and NHS (357 mg, 3.1 mmol, 1 eq) were added and the mixture was stirred for 3 h. **Halo-NH<sub>2</sub>** (803 mg, 3.1 mmol, 1 eq) and NEt<sub>3</sub> (860  $\mu\text{L}$ , 6.2 mmol, 2 eq) were added and the mixture was stirred for 16 h at rt. The reaction was quenched with a saturated aqueous solution of NH<sub>4</sub>Cl. The product was extracted three times with EtOAc, the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The product was purified with column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 20:1) to give compound **45** (670 mg, %).

**$^1\text{H}$ -NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$ : 7.34 (d,  $J$  = 8.50 Hz, 1H), 7.28-7.27 (m, 2H), 6.67 (t,  $J$  = 5.07 Hz, 1H), 6.51 (t,  $J$  = 1.27 Hz, 1H), 4.77 (d,  $J$  = 1.27 Hz, 2H), 4.31 (s, 2H), 4.25 (q,  $J$  = 6.73 Hz, 2H), 3.74 (t,  $J$  = 5.42 Hz, 1H), 3.64-3.44 (m, 12H), 1.79-1.72 (m, 2H), 1.63-1.56 (m, 2H), 1.46-1.33 (m, 4H), 1.28 (t,  $J$  = 6.73 Hz, 3H).

**$^{13}\text{C}$ -NMR** (100.6 MHz, CDCl<sub>3</sub>)  $\delta$ : 168.6, 161.0, 155.1, 154.3, 153.7, 145.4, 123.6, 121.7, 115.3, 113.7, 111.3, 71.4, 70.4, 70.1, 69.7, 63.0, 60.4, 53.6, 45.2, 39.6, 32.6, 31.1, 29.5, 26.8, 25.5, 14.6.

**Preparation of N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-((4-(hydroxymethyl)-2-oxo-2H-chromen-7-yl)amino)acetamide (46)**



Compound **45** (400 mg, 0.76 mmol, 1 eq) was dissolved in a 2 M aqueous NaOH solution (5 mL) and stirred at room temperature for 16 h. The reaction was acidified to pH 4, with 1 M aqueous HCl solution. The product was extracted three times with EtOAc, the combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent evaporated under reduced pressure. The crude product was purified by column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 10:1) to afford **46** (263 mg, 72%).

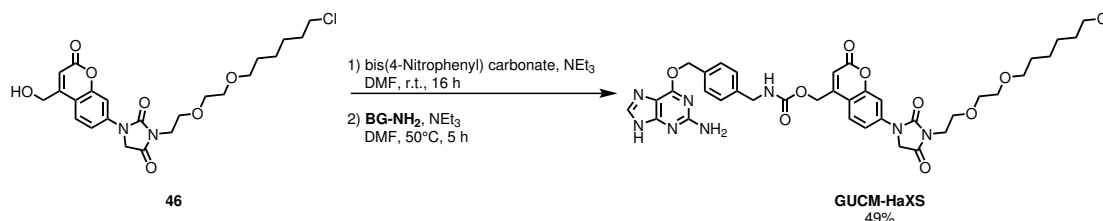
**$^1\text{H}$ -NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$ : 7.55 (m, 2H), 7.49 (d,  $J$  = 9.38 Hz, 1H), 6.53 (t,  $J$  = 1.58 Hz, 1H), 4.87 (d,  $J$  = 1.58 Hz, 2H), 4.33 (s, 2H), 3.86-3.83 (m, 2H), 3.80-3.77 (m, 2H), 3.68-3.64 (m,

## Experimental Section

2H), 3.58-3.55 (m, 2H), 3.51 (t,  $J = 6.68$  Hz, 2H), 3.43 (t,  $J = 6.68$  Hz, 2H), 1.77-1.70 (m, 2H), 1.59-1.52 (m, 2H), 1.44-1.29 (m, 4H).

$^{13}\text{C-NMR}$  (100.6 MHz,  $\text{CDCl}_3$ )  $\delta$ : 167.6, 160.8, 154.4, 153.6, 140.6, 124.5, 113.6, 113.4, 110.8, 105.6, 71.3, 70.1, 69.9, 67.2, 60.6, 49.5, 45.1, 38.4, 32.5, 29.5, 26.7, 25.4.

### Preparation of (7-((2-((2-((6-chlorohexyl)oxy)ethoxy)ethyl)amino)-2-oxoethyl)amino)-2-oxo-2H-chromen-4-yl)methyl 4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzylcarbamate (GUCM-HaXS)



Compound **46** (200 mg, 0.42 mmol, 1 eq) was added slowly to a solution of bis(4-nitrophenyl) carbonate (128 mg, 0.42 mmol, 1 eq) in  $\text{CH}_2\text{Cl}_2$  (10 mL) at  $0^\circ\text{C}$ .  $\text{NEt}_3$  (58  $\mu\text{L}$ , 0.42 mmol, 1 eq) was added and the mixture was stirred for 3 h at rt. The reaction was quenched with a saturated aqueous solution of  $\text{NH}_4\text{Cl}$ . The product was extracted three times with EtOAc the combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was evaporated under reduced pressure. Filtration over a short pad of silica gave the carbonate of **11**, which was directly used for the carbamate formation. Therefore, the carbonate was dissolved in DMF (2 mL), **BG-NH<sub>2</sub>** (110 mg, 0.42, 1 eq) and  $\text{NEt}_3$  (58  $\mu\text{L}$ , 0.42 mmol, 1 eq) were added and the solution was stirred for 3 h at  $60^\circ\text{C}$ . The reaction was quenched with water and the product was extracted with three times EtOAc. The combined organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was evaporated under reduced pressure. The crude product was purified by column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ , 10:1) to give compound **GUCM-HaXS** (160 mg, 49%).

$^1\text{H-NMR}$  (400 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 12.40 (s, 1H), 8.13 (t,  $J = 5.85$  Hz, 1H), 7.80 (s, 1H), 7.75 (d,  $J = 8.38$  Hz, 1H), 7.70-7.67 (m, 2H), 7.46 (d,  $J = 8.38$  Hz, 2H), 7.32 (d,  $J = 8.38$  Hz, 2H), 6.34 (t,  $J = 1.55$  Hz, 1H), 6.28 (s, 2H), 5.46 (s, 2H), 5.34 (s, 2H), 4.58 (s, 2H), 4.27 (d,  $J = 6.07$  Hz, 2H), 3.64-3.33 (m, 12H), 1.68-1.61 (m, 2H), 1.47-1.40 (m, 2H), 1.36-1.22 (m, 4H).

$^{13}\text{C-NMR}$  (100.6 MHz,  $\text{DMSO}-d_6$ )  $\delta$ : 168.6, 159.8, 159.8, 159.6, 155.6, 155.2, 154.2, 153.8, 151.4, 141.2, 139.2, 137.8, 135.5, 128.6, 127.1, 125.4, 114.0, 113.5, 112.2, 105.0, 70.2, 69.5, 69.4, 66.5, 66.4, 61.0, 49.8, 45.3, 43.8, 38.0, 32.0, 29.1, 26.1, 24.9.

Photophysical properties:  $\lambda_{\text{max}}$ : 330 nm,  $\epsilon_{360}$ :  $2364 \text{ M}^{-1}\text{cm}^{-1}$ ,  $\epsilon_{405}$ :  $1313 \text{ M}^{-1}\text{cm}^{-1}$ ,  $\phi_{360}$ : 0.007,  $\phi_{405}$ : 0.009.

### 5.3.2 NMR experiments to determine the mechanism

#### 5.3.2.1 Reference NMR spectra

Under a nitrogen atmosphere and in a standard NMR tube, **rapamycin** (2.2 mg, 2.4  $\mu\text{mol}$ ), phenyl carbamate (1.97 mg, 14.4  $\mu\text{mol}$ ), or  $\text{BF}_3\text{-Et}_2\text{O}$  (1.1  $\mu\text{L}$ , 7.2  $\mu\text{mol}$ ) were dissolved in 0.4 mL  $\text{CD}_2\text{Cl}_2$ . The tubes were transferred to the NMR spectrometer. For solutions containing **rapamycin**, or phenyl carbamate the  $^1\text{H}$ -,  $^{13}\text{C}$ -,  $^1\text{H}$ - $^1\text{H}$  COSY-,  $^1\text{H}$ - $^1\text{H}$  NOESY-,  $^1\text{H}$ - $^1\text{H}$  ROESY-,  $^1\text{H}$ - $^1\text{H}$  TOSCY-,  $^1\text{H}$ - $^{13}\text{C}$  HMQC-, and  $^1\text{H}$ - $^{13}\text{C}$  HMBC-NMR spectra were recorded at  $-30^\circ\text{C}$ . For the  $\text{BF}_3\text{-Et}_2\text{O}$  containing solution the  $^1\text{H}$ -,  $^{11}\text{B}$ -, and  $^{19}\text{F}$ -NMR spectra were recorded. After these measurements, THF (0.2 mL) was added to each tube and the same NMR experiments as described above were repeated.

#### 5.3.2.1 NMR spectra of rapamycin and $\text{BF}_3\text{-Et}_2\text{O}$ in $\text{CD}_2\text{Cl}_2$

Under a nitrogen atmosphere and in a standard NMR tube, **rapamycin** (2.2 mg, 2.4  $\mu\text{mol}$ , 1 eq) was dissolved in 0.4 mL  $\text{CD}_2\text{Cl}_2$ . The solution was cooled to  $-30^\circ\text{C}$  and  $\text{BF}_3\text{-Et}_2\text{O}$  (1.1  $\mu\text{L}$ , 7.2  $\mu\text{mol}$ , 4 eq) was added. The NMR tube was briefly shaken, quickly transferred to the NMR spectrometer, and the  $^1\text{H}$ -,  $^{13}\text{C}$ -,  $^{11}\text{B}$ -,  $^{19}\text{F}$ -,  $^1\text{H}$ - $^1\text{H}$  COSY-,  $^1\text{H}$ - $^1\text{H}$  NOESY-,  $^1\text{H}$ - $^1\text{H}$  ROESY-,  $^1\text{H}$ - $^1\text{H}$  TOSCY-,  $^1\text{H}$ - $^{13}\text{C}$  HMQC-, and  $^1\text{H}$ - $^{13}\text{C}$  HMBC-NMR spectra were recorded at  $-30^\circ\text{C}$ .

#### 5.3.2.1 NMR spectra of rapamycin and $\text{BF}_3\text{-Et}_2\text{O}$ in a mixture of $\text{CD}_2\text{Cl}_2$ and THF

Under a nitrogen atmosphere and in a standard NMR tube, **rapamycin** (2.2 mg, 2.4  $\mu\text{mol}$ , 1 eq) was dissolved in 0.4 mL  $\text{CD}_2\text{Cl}_2$ . The solution was cooled to  $-30^\circ\text{C}$  and  $\text{BF}_3\text{-Et}_2\text{O}$  (1.1  $\mu\text{L}$ , 7.2  $\mu\text{mol}$ , 4 eq) was added. The NMR tube was shaken for 3 minutes prior to the addition of THF (0.2 mL). The mixture was quickly transferred to the NMR spectrometer and the  $^1\text{H}$ -,  $^{13}\text{C}$ -,  $^{11}\text{B}$ -,  $^{19}\text{F}$ -,  $^1\text{H}$ - $^1\text{H}$  COSY-,  $^1\text{H}$ - $^1\text{H}$  NOESY-,  $^1\text{H}$ - $^1\text{H}$  ROESY-,  $^1\text{H}$ - $^1\text{H}$  TOSCY-,  $^1\text{H}$ - $^{13}\text{C}$  HMQC-, and  $^1\text{H}$ - $^{13}\text{C}$  HMBC-NMR spectra were recorded at  $-30^\circ\text{C}$ .

