

Novel photocleavable intracellular  
heterodimerizer to manipulate protein  
dynamics with high spatiotemporal precision

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Mirjam Sarah Zimmermann

aus Brugg AG, Switzerland

Basel, 2015

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel [edoc.unibas.ch](http://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/](http://creativecommons.org/licenses/by-nc-nd/3.0/ch/) eingesehen werden.

# Table of Contents

SUMMARY	1
1. INTRODUCTION	3
1.1 Manipulation of protein dynamics with CIDs	4
Indirect control of protein functions	5
CID-induced control of subcellular protein localization	
1.2 Most popular CID is based on rapamycin	9
Cellular effect of rapamycin and rapamycin analogs (rapalogs)	10
1.3 Alternative CIDs to control protein dynamics	11
Development of HaXS CID	10
1.4 Tools with increased spatiotemporal precision	12
Classic photo-caging approach to control protein dynamics	13
Genetically-encoded light-sensitive modules	14
Optogenetic dimerizer systems	14
1.5 Steps towards an optimal dimerizer tool	14
Intracellular manipulation of the chemical dimerizer molecule	15
Reversibility of induced dimerization	15
2. AIM OF THE THESIS	20
3. RESULTS	22
3.1 HaXS manuscript	23
3.2 MeNV-HaXS manuscript	47
3.3 Protocol Manuscript (in preparation)	64
3.4 pcRap manuscript (in preparation)	111
Comments to pcRap manuscript	119
4. MATERIALS & METHODS	123
5. DISCUSSION	136
New class of HaXS dimerizers	136
Light as regulatory trigger to increase precision	137
Modular synthetic strategy of HaXS	138
MeNV-HaXS unifies advantageous features of many CIDs	138
Simulation of HaXS-induced dimerization reactions	140
Expand toolbox of available CIDs	140
REFERENCES	142
ACKNOWLEDGMENTS	145
CURRICULUM VITAE	146

## Summary

Cells generate a large repertoire of signaling pathways out of a limited set of proteins through their ability to activate specific signaling proteins in a subcellular region for a short amount of time. A common mechanism to activate signaling pathways represents the translocation of proteins from the cytosol to internal membranes or the inner leaflet of the plasma membrane, where they meet their substrates and initiate downstream signaling pathways.

Chemical inducers of dimerization (CIDs) are powerful tools to manipulate proteins in a spatially and temporally confined region within the cell and to reproduce protein-protein and protein-membrane interactions, which trigger the activation of signaling pathways. These so-called “chemical dimerizers” are small organic molecules, which bind simultaneously two specific dimerizing domains. In the presence of the dimerizer, two proteins fused to these dimerizing domains are brought into close proximity, which initiates a cellular effect linked to the interaction of the protein partners.

We developed a new class of CIDs, based on an intracellular, covalent Halo- and SNAP-tag reactive dimerizer, called HaXS<sup>[1]</sup>. Its modular synthetic strategy enables the relatively simple introduction of novel functional groups into the core module linking the Halo- and the SNAP-tag substrates in order to generate HaXS derivatives with novel features. Through the introduction of a photocleavable methylnitroveratryl (MeNV) group, we developed the first, cell-permeable photocleavable heterodimerizer, called MeNV-HaXS<sup>[2]</sup>. Excitation at 360 nm cleaves MeNV-HaXS and reverses the MeNV-HaXS-induced SNAP- and HaloTag dimer complex.

HaXS unifies several important features, which are essential to mimic physiologic signaling pathways, such as the fast and selective induction of protein-protein interactions, the absent interferences with endogenous signaling pathways as well as the possibility to reverse an induced dimerization event. HaXS was successfully used to target tagged proteins to selected intracellular organelles such as endosomes, lysosomes, the plasma membrane, mitochondria, the nucleus and the actin skeleton, which can be exploited to study the function of a particular protein in different subcellular contexts. Furthermore, the manipulation of protein localizations can be used as a strategy to initiate signaling pathways at defined starting points and cellular locations. Through the HaXS-induced control of signaling protein localizations, which function upstream of the targeted signaling pathway, HaXS is able to selectively control the activation resp. inactivation of an isolated signal transduction branches in a complex signaling network. Additionally, the combination of chemical-induction and light-induced reversion of MeNV-HaXS-induced dimers enables to inactivate any protein of interest through sequestering it away from its functional compartment, followed by an optically guided cleavage of the dimer complex, which releases the sequestered proteins and restores its function. Since the release of anchored proteins occurs with high spatiotemporal precision ( $t <$

1 sec, from single vesicles), this simple experimental setup enables to study translocation kinetics of trapped proteins back to their normal localization in live cells.

To get further insights into the dimerization behavior of the HaXS CID or CIDs in general, we analyzed the chemical induced dimerization reactions with a modeling software called CellDesigner. We demonstrated how various parameters of a CID, such as the ratio of the rate constants of the dimerizing tags or the choice of the dimerizer concentration affects the speed as well as the efficiency of the dimer formation. This allows one to perform efficient dimerization experiments and to understand how parameters of a CID can be optimized to improve its dimerization performance.

The most widely used CID system is based on rapamycin, which induces a tight binding between FKBP12 and the FKBP rapamycin domain (FRB) of mammalian target of rapamycin (mTORC1). The rapamycin CID profits from excellent kinetics, but the cross reactions with mTORC1 diminishes the utility to study cellular events involved in cell growth and metabolism. Analogs of rapamycin (so-called rapalogs), which only react with an engineered version of the FRB domain but not with the endogenous FRB domain of mTORC1, are developed to overcome the immunosuppressive properties of rapamycin while retaining their dimerization ability. The synthesis of rapalogs is challenging, since already minute rapamycin or rapamycin by-products impurities are sufficient to inhibit mTORC1. We successfully established a new protocol for the synthesis and purification of a C16 phenyl carbamate (pcRap) rapalog, which induces dimerization of FKBP and FRB fusion proteins without interfering with the mTORC1 pathway.

Summing up, we provide three valid dimerizer molecules (HaXS8, MeNV-HaXS, pcRap) to the current toolbox of CIDs. Many important features necessary to reproduce physiological signaling pathways are unified in our CIDs and will enable to dissect many cellular signaling events. Furthermore, the combination of a HaXS CID with a rapamycin or rapalog based CID, enables the simultaneous and orthogonal control of two different proteins within a single cell and thus greatly improves the possibilities for cellular interrogations.



## 1. Introduction

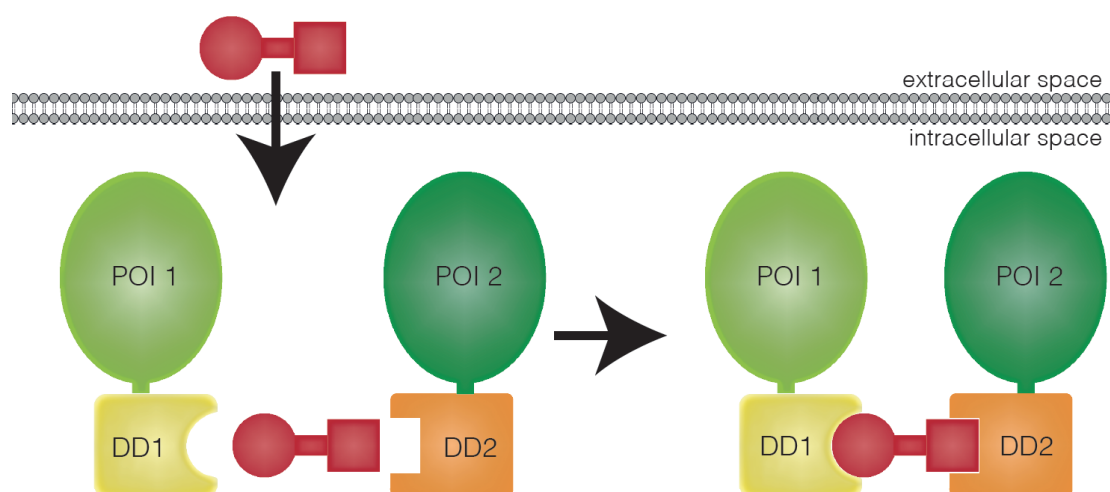
Cells are composed of subcellular compartments and organelles, which together comprise a large dynamic system. Within these highly complex structures, cells organize synthesis, transport and activation of proteins. Protein translocation from one compartment to another allows different cellular compartments to communicate and exchange material, which is essential for the regulation of cellular functions. Furthermore, the biological activity of proteins is often restricted to a subcellular compartment and depending on their intracellular localization, proteins can have different functions. Cells are able to generate a large repertoire of signaling pathways out of a limited set of signaling proteins through their ability to activate specific signaling proteins in a subcellular region for a limited time.

Overall, biological systems depend on precise spatial and temporal coordination of cellular events. The ability to precisely activate or inhibit a selected protein is essential to dissect complex signaling networks and the understanding of biological systems depends on tools that enable to manipulate a cellular process and to analyze the linked phenotypic response. Tools with the ability to specifically manipulate single protein activities with high spatiotemporal precisions and minimal off-target perturbations are required to study signal transduction pathways. Furthermore, the experimental perturbation should be easy to implement, predictable in its behavior and highly specific.

Perturbation of the function of a selected protein can in principle occur at three levels: at the DNA level (e.g. through knockouts or overexpression of constitutively active or inactive mutants), at the RNA level (e.g. through RNAi) or directly at the protein level. Strategies to indirectly target the function of a protein via manipulating the underlying DNA or RNA, profit from high selectivity but suffer from a low temporal resolution as well as from potential compensation effects. Additionally, some mutations can cause lethality early in development, which hampers studies in adult organisms<sup>[3]</sup>. In contrast, chemical biological strategies to directly target proteins at posttranslational level occur with high spatiotemporal precision. The use of small, organic chemicals offers many advantages, as these chemicals can be used in a rapid, dose dependent and even reversible manner. However, the use of inhibitors or activators is mainly limited by the availability of specific molecules for the targeted protein, as most proteins do not have high-affinity small-molecule binding partners. A good alternative to directly modulate protein activities in a more modular approach is based on chemical inducers of dimerization (CIDs).

## 1.1 Manipulation of protein-protein interactions through chemical inducers of dimerization

Schreiber, Crabtree and co-workers introduced the concept of CIDs more than 20 years ago. CID systems are based on small, organic molecules, which have the ability to bind two specific dimerizing domains simultaneously. In the presence of the dimerizer, two proteins of interest (POI), each fused to one of the two dimerizing domains are brought into close proximity. The induced proximity of two POIs results in an increased effective concentration of both proteins, which can physiologically interact and induce the biological effect linked to this interaction (Fig 1).

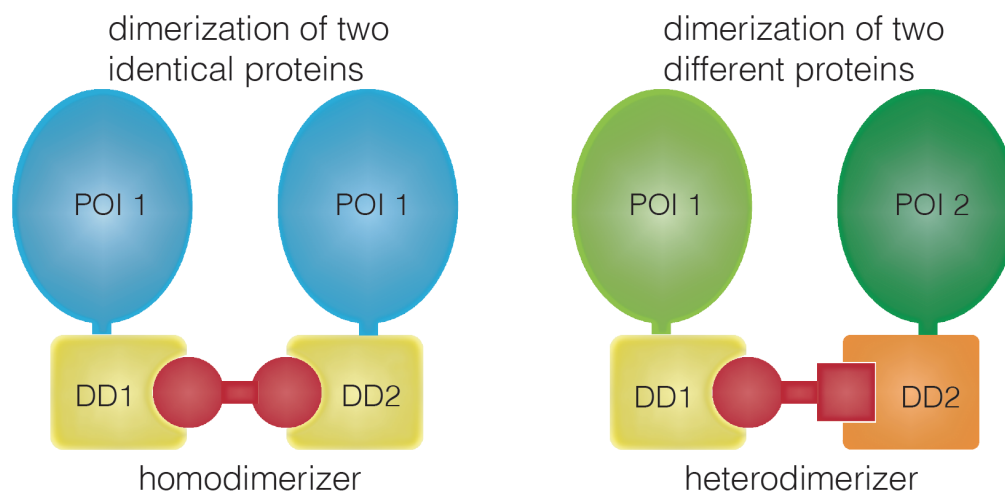


**Figure 1. Scheme describing the concept of CID.** A small, organic molecule (shown in red) enters the cell, where it specifically interacts with two dimerizing domains 1 and 2 (DD1 and DD2, shown in yellow and orange), which induces dimerization of two proteins of interests 1 and 2 (POIs) (shown in light and dark green) fused to the dimerizing domains. Induced proximity of the two POIs allows them to induce a biological effect.

The nature of the dimerizer can vary from naturally occurring molecules that simultaneously bind two proteins (such as coumermycin<sup>[4]</sup>, fusicoccin<sup>[5]</sup> or rapamycin) to naturally occurring molecules that induce a conformational change on a first protein, which then creates a binding surface for a second protein (such as S-(+)-abscisic acid<sup>[6]</sup> or gibberellin<sup>[7]</sup>). The majority of CIDs are based on synthetic, bi-functional molecules that consist of two protein substrates joined through a flexible linker (such as dexamethasone (DEX) conjugated to methotrexate (MTX) (DEX-MTX<sup>[8]</sup>) or to trimethoprim (TMP) (DEX-TMP<sup>[9]</sup>), TMP linked to a synthetic ligand of FKBP (SLF) (TMP-SLF<sup>[10]</sup>), benzylguanine (BG) linked to MTX (BG-MTX<sup>[11]</sup>) and many more).

In their original work, Schreiber and colleagues demonstrate how a small, synthetic molecule FK1012 could reproduce the ability of natural systems to use proximity as a way for activation<sup>[12]</sup>. Although FK1012 was often used to chemically induce proximity of proteins, its application is restricted to proteins whose function depends on self-dimerization, such as self-association of plasma membrane receptors in the absence of extracellular ligands, which leads to the initiation of intracellular signaling pathways<sup>[12]</sup>. Since this original work in the early nineties, many novel dimerizer molecules have been developed (see Introduction, Chapter

1.3). In contrast to the homodimerizer FK1012, many CID systems are based on heterodimerizers, which recognize two different dimerizing domains and thus enables dimerization of two different proteins (Fig 2).



**Figure 2. Scheme describing concept of a homodimerizer and heterodimerizer.** Homodimerizers recognize two identical dimerizing domains (DD1 and DD2) and induces dimerization of identical proteins, whereas heterodimerizers recognize two different dimerizing domains (DD1 and DD2) and induce dimerization of two different POIs.

CIDs can be used to indirectly or directly control the activities of selected proteins. In the first approach, CIDs are used to indirectly control protein activities through the control of the abundance of a selected protein in the cell, either through inducing their expression or degradation. In the second approach, the function of a protein is directly affected either through an induced protein-protein interaction or the induced translocation of the targeted protein.

Summing up, CID systems are versatile tools that can be used to control a wide range of cellular events, including transcriptional activation<sup>[13]</sup>, pre-RNA splicing<sup>[14]</sup>, translation initiation<sup>[15]</sup>, post-translational modification such as glycosylation<sup>[16][17]</sup>, induction of programmed cell death<sup>[18]</sup> or. However, since the clustering of cell surface receptors upon binding to extracellular growth factors and the subsequent recruitment of intracellular signaling proteins, which transfer information's through a biological system, together with many more signaling events are controlled through dynamic protein-protein interactions, the CID strategy is of particular importance for investigations of signal transduction pathways as was e.g. shown in<sup>[1][19][20]</sup>.

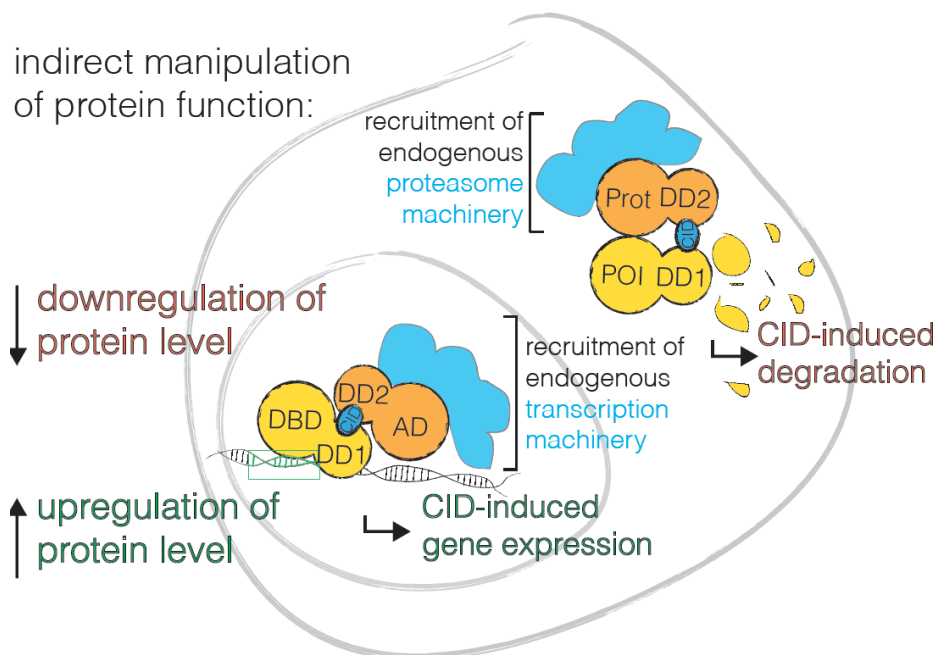
Indirect control of protein function through manipulating the protein abundance

An important application of a CID system is the induced control of the transcription of a specific gene. This is achieved if a transcription factor is split into two domains, the DNA binding domain (DBD) and the activation domain (AD). The DBD is fused to one of the dimerizing domains, while the AD of the same transcription factor is fused to the second

dimerizing domain. The presence of the CID results in the reconstitution of the transcription factor, resulting in the activation of genes under control of this specific transcription factor (Fig 2)<sup>[17]</sup> <sup>[21]</sup>. CID-induced control of gene expression provides an elegant way to control protein levels. However, since transcription takes hours to occur, the temporal resolution of this strategy is limited.

An alternative approach to control protein abundance in cells is the CID-induced degradation of a selected protein, so-called “chemical knockouts”. Most conditional protein degradation technologies take advantage of the ubiquitin-proteasome pathway, which is involved in the regulated proteolysis of intracellular proteins. The addition of several ubiquitin molecules is sufficient to target a protein to the proteasome, where it is degraded. CID-induced proximity of a POI to the E3 ubiquitin ligase causes ubiquitylation of the POI, followed by its targeting to the proteasome and degradation. Alternatively, CIDs can be used to directly target a POI to a subunit of the proteasome, where it is degraded, obviating the need for ubiquitylation. The Janse lab nicely demonstrated how the CID-induced dimerization of a POI and a component of the 26S proteasome (Rpn10) results in proximity of the POI with the proteasome, which is sufficient to induce its degradation<sup>[22]</sup>. In contrast to the control of protein expression, CID-induced degradation benefits from a higher temporal resolution. Unfortunately, this strategy is limited to proteins with access to the proteasome. Furthermore, the general robustness of this approach is limited and extensive construct optimization to achieve an efficient incorporation of the target protein into the degradation machinery is required<sup>[23]</sup>. In contrast to control protein abundance through the activation of the transcription machinery, the approach of CID-induced degradation does only affect the pool of overexpressed proteins fused to the dimerizing domain but not the endogenous protein levels, which could restrict the utility of this approach.

Another strategy for the conditional degradation of proteins is based on small molecules that directly induce degradation of a selected protein and thus enable to directly manipulate the endogenous protein levels. However, these small molecules have to be designed for every single targeted protein, whereas the approach of the CID-induced degradation is more generalizable and profits from the modular approach of a dimerizer system.



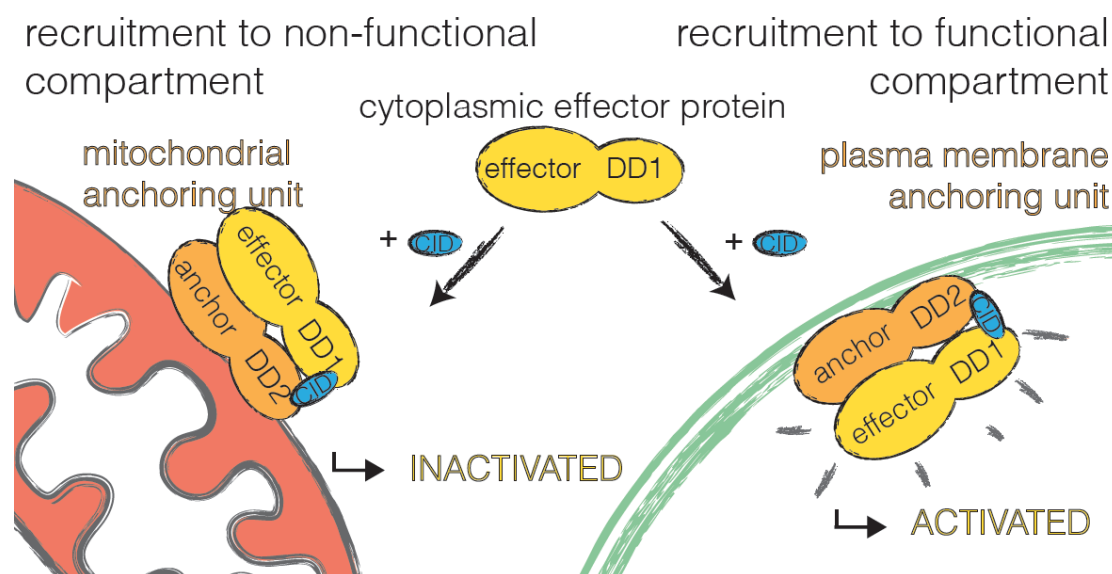
**Figure 2. Scheme describing CID-induced control of intracellular protein abundance.** a) CID-regulated transcription. Dimerizing domain 1 (DD1) is fused to the DNA binding domain (DBD, yellow) and dimerizing domain 2 (DD2) is fused to the activation domain (AD, orange). Addition of the CID initiates translocation of the DBD to the AD and reconstitutes the transcription factor. Recruitment of the endogenous transcription machinery (violet) to this transcription factor leads to the expression of the gene under control of the. b) CID-induced degradation. Dimerizing domain 1 (DD1) is fused to a selected protein (POI, yellow), whereas the dimerizing domain 2 (DD2) is fused to a component of the proteasome machinery (orange), which recruits the endogenous proteasome degradation machinery (violet). Presence of the CID results in the translocation of domain1-POI to the proteasome and induces degradation of the POI.

In contrast to the manipulation of intracellular protein abundance, the CID strategy can be used to directly control the function of a selected protein. This direct approach profits from higher temporal and spatial resolution, which is especially important for the analysis of cellular processes occurring on short time scales, as true for many physiological signal transduction pathways. Direct activation of protein function is achieved through the induced interaction of two proteins that require the presence of each other for their function. A prominent example is the dimerization of transmembrane surface receptors such as fibroblast growth factor receptors 1 and 2 (FGFR 1 and 2)<sup>[24]</sup>. Another strategy relies on proteins that are split into two inactive fragments. If properly engineered, chemical induced dimerization of these two fragments enables the reconstitution of a functional protein<sup>[25]</sup>. Nonetheless, the most common application of a CID is to control the localization of proteins in the cells and to exploit the subcellular localization of proteins as a strategy to control their activities.

#### CID induced control of subcellular localization as a strategy to control protein activity

Since the biological activity of a protein is often restricted to a subcellular compartment, CID-induced translocation of a protein towards (Fig 3, left) or away (Fig 3, right) from its functional locality can be employed to either activate or abolish a cellular event. CID-induced translocation of a selected protein is achieved through co-expression of an anchor unit and an effector unit. The anchor unit consists of one of the two dimerizing domains, which is anchored at a specific cellular localization, while the cytosolic effector unit contains the

second dimerizing domain fused to the POI. The presence of the dimerizer results in the translocation of the cytosolic POI to the anchor unit. Depending on whether the targeted subcellular compartment is a functional or a non-functional location for the POI, this approach can be employed as a localization-induced activation and inactivation of a selected protein. CID-induced activation of a POI through its translocation toward its functional compartment has been used to investigate various cellular events, such as intracellular calcium and cAMP signaling<sup>[26]</sup> as well as GPCR internalization<sup>[27]</sup>.



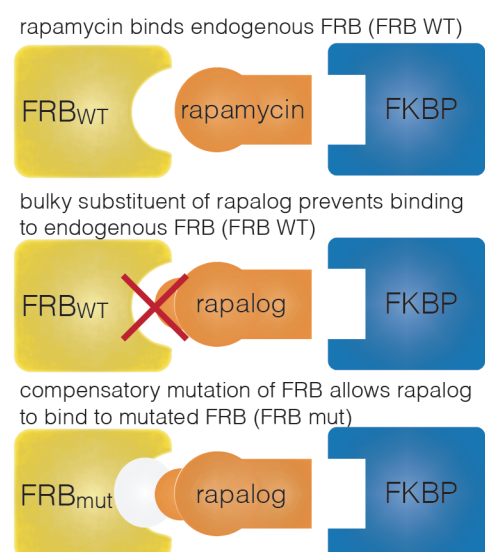
**Figure 3. CID controlled subcellular localization of proteins as a strategy to control protein activities.** CID induces the translocation of previously cytosolic effector protein fused to dimerizing domain 1 (DD1, shown in yellow) to the mitochondrial anchor unit (non-functional compartment, orange) composed of dimerizing domain 2 (DD2) fused to the mitochondrial anchoring sequence or to plasma membrane anchor unit (functional compartment) composed of the dimerizing domain 2 (DD2) fused to the plasma membrane targeting sequence. Translocation to a functional versus a non-functional compartment results in activation versus inactivation of the effector protein.

In the reverse approach, CID-induced translocation of a selected protein away from its functional location results in the inactivation of the selected protein. To keep it away from its functional compartment, the targeted protein is sequestered in a non-functional compartment. This conditional protein inactivation strategy occurs with high temporal precision provides an elegant way to dissect complex signaling networks. CID-induced inactivation of effector proteins through their mislocalization has been nicely demonstrated in yeast<sup>[28]</sup>. In this so-called “anchor-away” technique, rapamycin was used to rapidly inactivate FRB-tagged nuclear proteins, through their removal from their functional place and sequestration in the cytoplasm through a FKBP-tagged anchor unit. Robinson and colleagues established a similar approach in mammalian cells, called “knock-sideway”. They demonstrated the rapid depletion (on a time-scale of seconds to minutes) of adaptor proteins from FKBP-tagged clathrin-coated vesicles by anchoring them via mitochondrial-tagged FRB domains to the mitochondrial surface in mammalian cells<sup>[29]</sup>.

An alternative approach to rapidly induce inactivation of selected proteins through the sequestration of proteins away from their functional compartment is based on multimerizing and clustering of selected proteins, prevents them to work in their functional compartment and results in the reduced activity of these proteins<sup>[23]</sup>.

## 1.2 Most popular chemical inducer of dimerization is rapamycin

The naturally occurring immunosuppressant rapamycin represents the most thoroughly studied and implemented chemical dimerizer molecule. Rapamycin mediates heterodimerization of the chaperone FKBP12 and the FRB domain of mTORC1. In this trimeric complex, rapamycin acts as dimerizer, and brings two POIs genetically fused to FKBP12 and FRB, respectively, in close proximity<sup>[30]</sup> [21]. As true for most naturally occurring compounds, rapamycin display excellent biophysical properties, such as high cell permeability and good water solubility. Additionally, rapamycin-induced dimerization of FKBP12 and FRB fusion proteins occurs with excellent kinetics<sup>[31]</sup>. Although rapamycin CID is a useful tool, which has been extensively used to study many signaling pathways, especially those involving lipid metabolizing enzymes<sup>[32]</sup> or small GTPases<sup>[33]</sup>, it has to be considered that FKBP and FRB fusions proteins compete with endogenous FKBP12 and mTORC1 for rapamycin binding. Binding of FKBP12-rapamycin to the FRB domain of mTORC1, renders the mTORC1 complex enzymatically inactive. As mTORC1 is an important protein kinase involved in different signaling pathways controlling cell growth, proliferation as well as in the regulation of autophagy, the use of rapamycin may have unwanted side effects on the signaling network of the targeted cell and will eventually lead to cell-cycle arrest, which complicates *in vivo* investigations. However, as the experimental time is often shorter than the time of rapamycin to induce its potential toxic effect, the use of the rapamycin CID in cultured cells can still be justified. However, if longer experimental times are required or work in living animals is performed, it is advisable to use non-toxic rapamycin derivatives, known as rapalogs.



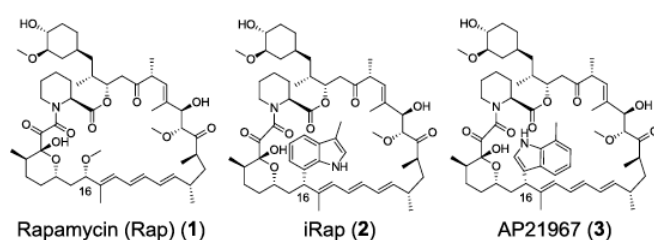
**Figure 4. Design of rapalogs using the “bump and hole” approach.** Rapamycin binds to the endogenous FRB domain (FRB WT). The part of rapamycin that interacts with the FRB domain is equipped with a bulky substituent. This abolishes binding of the rapalog with the endogenous FRB domain. In order to restore binding capacity of the rapalog with the FRB domain, the heterologously expressed FRB domain contains a compensatory mutation (FRB mut) that enlarges its rapamycin-binding pocket to accommodate the substituent.

Effect of rapamycin and rapamycin analogs (rapalogs) on cell function

In order to minimize the problem with potential off-target effects of rapamycin, non-toxic rapamycin analogues (so-called rapalogs) have been designed using the ‘bump-and-hole’ approach<sup>[34] [35]</sup>.

The analysis of the crystal structure of the trimeric FKBP12-rapamycin-FRB complex revealed that the C16-methoxy group of the rapamycin core directly points toward the FRB domain<sup>[36]</sup>. Conclusively, derivatization of this methoxy group with bulkier substituents can alter the FRB-binding affinity<sup>[37]</sup> and most currently available rapalogs are derivatized at this C16 position. The introduction of a bulky substituent at this position abolishes binding to the wild type FRB domain. To restore the dimerizing potency of this dubbed rapamycin, rapalogs are used in combination with an engineered version of FRB in which threonine at position 2098 is substituted by a leucine (T2098L). This mutation displays an enlarged rapamycin-binding pocket and this accommodates binding of the rapalogs with bulky substituents at the C16 position (Fig 4). Consequently, these rapalogs should be devoid of inhibitory effects towards mTORC1, but still be able to dimerize fusions of FKBP12 and the mutated FRB domain (T2098L)<sup>[38]</sup>. The successful use of several rapalogs such as iRap and AP21967 (Fig 5, [2] and [3]) are reported<sup>[39]</sup>. The effect on mTORC1 kinase activity<sup>[39]</sup> by these two rapalogs was evaluated through the determination of the IC<sub>50</sub> values in HEK293 cells. Surprisingly, the results revealed that both rapalogs still inhibit mTORC1, either caused by impurities of rapamycin or rapamycin-byproducts, or an undefined spatial orientation of the bulky group at the C16 position (IC<sub>50</sub> = 0.1 nM for rapamycin, IC<sub>50</sub> = 5 nM for iRap and IC<sub>50</sub> = 10 nM for AP21967)<sup>[40]</sup>. Further analysis by quantitative mass spectrometry revealed around 2.6% rapamycin contamination in the rapalog sample iRap<sup>[40]</sup>. These findings indicate that the synthesis and purification of rapalogs is very challenging, and already minor contaminations of rapamycin and rapamycin-byproducts (<<1%) dramatically affect the phosphoinositide 3-kinase (PI3K)/mTOR axis.

As true for all naturally occurring dimerizer molecules, derivatization of the rapamycin core can dramatically lower its cell permeability. The reduced cell permeability of the commercially available rapalog AP21967 (Clonetechn) requires the use of 25 times higher concentrations of this rapalog than of rapamycin in order to induce comparable dimerization effects<sup>[41]</sup>. Furthermore, only limited positions of the rapamycin core can be functionalized (C16, C28, C40), which makes the generation of rapamycin derivatives with novel features challenging.



**Figure 5. Structure of rapamycin and two commonly used rapalogs.** Structure of Rapamycin (Rap) [1], C16-(S)-3-methylindole-rapamycin (iRap) [2] and C16-(S)-7-methylindole-rapamycin (AP21967) [3].



Besides the challenge with the mTORC1 interference, the binding of rapamycin or rapalogs to members of the FKBP family has to be considered as well. FKBP are widely expressed immunophilins, which are involved in mediating an immunosuppressive effect. It has been shown that FKBP12 ligands such as FK506 or rapamycin interfere with the cellular function of FKBP12 proteins and for example interfere with proper regulation of the calcium release<sup>[42]</sup>. Additionally, the nonproductive interactions between rapamycin and abundant endogenous mammalian FKBP can sequester the added rapamycin or rapalog, and thus attenuate the dimerization performance of the rapamycin or rapalog CID, especially in cells with low endogenous levels of FKBP<sup>[43]</sup>.

### 1.3 Alternative chemical inducers of dimerization tools to control protein-protein interactions

The choice of an experimental strategy is generally a compromise between experimental requirements and available tools. Features such as specificity, robustness, reversibility, orthogonality and the control of cellular events with high temporal and spatial precision within the physiological context have to be counterbalanced with features such as simple handling and adaption to novel applications as well as low costs and the available equipment in the lab. Additionally, minimization of any off-targets is of high importance for all research tools used to perturb cellular processes.

The characterization of new protein-ligand pairs has led to the development of a number of alternative CID systems that have been extensively used to modulate protein-protein interactions in living cells. An overview of properties of dimerizer molecules of alternative homo- and heterodimerizers is summarized in Table 1. For a detailed description of currently available CIDs and their applications I refer to following reviews<sup>[44]</sup>.

Compound	Biochemical monitoring of complex	Cell permeant	Intracellular dimerization*			Estimated speed of dimerization	Interference with endogenous signaling pathways	Orthogonal use / applications
			y/n	mode	type			
Abcisic acid (ABA)	no	yes	yes	forced	hetero; non-covalent	slow (<h), high concentrations required	None reported, but toxic effects <i>in vivo</i> (Celik et al., 2007).	Compatible with rapalogs; regulation of coupled transcription processes (Liang et al., 2011; Hathaway et al., 2012).
CoDis (BG-BG)	yes	(yes)	yes	pre-ass.	homo; covalent	slow [h]	None reported. Tag substrates not expected to interfere with cellular signaling.	Detection of existing protein-protein interactions (Lemercier et al., 2007).
Coumemycin	no	yes	yes	forced	homo; non-covalent	fast [min]	None reported.	Translocation and activation of signaling molecules (Farrar et al., 1996; Farrar et al., 2000).
Dex-FK506	no	yes	yes	forced	hetero; non-covalent	very slow [days ?]	Not assessed in detail. Likely with nuclear receptors and FKBP.	Yeast three hybrid system (Licitra and Liu, 1996).
Dexamethasone -Methotrexate (Dex-Mtx)	no	yes	yes	forced	hetero; non-covalent	very slow [days ?]	Not assessed in detail. Likely with nuclear receptors.	Transcription reporter systems and yeast complementation screens (Lin et al., 2000; Baker et al., 2002).
Dexamethasone -Trimethoprim (Dex-TMP)	no	yes	yes	forced	hetero; non-covalent	very slow [days ?]	Not assessed in detail. Likely with nuclear receptors.	Transcription factor logics in a three hybrid system (Gallagher et al., 2007; Bronson et al., 2008).
Gibberellic acid derivatives (e.g. GA3-AM)	no	yes	yes	forced	hetero; non-covalent	very fast [s]	Cellular acidification at elevated concentrations (Celik et al., 2007).	Compatible with rapalog system; control of Jak/Stat signaling (Mohi et al., 1998; O'Farrell et al., 1998), and protein translocations (Miyamoto et al., 2012; Phua et al., 2012).
HaXS8	yes	yes	yes	forced	hetero; covalent	fast [min]	None observed. Tag substrates not expected to interfere with cellular signaling.	Forced protein translocation, activation of signaling pathways, orthogonal and multiplexed applications, compatible with rapalog system. Simple assessment of success of dimerization (Erhart et al., this work).
Rapalogs	no	yes	yes	forced	hetero; non-covalent	very fast [s]	Interference with mTOR signaling, and FKBP-dependent pathways. Proteins with low level rapamycin contamination.	Rapid translocation of signaling molecules to membranes, successful depletion of lipid pools (Libertes et al., 1997; Inoue et al., 2005; Varmai et al., 2006; Edwards and Wandless, 2007). Orthogonal applications possible (Bayle et al., 2006; Inoue and Meyer, 2008). For a review see (Putyrski and Schultz, 2012).
S-Cross (e.g. SC-Cy5)	yes	no	no	pre-ass.	hetero; covalent	slow [h]	Not expected, tag substrates not expected to interfere with cellular signaling.	Detection of protein-protein interactions in cell lysates; integration of fluorescent tracer (Gautier et al., 2009).
X-CrAsH	yes	yes	yes	pre-ass.	(homo); covalent	slow [h]	Redox state-dependent accumulation of FlAsH derivatives in mitochondria reported (Langhorst et al., 2006).	Dimerisation of pre-associated partners (Rutkowska et al., 2011).

\* forced intracellular dimerization triggers complex formation of two previously non-associated protein partners, while pre-ass. points to dimerization of preformed complexes. Complex formation can be mediated by symmetric interactions (homo-dimerization) or asymmetric interactions (hetero-dimerization), and can be mediated by non-covalent binding or covalent reactions.

**Table 1. Properties of alternative CID systems.** From Erhart et al., 2013<sup>[1]</sup>, Supplementary Figure S5.

Despite the great utility of these CIDs, none of them is without limitation. Either an expensive and complex multistep synthesis of the dimerizer itself or limited application in living animals due to unspecific binding profiles and/or affinities for endogenous proteins that lead to cytotoxic effects, restrict the use of some these CID systems to a number of limited applications. For example, the most commonly used CID is based on rapamycin, which is far from ideal due to its immunosuppressive properties. But its expediency and availability still make it the most often implemented CID.

In summary, each CID offers its own advantages and has been involved in solving specific problems in cell biology and related fields. However, none of these systems unifies all essential features of an optimal research tool suitable to efficiently investigate physiological signaling pathways, such as the possibility to reversibly induce intracellular dimerization with excellent kinetics and without interfering with endogenous signaling pathways.

#### Chemical development of an intracellular heterodimerizer: HaXS CID

We developed an alternative CID system in our lab, called HaXS. HaXS consists of a core module, which links two highly specific self-labeling tags, the HaloTag- and SNAP-tag substrate. An iterative approach in which the core module was chemically modified succeeded in optimizing HaXS for high cell permeability and fast reactivity. In contrast to other CIDs such as the most commonly used rapamycin CID, HaXS does not interfere with endogenous signaling pathways. In contrast to the covalent CID based on S-CROSS, which only dimerizes previously interacting proteins<sup>[45]</sup>, the HaXS CID is the first covalent dimerizer system, which can force proximity of previously non-interacting proteins. As true for all CIDs, the cloning effort is high and constructs containing the dimerizing domains need to be optimized for every single application. Since the formation of a covalent dimer complex allows simple quantification of the dimerization efficiency under denaturing conditions by immune-blotting, the process of construct design and optimization is greatly facilitated. Additionally, the straightforward analysis of the dimerization efficiency with immune-blotting experiments, enables the direct correlation of the induced amount of dimerization with the cellular output. Furthermore, the HaXS system can be used orthogonally with other CIDs, which allows the combined use of the HaXS CID with one or even more CIDs in a single cell. This enables to simultaneously control various signaling molecules at several locations in a single cell and thus to investigate more sophisticated signaling networks. A detailed description of the development and proof-of-concept experiments demonstrating the activation of a PI3K as well as a proof-of-principle application of the orthogonal use of the HaXS CID with the rapamycin CID can be found in the Results (see Results, Chapter 3.1, HaXS8 manuscript).

In contrast to naturally occurring dimerizer molecules, which are difficult to derivatize as already small changes of these naturally optimized molecules can have dramatic effects on their cell penetration performances, the core module of HaXS can be relatively easily

modified. This makes the HaXS dimerizer a good starting molecule for the development of further classes of dimerizers related to HaXS.

#### 1.4 Manipulation of protein dynamics with increased spatiotemporal precision

The HaXS CID and other CID systems were successfully used to regulate a number of protein-protein interactions on a time scale of minutes to hours. However, as many signaling pathways in living cells typically occur on a timescale of seconds, it is impossible to mimic physiological signaling pathways with tools based on chemical molecules. The time required for the chemical dimerizer to penetrate the cell membrane limits their temporal resolution, whereas the spatial resolution is limited by the intracellular diffusion of the chemical dimerizer. In order to improve the spatiotemporal precision of methods involved in the control of protein activities, recent efforts lie in the development of tools, which use light instead of a chemical molecule as the regulatory trigger. In contrast to chemical molecules, delivery of light is immediate, which enable the manipulation of protein activities with high temporal resolution. Additionally, light can be applied on single cells or even subcellular regions of a cell and thus allow manipulating molecular events not only with high temporal but also with high spatial precision. Furthermore, light has few off targets and relatively low toxicity. The intensity of the light can be modulated, thus the intensity of intracellular events activated by light can be varied. These advantageous features of light as a stimulus for biological interrogation raised great interest in the development and use of light-inducible tools in the last years.