#### 5.3.2.3 NMR spectra of phenyl carbamate and $\text{BF}_3\text{-Et}_2\text{O}$ in $\text{CD}_2\text{Cl}_2$

Under a nitrogen atmosphere and in a standard NMR tube, phenyl carbamate (1.97 mg, 14.4  $\mu\text{mol}$ , 1.5 eq) was dissolved in 0.4 mL  $\text{CD}_2\text{Cl}_2$ . The solution was cooled to  $-30^\circ\text{C}$  and  $\text{BF}_3\text{-Et}_2\text{O}$  (1.1  $\mu\text{L}$ , 7.2  $\mu\text{mol}$ , 1 eq) was added. The NMR tube was briefly shaken, quickly transferred to the NMR spectrometer, and the  $^1\text{H}$ -,  $^{13}\text{C}$ -,  $^{11}\text{B}$ -,  $^{19}\text{F}$ -,  $^1\text{H}$ - $^1\text{H}$  COSY-,  $^1\text{H}$ - $^1\text{H}$  NOESY-,  $^1\text{H}$ - $^1\text{H}$  ROESY-,  $^1\text{H}$ - $^1\text{H}$  TOSCY-,  $^1\text{H}$ - $^{13}\text{C}$  HMQC-, and  $^1\text{H}$ - $^{13}\text{C}$  HMBC-NMR spectra were recorded at  $-30^\circ\text{C}$ .

#### 5.3.2.4 NMR spectra of phenyl carbamate and $\text{BF}_3\text{-Et}_2\text{O}$ in a mixture of $\text{CD}_2\text{Cl}_2$ and THF

Under a nitrogen atmosphere and in a standard NMR tube, phenyl carbamate (1.97 mg, 14.4  $\mu\text{mol}$ , 1.5 eq) was dissolved in 0.4 mL  $\text{CD}_2\text{Cl}_2$ . The solution was cooled to  $-30^\circ\text{C}$  and  $\text{BF}_3\text{-Et}_2\text{O}$  (1.1  $\mu\text{L}$ , 7.2  $\mu\text{mol}$ , 1 eq) was added. The NMR tube was shaken for 3 minutes prior to the addition of THF (0.2 mL). The mixture was quickly transferred to the NMR spectrometer and the  $^1\text{H}$ -,  $^{13}\text{C}$ -,  $^{11}\text{B}$ -,  $^{19}\text{F}$ -,  $^1\text{H}$ - $^1\text{H}$  COSY-,  $^1\text{H}$ - $^1\text{H}$  NOESY-,  $^1\text{H}$ - $^1\text{H}$  ROESY-,  $^1\text{H}$ - $^1\text{H}$  TOSCY-,  $^1\text{H}$ - $^{13}\text{C}$  HMQC-, and  $^1\text{H}$ - $^{13}\text{C}$  HMBC-NMR spectra were recorded at  $-30^\circ\text{C}$ .

#### 5.3.2.5 NMR spectra of rapamycin, phenyl carbamate and $\text{BF}_3\text{-Et}_2\text{O}$ in $\text{CD}_2\text{Cl}_2$

Under a nitrogen atmosphere and in a standard NMR tube, **rapamycin** (2.2 mg, 2.4  $\mu\text{mol}$ , 1 eq), and phenyl carbamate (1.97 mg, 14.4  $\mu\text{mol}$ , 6 eq) were dissolved in 0.4 mL  $\text{CD}_2\text{Cl}_2$ . The solution was cooled to  $-30^\circ\text{C}$  and  $\text{BF}_3\text{-Et}_2\text{O}$  (1.1  $\mu\text{L}$ , 7.2  $\mu\text{mol}$ , 4 eq) was added. The NMR tube

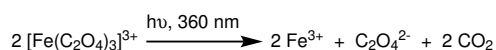
was briefly shaken, quickly transferred to the NMR spectrometer, and the  $^1\text{H}$ -,  $^{13}\text{C}$ -,  $^{11}\text{B}$ -, and  $^{19}\text{F}$ -NMR spectra were recorded at  $-30^\circ\text{C}$ .

### 5.3.2.6 NMR spectra of rapamycin, phenyl carbamate and $\text{BF}_3\text{-Et}_2\text{O}$ in a mixture of $\text{CD}_2\text{Cl}_2$ and THF

Under a nitrogen atmosphere and in a standard NMR tube, **rapamycin** (2.2 mg, 2.4  $\mu\text{mol}$ , 1 eq), and phenyl carbamate (1.97 mg, 14.4  $\mu\text{mol}$ , 6 eq) were dissolved in either 0.4 mL  $\text{CD}_2\text{Cl}_2$ . The solution was cooled to  $-30^\circ\text{C}$  and  $\text{BF}_3\text{-Et}_2\text{O}$  (1.1  $\mu\text{L}$ , 7.2  $\mu\text{mol}$ , 4 eq) were added. The NMR tube was shaken for 3 minutes prior to the addition of THF (0.2 mL). The mixture was quickly transferred to the NMR spectrometer and the  $^1\text{H}$ -,  $^{13}\text{C}$ -,  $^{11}\text{B}$ -, and  $^{19}\text{F}$ -NMR spectra were recorded at  $-30^\circ\text{C}$ .

### 5.3.3 Determination of photophysical properties

#### 5.3.3.1 Ferrioxalate actinometry to determine the quantum flow of the Lumos 43A photo reactor



Ferrioxalate actinometry was used to determine the quantum flow of the Lumos 43A photoreactor (LED 320 nm, 360 nm or 405 nm lamp). To this end, potassium ferrioxalate (147.5 mg, 0.34 mmol) was dissolved in 40 mL water. 40 mL of 1 N sulfuric acid were added and the solution was further diluted to 50 mL with water. 3 mL of this solution was pipetted into the absorption cell and irradiated for a given time. 2 mL of the irradiated solution was mixed with 2 mL of a 5.6 mM aqueous o-phenanthroline solution and 1 mL of a sodium acetate buffer. As reference 2 mL of non-irradiated ferrioxalate solution was treated with 2 mL of a 5.6 mM aqueous o-phenanthroline solution and 1 mL of a sodium acetate buffer. The 2 solutions were stored for 1 hour in absolute darkness prior to measure the absorbance of the solutions at 510 nm. The quantum flow was calculated using equation (1) and with a quantum yield of 1.26 at 360 nm and the molar extinction coefficient of the complex at 510 nm ( $1.11 \cdot 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ ).

$$\text{Quantum Flow [Einstein/s]} = \frac{A \cdot V1 \cdot V2}{t \cdot \phi \cdot I \cdot V2} \quad (1)$$

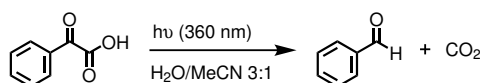
$A$  = absorbance of the irradiated solution at 510 nm  $V1$  = irradiated volume (3 mL)

$t$  = irradiation time in seconds  $V2$  = used volume of  $V1$  (2 mL)

$\phi$  = quantum yield (1.26 at 360 nm)  $V3$  = end volume (20 mL)

$\epsilon$  = extinction coefficient of the complex at 510 nm  $l$  = thickness of the cell (1 cm)

### 5.3.3.2 Phenylglyoxylic acid actinometry to determine the quantum flow of the Lumos 43A photo reactor



3 mL of a 0.05 M solution of phenylglyoxylic acid in a mixture of water and acetonitrile (3:1) were put in a standard absorbance quartz cell (1 cm path length) and irradiated in a LUMOS 43A photoreactor equipped with LED lamps at 360 nm for a given time. A UV/Vis absorbance spectrum was recorded every 5 minutes of irradiation. The decrease of the optical density at 390 nm ( $\text{OD} < 1$ ) was monitored and used to determine the number of moles photolyzed after each irradiation. As the absorbance of the product, benzaldehyde, is negligible at 390 nm, the conversion can be determined by dividing the measured optical density by the optical density of the starting solution, and multiplying with the starting amount (in mol). The number of converted mol was plotted as a function of the time. The slope of the regression line was divided by the published quantum yield of photolysis ( $\phi_{365} = 0.728$ ) to give a value of  $5.445.44 \cdot 10^{-7}$  E/min for the photon flux of the 360 nm LED lamp of the LUMOS 43A.

### 5.3.3.3 General procedure to determine the photolytic quantum yield

3 mL of a 0.5 mM of photocleavable HaXS molecules in a DMSO/water (1:10) solution were transferred into a standard absorbance cell and irradiated in the Lumos 43A photoreactor with light at specific wavelength (360 nm or 405 nm) for different time. The conversion was determined by UPLC-MS by integration of the absorption peak of the molecule. Disappearance of the molecule was plotted as a function of the time. The slope of the regression line corresponds to the rate constant ( $k$ ). The reaction rate was calculated using equation (2).

$$r = k[A] \quad (2)$$

With the measured reaction rates and the quantum flow of Lumos 43A photoreactor at specific wavelength, the photolysis quantum yield of photocleavable HaXS molecules was calculated with equation (3).

$$\text{Quantum Yield} = (\text{reaction rate})/(\text{quantum flow}) \quad (3)$$

### 5.3.3.4 UV/Vis spectra and determination of the molar absorption coefficient

The absorption spectra of different concentrated solutions of the photocleavable HaXS molecules in DMSO were recorded on a Perkin Elmer Lambda 40 UV/Vis spectrometer or a NanoDrop 2000/2000c spectrophotometer. The molar extinction coefficient at specific wavelength ( $\epsilon_{nm}$ ) were calculated according to the Beer-Lambert law (equation (4)) or determined graphically, plotting the absorption of the molecules at specific wavelength against the concentration. The slope of the resulting straight line corresponds to the molar extinction coefficient at the specific wavelength.

$$A = \epsilon \cdot l \cdot c \quad (4)$$



### 5.3.4 Biological procedures

#### 5.3.4.1 Cloning of expression constructs

The <sup>T2098L</sup>FRB coding sequence (kind gift from T. Inoue, Baltimore), FKBP12 coding sequence (kind gift from K. Johnsson, Lausanne), <sup>L273Y</sup>HaloTag7 coding sequence (Promega) SNAPf coding sequence (kind gift from K. Johnsson, Lausanne) mTFP1 (kind gift of O. Pertz, Basel), Golgi targeting sequence Giantin (kind gift from T. Inoue, Baltimore), LAMP1 (kind gift from T. Inoue, Baltimore), plasma membrane targeting sequence CAAX (kind gift of J. Downward, London) and mTq (kind gift of J. Goedhart Amsterdam) were amplified by PCR (Phusion polymerase, Finnzymes).

For <sup>T2098L</sup>FRB-RFP, FKBP12-GFP, <sup>L273Y</sup>HaloTag7-GFP, <sup>L273Y</sup>HaloTag7-RFP, and SNAPf-GFP expression constructs, <sup>T2098L</sup>FRB, FKBP12, <sup>L273Y</sup>HaloTag7 and SNAPf were transferred to pEGFP (Clontech), or pTag-RFP-N1 (Evrogen) expression vectors.

FKBP12-GFP-CAAX: the membran targeting sequence was introduced into the FKBP12-GFP expression vector.

SNAPf-mTFP1: GFP was exchanged in a SNAPf-GFP plasmid by the mTFP1 sequence.

<sup>L273Y</sup>HaloTag7-RFP-Giantin: the Giantin targeting sequence was introduced into the <sup>L273Y</sup>HaloTag7-RFP expression vector.

NLS-CFP-SNAP: FKBP1x from the NLS-CFP-FKBP1x expression vector (kind gift from S. Hbner, Wrzburg) was exchanged by the SNAPf sequence.