##### Classic photo-caging approach to control protein dynamics

In the field of chemical biology, the concept of regulating biological processes with light was introduced in the late seventies by the groups of Schläger and Hoffman. In a classic photo-caging approach, chemical and enzymatic methods such as *in vitro* translation<sup>[46]</sup> or chemical ligation<sup>[47]</sup> are used for the site-specific incorporation of a caging group *in vitro* (reviewed in Mayer and Heckel, 2006<sup>[48]</sup>). The derivatization of proteins with these photolabile groups, keeps them in an inactive status until photo-excitation releases the caging group, which results in the formation of the “uncaged” biologically active compound<sup>[49]</sup>. However, the spatial resolution of light-induced activation is limited, since the photo-released module is freely diffusible in the cell. The irreversible nature of the photo-uncaging reaction further limits this approach to few applications (reviewed in Binschik et al., 2011<sup>[50]</sup>). Additionally, photocaging is limited to specific targets and a time-consuming case-by-case optimization is required for each protein that is brought under light control. Additionally, since the chemical modification occurs *ex vivo*, the caged proteins have to be re-introduced into cells (for example via permeabilized cells<sup>[51]</sup>, via microinjection and electroporation<sup>[52]</sup>), which can be challenging and prevents the application of these approaches in living cells.

## Genetically-encoded light-sensitive modules to control protein dynamics

The combination of the specificity resulting from genetic manipulation together with the high spatiotemporal precision of light-induced control is a promising approach for the manipulation of protein dynamics within biological systems. Integration of genetically encoded light-controlled modules, which do not depend on extracellular chemical modifications, is one of the biggest advances in the field of light-controlled tools. The modules contain a naturally occurring photosensitive protein that undergoes a conformational change upon activation by illumination with light at a defined wavelength. One approach to use these genetically encoded light-sensitive protein modules, is to put a single protein under light control by preparing fusion proteins between a single POI and a light-inducible module. This was nicely demonstrated for the light-gated control of the GTPase Rac<sup>[53]</sup>. However, direct fusion of a photo-activable group on a POI either via genetic or chemical incorporation requires time-consuming case-by-case optimization.

## Optogenetic dimerizer systems

In a more modular approach, cellular events are controlled indirectly through recruiting or sequestering POIs via a light-dependent control of protein-protein interactions. This approach obviates the need to optimize the caging of individual proteins. A nice demonstration of this approach is the light-gated localization of a Rac guanine nucleotide exchange factor (GEF) to the membrane, which leads to the activation of Rac and subsequent actin polymerization<sup>[54]</sup>. While these genetically encoded light-activated dimerization systems are versatile tools, which profit from fast dimerization<sup>[55,56]</sup> as well as from reversibility of the dimerization reaction<sup>[54-57]</sup>, they all suffer from their own drawbacks as large photosensory protein tags<sup>[54][57]</sup>, requirement of an exogenous co-factor<sup>[57]</sup>, unprecedented speed of dimerization<sup>[54]</sup>, slow kinetics<sup>[57]</sup>, formation of unwanted homodimers<sup>[55]</sup>, limited localization differences in the illuminated versus dark state as well as sensitivity of accidental exposure to environmental light and a spectral overlap between wavelengths used for activation of a biological process and commonly used fluorescent reporter proteins such as GFP, CFP, YFP<sup>[55,56]</sup>. Additionally, a general restriction to all these light-inducible systems is the need of specialized microscopes and software's, which are necessary to spatially target the activatory light beam<sup>[53]</sup>.

### 1.5 Steps toward an optimal dimerizer tool

Although a wide variety of cellular events were successfully investigated with optogenetic dimerizer systems, the properties of the molecular components and the technologies of light delivery have to be improved to make these tools more accessible. In this context, a tool, which combines light-induction with a well-characterized dimerizer tool, which enables high spatiotemporal precision with a light sensitive tool that is based on a well-characterized chemical induced dimerization reaction will be advantageous.

---

### Intracellular manipulation of the chemical dimerizer molecule

Two strategies for the intracellular manipulation of the dimerizer in the commonly used rapamycin CID have been reported. The first is a photocleavable biotinylated rapamycin analog (cRb-A) that has been used to control the activity of small GTPases<sup>[58]</sup>. The biotin has a very high affinity to the protein streptavidin, which prevents the biotin-streptavidin conjugate from entering the cell. The need for the extracellular removal of the caging group to release rapamycin, which then can diffuse across the membrane, limits the application of this tool to investigations of POIs that are active at or very close to the plasma membrane. Karginov and colleagues present another photocaged rapamycin (pRap), which has to be used in combination with an engineered version of FKBP12 (iFKBP)<sup>[59]</sup>. From their published data it is difficult to conclude whether pRap is able to penetrate cells. Additionally, the possibility to induce dimerization in a subcellular region of the cell was not demonstrated, and the requirement to replace FKBP12 with iFKBP in existing constructs may hinder further applications. Additionally, in both approaches the caging group is located at the C40 position of the rapamycin core, which is the interaction site with FKBP12. Since rapamycin alone cannot bind to the FRB domain and only interacts with FRB if it has previously reacted with FKBP12, the uncaging reaction has to occur before rapamycin can react with a single protein domain. Conclusively, these approaches do not enable the intracellular manipulation of target proteins and rather produce a source of diffusible rapamycin.

Williams and colleagues synthesized a photocleavable homodimeriser based on the FKBP-FK506 CID<sup>[60]</sup>. This system nicely demonstrates the utility of photocleavable dimerizers. However, the relative long irradiation times required to induce photolysis as well as the fact that only homodimerization events can be controlled, limit the application of this CID.

### Reversibility of induced dimerization is required to mimic physiologic signaling events

Various studies have shown the successful CID-induced activation signaling events with high temporal resolution and specificity, but most of the available CIDs are not reversible and thus are not suited for the investigation of processes that are regulated in a reversible manner. However, physiological signaling events such as phosphorylation and de-phosphorylation events as well as molecular switches of small GTPases are often transient by nature. Consequently, tools that enable to reverse an induced status are essential to mimic physiological signaling events.

The dimerization of genetically encoded light-activated dimerization systems enable to control dimerization events in a reversible manner. The dimerization depends on a photoisomerizable chromophore, which undergoes a conformational change upon illumination with light at proper wavelength that relaxes back to its initial conformation in the dark. Thus the association and dissociation of the two dimerizing domains can be repeated, which would in principle enable to simulate oscillating signals, a well-known phenomenon in the phospholipid field<sup>[61]</sup> and many signal transduction events. The induction of the association of the two dimerizing

domains is generally fast for all of these systems, but the reversion of the process is slow and takes several minutes to hours<sup>[54][55-57]</sup>. Only with the system reported by Levskaya and colleagues the dimerization process can be reversed on a timescale of seconds. However, reversion is only achieved in the presence of a co-factor and upon illumination of light with a different wavelength than the one to induce association (750 nm versus 650 nm). The need for two different light sources makes this experimental setup less convenient and thus limits its applications. An overview of the most popular optogenetic dimerizer systems is summarized in Table 2.

	FK1F / GIGANTEA	PHYB / PIF	CIB1 / CRY2	LOV2 / ePDK1
Reference	Yazawa et al. 2009	Levskaya et al. 2009	Kennedy et al. 2010	Strickland et al. 2012
Association				
Principle	Optically switchable dimerizer systems based on conformational changes of LOV domain, blue light (450 nm) induces association of FK1F and GIGANTEA	Optically switchable systems based on conformational changes of phytochromes B, red light (650 nm) illumination to induce association of phytochrome B and PIF	Optically switchable systems based on conformational changes of cryptochrome 2, blue light (458-488 nm) induce dimerization of CIB1 and CRY2	Optically switchable systems based on conformational changes based on LOV, blue light (< 500 nm) induce dimerization of LOV2 and ePDZ1
Kinetics	Slow, 30 min, inefficient (30%)	Fast, seconds	Fast, seconds (100 ms)	Fast, seconds
Dissociation				
Principle	In absence of activatory light: conformational relaxation of LOV domain leads to dissociation of FK1F and GIGANTEA	Illumination with infrared light (750 nm) plus addition of a co-factor (PCB) induces dissociation of phytochrome B and PIF	In absence of activatory light: conformational relaxation of cryptochrome 2 leads to dissociation of CIB 1 and CRY1	In absence of activatory light: conformational relaxation of LOV leads to dissociation of LOV2 and ePDZ1
Kinetics	Slow, hours, almost irreversible	Slow, hours, under 750 nm light: fast, seconds (with addition of co-factor)	Medium, 10 min	Medium, 10 min
Multiple rounds	Yes	Yes	Yes	Yes
Application	Membrane recruitment of Rac induces lamellipodia formation, induction of transcription activation (Gal4 / VP16 )	Membrane recruitment of Rac GEFs results in cell protrusion	Reconstitute activation of split protein fragments (Cre recombinase), activation of Gal4 transcription activator system	Control of yeast mating pathway: recruitment of Ste5 to PM and yeast polarity by Cdc42 recruitment to PM

**Table 2. Overview of optogenetic dimerizer systems.** Properties of association and dissociation of the dimerizing modules of optogenetic dimerizer systems as well as an example of applications are listed.

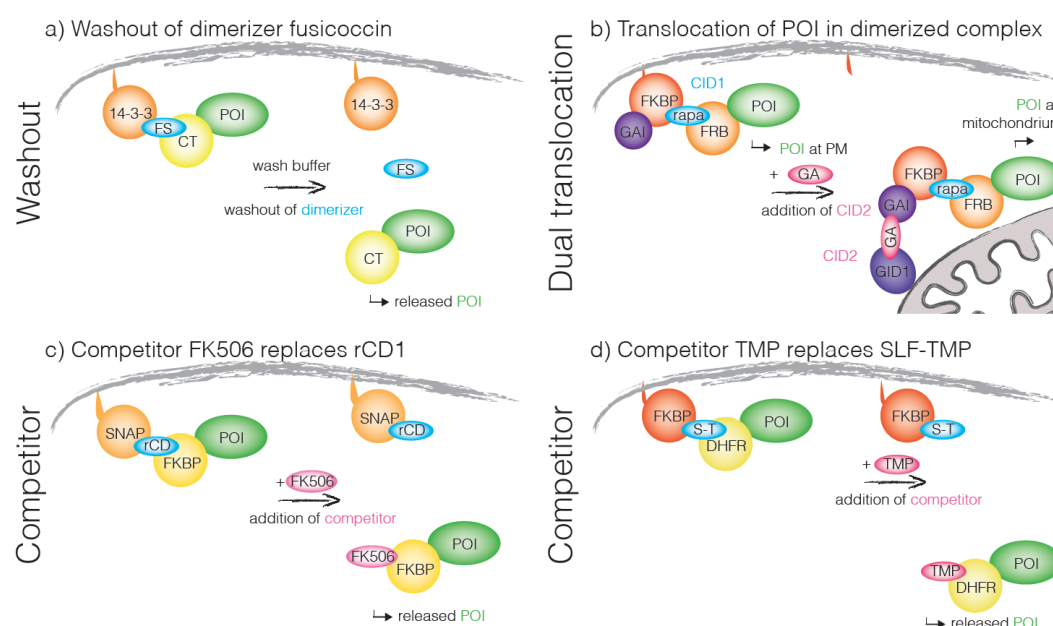
In contrast to these reversible optogenetic dimerizer systems, in most available CID systems the induced dimerization is not reversible, either due to covalent (HaXS CID) or high affinity (rapamycin CID) interactions of the dimerizer with the respective dimerizing domains. Rapamycin interacts with FKBP and FRB with a high affinity (FKBP–rapamycin:  $K_d=0.2$  nM, FKBP–rapamycin–FRB:  $K_d=12$  nM<sup>[31]</sup>, which makes the clearance of rapamycin from cells extremely slow. Neither extensive washout<sup>[62]</sup> nor an excess of a competitor<sup>[63]</sup> could reverse the rapamycin-induced dimerization on a biological relevant timescale.

Last years several strategies were reported to implement reversibility in available CIDs, which enable to use these tools to reversibly control signaling events and thus to mimic physiological signaling pathways. A fungal CID (fusiococcin) dimerizes fusion proteins of two plant-derived domains, 14–3-3 and C terminus (CT) of the H<sup>+</sup>-ATPase PMA2. The induced dimerization was shown to be reversible upon extensive washing steps. However this takes several minutes and thus is not suitable to mimic physiological signaling events<sup>[64]</sup> (Fig 6a). Lin et al reported reversibility of a chemical induced signal through a dual translocation strategy. Two orthogonal CID systems were combined (rapamycin and gibberellin CID) and reversion of an induced activation from the primary CID is achieved through the sequestration of the whole dimerized complex away from its site of action by redirecting the entire complex to another subcellular location. This dual translocation strategy was used to turn Rac/PI3K-dependent membrane ruffling on and off. As both of the used CID systems are irreversible, multiple rounds of recruitment and dissociation are not possible with this setup. Additionally, as certain proteins show activity at various membranes of organelles, this approach requires extensive control experiments to rule out the involvement of possible interfering activations and the sophisticated setup of combining two CID systems requires high effort in construct design and thus seems not to be convenient for many applications<sup>[63]</sup> (Fig 6b).

The group of Carsten Schultz<sup>[65]</sup> introduced a novel reversible CID system. In a first step, the reversible chemical dimerizer (rCD1) dimerizes two standard protein tags, the SNAP-tag and FKBP12. In a second step, the induced dimerization can be reversed through the addition of a commercially available competing ligand (FK506). The induced dimerization can be reversed sufficiently fast to determine kinetics of lipid metabolizing in living cells (half-time of deactivation around 15 s). The induction of PI3K activity at the plasma membrane activity followed by its inactivation was demonstrated. However, even the reversion of the induced dimerization is fast, the induction of the dimerization itself is slow (Fig 6c). Another novel reversible dimerizer system based on the same principle was introduced in the same year. The dimerizer contains a synthetic ligand of FKBP' (SLF') linked to trimethoprim (TMP). The SLF' moiety binds to the F36V mutant of FK506-binding protein (FKBP) and the TMP moiety binds to *E. coli* dihydrofolate reductase (eDHFR). SLF'-TMP-induced heterodimerization can be reversed upon addition of a high concentration (10  $\mu$ M) of the competitor TMP. This setup allows one to control multiple rounds of dimerization and dissociation. Additionally, the CID induced activation cannot only be reversed, but through the combination of the SLF'-TMP CID

with the rapamycin CID, a POI can be translocated between two subcellular sites. If the competitor and rapamycin are added simultaneously, TMP replaces SLF'-TMP and allow rapamycin to induces the translocation of an FKBP'-tagged POI to a second anchoring unit containing the FRB domain (Fig 6d). Valuability of the system was demonstrated by reversibly targeting a constitutively active Rac1 mutant to the plasma membrane in live cells, which resulted in the rapid and reversible formation of lamellipodia. Unfortunately dimerization and reversion of dimerization-induced process is rather slow (> 10 min) and this tool is limited to application occurring on slower time scales<sup>[66]</sup>.

In summary, these strategies to implement reversibility into CID systems have greatly enhanced the toolkit for controlling physiological signaling pathways. However, as all systems described suffer from disadvantages (see Table 3) such as inefficient washout effects, complicated experimental setups, slow dimerization and/or reversion of the induced dimerization, there is a need for a novel tool that allow to reversibly control dimerization events.



**Figure 6. Scheme demonstrating four strategies to implement reversibility in CID systems.** a) Fusicoccin dimerizes plant derived dimerizing domains 14-3-3 and CT. Extensive washout of the dimerizer leads to the reversion of the dimerization. b) Rapamycin induces dimerization between FKBP and FRB. Addition of gibberellin that induces dimerization between GID1 and GAI leads to the translocation of the dimerized complex at the first anchoring unit (PM) toward a second anchoring unit (mitochondrion). c) rCD1 induces dimerization between FKBP and SNAP. Addition of the competitor FK506 replaces rCD1 and reverses dimerization. d) SLF'-TMP dimerizes FKBP and DHFR. Addition of the competitor TMP replaces SLF'-TMP and reverses the dimerization.

In summary, three different strategies were used to implement reversibility to CID systems. The dual translocation strategy requires extensive cloning effort, which can be challenging for some applications. The washout of the dimerizer fusicoccin seems to be inefficient and dissociation proceeds at relatively slow rates, thus limiting this approach to investigations that occur on a slow time scale. More promising strategies are based on competitor molecules, which replace the dimerizer and thus reverse the induced dimerization. Replacement of the dimerizer SLF'-TMP with the competitor TMP takes several minutes, whereas the replacement of rCD1 with the competitor FK506 process within few seconds and thus enables to investigate physiological relevant signal inactivation events. However, as the induction of the dimerization in this CID is slow and takes several minutes ( $t > 20$  min), application of rCD1 CID is limited.



	Fusicoccin	Dual translocation	rCD1	SLF'-TMP
Reference	Skwarczynska et al. 2013	Lin et al. 2013	Feng et al. 2014	Liu et al. 2014
Association				
Principle	CID, fusicoccin induces dimerization of 14-3-3 and CT (plant derived, engineered)	Combination of 2 orthogonal CIDs, rapamycin induces dimerization of FKBP and FRB	CID, rCD1 induces dimerization of FKBP and SNAP-tag	CID, SLF'-TMP induces dimerization of FKBP' and DHFR
Kinetics	Slow, 20 min	Fast, seconds	Slow, 20 min	Slow, 20 min
Dissociation				
Principle	Extensive washing steps, washout of dimerizer	Dual translocation of two orthogonal CIDs, (GA3-AM and rapamycin)	Competitor FK506 replaces dimerizer	Competitor TMP replaces dimerizer
Kinetics	Slow, 15 min	Slow, 30 min	Fast, seconds	Slow, 15 min
Multiple rounds	Yes	No	No	Yes
Application	Modulate the subcellular localization of the transcriptional factor NF- $\kappa$ B to induce secretion of IL-8	Activation and inactivation of Rac-/PI3K-dependent membrane ruffling	Activation and inactivation of PI3K allow the measurement of PIP3 turnover and downstream effectors	Relocation of constitutively Rac to PM and mitochondrion, leads to rapid and reversible formation of lamellipodia

**Table 3. Overview of reversible CID systems.** Properties of association and dissociation as well as examples of applications are listed in the table.

## 2. Aim of the thesis

The topic of my thesis originated from the need of a tool that allows one to manipulate interactions of signaling proteins and to initiate activation of signaling pathways at defined starting points and cellular locations. Tools to force proximity of receptors and other signaling molecules with high spatiotemporal precision are required to modulate and investigate underlying signaling pathways. A promising approach to dynamically control protein dimerization and translocation is based on chemical inducers of dimerization (CIDs).

The development of a novel CID based on a Halo- and SNAP-tag reactive dimerizer was an ongoing project in the Wymann lab. Although, HaXS was successfully used to control the activation of an isolated signaling transduction branch in a complex signaling network (see Results, Chapter 3.1, HaXS8 manuscript), the HaXS-induced dimerization is relatively slow (15 min) and dimerization efficiency is moderate, which limits the utility of the HaXS CID to reproduce physiologic signaling pathways that occur on short time scales.

In order to expand the application range of the HaXS CID towards the study of fast cellular events, a first goal of my thesis project was to improve the speed of the dimerization of the HaXS CID (see Results, Chapter 3.3, Protocol Manuscript). A thorough analysis of the HaXS-induced dimerization reactions was performed with a modeling software in order to understand how parameters of the HaXS CID or CIDs in general can be modulated to optimize speed resp. efficiency of the dimer formation and to optimize the experimental conditions to get best results with the HaXS CID.

However, a tool based on chemical molecules will never be able to control cellular events on a time scale of seconds, as the time required for the chemical molecules to penetrate the cell membrane limits its temporal resolution. Additionally, the spatial resolution of a tool based on a chemical molecule is limited due to the intracellular diffusion of the chemical compound, which prevents the control of cellular events occurring in subcellular compartments. In contrast, light is an excellent regulatory trigger. Its delivery is immediate and can be applied on single cells or even subcellular regions, which enables the manipulation of protein dynamics with high temporal resolution.

A second goal of this thesis project was the development of a cell-permeable CID that can be manipulated intracellularly by light, which will provide the possibility to manipulate protein dynamics with increased spatiotemporal precision. In order to profit from the well-characterized dimerization reaction between Halo- and SNAP-tag fusion proteins, we set out to use the HaXS CID as a base for the development of a light-controllable CID. The modular synthetic strategy of HaXS8<sup>[1]</sup>, allows the relatively simple introduction of functional groups into the core module of HaXS8, and thus enables the development of a photocleavable derivative of the HaXS dimerizer (see Results, Chapter 3.2, MeNV-HaXS manuscript). The integration of the photocleavable group into the HaXS dimerizer will additionally enable to

reversibly induce a dimerization event; an important prerequisite for the reproduction of physiological signaling pathways and to study signalling molecule turnovers.

However, even we successfully improved the dimerization performance of the HaXS CID, the induced dimerization speed and efficiency cannot compete with the excellent dimerization kinetics of the rapamycin CID, which profits from high cell permeability of this naturally occurring dimerizer molecule as well as from the high rate constants of FKBP with rapamycin and of FKBP-rapamycin with FRB. However, the FKBP and FRB domains in the dimerizing constructs compete with the endogenous FKBP12 and mTORC1 for rapamycin binding. The binding of FKBP12-rapamycin to the endogenous FRB domain renders the mTORC1 enzymatically inactive and can induce unwanted side effects on the regulation of cell growth, proliferation as well as autophagy and thus limits the use of the rapamycin CID. To overcome the off-target effects of rapamycin, non-toxic rapamycin analogues (so-called rapalogs) were developed, which only bind a genetically engineered mutated version of the FRB, while binding to the wild type FRB domain is abolished.

However, interference with mTORC1 due to minute rapamycin impurities or toxic by-products as well as a low cell penetration capacity of these rapalogs limits the utility of as-yet developed molecules. The third goal of this thesis project was the development of a rapalog CID, which does not interfere with mTORC1 signaling. Additionally, the introduction of the bulky substituent at the interface of rapamycin, which interacts with the FRB domain, should not dramatically affect cell penetration capacities of this rapalog to make sure that excellent dimerization behavior is retained.

### 3. Results

Chapters 3.1 - 3.4 include all projects of this PhD thesis aiming at the development and optimization of heterodimerizer systems to manipulate signaling proteins with high spatiotemporal precision. Chapter 3.1 includes the published manuscript and supplementary information (excluding  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra) of the HaXS8 CID. The next chapter 3.2 contains the published manuscript and supplementary information (excluding  $^1\text{H}$  NMR and  $^{13}\text{C}$  NMR and HRMS spectra) of the photocleavable dimerizer (MeNV-HaXS). Chapter 3.3 includes a manuscript in preparation that describes a detailed protocol of the HaXS CID as well as a thorough analysis of dimerization reactions induced by HaXS or CIDs in general. The last chapter 3.4 includes a manuscript in preparation describing the novel synthetic route and a purification protocol for a non-mTORC1 interfering rapalog, called pcRap. Additional information's concerning the status of this manuscript are included.

## Chemical Development of Intracellular Protein Heterodimerizers

Dominik Erhart,<sup>1</sup> Mirjam Zimmermann,<sup>1</sup> Olivier Jacques,<sup>1</sup> Matthias B. Wittwer,<sup>2</sup> Beat Ernst,<sup>2</sup> Edwin Constable,<sup>3</sup> Marketa Zvelebil,<sup>4</sup> Florent Beaufigli,<sup>1,\*</sup> and Matthias P. Wymann<sup>1,\*</sup>

<sup>1</sup>Department of Biomedicine

<sup>2</sup>Department of Pharmaceutical Sciences

<sup>3</sup>Department of Chemistry

University of Basel, 4003 Basel, Switzerland

<sup>4</sup>The Institute of Cancer Research, London SW7 3RP, UK

\*Correspondence: florent.beaufigli@unibas.ch (F.B.), matthias.wymann@unibas.ch (M.P.W.)

<http://dx.doi.org/10.1016/j.chembiol.2013.03.010>

### SUMMARY

Cell activation initiated by receptor ligands or oncogenes triggers complex and convoluted intracellular signaling. Techniques initiating signals at defined starting points and cellular locations are attractive to elucidate the output of selected pathways. Here, we present the development and validation of a protein heterodimerization system based on small molecules cross-linking fusion proteins derived from HaloTags and SNAP-tags. Chemical dimerizers of HaloTag and SNAP-tag (HaXS) show excellent selectivity and have been optimized for intracellular reactivity. HaXS force protein-protein interactions and can translocate proteins to various cellular compartments. Due to the covalent nature of the HaloTag-HaXS-SNAP-tag complex, intracellular dimerization can be easily monitored. First applications include protein targeting to cytoskeleton, to the plasma membrane, to lysosomes, the initiation of the PI3K/mTOR pathway, and multiplexed protein complex formation in combination with the rapamycin dimerization system.

### INTRODUCTION

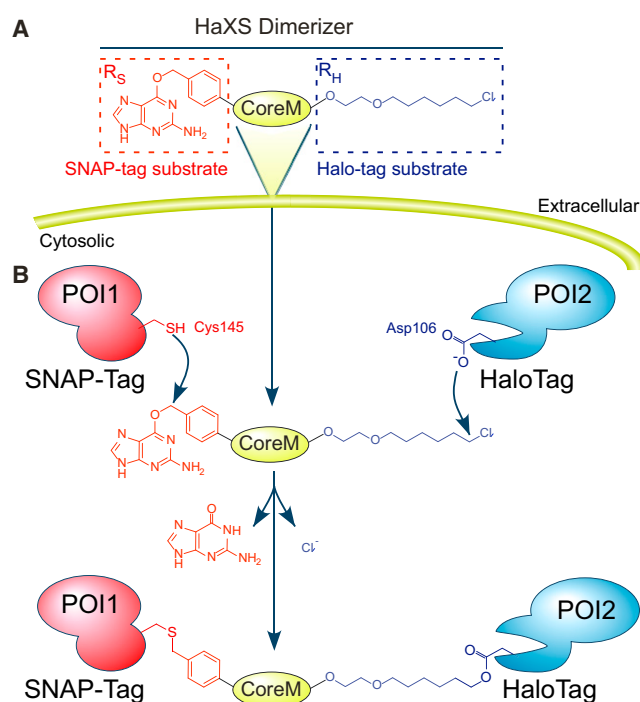
Chemical inducers of dimerization (CIDs) are powerful tools to specifically control protein homo- and heterodimerization (Corson et al., 2008). The most popular CID system currently uses rapamycin to dimerize FKBP12 and the FRB domain of mammalian target of rapamycin (mTOR), which results in the formation of a stable trimeric complex (Crabtree and Schreiber, 1996). Although the system displays excellent kinetics (Banaszynski et al., 2005), cross-reactions with the nutrient sensor mTOR complex 1 (TORC1) can occur. C16-derivatized rapalogs have been designed to interact only with mutated FRB domains, but not with endogenous mTOR (Liberles et al., 1997; Inoue et al., 2005; Edwards and Wandless, 2007). Using rapamycin and rapamycin derivatives (dubbed rapalogs), membrane translocation of lipid phosphatases was achieved

targeting the plasma membrane (Varnai et al., 2006) and endosomes (Fili et al., 2006), while Inoue and Meyer successfully triggered PI3K activation by translocation (Inoue and Meyer, 2008). Minor contamination with rapamycin or rapamycin by-products (<< 1%) can dramatically affect the PI3K/mTOR axis, which renders the rapamycin system less suitable for studies of growth control, immunity, and metabolism (Wullschleger et al., 2006).

The natural product coumermycin is a homodimerizer linking two amino-terminal subdomains of the B subunit of bacterial DNA gyrase (GyrB) on each side of the molecule (Mohi et al., 1998; O'Farrell et al., 1998). Several cellular signaling pathways were successfully targeted with the coumermycin system (Farrar et al., 1996; Mohi et al., 1998; Liu et al., 2000), but biological applications remain limited to homodimerization processes. CID based on dexamethasone (DEX) conjugated to methotrexate (MTX) or trimethoprim (TMP) have been introduced by the Cornish lab and were successfully used in a yeast three-hybrid system (Lin et al., 2000; Baker et al., 2002; Bronson et al., 2008). Phytohormone abscisic acid (ABA) stress response pathway proteins have recently been modified to control the proximity of cellular proteins (Liang et al., 2011). The ABA-induced CID system offers the advantage that it relies on binding partners that have no endogenous counterparts, but so far no rapid dimerization has been demonstrated. Furthermore, most of the systems discussed in Corson et al., 2008, including the ABA system, depend on reversible noncovalent interactions, which render the quantification of successful dimerization challenging.

The use of self-labeling technologies allows the covalent incorporation of a wide range of modifications at specific protein sites (Hinner and Johnsson, 2010). Mutants of the DNA repair protein O6-alkylguanine-DNA-alkyltransferase (SNAP-tag and CLIP-tag) transfer alkyl groups from their substrates, such as O6-benzylguanine (BG; Keppler et al., 2003) and O6-benzylcytosine (BC; Gautier et al., 2008), to a reactive Cys residue. Haloalkane dehalogenase (HaloTag) has been designed to covalently bind to synthetic chloroalkane ligands (Los et al., 2008). Covalent bond formation between these protein tags and their targeted compounds is highly specific, occurs rapidly under physiological conditions, and is essentially irreversible. Johnsson and co-workers have developed hetero-bisfunctional BG-MTX heterodimerizers, which were successfully used to control transcription





**Figure 1. Schematic Representation of Chemically Induced Dimerization of Intracellular Proteins by HaXS Molecules**

(A) HaloTag and SNAP-tag substrate moieties were integrated into bivalent reactive cross-linking molecules (HaXS). The O<sup>6</sup>-benzylguanine function (red, R<sub>S</sub> in subsequent figures) reacts with Cys145 in the SNAP-tag, while the chloroalkane (blue, R<sub>H</sub>) links to Asp106 in the HaloTag. (B) HaXS dimerizer molecules were rendered cell permeable by the integration of specifically designed features into the core module (CoreM) of HaXS dimerizers. Once HaXS enter cells, they heterodimerize exogenously expressed proteins of interest (POIs) fused to SNAP- and HaloTags. In contrary to other covalent chemical dimerizers, HaXS molecules link tagged proteins even if they are not pre-associated (see Figure S1 and Table S1).

in yeast (Gendreig et al., 2003). The same group produced a SNAP-homodimerizer (CoDis; Lemerrier et al., 2007) and bivalent fluorescent substrate S-CROSS, linking SNAP- and CLIP-tagged proteins (Gautier et al., 2009). Like for the covalent homodimerizer xCrAsH, which targets tetracysteine tags (Rutkowska et al., 2011), S-CROSS links exclusively pre-associated binding partners. Although xCrAsH is suited for intracellular applications, S-CROSS reactivity has only been demonstrated in cell lysates so far.

An alternative approach to control protein-protein interactions locally is provided by genetically encoded light-inducible protein dimerizing systems (Yazawa et al., 2009; Kennedy et al., 2010; Idevall-Hagren et al., 2012; Strickland et al., 2012). They offer an unprecedented speed of dimerization (Idevall-Hagren et al., 2012). Large protein tags (e.g., FKF1 with 1173 amino acids [Yazawa et al., 2009] or PhyB with 908 amino acids [Levska et al., 2009]), and sensitivity to accidental exposure to green/blue light, favor the application of these systems in imaging, but obstruct biochemical investigation of triggered interactions.

Here, we show the development and applications of an alternative dimerization strategy based on a class of covalent chemical cross-linkers fusing HaloTag and SNAP-tag substrates

(HaXS; Figure 1A; Figure S1 and Table S1 available online). This tag combination profits from a high reaction rate of the HaloTags and SNAP-tags with their respective substrates (Hinner and Johnsson, 2010) and attains high selectivity inside mammalian cells. The strategy unites efficient covalent protein dimerization with optimized cell permeability in HaXS molecules and provides a simple control and workup of CID-induced complexes (Figure 1B).

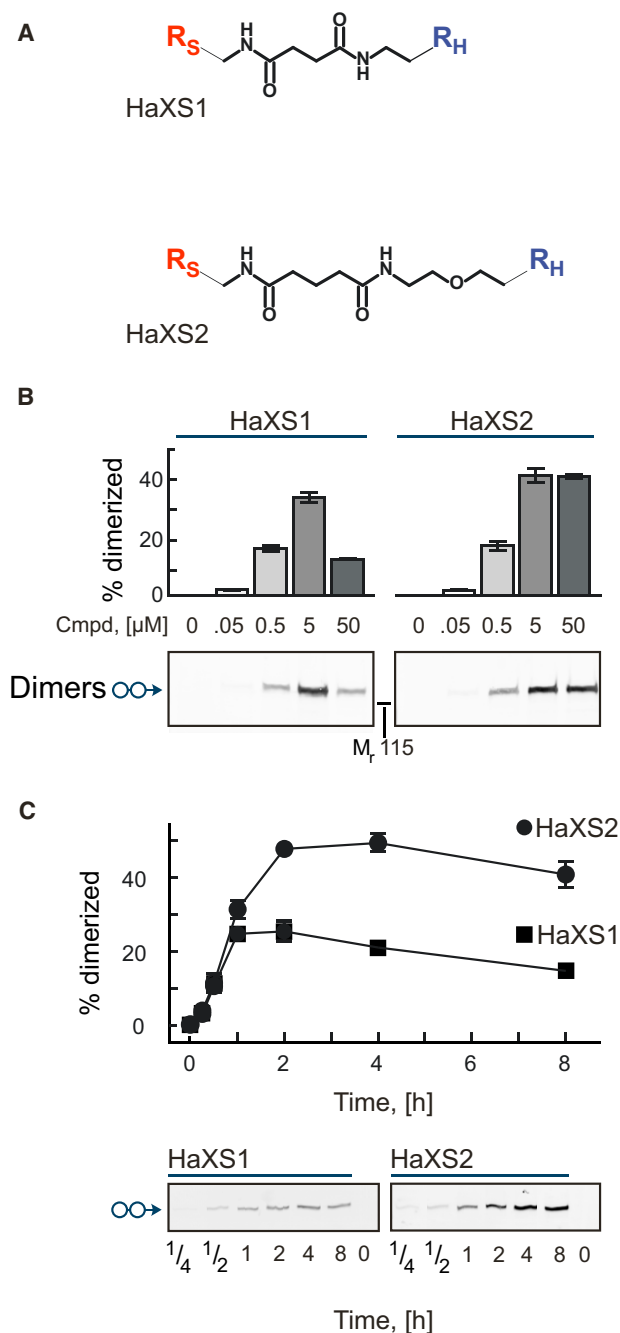
## RESULTS AND DISCUSSION

Intracellular cross-linkers need to combine tag reactivity and cell permeability. Because tag reactivity has been optimized before, the core module linking SNAP-tag and HaloTag substrates was left as a point of attack to modify physicochemical properties of HaXS molecules. A progressive chemical optimization of the core module included the integration of groups modulating water solubility and cell permeability, and of bulky or semirigid structures separating the two tag-reactive moieties (for schematics, see Figure 1B).

### Structural Analysis of Core Module Properties and Improvement of Cell Permeability of HaXS Molecules

To explore the required length of alkyl and PEG elements between HaloTag and SNAP-tag substrates, compounds HaXS1 and HaXS2 (Figure 2A) were tested in cellular systems. In HeLa cells transiently transfected with SNAP-GFP and Halo-GFP fusion proteins, optimal HaXS1 or HaXS2 concentrations (5 μM) yielded up to 40%–50% intracellular dimerization. The elongation of the core module in HaXS2, as compared to HaXS1, improved cross-linking at elevated concentrations and incubation times (Figures 2B and 2C). The observation that time-resolved dimerization curves of HaXS1 and HaXS2 match initially ( $t < 1$  hr, Figure 2C) and the fact that HaXS1 has a higher cell permeability than HaXS2 as determined in a permeability assay (see below) suggest that the faster diffusion rate of HaXS1 could initially compensate its inferior reactivity (compared to HaXS2), while the longer HaXS2 yields higher dimerization efficiencies at later time points (at 4 hr > 50% for HaXS2, 25% for HaXS1). The better cell permeability and inferior dimerization capacity of HaXS1 also explains why high concentrations of HaXS1 (50 μM) counteracted the formation of heterodimeric HaloTag-SNAP-tag complexes, most likely by masking reactions generating increasing levels of monomeric, reacted Halo- and SNAP-GFP species (Figure 2B), which was not observed with high concentrations of HaXS2.

These results suggest that intracellular dimerization reaction is best interpreted as a convolution of substrate/tag reactivity and limited diffusion of the compounds into intracellular space. This view is also supported by the extended time required to reach relevant yields of intracellular dimerization with this first-generation HaXS molecules, as compared to previously reported single sided reaction rates of HaloTag and SNAP-Tag with their specific substrates ( $3 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$  and  $3 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ , respectively; Hinner and Johnsson, 2010). These pilot studies with HaXS1 and HaXS2 made it clear that a minimal core module size is crucial to efficiently cross-link the Halo-GFP and SNAP-GFP fusion proteins, and that cell permeability needs to be improved for faster intracellular reactivity.



**Figure 2. Intracellular Heterodimerization Induced by First-Generation HaXS Molecules**

(A) Abbreviated structure (for  $R_S$  and  $R_H$  see Figure 1) of the bisfunctional HaloTag and SNAP-tag reactive molecules HaXS1 and HaXS2.

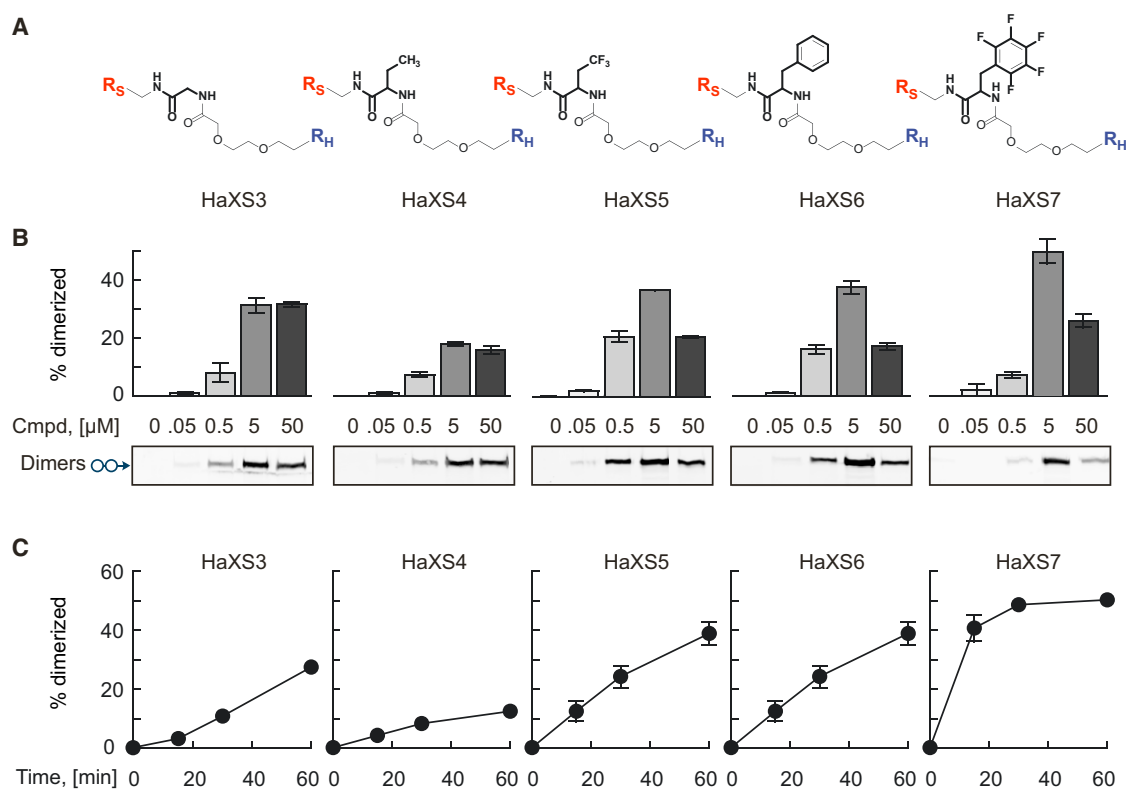
(B) HeLa cells transfected with expression constructs for SNAP-GFP and Halo-GFP (see Supplemental Experimental Procedures) were exposed to HaXS1 and HaXS2 at the indicated concentrations for 1 hr at 37°C in complete cell culture medium, before cells were lysed and proteins were subjected to SDS-PAGE and immune-blotting for SNAP/HaloTag dimers using anti-GFP (primary) and fluorescently-labeled (secondary) antibodies to generate a readout on an infrared imaging system.

(C) HeLa cells (cotransfected with SNAP-GFP and Halo-GFP as in (B)) were incubated with the indicated HaXS compound (5  $\mu$ M), and lysed at the indicated time for analysis of dimers as in (B). Quantifications of the concentration

and time-dependent heterodimerization are means  $\pm$  SEM of three independent experiments. Error bars were removed where smaller than the symbols used. A detailed description of the synthesis of HaXS molecules can be found in the Supplemental Experimental Procedures.