<sup>L273Y</sup>HaloTag7-RFP-Rheb15: <sup>L273Y</sup>HaloTag7-RFP was amplified by PCR and fused to a Rheb15 sequence to be inserted into the pTag-RFP-N1 backbone with RFP was excised.

Halo-RFP-Rheb15: HA-Raptor was replaced by <sup>L273Y</sup>HaloTag7-RFP in a HA-Raptor-Rheb15 expression vector kindly obtained from Anna Melone.

LAMP-RFP-<sup>L273Y</sup>HaloTag7: the LAMP1 targeting sequence was inserted into the RFP-<sup>L273Y</sup>HaloTag7 expression vector.

Mito-SYFP-SNAPf: FRB was exchanged into a Mito-SYFP-FRB expression construct (kind gift of P. Scheiffele, Basel) by the SNAPf sequence.

LifeAct-mTFP1-SNAPf: mTFP1 was fused to LifeAct by PCR, and inserted into the GFP-SNAPf expression construct (where GFP was excised).

<sup>L273Y</sup>HaloTag7-mTq: GFP was exchanged by the mTq sequence. Maps and expression vector sequences can be obtained from the authors upon request.

#### 5.3.4.2 Cell Culture and Transfection

HeLa, HEK293, and A2058 cells were cultured in complete Dulbeccos's modified Eagle medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum (HIFCS), 2 mM L-glutamine (Gln), 100 units/ml penicillin, 100 µg/ml streptomycin (PEST), at 37°C, 5% CO<sub>2</sub>. Transfections were carried out with JetPEI<sup>TM</sup> (Polyplus-transfection), according to manufacturer's guidelines.

#### 5.3.4.3 Protein denaturation, cell lysis, and immune-blotting

Cells were washed with ice cold PBS and lysed in a NP-40 lysis buffer [1% NP-40, 20 mM Tris-HCl pH 8.0, 138 mM NaCl, 2.7 mM KCl, 5% glycerol, 40 mM NaF, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 20 µM Leupeptin, 18 µM Pepstatin, 5 µM Aprotinin, 1 mM PMSF, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 5 mM EDTA]. Cell lysates were cleared by centrifugation at 13000 rpm for 15 min and proteins were denatured by the addition of 5x sample buffer [312.5 mM Tris-HCl (pH 6.8), 10% SDS, 25% β-mercaptoethanol, 50% glycerol, bromophenol blue] and cooking for 6 min. Proteins were separated by SDS-PAGE and transferred to Immobilon PVDF membranes (Millipore). Proteins labeled with primary and secondary antibodies were visualised enhanced chemiluminescence (Millipore) and a CCD camera (Fusion Fx7, Vilber).

For rapalogs: the phosphorylation of p70<sup>S6K</sup> was analysed using mouse monoclonal antibody (mAb) to pThr389-S6K, rabbit mAb to S6K1, mouse mAb to pSer473-PKB/Akt and to pThr308-PKB/Akt (all from Cell Signaling Technology, Danvers) and horseradish peroxidase (HRP)-coupled secondary goat anti-mouse or anti-rabbit antibodies.

For HaXS molecules: GFP-fusion proteins were analysed, using mAb to GFP (Roche Diagnostics) and horseradish peroxidase (HRP-conjugated goat anti-mouse IgG) as secondary antibody.

**5.3.4.4 *Statistic analysis*** Statistical Analysis Statistical analysis was performed with GraphPad Prism v6. For Student's test (two sided, non-paired with Welch correction,  $p < 0.05$ )  $\geq 3$  independent experiments were compared.

**5.3.4.5 *Determination of the effect of pcRap on mTORC1 pathway***

HEK293 cells were exposed to DMSO, rapamycin, (*S*)-pcRap and ((*R*)-pcRap for 15 min at 37°C, before cells were lysed, and proteins were separated by SDS-PAGE and analysed by immune-blotting.

**5.3.4.6 *In cell Western to determine the IC<sub>50</sub> value of (R)-pcRap***

A2058 cells were exposed to rapamycin and (*R*)-pcRap for 1 h at 37°C, before cells were lysed, and proteins were separated by SDS-Page. The phosphorylation of p70<sup>S6K</sup> was analysed by in-cell Western using goat anti-rabbit IRDye800 and goat anti-mouse IRDye680 secondary antibodies.

**5.3.4.7 *(R)-pcRap-induced membrane translocation***

For microscopy, HeLa cells grown on 12 mm cover slips (Menzel) were transfected with expression constructs for FKBP12-GFP-CAAX and mutated T2098L FRB-RFP fusion proteins. Cells were exposed to DMSO or (*R*)-pcRap for 15 min at 37°C in fully supplemented, complete DMEM medium 24 h after transfection, were then washed twice with PBS, fixed with 4% *p*-formaldehyde (PFA) in PBS, and mounted in Mowiol (Plss-Stauffer) containing 1% Propyl gallate.

**5.3.4.8 *In vivo heterodimerization of SNAP-GFP and Halo-GFP with HaXS molecules***

HeLa cells were grown in 6-well cell culture plates (Falcon), and were transfected with expression constructs for SNAP-GFP and Halo-GFP. After 24 h, cells were exposed to HaXS dimerisers for 5, 15, 30, 60 or 300 minutes at 37°C. Cells were washed twice in 1 x PBS after treatment with HaXS molecules. Cells were lysed, proteins were separated by SDS-PAGE, and analysed by immune-blotting as described above.

**5.3.4.9 *In vivo photocleavage of SNAP-GFP-Halo-GFP dimers***

HeLa cells were grown in 6-well cell culture plates (Falcon), and were transfected with expression constructs for SNAP-GFP and Halo-GFP. After 24 h, cells were exposed to HaXS dimerisers for 15 minutes at 37°C. Cells were washed twice in 1 x PBS after treatment with HaXS molecules. Then 600  $\mu$ l 1x PBS was added to each well. Culture plates were put on ice and irradiated for 2 or 10 minutes with a high-intensity UV lamp (BlakRay B-100A high intensity UV lamp; 100 Watt, 365 nm, UVP) at a distance of 5 cm. Cells were lysed, proteins were separated by SDS-PAGE, and analysed by immune-blotting as described above.

### ***5.3.4.10 SNAP-mTFP1 fusion protein translocation to the the Golgi and subsequent release***

HeLa cells grown on 12 mm cover slips (Menzel) were transfected with SNAP-mTFP1 and  $L^{273Y}$ HaloTag7-RFP-Giantin. After 24 h, cells were treated with HaXS dimerisers for 15 minutes. Cells were mounted in Ludin chambers (Live Imaging Services) in the closed confirmation with complete cell culture medium without phenol red. Pictures were taken every second before illumination of the cells and every 1 or 3 second after illumination. Single spots within the indicated region of interest were illuminated for 40 ms with a FRAP scanning laser or for 20 seconds with a mercury halide lamp using a standard DAPI filter. Movies were assembled and analysed with Fiji (ImageJA, 1.44b). mTFP1 fluorescence at vesicles was calculated according to the formula  $mTFP1 = F(mTFP1, vesicle) / F(RFP, whole cell) - F(mTFP1, cytoplasm) / F(RFP, whole cell)$  for every frame. mTFP1 fluorescence in the cytoplasm was calculated according to the formula  $mTFP1 = F(mTFP1, vesicle) / F(RFP, whole cell)$ . To illustrate the fluorescence decrease from vesicles and to measure fluorescence increase in the cytoplasm, mTFP1 fluorescence intensity was plotted over time.

### ***5.3.4.11 NLS-CFP-SNAP fusion protein translocation to the Golgi, subsequent release and relocalitation to the nucleus***

HeLa cells grown on 12 mm cover slips (Menzel) were transfected with NLS-CFP-SNAP and  $L^{273Y}$ HaloTag7-RFP-Giantin. After 24 h, cells were treated with HaXS dimerisers for 15 minutes. Cells were mounted in Ludin chambers (Live Imaging Services) in the closed confirmation with complete cell culture medium without phenol red. Pictures were taken every second before illumination of the cells and every 1 or 3 second after illumination. Single spots within the indicated region of interest were illuminated for 40 ms with a FRAP scanning laser. Movies were assembled and analysed with Fiji (ImageJA, 1.44b). CFP fluorescence at vesicles was calculated according to the formula  $CFP = F(CFP, vesicle) / F(RFP, whole cell) - F(CFP, cytoplasm) / F(RFP, whole cell)$  for every frame. Nuclear CFP fluorescence intensity was calculated according to the formula  $CFP = F(CFP, nucleus) / F(RFP, whole cell)$  for every image. To illustrate the fluorescence decrease from vesicle and to measure fluorescence increase in the nucleus, CFP fluorescence intensity was plotted over time.

## 6 Appendix

### 6.1 Contributions

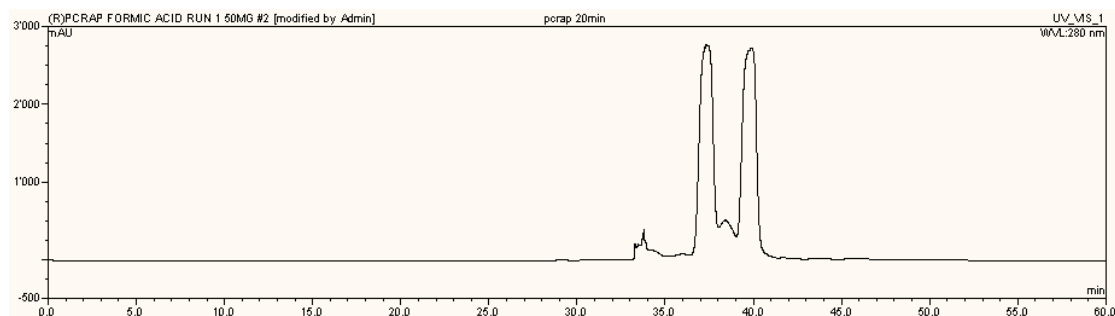
All molecules in this thesis except of 7-((18-chloro-3,6,9,12-tetraoxaoctadecyl)oxy)-4-(hydroxymethyl)-2H-chromen-2-one (**23**) and **HCM-HaXS**, which were synthesised by Dr. Florent Beaufils, were synthesised and characterised by Ruben Cal. The synthesis of MeNV-HaXS was developed by Viktor Hoffmann, in the course of his Master thesis and optimised by Ruben Cal.

NMR spectra to determine the mechanism of the Lewis acid-mediated C16 substitution of rapamycin and NMR spectra to determine the absolute configuration of rapalogs were recorded by Kaspar Zimmermann in the group of PD. Dr. Daniel Hussinger at the University of Basel. The spectra were subsequently analysed by Kaspar Zimmermann, PD Dr. Daniel Hussinger and Ruben Cal.

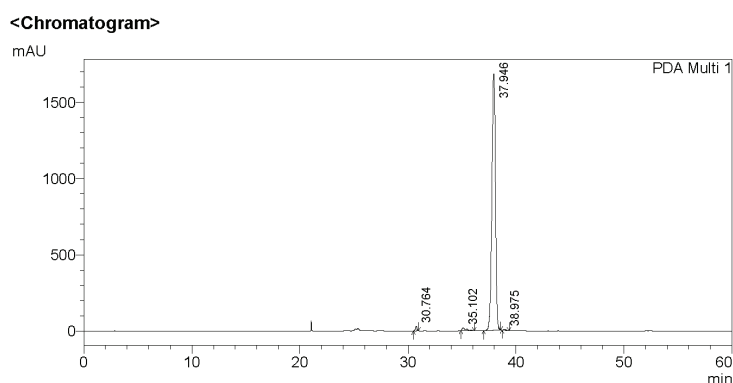
All in vitro photolysis experiments were performed and analysed by Elia Janett in the group of Prof. Dr. Christian Bochet at the University of Fribourg. HRMS and ESI-MS spectra were recorded by the mass spectroscopy service at the University of Bern.