Second-generation HaXS molecules were designed on the basis of a HaXS2 scaffold aimed to maintain comparable core module size and low molecular weight, while introducing features to improve cell permeability. Along these lines, HaXS3 was designed to replace the central glycine moiety with functionalized amino acids, and to generate a HaXS library (Figure 3A). Because fluorinated groups have been reported to modulate physicochemical properties of molecules considerably (Böhm et al., 2004), we produced alkylated and fluorinated analogs of HaXS3 to integrate a lipophilic surface as an entry point into cellular membranes. To further investigate the influence of fluorine substitutions on cell penetration of HaXS3 derivatives, fluorinated and nonfluorinated analogs were produced in parallel: alkyl-analog HaXS4 and fluorinated HaXS5 were synthesized by the exchange of the glycine moiety included in the HaXS3 core structure by methylalanine or trifluoromethylalanine groups, respectively. The phenyl-analogs HaXS6 and the corresponding fluorinated HaXS7 were made according to the same strategy from Fmoc-phenylalanine or Fmoc-pentafluorophenylalanine (Figure 3A). HaXS3 has comparable intracellular dimerization efficiency compared to first-generation HaXS molecules. The simple alkyl chain introduced in HaXS4 led to a reduction of intracellular dimerization as compared to HaXS3 and HaXS2, but the trifluoromethylated HaXS5 rescued dimerization of Halo-GFP and SNAP-GFP fusion protein back to 40% within 1 hr. The lipophilic nature of the phenyl group in HaXS6 and the pentafluorophenyl group in HaXS7 improved cell penetration even further. The nonfluorinated aryl derivative HaXS6 showed characteristics matching those of the fluorinated HaXS5 molecule. Significant progress in rate and efficiency in intracellular dimerization was achieved with the pentafluorophenyl derivative HaXS7 (Figures 3B and 3C).

Based on results from the above HaXS series, a third-generation dimerizer was envisaged: assuming that HaXS molecules are delivered inside cells via passive diffusion through cell membranes, we next targeted a dimerizer with a lower molecular weight, and thus synthesized a flat derivative of HaXS7. One of the emerging strategies was to integrate the polyfluorophenyl substituent from HaXS7 into the core module structure. HaXS8 was synthesized in six steps from tetrafluorohydroquinone, tetraethylene glycol, BG-NH<sub>2</sub>, and 6-chloro-1-iodohexane (Figure 4A) with a good overall yield (1 g produced in total; see Supplemental Experimental Procedures available online). Additionally, minimization of core module flexibility aimed to prevent steric clashes during the formation of the protein complex, hereby allowing higher yields of Halo-GFP and SNAP-GFP dimerization. Molecular modeling was performed to eliminate the possibility that the designed molecules would link the Halo-Tag and the SNAP-tag in a way to generate unfavorable steric clashes between the two tags. As shown in Figure 4B, HaXS8 provided a fair amount of degrees of freedom to arrange the HaloTag and the SNAP-tag proteins in the docked complex, suggesting that HaXS8 provided sufficient core module size and flexibility for optimal docking.



**Figure 3. Intracellular Dimerization Triggered by Second-Generation HaXS Molecules—Influence of Concentration and Time of Exposure**

(A) Abbreviated structures of fluorinated and nonfluorinated bisfunctional Halo- and SNAP-tag reactive molecules. (B) HeLa cells co-expressing SNAP-GFP and Halo-GFP were exposed to the indicated amounts of HaXS compounds for 1 hr at 37°C (for details see Figure 2C and Supplemental Experimental Procedures). (C) HeLa cells expressing SNAP-GFP and Halo-GFP fusion proteins were incubated with 5 μM of the indicated HaXS molecules, before they were lysed at the given times for detection of dimerization by immune-blotting (as in Figure 2C). Values represent the mean of three independent experiments ± SEM.

In HeLa cells, intracellular dimerization of Halo-GFP and SNAP-GFP fusion proteins reached > 65% with HaXS8, notably already at 10-fold lower concentrations of the compound as compared to HaXS2 and HaXS7 (Figure 4C). Differences between HaXS8 and earlier generation HaXS molecules became even more prominent at lower concentrations, as HaXS8 displayed significant intracellular dimerization as low as 50 nM, where none of the other HaXS molecules displayed significant dimerization. Moreover, HaXS8 was capable of dimerizing > 50% of tagged proteins over a wide concentration range. This exceeds values typically achieved for surface receptor/signaling chain interactions by a factor of around 10×, as receptors recruit cytosolic components in the low percentage range (Deswal et al., 2011).

To correlate cell permeability and the performance in intracellular dimerization reactions directly, we performed a parallel artificial membrane permeability assay (PAMPA; for a review see Faller, 2008) with all HaXS molecules (Figure 4E). A correlation between the obtained permeability values ( $P_e$ ) and the intracellular dimerization obtained for each HaXS compound showed that cell permeability is indeed a key factor for an efficient heterodimerization process in cells. HaXS1 is the only outlier in this analysis, as the compound has the lowest molecular weight contributing to improved membrane

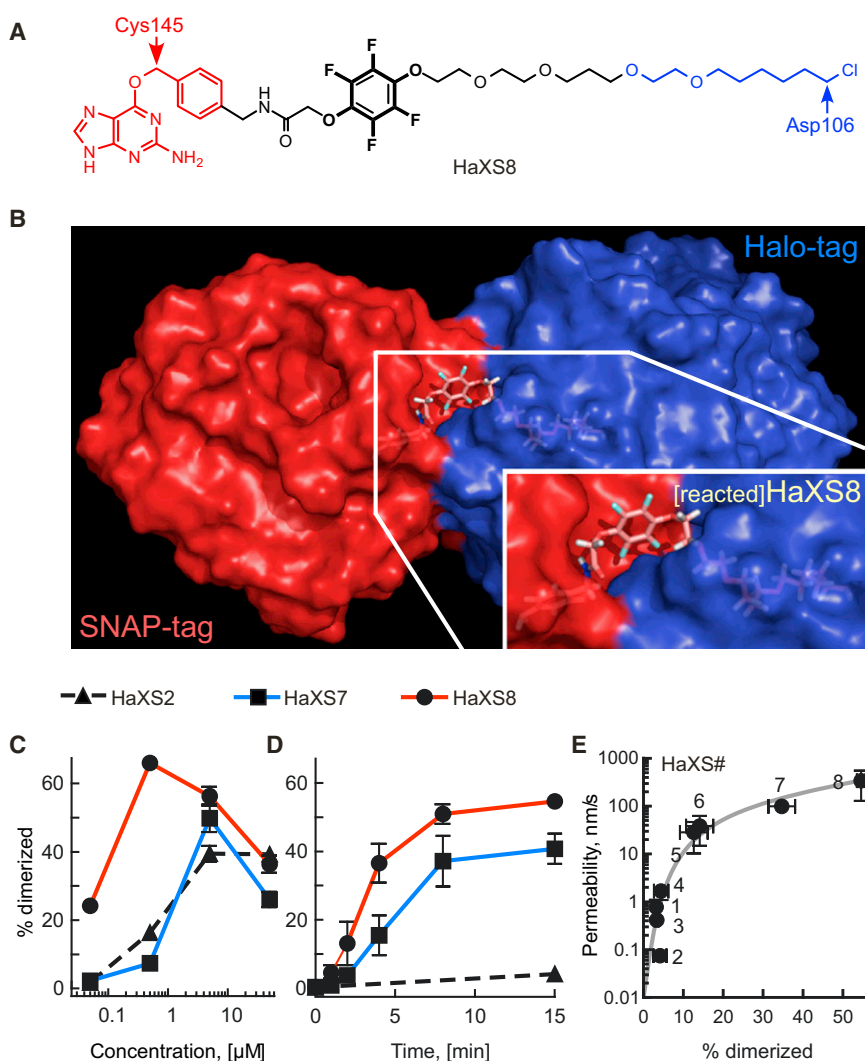
permeability, but reduced coupling reactivity due to steric constraints.

#### Protein Translocation to Selected Cellular Compartments

Using HaXS8, we next aimed to target tagged proteins to specific locations inside living cells. As a first approach, tagged fusion proteins of actin (SNAP-Actin) and GFP (Halo-GFP) were co-expressed in NIH 3T3 cells. These cells form stress fibers, which contain filamentous actin (F-actin), and can be visualized with the F-actin stain rhodamine-phalloidin. Upon addition of HaXS8, Halo-GFP was translocated to stress fibers, and colocalized with rhodamine-phalloidin staining, which was confirmed by image analysis of cross-linked Halo-GFP/SNAP-Actin complexes (Figure 5A).

In parallel, HaXS8 controlled the translocation of fluorescent fusion proteins from the cytoplasm to the plasma membrane in MDCK epithelial cells: cytosolic red fluorescent protein (monomeric RFP; TagRFP) was expressed as HaloTag fusion protein (Halo-RFP), and a membrane anchor was constructed fusing a SNAP-tag, GFP, and the isoprenylation sequence from KRas4B (the CAAX box), which targeted the resulting SNAP-GFP-CAAX fusion protein exclusively to the plasma membrane. When MDCK cells, expressing SNAP-GFP-CAAX and Halo-RFP,





**Figure 4. The Structurally Optimized Dimerizer HaXS8 in Comparison with the First- and Second-Generation HaXS Compounds**

(A) Chemical structure of the optimized bisfunctional HaloTag and SNAP-tag reactive molecule HaXS8.

(B) Modeled structure of a covalently linked HaloTag-[reacted]HaXS8-SNAP-tag adduct. The [reacted]HaXS8 molecule (elimination of guanine and  $\text{Cl}^-$ , see Figure 1) was linked to Cys145 in the SNAP-tag (red) and to Asp106 in the HaloTag (blue) crystal structures, before randomized structural starting points were put through energy minimization (see Supplemental Experimental Procedures). Depicted is a sample structure representing a tight fit between HaloTag and SNAP-tag, which leaves the alkane chain within the HaloTag entry tunnel in a relaxed conformation.

(C) HeLa cells co-expressing SNAP-GFP and Halo-GFP were incubated in complete medium with HaXS8, HaXS7, or HaXS2 at the indicated concentrations for 1 hr at 37°C before cell lysis (see Figure 2C).

(D) HeLa cells as in (C) were exposed to 5 μM of HaXS8, HaXS7, or HaXS2 for the indicated time, before cells were lysed and probed for HaloTag and SNAP-tag dimerization (all values are means of three individual experiments ± SEM).

(E) All HaXS compounds presented here were subjected to an artificial membrane permeability assay (PAMPA; see Supplemental Experimental Procedures) to assess cell permeability (expressed as  $P_e$ , [nm/s]).  $P_e$  was plotted logarithmically against the percentage of intracellular dimerization achieved after a 15 min exposure of HeLa cells (as in C) to the indicated HaXS compounds at 5 μM (labeled 1–8;  $n = 3$ ; mean ± SEM).

were incubated with HaXS8, the initially cytosolic Halo-RFP protein was relocated to the plasma membrane (Figure 5B).

#### Chemical Induction of Intracellular Signaling

Protein/membrane interactions play a major role in the activation of cellular signaling cascades in physiology and disease (Wymann and Schneider, 2008). Overactivation of phosphoinositide 3-kinase (PI3K) has gained much attention because it contributes to the promotion of cancer and inflammation (Cantley, 2002; Wymann and Marone, 2005; Zoncu et al., 2011; Wymann, 2012). Class IA PI3Ks are composed of a regulatory p85 and a catalytic p110 kD subunit. The p85 subunit docks with two src-homology 2 (SH2) domains to phosphorylated tyrosines on activated growth factor receptors and drags the catalytic p110 subunit via a coiled-coil domain located between the two SH2 domains (dubbed inter-SH2, or iSH2; for structural schemes see Wymann et al., 2003) to the membrane. PI3K then produces phosphatidylinositol(3,4,5)-trisphosphate [ $\text{PtdIns}(3,4,5)P_3$ ], a lipid that serves as a plasma membrane docking site for signaling enzymes containing pleckstrin homology (PH) domains. Phosphoinositide-dependent kinase-1 (PDK1) and

protein kinase B (PKB/Akt) both contain  $\text{PtdIns}(3,4,5)P_3$ -binding PH domains. PDK1 directly phosphorylates PKB/Akt on Thr308, while a secondary site on PKB/Akt (Ser473) is phosphorylated by mTOR complex 2 (TORC2). When fully activated, PKB/Akt takes a major role in the activation of TORC1, and active TORC1 phosphorylates p70 S6 kinase ( $p70^{\text{S6K}}$ ) on Thr389.

Rapamycin derivatives have been reported to boost a feedback leading to the amplification of PI3K signaling upstream of TORC1 (O'Reilly et al., 2006), which provides a reason to validate the HaXS strategy in the PI3K/mTOR pathway: a SNAP-GFP-CAAX anchor in the plasma membrane (see above) served as the docking site for a HaloTag protein fused to the inter-SH2 domain of p85 (Halo-iSH2-GFP; Figure 6A). The addition of HaXS8 to HEK293 cells induced a rapid and efficient cross-linking of the membrane anchor and the iSH2 construct, and triggered the activation of the downstream targets PKB/Akt (as monitored by Thr308 and Ser473 phosphorylation on PKB/Akt) and mTOR (monitored by phosphorylation of T389 of p70 S6 kinase,  $p70^{\text{S6K}}$  by TORC1). Although typical growth factor stimulation also induces an activation of the MAPK pathway, HaXS8

did not upregulate MAPK phosphorylation, demonstrating that the system can be used to selectively induce isolated signal transduction branches (Figure 6B).

Control experiments forcing the dimerization of a cytosolic SNAP-GFP with the Halo-iSH2-GFP did not activate the PI3K/mTOR pathway, illustrating that membrane proximity of the iSH2/p110 PI3K complex was crucial to relay downstream signals. Moreover, HaXS8 did not interfere with mTOR signaling induced by serum stimulation. Rapalog-based CID systems can generate a distorted output in this system (see Figure S2 and Supplemental Experimental Procedures). Although HaXS8 does not yet match the dimerization rates attained with rapalog CID systems, it offers the advantage that numerous connections within the PI3K/mTOR pathway can be explored without interference during long term stimulation.

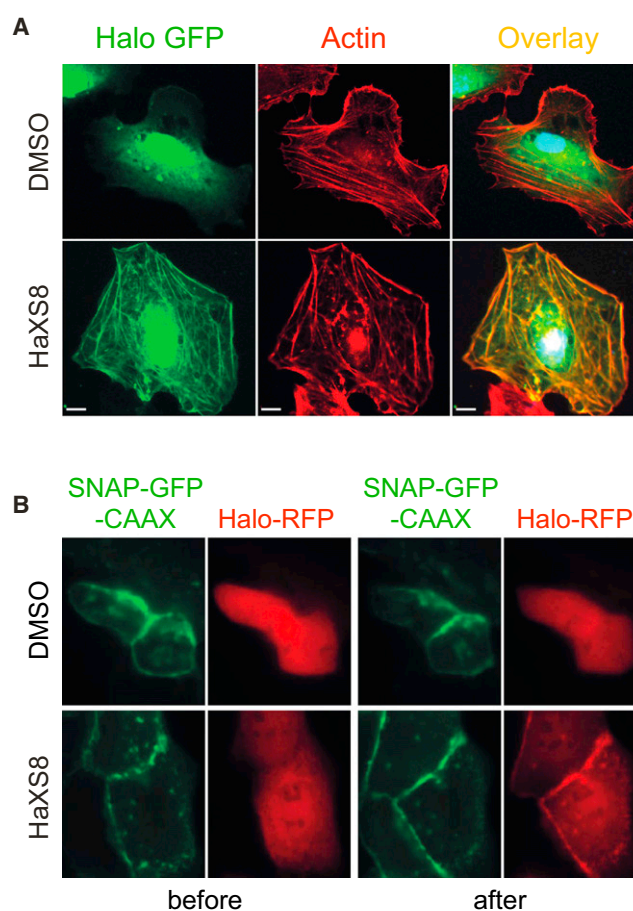
#### Multiplexing Applications

It can also be envisaged that the HaXS dimerizers can be applied orthogonally in conjunction with other CID systems, as this has been elegantly demonstrated for the combination of rapamycin and gibberellin analogs recently (Miyamoto et al., 2012). Such combinations allow the construction and alignment of sophisticated enzymatic reaction chains that can be utilized to probe complex signaling pathways. A proof-of-concept approach combining a pre-dimerization step mediated by HaXS8, followed by a rapamycin-induced translocation of the preformed Halo/SNAP-tag complex to a lysosomal anchor protein containing LAMP1 (LAMP-CFP-FRB), is shown in Figure 7. Successful translocation of a cytosolic Halo-RFP fusion protein to lysosomes was only observed in the presence of HaXS8 and rapamycin, as a FKBP-SNAP protein acted as a bridge from Halo-RFP to the membrane anchor (see Figure 7B for a scheme of the docked complexes). Although only two fluorescent proteins (CFP and RFP) were used as cargo to monitor selective translocation, up to three cargo proteins or enzymes can be easily integrated into the setup.

As for other CID systems, tag fusions and probe constructs need to be considered carefully when using HaXS8 molecules to dimerize HaloTag and SNAP-tag fusion proteins. While non-covalent CID assemblies can only be rated and modified by their expected cellular output, covalent complexes can be easily detected in the HaXS system, and success of dimerization can be validated, and correlated with signal output and subcellular targeting directly. In contrast to other covalent systems (see S-CROSS and xCrAsH), HaXS8 can force dimerization of previously noninteracting proteins, and has been specifically optimized for cellular permeability.

#### SIGNIFICANCE

The cell-permeable HaXS molecules promote a covalent intracellular dimerization of HaloTag and SNAP-tagged proteins of interest. The covalent and irreversible nature of the reaction of the chemical dimerizer with HaloTag and SNAP-tag allows easy monitoring of the dimerization process even under denaturing conditions. The design and chemical development of HaXS compounds focused on the optimization of cellular availability and reactivity. Although previously described covalent protein dimerizers

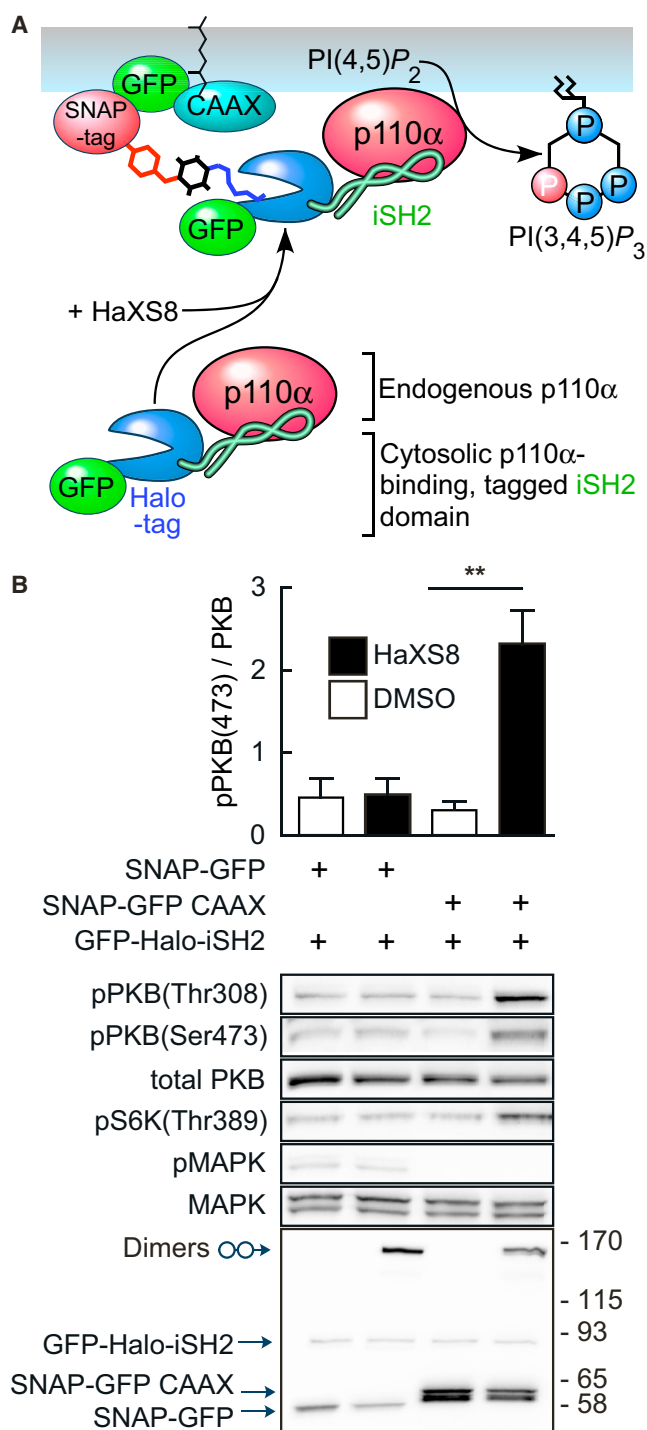


**Figure 5. Protein Translocation to Specific Cellular Compartments**

(A) NIH 3T3 cells expressing SNAP-Actin and Halo-GFP (green) fusion proteins were incubated with DMSO or 0.5  $\mu$ M HaXS8 for 1 hr at 37°C. Rhodamine-phalloidin was used to visualize F-actin in stress fibers (Actin, red). Translocation of GFP to the actin cytoskeleton was imaged by confocal microscopy on live cells.

(B) MDCK cells expressing a SNAP-GFP-CAAX membrane anchor (green [-CAAX is the polybasic isoprenylation sequence from KRas-4B]) and Halo-RFP (red) fusion protein were incubated with DMSO or 0.5  $\mu$ M HaXS8 in complete medium for 40 min at 37°C. Translocation of Halo-RFP to the plasma membrane was imaged by confocal microscopy on live cells.

were used to confirm well-known protein-protein interactions, HaXS molecules have the ability to force protein-protein interactions. As shown here, forced protein complex formation can also be exploited to promote protein translocation to different cellular compartments, and to activate distinct cellular signaling pathways. In contrary to the widely used rapalog CID systems, HaXS compounds can be utilized to trigger PI3K/mTOR signaling pathways, without interference with endogenous signaling molecules and induction of feedback mechanisms. HaXS molecules can be used in combination with other (noncovalent) dimerization systems, and thus extend the possibilities to devise multiplexing approaches, and to chemically control the assembly of elaborate protein complexes and signalosomes.



**Figure 6. Activation of the PI3K/mTOR Pathway by Chemically Induced Membrane Targeting**

(A) Schematic representation of the translocation process recruiting a cytosolic iSH2 domain (derived from the PI3K regulatory PI3K subunit p85, see text) fused to the HaloTag to a membrane anchored SNAP-tag using HaXS8. PI3Ks activity at the membrane forms PtdIns(3,4,5)P<sub>3</sub> [PI(3,4,5)P<sub>3</sub>] from PI(4,5)P<sub>2</sub>, followed by specific activation of PI3K/PKB/mTOR pathway.

(B) HEK293 cells were cotransfected with SNAP-GFP-CAAX (from KRas4B) and iSH2 fused to Halo-GFP. HEK293 cells expressing SNAP-GFP-CAAX and iSH2-Halo-GFP were starved overnight, and exposed to DMSO or 0.5 μM

## EXPERIMENTAL PROCEDURES

### Synthesis

Detailed synthetic procedures, materials and reagents, and characterizations for all compounds are described in the [Supplemental Experimental Procedures](#).

### Protein Denaturation, Cell Lysis, and Immunoblotting

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 13,000 rpm for 15 min and proteins were denatured by the addition of 5× sample buffer (312.5 mM Tris-HCl [pH 6.8], 10% SDS, 25% β-mercaptoethanol, 50% glycerol, bromophenol blue) and boiling for 6 min. Proteins were separated by SDS-PAGE, and transferred to Immobilon PVDF membranes (Millipore). 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), mouse mAb to PKB (kind gift of E. Hirsch, Turin, Italy), mouse mAb to pMAPK and rabbit mAb to MAPK (both from Sigma-Aldrich), mouse mAb to HA (HA.11, Babco), and GFP (Roche Diagnostics) were used to detect proteins by immunoblotting. Secondary antibodies were either labeled with Alexa Fluor 680 (LI-COR) for detection on an Odyssey (LI-COR) infrared imaging system, or were horseradish peroxidase (HRP)-conjugated goat antimouse IgG and goat anti-rabbit IgG (Sigma) for visualization using enhanced chemiluminescence (Millipore) detected by a CCD camera system (Fusion Fx7, Vilber).

### Cloning and Expression of Recombinant Proteins

The HaloTag7 coding sequence (Promega), SNAP-tag (pSS26m) coding sequence (Covalys), iSH2 domain coding sequence (Addgene), actin coding sequence (Clontech), and EGFP coding sequence (Clontech) were amplified by PCR (Phusion polymerase, Finnzymes) and transferred to pcDNA3 (Invitrogen), pTagRFP-N1 (Evrogen; expression vector for a monomeric TagRFP, here short RFP from sea anemone *Entacmaea quadricolor* [Merzlyak et al., 2007]), pEGFP (Clontech) with excised GFP or pEGFP-C3 (Clontech) vectors for expression. For recombinant protein production, Halo-EGFP and SNAP-EGFP were cloned into pTriEx-4 (Novagen) and expressed as N-terminal (His)<sub>6</sub> fusion proteins, and purified on Ni<sup>2+</sup>-NTA beads (QIAGEN) according to the manufacturer's instructions.

To generate a Halo-FRB expression construct, CLIP was exchanged in a CLIP-FRB plasmid (Gautier et al., 2009) by the Halo sequence. The Halo-FRB and SNAP-FKBP (Gautier et al., 2009) cassettes were then transferred into a pcDNA3 (Invitrogen) backbone containing an N-terminal HA-tag in the multicloning site. The LAMP-CFP-FRB plasmid is described in (Komatsu et al., 2010).

### Cell Culture and Transfection

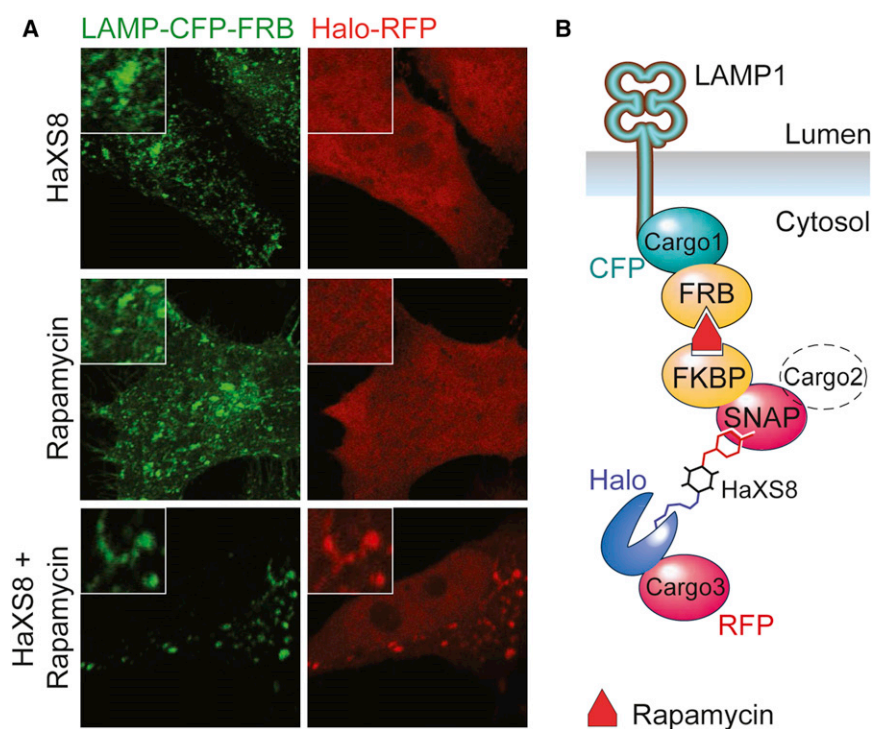
HeLa, HEK293, MDCK, and NIH 3T3 (originally from ATCC) were cultured in complete Dulbecco's modified Eagle medium (DMEM) with 10% heat-inactivated fetal calf serum (HIFCS), 2 mM L-glutamine (Gln), 1% penicillin-streptomycin solution (PEST) at 37°C, and 5% CO<sub>2</sub>. Transfections were carried out with JetPEI (Brunschwig) according to the manufacturer's guidelines.

### Cellular Heterodimerization and Biological Induction

One day after transfection of HeLa cells with expression constructs for SNAP-GFP, SNAP-GFP-CAAX, Halo-RFP, Halo-GFP, LAMP-CFP-FRB, HA-SNAP-FKBP, or SNAP-Actin, cells were exposed to HaXS dimerizers at the indicated concentrations for indicated times at 37°C in fully supplemented complete

medium for 40 min, before cell lysis and signaling pathway analysis using the indicated phosphor-specific antibodies. Control experiments included cytosolic SNAP-GFP constructs. SNAP/HaloTag dimers were detected as described previously. Quantifications of signal intensities represent the mean of ± SEM of two independent experiments. HaXS8 does not interfere with PI3K/mTOR signaling (see [Figure S2](#)).





**Figure 7. An Orthogonal Application of the HaXS System in Combination with Rapamycin-Mediated Protein Dimerization**

(A) HeLa cells co-expressing LAMP-CFP-FRB (green), SNAP-FKBP, and Halo-RFP (red) were first exposed to 0.5  $\mu$ M HaXS8 for 1 hr, followed by 0.5  $\mu$ M rapamycin for 5 min, where indicated. Subsequently, cells were fixed and pictured by confocal microscopy.

(B) Schematic representation of the docked and assembled multimeric complex: cytosolic Halo-Tag fused to RFP (Cargo3) can form a covalent link to a SNAP-FKBP fusion protein via the reaction with HaXS8. This complex (which could be extended by Cargo2) is still cytosolic. Once rapamycin is added, the complex is targeted to a FRB-tagged anchor (here LAMP-CFP-FRB) at specific cellular locations (here lysosomal membranes) via the FKBP-rapamycin-FRB interaction.

DMEM medium. For immunoblot analysis, cells were lysed, and proteins were separated by SDS-PAGE. SNAP-tag/HaloTag dimers were detected using anti-GFP (primary) and fluorescently labeled (secondary) antibodies, and measured on the Odyssey infrared imaging system. For microscopy transfected cells were grown on 12 mm coverslips (Menzel), treated with DMSO or HaXS8, washed twice with PBS, fixed with 4% *p*-formaldehyde (PFA) in PBS, and mounted in Mowiol (Plüss-Stauffer) containing 1% propyl gallate (Sigma-Aldrich). For staining of F-actin, cells were permeabilized in PBS, 1% BSA, 0.1% Triton X-100, and incubated with rhodamine-phalloidin (Molecular Probes).

For the studies of PI3K/mTOR signaling, HEK293 cells were cotransfected with SNAP-GFP-CAAX and iSH2 fused to Halo-GFP. One day after transfection, cells were serum-starved overnight. After a 40 min exposure with 0.5  $\mu$ M HaXS8 cells were lysed for the analysis of signal pathway induction, and the formation of HaloTag/SNAP-tag dimers.

For live cell microscopy, transfected cells grown on coverslips were mounted in life-microscopy chambers (Life Imaging Services) in a closed confirmation with complete medium. Images were acquired on an Axiovert 200 M microscope (Zeiss) fitted with a Plan-Achromat 63 $\times$ /1.4 oil objective and an Orca ER II camera (Hamamatsu), and operated by OpenLab software (Perkin Elmer).

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, two figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.chembiol.2013.03.010>.

#### ACKNOWLEDGMENTS

We thank T. Wandless for iRap, and B. Giese for valuable advice and discussions, T. Inoue for the LAMP-CFP-FRB expression plasmid, and K. Johnsson for FRB- and FKBP domain-containing plasmids. This work was supported by Swiss National Science Foundation grants 205320-138302 and 31EM30-126143, the ESF EuroMEMBRANE programme grant FP-018, and the Novartis (Jubilée) Foundation.

Received: January 8, 2013

Revised: March 13, 2013

Accepted: March 20, 2013

Published: April 18, 2013

#### REFERENCES

- Baker, K., Bleczynski, C., Lin, H., Salazar-Jimenez, G., Sengupta, D., Krane, S., and Cornish, V.W. (2002). Chemical complementation: a reaction-independent genetic assay for enzyme catalysis. *Proc. Natl. Acad. Sci. USA* 99, 16537–16542.
- Banaszynski, L.A., Liu, C.W., and Wandless, T.J. (2005). Characterization of the FRB.rapamycin.FRB ternary complex. *J. Am. Chem. Soc.* 127, 4715–4721.
- Böhm, H.J., Banner, D., Bendels, S., Kansy, M., Kuhn, B., Müller, K., Obst-Sander, U., and Stahl, M. (2004). Fluorine in medicinal chemistry. *ChemBioChem* 5, 637–643.
- Bronson, J.E., Mazur, W.W., and Cornish, V.W. (2008). Transcription factor logic using chemical complementation. *Mol. Biosyst.* 4, 56–58.
- Cantley, L.C. (2002). The phosphoinositide 3-kinase pathway. *Science* 296, 1655–1657.
- Corson, T.W., Aberle, N., and Crews, C.M. (2008). Design and applications of bifunctional small molecules: why two heads are better than one. *ACS Chem. Biol.* 3, 677–692.
- Crabtree, G.R., and Schreiber, S.L. (1996). Three-part inventions: intracellular signaling and induced proximity. *Trends Biochem. Sci.* 21, 418–422.
- Deswal, S., Schulze, A.K., Höfer, T., and Schamel, W.W. (2011). Quantitative analysis of protein phosphorylations and interactions by multi-colour IP-FCM as an input for kinetic modelling of signalling networks. *PLoS ONE* 6, e22928.
- Edwards, S.R., and Wandless, T.J. (2007). The rapamycin-binding domain of the protein kinase mammalian target of rapamycin is a destabilizing domain. *J. Biol. Chem.* 282, 13395–13401.

- Faller, B. (2008). Artificial membrane assays to assess permeability. *Curr. Drug Metab.* 9, 886–892.
- Farrar, M.A., Alberol-Ila, J., and Perlmutter, R.M. (1996). Activation of the Raf-1 kinase cascade by coumermycin-induced dimerization. *Nature* 383, 178–181.
- Fili, N., Calleja, V., Woscholski, R., Parker, P.J., and Larjani, B. (2006). Compartmental signal modulation: Endosomal phosphatidylinositol 3-phosphate controls endosome morphology and selective cargo sorting. *Proc. Natl. Acad. Sci. USA* 103, 15473–15478.
- Gautier, A., Juillerat, A., Heinis, C., Corrêa, I.R.J., Jr., Kindermann, M., Beauvais, F., and Johnsson, K. (2008). An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* 15, 128–136.
- Gautier, A., Nakata, E., Lukinavicius, G., Tan, K.T., and Johnsson, K. (2009). Selective cross-linking of interacting proteins using self-labeling tags. *J. Am. Chem. Soc.* 131, 17954–17962.
- Gendreizig, S., Kindermann, M., and Johnsson, K. (2003). Induced protein dimerization in vivo through covalent labeling. *J. Am. Chem. Soc.* 125, 14970–14971.
- Hinner, M.J., and Johnsson, K. (2010). How to obtain labeled proteins and what to do with them. *Curr. Opin. Biotechnol.* 21, 766–776.
- Idevall-Hagren, O., Dickson, E.J., Hille, B., Toomre, D.K., and De Camilli, P. (2012). Optogenetic control of phosphoinositide metabolism. *Proc. Natl. Acad. Sci. USA* 109, E2316–E2323.
- Inoue, T., and Meyer, T. (2008). Synthetic activation of endogenous PI3K and Rac identifies an AND-gate switch for cell polarization and migration. *PLoS ONE* 3, e3068.
- Inoue, T., Heo, W.D., Grimley, J.S., Wandless, T.J., and Meyer, T. (2005). An inducible translocation strategy to rapidly activate and inhibit small GTPase signaling pathways. *Nat. Methods* 2, 415–418.
- Kennedy, M.J., Hughes, R.M., Peteya, L.A., Schwartz, J.W., Ehlers, M.D., and Tucker, C.L. (2010). Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods* 7, 973–975.
- Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* 21, 86–89.
- Komatsu, T., Kukelyansky, I., McCaffery, J.M., Ueno, T., Varela, L.C., and Inoue, T. (2010). Organelle-specific, rapid induction of molecular activities and membrane tethering. *Nat. Methods* 7, 206–208.
- Lemercier, G., Gendreizig, S., Kindermann, M., and Johnsson, K. (2007). Inducing and sensing protein–protein interactions in living cells by selective cross-linking. *Angew. Chem. Int. Ed. Engl.* 46, 4281–4284.
- Levskaia, A., Weiner, O.D., Lim, W.A., and Voigt, C.A. (2009). Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* 461, 997–1001.
- Liang, F.S., Ho, W.Q., and Crabtree, G.R. (2011). Engineering the ABA plant stress pathway for regulation of induced proximity. *Sci. Signal.* 4, rs2.
- Liberles, S.D., Diver, S.T., Austin, D.J., and Schreiber, S.L. (1997). Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen. *Proc. Natl. Acad. Sci. USA* 94, 7825–7830.
- Lin, H., Abida, W.M., Sauer, R.T., and Cornish, V.W. (2000). Dexamethasone-Methotrexate: An Efficient Chemical Inducer of Protein Dimerization. *J. Am. Chem. Soc.* 122, 4247–4248.
- Liu, R., Liu, C.B., Mohi, M.G., Arai, K., and Watanabe, S. (2000). Analysis of mechanisms involved in the prevention of gamma irradiation-induced apoptosis by hGM-CSF. *Oncogene* 19, 571–579.
- Los, G.V., Encell, L.P., McDougall, M.G., Hartzell, D.D., Karassina, N., Zimprich, C., Wood, M.G., Learish, R., Ohana, R.F., Urh, M., et al. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* 3, 373–382.
- Merzlyak, E.M., Goedhart, J., Shcherbo, D., Bulina, M.E., Shcheglov, A.S., Fradkov, A.F., Gaintzeva, A., Lukyanov, K.A., Lukyanov, S., Gadella, T.W., and Chudakov, D.M. (2007). Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nat. Methods* 4, 555–557.
- Miyamoto, T., DeRose, R., Suarez, A., Ueno, T., Chen, M., Sun, T.P., Wolfgang, M.J., Mukherjee, C., Meyers, D.J., and Inoue, T. (2012). Rapid and orthogonal logic gating with a gibberellin-induced dimerization system. *Nat. Chem. Biol.* 8, 465–470.
- Mohi, M.G., Arai, K., and Watanabe, S. (1998). Activation and functional analysis of Janus kinase 2 in BA/F3 cells using the coumermycin/gyrase B system. *Mol. Biol. Cell* 9, 3299–3308.
- O'Farrell, A.M., Liu, Y., Moore, K.W., and Mui, A.L. (1998). IL-10 inhibits macrophage activation and proliferation by distinct signaling mechanisms: evidence for Stat3-dependent and -independent pathways. *EMBO J.* 17, 1006–1018.
- O'Reilly, K.E., Rojo, F., She, Q.B., Solit, D., Mills, G.B., Smith, D., Lane, H., Hofmann, F., Hicklin, D.J., Ludwig, D.L., et al. (2006). mTOR inhibition induces upstream receptor tyrosine kinase signaling and activates Akt. *Cancer Res.* 66, 1500–1508.
- Rutkowska, A., Haering, C.H., and Schultz, C. (2011). A FIAsh-based cross-linker to study protein interactions in living cells. *Angew. Chem. Int. Ed. Engl.* 50, 12655–12658.
- Strickland, D., Lin, Y., Wagner, E., Hope, C.M., Zayner, J., Antoniou, C., Sosnick, T.R., Weiss, E.L., and Glotzer, M. (2012). TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods* 9, 379–384.
- Varnai, P., Thyagarajan, B., Rohacs, T., and Balla, T. (2006). Rapidly inducible changes in phosphatidylinositol 4,5-bisphosphate levels influence multiple regulatory functions of the lipid in intact living cells. *J. Cell Biol.* 175, 377–382.
- Wullschleger, S., Loewith, R., and Hall, M.N. (2006). TOR signaling in growth and metabolism. *Cell* 124, 471–484.
- Wymann, M.P. (2012). PI3Ks-Drug Targets in Inflammation and Cancer. *Subcell. Biochem.* 58, 111–181.
- Wymann, M.P., and Marone, R. (2005). Phosphoinositide 3-kinase in disease: timing, location, and scaffolding. *Curr. Opin. Cell Biol.* 17, 141–149.
- Wymann, M.P., and Schneider, R. (2008). Lipid signalling in disease. *Nat. Rev. Mol. Cell Biol.* 9, 162–176.
- Wymann, M.P., Zvelebil, M., and Laffargue, M. (2003). Phosphoinositide 3-kinase signalling—which way to target? *Trends Pharmacol. Sci.* 24, 366–376.
- Yazawa, M., Sadaghiani, A.M., Hsueh, B., and Dolmetsch, R.E. (2009). Induction of protein-protein interactions in live cells using light. *Nat. Biotechnol.* 27, 941–945.
- Zoncu, R., Efeyan, A., and Sabatini, D.M. (2011). mTOR: from growth signal integration to cancer, diabetes and ageing. *Nat. Rev. Mol. Cell Biol.* 12, 21–35.