All biological experiments except for the determination of the IC<sub>50</sub> value of rapamycin, which were performed by Anna Melone, were performed by Mirjam Zimmermann.

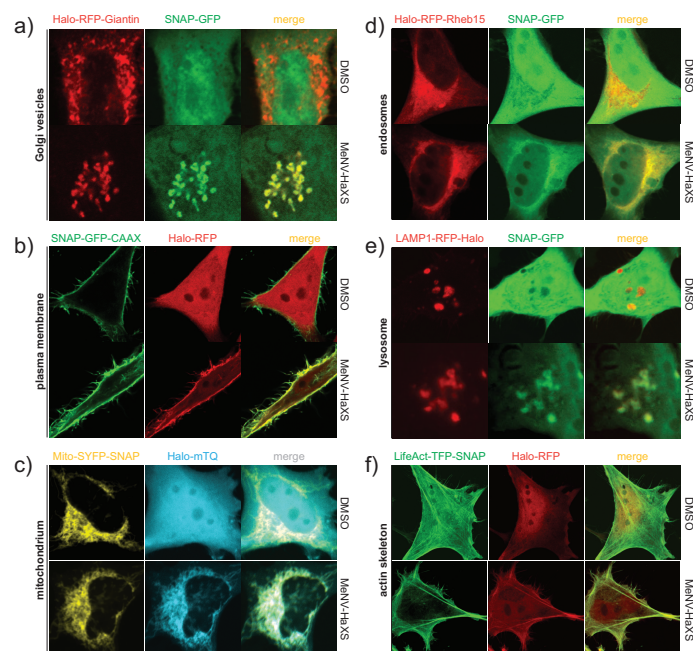
### 6.2 Supplementary figures



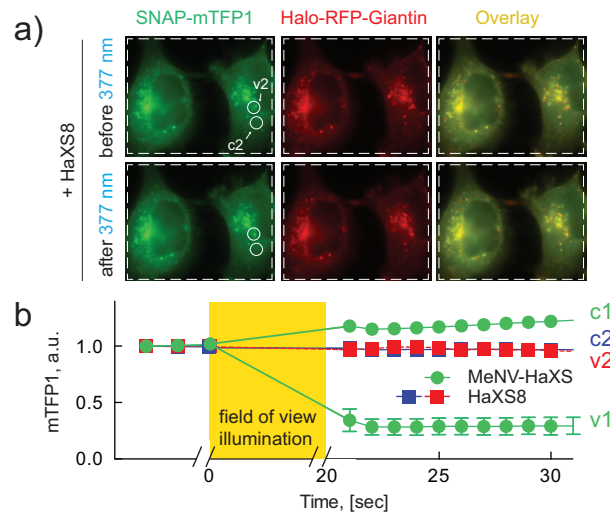
**Figure S 1:** The UV/Vis spectra of the crude mixture of **pcRap** separated by preparative HPLC. 10 mg of the crude **pcRap** were dissolved in 2 mL CH<sub>2</sub>Cl<sub>2</sub> and injected into a Dionex P680 HPLC instrument with a Reprosil 100 Si, 5 µm column (250 x 40 mm). Sample was eluted using a linear gradient of 100:0 to 94:6 (CH<sub>2</sub>Cl<sub>2</sub>:methanol) and a flow rate of 30 mL·min<sup>-1</sup> over 90 minutes



**Figure S 2:** The UV/Vis spectra of pure (*R*)**pcRap** separated by analytical HPLC. 50  $\mu$ L of a 100  $\mu$ M solution of product in  $\text{CH}_2\text{Cl}_2$  were injected into Shimadzu LC-20AT instrument with a Reprospher 100 Si, 5  $\mu$ m column (100  $\times$  10 mm) and eluted using a linear gradient of 100:0 to 96:4 ( $\text{CH}_2\text{Cl}_2$ :methanol) and a flow rate of 2  $\text{mL}\cdot\text{min}^{-1}$  over 60 minutes



**Figure S 3:** Cytosolic protein translocation to different cellular compartments: a) Giantin was used to target the Halo-RFP-Giantin to the Golgi, b) the CAAX box of K-Ras targeted SNAP-GFP-CAAX to the plasma membrane, c) Mito-SYFP-SNAP localized on mitochondria, d) the Rheb15-tagged Halo-RFP-Rheb15 was localized on early and late endosomes, e) LAMP1 was on lysosomes, and f) LifeAct served as an anchor on the F-actin cytoskeleton. HeLa cells expressing the indicated organelle anchors and the indicated cytosolic cargo proteins were grown on 12 mm coverslips (Menzel), before they were incubated with **MeNV-HAXS** (37°C, 15 min), washed twice with PBS, and fixed with 4% *p*-formaldehyde (PFA, in PBS), and mounted in Mowiol (Plss-Stauffer) containing 1% propyl gallate (Sigma-Aldrich). Translocation of cytosolic Halo-RFP, Halo-mTq or Halo-RFP fusion proteins (cargo) to the respective anchors is documented.



**Figure S 4:** Release of Golgi-trapped SNAP-mTFP1 protein after UV illumination. a) At  $t = >15$  min, HeLa cells expressing SNAP-mTFP1 and Halo-RFP-Giantin were exposed to  $5 \mu\text{M}$  **HaXS8** ( $5 \mu\text{M}$  **MeNV-HaXS** see figure 22, main part), which caused the translocation of cytosolic SNAP-mTFP1 to the Golgi (labeled as "before 377 nm"). SNAP-mTFP1 intensity was monitored within the indicated circular regions of interest by live cell microscopy. At  $t = 0$  illumination with UV light using a standard DAPI filter set on a conventional fluorescence microscope ( $t = 20$  sec,  $377 \pm 25$  nm) was initiated for 20 s, and cells are shown after illumination (after 377 nm). b) Quantification of mTFP1 fluorescence intensity of selected regions of interest (circles) at Golgi-derived vesicles (v) and in cytoplasm (c) before and after illumination as described in a) are shown (values represent means  $\pm$  SEM,  $n = 10$ , error bars removed where smaller than symbols used). Curves obtained with **MeNV-HaXS** (see figure 22) are shown in green.

## 7 References

- (1) Klemm, J. D.; Schreiber, S. L.; Crabtree, G. R. *Annual Review of Immunology* **1998**, *16*, 569–592.
- (2) Chen, Y.-X.; Triola, G.; Waldmann, H. *Accounts of Chemical Research* **2011**, *44*, 762–773.
- (3) Rutkowska, A.; Schultz, C. en *Angewandte Chemie International Edition* **2012**, *51*, 8166–8176.
- (4) Bishop, A.; Buzko, O.; Heyeck-Dumas, S.; Jung, I.; Kraybill, B.; Liu, Y.; Shah, K.; Ulrich, S.; Witucki, L.; Yang, F.; Zhang, C.; Shokat, K. M. *Annual Review of Biophysics and Biomolecular Structure* **2000**, *29*, 577–606.
- (5) Banaszynski, L. A.; Chen, L.-c.; Maynard-Smith, L. A.; Ooi, A. G. L.; Wandless, T. J. *Cell* **2006**, *126*, 995–1004.
- (6) Sletten, E.; Bertozzi, C. en *Angewandte Chemie International Edition* **2009**, *48*, 6974–6998.
- (7) Foley, T. L.; Burkart, M. D. *Current Opinion in Chemical Biology* **2007**, *11*, 12–19.
- (8) Baker, K.; Blecinski, C.; Lin, H.; Salazar-Jimenez, G.; Sengupta, D.; Krane, S.; Cornish, V. W. en *Proceedings of the National Academy of Sciences* **2002**, *99*, 16537–16542.
- (9) Walsh, D. P.; Chang, Y.-T. *Chemical Reviews* **2006**, *106*, 2476–2530.
- (10) Schultz, C. *HFSP Journal* **2007**, *1*, 230–248.
- (11) Hwang, Y. W.; Miller, D. L. eng *The Journal of Biological Chemistry* **1987**, *262*, 13081–13085.
- (12) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. en *Science* **2001**, *292*, 498–500.
- (13) Lang, K.; Chin, J. W. *Chemical Reviews* **2014**, *114*, 4764–4806.
- (14) Xie, J.; Schultz, P. G. en *Nature Reviews Molecular Cell Biology* **2006**, *7*, 775–782.
- (15) Wu, Y.-W.; Goody, R. S. en *Journal of Peptide Science* **2010**, *16*, 514–523.
- (16) Wombacher, R.; Cornish, V. W. en *Journal of Biophotonics* **2011**, *4*, 391–402.
- (17) Jing, C.; Cornish, V. W. *Accounts of Chemical Research* **2011**, *44*, 784–792.
- (18) Johnsson, N.; Johnsson, K. *ACS Chemical Biology* **2007**, *2*, 31–38.
- (19) Banaszynski, L. A.; Wandless, T. J. *Chemistry & Biology* **2006**, *13*, 11–21.
- (20) Hinner, M. J.; Johnsson, K. *Current Opinion in Biotechnology* **2010**, *21*, 766–776.
- (21) Liu, C. C.; Schultz, P. G. *Annual Review of Biochemistry* **2010**, *79*, 413–444.
- (22) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. en *Science* **1998**, *281*, 269–272.
- (23) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. *Journal of the American Chemical Society* **2002**, *124*, 6063–6076.
- (24) Cao, H.; Xiong, Y.; Wang, T.; Chen, B.; Squier, T. C.; Mayer, M. U. *Journal of the American Chemical Society* **2007**, *129*, 8672–8673.
- (25) Lelek, M.; Nunzio, F. D.; Henriques, R.; Charneau, P.; Arhel, N.; Zimmer, C. en *Proceedings of the National Academy of Sciences* **2012**, *109*, 8564–8569.
- (26) Pereira, C. F.; Ellenberg, P. C.; Jones, K. L.; Fernandez, T. L.; Smyth, R. P.; Hawkes, D. J.; Hijnen, M.; Vivet-Boudou, V.; Marquet, R.; Johnson, I.; Mak, J. *PLoS ONE* **2011**, *6*, e17016.

- 
- (27) Cottet, M.; Faklaris, O.; Trinquet, E.; Pin, J.-P.; Durroux, T. en In *Fluorescent Methods to Study Biological Membranes*, Mély, Y., Duportail, G., Eds.; Springer Series on Fluorescence 13; Springer Berlin Heidelberg: 2013, pp 389–415.
- (28) Nakanishi, J.; Takarada, T.; Yunoki, S.; Kikuchi, Y.; Maeda, M. *Biochemical and Biophysical Research Communications* **2006**, *343*, 1191–1196.
- (29) Lohse, M. J.; Nuber, S.; Hoffmann, C. en *Pharmacological Reviews* **2012**, *64*, 299–336.
- (30) Roberti, M. J.; Bertoncini, C. W.; Klement, R.; Jares-Erijman, E. A.; Jovin, T. M. en *Nature Methods* **2007**, *4*, 345–351.
- (31) Ignatova, Z.; Gierasch, L. M. *Biochemistry* **2005**, *44*, 7266–7274.
- (32) Chaumont, S.; Khakh, B. S. en *Proceedings of the National Academy of Sciences* **2008**, *105*, 12063–12068.
- (33) Andresen, M.; Schmitz-Salue, R.; Jakobs, S. en *Molecular Biology of the Cell* **2004**, *15*, 5616–5622.
- (34) Pace, C. J.; Huang, Q.; Wang, F.; Palla, K. S.; Fuller, A. A.; Gao, J. en *ChemBioChem* **2011**, *12*, 1018–1022.
- (35) Dyachok, O.; Isakov, Y.; Sâgetorp, J.; Tengholm, A. en *Nature* **2006**, *439*, 349–352.
- (36) Panchal, R. G.; Ruthel, G.; Kenny, T. A.; Kallstrom, G. H.; Lane, D.; Badie, S. S.; Li, L.; Bavari, S.; Aman, M. J. en *Proceedings of the National Academy of Sciences* **2003**, *100*, 15936–15941.
- (37) Irtegun, S.; Wood, R.; Lackovic, K.; Schweiggert, J.; Ramdzan, Y. M.; Huang, D. C. S.; Mulhern, T. D.; Hatters, D. M. *ACS Chemical Biology* **2014**, *9*, 1426–1431.
- (38) Spagnuolo, C.; Joselevich, M.; Leskow, F. C.; Jares-Erijman, E. A. en In *Advanced Fluorescence Reporters in Chemistry and Biology III*, Demchenko, A. P., Ed.; Springer Series on Fluorescence 113; Springer Berlin Heidelberg: Jan. 2011, pp 263–295.
- (39) Li, Y.; Shen, L.; Li, C.; Huang, J.; Zhao, B.; Sun, Y.; Li, S.; Luo, Y.; Qiu, H.-J. *Virus Research* **2014**, *189*, 67–74.
- (40) Scheck, R. A.; Schepartz, A. *Accounts of Chemical Research* **2011**, *44*, 654–665.
- (41) Hoffmann, C.; Gaietta, G.; Bünemann, M.; Adams, S. R.; Oberdorff-Maass, S.; Behr, B.; Vilardaga, J.-P.; Tsien, R. Y.; Ellisman, M. H.; Lohse, M. J. en *Nature Methods* **2005**, *2*, 171–176.
- (42) Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A. *Journal of the American Chemical Society* **2009**, *131*, 438–439.
- (43) Guignet, E. G.; Hovius, R.; Vogel, H. en *Nature Biotechnology* **2004**, *22*, 440–444.
- (44) Hauser, C. T.; Tsien, R. Y. en *Proceedings of the National Academy of Sciences* **2007**, *104*, 3693–3697.
- (45) Ojida, A.; Honda, K.; Shinmi, D.; Kiyonaka, S.; Mori, Y.; Hamachi, I. *Journal of the American Chemical Society* **2006**, *128*, 10452–10459.
- (46) Franz, K. J.; Nitz, M.; Imperiali, B. en *ChemBioChem* **2003**, *4*, 265–271.
- (47) Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufls, F.; Johnsson, K. *English Chemistry & Biology* **2008**, *15*, 128–136.
- (48) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. en *Nature Biotechnology* **2003**, *21*, 86–89.
-



- (49) Juillerat, A.; Gronemeyer, T.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K. *Chemistry & Biology* **2003**, *10*, 313–317.
- (50) Gronemeyer, T.; Chidley, C.; Juillerat, A.; Heinis, C.; Johnsson, K. en *Protein Engineering Design and Selection* **2006**, *19*, 309–316.
- (51) Jansen, L. E. T.; Black, B. E.; Foltz, D. R.; Cleveland, D. W. en *The Journal of Cell Biology* **2007**, *176*, 795–805.
- (52) Farr, G. A.; Hull, M.; Mellman, I.; Caplan, M. J. *The Journal of Cell Biology* **2009**, *186*, 269–282.
- (53) Milenkovic, L.; Scott, M. P.; Rohatgi, R. *The Journal of Cell Biology* **2009**, *187*, 365–374.
- (54) Maurel, D.; Comps-Agrar, L.; Brock, C.; Rives, M.-L.; Bourrier, E.; Ayoub, M. A.; Bazin, H.; Tinel, N.; Durroux, T.; Prézeau, L.; Trinquet, E.; Pin, J.-P. en *Nature Methods* **2008**, *5*, 561–567.
- (55) Hein, B.; Willig, K. I.; Wurm, C. A.; Westphal, V.; Jakobs, S.; Hell, S. W. *Biophysical Journal* **2010**, *98*, 158–163.
- (56) Jones, S. A.; Shim, S.-H.; He, J.; Zhuang, X. en *Nature Methods* **2011**, *8*, 499–505.
- (57) Klein, T.; Löschberger, A.; Proppert, S.; Wolter, S.; van de Linde, S.; Sauer, M. en *Nature Methods* **2011**, *8*, 7–9.
- (58) Malkusch, S.; Muranyi, W.; Müller, B.; Kräusslich, H.-G.; Heilemann, M. eng *Histochemistry and Cell Biology* **2013**, *139*, 173–179.
- (59) Los, G. V. et al. *ACS Chemical Biology* **2008**, *3*, 373–382.
- (60) Hughes, B. G.; Fan, X.; Cho, W. J.; Schulz, R. en *American Journal of Physiology - Heart and Circulatory Physiology* **2014**, *306*, H764–H770.
- (61) Appelhans, T.; Richter, C. P.; Wilkens, V.; Hess, S. T.; Piehler, J.; Busch, K. B. *Nano Letters* **2012**, *12*, 610–616.
- (62) Ose, R.; Oharaa, O.; Nagase, T. *Current Chemical Genomics* **2012**, *6*, 18–26.
- (63) Tae, H. S.; Sundberg, T. B.; Neklesa, T. K.; Noblin, D. J.; Gustafson, J. L.; Roth, A. G.; Raina, K.; Crews, C. M. en *ChemBioChem* **2012**, *13*, 538–541.
- (64) Neklesa, T. K.; Tae, H. S.; Schneekloth, A. R.; Stulberg, M. J.; Corson, T. W.; Sundberg, T. B.; Raina, K.; Holley, S. A.; Crews, C. M. en *Nature Chemical Biology* **2011**, *7*, 538–543.
- (65) Raina, K.; Noblin, D. J.; Serebrenik, Y. V.; Adams, A.; Zhao, C.; Crews, C. M. en *Nature Chemical Biology* **2014**, *10*, 957–962.
- (66) Chidley, C.; Mosiewicz, K.; Johnsson, K. *Bioconjugate Chemistry* **2008**, *19*, 1753–1756.
- (67) Mizukami, S.; Watanabe, S.; Hori, Y.; Kikuchi, K. *Journal of the American Chemical Society* **2009**, *131*, 5016–5017.
- (68) Hori, Y.; Ueno, H.; Mizukami, S.; Kikuchi, K. *Journal of the American Chemical Society* **2009**, *131*, 16610–16611.
- (69) Hori, Y.; Nakaki, K.; Sato, M.; Mizukami, S.; Kikuchi, K. en *Angewandte Chemie International Edition* **2012**, *51*, 5611–5614.
- (70) Hori, Y.; Norinobu, T.; Sato, M.; Arita, K.; Shirakawa, M.; Kikuchi, K. *Journal of the American Chemical Society* **2013**, *135*, 12360–12365.

- 
- (71) Amara, J. F.; Clackson, T.; Rivera, V. M.; Guo, T.; Keenan, T.; Natesan, S.; Pollock, R.; Yang, W.; Courage, N. L.; Holt, D. A.; Gilman, M. en *Proceedings of the National Academy of Sciences* **1997**, *94*, 10618–10623.
- (72) Clackson, T.; Yang, W.; Rozamus, L. W.; Hatada, M.; Amara, J. F.; Rollins, C. T.; Stevenson, L. F.; Magari, S. R.; Wood, S. A.; Courage, N. L.; Lu, X.; Cerasoli, F.; Gilman, M.; Holt, D. A. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 10437–10442.
- (73) Rollins, C. T.; Rivera, V. M.; Woolfson, D. N.; Keenan, T.; Hatada, M.; Adams, S. E.; Andrade, L. J.; Yaeger, D.; Schravendijk, M. R. v.; Holt, D. A.; Gilman, M.; Clackson, T. en *Proceedings of the National Academy of Sciences* **2000**, *97*, 7096–7101.
- (74) Rivera, V. M.; Wang, X.; Wardwell, S.; Courage, N. L.; Volchuk, A.; Keenan, T.; Holt, D. A.; Gilman, M.; Orci, L.; Cerasoli, F.; Rothman, J. E.; Clackson, T. en *Science* **2000**, *287*, 826–830.
- (75) Bonger, K. M.; Chen, L.-c.; Liu, C. W.; Wandless, T. J. en *Nature Chemical Biology* **2011**, *7*, 531–537.
- (76) Miller, L. W.; Cai, Y.; Sheetz, M. P.; Cornish, V. W. en *Nature Methods* **2005**, *2*, 255–257.
- (77) Gallagher, S. S.; Sable, J. E.; Sheetz, M. P.; Cornish, V. W. *ACS Chemical Biology* **2009**, *4*, 547–556.
- (78) Iwamoto, M.; Björklund, T.; Lundberg, C.; Kirik, D.; Wandless, T. J. *Chemistry & Biology* **2010**, *17*, 981–988.
- (79) Long, M. J. C.; Gollapalli, D. R.; Hedstrom, L. *Chemistry & biology* **2012**, *19*, 629–637.
- (80) Ishida, M.; Watanabe, H.; Takigawa, K.; Kurishita, Y.; Oki, C.; Nakamura, A.; Hamachi, I.; Tsukiji, S. *Journal of the American Chemical Society* **2013**, *135*, 12684–12689.
- (81) Howarth, M.; Takao, K.; Hayashi, Y.; Ting, A. Y. en *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 7583–7588.
- (82) De Boer, E.; Rodriguez, P.; Bonte, E.; Krijgsveld, J.; Katsantoni, E.; Heck, A.; Grosveld, F.; Strouboulis, J. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 7480–7485.
- (83) Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. en *Nature Methods* **2005**, *2*, 99–104.
- (84) Slavoff, S. A.; Chen, I.; Choi, Y.-A.; Ting, A. Y. *Journal of the American Chemical Society* **2008**, *130*, 1160–1162.
- (85) Fernandez-Suarez, M.; Baruah, H.; Martinez-Hernandez, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nature biotechnology* **2007**, *25*, 1483–1487.
- (86) Uttamapinant, C.; White, K. A.; Baruah, H.; Thompson, S.; Fernández-Suárez, M.; Puthenveetil, S.; Ting, A. Y. en *Proceedings of the National Academy of Sciences* **2010**, *107*, 10914–10919.
- (87) Cohen, J. D.; Thompson, S.; Ting, A. Y. *Biochemistry* **2011**, *50*, 8221–8225.
- (88) Uttamapinant, C.; Tangpeerachaikul, A.; Grecian, S.; Clarke, S.; Singh, U.; Slade, P.; Gee, K. R.; Ting, A. Y. en *Angewandte Chemie International Edition* **2012**, *51*, 5852–5856.
- (89) Cohen, J. D.; Zou, P.; Ting, A. Y. en *ChemBioChem* **2012**, *13*, 888–894.
-