## Supplemental Information

### Chemical Development

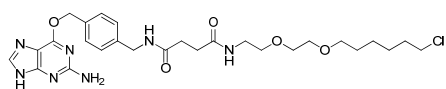
#### of Intracellular Protein Heterodimerizers

Dominik Erhart, Mirjam Zimmermann, Olivier Jacques, Matthias B. Wittwer, Beat Ernst, Edwin Constable, Marketa Zvelebil, Florent Beaufigli, and Matthias P. Wymann

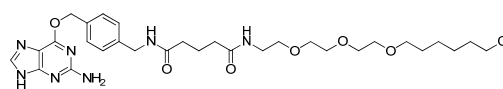
#### Inventory of Supplemental Information

- |  |         |
|--|---------|
| 1. Structures of Key Compounds Used in this Study        | page 2  |
| 2. Supplemental Experimental Procedures and Conditions   | page 2  |
| 3. Chemical Synthesis and Characterization of Compounds  | page 4  |
| 4. Supplemental Data                                     | page 12 |
| Figure S1, related to Figure 1                           |         |
| Figure S2, related to Figure 6                           |         |
| 5. Supplemental Table 1, related to Figure 1             | page 13 |
| 6. Supplemental References                               | page 14 |
| 7. Spectra ( <sup>1</sup> H NMR and <sup>13</sup> C NMR) | page 16 |

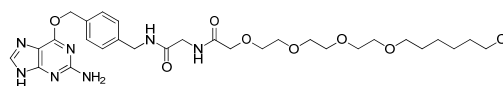
## 1. Structures of Key Compounds Used in this Study



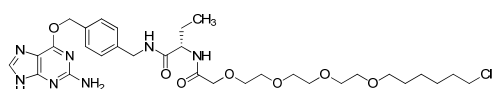
HaXS1



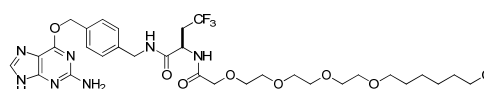
HaXS2



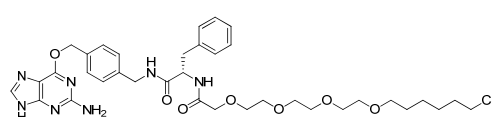
HaXS3



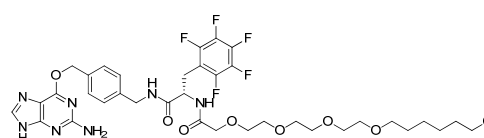
HaXS4



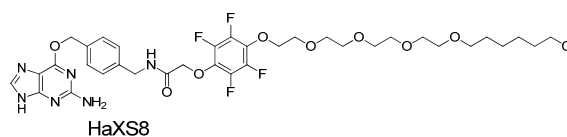
HaXS5



HaXS6



HaXS7



HaXS8

## 2. Supplemental Experimental Procedures and Conditions

### Chemical synthesis and characterization

Materials and reagents were of the highest commercially available grade, and used without further purification. Reactions were monitored by thin layer chromatography using Merck silica gel 60 F254 plates. Compounds were visualized by UV, ceric ammonium molybdate (CAM),  $\text{KMnO}_4$  and ninhydrin. Flash chromatographies were performed using Merck silica gel 60, particle size 40 - 63  $\mu\text{m}$ .  $^1\text{H}$ ,  $^{19}\text{F}$  and  $^{13}\text{C}$  NMR spectra were recorded on a Bruker AV-400 or a DRX-600 NMR spectrometer. Chemical shifts are reported in ppm using the solvent residue signals as reference. All solvents used for reactions were purchased as anhydrous grade from Fluka. Solvents for extractions, column chromatography and TLC were commercial grade. Mass spectra were recorded with a VG70-250 (FAB), Finnigan MAT MS 312 (EI) or Finnigan MAT LCQ (ESI) spectrometer. High resolution mass spectra were recorded with a thermo Fisher Scientific LTQ Orbitrap XL, nanoelectrospray ion source. For spectral data see section 6.

### PAMPA

$P_e$  (expressed as nm/s) was determined with a parallel artificial membrane permeability assay (PAMPA; (Kansy et al., 1998) for a review see (Faller, 2008)). For each compound, measurements were performed in quadruplicate at pH 7.4 according to the manufacturers procedures (Pion, Billerica). Each well of a deep well plate was filled with 650  $\mu\text{l}$  of pH-adjusted System Solution (Pion, P/N 110151). 150  $\mu\text{l}$  of each well were transferred to an UV-plate and analyzed by UV-spectroscopy to determine the blank spectra. 2.5  $\mu\text{L}$  of the according 2 mM compound stock solution in DMSO were added to the remaining System Solution in each well and mixed. To exclude precipitation, the optical density was measured at 650 nm, with 0.01 being the threshold value. Samples of 150  $\mu\text{L}$  were transferred from the deep well plate to another UV-plate to determine the reference spectra. Further

200  $\mu\text{L}$  were transferred to each well of the donor plate of the PAMPA sandwich. The filter membranes at the bottom of the acceptor plate were impregnated with 5  $\mu\text{L}$  of GIT-0 Lipid Solution (Pion, P/N 110669) and 200  $\mu\text{L}$  of Acceptor Sink Buffer (Pion, P/N 110139) were filled into each acceptor well. To start the experiment, the sandwich was assembled, then placed in the GutBox™ and stirred for 30 minutes. To finish the assay, the sandwich was disassembled and 150  $\mu\text{L}$  from each donor and acceptor well were transferred to UV-plates. Quantification was performed by both UV-spectroscopy and LC-MS.  $\log P_e$ -values were calculated based on the LC-MS results and with the aid of the PAMPA Explorer Software (Pion, version 3.5).

#### **Statistical Analysis**

Statistical analysis was performed using GraphPad Prism v5. For Student's t test (two sided, non-paired) at least 3 independent experiments were compared. Where differences are claimed, and not indicated otherwise, p was  $<0.01$ .

#### **Molecular modeling**

Starting points for the modeling process were generated from crystal structures of the Halo-tag (Newman et al., 1999) and the SNAP-tag (Mollwitz et al., 2012). An initial manual docking of HaXS substrates was refined using the genetic algorithm application GOLD (Jones et al., 1997). Covalently linked HaXS substrates and protein structures were subsequently subjected to energy minimization using the program Yasara (Krieger et al., 2004).

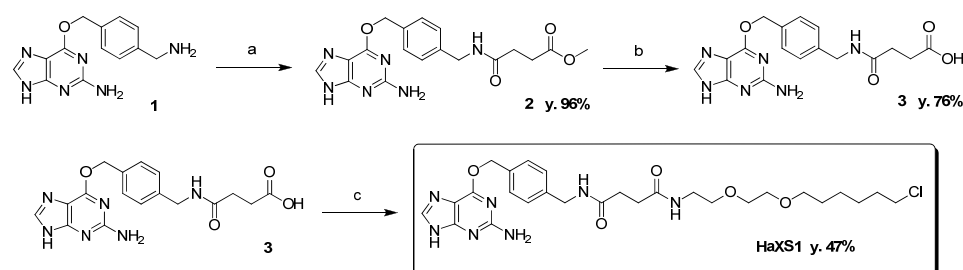


### 3. Chemical Synthesis and Characterization of Compounds

#### General procedure 1 (GP1) for peptide coupling with PYBOP.

PYBOP (0.2 mmol) is added to a solution of carboxylic acid derivative (0.2 mmol) in DMF (2 ml) at rt. The solution is stirred at rt for 2h. Then, amine substrate (0.2 mmol) and DIPEA (0.2 mmol) are added and the solution is stirred at rt overnight. Then the crude mixture is poured onto water (70mL) and aqueous phase is extracted with AcOEt (2\*50mL). The organic layer is washed with water (20mL), dried over sodium sulfate and evaporated under reduced pressure. Flash chromatography (FC) gives the desired compound.

#### Synthesis of HaXS1.



Reaction conditions: a) Methyl succinyl chloride, DIPEA, DMF, 65°C, 16h; b) 2 M NaOH, THF, rt, 1h; c) (1.) PYBOP, DMF, rt, 3h, (2.) 2-(2-(6-chlorohexyloxy)ethoxy)ethanamine hydrochloride, DIPEA, rt, 16h.

#### Scheme S1. Synthetic route to HaXS1.

#### Preparation of methyl 4-(4-((2-amino-7H-purin-6-yloxy)methyl)benzylamino)-4-oxobutanoate (2)

O6-aminomethylbenzylguanine (1) was prepared according to the literature. (Keppler et al., 2003) According to GP1, starting from PYBOP (481 mg, 0.92 mmol), mono-methyl succinate (122 mg, 0.92 mmol), O6-aminomethylbenzylguanine 1 (250 mg, 0.92 mmol) and DIPEA (159  $\mu$ L, 0.92 mmol). The crude mixture was poured into water. The resulting precipitate was collected, washed with water and dried under vacuum to give compound 2 (340 mg, 96%);  $^1\text{H NMR}$  (400 MHz, DMSO- $d_6$ ):  $\delta$  8.38 (t,  $J=5.5\text{Hz}$ , 1H), 7.79 (s, 1H), 7.42 (d,  $J=8.0\text{Hz}$ , 2H), 7.23 (d,  $J=8.0\text{Hz}$ , 2H), 6.55 (s, 2H), 5.43 (s, 2H), 4.23 (d,  $J=5.8\text{Hz}$ , 2H), 3.54 (s, 3H), 2.50 (t,  $J=6.9\text{Hz}$ , 2H), 2.40 (t,  $J=6.9\text{Hz}$ , 2H);  $^{13}\text{C NMR}$  (100.6 MHz, DMSO- $d_6$ ):  $\delta$  172.8, 170.6, 159.3, 155.4, 139.5, 138.9, 134.8, 128.5, 127.1, 66.9, 51.3, 41.8, 29.7, 28.7; **MS** (FAB, NBA):  $m/z$  (%): 385 ( $[\text{M}+\text{H}]^+$ , 100), 234 (26), 202 (16), 152 (43); **HRMS**  $\text{C}_{18}\text{H}_{21}\text{N}_6\text{O}_6$   $[\text{M}+\text{H}]^+$  calcd: 385.1619, found: 385.1614

#### Preparation of 4-(4-((2-amino-7H-purin-6-yloxy)methyl)benzylamino)-4-oxobutanoic acid (3)

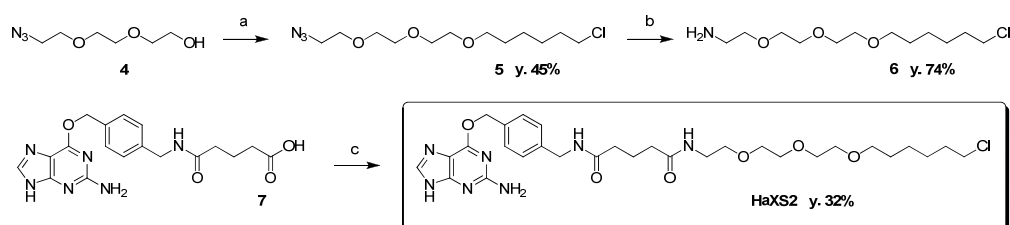
A solution of compound 2 (410 mg, 1.07 mmol) in THF (12 mL) was stirred with NaOH (2 mL, 1M) for 1 h at rt. THF was removed under reduce pressure and pH adjusted to 4.5 by slow addition of acetic acid. The resulting suspension was filtrated, the solid collected and dried under reduce pressure to yield the compound 3 which was used without further purification (300 mg, 76%);  $^1\text{H NMR}$  (400 MHz, DMSO- $d_6$ ):  $\delta$  12.34 (bs, 1H), 8.39 (m, 1H), 7.84 (s, 1H), 7.44 (d,  $J=7.8\text{Hz}$ , 2H), 7.27 (d,  $J=7.8\text{Hz}$ , 2H), 6.29 (s, 2H), 5.46 (s, 2H), 4.27 (m, 2H), 2.48-2.37 m, 4H);  $^{13}\text{C NMR}$  (100.6 MHz, DMSO- $d_6$ ):  $\delta$  174.0, 171.1, 159.7, 139.5, 135.2, 128.5, 127.2, 66.6, 41.9, 30.1, 29.2; **MS** (EI):  $m/z$  (%): 393 ( $[\text{M}+\text{Na}]^+$ , 100), 371 (84), 342 (15); **HRMS**  $\text{C}_{17}\text{H}_{19}\text{N}_6\text{O}_4$   $[\text{M}+\text{H}]^+$  calcd: 371.1462, found: 371.1462

#### Preparation of HaXS1

2-(2-(6-chlorohexyloxy)ethoxy)ethanamine hydrochloride was prepared according to the literature. (Los et al., 2008) According to GP1, starting from PYBOP (419 mg, 0.81 mmol), compound 3 (300 mg, 0.81 mmol), 2-(2-(6-chlorohexyloxy)ethoxy)ethanamine hydrochloride (209 mg, 0.81 mmol) and DIPEA (0.279 mL, 1.62 mmol). FC ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ , 40:1 then 20:1) gave compound HaXS1 (210 mg, 47%).  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3 + 10\% \text{CD}_3\text{OD}$ ):  $\delta$  7.33 (d,  $J=8.3\text{Hz}$ , 2H), 7.30 (s,

1H), 7.17 (d,  $J = 7.8$  Hz, 2H), 5.35 (s, 2H), 4.30 (s, 2H), 3.62 (m, 4H), 3.49 (m, 4H), 3.45-3.42 (m, 4H), 3.38-3.34 (m, 2H), 3.31-3.27 (m, 2H), 2.50-2.43 (m, 4H), 1.71-1.64 (m, 2H), 1.55-1.48 (m, 2H), 1.40-1.23 (m, 4H);  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3 + 10\% \text{CD}_3\text{OD}$ ):  $\delta$  173.0, 172.9, 159.5, 138.5, 135.0, 128.8, 127.7, 71.3, 70.1, 69.1, 69.6, 68.1, 45.0, 43.2, 39.2, 32.4, 31.4, 31.3, 29.2, 26.6, 25.3; **MS** (FAB, NBA):  $m/z$  (%): 576 ( $[\text{M}+\text{H}]^+$ , 100), 306 (54), 202 (44), 152 (94); **HRMS**  $\text{C}_{27}\text{H}_{38}\text{N}_7\text{O}_5\text{ClNa}$   $[\text{M}+\text{Na}]^+$  calcd: 598.2515, found: 598.2509.

### Synthesis of HaXS2.



Reaction conditions: a) (1.) NaH, THF, (2.) 6-chloro-1-iodohexane, rt, 16h; b)  $\text{H}_2$ , Pd/C, MeOH, rt, 2h; c) (1.) PYBOP, DMF, rt, 3h, (2.) **6**, DIPEA, rt, 16h.

### Scheme S2. Synthetic route to HaXS2.

#### Preparation of 1-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)-6-chlorohexane (5)

Compound **4** was prepared according to the literature. (Rensen et al., 2004) Sodium hydride (60% in mineral oil, 914 mg, 22.8 mmol) was added portionwise to a solution of compound **4** (2 g, 11.4 mmol) in THF/DMF (30 mL/10 mL) mixture at 0 °C. After stirring for 30 min, 6-chloro-1-iodohexane (2.6 mL, 17.1 mmol) was added at 0 °C. The mixture was stirred overnight at rt. The excess of sodium hydride was carefully quenched with saturated solution of  $\text{NH}_4\text{Cl}$  (10 mL) and the crude mixture was poured into water (200 mL) and extracted twice with AcOEt (50 mL). The combined organic phase were dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduce pressure. FC (hexane/AcOEt 1:3 then 1:1) yielded compound **5** (1.5 g, 45%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  3.66-3.61 (m, 8H), 3.56-3.48 (m, 2H), 3.50 (t,  $J = 6.6$  Hz, 2H), 3.43 (t,  $J = 6.6$  Hz, 2H), 3.36 (t,  $J = 5.1$  Hz, 2H), 1.78-1.71 (m, 2H), 1.60-1.53 (m, 2H), 1.45-1.31 (m, 4H);  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ ):  $\delta$  71.6, 71.1, 71.1, 71.0, 70.5, 70.5, 51.1, 45.4, 32.9, 30.0, 27.1, 25.8, **MS** (ESI-MS):  $m/z$  (%): 316 ( $[\text{M}+\text{Na}]^+$ , 100); **HRMS**  $\text{C}_{12}\text{H}_{24}\text{ClN}_3\text{O}_3\text{Na}$   $[\text{M}+\text{Na}]^+$  calcd: 316.1398, found: 316.1402.

#### Preparation of 2-(2-(2-(6-chlorohexyloxy)ethoxy)ethoxy)ethanamine (6)

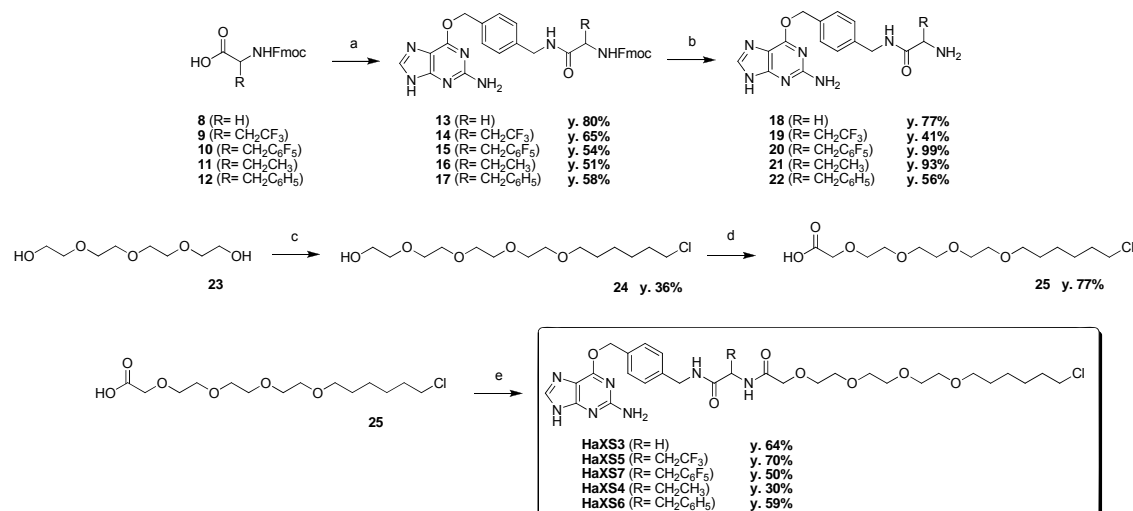
To a solution of compound **5** (1.8 g, 6.1 mmol) in MeOH (20 mL) was added palladium on charcoal (120 mg). The resulting mixture was stirred at rt for 4h under  $\text{H}_2$  atmosphere. The crude mixture was filtrated over celite and concentrated under reduce pressure. FC (DCM/MeOH + 0.5%  $\text{Et}_3\text{N}$ ; 10:1 then 5:1) yielded compound **6** (1.2 g, 74%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  3.66-3.62 (m, 6H), 3.59-3.57 (m, 2H), 3.49-3.51 (m, 4H), 3.46 (t,  $J = 6.6$  Hz, 2H), 2.88 (t,  $J = 6.6$  Hz, 2H), 1.93 (s, 2H), 1.79-1.75 (m, 2H), 1.60-1.57 (m, 2H), 1.46-1.34 (m, 4H);  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CDCl}_3$ ):  $\delta$  71.6, 71.1, 71.0, 70.9, 70.8, 70.5, 70.4, 45.4, 32.9, 29.8, 27.0, 25.8, **MS** (MALDI-TOF):  $m/z$  (%): 268 ( $[\text{M}+\text{H}]^+$ , 100), 290 ( $[\text{M}+\text{Na}]^+$ , 95) **HRMS**  $\text{C}_{12}\text{H}_{27}\text{ClNO}_3$   $[\text{M}+\text{H}]^+$  calcd: 268.1674, found: 268.1675.

#### Preparation of N1-(4-((2-amino-9H-purin-6-yloxy)methyl)benzyl)-N5-(2-(2-(2-(6-chlorohexyloxy)ethoxy)ethoxy)ethyl)glutaramide HaXS2

Compound **7** was prepared according to the literature. (Lemercier et al., 2007) According to GP1, starting from PYBOP (203 mg, 0.39 mmol), compound **7** (150 mg, 0.39 mmol), compound **6** (104 mg, 0.39 mmol) and DIPEA (68  $\mu\text{L}$ , 0.39 mmol). FC ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ , 10:1 then 7.5:1) gave compound **HaXS2** (80 mg, 32%);  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  7.85 (s, 1H), 7.47 (d,  $J = 7.8$  Hz, 2H), 7.28 (d,  $J = 7.8$  Hz, 2H), 5.52 (s, 2H), 4.36 (s, 2H), 3.58-3.49 (m, 12H), 3.42 (t,  $J = 6.6$  Hz, 2H), 3.5-3.30 (m, 3H), 2.25 (m, 4H), 1.91 (m, 2H), 1.72 (m, 2H), 1.54 (m, 2H), 1.44-1.32 (m, 4H);  $^{13}\text{C}$  NMR (100.6 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  174.5, 174.3, 160.6, 160.3, 139.0, 135.9, 128.7, 127.7, 117.7, 110.8, 71.1, 70.6,

70.5, 70.2, 70.1, 69.5, 67.6, 44.7, 42.9, 39.4, 35.1, 29.5, 26.7, 25.5, 22.2; **MS** (ESI-MS):  $m/z$  (%): 634 ( $[M+H]^+$ , 100), 483 (20), 364 (20), 216 (30), 140 (45); **HRMS**  $C_{30}H_{45}ClN_7O_6$   $[M+H]^+$  calcd: 634.3114, found: 634.3125.

### Synthesis of HaXS3, HaXS4, HaXS5, HaXS6 and HaXS7.



Reaction conditions: a) (1.) **8** or **9** or **10** or **11** or **12**, PYBOP, DMF, rt, 3h, (2.) **1**, DIPEA, rt, 16h; b) Diethylamine, DMF, rt, 3h; c) (1.) NaH, THF, (2.) 6-chloro-1-iodohexane, rt, 16h; d) Jones reagent, Acetone, rt, 2h; e) (1.) **25**, PYBOP, DMF, rt, 3h, (2.) **18** or **19** or **20** or **21** or **22**, DIPEA, rt, 16h.

**Scheme S3.** Synthetic route to HaXS3, HaXS4, HaXS5, HaXS6 and HaXS7.

### Preparation of (9H-fluoren-9-yl)methyl-2-(4-((2-amino-9H-purin-6-yloxy)methyl)benzylamino)-2-oxoethylcarbamate (**13**)

According to **GP1**, starting from PYBOP (289 mg, 0.55 mmol), Fmoc-Gly-OH **8** (165 mg, 0.55 mmol), O6-aminomethylbenzylguanine (150 mg, 0.55 mmol) and DIPEA (57  $\mu$ L, 0.55 mmol). FC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1 then 10:1) gave compound **13** (246 mg, 80%). **<sup>1</sup>H NMR** (400MHz, DMSO-d<sub>6</sub>)  $\delta$ : 8.32 (m, 1H), 7.86-7.80 (m, 5H), 7.43-7.38 (m, 4H), 7.36-7.26 (m, 4H), 6.91 (m, 1H), 6.29 (s, 2H), 6.25 (s, 2H), 5.45 (s, 2H), 4.28 (m, 2H), 3.58 (m, 1H); **<sup>13</sup>C NMR** (100.6 MHz, DMSO-d<sub>6</sub>)  $\delta$ : 170.6, 160.5, 158.5, 143.4, 140.3, 140.2, 138.3, 136.0, 129.8, 129.3, 128.2, 122.2, 120.9, 110.6, 67.4, 44.7, 42.7 **MS** (ESI-MS):  $m/z$  (%): 550 ( $[M+H]^+$ , 100), 281 (55); **HRMS** for  $C_{30}H_{28}N_7O_4$   $[M+H]^+$  calcd: 550.2197, found 550.2203.

### Preparation of (9H-fluoren-9-yl)methyl-1-(4-((2-amino-9H-purin-6-yloxy)methyl)benzylamino)-4,4,4-trifluoro-1-oxobutan-2-ylcarbamate (**14**)

According to **GP1**, starting from PYBOP (288 mg, 0.55 mmol), (R,S)-Fmoc-2-amino-4,4,4-trifluoro-butiric acid **9** (210 mg, 0.55 mmol), O6-aminomethylbenzylguanine (150 mg, 0.55 mmol) and DIPEA (96  $\mu$ L, 0.55 mmol). FC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1 then 10:1) gave compound **14** (226 mg, 65%). **<sup>1</sup>H NMR** (400MHz, CDCl<sub>3</sub>)  $\delta$ : 12.42 (m, 1H), 8.65 (t,  $J = 6.1$ Hz, 1H), 7.86 (d,  $J = 7.2$ , 2H), 7.79 (m, 1H), 7.68 (d,  $J = 7.7$ , 2H), 7.42-7.37 (m, 4H), 7.31-7.24 (m, 4H), 6.27 (s, 2H), 5.43 (s, 2H), 4.40-4.18 (m, 5H), 3.26 (d,  $J = 5.5$ , 1H), 2.82 (m, 1H), 2.65 (m, 1H); **<sup>19</sup>F NMR** (100.6 MHz, CDCl<sub>3</sub>)  $\delta$ : -62.5; **<sup>13</sup>C NMR** (100.6 MHz, CDCl<sub>3</sub>)  $\delta$ : 170.5, 160.7, 160.5, 156.6, 156.1, 144.6, 144.5, 141.6, 139.7, 129.3, 128.5, 128.1, 127.90, 126.1, 126.1, 121.0, 114.4, 67.3, 66.7, 50.1, 49.5, 47.5, 43.1; **MS** (ESI-MS):  $m/z$  (%): 632 ( $[M+H]^+$ , 100), 281 (45); **HRMS** for  $C_{32}H_{29}F_3N_7O_4$   $[M+H]^+$  calcd: 632.2228, found: 632.2235.

### Preparation of (S)-(9H-fluoren-9-yl)methyl-1-(4-((2-amino-9H-purin-6-yloxy)methyl)benzylamino)-1-oxo-3-(perfluorophenyl)propan-2-ylcarbamate (**15**)

According to **GP1**, starting from PYBOP (96 mg, 0.19 mmol), (R,S)-Fmoc-2-amino-4,4,4-trifluoro-butiric acid **10** (88 mg, 0.19 mmol), O6-aminomethylbenzylguanine (50 mg, 0.19 mmol) and

DIPEA (32  $\mu$ L, 0.19 mmol). FC ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ , 20:1 then 10:1) gave compound **15** (75 mg, 54%).  **$^1\text{H NMR}$**  (400MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 12.43 (m, 1H), 8.35 (t,  $J=5.3$  Hz, 1H), 7.86 (m, 2H), 7.79 (s, 1H), 7.70 (m, 2H), 7.56 (t,  $J=6.1$  Hz, 1H), 7.44-7.37 (m, 4H), 7.32-7.25 (m, 4H), 6.30 (s, 2H), 5.44 (s, 2H), 4.29 (m, 4H), 4.28 (m, 1H), 3.65 (d,  $J=5.8$  Hz, 2H);  **$^{19}\text{F NMR}$**  (100.6 MHz,  $\text{DMSO-d}_6$ )  $\delta$ : -68.80, -70.69, -73.92;  **$^{13}\text{C NMR}$**  (100.6 MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 170.0, 160.7, 160.5, 157.4, 156.1, 144.7, 141.6, 140.1, 138.7, 136.2, 129.4, 128.5, 128.2, 128.0, 126.1, 121.0, 114.4, 67.4, 66.6, 47.5, 44.4, 42.8; **MS** (ESI-MS):  $m/z$  (%): 730 ( $[\text{M}+\text{H}]^+$ , 100), 280 (70), 258 (97); **HRMS** for  $\text{C}_{37}\text{H}_{29}\text{F}_5\text{N}_7\text{O}_4$   $[\text{M}+\text{H}]^+$  calcd: 730.2196, found: 730.2201.

#### Preparation of (S)-(9H-fluoren-9-yl)methyl-1-(4-((2-amino-9H-purin-6-yloxy)methyl)benzylamino)-1-oxobutan-2-ylcarbamate (16)

According to **GP1**, starting from PYBOP (231 mg, 0.44 mmol), Fmoc-Abu-OH **11** (145 mg, 0.44 mmol), O6-aminomethylbenzylguanine (120 mg, 0.44 mmol) and DIPEA (77  $\mu$ L, 0.44 mmol). FC ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ , 20:1 then 10:1) gave compound **16** (130 mg, 51%).  **$^1\text{H NMR}$**  (400MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 8.43 (t,  $J = 5.8$  Hz, 1H), 7.86 (m, 3H), 7.72 (m, 2H), 7.50-7.37 (m, 4H), 7.38-7.25 (m, 4H), 6.28 (s, 2H), 5.45 (s, 2H), 4.30-4.15 (m, 5H), 3.94 (m, 1H), 3.17 (d,  $J = 3.0$  Hz, 1H), 1.69 (m, 1H), 1.58 (m, 1H), 0.86 (t,  $J = 7.4$  Hz, 3H);  **$^{13}\text{C NMR}$**  (100.6 MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 173.2, 172.7, 160.5, 156.9, 144.8, 144.7, 143.4, 141.6, 140.3, 140.2, 138.3, 136.1, 129.8, 129.4, 129.3, 128.5, 128.2, 128.1, 128.0, 127.9, 126.2, 122.2, 120.9, 120.8, 110.6, 67.4, 66.4, 57.0, 49.5, 47.5, 26.2, 11.4; **MS** (ESI-MS):  $m/z$  (%): 578 ( $[\text{M}+\text{H}]^+$ , 100), 281 (55); **HRMS** for  $\text{C}_{32}\text{H}_{32}\text{N}_7\text{O}_4$   $[\text{M}+\text{H}]^+$  calcd: 578.2510, found: 578.2520.

#### Preparation of (S)-(9H-fluoren-9-yl)methyl-1-(4-((2-amino-9H-purin-6-yloxy)methyl)benzylamino)-1-oxo-3-phenylpropan-2-ylcarbamate (17)

According to **GP1**, starting from PYBOP (96 mg, 0.19 mmol), Fmoc-Phe-OH **12** (72 mg, 0.19 mmol), O6-aminomethylbenzylguanine (50 mg, 0.19 mmol) and DIPEA (32  $\mu$ L, 0.19 mmol). FC ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$ , 20:1 then 10:1) gave compound **17** (70 mg, 58%).  **$^1\text{H NMR}$**  (400MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 12.42 (m, 1H), 8.55 (t, 1H,  $J = 5.5$ Hz), 7.86 (d,  $J = 7.5$  Hz, 2H), 7.82 (s, 1H), 7.63-7.70 (m, 3H), 7.43-7.38 (m, 4H), 7.31-7.16 (m, 8H), 6.31 (s, 2H), 5.47 (s, 2H), 4.34-4.27 (m, 3H), 4.21-4.11 (m, 3H), 3.02 (dd,  $J = 4.5$  Hz; 13.6 Hz, 1H), 3.02 (dd,  $J = 13.8$  Hz; 10.6 Hz, 1H);  **$^{13}\text{C NMR}$**  (100.6 MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 172.2, 160.1, 159.9, 156.7, 155.8, 144.6, 141.5, 141.5, 140.2, 139.0, 135.6, 130.1, 130.0, 128.9, 128.5, 128.1, 127.9, 127.1, 126.2, 126.1, 120.9, 68.1, 66.5, 57.2, 47.5, 42.8, 38.5; **MS** (ESI-MS):  $m/z$  (%): 640 ( $[\text{M}+\text{H}]^+$ , 100), 281 (40); **HRMS** for  $\text{C}_{37}\text{H}_{34}\text{N}_7\text{O}_4$   $[\text{M}+\text{H}]^+$  calcd: 640.2667, found: 640.2673.

#### General procedure 2 (GP2) for Fmoc deprotection with diethylamine.

To a solution of Fmoc derivative (0.39 mmol) in DMF (2mL) is added diethylamine (200 $\mu$ L). The resulting mixture is stirred at rt for 1h. Then the crude mixture is poured onto water (70mL) and aqueous phase is extracted with AcOEt (2\*50mL). The organic layer is washed with water (20mL), dried over sodium sulfate and evaporated under reduced pressure. Then, desired compound is purified by flash chromatography (FC) or precipitated in water.

#### Preparation of 2-amino-N-(4-((2-amino-9H-purin-6-yloxy)methyl)benzyl)acetamide (18)

According to **GP2**, starting from compound **13** (212 mg, 0.39 mmol) and diethylamine (200  $\mu$ L). FC ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  with  $\text{Et}_3\text{N}$  (1%), 10:1 then 5:1) gives the desired compound **18** (96 mg, 77%).  **$^1\text{H NMR}$**  (400MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 8.32 (m, 1H), 7.81 (s, 1H), 7.42 (d,  $J=7.8$  Hz, 2H), 7.26 (d,  $J=7.6$  Hz, 2H), 6.27 (s, 2H), 5.73 (s, 2H), 4.27 (m, 2H), 3.13 (m, 2H).  **$^{13}\text{C NMR}$**  (100.6 MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 160.5, 140.3, 139.3, 136.1, 129.4, 128.2, 67.4, 45.6, 42.5. **MS** (ESI-MS):  $m/z$  (%): 328 ( $[\text{M}+\text{H}]^+$ , 100), 177 (25); **HRMS** for  $\text{C}_{15}\text{H}_{18}\text{N}_7\text{O}_2$   $[\text{M}+\text{H}]^+$  calcd: 328.1516, found: 328.1520.

#### Preparation of (S)-2-amino-N-(4-((2-amino-9H-purin-6-yloxy)methyl)benzyl)-4,4,4-trifluorobutanamide (19)

According to **GP2**, starting from compound **14** (226 mg, 0.36 mmol) and diethylamine (400  $\mu$ L). FC ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  with  $\text{Et}_3\text{N}$  (1%), 10:1 then 5:1) gives the desired compound **19** (60 mg, 41%).  **$^1\text{H NMR}$**  (400MHz,  $\text{DMSO-d}_6$ )  $\delta$ : 12.39 (s, 1H), 8.58 (t,  $J=5.7$  Hz, 1H), 7.78 (s, 1H), 7.42 (d,  $J=7.6$  Hz, 2H), 7.26 (d,  $J=8.1$  Hz, 2H), 6.27 (s, 2H), 5.43 (s, 2H), 4.27 (d,  $J=5.8$  Hz, 2H), 3.48 (dd,  $J = 5.0$  Hz,  $J$

=7.6 Hz 1H), 2.67 (m, 1H), 2.34 (m, 1H); <sup>19</sup>F NMR (100.6 MHz, DMSO-d<sub>6</sub>) δ: -61.73; <sup>13</sup>C NMR (100.6 MHz, DMSO-d<sub>6</sub>) δ: 173.8, 160.5, 140.0, 136.2, 129.4, 129.2, 128.2, 126.4, 67.9, 50.8, 49.5, 42.8, 38.8. **MS** (ESI-MS): *m/z* (%): 410 ([M+H]<sup>+</sup>, 100), 259 (25), 242 (30); **HRMS** for C<sub>17</sub>H<sub>19</sub>F<sub>3</sub>N<sub>7</sub>O<sub>2</sub> [M+H]<sup>+</sup> calcd: 410.1547, found: 410.1549.

#### Preparation of (S)-2-amino-N-(4-((2-amino-9H-purin-6-yloxy)methyl)benzyl)-3-(perfluorophenyl)propanamide (20)

According to **GP2**, starting from compound **15** (71 mg, 0.1 mmol) and diethylamine (200 μL). FC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH with Et<sub>3</sub>N (1%), 10:1 then 5:1) gives the desired compound **20** (49 mg, 99%). <sup>1</sup>H NMR (400MHz, DMSO-d<sub>6</sub>) δ: 12.40 (s, 1H), 8.44 (m, 1H), 7.80 (s, 1H), 7.41 (d, *J*=7.8 Hz, 2H), 7.20 (d, *J*=8.1 Hz, 2H), 6.26 (s, 2H), 5.42 (s, 2H), 4.23 (m, 2H), 3.42 (s, 2H), 2.98 (dd, *J*=6.6 Hz, *J*=13.6 Hz 1H), 2.81 (dd, *J*=7.9 Hz, *J*=13.4 Hz 1H); <sup>19</sup>F NMR (100.6 MHz, DMSO-d<sub>6</sub>) δ: -141.7, -157.9, -163.4; <sup>13</sup>C NMR (100.6 MHz, DMSO-d<sub>6</sub>) δ: 173.8, 160.5, 160.5, 139.9, 136.2, 129.3, 128.2, 113.4, 67.6, 55.1, 42.7, 29.0; **MS** (ESI-MS): *m/z* (%): 508 ([M+H]<sup>+</sup>, 100), **HRMS** for C<sub>22</sub>H<sub>19</sub>F<sub>5</sub>N<sub>7</sub>O<sub>2</sub> [M+H]<sup>+</sup> calcd: 508.1515, found: 508.1516.

#### Preparation of (S)-2-amino-N-(4-((2-amino-9H-purin-6-yloxy)methyl)benzyl)butanamide (21)

According to **GP2**, starting from compound **16** (130 mg, 0.22 mmol) and diethylamine (400 μL). The resulting mixture is precipitated in water (20mL), collected and dried under reduced vacuum to give the desired compound **21** (73 mg, 93%). <sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>) δ: 8.36 (t, *J*=5.5 Hz, 1H), 7.84 (s, 1H), 7.45 (m, 2H), 7.28 (m, 2H), 6.29 (s, 2H), 5.46 (s, 2H), 4.30 (m, 2H), 3.13 (m, 1H), 1.59 (m, 1H), 1.43 (m, 1H), 0.84 (t, *J* = 7.7, 3H). <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) δ: 175.8, 160.5, 140.4, 136.0, 129.4, 128.1, 67.4, 56.9, 49.4, 42.5, 28.9, 11.0. **MS** (ESI-MS): *m/z* (%): 356 ([M+H]<sup>+</sup>, 100), 267 (25); **HRMS** for C<sub>17</sub>H<sub>22</sub>N<sub>7</sub>O<sub>2</sub> [M+H]<sup>+</sup> calcd: 356.1829, found: 356.1841.

#### Preparation of (S)-2-amino-N-(4-((2-amino-9H-purin-6-yloxy)methyl)benzyl)-3-phenylpropanamide (22)

According to **GP2**, starting from compound **17** (188 mg, 0.29 mmol) and diethylamine (400 μL). FC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH with Et<sub>3</sub>N (1%), 10:1 then 5:1) gives the desired compound **22** (68 mg, 56%). <sup>1</sup>H NMR (400MHz, DMSO-d<sub>6</sub>) δ: 8.39 (t, *J* = 6.2, 1H), 7.85 (s, 1H), 7.41 (d, *J* = 8.1, 2H), 7.26-7.14 (m, 7H), 6.31 (s, 2H), 5.47 (s, 2H), 4.26 (m, 2H), 3.50 (dd, *J*=7.7Hz; *J*=13.6Hz, 1H), 2.94 (dd, *J*=5.7Hz; *J*=13.2Hz, 1H), 2.71 (dd, *J*=7.4 Hz, *J*=13.0 Hz, 1H); <sup>13</sup>C NMR (100.6 MHz, DMSO-d<sub>6</sub>) δ: 174.8, 160.5, 140.1, 139.3, 136.0, 130.2, 129.4, 129.00, 128.1, 127.0, 67.4, 57.1, 49.5, 42.6, 41.8. **MS** (ESI-MS): *m/z* (%): 418 ([M+H]<sup>+</sup>, 100), 267 (20), 209 (25); **HRMS** for C<sub>22</sub>H<sub>24</sub>N<sub>7</sub>O<sub>2</sub> [M+H]<sup>+</sup> calcd: 418.1986, found: 418.1992.

#### Preparation of 18-chloro-3,6,9,12-tetraoxaoctadecan-1-ol (24)

Sodium hydride (60% in mineral oil, 2.46 g, 60 mmol) was added portionwise to a solution of tetraethyleneglycol **23** (10 g, 51.4 mmol) in THF (500mL) at 0 °C. After stirring for 30 min, 6-chloro-1-iodohexane (8.6 mL, 56 mmol) was added at 0 °C. The mixture was stirred overnight at rt. The excess of sodium hydride was carefully quenched with a saturated solution of NH<sub>4</sub>Cl (10 mL). Then solvent was evaporated. The crude mixture was poured into water (500 mL) and extracted twice with AcOEt (100 mL). The combined organic phase were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduce pressure. The crude oil was purified by flash column chromatography (DCM/MeOH 30:1) to yield compound **24** (5.8 g, 36%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) : δ 3.7 (s, 1H), 3.63 (s, 2H), 3.58-3.50 (m, 14 H) 3.45 (t, *J* = 6.8Hz, 2H), 3.38 (t, *J* = 6.8Hz, 2H), 2.79 (m, 1H), 1.73-1.66 (m, 2H), 1.56-1.49 (m, 2H), 1.41-1.26 (m, 4H); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>) : δ 73.1, 70.9, 70.8, 70.8, 70.4, 70.4, 61.9, 45.4, 32.9, 29.6, 27.0, 25.7. **MS** (ESI-MS): *m/z* (%): 335 ([M+Na]<sup>+</sup>, 100), 299 (35); **HRMS** C<sub>14</sub>H<sub>29</sub>ClO<sub>5</sub>Na [M+Na]<sup>+</sup> calcd: 335.1596, found: 335.1594.