## References

---

- (90) Hauke, S.; Best, M.; Schmidt, T. T.; Baalman, M.; Krause, A.; Wombacher, R. *Bioconjugate Chemistry* **2014**, *25*, 1632–1637.
- (91) Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y. *Journal of the American Chemical Society* **2012**, *134*, 792–795.
- (92) Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. en *Proceedings of the National Academy of Sciences* **2009**, *106*, 3000–3005.
- (93) Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. en *Nature Chemical Biology* **2007**, *3*, 321–322.
- (94) Rush, J. S.; Bertozzi, C. R. *Journal of the American Chemical Society* **2008**, *130*, 12240–12241.
- (95) Lin, C.-W.; Ting, A. Y. *Journal of the American Chemical Society* **2006**, *128*, 4542–4543.
- (96) Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O. en *Science* **1999**, *285*, 760–763.
- (97) Popp, M. W.-L.; Antos, J. M.; Ploegh, H. L. en In *Current Protocols in Protein Science*; John Wiley & Sons, Inc.: 2009; Vol. 56; Chapter 15.3.1–15.3.9.
- (98) Tanaka, T.; Yamamoto, T.; Tsukiji, S.; Nagamune, T. en *ChemBioChem* **2008**, *9*, 802–807.
- (99) Strijbis, K.; Spooner, E.; Ploegh, H. L. en *Traffic* **2012**, *13*, 780–789.
- (100) George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K. *Journal of the American Chemical Society* **2004**, *126*, 8896–8897.
- (101) Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *Journal of the American Chemical Society* **2004**, *126*, 7754–7755.
- (102) Clarke, K. M.; Mercer, A. C.; La Clair, J. J.; Burkart, M. D. *Journal of the American Chemical Society* **2005**, *127*, 11234–11235.
- (103) Yin, J.; Straight, P. D.; McLoughlin, S. M.; Zhou, Z.; Lin, A. J.; Golan, D. E.; Kelleher, N. L.; Kolter, R.; Walsh, C. T. en *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 15815–15820.
- (104) Zhou, Z.; Cironi, P.; Lin, A. J.; Xu, Y.; Hrvatin, S.; Golan, D. E.; Silver, P. A.; Walsh, C. T.; Yin, J. *ACS Chemical Biology* **2007**, *2*, 337–346.
- (105) Duckworth, B. P.; Zhang, Z.; Hosokawa, A.; Distefano, M. D. en *ChemBioChem* **2007**, *8*, 98–105.
- (106) Dursina, B.-E.; Reents, R.; Niculae, A.; Veligodsky, A.; Breitling, R.; Pyatkov, K.; Waldmann, H.; Goody, R. S.; Alexandrov, K. *Protein Expression and Purification* **2005**, *39*, 71–81.
- (107) Wang, Y.-C.; Dozier, J. K.; Beese, L. S.; Distefano, M. D. *ACS Chemical Biology* **2014**, *9*, 1726–1735.
- (108) Gangopadhyay, S. A.; Losito, E. L.; Hougland, J. L. *Biochemistry* **2014**, *53*, 434–446.
- (109) Dozier, J. K.; Khatwani, S. L.; Wollack, J. W.; Wang, Y.-C.; Schmidt-Dannert, C.; Distefano, M. D. *Bioconjugate Chemistry* **2014**, *25*, 1203–1212.
- (110) Sinz, A. en *Journal of Mass Spectrometry* **2003**, *38*, 1225–1237.
- (111) Kalkhof, S.; Sinz, A. en *Analytical and Bioanalytical Chemistry* **2008**, *392*, 305–312.
- (112) Kelly, K. A.; Bardeesy, N.; Anbazhagan, R.; Gurumurthy, S.; Berger, J.; Alencar, H.; DePinho, R. A.; Mahmood, U.; Weissleder, R. *PLoS Med* **2008**, *5*, e85.

- 
- (113) Hino, N.; Okazaki, Y.; Kobayashi, T.; Hayashi, A.; Sakamoto, K.; Yokoyama, S. en *Nature Methods* **2005**, *2*, 201–206.
- (114) Suchanek, M.; Radzikowska, A.; Thiele, C. en *Nature Methods* **2005**, *2*, 261–268.
- (115) Luedtke, N. W.; Dexter, R. J.; Fried, D. B.; Schepartz, A. *Nature chemical biology* **2007**, *3*, 779–784.
- (116) Krishnan, B.; Gierasch, L. M. *Chemistry & biology* **2008**, *15*, 1104–1115.
- (117) Rutkowska, A.; Haering, C. H.; Schultz, C. en *Angewandte Chemie International Edition* **2011**, *50*, 12655–12658.
- (118) Gautier, A.; Nakata, E.; Lukinavičius, G.; Tan, K.-T.; Johnsson, K. *Journal of the American Chemical Society* **2009**, *131*, 17954–17962.
- (119) Lukinavičius, G.; Lavogina, D.; Orpinell, M.; Umezawa, K.; Reymond, L.; Garin, N.; Gönczy, P.; Johnsson, K. *Current Biology* **2013**, *23*, 265–270.
- (120) Sun, X.; Dusserre-Bresson, F.; Baker, B.; Zhang, A.; Xu, P.; Fibbe, C.; Noren, C. J.; Corrêa Jr., I. R.; Xu, M.-Q. *European Journal of Medicinal Chemistry* **2014**, *88*, 34–41.
- (121) Lemercier, G.; Gendreizig, S.; Kindermann, M.; Johnsson, K. en *Angewandte Chemie* **2007**, *119*, 4359–4362.
- (122) Baruah, H.; Puthenveetil, S.; Choi, Y.-A.; Shah, S.; Ting, A. en *Angewandte Chemie International Edition* **2008**, *47*, 7018–7021.
- (123) Slavoff, S. A.; Liu, D. S.; Cohen, J. D.; Ting, A. Y. *Journal of the American Chemical Society* **2011**, *133*, 19769–19776.
- (124) Yan, P.; Wang, T.; Newton, G. J.; Knyushko, T. V.; Xiong, Y.; Bigelow, D. J.; Squier, T. C.; Mayer, M. U. en *ChemBioChem* **2009**, *10*, 1507–1518.
- (125) Corson, T. W.; Aberle, N.; Crews, C. M. *ACS Chemical Biology* **2008**, *3*, 677–692.
- (126) Liang, F.-S.; Ho, W. Q.; Crabtree, G. R. en *Science Signaling* **2011**, *4*, rs2–rs2.
- (127) Farrar, M. A.; Alberola-Ila, J.; Perlmutter, R. M. en *Nature* **1996**, *383*, 178–181.
- (128) Erhart, D.; Zimmermann, M.; Jacques, O.; Wittwer, M. B.; Ernst, B.; Constable, E.; Zvelebil, M.; Beaufils, F.; Wymann, M. P. *Chemistry & Biology* **2013**, *20*, 549–557.
- (129) Haruki, H.; Nishikawa, J.; Laemmli, U. K. *Molecular Cell* **2008**, *31*, 925–932.
- (130) Robinson, M. S.; Sahlender, D. A.; Foster, S. D. *Developmental Cell* **2010**, *18*, 324–331.
- (131) Inoue, T.; Heo, W. D.; Grimley, J. S.; Wandless, T. J.; Meyer, T. en *Nature Methods* **2005**, *2*, 415–418.
- (132) Castellano, F.; Montcourrier, P.; Chavrier, P. en *Journal of Cell Science* **2000**, *113*, 2955–2961.
- (133) Gestwicki, J. E.; Crabtree, G. R.; Graef, I. A. en *Science* **2004**, *306*, 865–869.
- (134) Braun, P. D.; Barglow, K. T.; Lin, Y.-M.; Akompong, T.; Briesewitz, R.; Ray, G. T.; Haldar, K.; Wandless, T. J. *Journal of the American Chemical Society* **2003**, *125*, 7575–7580.
- (135) Sellmyer, M. A.; Stankunas, K.; Briesewitz, R.; Crabtree, G. R.; Wandless, T. J. *Bioorganic & Medicinal Chemistry Letters* **2007**, *17*, 2703–2705.
- (136) Pollock, R.; Clackson, T. *Current Opinion in Biotechnology* **2002**, *13*, 459–467.
-

## References

---

- (137) Liu, P.; Calderon, A.; Konstantinidis, G.; Hou, J.; Voss, S.; Chen, X.; Li, F.; Banerjee, S.; Hoffmann, J.-E.; Theiss, C.; Dehmelt, L.; Wu, Y.-W. en *Angewandte Chemie International Edition* **2014**, *53*, 10049–10055.
- (138) Janse, D. M.; Crosas, B.; Finley, D.; Church, G. M. en *Journal of Biological Chemistry* **2004**, *279*, 21415–21420.
- (139) Buckley, D. L.; Crews, C. M. en *Angewandte Chemie International Edition* **2014**, *53*, 2312–2330.
- (140) Pratt, M. R.; Schwartz, E. C.; Muir, T. W. en *Proceedings of the National Academy of Sciences* **2007**, *104*, 11209–11214.
- (141) Bayle, J. H.; Grimley, J. S.; Stankunas, K.; Gestwicki, J. E.; Wandless, T. J.; Crabtree, G. R. *Chemistry & Biology* **2006**, *13*, 99–107.
- (142) Stankunas, K.; Bayle, J. H.; Gestwicki, J. E.; Lin, Y.-M.; Wandless, T. J.; Crabtree, G. R. *Molecular Cell* **2003**, *12*, 1615–1624.
- (143) Mootz, H. D.; Muir, T. W. *Journal of the American Chemical Society* **2002**, *124*, 9044–9045.
- (144) Mootz, H. D.; Blum, E. S.; Tyszkiewicz, A. B.; Muir, T. W. *Journal of the American Chemical Society* **2003**, *125*, 10561–10569.
- (145) Spencer, D. M.; Wandless, T. J.; Schreiber, S. L.; Crabtree, G. R. en *Science* **1993**, *262*, 1019–1024.
- (146) Keenan, T.; Yaeger, D. R.; Courage, N. L.; Rollins, C. T.; Pavone, M. E.; Rivera, V. M.; Yang, W.; Guo, T.; Amara, J. F.; Clackson, T.; Gilman, M.; Holt, D. A. *Bioorganic & Medicinal Chemistry* **1998**, *6*, 1309–1335.
- (147) Yang, W.; Keenan, T. P.; Rozamus, L. W.; Wang, X.; Rivera, V. M.; Rollins, C. T.; Clackson, T.; Holt, D. A. *Bioorganic & Medicinal Chemistry Letters* **2003**, *13*, 3181–3184.
- (148) Cheng, J.; Yu, L.; Zhang, D.; Huang, Q.; Spencer, D.; Su, B. en *Journal of Biological Chemistry* **2005**, *280*, 13477–13482.
- (149) Spencer, D. M.; Belshaw, P. J.; Chen, L.; Ho, S. N.; Randazzo, F.; Crabtree, G. R.; Schreiber, S. L. *Current Biology* **1996**, *6*, 839–847.
- (150) Thomis, D. C.; Marktel, S.; Bonini, C.; Traversari, C.; Gilman, M.; Bordignon, C.; Clackson, T. en *Blood* **2001**, *97*, 1249–1257.
- (151) MacCorkle, R. A.; Freeman, K. W.; Spencer, D. M. en *Proceedings of the National Academy of Sciences* **1998**, *95*, 3655–3660.
- (152) Fan, L.; Freeman, K. W.; Khan, T.; Pham, E.; Spencer, D. M. *Human Gene Therapy* **1999**, *10*, 2273–2285.
- (153) Mallet, V. O.; Mitchell, C.; Guidotti, J.-E.; Jaffray, P.; Fabre, M.; Spencer, D.; Arnoult, D.; Kahn, A.; Gilgenkrantz, H. en *Nature Biotechnology* **2002**, *20*, 1234–1239.
- (154) Carlotti, F.; Zaldumbide, A.; Martin, P.; Boulukos, K. E.; Hoeben, R. C.; Pognonec, P. en *Cancer Gene Therapy* **2005**, *12*, 627–639.
- (155) Smith, K. M.; Etten, R. A. V. en *Journal of Biological Chemistry* **2001**, *276*, 24372–24379.
- (156) Je, H.-S.; Lu, Y.; Yang, F.; Nagappan, G.; Zhou, J.; Jiang, Z.; Nakazawa, K.; Lu, B. en *The Journal of Neuroscience* **2009**, *29*, 6761–6766.