#### Preparation of 18-chloro-3,6,9,12-tetraoxaoctadecan-1-oic acid (25)

To an ice cooled solution of the alcohol **24** (1.5 g, 4.7 mmol) in acetone (70 mL) was added dropwise a solution of Jones's reagent (7.5 mL prepared according to the procedure described below). The solution was then stirred at rt for 4h. The crude mixture was poured into water (100 mL) and

extracted twice with AcOEt (2x50 mL). The combined organic phase were combined and washed with NaHCO<sub>3</sub> solution. The aqueous phase was acidified to pH 1 with HCl (1M), and was extracted with AcOEt (3x50 mL). The combined organic phase were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduce pressure to yield compound **25** (1.2 g, 77%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) : δ 10.15 (s, 1H), 4.14 (s, 2H), 3.73-3.71 (m, 2H), 3.66-3.60 (m, 8H) 3.57-3.55 (m, 2H), 3.49 (t, *J* = 6.6Hz, 2H), 3.44 (t, *J* = 6.6Hz, 2H), 1.77-1.70 (m, 2H), 1.60-1.53 (m, 2H), 1.45-1.29 (m, 4H); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>): δ 173.8, 71.7, 71.0, 70.9, 70.8, 70.8, 70.8, 70.5, 69.2, 45.4, 32.9, 29.7, 27.0, 25.7. **MS** (ESI-MS): *m/z* (%): 359 ([M+Na]<sup>+</sup>, 100), 313 (20); **HRMS** C<sub>14</sub>H<sub>27</sub>ClO<sub>6</sub>Na [M+Na]<sup>+</sup> calcd: 349.1388, found: 349.1383. Preparation of Jones reagent: Concentrated H<sub>2</sub>SO<sub>4</sub> (230 mL) was added under cooling with an ice bath to CrO<sub>3</sub> (267 g, 2.67 mol) in H<sub>2</sub>O (400 mL). The cold solution is diluted with H<sub>2</sub>O up to 1 L to form the 8 N Jones reagent ready for use.

#### Preparation of HaXS3

According to **GP1**, starting from PYBOP (148 mg, 0.28 mmol), compound **18** (93 mg, 0.28 mmol), compound **25** (93 mg, 0.28 mmol) and DIPEA (50 μL, 0.28 mmol). FC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1 then 10:1) gave compound **HaXS3** (117 mg, 64%). <sup>1</sup>H NMR (400MHz, CD<sub>3</sub>OD+10%CDCl<sub>3</sub>) δ: 7.80 (s, 1H), 7.42 (d, *J*=7.9 Hz, 2H), 7.26 (d, *J*=7.9 Hz, 2H), 5.48 (s, 2H), 4.38 (s, 2H), 4.03 (s, 2H), 3.95 (m, 2H), 3.67-3.65 (m, 2H), 3.62-3.47 (m, 11H), 3.41-3.37 (m, 2H), 1.77-1.66 (m, 2H), 1.56-1.49 (m, 2H), 1.42-1.24 (m, 4H); <sup>13</sup>C NMR (100.6 MHz, CD<sub>3</sub>OD+10%CDCl<sub>3</sub>) δ:172.5, 170.3, 160.6, 138.7, 135.9, 128.7, 127.7, 71.2, 71.0, 70.5, 70.4, 70.2, 70.1, 67.7, 44.8, 42.9, 42.2, 32.7, 29.5, 26.8, 25.5. **MS** (ESI-MS): *m/z* (%): 636 ([M+H]<sup>+</sup>, 100), 404 (45); **HRMS** for C<sub>29</sub>H<sub>43</sub>ClN<sub>7</sub>O<sub>7</sub> [M+H]<sup>+</sup> calcd: 636.2907, found: 636.2918.

#### Preparation of HaXS5

According to **GP1**, starting from PYBOP (33 mg, 0.06 mmol), compound **25** (21 mg, 0.06 mmol), compound **19** (26 mg, 0.06 mmol) and DIPEA (12 μL, 0.06 mmol). FC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1 then 10:1) gave compound **HaXS5** (32 mg, 70%). <sup>1</sup>H NMR (400MHz, CD<sub>3</sub>OD) δ: 7.83 (s, 1H), 7.42 (d, *J*=8.1 Hz, 2H), 7.26 (d, *J*=8.1 Hz, 2H), 5.50 (s, 2H), 4.82 (dd, *J*=4.0 Hz, *J*=9.6 Hz 1H), 4.37 (d, *J*=15.0 Hz, 1H), 4.33 (d, *J*=15.0 Hz, 1H), 4.05 (d, *J*=15.7 Hz, 1H), 3.99 (d, *J*=15.7 Hz, 1H), 3.64-3.48 (m, 12H), 3.38 (m, 2H), 2.88 (m, 1H), 2.70 (m, 1H), 1.70 (m, 2H), 1.51 (m, 2H), 1.43-1.29 (m, 4H). <sup>19</sup>F NMR (100.6 MHz, CD<sub>3</sub>OD) δ: -65.10. <sup>13</sup>C NMR (100.6 MHz, CD<sub>3</sub>OD) δ:171.9, 170.2, 160.6, 138.6, 136.0, 128.7, 127.7, 127.6, 125.3, 71.1, 70.9, 70.5, 70.4, 70.4, 70.3, 70.2, 70.1, 67.6, 46.3, 44.7, 43.1, 35.1, 32.7, 29.5, 26.7, 26.3 25.5. **MS** (ESI-MS): *m/z* (%): 718 ([M+H]<sup>+</sup>, 100), 258 (20); **HRMS** for C<sub>31</sub>H<sub>44</sub>ClF<sub>3</sub>N<sub>7</sub>O<sub>7</sub> [M+H]<sup>+</sup> calcd: 718.2937, found: 718.2939.

#### Preparation of HaXS7

According to **GP1**, starting from PYBOP (52 mg, 0.1 mmol), compound **25** (32 mg, 0.1 mmol), compound **20** (50 mg, 0.1 mmol) and DIPEA (17 μL, 0.1 mmol). FC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1 then 10:1) gave compound **HaXS7** (40 mg, 50%). <sup>1</sup>H NMR (400MHz, CD<sub>3</sub>OD) δ: 7.82 (s, 1H), 7.43 (d, *J*=7.8 Hz, 2H), 7.23 (d, *J*=7.8 Hz, 2H), 5.50 (s, 2H), 4.74 (t, *J*=7.1.0 Hz, 1H), 4.34 (s, 2H), 3.97 (dd, *J*=15.7 Hz, *J*=15.7 Hz, 2H), 3.61 (s, 4H), 3.59-3.56 (m, 6H), 3.50 (m, 4H), 3.40 (m, 2H), 3;30 (m, 1H), 3.10 (m, 1H), 1.67 (m, 2H), 1.51 (m, 2H), 1.43-1.29 (m, 4H). <sup>19</sup>F NMR (100.6 MHz, CD<sub>3</sub>OD) δ: -143.3, -158.6, -164.8; <sup>13</sup>C NMR (100.6 MHz, CD<sub>3</sub>OD) δ: 171.6, 170.5, 160.6, 160.3, 147.2, 144.8, 144.7, 139.0, 138.5, 136.1, 128.7, 127.9, 111.3, 71.1, 71.0, 70.5, 70.5, 70.5, 70.4, 70.1, 67.6, 52.0, 44.7, 43.0, 32.7, 29.5, 26.7, 25.5, 25.4. **MS** (ESI-MS): *m/z* (%): 816 ([M+H]<sup>+</sup>, 100), 404 (20), 258 (35); **HRMS** for C<sub>36</sub>H<sub>44</sub>ClF<sub>5</sub>N<sub>7</sub>O<sub>7</sub> [M+H]<sup>+</sup> calcd: 816.2905, found: 816.2922.

#### Preparation of HaXS4

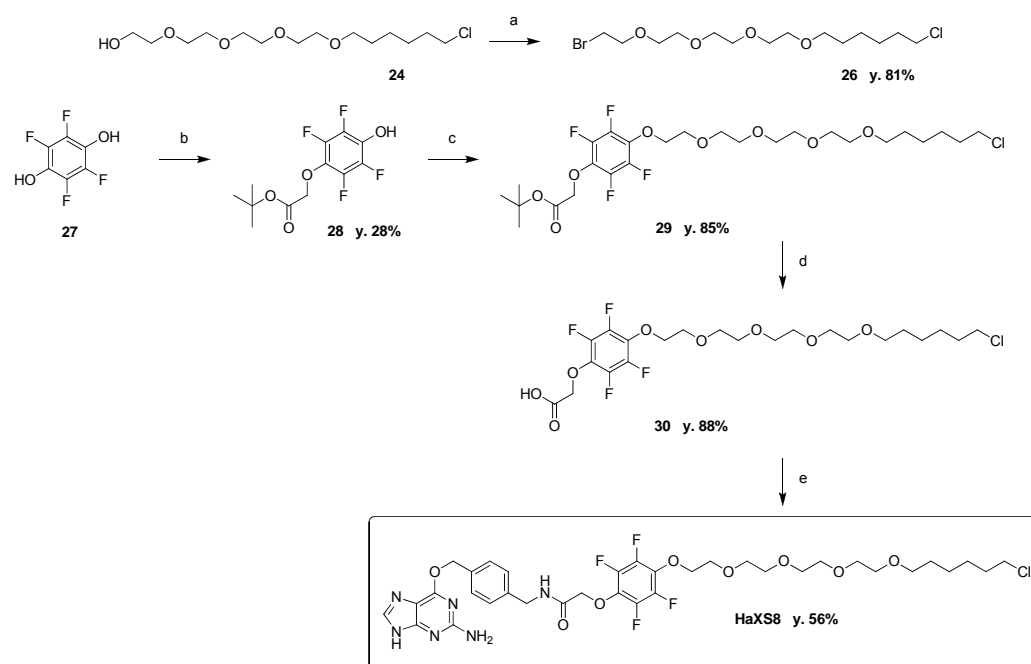
According to **GP1**, starting from PYBOP (107 mg, 0.20 mmol), compound **25** (67 mg, 0.20 mmol), compound **21** (73 mg, 0.20 mmol) and DIPEA (36 μL, 0.20 mmol). FC (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1 then 10:1) gave compound **HaXS4** (40 mg, 30%). <sup>1</sup>H NMR (400MHz, CD<sub>3</sub>OD) δ: 7.83 (s, 1H), 7.45 (d, *J*=7.8 Hz, 2H), 7.27 d, *J*=7.8 Hz, 2H), 5.50 (s, 2H), 4.38 (s, 2H), 4.49 (dd, *J*=5.8 Hz; *J*=8.1 Hz, 1H), 4.02 (s, 2H), 3.67-3.48 (m, 14H), 3.42 (m, 2H), 3.30 (m, 1H), 1.86 (m, 1H), 1.71 (m, 3H), 1.52 (m, 2H), 1.45-1.30 (m, 4H), 0.94 (t, *J*=7.5, 3H); <sup>13</sup>C NMR (100.6 MHz, CD<sub>3</sub>OD) δ: 172.8, 171.7, 160.7, 138.8, 136.0, 128.7, 127.7, 71.1, 71.0, 70.5, 70.5, 70.5, 70.4, 70.1, 67.6, 54.8, 44.8, 42.8, 32.7, 29.5, 26.7,

25.7, 25.5, 9.7; **MS** (ESI-MS):  $m/z$  (%): 664 ( $[M+H]^+$ , 100), 620 (20); **HRMS** for  $C_{31}H_{47}ClN_7O_7$   $[M+H]^+$  calcd: 664.3220, found: 664.3228.

#### Preparation of HaXS6

According to **GP1**, starting from PYBOP (81 mg, 0.16 mmol), compound **25** (51 mg, 0.16 mmol), compound **22** (65 mg, 0.16 mmol) and DIPEA (27  $\mu$ L, 0.16 mmol). FC ( $CH_2Cl_2/MeOH$ , 20:1 then 10:1) gave compound **HaXS6** (69 mg, 59%). **<sup>1</sup>H NMR** (400MHz,  $DMSO-d_6$ )  $\delta$ : 12.43 (s, 1H), 8.59 (t,  $J=5.4$ , 1H), 7.81 (s, 1H), 7.70 (d,  $J=8.5$  Hz, 2H), 7.43 (d,  $J=8.0$  Hz, 2H), 7.25-7.16 (m, 7H), 6.30 (s, 2H), 5.46 (s, 2H), 4.62 (m, 1H), 4.29 (m, 2H), 3.88 (d,  $J=15.3$  Hz, 1H), 3.77 (d,  $J=15.5$  Hz, 1H), 3.59 (m, 2H), 3.52 (m, 10H), 3.36 (m, 3H), 3.17 (d,  $J=4.9$ , 1H), 3.03 (dd,  $J=5.0$  Hz;  $J=13.6$  Hz, 1H), 2.90 (dd,  $J=9.1$  Hz;  $J=13.6$  Hz, 1H), 1.67 (m, 2H), 1.46 (m, 2H), 1.35 (m, 2H), 1.27 (m, 2H); **<sup>13</sup>C NMR** (100.6 MHz,  $DMSO-d_6$ )  $\delta$ : 171.5, 169.8, 160.5, 139.8, 138.4, 136.1, 130.1, 130.1, 129.4, 128.9, 128.1, 127.2, 71.0, 71.0, 70.7, 70.6, 70.5, 70.4, 70.3, 67.3, 54.2, 49.5, 46.2, 42.8, 38.6, 32.9, 29.9, 27.0, 25.8; **MS** (ESI-MS):  $m/z$  (%): 726 ( $[M+H]^+$ , 100), 428 (15); **HRMS** for  $C_{36}H_{49}ClN_7O_7$   $[M+H]^+$  calcd: 726.3377, found: 726.3384.

#### Synthesis of HaXS8



Reaction conditions: a)  $PPh_3$ ,  $CBr_4$ , THF, rt, ov.; b)  $K_2CO_3$ , tert-butyl bromoacetate, DMF, rt, 4h; c)  $K_2CO_3$ , compound **34**, DMF, 50°C, ov.; d) TFA,  $CH_2Cl_2$ , rt, 1h; e) (1.) PYBOP, DMF, rt, 3h, (2.) **38**, DIPEA, rt, 16h.

**Scheme S4.** Synthetic route to HaXS8.

#### Preparation of 1-bromo-18-chloro-3,6,9,12-tetraoxaoctadecane (26)

To a solution of compound **24** (3.7 g, 11.9 mmol) in THF (80 mL) at 0°C are added portionwise triphenylphosphine (3.6 g, 13.6 mmol) and then, carbon tetrabromide (4.5 g, 13.6 mmol). The resulting mixture is stirred at rt overnight. Solvent is evaporated under reduced pressure and the crude oil is purified by flash column chromatography. FC (cyclohexane/AcOEt 3:1) gave compound **26** (3.6 g, 81%). **<sup>1</sup>H NMR** (400 MHz,  $CDCl_3$ )  $\delta$ : 3.80 (m, 2H), 3.68-3.62 (m, 10H), 3.59-3.46 (m, 8H), 1.77 (m, 2H), 1.58 (m, 2H), 1.36-1.52 (m, 4H); **<sup>13</sup>C NMR** (100.6 MHz,  $CDCl_3$ )  $\delta$ : 72.3, 72.1, 71.6, 71.5, 71.5, 71.5, 71.4, 71.1, 45.7, 33.7, 31.4, 30.5, 27.7, 26.5; **MS** (ESI-MS):  $m/z$  (%): 399.07 ( $[M+Na]^+$ , 100); **HRMS**  $C_{14}H_{28}BrClO_4Na$   $[M+Na]^+$  calcd: 399.0752, found: 399.0746.

#### Preparation of tert-butyl 2-(2,3,5,6-tetrafluoro-4-hydroxyphenoxy)acetate (28)

To a mixture of tetrafluorohydroquinone **27** (4 g, 22 mmol) and  $K_2CO_3$  (3 g, 22 mmol) in DMF (100 mL) was added dropwise a solution of tert-butyl bromoacetate (3.2 mL, 22 mmol). The resulting mixture was stirred at rt for 4h. The crude mixture was poured into a saturated solution of  $NH_4Cl$  (900 mL) and extracted twice with AcOEt (200 mL). The combined organic phase were washed with NaCl solution, dried over  $Na_2SO_4$  and concentrated under reduce pressure. FC (cyclohexane/AcOEt 3:1) gave compound **28** (1.83 g, 28%).  $^1H$  NMR (400 MHz, DMSO- $d_6$ ):  $\delta$  11.00 (s, 1H), 4.68 (s, 2H), 1.40 (m, 9H);  $^{13}C$  NMR (100.6 MHz, DMSO- $d_6$ ):  $\delta$  167.1, 81.8, 70.3, 27.6;  $^{19}F$  NMR (100.6 MHz, DMSO- $d_6$ ):  $\delta$ : -158.6, -162.4; MS (ESI-MS):  $m/z$  (%): 319.05 ( $[M+Na]^+$ , 100); HRMS  $C_{12}H_{12}F_4O_4Na$   $[M+Na]^+$  calcd: 319.0564, found: 319.0571.

#### Preparation of tert-butyl-2-(4-(18-chloro-3,6,9,12-tetraoxaocadecyloxy)-2,3,5,6-tetrafluorophenoxy)acetate (29)

A mixture of compound **28** (1.29 g, 4.8 mmol),  $K_2CO_3$  (603 mg, 4.8 mmol) and compound **26** (1.64 g, 4.8 mmol) in DMF (5 mL) was stirred at rt overnight. The crude mixture was poured into a saturated solution of  $NH_4Cl$  (100 mL) and extracted twice with AcOEt (150 mL). The combined organic phase were combined and washed with NaCl solution, dried over  $Na_2SO_4$  and concentrated under reduce pressure. FC (cyclohexane/AcOEt 1:1) gave compound **29** (2.2 g, 85%).  $^1H$  NMR (400 MHz,  $CD_3OD$ ):  $\delta$  4.72 (s, 2H), 4.30 (m, 2H), 3.81 (m, 2H), 3.67-3.60 (m, 10H), 3.59-3.55 (m, 4H), 3.50-3.45 (m, 2H), 1.76 (m, 2H), 1.57 (m, 2H), 1.48 (m, 9H), 1.35-1.53 (m, 4H);  $^{13}C$  NMR (100.6 MHz,  $CD_3OD$ ):  $\delta$  168.8, 83.6, 75.6, 72.1, 71.7, 71.6, 71.5, 71.5, 71.4, 71.2, 71.2, 71.1, 45.7, 33.7, 30.5, 28.3, 27.7, 26.5;  $^{19}F$  NMR (100.6 MHz,  $CD_3OD$ ):  $\delta$ : -161.3, -161.5; MS (ESI-MS):  $m/z$  (%): 613.21 ( $[M+Na]^+$ , 100); HRMS  $C_{26}H_{39}F_4ClO_8Na$   $[M+Na]^+$  calcd: 613.2162, found: 613.2167.

#### Preparation of 2-(4-(18-chloro-3,6,9,12-tetraoxaocadecyloxy)-2,3,5,6-tetrafluorophenoxy)acetic acid (30)

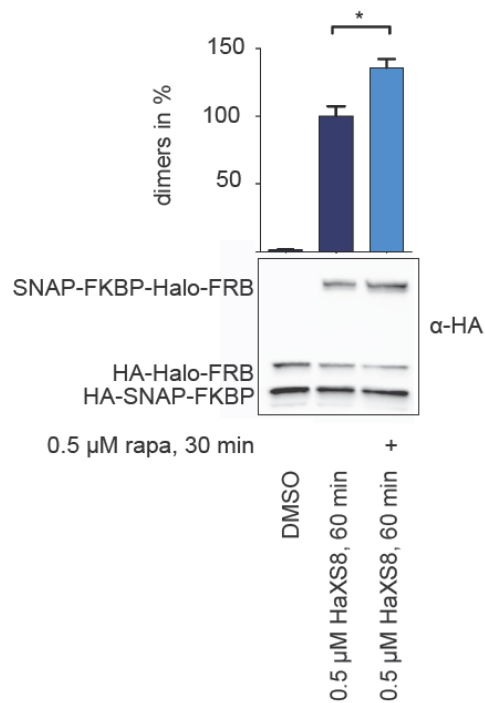
To a solution of compound **29** (2.2g, 3.7 mmol) in DCM (5 mL) was added TFA (5 mL). The mixture was stirred at rt for 2h. Then, solvent was removed under reduce pressure. FC (DCM/MeOH 20:1) gave compound **30** (1.75 g, 88%).  $^1H$  NMR (400 MHz,  $CD_3OD$ ):  $\delta$ : 4.80 (s, 2H), 4.30 (m, 2H), 3.81 (m, 2H), 3.65-3.59 (m, 10H), 3.59-3.55 (m, 4H), 3.47 (m, 2H), 1.76 (m, 2H), 1.58 (m, 2H), 1.35-1.52 (m, 4H);  $^{13}C$  NMR (100.6 MHz,  $CD_3OD$ ):  $\delta$  171.3, 75.6, 72.1, 71.7, 71.6, 71.5, 71.4, 71.1, 70.6, 45.7, 33.7, 30.5, 27.7, 26.5;  $^{19}F$  NMR (100.6 MHz,  $CD_3OD$ ):  $\delta$ : -159.8, -160.0; MS (ESI-MS):  $m/z$  (%): 535.17 ( $[M+H]^+$ , 100); HRMS  $C_{22}H_{32}F_4ClO_8$   $[M+H]^+$  calcd: 535.1716, found: 535.1724.

#### Preparation of HaXS8

A mixture of compound **30** (1.35 g, 2.3 mmol), O6-aminomethylbenzylguanine **1** (626 mg, 2.3 mmol), 1-hydroxybenzotriazole hydrate (313 mg, 2.3 mmol), N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (446 mg, 2.3 mmol) and DIPEA (404  $\mu$ L, 2.3 mmol) in DMF (20 mL) was stirred at rt overnight. The crude mixture was poured onto water (150mL) and aqueous phase is extracted with AcOEt (2\*80mL). The combined organic phase were combined and washed with NaCl solution, dried over  $Na_2SO_4$  and concentrated under reduce pressure. FC ( $CH_2Cl_2$ /MeOH, 20:1 then 10:1) gave **HaXS8** (1.02 g, 56%).  $^1H$  NMR (400 MHz,  $CD_3OD$ ):  $\delta$  7.85 (s, 1H), 7.50 (m, 2H), 7.31 (m, 2H), 5.53 (s, 2H), 4.67 (s, 2H), 4.48 (s, 2H), 4.29 (m, 2H), 3.78 (m, 2H), 3.66-3.60 (m, 10H), 3.58 (m, 4H), 3.47 (m, 2H), 1.74 (m, 2H), 1.54 (m, 2H), 1.31-1.48 (m, 4H);  $^{13}C$  NMR (100.6 MHz,  $CD_3OD$ ):  $\delta$  169.8, 161.6, 139.6, 137.0, 129.7, 128.7, 75.7, 73.9, 72.1, 71.7, 71.6, 71.5, 71.1, 68.7, 45.7, 43.5, 33.8, 30.6, 27.7, 26.5;  $^{19}F$  NMR (100.6 MHz,  $CD_3OD$ ):  $\delta$ : -159.2, -159.6; MS (ESI-MS):  $m/z$  (%): 787.28 ( $[M+H]^+$ , 100); HRMS  $C_{35}H_{44}F_4N_6ClO_8$   $[M+H]^+$  calcd: 787.2840, found: 787.2825.

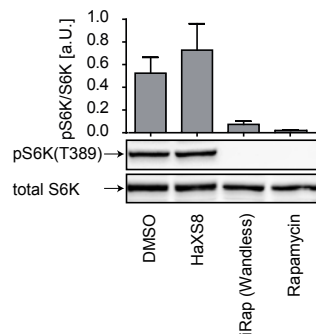


#### 4. Supplemental Data



**Supplemental Figure S1**, related to Figure 1 and Figure 4. HeLa cells co-expressing HA-SNAP-FKBP and HA-Halo-FRB were pretreated with either DMSO or 0.5 μM rapamycin for 30 min., before DMSO or 0.5 μM HaXS8 were added for 60 min. Cells were subsequently analyzed for the intracellular formation of HA-tagged Halo/SNAP-tag dimers using anti-HA antibodies. Dimers forced to form by HaXS8 in the absence of rapamycin were normalized to 100% (Mean ± SEM, one tailed Student's t-test, \* $p < 0.05$ ,  $n = 3$ ).

This experiment demonstrates once more that HaXS8 is capable to force the proximity of previously non-interacting proteins. If the HA-SNAP-FKBP and HA-Halo-FRB fusion proteins are pre-associated by the addition of rapamycin (to form a HA-SNAP-FKBP/rapamycin/HA-Halo-FRB complex), before HaXS8 is added to link the Halo- and SNAP-tags covalently, a modest, but significant increase in dimerization was observed.



**Supplemental Figure S2**, related to Figure 6. HEK293 cells were incubated with DMSO, 0.5 μM HaXS8, iRap (kindly provided by Tom Wandless, Stanford) or rapamycin for 1 h, before cells were lysed, and proteins were subjected to SDS-PAGE and immuno-blotting using antibodies detecting total p70<sup>S6K</sup> (S6K), and phosphorylated S6K (pS6K, T389 residue phosphorylated by TORC1).

5. Supplemental Table 1 – Properties of Substrates Used for Chemically Induced Dimerization

Compound	Biochemical monitoring of complex	Cell permeant	Intracellular dimerization*			Estimated speed of dimerization	Interference with endogenous signaling pathways	Orthogonal use / applications
			y/n	mode	type			
Abscisic acid (ABA)	no	yes	yes	forced	hetero; non-covalent	slow [ $<h$ ], high concentrations required	None reported, but toxic effects <i>in vivo</i> (Celik et al., 2007).	Compatible with rapalogs; regulation of coupled transcription processes (Liang et al., 2011; Hathaway et al., 2012).
CoDis (BG-BG)	yes	(yes)	yes	pre-ass.	homo; covalent	slow [h]	None reported. Tag substrates not expected to interfere with cellular signaling.	Detection of existing protein-protein interactions (Lemercier et al., 2007).
Coumermycin	no	yes	yes	forced	homo; non-covalent	fast [min]	None reported.	Translocation and activation of signaling molecules (Farrar et al., 1996; Farrar et al., 2000).
Dex-FK506	no	yes	yes	forced	hetero; non-covalent	very slow [days ?]	Not assessed in detail. Likely with nuclear receptors and FKBP.	Yeast three hybrid system (Licitra and Liu, 1996).
Dexamethasone-Methotrexate (Dex-Mtx)	no	yes	yes	forced	hetero; non-covalent	very slow [days ?]	Not assessed in detail. Likely with nuclear receptors.	Transcription reporter systems and yeast complementation screens (Lin et al., 2000; Baker et al., 2002).
Dexamethasone-Trimethoprim (Dex-TMP)	no	yes	yes	forced	hetero; non-covalent	very slow [days ?]	Not assessed in detail. Likely with nuclear receptors.	Transcription factor logics in a three hybrid system (Gallagher et al., 2007; Bronson et al., 2008).
Gibberellic acid derivatives (e.g. GA3-AM)	no	yes	yes	forced	hetero; non-covalent	very fast [s]	Cellular acidification at elevated concentrations (Celik et al., 2007).	Compatible with rapalog system; control of Jak/Stat signaling (Mohi et al., 1998; O'Farrell et al., 1998), and protein translocations (Miyamoto et al., 2012; Phua et al., 2012).
HaXS8	yes	yes	yes	forced	hetero; covalent	fast [min]	None observed. Tag substrates not expected to interfere with cellular signaling.	Forced protein translocation, activation of signaling pathways, orthogonal and multiplexed applications, compatible with rapalog system. Simple assessment of success of dimerization (Erhart et al., this work).
Rapalogs	no	yes	yes	forced	hetero; non-covalent	very fast [s]	Interference with mTOR signaling, and FKBP-dependent pathways. Problems with low level rapamycin contamination.	Rapid translocation of signaling molecules to membranes, successful depletion of lipid pools (Liberles et al., 1997; Inoue et al., 2005; Varnai et al., 2006; Edwards and Wandless, 2007). Orthogonal applications possible (Bayle et al., 2006; Inoue and Meyer, 2008). For a review see (Putyrski and Schultz, 2012).
S-Cross (e.g. SC-Cy5)	yes	no	no	pre-ass.	hetero; covalent	slow [h]	Not expected, tag substrates not expected to interfere with cellular signaling.	Detection of protein-protein interactions in cell lysates; integration of fluorescent tracer (Gautier et al., 2009).
X-CrAsH	yes	yes	yes	pre-ass.	(homo); covalent	slow [h]	Redox state-dependent accumulation of FIAsH derivatives in mitochondria reported (Langhorst et al., 2006).	Dimerisation of pre-associated partners (Rutkowska et al., 2011).

\* *forced* intracellular dimerization triggers complex formation of two previously non-associated protein partners, while *pre-ass.* points to dimerization of preformed complexes. Complex formation can be mediated by symmetric interactions (homo-dimerization) or asymmetric interactions (hetero-dimerization), and can be mediated by *non-covalent* binding or *covalent* reactions.

## 6. Supplemental References

- Baker, K., Blecziński, C., Lin, H., Salazar-Jimenez, G., Sengupta, D., Krane, S., Cornish, V. W. (2002). Chemical complementation: a reaction-independent genetic assay for enzyme catalysis. *Proc Natl Acad Sci U S A* 99, 16537-16542.
- Bayle, J. H., Grimley, J. S., Stankunas, K., Gestwicki, J. E., Wandless, T. J., Crabtree, G. R. (2006). Rapamycin analogs with differential binding specificity permit orthogonal control of protein activity. *Chem Biol* 13, 99-107.
- Bronson, J. E., Mazur, W. W., Cornish, V. W. (2008). Transcription factor logic using chemical complementation. *Mol Biosyst* 4, 56-58.
- Celik, I., Tuluçe, Y., Isik, I. (2007). Evaluation of toxicity of abscisic acid and gibberellic acid in rats: 50 days drinking water study. *J Enzyme Inhib Med Chem* 22, 219-226.
- Edwards, S. R., Wandless, T. J. (2007). The rapamycin-binding domain of the protein kinase mammalian target of rapamycin is a destabilizing domain. *J Biol Chem* 282, 13395-13401.
- Faller, B. (2008). Artificial membrane assays to assess permeability. *Curr Drug Metab* 9, 886-892.
- Farrar, M. A., Alberol-Ila, J., Perlmutter, R. M. (1996). Activation of the Raf-1 kinase cascade by coumermycin-induced dimerization. *Nature* 383, 178-181.
- Farrar, M. A., Olson, S. H., Perlmutter, R. M. (2000). Coumermycin-induced dimerization of GyrB-containing fusion proteins. *Methods Enzymol* 327, 421-429.
- Gallagher, S. S., Miller, L. W., Cornish, V. W. (2007). An orthogonal dexamethasone-trimethoprim yeast three-hybrid system. *Anal Biochem* 363, 160-162.
- Gautier, A., Nakata, E., Lukinavicius, G., Tan, K. T., Johnsson, K. (2009). Selective cross-linking of interacting proteins using self-labeling tags. *J Am Chem Soc* 131, 17954-17962.
- Hathaway, N. A., Bell, O., Hodges, C., Miller, E. L., Neel, D. S., Crabtree, G. R. (2012). Dynamics and memory of heterochromatin in living cells. *Cell* 149, 1447-1460.
- Inoue, T., Heo, W. D., Grimley, J. S., Wandless, T. J., Meyer, T. (2005). An inducible translocation strategy to rapidly activate and inhibit small GTPase signaling pathways. *Nat Methods* 2, 415-418.
- Inoue, T., Meyer, T. (2008). Synthetic activation of endogenous PI3K and Rac identifies an AND-gate switch for cell polarization and migration. *PLoS One* 3, e3068.
- Jones, G., Willett, P., Glen, R. C., Leach, A. R., Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 267, 727-748.
- Kansy, M., Senner, F., Gubernator, K. (1998). Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J Med Chem* 41, 1007-1010.
- Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat Biotechnol* 21, 86-89.
- Krieger, E., Darden, T., Nabuurs, S. B., Finkelstein, A., Vriend, G. (2004). Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins* 57, 678-683.
- Langhorst, M. F., Genisyuerk, S., Stuermer, C. A. (2006). Accumulation of FIAsh/Lumio Green in active mitochondria can be reversed by beta-mercaptoethanol for specific staining of tetracysteine-tagged proteins. *Histochem Cell Biol* 125, 743-747.
- Lemercier, G., Gendreizig, S., Kindermann, M., Johnsson, K. (2007). Inducing and sensing protein-protein interactions in living cells by selective cross-linking. *Angew Chem Int Ed Engl* 46, 4281-4284.
- Liang, F. S., Ho, W. Q., Crabtree, G. R. (2011). Engineering the ABA plant stress pathway for regulation of induced proximity. *Sci Signal* 4, rs2.
- Liberles, S. D., Diver, S. T., Austin, D. J., Schreiber, S. L. (1997). Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen. *Proc Natl Acad Sci U S A* 94, 7825-7830.
- Licitra, E. J., Liu, J. O. (1996). A three-hybrid system for detecting small ligand-protein receptor interactions. *Proc Natl Acad Sci U S A* 93, 12817-12821.
- Lin, H., Abida, W. M., Sauer, R. T., Cornish, V. W. (2000). Dexamethasone-Methotrexate: An Efficient Chemical Inducer of Protein Dimerization. *J. Am. Chem. Soc.* 122, 4247-4248.
- Los, G. V. et al. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* 3, 373-382.
- Miyamoto, T., DeRose, R., Suarez, A., Ueno, T., Chen, M., Sun, T. P., Wolfgang, M. J., Mukherjee, C., Meyers, D. J., Inoue, T. (2012). Rapid and orthogonal logic gating with a gibberellin-induced dimerization system. *Nat Chem Biol* 8, 465-470.

- Mohi, M. G., Arai, K., Watanabe, S. (1998). Activation and functional analysis of Janus kinase 2 in BA/F3 cells using the coumermycin/gyrase B system. *Mol Biol Cell* 9, 3299-3308.
- Mollwitz, B., Brunk, E., Schmitt, S., Pojer, F., Bannwarth, M., Schiltz, M., Rothlisberger, U., Johnsson, K. (2012). Directed evolution of the suicide protein O(6)-alkylguanine-DNA alkyltransferase for increased reactivity results in an alkylated protein with exceptional stability. *Biochemistry* 51, 986-994.
- Newman, J., Peat, T. S., Richard, R., Kan, L., Swanson, P. E., Affholter, J. A., Holmes, I. H., Schindler, J. F., Unkefer, C. J., Terwilliger, T. C. (1999). Haloalkane dehalogenases: structure of a *Rhodococcus* enzyme. *Biochemistry* 38, 16105-16114.
- O'Farrell, A. M., Liu, Y., Moore, K. W., Mui, A. L. (1998). IL-10 inhibits macrophage activation and proliferation by distinct signaling mechanisms: evidence for Stat3-dependent and -independent pathways. *EMBO J* 17, 1006-1018.
- Phua, S. C., Pohlmeier, C., Inoue, T. (2012). Rapidly relocating molecules between organelles to manipulate small GTPase activity. *ACS Chem Biol* 7, 1950-1955.
- Putyrski, M., Schultz, C. (2012). Protein translocation as a tool: The current rapamycin story. *FEBS Lett* 586, 2097-2105.
- Rensen, P. C., van Leeuwen, S. H., Sliedregt, L. A., van Berkel, T. J., Biessen, E. A. (2004). Design and synthesis of novel N-acetylgalactosamine-terminated glycolipids for targeting of lipoproteins to the hepatic asialoglycoprotein receptor. *J Med Chem* 47, 5798-5808.
- Rutkowska, A., Haering, C. H., Schultz, C. (2011). A FIAsh-based cross-linker to study protein interactions in living cells. *Angew Chem Int Ed Engl* 50, 12655-12658.
- Varnai, P., Thyagarajan, B., Rohacs, T., Balla, T. (2006). Rapidly inducible changes in phosphatidylinositol 4,5-bisphosphate levels influence multiple regulatory functions of the lipid in intact living cells. *J Cell Biol* 175, 377-382.

## Cell-Permeant and Photocleavable Chemical Inducer of Dimerization\*\*

Mirjam Zimmermann, Ruben Cal, Elia Janett, Viktor Hoffmann, Christian G. Bochet, Edwin Constable, Florent Beauflis,\* and Matthias P. Wymann\*

**Abstract:** Chemical inducers of dimerization (CIDs) have been developed to orchestrate protein dimerization and translocation. Here we present a novel photocleavable HaloTag- and SNAP-tag-reactive CID (MeNV-HaXS) with excellent selectivity and intracellular reactivity. Excitation at 360 nm cleaves the methyl-6-nitroveratryl core of MeNV-HaXS. MeNV-HaXS covalently links HaloTag- and SNAP-tag fusion proteins, and enables targeting of selected membranes and intracellular organelles. MeNV-HaXS-mediated translocation has been validated for plasma membrane, late endosomes, lysosomes, Golgi, mitochondria, and the actin cytoskeleton. Photocleavage of MeNV-HaXS liberates target proteins and provides access to optical manipulation of protein relocation with high spatiotemporal and subcellular precision. MeNV-HaXS supports kinetic studies of protein dynamics and the manipulation of subcellular enzyme activities, which is exemplified for Golgi-targeted cargo and the assessment of nuclear import kinetics.

Localization of signaling enzymes is key to controlling protein and lipid kinase cascades in physiology and disease.<sup>[1]</sup> Control of protein localization and enzyme activity by illumination provides unique access to the manipulation of biological processes in living cells with high spatiotemporal

precision. Caged small molecules and enzyme substrates have been developed for a number of applications.<sup>[2]</sup>

Naturally occurring light-sensitive protein domains have been used to design genetically encoded light-controlled protein–protein interaction modules. These so-called optogenetic systems contain a photoisomerizable chromophore, which undergoes a conformational change upon illumination at a defined wavelength. Optogenetic systems have been used to control the activation of single signaling proteins by protein caging (light-inducible GTPase Rac),<sup>[3]</sup> or in a more modular approach to indirectly manipulate cellular signaling, through the light-dependent dimerization of two protein modules.<sup>[4]</sup> Optogenetic light-activated dimerization systems are versatile tools, but suffer from several drawbacks such as large photosensory protein tags,<sup>[4,5]</sup> the requirement of exogenous cofactors,<sup>[4]</sup> slow kinetics,<sup>[5]</sup> formation of unwanted homodimers,<sup>[6]</sup> and sensitivity to environmental light, and/or overlap with excitation wavelength of popular fluorescent reporter proteins.<sup>[6,7]</sup> Another approach to control protein localization and enzyme activity are chemical inducers of dimerization (CIDs)<sup>[8]</sup> and self-localizing ligands,<sup>[9]</sup> which have been successfully used to manipulate signaling pathways including phosphoinositide turnover,<sup>[10]</sup> and small GTPases.<sup>[11]</sup>

Presently, cell-permeable CIDs that can be efficiently manipulated intracellularly have not been reported.<sup>[8]</sup> Some spatial selectivity has been achieved with photocleavable, biotinylated  $\alpha$ -methylnitrobenzylrapamycin, which has been used to control small GTPase activity.<sup>[12]</sup> This caged rapamycin was targeted to an extracellular location by means of its biotin moiety, required, however, extracellular photolytic removal of the caging group before rapamycin was released to diffuse across the cell membrane.<sup>[12]</sup> Another photocaged rapamycin derivative is pRap.<sup>[13]</sup> Both of these noncovalent, photocleavable CIDs provide a source of highly diffusible dimerizer, limiting local target manipulation.

Here we present a novel photocleavable CID, which forms a covalent link between HaloTag-<sup>[14]</sup> and SNAP-tag<sup>[15]</sup>-fused proteins. The photocleavable methyl-6-nitroveratryl (MeNV) group was introduced into the core module linking the HaloTag-reactive chloroalkane ligand and the SNAP-tag-reactive O6-benzylguanine, and the cell permeability of the resulting CID molecule is retained (dubbed MeNV-HaXS; Figure 1). The combination of chemical-induced dimerization and the possibility of a subsequent light-induced reversal of the protein–protein interaction combines the advantage of a modular approach of genetically encodable tags with a highly specific spatiotemporal control by light.

As depicted schematically in Figure 1, MeNV-HaXS penetrates cells and induces the dimerization of HaloTag

[\*] M. Zimmermann,<sup>[1]</sup> R. Cal,<sup>[1]</sup> V. Hoffmann, Dr. F. Beauflis, Prof. Dr. M. P. Wymann  
University of Basel, Department of Biomedicine  
Mattenstrasse 28, Basel (Switzerland)  
E-mail: Florent.Beauflis@UniBas.CH  
Matthias.Wymann@UniBas.CH

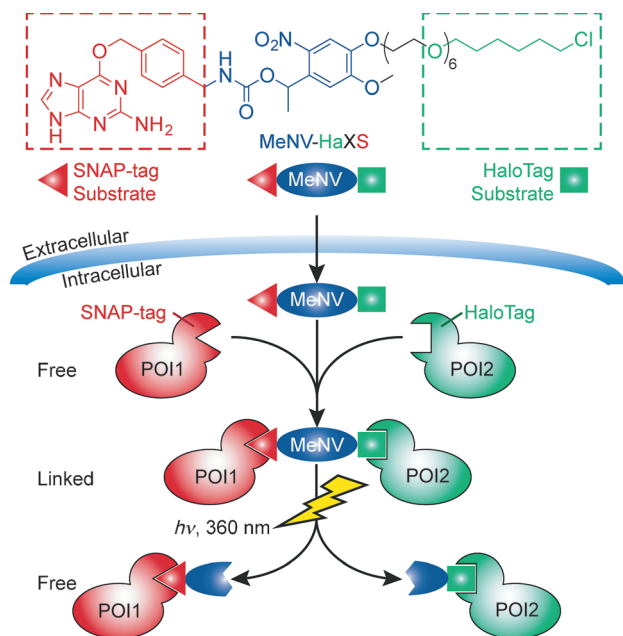
E. Janett, Prof. Dr. C. G. Bochet  
University of Fribourg, Department of Chemistry  
Chemin du Musée 9, Fribourg (Switzerland)  
Prof. Dr. E. Constable  
University of Basel, Department of Chemistry  
Spitalstrasse 51, Basel (Switzerland)

[†] These authors contributed equally to this work.

[\*\*] We thank Takanari Inoue for FRB-YFP-Giantin and Stephan Hübner for NLS-CFP-FKBP expression constructs. This work was supported by Swiss National Science Foundation (205320-138302, 205320-143699), the ESF EuroMEMBRANE Programme (31EM30-126143), and the Novartis (Jubilée) Foundation.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201310969>.

© 2014 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

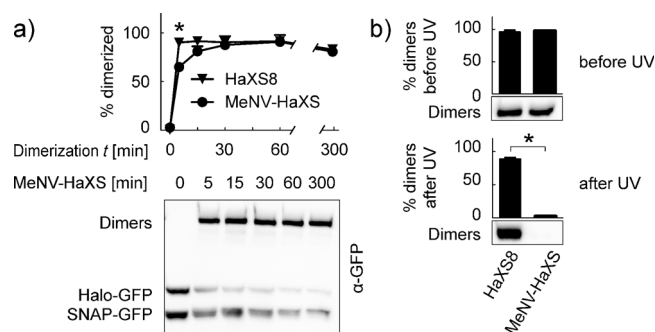


**Figure 1.** A photocleavable, cell-permeable HaloTag- and SNAP-tag-reactive CID with a methyl-6-nitroveratryl (MeNV) core was generated (MeNV-HaXS). After cell entry, MeNV-HaXS dimerizes HaloTag- and SNAP-tag-fused proteins of interest (POI). Illumination of MeNV-HaXS (360 nm;  $\epsilon = 4058 \text{ M}^{-1} \text{ cm}^{-1}$ ; quantum yield = 0.075) cleaves the link between the POIs, and releases them from the covalent complex. For the synthesis of MeNV-HaXS see the Supporting Information.

and SNAP-Tag fusion proteins to form a covalently stabilized complex. Upon illumination (360 nm), the MeNV group undergoes photolysis, which triggers the cleavage of the protein dimer and the release of cargo proteins.

MeNV-HaXS was optimized to match the cell permeability of the noncleavable dimerizer of HaloTag and SNAP-tag fusion proteins called HaXS8.<sup>[16]</sup> Time-dependent dimerization of HaloTag-GFP and SNAP-tag-GFP fusion proteins expressed in HeLa cells was studied in response to the addition of MeNV-HaXS and HaXS8, and we found that MeNV-HaXS and HaXS8 produced dimers at a similar rates. Slight differences were measurable at earlier time points ( $t < 10$  min, Figure 2a), which became insignificant after 15 min of treatment. These slight differences may be explained by the increased molecular weight and higher polarity of MeNV-HaXS, resulting from the incorporation of the PEG6 element. MeNV-HaXS-induced HaloTag-SNAP-tag dimers were stable for  $> 5$  h, and exposure to ambient light did not affect the stability of dimers.

Due to the matched dimerization properties of MeNV-HaXS and HaXS8, the noncleavable HaXS8 is a valuable control compound to monitor the efficiency of photocleavage, and to detect potential side effects exerted by UV irradiation. The efficiency of the intracellular photocleavage of MeNV-HaXS was tested in HeLa cells expressing HaloTag-GFP and SNAP-tag-GFP fusion proteins. A HaloTag-SNAP-tag complex preformed with MeNV-HaXS could be cleaved in bulk by illumination with a 360 nm lamp (Blak-Ray, B-100A, UVP). Quantification of protein dimers before and after



**Figure 2.** MeNV-HaXS induces the formation of intracellular dimers of HaloTag and SNAP-tag fusion proteins, which are cleaved upon UV illumination. a) HeLa cells transfected with expression constructs for SNAP-tag-GFP (SNAP-GFP) and HaloTag-GFP (Halo-GFP) fusion proteins were exposed to  $5 \mu\text{M}$  MeNV-HaXS or  $5 \mu\text{M}$  light-insensitive HaXS8 for the indicated times in cell culture medium at  $37^\circ\text{C}$ . 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  $\pm$  SEM,  $n = 3$ ). b) HeLa cells expressing SNAP-GFP and Halo-GFP as in (a) were incubated with  $5 \mu\text{M}$  MeNV-HaXS or HaXS8 for 15 min. Cells were then washed and submerged in phosphate-buffered saline (PBS) to remove unreacted compounds, and illuminated with a high-intensity UV lamp (100 W, 5 cm distance) for 10 min. Analysis of dimerization products was performed as in (a); values represent means  $\pm$  SEM,  $n = 3$ ; \* indicates  $p < 0.05$ .

exposure to UV light revealed that MeNV-HaXS-induced dimers were quantitatively cleaved after 10 min of illumination, whereas HaXS8-containing dimers remained intact (Figure 2b).

MeNV-HaXS thus offers the possibility to trigger covalent dimerization, and to subsequently release associated proteins in a controlled way. This allows the manipulation of protein localization, which can be exploited to mimic cellular signaling events in a timed and localized fashion. Many signaling events take place at defined intracellular locations. The combination of a tagged anchor protein and MeNV-HaXS exposure permits the depletion of tagged signaling enzymes from their productive sites. Presently, MeNV-HaXS has been validated for the targeting of tagged proteins to intracellular organelles such as Golgi (Figure 3a), plasma membrane, lysosomes, mitochondria, and the actin skeleton (Figure S2).

Targeted irradiation of MeNV-HaXS-anchored protein complexes using a microscope equipped with an XY scanning excitation laser for FRAP (fluorescence recovery after photobleaching; 355 nm) indeed released tagged proteins from the illuminated spots. HeLa cells were cotransfected with cytosolic teal fluorescent (cyan) SNAP-tag fusion protein (SNAP-mTFP1) and the Golgi anchor Halo-RFP-Giantin, which was constructed by the fusion of a red fluorescent protein (monomeric RFP; TagRFP) to a HaloTag and a C-terminal Golgi-targeting motif derived from giantin.<sup>[17]</sup>

Incubation with  $5 \mu\text{M}$  of MeNV-HaXS or  $5 \mu\text{M}$  of HaXS8 efficiently translocated cytosolic SNAP-mTFP1 to the cytosolic surface of the Golgi membrane (Figure 3a). After an illumination pulse ( $8 \times 5$  ms at 355 nm) of a subset of Golgi-



derived vesicles, SNAP-mTFP1 was efficiently and selectively released from illuminated, but not from the non-illuminated vesicular compartments (Figure 3a,c). No significant loss of fluorescence intensity was observed when the light-insensitive HaXS8 was used to anchor SNAP-mTFP1 at Golgi membranes (Figure 3b,c), confirming that the observed light-triggered decrease in fluorescence of SNAP-mTFP1 tethered with MeNV-HaXS results from the release, and not from the photobleaching of SNAP-mTFP1 fluorescence. Moreover, membrane-anchored SNAP-mTFP1 detaches rapidly from

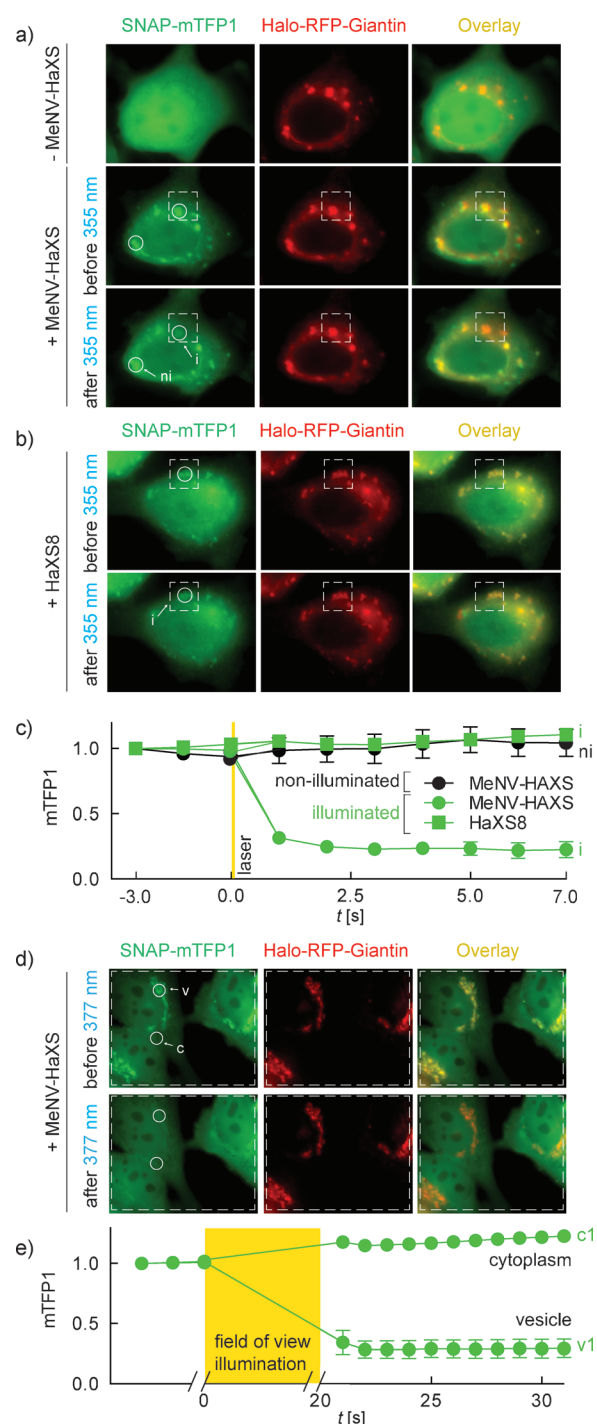
membranes after a laser pulse ( $t < 1$  s). This demonstrates that MeNV-HaXS can be utilized to manipulate protein localization with excellent subcellular precision on a timescale of seconds.

As a scanning FRAP laser is not part of standard fluorescence microscope equipment, we explored the possibility of using global field of view illumination with standard DAPI excitation filters ( $377 \pm 25$  nm; standard mercury halide lamp) to induce the photocleavage of MeNV-HaXS-induced protein complexes. We found that an illumination time of  $< 20$  s was sufficient to completely liberate SNAP-mTFP1 from Golgi membranes (Figure 3d,e). A corresponding increase of mTFP1 fluorescence intensity in the cytoplasm demonstrates that the light-induced fluorescence decrease of SNAP-mTFP1 at vesicles results from the release and not from global photobleaching of SNAP-mTFP1 fluorescence.

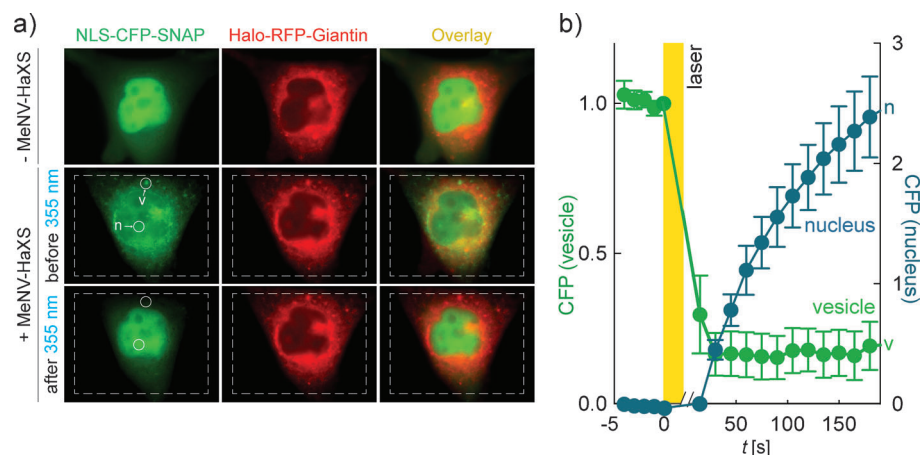
Although slower due to the limited excitation energy, photocleavage using DAPI excitation filters on a conventional fluorescence microscope greatly expands the range of applications offered by the MeNV-HaXS system.

In contrast to previously reported CIDs, MeNV-HaXS can instantaneously release its cargo when illuminated. To further evaluate this concept, a nuclear probe was forced to dock on a Golgi anchor. Expressed alone in HeLa cells, the nuclear localization sequence (NLS)-containing cyan fluorescent probe (NLS-CFP-SNAP) accumulated in the nucleus. In the presence of co-expressed Golgi anchor (Giantin-RFP-Halo) and added MeNV-HaXS, NLS-CFP-SNAP was trapped at perinuclear sites on the Golgi (Figure 4a). Subsequent irradiation of cells triggered the release of NLS-CFP-SNAP from the Golgi to the cytosol within seconds (Figure 4b). Nuclear import of liberated NLS-CFP-SNAP was delayed (Figure 4b), reflecting nuclear import processes.<sup>[18]</sup> This illustrates that photocleavage of MeNV-HaXS is rapidly achieved, and permits the study of nuclear import kinetics in real time in a simple experimental setup.

In summary, MeNV-HaXS is the first cell-permeant CID giving rise to the formation of covalently linked and photo-



**Figure 3.** Translocation of cytosolic SNAP-mTFP1 proteins to the Golgi followed by their release upon UV illumination. a,b) HeLa cells expressing SNAP-mTFP1 and Halo-RFP-Giantin were exposed to a)  $5 \mu\text{M}$  MeNV-HaXS or b)  $5 \mu\text{M}$  HaXS8 in cell culture medium for 15 min at  $37^\circ\text{C}$ , both of which induced translocation of cytosolic SNAP-mTFP1 to the Golgi. SNAP-mTFP1 intensity 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  $\times$  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); values are means  $\pm$  SEM,  $n=10$ , error bars not shown when smaller than symbols. d) HeLa cells as in (a) were exposed to  $5 \mu\text{M}$  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 (v), and in the cytoplasm (c) before and after illumination as described in (d); values are means  $\pm$  SEM,  $n=10$ , error omitted when smaller than symbols; for more controls see Figure S3.



**Figure 4.** Control of nuclear export and import by MeNV-HaXS. Translocation of the NLS-CFP-SNAP probe from the nucleus to the Golgi was controlled by MeNV-HaXS and reversed by subsequent UV illumination. a) HeLa cells expressing Halo-RFP-Giantin and NLS-CFP-SNAP were exposed to 5  $\mu\text{m}$  MeNV-HaXS in cell-culture medium for 15 min at 37°C, which induced export of NLS-CFP-SNAP from the nucleus to the Golgi. 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) with a scanning FRAP laser (150 areas  $\times$  5 ms at 355 nm). 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  $\pm$  SEM,  $n=7$  cells, error bars not shown where smaller than symbols used.

cleavable protein complexes. Like other CID approaches, the design of fusion proteins bearing interacting functional protein domains must be carefully considered. The covalent link formed by MeNV-HaXS simplifies monitoring and validation of protein complexes greatly. MeNV-HaXS is therefore a novel valuable tool that combines intracellular protein dimerization with photocleavage triggered by a wavelength compatible with widely used fluorescent reporter proteins. The possibility to control protein localization by two independent events can be exploited to sequester any protein of interest away from its functional compartment. CID-dependent trapping of enzymes to nonfunctional sites has been established as an elegant approach to interrupt signaling pathways.<sup>[19,20]</sup> Optically guided cleavage of MeNV-HaXS can release anchored proteins and restore their function. Many more scenarios are possible, for example the deployment of on-off-on and off-on-off protocols and the orthogonal use of MeNV-HaXS with other CIDs. The choice between global and local illumination triggering dissociation of the CID-mediated complex opens a wide range of applications, such as investigation of cell-compartment-associated signaling, and the simulation of cell-wide physiological and pathological signaling dynamics.

Received: December 18, 2013  
Published online: March 26, 2014

**Keywords:** chemical inducers of dimerization · fusion proteins · photocleavable linkers · photolysis · protein–protein interactions

- [1] M. P. Wymann, R. Schneider, *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 162.
- [2] G. Mayer, A. Heckel, *Angew. Chem.* **2006**, *118*, 5020; *Angew. Chem. Int. Ed.* **2006**, *45*, 4900; C. Brieke, F. Rohrbach, A. Gottschalk, G. Mayer, A. Heckel, *Angew. Chem.* **2012**, *124*, 8572; *Angew. Chem. Int. Ed.* **2012**, *51*, 8446.
- [3] Y. I. Wu, X. Wang, L. He, D. Montell, K. M. Hahn, *Methods Enzymol.* **2011**, *497*, 393.
- [4] A. Levskaya, O. D. Weiner, W. A. Lim, C. A. Voigt, *Nature* **2009**, *461*, 997.
- [5] M. Yazawa, A. M. Sadaghiani, B. Hsueh, R. E. Dolmetsch, *Nat. Biotechnol.* **2009**, *27*, 941.
- [6] M. J. Kennedy, R. M. Hughes, L. A. Peteya, J. W. Schwartz, M. D. Ehlers, C. L. Tucker, *Nat. Methods* **2010**, *7*, 973.
- [7] D. Strickland, Y. Lin, E. Wagner, C. M. Hope, J. Zayner, C. Antoniou, T. R. Sosnick, E. L. Weiss, M. Glotzer, *Nat. Methods* **2012**, *9*, 379.
- [8] T. W. Corson, N. Aberle, C. M. Crews, *ACS Chem. Biol.* **2008**, *3*, 677.
- [9] M. Ishida, H. Watanabe, K. Takigawa, Y. Kurishita, C. Oki, A. Nakamura, I. Hamachi, S. Tsukiji, *J. Am. Chem. Soc.* **2013**, *135*, 12684.
- [10] P. Varnai, B. Thyagarajan, T. Rohacs, T. Balla, *J. Cell Biol.* **2006**, *175*, 377; T. Inoue, T. Meyer, *PLoS One* **2008**, *3*, e3068; B. C. Suh, T. Inoue, T. Meyer, B. Hille, *Science* **2006**, *314*, 1454.
- [11] F. Castellano, P. Montcourrier, P. Chavrier, *J. Cell Sci.* **2000**, *113*, 2955; T. Inoue, W. D. Heo, J. S. Grimley, T. J. Wandless, T. Meyer, *Nat. Methods* **2005**, *2*, 415.
- [12] N. Umeda, T. Ueno, C. Pohlmeier, T. Nagano, T. Inoue, *J. Am. Chem. Soc.* **2011**, *133*, 12.
- [13] A. V. Karginov, Y. Zou, D. Shirvanyants, P. Kota, N. V. Dokholyan, D. D. Young, K. M. Hahn, A. Deiters, *J. Am. Chem. Soc.* **2011**, *133*, 420.
- [14] G. V. Los, L. P. Encell, M. G. McDougall, D. D. Hartzell, N. Karassina, C. Zimprich, M. G. Wood, R. Learish, R. F. Ohane, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. H. Klaubert, R. F. Balleit, K. V. Wood, *ACS Chem. Biol.* **2008**, *3*, 373.
- [15] M. J. Hinner, K. Johnsson, *Curr. Opin. Biotechnol.* **2010**, *21*, 766.
- [16] D. Erhart, M. Zimmermann, O. Jacques, M. B. Wittwer, B. Ernst, E. Constable, M. Zvelebil, F. Beaufils, M. P. Wymann, *Chem. Biol.* **2013**, *20*, 549.
- [17] T. Komatsu, I. Kukelyansky, J. M. McCaffery, T. Ueno, L. C. Varela, T. Inoue, *Nat. Methods* **2010**, *7*, 206.
- [18] M. Stewart, *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 195.
- [19] M. S. Robinson, D. A. Sahlender, S. D. Foster, *Dev. Cell* **2010**, *18*, 324.
- [20] H. Haruki, J. Nishikawa, U. K. Laemmli, *Mol. Cell* **2008**, *31*, 925.



Supporting Information

© Wiley-VCH 2014

69451 Weinheim, Germany

**Cell-Permeant and Photocleavable Chemical Inducer of  
Dimerization\*\***

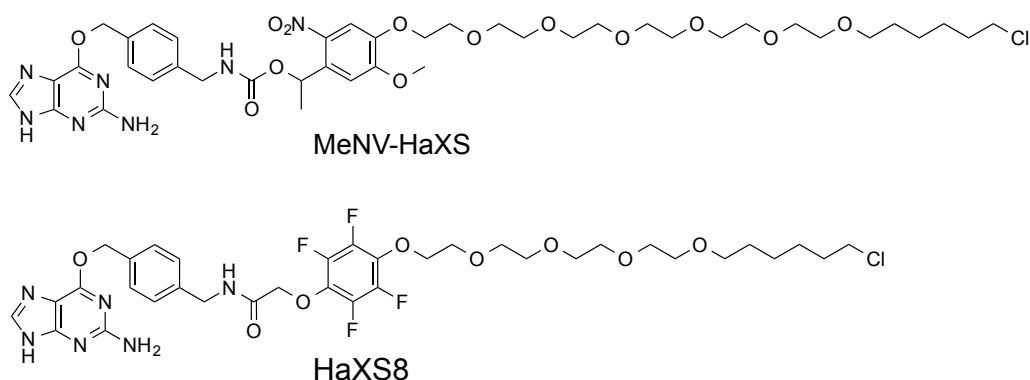
*Mirjam Zimmermann, Ruben Cal, Elia Janett, Viktor Hoffmann, Christian G. Bochet,  
Edwin Constable, Florent Beaufils,\* and Matthias P. Wymann\**

anie\_201310969\_sm\_miscellaneous\_information.pdf

## Table of Contents

1. Structure of key compound used in this study	page 2
2. Supplementary experimental procedures and conditions	page 3
3. Chemical synthesis and characterization of compounds	page 6
4. Supplementary experiments and data	page 9
5. Spectra of intermediates and MeNV-HaXS	
a. NMR spectra ( <sup>1</sup> H NMR and <sup>13</sup> C NMR)	page 14
b. HRMS spectra	page 22

## 1. Structures of key compounds used in this study



MeNV-HaXS (top):

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

HaXS8 (bottom):

*N*-(4-(((2-amino-9H-purin-6-yl)oxy)methyl)benzyl)-2-(4-((18-chloro-3,6,9,12-tetraoxaoctadecyl)oxy)-2,3,5,6-tetrafluorophenoxy)acetamide (see reference [16] in main text).

## 2. Supplementary experimental procedures and conditions

### Chemical synthesis and characterization

MeNV-HaXS was synthesized in 8 steps starting from acetovanillone, hexaethylene glycol, O6-aminomethylbenzylguanine and 6-chloro-1-iodohexane with good overall yield (>800 mg produced). To balance cell permeability and solubility in water, a PEG6 element was introduced. The methyl-6-nitroveratryl group (MeNV) was selected as a photocleavable group because it absorbs light at >360 nm. This minimizes cell damage, matches initiation of photocleavage to widely available excitation equipment and filters, and MeNV provides a sufficient quantum yield for intracellular photocleavage. Last but not least, intermediates and by-products emerging from the photolytic process do not interfere with living cells. The photophysical properties of MeNV-HaXS were determined *in vitro* as described below. MeNV-HaXS is not sensitive to ambient light, displays excellent cleavage efficiency at 360 nm with an absorption coefficient of  $4058 \text{ M}^{-1}\text{cm}^{-1}$ , and a quantum yield of 0.075. No special precautions concerning light exposure were required during synthesis.

Materials and reagents were of the highest commercially available grade and used without further purification. Reactions were monitored by thin layer chromatography (TLC) using Merck silica gel 60 F254 plates. Compounds were visualized by UV, ceric ammonium molybdate (CAM),  $\text{KMnO}_4$  and ninhydrin. Flash chromatography was performed using Merck silica gel 60, particle size 40 - 63  $\mu\text{m}$ .

$^1\text{H}$ ,  $^{19}\text{F}$  and  $^{13}\text{C}$  NMR spectra were recorded on a Bruker AV-400 or a DRX-600 NMR spectrometer. Chemical shifts are reported in ppm using the solvent residue signals as reference. All solvents used for reactions were purchased as anhydrous grade from Fluka. Solvents for extractions, flash chromatography and TLC were commercial grade. Mass spectra were recorded with a VG70-250 (FAB), Finnigan MAT MS 312 (EI) or Finnigan MAT LCQ (ESI) spectrometer. High resolution mass spectra were recorded with a Thermo Fisher Scientific LTQ Orbitrap XL, nanoelectrospray ion source.

### Measurement of the absorption spectra

UV/Vis spectra of  $10^{-5}$  M solution of MeNV-HaXS in DMSO was recorded on a Perkin Elmer Lambda 40 UV/Vis spectrometer using quartz standard absorption cells. The extinction coefficients were calculated for 360 nm.

The emission of the transilluminator lamps (RPR-3500 Å, RPR-3600 Å and RPR-4190 Å) was recorded with an AVA SPEC 2048 spectrometer. With the same spectrometer an absorption spectrum of the glass cover slip was recorded to verify that there is no absorbance in the irradiation wavelengths range.

### General procedure for the analysis of photolysis

The irradiated solution was transferred into an HPLC microvial and injected into an Aquity H-Class UPLC system equipped with an ESI-SQD mass spectrometer. The photolysis yield was determined by integration of the Single Ion Monitoring (SIM) signal of the starting product.

### 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  $\text{Na}_3\text{VO}_4$ , 20  $\mu\text{M}$  Leupeptin, 18  $\mu\text{M}$  Pepstatin, 5  $\mu\text{M}$  Aprotinin, 1 mM PMSF, 1 mM  $\text{MgCl}_2$ , 1 mM  $\text{CaCl}_2$ , 5 mM EDTA]. Cell lysates were cleared by centrifugation at 13'000 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%  $\beta$ -mercaptoethanol, 50% glycerol, bromphenol blue] and cooking for 6 min. Proteins were separated by SDS-PAGE and transferred to Immobilon PVDF membranes (Millipore). Mouse monoclonal antibody (mAb) to GFP (Roche Diagnostics) was used to detect GFP-fusion proteins. Secondary antibodies labeled with horseradish peroxidase (HRP-conjugated goat anti-mouse IgG) were visualized using enhanced chemiluminescence (Millipore) and a CCD camera (Fusion Fx7, Vilber).

### Cloning of expression constructs

The HaloTag7 L273Y coding sequence (Promega) and 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 HaloTag L273Y L273Y-GFP, HaloTag7 L273Y-RFP, and SNAPf-GFP expression constructs, HaloTag7 L273Y and SNAPf were transferred to pEGFP (Clontech), or pTag-RFP-N1 (Evrogen) expression vectors.

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

HaloTag7 L273Y-RFP-Giantin: the Giantin targeting sequence was introduced into the HaloTag7 L273Y-RFP expression vector.

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

HaloTag7 L273Y-RFP-Rheb15: HaloTag7 L273Y-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 HaloTag7 L273Y-RFP in a HA-Raptor-Rheb15 expression vector kindly obtained from Anna Melone.

LAMP-RFP-HaloTag7 L273Y: the LAMP1 targeting sequence was inserted into the RFP-HaloTag7 L273Y 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).

HaloTag7 L273Y-mTq: GFP was exchanged by the mTq sequence.

*Maps and expression vector sequences can be obtained from the authors upon request.*

### Statistical Analysis

Statistical analysis was performed with GraphPad Prism v6. For Student's t test (two sided, non-paired with Welch correction,  $p < 0.05$ )  $\geq 3$  independent experiments were compared.

### Cell culture and transfection

HeLa cells (ATCC) were cultured in complete Dulbecco's modified Eagle medium (DMEM) with 10% heat-inactivated fetal calf serum (HIFCS), 2 mM L-glutamine (Gln), 1% penicillin-streptomycin solution (PEST) at 37°C and 5% CO<sub>2</sub>. Transfections were carried out with JetPEI (Brunschiwig) according to manufacturer's guidelines.

### Cellular heterodimerization and intracellular cleavage

HeLa cells were grown in 6-well cell culture plates (Falcon), and were transfected with expression constructs for SNAP-GFP, Halo-GFP, Halo-RFP-Giantin, SNAP-mTFP1 or NLS-CFP-SNAP. After 24 h, cells were exposed to MeNV-HaXS or HaXS8 dimerizer as described in complete cell culture medium (at 37°C). Before bulk photocleavage experiments, cells were washed twice in 1 x PBS after treatment with HaXS8 or MeNV-HaXS. Then 600  $\mu$ l 1x PBS was added to each well, and the cell cells in the culture plate on ice was illuminated for 10 min with a high-intensity UV lamp (Blak-Ray B-100A high intensity UV lamp; 100 Watt, 365 nm, UVP) at a distance of 5 cm.

For Western Blot analysis, cells were lysed, and proteins were separated by SDS-PAGE. SNAP/Halo-tag dimers were detected using anti-GFP (primary) and horseradish peroxidase (HRP)-conjugated (secondary) antibodies to visualize dimer formation using enhanced chemiluminescence (Millipore) and a CCD camera (Fusion Fx7, Vilber).

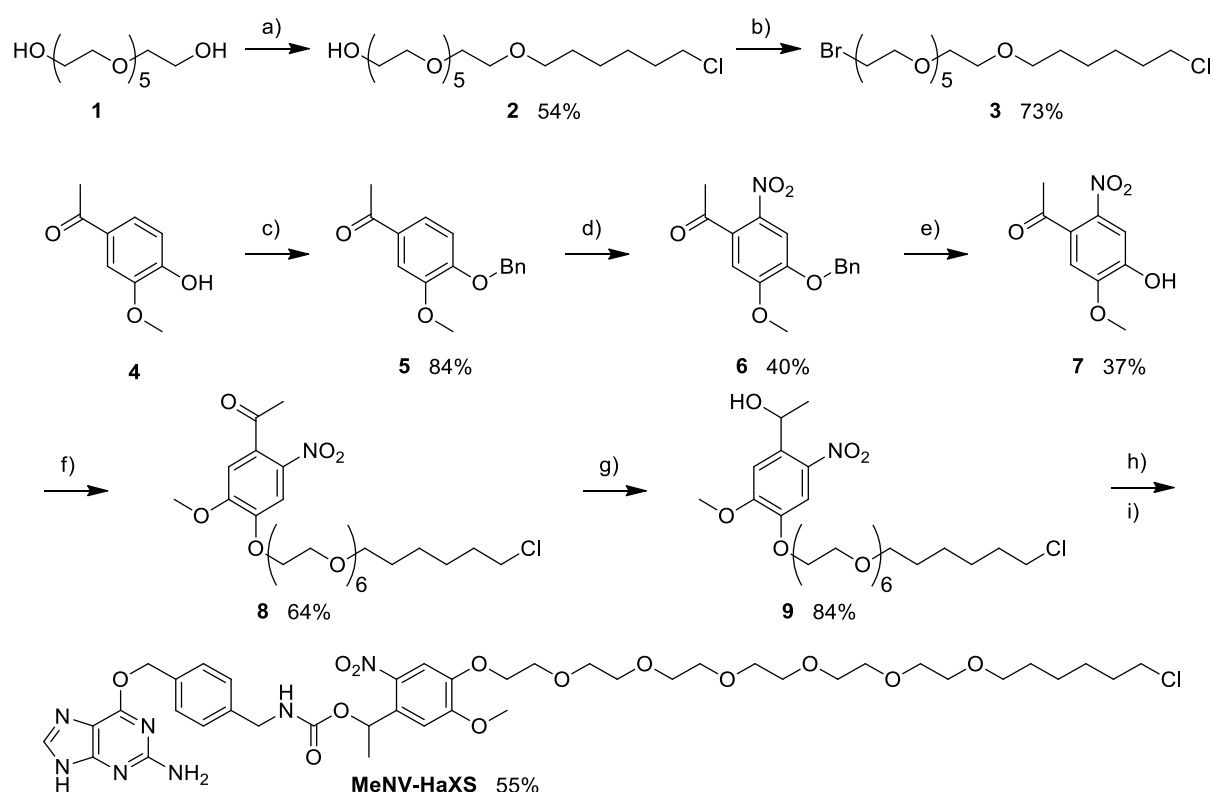
For live cell microscopy, transfected and MeNV-HaXS resp. HaXS8 treated cells grown on 12 mm cover slips (Menzel) were mounted in Ludin chambers (Live Imaging Services) in the closed confirmation with complete cell culture medium without phenol red. Pictures were taken every 1 seconds before illumination of the cells and every 1 sec for Figure 3 resp. every 15 sec for Figure 4. Images were acquired on a Leica Live Imaging Microscope fitted with a HCX Plan-Fluotar 63x/1.4 oil objective and a Photometrics CCD Camera CoolSnap HQ2 with Metamorph 7.1 software (Molecular devices). Cleavage in Fig. 3a and 3b and Fig. 4a was performed with a 50 mW FRAP scanning laser (UV-Diode Laser, 355 nm, 30 Hz). Single spots within the indicated region of interest were irradiated for the indicated times. For photocleavage experiments as shown in Fig. 3d, standard DAPI excitation filters were used on a Leica EL600 microscope (equipped with a mercury metal halide lamp; filter cube: ex: 377 $\pm$ 25 nm, em: 447 $\pm$ 30 nm, dichroic mirror: 409 nm; 63x objective: 63xHCX Plan-Fluotar, NA: 1.25 [oil]).

Movies were assembled and analyzed with Fiji (ImageJA, 1.44b). In Figure 3a 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. In Figure 3b mTFP1

fluorescence in the cytoplasm was calculated according to the formula  $mTFP1 = F(mTFP1, vesicle) / F(RFP, whole cell)$ . In Figure 4 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 vesicles and to measure fluorescence increase in the cytoplasm or nucleus, mTFP1 or CFP fluorescence intensity was plotted over time.

## 3. Chemical synthesis and characterization of compounds

## Synthesis of MeNV-HaXS



Reaction conditions: a) NaH, 6-chloro-1-iodohexane, THF/DMF, 0 °C, 16 h; b) PPh<sub>3</sub>, CBr<sub>4</sub>, rt, 16 h; c) K<sub>2</sub>CO<sub>3</sub>, benzyl bromide, DMF, 80 °C, 16 h; d) Acetic acid, acetic anhydride, HNO<sub>3</sub>, rt, 16 h; e) Acetic acid, HBr, 85 °C, 1.5 h; f) K<sub>2</sub>CO<sub>3</sub>, **3**, DMF, 60 °C, 16 h; g) NaBH<sub>4</sub>, MeOH/Dioxane, rt, 2 h; h) bis *p*-nitrophenyl carbonate, triethylamine, DMF, rt, 16 h; i) O6-aminomethylbenzylguanine, 60 °C, 6 h.

## Scheme S1. Synthetic route to MeNV-HaXS.

Preparation of 24-chloro-3,6,9,12,15,18-hexaoxatetracosan-1-ol (**2**)

Hexaethyleneglycol (**1**) (2.1 mL, 8.4 mmol) was dissolved in a THF/DMF mixture (3:1). Sodium hydride (60% in mineral oil, 370 mg, 9.2 mmol) was added portionwise at 0 °C. After stirring for 30 min at rt, 6-chloro-1-iodohexane (6.4 mL, 8.4 mmol) was added dropwise at 0 °C. The mixture was stirred at rt for 16 h. The excess of sodium hydride was carefully quenched with water and the crude mixture was poured into water and extracted twice with AcOEt. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude oil was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1) to yield compound **2** (5.3 g, 54%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD): δ 3.66 - 3.61 (m, 18H), 3.58 - 3.54 (m, 6H), 3.47 (t, J = 6.80 Hz, 2H), 3.30 (m, 2H), 1.77 (m, 2H), 1.58 (m, 2H), 1.50-1.38 (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 1-bromo-24-chloro-3,6,9,12,15,18-hexaoxatetracosane (**3**)

24-chloro-3,6,9,12,15,18-hexaoxatetracosan-1-ol **2** (950 mg, 2.37 mmol) was dissolved in THF. Triphenylphosphine (721 mg, 2.76 mmol) and carbon tetrabromide (900 mg, 2.76 mmol) were added portionwise at 0 °C. The resulting mixture was stirred at rt for 16 h. The solvent is evaporated

under reduced pressure and the crude oil is purified by flash chromatography (cyclohexane/EtOAc, 3:1) to yield compound **3** (800 mg, 73%). <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD): δ 3.80 (t, J = 6.02 Hz, 2H), 3.66-3.61 (m, 18H), 3.58-3.54 (m, 4H), 3.51 (t, J = 6.04 Hz, 2H), 3.48 (t, J = 6.52), 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): δ 72.3, 72.1, 71.6, 71.4, 71.2, 45.7, 33.8, 31.4, 30.6, 27.7, 26.5.

#### Preparation of 1-(4-(benzyloxy)-3-methoxyphenyl)ethanone (**5**)

Acetovanillone (**4**) (5 g, 30.1 mmol) and benzylbromide (5.15 g, 30.1 mmol) were dissolved in 50 mL DMF. K<sub>2</sub>CO<sub>3</sub> (4.15 g, 30.1 mmol) was added and the mixture was stirred at 80°C for 16 h. The reaction was quenched with sat. NH<sub>4</sub>Cl (2 L) and extracted with AcOEt three times. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The resulting product **5** (6.5 g, 84%) was used without further purification. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 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 (m, 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, CDCl<sub>3</sub>): δ 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 C<sub>16</sub>H<sub>17</sub>O<sub>3</sub> [M+H]<sup>+</sup> calcd: 257.1172, found: 257.1166.

#### Preparation of 1-(4-(benzyloxy)-5-methoxy-2-nitrophenyl)ethanone (**6**)

Nitric acid (1.4 mL, 69%) was added dropwise to a solution of *o*-benzylacetovanillone (**5**) (1.5 g, 5.8 mmol) in acetic acid (4.9 mL) and acetic anhydride (1.4 mL) at 0°C. The mixture was stirred at rt overnight. The reaction was poured into ice water and extracted with AcOEt three times. The combined organic layers were washed with saturated NaHCO<sub>3</sub>, brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated under reduced pressure and the crude product was purified by flash chromatography (cyclohexane/EtOAc 4:1) to yield the desired product **6** (718 mg, 40%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 7.67 (s, 1H), 7.46-7.29 (m, 5H), 6.77 (s, 1H), 5.22 (s, 2H), 3.98 (s, 3H), 2.49 (s, 3H); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>): δ 200.1, 154.6, 148.6, 138.3, 135.3, 133.1, 128.9, 128.6, 127.6, 108.8, 71.4, 56.7, 30.4; HRMS C<sub>16</sub>H<sub>16</sub>O<sub>5</sub>N [M+H]<sup>+</sup> calcd: 302.1023, found: 302.1017.

#### Preparation of 1-(4-hydroxy-5-methoxy-2-nitrophenyl)ethanone (**7**)

6-Nitro-*o*-benzylacetovanillone (**6**) (4.3 g, 14.3 mmol) was dissolved in acetic acid (30 mL, 99%), and heated to 85°C. HBr (15 mL, 48%) was added to the mixture and stirred for 1.5 h. The mixture was poured into ice water and extracted with AcOEt three times. The combined organic layers were washed with saturated NaHCO<sub>3</sub>, brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated under reduced pressure. The crude product was purified by flash chromatography (cyclohexane/EtOAc, 2:1) to yield the desired compound **7** (1.1 g, 37%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 9.76 (s, 1H), 7.60 (s, 1H), 6.81 (s, 1H), 3.98 (s, 3H), 2.48 (s, 3H); <sup>13</sup>C NMR (100.6 MHz, CDCl<sub>3</sub>): δ 199.7, 152.3, 147.8, 138.8, 130.4, 110.8, 108.8, 56.2, 29.9; HRMS C<sub>9</sub>H<sub>10</sub>O<sub>5</sub>N [M+H]<sup>+</sup> calcd: 212.0553, found: 212.0553.

#### Preparation of 1-(4-((24-chloro-3,6,9,12,15,18-hexaoxatetracosyl)oxy)-5-methoxy-2-nitrophenyl)ethanone (**8**)

*o*-Hydroxy-6-nitrovanillone (**7**) (455 mg, 2.2 mmol) and K<sub>2</sub>CO<sub>3</sub> (297 mg, 2.2 mmol) were added to a solution of **3** (1.0 g, 2.2 mmol) in DMF (6 mL). The mixture was stirred at 60°C overnight. The reaction mixture was poured into a solution of sat. NH<sub>4</sub>Cl (150 mL) and extracted with EtOAc three times and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was evaporated under reduced pressure and the crude product was purified by flash chromatography (cyclohexane/EtOAc, 1:3) to yield the desired product **8** (830 mg, 64%). <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>ClNO<sub>11</sub> [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 (9)**

Compound **8** (675 mg, 1.1 mmol) was dissolved in MeOH/Dioxane mixture (7 mL / 7 mL). NaBH<sub>4</sub> (51 mg, 2.0 mmol) was added portionwise at 0°C. The mixture was stirred at rt for 2 h. Then, the mixture was poured into water, neutralized with a 1 M solution of HCl and extracted with CH<sub>2</sub>Cl<sub>2</sub> three times. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 20:1) to yield compound **9** (568 mg, 84%). <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>ClNO<sub>11</sub>Na [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)benzylcarbamate (MeNV-HaXS)**

Compound **9** (829 mg, 1.4 mmol) was added slowly to a solution of bis *p*-nitrophenyl carbonate (423 mg, 1.4 mmol) in 10 mL of DMF at 0°C. Triethylamine (193 μL, 1.4 mmol) was added and the mixture was stirred at rt overnight. O6-aminomethylbenzylguanine (379 mg, 1.4 mmol) was added portionwise. The solution was stirred at 70°C for 6 h. The reaction was quenched with H<sub>2</sub>O and extracted with EtOAc three times. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was evaporated under reduced pressure. The crude product was purified by flash 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, DMSO-d<sub>6</sub>): 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); <sup>13</sup>C NMR (100.6 MHz, DMSO-d<sub>6</sub>): δ 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 C<sub>41</sub>H<sub>59</sub>ClN<sub>7</sub>O<sub>13</sub> [M+H]<sup>+</sup> calcd: 892.3854, found: 892.3842.