- 
- (157) Farrar, M. A.; Olson, S. H.; Perlmutter, R. M. In *Methods in Enzymology*, Jeremy Thorner, S. D. E. a. J. N. A., Ed.; Applications of Chimeric Genes and Hybrid Proteins - Part B: Cell Biology and Physiology 421–429, Vol. Volume 327; Academic Press: 2000, 421–IN5.
- (158) Farrar, M. A.; Tian, J.; Perlmutter, R. M. en *Journal of Biological Chemistry* **2000**, *275*, 31318–31324.
- (159) Mohi, M. G.; Arai, K.-i.; Watanabe, S. en *Molecular Biology of the Cell* **1998**, *9*, 3299–3308.
- (160) Mizuguchi, R.; Hatakeyama, M. en *Journal of Biological Chemistry* **1998**, *273*, 32297–32303.
- (161) O'Farrell, A. M.; Liu, Y.; Moore, K. W.; Mui, A. L. *The EMBO Journal* **1998**, *17*, 1006–1018.
- (162) O'Farrell, A.-M.; Parry, D. A.; Zindy, F.; Roussel, M. F.; Lees, E.; Moore, K. W.; Mui, A. L.-F. en *The Journal of Immunology* **2000**, *164*, 4607–4615.
- (163) Liu R Liu C.-H., M. M. G. *Oncogene* **2000**, *19*, 571–579.
- (164) Mizuguchi, R.; Noto, S.; Yamada, M.; Ashizawa, S.; Higashi, H.; Hatakeyama, M. eng *Japanese Journal of Cancer Research: Gann* **2000**, *91*, 527–533.
- (165) Boyd, Z. S.; Kriatchko, A.; Yang, J.; Agarwal, N.; Wax, M. B.; Patil, R. V. en *Investigative Ophthalmology & Visual Science* **2003**, *44*, 5206–5211.
- (166) Li, X.; Steeber, D. A.; Tang, M. L. K.; Farrar, M. A.; Perlmutter, R. M.; Tedder, T. F. en *The Journal of Experimental Medicine* **1998**, *188*, 1385–1390.
- (167) Kume, A.; Ito, K.; Ueda, Y.; Hasegawa, M.; Urabe, M.; Mano, H.; Ozawa, K. *Biochemical and Biophysical Research Communications* **1999**, *260*, 9–12.
- (168) E. Knight A. Warner, A. M. *Oncogene* **2000**, *19*, 5398–5405.
- (169) Perron, A.; Chen, Z.-g.; Gingras, D.; Dupré, D. J.; Stanková, J.; Rola-Pleszczynski, M. en *Journal of Biological Chemistry* **2003**, *278*, 27956–27965.
- (170) Gotoh, Y.; Cooper, J. A. en *Journal of Biological Chemistry* **1998**, *273*, 17477–17482.
- (171) Sanz, L.; Diaz-Meco, M. T.; Nakano, H.; Moscat, J. *The EMBO Journal* **2000**, *19*, 1576–1586.
- (172) Ung, T. L.; Cao, C.; Lu, J.; Ozato, K.; Dever, T. E. *The EMBO Journal* **2001**, *20*, 3728–3737.
- (173) Zhao, H.-F.; Boyd, J.; Jolicoeur, N.; Shen, S.-H. *Human Gene Therapy* **2003**, *14*, 1619–1629.
- (174) Belshaw, P. J.; Schoepfer, J. G.; Liu, K.-Q.; Morrison, K. L.; Schreiber, S. L. en *Angewandte Chemie International Edition in English* **1995**, *34*, 2129–2132.
- (175) Belshawl, P. J.; Spencer, D. M.; Crabtree, G. R.; Schreiber, S. L. *Chemistry & Biology* **1996**, *3*, 731–738.
- (176) Kopytek, S. J.; Standaert, R. F.; Dyer, J. C.; Hu, J. C. *Chemistry & Biology* **2000**, *7*, 313–321.
- (177) Chen, J; Zheng, X. F.; Brown, E. J.; Schreiber, S. L. *Proceedings of the National Academy of Sciences of the United States of America* **1995**, *92*, 4947–4951.
- (178) Choi, J.; Chen, J.; Schreiber, S. L.; Clardy, J. eng *Science (New York, N.Y.)* **1996**, *273*, 239–242.
-

## References

---

- (179) Rivera, V. M.; Clackson, T.; Natesan, S.; Pollock, R.; Amara, J. F.; Keenan, T.; Magari, S. R.; Phillips, T.; Courage, N. L.; Cerasoli, F.; Holt, D. A.; Gilman, M. en *Nature Medicine* **1996**, *2*, 1028–1032.
- (180) Putyrski, M.; Schultz, C. *FEBS Letters* **2012**, *586*, 2097–2105.
- (181) Luengo, J. I. et al. *Chemistry & Biology* **1995**, *2*, 471–481.
- (182) Liberles, S. D.; Diver, S. T.; Austin, D. J.; Schreiber, S. L. en *Proceedings of the National Academy of Sciences* **1997**, *94*, 7825–7830.
- (183) Banaszynski, L. A.; Liu, C. W.; Wandless, T. J. *Journal of the American Chemical Society* **2005**, *127*, 4715–4721.
- (184) Kozany, C.; März, A.; Kress, C.; Hausch, F. en *ChemBioChem* **2009**, *10*, 1402–1410.
- (185) Umeda, N.; Ueno, T.; Pohlmeier, C.; Nagano, T.; Inoue, T. *Journal of the American Chemical Society* **2011**, *133*, 12–14.
- (186) Karginov, A. V.; Zou, Y.; Shirvanyants, D.; Kota, P.; Dokholyan, N. V.; Young, D. D.; Hahn, K. M.; Deiters, A. *Journal of the American Chemical Society* **2011**, *133*, 420–423.
- (187) Terrillon, S.; Bouvier, M. en *The EMBO Journal* **2004**, *23*, 3950–3961.
- (188) Putyrski, M.; Schultz, C. *Chemistry & Biology* **2011**, *18*, 1126–1133.
- (189) Heo, W. D.; Inoue, T.; Park, W. S.; Kim, M. L.; Park, B. O.; Wandless, T. J.; Meyer, T. en *Science* **2006**, *314*, 1458–1461.
- (190) Suh, B.-C.; Inoue, T.; Meyer, T.; Hille, B. en *Science* **2006**, *314*, 1454–1457.
- (191) Szentpetery, Z.; Várnai, P.; Balla, T. en *Proceedings of the National Academy of Sciences* **2010**, *107*, 8225–8230.
- (192) Klemm, J. D.; Beals, C. R.; Crabtree, G. R. *Current Biology* **1997**, *7*, 638–644.
- (193) Geda, P.; Patury, S.; Ma, J.; Bharucha, N.; Dobry, C. J.; Lawson, S. K.; Gestwicki, J. E.; Kumar, A. en *Yeast* **2008**, *25*, 577–594.
- (194) Busch, A.; Kiel, T.; Hübner, S. en *Traffic* **2009**, *10*, 1221–1227.
- (195) Xu, T.; Johnson, C. A.; Gestwicki, J. E.; Kumar, A. en *Nature Protocols* **2010**, *5*, 1831–1843.
- (196) Miyamoto, T.; DeRose, R.; Suarez, A.; Ueno, T.; Chen, M.; Sun, T.-p.; Wolfgang, M. J.; Mukherjee, C.; Meyers, D. J.; Inoue, T. en *Nature Chemical Biology* **2012**, *8*, 465–470.
- (197) Skwarczynska, M.; Molzan, M.; Ottmann, C. en *Proceedings of the National Academy of Sciences* **Jan. 2013**, *110*, E377–E386.
- (198) Licitra, E. J.; Liu, J. O. en *Proceedings of the National Academy of Sciences* **1996**, *93*, 12817–12821.
- (199) Belshaw, P. J.; Ho, S. N.; Crabtree, G. R.; Schreiber, S. L. en *Proceedings of the National Academy of Sciences* **1996**, *93*, 4604–4607.
- (200) Czlapinski, J. L.; Schelle, M. W.; Miller, L. W.; Laughlin, S. T.; Kohler, J. J.; Cornish, V. W.; Bertozzi, C. R. *Journal of the American Chemical Society* **2008**, *130*, 13186–13187.
- (201) Feng, S.; Laketa, V.; Stein, F.; Rutkowska, A.; MacNamara, A.; Depner, S.; Klingmüller, U.; Saez-Rodriguez, J.; Schultz, C. en *Angewandte Chemie International Edition* **2014**, *53*, 6720–6723.
- (202) Lin, H.; Abida, W. M.; Sauer, R. T.; Cornish, V. W. *Journal of the American Chemical Society* **2000**, *122*, 4247–4248.

- 
- (203) Abida, W. M.; Carter, B. T.; Althoff, E. A.; Lin, H.; Cornish, V. W. en *ChemBioChem* **2002**, *3*, 887–895.
- (204) Carter, B. T.; Lin, H.; Goldberg, S. D.; Althoff, E. A.; Raushel, J.; Cornish, V. W. en *ChemBioChem* **2005**, *6*, 2055–2067.
- (205) Gallagher, S. S.; Miller, L. W.; Cornish, V. W. *Analytical Biochemistry* **2007**, *363*, 160–162.
- (206) Athavankar, S.; Peterson, B. R. *Chemistry & Biology* **2003**, *10*, 1245–1253.
- (207) Weber, W.; Stelling, J.; Rimann, M.; Keller, B.; Daoud-El Baba, M.; Weber, C. C.; Aubel, D.; Fussenegger, M. eng *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 2643–2648.
- (208) Yoshimura, A.; Mizukami, S.; Mori, Y.; Yoshioka, Y.; Kikuchi, K. *Chemistry Letters* **2014**, *43*, 219–221.
- (209) Hussey, S. L.; Muddana, S. S.; Peterson, B. R. *Journal of the American Chemical Society* **2003**, *125*, 3692–3693.
- (210) Muddana, S. S.; Peterson, B. R. *Organic Letters* **2004**, *6*, 1409–1412.
- (211) Damoiseaux, R.; Keppler, A.; Johnsson, K. en *ChemBioChem* **2001**, *2*, 285–287.
- (212) Gendreizig, S.; Kindermann, M.; Johnsson, K. *Journal of the American Chemical Society* **2003**, *125*, 14970–14971.
- (213) Etoc, F.; Lisse, D.; Bellaiche, Y.; Piehler, J.; Coppey, M.; Dahan, M. en *Nature Nanotechnology* **2013**, *8*, 193–198.
- (214) Brieke, C.; Rohrbach, F.; Gottschalk, A.; Mayer, G.; Heckel, A. en *Angewandte Chemie International Edition* **2012**, *51*, 8446–8476.
- (215) Wilkins, B. J.; Yang, X.; Cropp, T. A. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 4296–4298.
- (216) Sadowski, O.; Jaikaran, A. S. I.; Samanta, S.; Fabian, M. R.; Dowling, R. J. O.; Sonenberg, N.; Woolley, G. A. *Bioorganic & Medicinal Chemistry* **2010**, *18*, 7746–7752.
- (217) Banala, S.; Arnold, A.; Johnsson, K. en *ChemBioChem* **2008**, *9*, 38–41.
- (218) Ballister, E. R.; Aonbangkhen, C.; Mayo, A. M.; Lampson, M. A.; Chenoweth, D. M. en *Nature Communications* **2014**, *5*, 5475.
- (219) Mayer, G.; Heckel, A. en *Angewandte Chemie International Edition* **2006**, *45*, 4900–4921.
- (220) Ahmed, S.; Xie, J.; Horne, D.; Williams, J. C. en *Journal of Biological Chemistry* **2014**, *289*, 4546–4552.
- (221) England, P. M.; Lester, H. A.; Davidson, N.; Dougherty, D. A. en *Proceedings of the National Academy of Sciences* **1997**, *94*, 11025–11030.
- (222) Endo, M.; Nakayama, K.; Kaida, Y.; Majima, T. en *Angewandte Chemie International Edition* **2004**, *43*, 5643–5645.
- (223) Hahn, M. E.; Muir, T. W. en *Angewandte Chemie International Edition* **2004**, *43*, 5800–5803.
- (224) Pellois, J.-P.; Hahn, M. E.; Muir, T. W. *Journal of the American Chemical Society* **2004**, *126*, 7170–7171.
- (225) Peters, F. B.; Brock, A.; Wang, J.; Schultz, P. G. *Chemistry & Biology* **2009**, *16*, 148–152.
- (226) Jay, D. G. en *Proceedings of the National Academy of Sciences* **1988**, *85*, 5454–5458.
-