#### 4. Supplementary experiments and data

##### Determination of the quantum yield of MeNV-HaXS at 360 nm:

The quantum yield ( $\phi$ ) of MeNV-HaXS was determined by dividing the reaction rate of the photolysis ( $r$ ) by the number of photons entering the photoreactor per unit of time (quantum flow).

$$\phi = \frac{\text{reaction rate}}{\text{quantum flow}} \quad (1)$$

##### Determination of the quantum flow of the Lumos 43A photoreactor:

Ferrioxalate actinometry was used to determine the quantum flow of the Lumos 43A photoreactor (LED 360 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 (Figure S1 A). The quantum flow was calculated using equation (2).

$$\text{Quantum flow [Einstein/s]} = \frac{A \cdot V_1 \cdot V_3}{t \cdot \phi \cdot \epsilon \cdot l \cdot V_2} \quad (2)$$

A = absorbance of the irradiated solution at 510 nm

t = irradiation time in seconds

$\phi$  = quantum yield at the irradiation wavelength (1.26 at 360 nm)

$\epsilon$  = extinction coefficient of the complex at 510 nm  
( $1.11 \cdot 10^4$  L/Mol/cm)

V<sub>1</sub> = irradiated volume (3 mL)

V<sub>2</sub> = used volume of V<sub>1</sub> (2 mL)

V<sub>3</sub> = end volume (20 mL)

l = thickness of the cell (1 cm)

The dose of photon emitted by the lamp was determined and plotted against time (Figure S1 B). The slope of the regression line, corresponding to the quantum flow of the Lumos 43A photoreactor is  $1.33 \cdot 10^{-8}$  E/s or  $4.79 \cdot 10^{-5}$  E/h.

##### Determination of the photolysis rate of MeNV-HaXS:

3 mL of a 0.5 M of a MeNV-HaXS in a DMSO/water (1:10) solution were transferred into a standard absorbance cell and irradiated in the Lumos 43A photoreactor at 360 nm for 1, 2, 3.5 and 5 minutes. The conversion was determined by UPLC-MS by integration of the absorption peak of **MeNV-HaXS**. Disappearance of the **MeNV-HaXS** was plotted as a function of the time (Figure S1 C). The slope of the regression line corresponds to the rate constant ( $k$ ). The reaction rate was calculated using equation (3). The reaction rate was quantified as  $6 \cdot 10^{-8}$  mol/min.

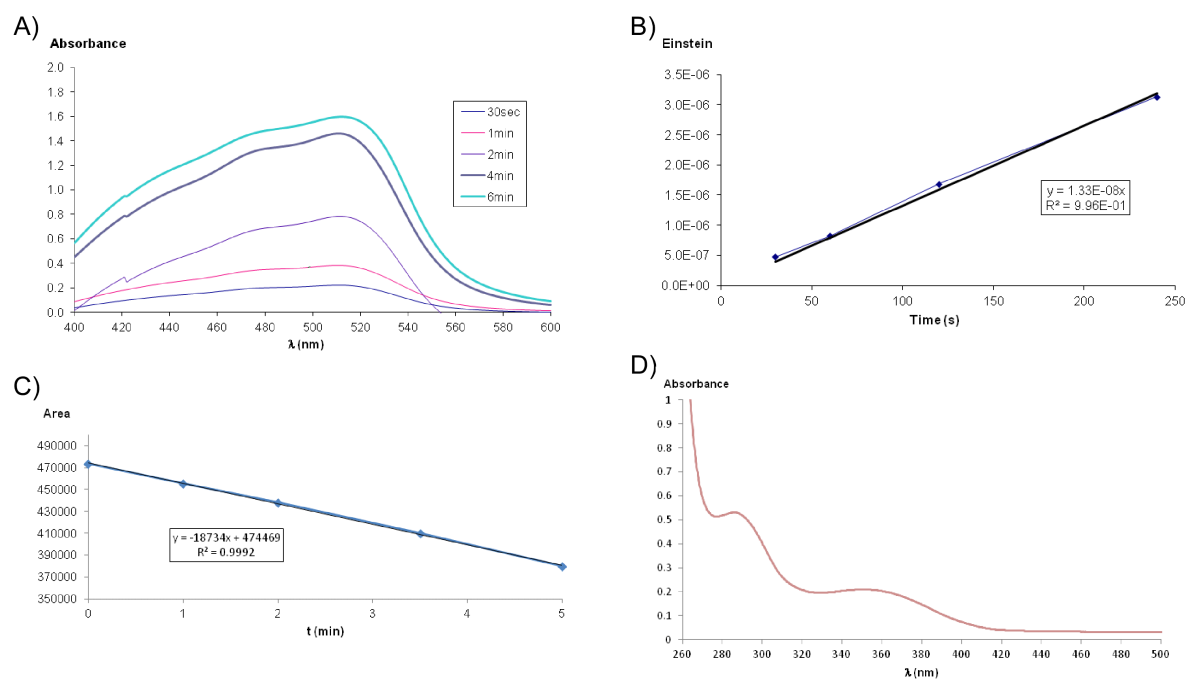
$$r = k[A] \quad (3)$$

The reaction rate was quantified as  $6 \cdot 10^{-8}$  mol/min. Using equation (1) with the reaction rate  $6 \cdot 10^{-8}$  mol/min and the quantum flow of  $1.33 \cdot 10^{-8}$  E/s, the photolysis quantum yield of MeNV-HaXS was calculated to be 0.075.

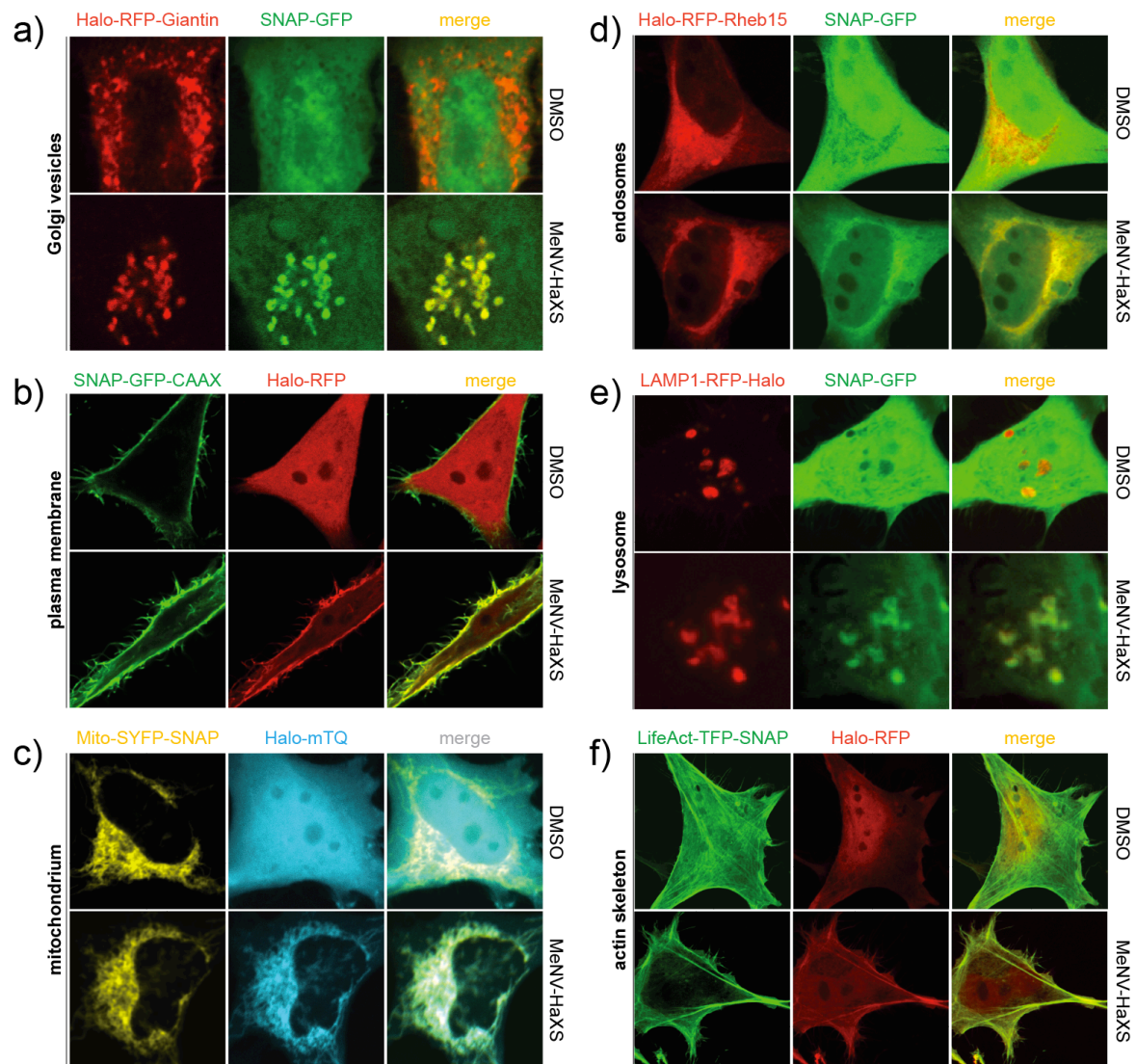
##### Determination of the molar absorption coefficient of MeNV-HaXS at 360 nm

A  $10^{-5}$  M solution of MeNV-HaXS in DMSO was recorded on a Perkin Elmer Lambda 40 UV/Vis spectrometer using quartz standard absorption cells (l = 1cm) (Figure S1 D). The extinction coefficient at 360 nm was calculated using the Beer-Lambert law (4).

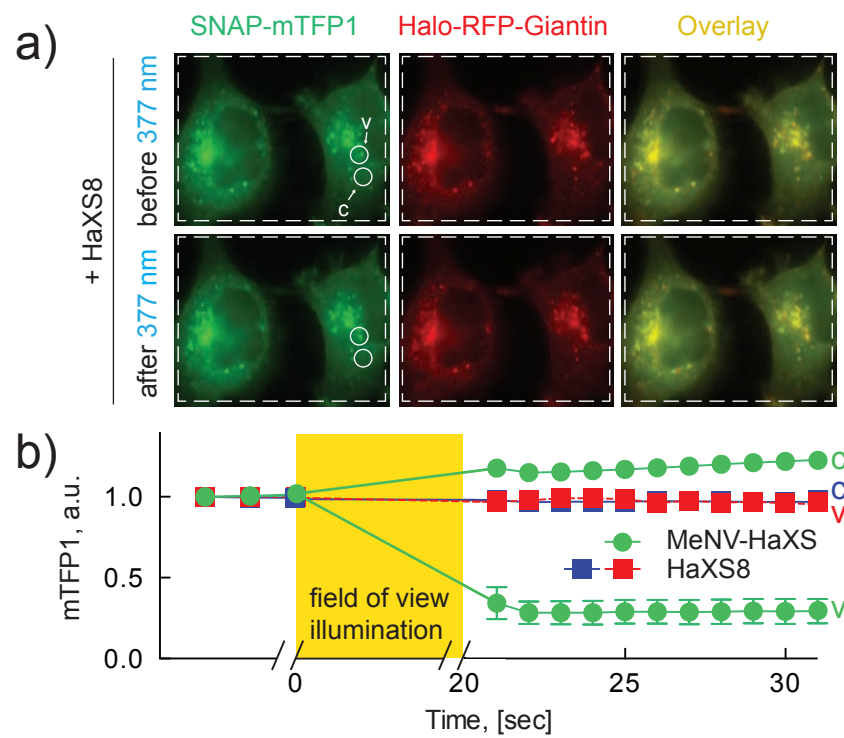
$$A = \epsilon \cdot l \cdot c \quad (4)$$



**Figure S1:** A) UV/Vis spectrum of ferrous *o*-phenanthroline complex formed after irradiation of Potassium ferrioxalate B) Dose of photons emitted by the lamp plotted against time to determine the quantum flow (E/s) of the Lumos 43A photoreactor (LED 360 nm lamp) C) Amount of MeNV-HaXS as determined with UPLC-MS plotted against time to determine the photolysis rate of MeNV-HaXS D) UV/Vis spectrum of MeNV-HaXS.

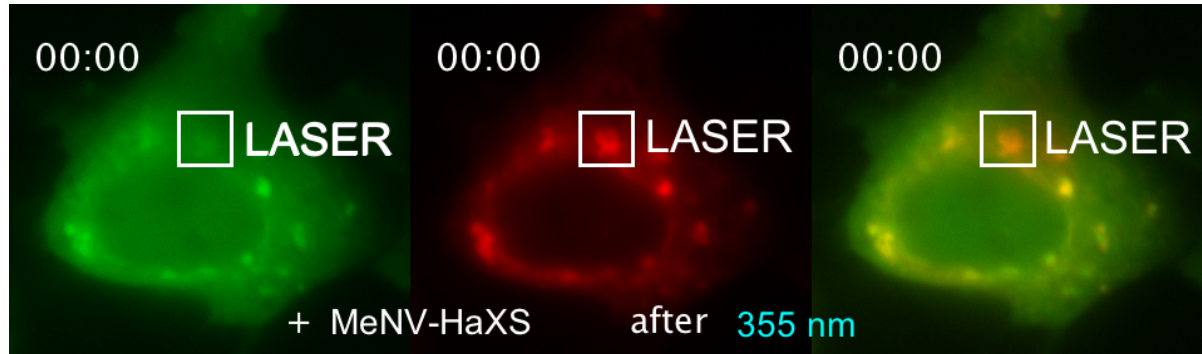


**Figure S2:** Cytosolic protein translocation to different cellular compartments: a) Giantin was used to target the Halo-RFP-Giantin to the Golgi, b) the CAAX box of K-Ras targeted SNAP-GFP-CAAX to the plasma membrane, c) Mito-SYFP-SNAP localized on mitochondria, d) the Rheb15-tagged Halo-RFP-Rheb15 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 (Plüss-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.

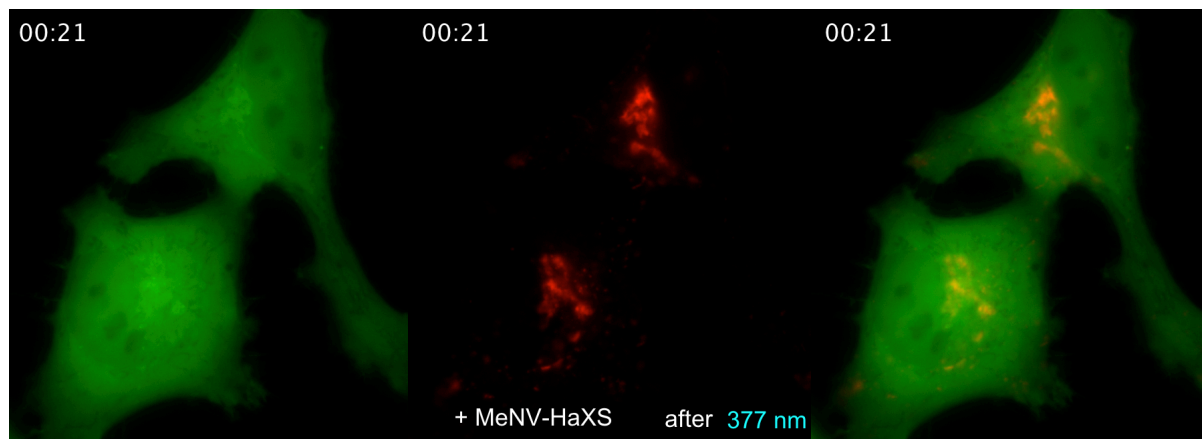


**Figure S3:** Release of Golgi-trapped SNAP-mTFP1 protein after UV illumination. a) At  $t = -15$  min, HeLa cells expressing SNAP-mTFP1 and Halo-RFP-Giantin were exposed to  $5 \mu\text{M}$  HaXS8 ( $5 \mu\text{M}$  MeNV-HaXS see Figure 3, main part), which caused the translocation of cytosolic SNAP-mTFP1 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 3) are shown in green.

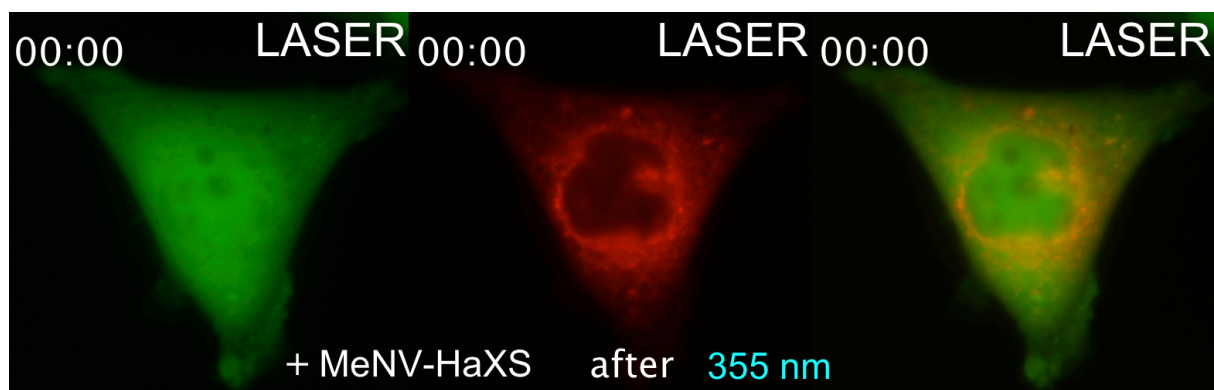
## Legends to movie files



**Movie S1:** HeLa cells expressing SNAP-mTFP1 and Halo-RFP-Giantin (Golgi anchor) were exposed to 5  $\mu\text{M}$  MeNV-HaXS ( $t = -15:00$  min). MeNV-HaXS induced translocation of cytosolic SNAP-mTFP1 to the Golgi as imaged by live cell microscopy is shown in Movie S1. UV illumination at  $t = 00:00$  of a subset of vesicles with a scanning FRAP laser (area marked by white square, LASER) released anchored SNAP-mTFP1 into the cytosol.



**Movie S2:** HeLa cells expressing SNAP-mTFP1 and Halo-RFP-Giantin were exposed to 5  $\mu\text{M}$  MeNV-HaXS. Illumination of cells starting at  $t = 00:00$  (for 20s, DAPI excitation filter, see main text) released Golgi-anchored SNAP-mTFP1.



**Movie S3:** HeLa cells expressing the nuclear NLS-CFP-SNAP probe and Halo-RFP-Giantin were exposed to 5  $\mu\text{M}$  MeNV-HaXS. The MeNV-HaXS induced translocation of nuclear NLS-CFP-SNAP to the Golgi was monitored by live cell microscopy. Illumination of the cell body was initiated at  $t = 00:00$  (150 areas \* 5 ms at 355 nm) using a scanning FRAP laser, leading to the immediate release of NLS-CFP-SNAP from Golgi vesicles into the cytosol. The reimport into the nucleus displays slower kinetics due to processes controlled by the nuclear import machinery.

# Photocleavable Chemical Inducer Of Dimerization To Manipulate Protein Dynamics With High Spatiotemporal Precision

Mirjam Zimmermann, Ruben Cal, Viktor Hoffmann, Marketa Zvelebil, Edwin Constable, Florent Beaufilets\* and Matthias P. Wymann\*

## Abstract

Chemical inducers of dimerization (CIDs) are popular tools to control protein localizations and activities. To combine the advantage of a modular dimerization system with the possibility to access protein manipulation with high spatiotemporal precision, we integrated a photocleavable methylnitroveratryl (MeNV) group into the core module of a well characterized SNAP- and HaloTag reactive dimerizer (HaXS8) and succeeded in the development of the first cell permeable and photocleavable heterodimerizer, called MeNV-HaXS. The combination of the fast MeNV-HaXS-induced dimerization of SNAP- and HaloTag fusion proteins with the immediate release of the covalent complex upon cleavage of MeNV-HaXS with 360 nm light, offers a wide variety of possible reaction schemes to manipulate protein dynamics.

Here we present a detailed protocol of the MeNV-HaXS method that has been successfully used to control intracellular transport kinetics and to activate isolated signal transduction branches in complex signaling networks. Furthermore, we simulated the dimerization reactions of MeNV-HaXS with a modeling software, which revealed important insights into the dimerization behavior of MeNV-HaXS and CIDs in general. The predictions of simulated dimerization reactions under various conditions enable to perform efficient dimerization experiments and to understand how parameters can be modulated in order to optimize speed and efficiency of the dimer formation.

## Key words

Chemical inducer of dimerization – photocleavable dimerizer – HaloTag – SNAP-tag – manipulation of protein dynamics with high spatiotemporal precision – simulation of dimerization reactions

## Introduction

The biological activity of proteins is often restricted to a subcellular compartment and depending on their intracellular localizations, proteins can have different functions. The ability



of cells to activate specific signaling proteins in a subcellular region for a limited time, explains how cells are able to generate a large repertoire of signaling pathways out of a limited set of signaling proteins. A common mechanism to induce the activation of a signaling protein and its downstream signaling pathways represents the translocation of signaling proteins from the cytosol to internal membranes or the inner leaflet of the plasma membrane.

Investigation of mechanisms lying beyond a cellular event depends on tools, which allow one to systematically perturb single proteins within a complex cellular system and to analyze the linked phenotypic response. A popular approach to control localization and activation of single proteins is based on chemical inducers of dimerization (CIDs). The simultaneous interaction of the dimerizer with two proteins of interest (POIs) induces their proximity, which can be exploited to manipulate localization and activity of POIs and to modulate the function of associated cellular processes. CIDs are either used to induce dimerization of two POIs (integrated in effector unit 1 and 2) that require the presence of each other for their function. Examples for this so-called “mutual dimerization” approach are proteins that form cytosolic complexes or transmembrane surface receptors such as receptors from the PDGF<sup>[1]</sup> or FGF family<sup>[2]</sup>. Another common application of CIDs is the induced translocation of a POI (integrated in the effector unit) towards an anchor unit. The CID-induced translocation of a selected protein in this so-called “effector translocation” approach is achieved through the co-expression of an anchor unit, which determines the targeted subcellular localization, and an effector unit, which contains the POI. Presence of the dimerizer results in the translocation of the cytosolic POI in the effector unit towards the anchor unit.

### Development of the HaXS method

Over the last years, various CIDs have been developed and successfully used to investigate diverse aspects of cell biology, such as intracellular calcium and cAMP signaling<sup>[2]</sup>, T cell receptor activation<sup>[4]</sup>, modulation of phospholipids on various intracellular membranes<sup>[5,6]</sup>, activation of ERK signaling or induction of GPCR internalization<sup>[7]</sup>. For a nice overview of available CIDs we refer to following reviews<sup>[8][9]</sup>.

We developed an intracellular, covalent dimerizer of HaloTag- and SNAP-tag fusion proteins, called HaXS8<sup>[10]</sup>. The SNAP-tag is a small protein (20 kDa) derived from mammalian O<sup>6</sup>-alkylguanine-DNA-alkyltransferase (AGT), which was modified in a way that it covalently attaches O<sup>6</sup>-benzylguanine derivatives to one of its cysteine residues (Cys145)<sup>[11]</sup>. The HaloTag (34 kDa) is based on a modified bacterial haloalkane dehalogenase designed to form a covalent bond with chloroalkanes<sup>[12]</sup>. HaXS8 induces protein-protein interactions and translocates proteins to various cellular compartments. Additionally, HaXS8 enables to initiate the PI3K/mTOR pathway and to induce multiplexed protein complex formation in combination with the rapamycin CID. The modular synthetic strategy of HaXS8, allows the relatively simple introduction of novel functional groups into the core module linking the Halo- and the SNAP-tag substrate in order to generate dimerizers with novel features. Through substitution of the

tetrafluorohydroquinone group in HaXS8 with the photocleavable methylnitroveratryl- (MeNV-) group, we developed a new photocleavable CID called MeNV-HaXS,<sup>[13]</sup> which displays comparable intracellular dimerization properties as HaXS8. Excitation at 360 nm cleaves MeNV-HaXS and enables to reverse the induced dimerization of the SNAP- and HaloTag fusion proteins (see Scheme 1).

In summary, MeNV-HaXS allows to control protein localizations by two independent events in a sequential manner that can be exploited for a wide range of experimental setups. Additionally, the combination of chemical-induction and light-induced cleavage of induced dimers combines the advantages of a modular approach of a genetically encodable dimerizer system with the possibility to manipulate protein dynamics with high spatiotemporal precision by light.

*Terminology.* The expression HaXS comprises both Halo- and SNAP- reactive dimerizer molecules, the photocleavable MeNV-HaXS and the non-cleavable HaXS8. It is specified if only MeNV-HaXS or HaXS8 is addressed.

#### Functional and non-functional reaction path of HaXS

HaXS follows a non-directed reaction path, meaning that HaXS can either first react with the SNAP monomer resulting in the formation of a saturated SNAP-monomer (HaXS-SNAP) or with the Halo monomer resulting in the formation of a saturated Halo monomer (HaXS-Halo). Sequential incubation of recombinant SNAP- and HaloTag fusion proteins with HaXS, revealed that only HaXS dimerizers, which have first reacted with SNAP-tag proteins can induce efficient dimer formation through the further reaction with HaloTag proteins. In contrast, no dimer formation is observed, if HaXS has first reacted with the HaloTag (Fig 1a), suggesting that the reaction of saturated HaXS-Halo proteins with SNAP-tag proteins is inefficient. Comparison of two models with HaXS8 covalently linked to either the monomeric HaloTag only (Fig 1b, left) or simultaneously to the SNAP-tag and the HaloTag (Fig 1b, right) reveals a different configuration of HaXS8 depending on its protein environment (modeling performed by M. Zvelebil, London). If HaXS8 is only linked to the HaloTag, the dimerizer molecule gets incorporated into the channel and active side cavity of the HaloTag. We assume that this integrated HaXS8 molecule is obscured from its further reaction with the SNAP-tag and thus does not result in the formation of a dimer complex. This hypothesis is consistent with results from *in vitro* experiments, which revealed that HaXS-Halo does not react with SNAP-tag proteins (Fig 1a).

Conclusively, only the reaction path in which HaXS first reacts with the SNAP monomer to form the saturated HaXS-SNAP monomer and the subsequent reaction with the unsaturated Halo monomer, results in an efficient dimer formation (functional reaction path), whereas the opposite reaction pathway results in the formation of saturated HaXS-Halo monomers that cannot further react with SNAP monomers (non-functional reaction path) (Fig 1c). The formation of these saturated HaXS-Halo monomers lower the dimerization efficiency of the



HaXS CID through limiting the amount of available free Halo monomers that can react with HaXS-SNAP monomers to form the dimer complex through the functional reaction path.

#### Improve dimerization efficiency of HaXS CID

The *in vitro* value for the rate constant of the SNAP-tag (SNAP26m) to its substrate O<sup>6</sup>-benzylguanine is around 100 times lower than the one of the HaloTag (HT7) to its chloroalkane substrate<sup>[11,12]</sup>. Even these values were measured *in vitro*, we assumed that the intracellular reaction of the SNAP-tag with its substrate occurs slower than the one of the HaloTag to its substrate. Accordingly, we assumed that the intracellular reaction of HaXS with the HaloTag that follows the non-functional reaction pathway is faster than the reaction of HaXS with the SNAP-tag, which leads to efficient dimer formation.

We hypothesized that increasing the reaction rate of the SNAP-tag and simultaneously decreasing the reaction rate of the HaloTag may improve the overall dimerization performance of HaXS dimerizers. To achieve this, we integrated a faster SNAP-tag combined with a slower HaloTag variant in the HaXS CID and expected that relative amount of HaXS dimerizers that follow the functional reaction path compared to the non-functional reaction path will be increased.

To achieve a faster reaction of HaXS with the SNAP site, we replaced the SNAPtag (SNAP26m) by a faster SNAP variant, called SNAPf<sup>[14]</sup>. SNAPf shows an up to tenfold increase in its *in vitro* reactivity towards benzylguanine substrates compared to SNAP26m<sup>[14]</sup>. In addition to SNAP26m, which carries 19 amino acids substitutions and a C-terminal deletion compared to the wild type human DNA repair protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase (hAGT), SNAPf carries ten extra mutations compared to SNAP26m<sup>[14]</sup> (Supplementary Figure 1).

To decrease the reaction rate of HaXS with the Halo site, we integrated a slower HaloTag variant. The analysis of a structure model of the HaloTag with the reacted HaloTag substrate revealed that a subset of amino acids (175, 176 and 273) has the most contact with the HaloTag substrate and thus seems to be critical for the reaction. During the optimization process of the nowadays commercially available HaloTag variant HT7 at Promega, small amino acids were integrated at position 175, 176 and 273 to achieve an opening up of the channel and the active side cavity of the HaloTag protein, which facilitated the entry of the HaloTag substrate into this tunnel. These improvements ultimately resulted in a faster reaction of the ligand with the HaloTag<sup>[15]</sup>. We speculated that amino acids with larger side chains at these critical positions 175, 176 and 273 will prevent fast and efficient ligand entry into the channel as well as an efficient interaction of the HaloTag ligand with the catalytic important amino acids at position 106 and 272 in the active side cavity, which decreases the reaction rate of the HaloTag.

Indeed, the HaloTag containing the relatively big tyrosine at position 273 (Fig 2a, amino acid 273 shown in orange, bottom) has a lower rate constant than the original HT7 with the

relatively small amino acid leucine at position 273 (Fig 2a, amino acid 273 shown in orange, top)<sup>[15]</sup>. The phenylalanine at position 272 is an important amino acid involved in the enzymatic reaction through stabilizing the covalent, alkyl-enzyme intermediate that forms between HaloTag and chloroalkanes<sup>[15]</sup>. It is hypothesized that the big side chain of tyrosine at position 273 induces a repositioning of adjacent side chains. The repositioning of the phenylalanine side chain (272) into the tunnel of the HaloTag requires its  $\sim 45^\circ$  rotation to enable entry of the HaloTag substrate and to enable an efficient interaction of the HaloTag substrate with the catalytic important amino acids 106 and 272 (Fig 2a, catalytic important amino acids 106 and 272 are shown in yellow and red) and thus explains the lower rate constant of this HaloTag variant (HT7(L273Y))<sup>[15]</sup> (Fig 2b).

From our studies with eight HaXS dimerizers (HaXS1 to HaXS8)<sup>[10]</sup> bearing different cell penetration capacities and substrate/tag reactivities we learnt that intracellular dimerization reaction of HaXS with SNAP-tag and HaloTag fusion proteins is best interpreted as a combination of substrate/tag reactivity and limited diffusion of the dimerizer into intracellular space. The fact that extended time (minutes) is required to reach relevant yields of intracellular dimers as compared to reported single sided reaction rates of the HaloTag ( $3 \cdot 10^6 \text{ M}^{-1} \text{ s}^{-1}$ ) and SNAP-tag ( $3 \cdot 10^4 \text{ M}^{-1} \text{ s}^{-1}$ ) with their specific substrates<sup>[16]</sup>, additionally supports the view that diffusion is an important limitation of the dimerization reaction. Since the diffusion force of HaXS from the extracellular space into the intracellular space is directly dependent on the extracellular concentration of HaXS, the extracellular HaXS concentration directly affects the dimerization efficiency. Thus, dimerization performance of the HaXS CID cannot only be improved through optimizing the combination of tag variants, but also through choosing the optimal dimerizer concentration.

To confirm the hypothesis that dimerization performance of the HaXS CID can be improved through the use of a slower HaloTag in combination with a faster SNAP-tag and to find the optimal concentration to induce most efficient dimer formation, we incubated HeLa cells co-expressing SNAPf-GFP and HT7(L273Y)-GFP (optimized tag combination) resp. SNAP26m-GFP and HT7-GFP (standard tag combination) with an increasing concentration of HaXS (50 nm to 100  $\mu\text{M}$ ) for 5 min before cells were lysed and dimer formation was analyzed by SDS-PAGE and Western Blotting. Quantification of dimerization efficiencies revealed that the optimized tag combination (Fig 2a, circles) dramatically increases the dimerization efficiency compared to the standard tag combination (Fig 2b, squares), confirming that the increased reaction rate of the HaXS with the SNAP site combined with the decreased reaction rate of HaXS with the Halo site improves the dimerization performance of the HaXS CID. Furthermore, the quantification reveals that a concentration of around 5  $\mu\text{M}$  is optimal to induce most efficient dimer formation after a short (5 min) incubation with HaXS. An improved dimerization performance with the optimized tag combination is also observed if dimerization is induced with 0.5  $\mu\text{M}$  HaXS for 15 min (Supplementary Figure 1d).

## Simulation of HaXS-induced dimerization with CellDesigner

### Validation of model with results from intracellular dimerization experiments

We used CellDesigner 4.3, a software package that enables to simulate molecular interactions, to model the HaXS-induced dimerization reactions in order to improve the understanding of the dimerization behavior of the HaXS CID and CIDs in general. The graphical setup of the dimerization reactions and kinetic laws to simulate the HaXS-induced dimerization are shown (Supplementary Figure 2). The transport of HaXS into the intracellular space cells is described as diffusion across the membrane [ $\text{Diff} = P_{\text{Diff}} * (\text{Halo}_{\text{in}} - \text{Halo}_{\text{out}})$ ], whereas reactions of HaXS with the SNAP resp. the HaloTag are described with mass-action kinetics that defines chemical reaction rates as a product of a rate constant and the concentrations of the reactants (Supplementary Figure 2b).

Model 1 simulates the intracellular dimerization efficiency between SNAP and Halo monomers with the standard tag combination (SNAP26m and HT7). The value for the intracellular SNAP rate constant ( $k_{\text{S, std}}$ ) was set to 1800, and the Halo rate constant ( $k_{\text{H, std}}$ ) was set to 4500. The resulting ratio of the intracellular rate constants is 0.4 ( $r_{\text{std}} = k_{\text{S, std}} / k_{\text{H, std}} = 0.4$ ). Model 2 simulates the intracellular dimerization of SNAP and Halo monomers with the optimized tag combination (SNAPf, HT7(L273Y)). The intracellular rate constants of the faster SNAP ( $k_{\text{S, opt}}$ ) was increased by a factor of 10 and set to 18'000, whereas rate constant of slower Halo ( $k_{\text{H, opt}}$ ) was decreased by a factor 5 and set to 900. The resulting ratio of the intracellular rate constants is 20 ( $r_{\text{opt}} = k_{\text{S, opt}} / k_{\text{H, opt}} = 20$ ).

To confirm the validity of Model 1 and Model 2, the simulated dimerization efficiencies were compared with experimental data received from intracellular analysis of dimerization efficiencies of SNAP- and HaloTag fusion proteins with the standard tag combination (SNAP26m, HT7) resp. with the optimized tag combination (SNAPf, HT7(L273Y)). The comparison of experimental and model data revealed a relatively good match of modeled and experimental dimerization efficiencies with both tag combinations (Fig 3a, compare black curve (circles) with red dots for optimized tag combination and black curve (squares) with blue dots for standard tag combination), suggesting that these models can be used to predict how parameters of the HaXS CID can be modulated in order to optimize speed and efficiency of dimer formation.

### Improve dimerization efficiency of HaXS CID through modulating kinetic parameters of SNAP- and HaloTag

First, we analyzed whether the use of tag variants that display a higher ratio of their rate constants than  $r_{\text{opt}} = 20$  (as defined in Model 2) can further improve the dimerization performance of the HaXS CID (after 5 min HaXS incubation).

The ratio of the rate constants  $r_{\text{opt}}$  can be increased either through increasing the rate constant of SNAP ( $k_{\text{S}}$ ) or through decreasing the rate constant of Halo ( $k_{\text{H}}$ ). The correlation of the modeled dimerization efficiencies with an increasing ratio  $r_{\text{opt}}$ , while keeping  $k_{\text{S, opt}}$

constant ( $k_{\text{SNAP, opt}} = 18'000$ ) and varying the rate constant  $k_{\text{H, opt}}$  (Fig 3b, vertical half filled circles) was analyzed separately from the correlation of the modeled dimerization efficiencies with an increasing ratio  $r_{\text{opt}}$ , while keeping  $k_{\text{H, opt}}$  constant ( $k_{\text{H, opt}} = 900$ ) and varying the rate constant  $k_{\text{S, opt}}$  (Fig 3b, horizontal half filled circles). The analysis of the modeled curve with constant  $k_{\text{S, opt}}$  revealed that an increased ratio  $r_{\text{opt}}$  resulting from a decreasing rate constant  $k_{\text{H, opt}}$  ( $k_{\text{H, opt}}$  from 18'000 to 900, vertical half filled circles) initially increased the dimerization efficiency, until  $k_{\text{H, opt}}$  reached values lower than 900, after which dimerization efficiency is decreased. This is explained through the fact, that the rate constant of the Halo affects dimerization efficiency through two opposing mechanisms. First, a slower Halo decreases the amount of HaXS dimerizer that follows the non-functional reaction path, which results in less saturated and blocked Halo monomers and more free Halo monomers that are available to induce dimer formation through the reaction with saturated HaXS-SNAP monomers and thus results in increased dimerization efficiency. Second, at the same time a slower Halo negatively affects the overall dimerization performance as the lower rate constant  $k_{\text{H, opt}}$  decreases the speed of the reaction of HaXS-SNAP with Halo. Consequently, a compromise between a decreased amount of blocked Halo monomers, which result in higher dimer formation due to more available free Halo monomers that can be integrated in the functional reaction pathway and a slower reaction of HaXS-SNAP with Halo, which results in slower dimer formation has to be found. The second possibility to increase the ratio  $r_{\text{opt}}$  through keeping  $k_{\text{H, opt}}$  constant ( $k_{\text{H, opt}} = 900$ ) and varying the rate constant  $k_{\text{S, opt}}$  circumvents the problem of lowering the reaction speed of HaXS-SNAP with Halo through integration of a slower Halo and thus seems to be more promising for improving the overall dimerization performance of the HaXS CID. The analysis of the curve with the fixed Halo rate constant ( $k_{\text{H, opt}} = 900$ , horizontal half filled circles), reveals that increasing the ratio  $r_{\text{opt}}$  through increasing of  $k_{\text{S, opt}}$  increases dimerization efficiency until a plateau is reached, after which further increase of  $k_{\text{S, opt}}$  only result in very modest increase of the dimerization efficiency.

We hypothesized that combinations of slower Halo variants ( $k_{\text{H, opt}} < 900$ ) together with faster SNAP variants ( $k_{\text{S, opt}} > 18000$ ), which allow more HaXS molecules to follow the functional reaction path and at the same time counteracts the effect of the slower reaction of Halo with SNAP-HaXS through the use of a faster SNAP, will improve the dimerization performance of the HaXS CID. However, the simulated dimerization efficiencies of various combinations of slower Halo variants with faster SNAP variants did not result in significantly increased dimerization efficiencies (modeling data not shown), suggesting that under the tested conditions (concentrations decreased to  $0.5 \mu\text{M}$ , incubation time decreased to 1 min) a further improvement of the dimerization performance through modulation of kinetic parameters of SNAP- and HaloTags is not expected and that the limits of the HaXS CID are reached with the optimized tag variants.

### Improve dimerization efficiency through preventing the incorporation of HaXS into monomeric HaloTag

The reaction of the HaloTag with free HaXS dimerizer that has not previously reacted with the SNAP-tag, results in the incorporation of the whole molecule into the HaloTag channel, which prevents the completion of the dimer formation. This integrated HaXS in the HaloTag lowers the dimerization performance of the HaXS CID (Fig 1a). Next we analyzed whether the dimer formation of the HaXS CID could be improved through eliminating these masked HaXS in the HaloTag, which then enables HaXS to induce dimer formation through both reactions pathways. To prevent the integration of HaXS into the HaloTag, the amino acid 176, which directly lies in the HaloTag channel, could be substituted by a bulkier amino acid. We hypothesize that the resulting narrower channel could prevent the integration of the HaXS molecule into the channel and thus result in HaXS-Halo monomers that can further react with SNAP monomers. As for Fig 6b the correlations of modeled dimerization efficiencies with an increasing ratio of the rate constants  $r_{opt}$ , either while keeping  $k_{H, opt}$  constant (Fig 6c, left, horizontal half filled circles) while varying the rate constant of  $k_{S, opt}$  or while keeping  $k_{S, opt}$  constant (Fig 6c, right, vertical half filled circles) and varying the rate constant of  $k_{H, opt}$  were analyzed separately. The analysis revealed that solving the problem with the non-functional reaction path, only increases the dimerization efficiency if the ratio  $r_{opt}$  is below 20 (as defined in Model 2), whereas only little improvement of dimerization efficiencies are observed if the ratio of the rate constants  $r_{opt}$  is higher than 20. Conclusively, we assume that preventing the integration of HaXS into the channel of the HaloTag, does not dramatically increase dimerization efficiency of the HaXS CID with the optimized tag combination presently integrated in the HaXS CID and under these tested conditions (2.5  $\mu$ M HaXS for 5 min).