## References

---

- (227) Surrey, T.; Elowitz, M. B.; Wolf, P.-E.; Yang, F.; Nedelec, F.; Shokat, K.; Leibler, S. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 4293–4298.
- (228) Marek, K. W.; Davis, G. W. *Neuron* **2002**, *36*, 805–813.
- (229) Marks, K. M.; Braun, P. D.; Nolan, G. P. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 9982–9987.
- (230) Keppler, A.; Ellenberg, J. *ACS Chemical Biology* **2009**, *4*, 127–138.
- (231) Takemoto, K.; Matsuda, T.; McDougall, M.; Klaubert, D. H.; Hasegawa, A.; Los, G. V.; Wood, K. V.; Miyawaki, A.; Nagai, T. *ACS Chemical Biology* **2011**, *6*, 401–406.
- (232) Jacobson, K.; Rajfur, Z.; Vitriol, E.; Hahn, K. English *Trends in Cell Biology* **2008**, *18*, 443–450.
- (233) Sano, Y.; Watanabe, W.; Matsunaga, S. en *Journal of Cell Science* **2014**, *127*, 1621–1629.
- (234) Wojtovich, A. P.; Foster, T. H. *Redox Biology* **2014**, *2*, 368–376.
- (235) Banghart, M. R.; Volgraf, M.; Trauner, D. *Biochemistry* **2006**, *45*, 15129–15141.
- (236) Beharry, A. A.; Woolley, G. A. en *Chemical Society Reviews* **2011**, *40*, 4422–4437.
- (237) Zhang, Y.; Erdmann, F.; Fischer, G. en *Nature Chemical Biology* **2009**, *5*, 724–726.
- (238) Shimizu-Sato, S.; Huq, E.; Tepperman, J. M.; Quail, P. H. en *Nature Biotechnology* **2002**, *20*, 1041–1044.
- (239) Tyszkiewicz, A. B.; Muir, T. W. en *Nature Methods* **2008**, *5*, 303–305.
- (240) Levskaya, A.; Weiner, O. D.; Lim, W. A.; Voigt, C. A. en *Nature* **2009**, *461*, 997–1001.
- (241) Yazawa, M.; Sadaghiani, A. M.; Hsueh, B.; Dolmetsch, R. E. en *Nature Biotechnology* **2009**, *27*, 941–945.
- (242) Motta-Mena, L. B.; Reade, A.; Mallory, M. J.; Glantz, S.; Weiner, O. D.; Lynch, K. W.; Gardner, K. H. en *Nature Chemical Biology* **2014**, *10*, 196–202.
- (243) Wang, X.; Chen, X.; Yang, Y. en *Nature Methods* **2012**, *9*, 266–269.
- (244) Wu, Y. I.; Frey, D.; Lungu, O. I.; Jaehrig, A.; Schlichting, I.; Kuhlman, B.; Hahn, K. M. en *Nature* **2009**, *461*, 104–108.
- (245) Kennedy, M. J.; Hughes, R. M.; Peteya, L. A.; Schwartz, J. W.; Ehlers, M. D.; Tucker, C. L. en *Nature Methods* **2010**, *7*, 973–975.
- (246) Zhou, X. X.; Chung, H. K.; Lam, A. J.; Lin, M. Z. en *Science* **2012**, *338*, 810–814.
- (247) Binkowski, B. F.; Miller, R. A.; Belshaw, P. J. eng *Chemistry & Biology* **2005**, *12*, 847–855.
- (248) Jullien, N.; Goddard, I.; Selmi-Ruby, S.; Fina, J.-L.; Cremer, H.; Herman, J.-P. *PLoS ONE* **2007**, *2*, e1355.
- (249) Kuo, C. J.; Chung, J.; Fiorentino, D. F.; Flanagan, W. M.; Blenis, J.; Crabtree, G. R. en *Nature* **1992**, *358*, 70–73.
- (250) Klán, P.; Šolomek, T.; Bochet, C. G.; Blanc, A.; Givens, R.; Rubina, M.; Popik, V.; Kostikov, A.; Wirz, J. *Chemical Reviews* **2013**, *113*, 119–191.
- (251) Walker, J. W.; McCray, J. A.; Hess, G. P. *Biochemistry* **1986**, *25*, 1799–1805.
- (252) Patchornik, A.; Amit, B.; Woodward, R. B. *Journal of the American Chemical Society* **1970**, *92*, 6333–6335.

- 
- (253) Görner, H. eng *Photochemical & Photobiological Sciences: Official Journal of the European Photochemistry Association and the European Society for Photobiology* **2005**, *4*, 822–828.
- (254) Šolomek, T.; Mercier, S.; Bally, T.; Bochet, C. G. en *Photochemical & Photobiological Sciences* **2012**, *11*, 548–555.
- (255) Eckardt, T.; Hagen, V.; Schade, B.; Schmidt, R.; Schweitzer, C.; Bendig, J. en *The Journal of Organic Chemistry* **2002**, *67*, 703–710.
- (256) So, M.-k.; Yao, H.; Rao, J. *Biochemical and Biophysical Research Communications* **2008**, *374*, 419–423.
- (257) Zimmermann, M.; Cal, R.; Janett, E.; Hoffmann, V.; Bochet, C. G.; Constable, E.; Beau-fils, F.; Wymann, M. P. en *Angewandte Chemie International Edition* **Apr. 2014**, *53*, 4717–4720.
- (258) Tsai, S.-C.; Klinman, J. P. *Bioorganic Chemistry* **2003**, *31*, 172–190.
- (259) Komatsu, T.; Kukelyansky, I.; McCaffery, J. M.; Ueno, T.; Varela, L. C.; Inoue, T. en *Nature Methods* **2010**, *7*, 206–208.
- (260) Stewart, M. en *Nature Reviews Molecular Cell Biology* **2007**, *8*, 195–208.
- (261) Hagen, V.; Dekowski, B.; Nache, V.; Schmidt, R.; Geißler, D.; Lorenz, D.; Eichhorst, J.; Keller, S.; Kaneko, H.; Benndorf, K.; Wiesner, B. en *Angewandte Chemie International Edition* **2005**, *44*, 7887–7891.
- (262) Mannekutla, J. R.; Mulimani, B. G.; Inamdar, S. R. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2008**, *69*, 419–426.
- (263) Sakata, T.; Kawashima, Y.; Nakano, H. en *International Journal of Quantum Chemistry* **2009**, *109*, 1940–1949.
- (264) Lin, Q.; Bao, C.; Yang, Y.; Liang, Q.; Zhang, D.; Cheng, S.; Zhu, L. *Adv Mater* **2013**, *25*, 1981–6.
- (265) Cürten, B.; Kullmann, P. H. M.; Bier, M. E.; Kandler, K.; Schmidt, B. F. en *Photochemistry and Photobiology* **May 2005**, *81*, 641–648.
- (266) Givens, R. S.; Rubina, M.; Wirz, J. *Photochem Photobiol Sci* **2012**, *11*, 472–88.
- (267) Pelliccioli, A. P.; Wirz, J. *Photochem Photobiol Sci* **2002**, *1*, 441–58.
- (268) Kim, M. S.; Diamond, S. L. *Bioorg Med Chem Lett* **2006**, *16*, 4007–10.

## References

---

### Personal data

Name	Ruben Cal
Address	Flughafenstrasse 8 4056 Basel, Switzerland
Phone	+41 (0) 61 381 68 59
Mobile	+41 (0) 79 847 71 39
E-Mail	ruben.cal@me.com
Date of birth	2.10.1985
native place	Le Noirmont (JU)
Marital status	single

### Education

---

- |                     |   |
|---------------------|---|
| 01.05.10 - 07.05.15 | <b>Ph.D. studies in organic chemistry at the University Basel</b><br>Conducted in the group of Prof. Dr. Matthias P. Wymann<br>Involved in the synthesis of organic molecules to investigate the role of proteins in the metabolic pathway<br>Keywords: Total synthesis, reaction optimization, Photo-chemistry, mechanistic studies, computer assisted research (Maestro, PyMol), scientific communication (written and oral)  |
| 01.09.08 - 16.12.09 | <b>Master of Science (M Sc) at the University Basel</b><br>Master thesis conducted in the group Prof. Dr. Thomas R. Ward<br>Responsible for the development of a novel methodology to regioselectively modify proteins<br>Keywords: Organometallic catalysis, regioselective catalysis, modification of biomolecules, artificial metalloenzymes<br><br>Internship in the research group of Prof. Dr. Helma Wennemers<br>Involved in the synthesis and screening of peptide libraries<br>Keywords: Organic catalysts, combinatorial chemistry, peptide synthesis, protecting group chemistry |
| 01.10.05 - 31.08.08 | <b>Bachelor of Science (B Sc) at the University Basel</b>   |

---

**Publications**


---

Ruben Cal, Mirjam Zimmermann, Edwin Constable, Florent Beaufils, und Matthias P. Wymann, manuscript in preparation

Mirjam Zimmermann, Ruben Cal, Viktor Hoffmann, Marketa Zvelebil, Edwin Constable, Florent Beaufils, and Matthias P. Wymann, manuscript in preparation

Ruben Cal, M. Zimmermann, E. Janett, V. Hoffmann, C. G. Bochet, E. Constable, F. Beaufils, and M. P. Wymann "Cell-Permeant and Photocleavable Chemical Inducer of Dimerization" Angew. Chem. Int. Ed., 2014, 53, 4717-4720 (Highlighted in CHIMIA, 2014, 68, 452)

---

**Selected Presentations**


---

**13th Belgian Organic Synthesis Symposium**, 2012, Leuven, Belgium Poster presentation: "Synthesis of photocleavable Chemical Inducer of Dimerization"

**USGEB 2012 Molecular Recognition: When Biology meets Chemistry**, 2012, Lausanne, Switzerland Poster presentation: "Synthesis of novel Rapalogs with no endogenous target"

**TOR, PI3K and Akt conference**, 2011, Basel, Switzerland Poster presentation: "Development of mTOR12 safe Rapalogs"

**Workshop on "Fluorescent and lipid probe development"**, 2010, San Siro, Italy Oral presentation: "Fast and simple synthesis of carbamate bearing rapalogs"

---

**Teaching experience**


---

03.09.12-12.05.13	Supervision of a Master thesis
06.05.12-02.09.12	Supervision of three Master internships
22.03.05-15.11.14	Private lessons in Chemistry

---

**Languages**


---

German	Native
Spanish	Native
English	Good knowledge in writing and speaking
French	Good knowledge in writing and speaking

## References

---

### IT Knowledge

---

MS Office	Very good knowledge
ChemOffice	Very good knowledge
Maestro(Schrödingier)	Very good knowledge
PyMol	Good knowledge
Latex	Good knowledge

### Other activities

---

2010-Now	Member of FC Telegraph
2012-Now	Captain of the first team of FC Telegraph
2012-Now	Event manager of FC Telegraph
2005-Now	Military Service - NBC Defense

### References

---

Prof. Matthias P. Wymann  
Department of Biomedicine  
University of Basel  
Mattenstrasse 28  
CH - 4058 Basel  
+41 (0)61 695 30 46  
matthias.wymann@unibas.ch

Prof. Thomas Ward  
Department of Chemistry  
University of Basel  
Spitalstrasse 51  
CH - 4056 Basel  
+41 (0)61 267 0 04  
thomas.ward@unibas.ch