### Model effect of dimerizer concentration on dimerization efficiency

As discussed above, not only substrate/tag reactivities, but also the diffusion of HaXS into the intracellular compartment affects the dimerization efficiency. Quantification of the intracellular dimerization efficiency revealed that maximum dimerization efficiency is only achieved with an optimal HaXS concentration (Fig 2b), whereas concentrations below or above the optimal concentration negatively affect the dimerization efficiency. We next modeled the influence of the HaXS concentration on the dimerization efficiency. The analysis revealed that dimer formation occurs slowly at concentrations below an optimal concentration (Fig 3d, left, yellow curves), likely due to the diffusion of the dimerizer into the intracellular space, which limits the dimerizer availability. In contrast, too high concentrations of the HaXS (Fig 3d, right, orange curves) result in a limited overall dimerization efficiency, likely due to the generation of a high number of saturated HaXS-SNAP- and HaXS-Halo monomers. These two models confirm the importance of determining the optimal dimerizer concentration at which the dimerization reaction is faster than the competing saturation reaction of the monomeric tags and at which dimer formation occurs sufficiently fast.

In summary, the relatively good fit of the cellular and model data suggest that the model (see Supplementary Information, HaXS\_Model\_CellDesigner.xml) can be used for further predictions dealing with the performance of the HaXS CID or CIDs in general.

*Terminology.* For the following text and figures, Halo and SNAP refers to the optimized tag combination, SNAPf with HT7(L273Y) and it is specifically indicated if the standard tag combination is addressed.

### Applications of the MeNV-HaXS method

Signaling events operate at defined intracellular locations and depending on their cellular context, signaling proteins can have different functions. A prominent example is the activation of Ras proteins at the plasma membrane that induces membrane ruffling, while the activation of Ras at the Golgi has no effect on cell morphology<sup>[17]</sup>. We validated the HaXS CID to target tagged proteins to various intracellular localizations. HeLa cells co-expressing an effector unit and an anchor unit containing a targeting motif, which attaches this anchor unit to the cytoplasmic face of the targeted subcellular organelle or compartment were exposed to HaXS, which results in the HaXS-induced translocation of the effector unit towards the anchor unit. HeLa cells co-expressing the Golgi anchor unit Halo-RFP-Giantin, and the cytosolic effector unit SNAP-mTFP1 were treated with HaXS, which efficiently induced translocation of the effector unit from the cytoplasm towards the cytosolic surface of the Golgi membrane (Fig 4a). Accordingly, HaXS-induced translocation of cytosolic effector units towards endosomes, lysosomes, the plasma membrane, mitochondria and the actin skeleton is shown (Fig 4b), demonstrating that HaXS-induced protein translocation enables to study the function of selected proteins in various subcellular compartments. Furthermore, HaXS-induced manipulation of protein localization can be used as strategy to activate or inactivate a signaling transduction branch through controlling the subcellular localization of a signaling protein, which functions upstream of the targeted signaling pathway. HaXS-induced translocation of a signaling protein away from their functional cellular compartment towards a non-functional compartment (Fig 5a, left) resulting in the inactivation of the downstream signaling pathway or vice versa, the translocation of a signaling protein away from a non-functional cellular compartment towards a functional place (Fig 5a, right) resulting in the activation of the downstream signaling pathway, was validated as a proof-of-concept for the HaXS-induced manipulation of isolated signaling transduction branches in complex signaling networks, without affecting other signaling pathways.

The constitutively active Ras GTPases (RasV12) leads to the activation of the downstream MAPK signaling pathway independent of growth factors and other upstream stimuli. The hypervariable region (HVR) at the C-terminus of Ras proteins determines their intracellular localization. Ras proteins display asymmetrically distributions over cellular membranes and a shuttle between endomembranes and the plasma membrane, where they regulate most important signaling functions. At steady state, H-Ras is strongly enriched at the plasma

membrane and endosomes<sup>[18]</sup>. Exposure of HaXS to HEK293 cells co-expressing HRasV12 fused to a SNAP-tag, a GFP and a HA-tag together with a mitochondrial anchoring unit (Mito-Halo<sub>2x</sub>) induced the translocation of correctly localized hyperactive SNAP-GFP-HA-HRasV12 away from its place of function towards a non-functional compartment, resulting in the inactivation of the MAPK pathway as monitored by the decrease of MAPK phosphorylation (see Fig 5b, left). In the reverse approach, the dimerizer-induced activation of a signaling protein was demonstrated through the HaXS-induced translocation of p110 $\alpha$  to the plasma membrane, which resulted in the activation of the PI3K pathway. HEK293 cells co-expressing SNAP-GFP-CAAX, which serves as plasma membrane anchor unit and a HaloTag protein fused to a GFP and the inter-SH2 domain of p85 of PI3K (Halo-iSH2-GFP) were exposed to HaXS resulting in a rapid and efficient translocation of Halo-iSH2-GFP to the plasma membrane. The inter-SH2 domain recruits the endogenous p110 $\alpha$  domain to the plasma membrane, where p110 $\alpha$  meets its substrate phosphatidylinositol(3,4,)P<sub>2</sub> and converts it into phosphatidylinositol(3,4,5)-trisphosphate [PtdIns(3,4,5)P<sub>3</sub>], a lipid that serves as a plasma membrane docking site for downstream signaling enzymes. HaXS-induced translocation of p110 $\alpha$  to the plasma membrane initiated the PI3K/mTOR pathway as monitored by Thr308 and Ser473 phosphorylation on PKB<sup>[10]</sup> (see Fig 5b, right), whereas MAPK phosphorylation was not increased (data not shown)<sup>[10]</sup>. In summary, we demonstrated that the HaXS CID enables the selective manipulation of isolated signaling branches out of complex signal networks.

The photocleavable MeNV-group in the dimerizer is efficiently cleaved upon exposure to UV light, which results in an instantaneous release of dimerized proteins. The combination of the chemical induced dimerization followed by the subsequent UV light-induced release of the formed dimers, can be exploited to sequester a POI to any organelle away from its place of function, followed by subsequent UV illumination of the formed dimer complex, which immediately releases the trapped POIs. This enables to investigate translocation kinetics and to analyze the associated cellular effects upon their translocation back to their functional compartment. HeLa cells co-expressing a CFP-SNAP fused to a nuclear localization sequence (NLS-CFP-SNAP) and DsRed-Halo fused to a nuclear export sequence (NES-DsRed-Halo) were exposed to MeNV-HaXS. Before MeNV-HaXS treatment, NLS-CFP-SNAP shuttles between the nucleus and the cytoplasm. Upon its dimerization with NES-DsRed-Halo the formed dimer exceeds the size for passive nuclear import resulting in a retention NLS-CFP-SNAP in the cytoplasm. UV-induced cleavage of the formed dimer complex results in the immediate release of NLS-CFP-SNAP, which enables the observation of the nuclear re-import kinetics in real time (Fig 6). In contrast to most kinetic studies of nuclear transports, which have been performed *in vitro* using permeabilized cell system, this simple experimental setup allows to study nuclear import rates in living cells.

## Comparison of the HaXS CID with other CIDs

A wide variety of CIDs have been developed and were successfully used to investigate diverse aspects of cell biology. However, none of them works without limitations and the choice of the CID that suits best to your planned application is critical. Table 1 summarizes all advantageous features that are unified in the HaXS CID and alternative CIDs with the same feature are indicated. In the second row, examples of CIDs that display opposite features are listed.

Advantageous features unified in HaXS CID	Disadvantageous features of other CIDs
Cell-permeable dimerizer	Non-cell permeable dimerizer
→ Investigation of intracellular events	→ Investigation of extracellular events or in cell lysates
Alternatives to HaXS CID: most CIDs	Examples: S-CROSS <sup>[11]</sup>
Induces covalent interaction	Induces non-covalent interaction
<ul style="list-style-type: none"> <li>→ Simple monitoring of dimerization efficiency under denaturing conditions (immune blot analysis)</li> <li>→ Efficiently shortens time for construct optimization (most time consuming part of CID experiments)</li> </ul>	<ul style="list-style-type: none"> <li>→ Monitoring of dimerization efficiency only by analysis of microscopic translocation experiments or through a time-consuming analysis of expected cellular outputs</li> </ul>
Alternatives to HaXS CID: S-CROSS <sup>[11]</sup>	Examples: most CIDs
Dimerization of freely diffusible proteins	Dimerization only of pre-associated proteins
→ Induction of protein-protein interactions	<ul style="list-style-type: none"> <li>→ No induction of protein-protein interactions,</li> <li>→ Works as proximity sensor only</li> </ul>
Alternatives to HaXS CID: most CIDs	Examples: S-CROSS <sup>[11]</sup> , CoDiS <sup>[19]</sup> , X-CrAsH <sup>[20]</sup>
No interference with endogenous proteins	Interference with endogenous proteins
<ul style="list-style-type: none"> <li>→ No interference with process under investigation expected</li> <li>→ Enables the investigation of cellular events without disturbance induced by the dimerizer</li> </ul>	<ul style="list-style-type: none"> <li>→ Disturbance of cellular events due to high affinities of dimerizer with endogenous proteins</li> <li>→ Cytotoxic effects</li> <li>→ Restricted use to a number of limited applications (e.g. short time treatment)</li> </ul>
Alternatives to HaXS CID: most CIDs	Examples: rapamycin CID, some rapalog CIDs (iRap), ABA CID <sup>[21]</sup>



Fast dimerization on a timescale of second to minutes	Slow dimerization on a timescale of minutes to hours
→ Allows analysis of cellular events occurring on a short timescale	→ Enables mimicking of slower cellular events only
Alternatives to HaXS CID: rapamycin CID, gibberlin analogue GA3-AM CID <sup>[22]</sup>	Examples: ABA CID <sup>[21]</sup>
Reversibility of MeNV-HaXS induced dimerization	Non-reversible dimerization
→ Reversible control signaling events, enables to mimic physiological signaling pathways (e.g. phosphorylation) → Versatile experimental setups (such as re-inducible knock sideways)	→ No mimicking of reversible signaling events
Alternatives to HaXS CID: → Dual translocation strategy <sup>[23]</sup> (complicated construct setup) → Competitor of dimerizer <sup>[24][25]</sup> (slow dimerization and/or reversion of dimerization, > 10 min)	Examples: most CIDs, either due to covalent (S-CROSS) or high affinity (rapamycin) interactions of the dimerizer with the respective dimerizing domains

#### Comparison of the HaXS CID with most popular CID based on rapamycin

FKBP and FKBP-rapamycin-binding (FRB) domain of mammalian target of rapamycin (mTORC1) domains in the dimerizing constructs of the rapamycin CID compete with endogenous FKBP12 and mTORC1 for rapamycin binding. Binding of FKBP12-rapamycin to the FRB domain of mTORC1, renders the mTORC1 complex enzymatically inactive. Conclusively, the use of rapamycin can have unwanted side-effects on the regulation of cell growth, proliferation or autophagy, rendering the analysis of signaling pathways challenging. Additionally, feedback actions as the activation of PKB upon mTORC1 inhibition by rapamycin<sup>[26]</sup> can further complicate analysis of rapamycin-induced dimerization effects in living cells. However, the use of the rapamycin CID in cultured cells can still be justified, if the experimental time is shorter than the time required for rapamycin to induce its potential toxic effects. If longer experimental times are required or work in living animals is performed, it is advisable to use non-toxic rapamycin derivatives, known as rapalogs. Most rapalogs are derivatized at the C16 position of the rapamycin core through the introduction of groups with bulkier substituents, which prevents the interaction with the wild type FRB domain of mTORC1<sup>[27]</sup>, but enables to bind a modified FRB domain (FRB T2098L). However, 25 times higher concentrations of the commercially available rapalog (AP21967, Clontech) than of rapamycin have to be used in order to achieve comparable dimerization efficiency<sup>[28]</sup>, likely

due to the lowered cell penetration capacity of AP21967 resulting from the derivatization of the rapamycin core.

In contrast, the HaloTag originates from a genetically modified derivative of a hydrolase protein of prokaryotic origin, thus endogenous activities are absent in mammalian cells. The SNAP-tag substrate is designed to interact with a mutated version of the human DNA repair protein O<sup>6</sup> (hAGT), which prevents it from interacting with endogenous proteins. Conclusively, due to the tag choice in the HaXS CID no interference with endogenous signaling pathways is expected.

Despite the cytotoxic properties of rapamycin, the rapamycin CID with its excellent binding kinetics is still the most widely implemented CID. The high rate constants of the rapamycin reaction with FKBP12 ( $k_a = 5.8 \times 10^6 \text{ M}^{-1} \cdot \text{sec}^{-1}$ <sup>[29]</sup>) and FKBP-rapamycin-binding to FRB ( $k_a = 1.7 \times 10^6 \text{ M}^{-1} \cdot \text{sec}^{-1}$ <sup>[29]</sup>) allow rapamycin to induce dimerization between FKBP12 and FRB with excellent kinetics and efficiency. Additionally, as the penetration of rapamycin through cellular membranes is very fast, low concentrations (200 nM) are sufficient to induce fast and efficient dimerization of FKBP and FRB fusion proteins.

Overall, the speed as well as the efficiency of HaXS-induced dimerization is lower compared to rapamycin-induced dimerization. Even though the amount of HaXS-induced dimerization exceeds values required for many applications, as for example for a surface receptor, which recruits cytosolic signaling molecules<sup>[30]</sup>, some applications demand for a faster and more efficient dimerization, and the use of the rapamycin CID is recommended. In contrast to the HaXS CID, in which too high concentrations of the dimerizer prevent an efficient dimer formation (Fig 2b), increasing rapamycin concentrations result in an increasing amount of formed dimers until a plateau is reached and too high rapamycin concentration will never result in a lowered dimerization efficiency. This is explained by the directed reaction mechanism of rapamycin. Rapamycin can only bind to the FRB domain if it has previously reacted with FBKP12. The directed reaction path of rapamycin only allows the formation of saturated-FKBP monomers, but no saturated-FRB monomers. Too high amounts of saturated monomers in a CID are only problematic if both species of monomers can be saturated, which prevents them from dimerizing with each other and thus lower the dimerization efficiency (Supplementary Figure 3).

#### Limitations of the HaXS method

As true for all CID systems, the design of constructs bearing the dimerizing domains is time consuming and requires careful consideration and testing for every single application, which requires a certain time to be invested until constructs are ready before performing the actual HaXS dimerization experiment. Particularly, experiments in which HaXS is used to mimic a physiological interaction between two POIs (such as in the “mutual dimerization” approach) or a POI and a targeted organelle/membrane (such as in the “effector translocation” approach) are challenging and many constructs have to be tested until they are suited for a functional

interaction. Sterical constraints resulting from the presence of proteins at the targeted membrane or in the Halo- and SNAP-tag domain constructs can affect the functionality of the interaction. Less challenging is the design of constructs in the „effector translocation” approach used to inactivate proteins through their translocation away from their functional compartment, as there is no need to mimic a functional interaction.

The endogenous version of the overexpressed POI that is targeted by the HaXS CID or any other CID in general can hamper the performance and analysis of the experimental outcome in some functional studies. For some applications it can be sufficient to overexpress the POI in the effector unit, which outcompetes the endogenous version of the POI. However, the high expression level of the effector unit can either have toxic effects on the cells or lead to a saturation of the anchor unit resulting in inconclusive results. In some application, the use of constitutively active mutants is sufficient to outcompete the effects of the endogenous POI. A more physiologic approach is to control a regulatory protein of the POI with the CID (such as putting the GAP or GEF of G proteins under HaXS control), which enables to indirectly control the endogenous levels of the POI. For some applications it can be important to deplete the endogenous version of the POI. One strategy relies on the knockdown of the endogenous POI through the transfection of a siRNA constructs against the endogenous POI while using a siRNA-resistant POI in the effector unit construct.

Another limitation of the HaXS CID or any other CID in general results from unwanted background activation resulting from the access of effector units to the target prior to the HaXS treatment. The overexpressed effector units can have access to their functional compartments or to another effector units as a consequence of their free diffusion in the cytoplasm. This leads to an elevated background activity before chemical induction, which perturbs the cellular basal state. One strategy to restrict the free diffusion of the effector units in the cytosol is to trap them inside the nucleus until the dimerizer is applied. However, this strategy suffers from a slower onset of the dimerization-induced activation, as the effector unit first has to be exported from the nucleus<sup>[28]</sup>. Another strategy to circumvent the unwanted background activity without compromising the extent of induced activation was demonstrated in the lab of T. Inoue. They initially sequestered the POI to a non-functional organelle through exploiting a weak interaction between a membrane lipid and its binding protein; particularly they used the PH domain of a family of four-phosphate-adaptor proteins (FAPPs), which binds to phosphatidylinositol 4-phosphate (PI4P) lipid that is enriched at the trans-Golgi. As FAPP binds to PI(4)P with a dissociation constant of 230 nM, most of the effectors units fused with the PH(FAPP) domain remain on the Golgi surface until the presence of the dimerizer rapamycin, which subsequently traps the effectors at the FRB anchor unit at the plasma membrane<sup>[29]</sup>.

As seen in Figure 1a, the dimerization efficiency of recombinant Halo- and SNAP-tag fusion proteins is very low compared to the excellent intracellular dimerization efficiencies. Optimization of the buffer conditions (pH, salt concentrations, detergents, reducing agents)

dimerization efficiency of the HaXS CID. Thus, HaXS CID is not suitable for *in vitro* experiments.

Furthermore, limited expression levels of one or both dimerizing constructs directly affect the outcome of the HaXS CID or CIDs in general. Thus, cell lines that are difficult to transfect are not recommended to use with the HaXS CID.

### Protocol overview

The workflow of HaXS experiments is divided into 4 main steps (see STEP 1 to 4). To achieve optimal results four parameters of the HaXS CID need to be optimized: the design of the constructs containing the dimerizing domains (STEP 1), the cellular system (STEP 2), the MeNV-HaXS treatment (STEP 3) and the conditions to induce photocleavage of MeNV-HaXS-dimers (STEP 4). In order to analyze the success of each single parameter, protocols for simple test experiments are described (see TEST EXPERIMENT 1 to 4). Furthermore, troubleshooting strategies to optimize each step are provided (see TROUBLESHOOTING TABLE).

*Comment. The photocleavable dimerizer MeNV-HaXS is only implemented in experiments, if induced-dimers are subsequently cleaved by UV light. To perform control experiments or the test experiments and for all applications with dimers that will not be cleaved, the light-insensitive HaXS8 can be used.*

### STEP 1. Creation of SNAP- and HaloTag expression constructs

One constructs containing the SNAP-tag and a second constructs containing the HaloTag need to be designed. Depending on the targeted approach, either two effector units (“mutual dimerization”) or an effector unit and an anchor unit (“effector translocation”) are required. The anchor unit consists of one of the two dimerizing domains, a fluorescent reporter for the visual control of the localization of the anchor unit and to perform co-localization studies with the effector unit as well as an anchoring motif, which controls the localization of the anchor unit through the attachment to the cytoplasmic face of the targeted organelle. The effector unit consists of one of the two dimerizing domains (SNAP- or HaloTag protein), the POI as well as a fluorescent reporter to track the localization and the HaXS-induced translocation.

### Library of ready-to-use anchor units

We designed a small library of ready-to-use plasmids encoding anchor units targeting various intracellular organelles such as the Golgi, endosomes, lysosomes, plasma membrane, mitochondrion, actin skeleton as well as subcellular compartments as the nucleus and space outside the nucleus (Fig 7). The Golgi targeting motif derived from Giantin (C-terminal 128 aa fragment of Giantin, 3131-3259) was used to target the Halo-RFP-Giantin to the Golgi<sup>[17]</sup>, the last amino acids of Rheb1 (C-terminal 15 aa fragment of Rheb1, called Rheb15) were used to target early and late endosomes (Rheb15-Halo-RFP), the lysosomal-associated membrane protein (LAMP1) is used to target Halo-RFP-Lamp1 to the surface of lysosomes<sup>[17]</sup>, the CAAX

box (polybasic isoprenylation sequence) of K-Ras4B targets SNAP-GFP-CAAX to the plasma membrane, the import signal of the yeast mitochondrial outer membrane protein Tom70p target Mito-SYFP-SNAP on mitochondria, the small peptide LifeAct (N terminal 17 aa fragment of actin binding protein Abp140, 1 - 17) serves as an anchor for the filamentous-actin (F-actin) structures, CFP-SNAP equipped with a classical nuclear localization sequence (NLS) from SV40 T large antigen<sup>[33]</sup> (NLS-CFP-SNAP) is retained in the nuclear compartment whereas DsRed-Halo equipped with a nuclear export sequence (NES) from HIV I Rev is excluded from the nuclear compartment. An overview of these anchor units used in combination with the rapamycin or the HaXS CID system is shown in Table 2.

Organelle	Targeting motif	Rapamycin CID	HaXS CID	Source
Golgi	128 aa C-terminal fragment of Giantin (3131 – 3259) <sup>[34]</sup>	FRB-YFP-Giantin	Giantin-Halo-RFP <sup>[13]</sup>	FRB-YFP-Giantin from Takanari Inoue, Stanford
Lysosome	Lysosomal-associated membrane protein 1 (LAMP1), (1-417)	LAMP1-ECFP-FRB <sup>[17]</sup>	LAMP1-RFP-Halo <sup>[13]</sup> LAMP1-CFP-FRB <sup>[10]</sup>	LAMP1-ECFP-FRB from Takanari Inoue, Stanford
Early endosomes	15 aa C-terminal fragment of Rheb1, Rheb15 <sup>[35]</sup>	not used in combination with rapamycin CID	Halo-RFP-Rheb15 <sup>[13]</sup>	Rheb1 from Mike Hall, Basel
Plasma membrane	polybasic isoprenylation sequence from KRas-4B	FRB-CAAX <sup>[6]</sup>	Halo-GFP-CAAX <sup>[13]</sup> SNAP-GFP-CAAX <sup>[10]</sup>	Labaratroy plasmid collection, Matthias Wymann, Basel
Mitochondrium	Tom20 (1-33) <sup>[28]</sup>	Mito-YFP-FRB <sup>[28]</sup>	Tom20-Halo, Tom20-Halo <sub>2x</sub> , Tom20-GFP-Halo, Tom20-GFP-Halo <sub>2x</sub> , unpublished data, Mito-RFP-Halo <sup>[13]</sup>	Tom20-Halo, Tom20-Halo <sub>2x</sub> , Tom20-GFP-Halo, Tom20-GFP-Halo <sub>2x</sub> , all constructs from Dominik Buser, Basel, Mito-YFP-FRB from Takanari Inoue, Standfors
Actin skeleton	LifeAct, 17 aa N-terminal fragment of actin binding protein Abp140 (1-17) <sup>[36]</sup> Actin coding sequence	not used	LifeAct-mTFP1-SNAP <sup>[13]</sup> SNAP-actin <sup>[10]</sup>	LifeAct and actin targeting sequences from Olivier Pertz, Basel
Nucleus	nuclear localization sequence (NLS) from SV40 T large antigen <sup>[33][37]</sup>	NLS-CPF-FKBP1x and NLS-GFP-FKBP <sub>3x</sub> <sup>[38]</sup>	NLS-CPF-SNAP <sup>[13]</sup>	NLS-CFP-FKBP from Stefan Hübner, Würzburg

Nuclear exclusion	nuclear export sequence (NES) from HIV I Rev <sup>[39]</sup>	NES-DsRed-FKBP12 <sup>[38]</sup>	NES-DsRed-Halo, this study	NES-DsRed-Halo from Stefan Hübner, Würzburg
-------------------	--	----------------------------------	----------------------------	---

#### Sequester POIs at mitochondrion to inactivate them

It has been demonstrated that the mitochondrion tolerates foreign proteins relatively well<sup>[40]</sup>. Additionally, the relatively large surface and the little involvement of the outer mitochondrial membrane in signaling events, makes the mitochondrion an optimal organelle to trap proteins away from their functional compartment in order to prevent them from interacting with their downstream signaling components. In contrast to genetic perturbation methods, rerouting to mitochondria is completed within minutes, thus these so-called knocksideways are around 3 to 4 orders of magnitude faster than conventional knockdowns<sup>[28]</sup>. Additionally, since proteins are not inactivated through their destruction but through their removal from their place of function, these proteins can in principle be re-activated through the UV light-induced cleavage of MeNV-HaXS, which release the sequestered proteins and enables them to translocate back to their place of function.

#### Box 1: Choice of optimal linker

The choice of the linker used to join the different protein domains has an important impact on the behavior of the constructs. Direct fusion or too short and rigid linkers between two domains can lead to misfolding of the fusion proteins<sup>[41]</sup>, impaired activity of the POI<sup>[42]</sup> or low expression levels<sup>[43]</sup>. Additionally, too short and rigid linkers can prevent efficient association of the anchor unit to the targeted organelle or prevent efficient dimerization between an anchor unit and effector unit or two effector units due to the missing flexibility required to enable an efficient dimerization between the SNAP- and HaloTag fusion proteins. The linkers between two domains must be long and flexible enough to guarantee the independent function of the two connected parts as well as to support dimerization efficiency of the anchor- and effector units. However, too long linkers in the anchor unit or in the effector unit can cause instability of the construct and can prevent a functional interaction of the POI in the effector unit with the targeted organelle. Good examples for flexible linkers are those composed of the small amino acids glycine and serine (e.g. (GGGS)<sub>3</sub><sup>[44]</sup> or (G)<sub>8</sub><sup>[45]</sup>). The small size of these amino acids allow the linkers to be flexible, whereas the polarity and hydrophilicity allow the formation of hydrogen bonds with water molecules, which increases the stability of these linkers in aqueous solutions and prevents them from unfavorable interactions with the protein domains. Beside considering the length and flexibility, the linkers should not have an influence on the connected parts, thus amino acids, which do not interact with other amino acid residues are advisable to use. The linker with putative sites for modification (such as phosphorylation or glycosylation sites), too rigid geometry (such as prolines or bulky aromatics) and excessive charged amino acids should be avoided. Too

hydrophobic amino acids are not advisable to use as they can promote self-interaction in aqueous solution.

#### STEP 2. Creation of cells co-expressing SNAP- and HaloTag expression constructs

HaXS CID has been used in HeLa and HEK293 cells<sup>[10,13]</sup> as well as in MDCK epithelial and NIH 3T3 cells<sup>[10]</sup>. In principle any cells that can be transfected with the anchor and effector unit can be used to perform HaXS experiments. If a new cellular system is used, it is important to include control experiments in which the HaXS dimerizer is added to untransfected cells and to rule out effects of the dimerizer on the cell. Transient transfections were performed with JetPei (Brunschiwig) according to manufacture's protocol.

*Alternatively, any other transfection reagents can be used to transfect the SNAP- and HaloTag expression constructs.*

#### Expression levels of anchor and effector units

For all applications that do not require the possibility to vary and adapt the expression levels of the two plasmids (such as for the "mutual dimerization" approach), it is best to insert both ORFs encoding the SNAP- and HaloTag dimerizing units in a single, bi-cistronic plasmid. Expression from a single plasmid guarantees co-transfection of both units in all cells and equal expression levels of both units. In contrast, if two plasmids are transfected, some cells are co-transfected, whereas some cells only express one of the two plasmids and the relative amount of the two expressed proteins shows high cell-to-cell variation. This results in a high variability of the amount of induced dimers and the correlated cellular effect, which makes especially single-cell experiments difficult to analyze and compare. Additionally, if a single bi-cistronic plasmid is used, in principle only a single fluorescent reporter has to be integrated to control expression levels, which simplifies construct design (Fig 8a, left).

In contrast, in the "effector translocation" approach it is important to have the possibility to vary and adapt the expression levels of the anchor and effector unit in order to optimize the ratio of the anchor unit to the effector unit, which can be important to make sure that all effector units can be translocated towards the anchor unit. Particularly, in the approach of HaXS-induced inactivation of an effector unit, the incomplete integration of effector units would result in incomplete inactivation of the POI. Thus, for the "effector translocation" approach, two separate plasmids, one encoding for the anchor unit and the other for the effector is advisable to use, as expression levels of both plasmid can be varied and adapted to experimental requirements (Fig 8a, right).

#### Single plasmid expressing both effector units

Plasmids containing a "self-cleaving" T2A peptides, allow the bi-cistronic expression of two ORFs<sup>[46]</sup>. T2A peptide sequences are derived from viruses and are used to mediate protein cleavage through a ribosomal skip mechanism, which results in the expression of two separate proteins in equimolar amounts. Major advantages of the 2A system are the small

size of the peptide (54–174 bp) and the equal expression of both proteins. Additionally, the separation activity only depends on eukaryotic ribosomes, which are highly conserved among all eukaryotes and thus works efficiently in various cell types. To analyze performance of this T2A plasmid, HeLa cells expressing SNAP-GFP-T2A-Halo-GFP were exposed to HaXS before cells were lysed. Analysis revealed that expression level of SNAP-GFP-T2A-Halo-GFP is high and dimerization efficiency between SNAP-GFP and Halo-GFP is comparable to dimerization efficiency induced by cells co-expressing two separate plasmids (Fig 8b), demonstrating that the insertion of the T2A sequence does neither negatively influence expression level nor dimerization efficiency. Analysis of the ratio of monomers (amount of Halo-GFP monomers divided by the amount of SNAP-GFP monomers, in DMSO treated cells) reveals equal expression of both monomers (Fig 8b). A weak band slightly below the size of the dimer band is visible in DMSO control cells, likely representing uncleaved dimer. However, as the ratio of induced dimer to uncleaved dimer is very low ( $0.031 \pm 0.027$ ), no problems resulting from a background activation induced by this uncleaved dimer before HAXS treatment is expected. In order to make the cleavage more efficient, two T2A sequences were integrated in the bi-cistronic backbone system. However, expression level of this 2x T2A plasmid was very low and the ratio of the induced dimer to uncleaved dimer was not improved (data now shown, all experiment with T2A plasmids were performed by Sandra Dehn).

#### Two separate plasmids to express anchor and effector units separately

If the ratio of the expression level of both plasmids has to be variable, expression of two separate plasmids, each expressing one of the two SNAP- and HaloTag dimerizing units are advisable to use. In the “effector translocation” approach, complete integration of the effector unit into the formed dimer complex is important. To analyze the possibility to vary the ratio of anchor unit to effector unit expression levels, HeLa cells were transfected with equal amounts of Halo-GFP and SNAP-GFP DNA (first two lanes, see Fig 8c, top) and increasing amount of Halo-GFP DNAP with decreasing amount of SNAP-GFP DNA (lane 3 to 8, see Fig 8c, top). 24 h later cells were exposed to HaXS, lysed and expression levels of SNAP-GFP and Halo-GFP monomers was analyzed by immune blot. The DNA ratio used for the transfection is reflected in the protein expression level and thus the variation of DNA levels enables the adaption of expression levels of the anchor- and effector unit (Fig 8c, top). The percentage of non-incorporated SNAP-GFP monomers in the dimer complex upon HaXS treatment is shown in the quantification (Fig 8c, bottom). High ratios of Halo-GFP to SNAP-GFP result in a almost complete incorporation of SNAP-GFP in the dimer complex, which is for example important for the “effector translocation” approach used to inactivate a POI through its removal from its functional compartment.



### STEP 3. HaXS treatment of cells co-expressing SNAP- and Halo expression constructs

As discussed previously, the HaXS-induced dimerization is best interpreted as a combination of substrate/tag reactivity and diffusion of HaXS into intracellular space. Since the diffusion force directly correlates with extracellular concentration of HaXS, it is obvious that maximum dimerization efficiency is only achieved at an optimal HaXS concentration. Intracellular dimerization experiments revealed that around 5  $\mu$ M HaXS induces most efficient dimerization between SNAPf-GFP and HT7(L273Y)-GFP fusion proteins and is able to induce efficient dimerization within 5 min. However, the size and conformation of the anchor and effector unit can have an effect on the dimerization speed, and longer incubation times with HaXS could be necessary in order to achieve efficient dimerization. Thus, the optimal length of the HaXS incubation has to be determined empirically for every single combination of anchor and effector units.

### STEP 4. Cleavage of MeNV-HaXS-induced dimers

As the photocleavage of MeNV-HaXS is induced with UV light, which can damage cells if the energy is too high, it is important to detect unwanted side effects exerted by UV irradiation. MeNV-HaXS was developed through the insertion of the photocleavable group methyl-6-nitroveratryl (MeNV-) into the core module linking the HaloTag- and SNAP tag reactive groups in the light insensitive HaXS8 dimerizer<sup>[10]</sup>. As expected due to the very similar chemical setup of MeNV-HaXS and HaXS8, the comparison of their intracellular performance revealed very similar dimerization behavior of both molecules. Conclusively, the light insensitive HaXS8 dimerizer is a valuable control compound to monitor potential side effects induced by UV illumination, while performing photo cleavage experiments with MeNV-HaXS. To exclude that UV light induce any MeNV-HaXS independent effects on cells, it is important to include the HaXS8 control in all experiments, especially if a new light source is used to induce cleavage. This allows for example to distinguish whether a detected light-induced loss of fluorescence results from the release of anchored proteins due to the cleavage of MeNV-HaXS or from photobleaching of the fluorescent protein.

### MeNV-HaXS is not sensitive to ambient light

Last years various dimerization systems that are based on a genetically encoded, naturally occurring photosensitive that undergo a conformational change upon illumination with light at a defined wavelength proteins were developed and proved to be very useful tools with a wide variety of applications<sup>[47-49]</sup>. However, some of these optogenetic systems suffer from the drawback that they are sensitive to accidental exposure to environmental light and experiments have to be performed in dark, which makes performance of these experiments challenging. To investigate sensitivity of MeNV-HaXS to environmental light, stock solution of MeNV-HaXS, either in DMEM medium or in DMSO, were exposed to environmental light for the times indicated (Supplementary Figure 4a). HeLa cells co-expressing SNAP-GFP and Halo-GFP were then incubated with these light-exposed MeNV-HaXS dimerizers to induce

dimer formation. HaXS8 was used as control compound. Analysis of dimer formation reveals comparable dimerization efficiencies of light-exposed HaXS8 and MeNV-HaXS, suggesting that MeNV-HaXS is stable even after extensive exposure to environmental light (Supplementary Figure 4a), demonstrating that MeNV-HaXS is not sensitive to ambient light, thus no special precautions concerning light exposure during performance of experiments are required.

#### Choice of fluorescent reporters in MeNV-HaXS system

In three commonly used optogenetic dimerizer systems, blue light is induced to induce dimerization between FKF1 and GIGANTEA (GI)<sup>[49]</sup>, between CIB1 (basic helix-loop-helix protein *Arabidopsis*) and CYR2 (cryptochrome 2)<sup>[47]</sup> or between AsLOLOV2 (LOV2 domain of *Avena sativa* phototropin 1) and the engineered PDZ domain (ePDZ)<sup>[48]</sup>. Thus a critical challenge in these systems is the choice of fluorescent reporters as many commonly used fluorescent proteins has a significant spectral overlap between the excitation wavelength for imaging and the wavelength used for the activation of the dimerization process. In contrast, the 360 nm light required to induce photocleavage of MeNV-HaXS is compatible with most fluorescent reporter proteins, thus little restriction concerning choice of fluorescent reporters exist.

#### Light sources to induce cleavage of MeNV-HaXS in living cells

General restrictions to light-inducible systems are the specialized microscopes and softwares, which are required to induce light-activated effects<sup>[50]</sup>. MeNV-HaXS offers the possibility to induce intracellular cleavage with different light sources, depending on available microscopic equipment and the targeted application. If cleavage has to be induced on single cells or subcellular region of cells with high temporal resolution, the use of microscopes equipped with an XY scanning excitation laser for FRAP (fluorescence recovery after photobleaching; 355 nm) or a 405 nm laser line from a conventional confocal microscope can be used (Supplementary Figure 4b). If neither a scanning FRAP laser nor 405 nm laser from a confocal microscope is available or if photocleavage has to be induced on a whole cell population, a global field of view illumination setup with standard DAPI excitation filters (377±25 nm; from a standard fluorescence microscope can be alternatively used to induce photocleavage and thus make the MeNV-HaXS system applicable to many labs, including them equipped with a standard fluorescence microscope only. As the energy from a fluorescence lamp is limited, the illumination time to achieve efficient cleavage has to be extended. However, an illumination time of  $t < 20$  sec was shown to be sufficient to induce efficient photocleavage<sup>[13]</sup>. Depending on the strength of the laser or the intensity of the fluorescence lamp, the optimal time required to induce efficient photocleavage varies.

### Detection of SNAP- and HaloTag fusion proteins

Due to the covalent interaction of HaXS with the SNAP- and HaloTag proteins, the intracellular dimer formation of these covalent complexes can be easily monitored by immune blotting experiments. This allows the direct correlation of the amount of formed dimers with the induced cellular output. Additionally, since constructs have to be optimized for every single application, which is the most time consuming part while performing CID experiments, a fast and simple validation of dimerization efficiencies of newly designed constructs greatly simplifies and shortens the process of construct design, and thus cuts down the whole experimental time. In contrast, many CIDs (such as the rapamycin CID) are based on non-covalent interactions, and thus render quantification of dimerization efficiency challenging, as the dimer formation can only be rated by their expected cellular output or by microscopic co-localization analysis, which demand a challenging analysis strategy to achieve quantitative results. Antibodies for HaloTag and SNAP-tag domains are commercially available. Anti-HaloTag® polyclonal Antibody (Promega, G9281) detects HaloTag fusion proteins in Western blot experiments with high sensitivity and specificity, whereas a relatively high background signal is observed while performing immune fluorescence experiments. Polyclonal Anti-SNAP-tag® Antibody (NEB, P9310S) only works in Western blot experiments but not in immune fluorescence experiments. However, while performing Western Blot experiments only high concentrations (1:500) of the primary antibody result in detectable signals and signal to noise ratio is low (personal communication Dominik Buser, Basel). Alternatively, if these SNAP- and HaloTag antibodies are not available, alternative tags (e.g. myc, HA, or fluorescent reporters) can be put on the fusion proteins and antibodies against these tags can be used for the detection. For the detection of SNAP- and HaloTag proteins before they have reacted with the HaXS dimerizer, e.g. to control expression levels or to confirm subcellular localization of a tagged fusion protein, commercially available fluorescently labeled dyes (e.g. SNAP-Cell® TMR-Star (NEB, S9105S) for SNAP-tag; HaloTag® TMR Ligand TMR Halo for HaloTag (Promega, G8251)) can be used.

## Materials

### Dimerizer molecules and solutes

HaXS8	synthesized by Olivier Jacques and Florent Beauflis, stock solution in DMSO, chemical synthesis and characterization described in Erhart et al., 2013
MeNV-HaXS	synthesized by Viktor Hofmann, Ruben Cal and Florent Beauflis, stock solution in DMSO, chemical synthesis and characterization described in Zimmermann et al., 2014
Dimethylsulfoxide (DMSO)	(Sigma-Aldrich, cat. nr. 34869-100ML)

### Kits and enzymes

Phusion DNA Polymerase	(New England Biolabs, cat. nr. 300338)
Bio-Rad Protein Assay	(Bio-Rad, cat. nr. 500-0006)
GenElute Gel Extraction Kit	(Sigma Aldrich, cat. nr. NA1111-1KT)
GenElute HP Endotoxin-Free Plasmid Maxiprep Kit	(Sigma-Aldrich, cat. nr. NA0310)
GenElute PCR Clean Up Kit	(Sigma-Aldrich, cat. nr. NA1020-1KT)
GenElute Plasmid Miniprep Kit	(Sigma-Aldrich, cat. nr. PLN-350-1KT)
Immobilion Western, Chemiluminescent HRP Substrate	(Millipore, cat. nr. WBKLS0500)
In-fusion Cloning Kit	(Clontech, cat. nr. 639649)
JetPEI DNA transfection agent	(Polyplus, cat. nr. 101B-10)
Phusion DNA Polymerase and PCR buffer	(New England Biolabs, cat. nr. 300338)
Restriction endonucleases and NEB buffers	(New England Biolabs)

### Solutions

Standard 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, 5% glycerol
10 x electrode buffer (Tris-Glycine)	30.3 g Tris-HCl (250 mM Tris-HCl), 144.2 g glycine (1.92 M), 10 g SDS (1% SDS)
10 x transfer buffer	250 mM Tris-HCl, 1.92 M glycine
1x transfer buffer	methanol added to a final concentration of 20% (v/v)
3% paraformaldehyde (PFA) in PBS (20 ml)	18 ml ddH <sub>2</sub> O and 7.5 μl 1 M NaOH added to 600 mg PFA, stirring in a hot plate until PFA is dissolved, 2 ml 10 x PBS pH 6.55 added, mixed, cooled to 37°C and pH adjusted to pH 7.2) <b>!CAUTION</b> PFA is a known human carcinogen. It is toxic on inhalation.

### Media and supplements

LB Miller	10 g NaCl, 5 g yeast extract, 10 g Bacto-Tryptone, 5 ml 1 M NaOH, filled up to 1 liter with ddH <sub>2</sub> O and sterilized by autoclaving.
LB Miller agar	LB Miller and 12.5 g agar, after autoclaving, medium cooled down to 60°C, appropriate antibiotics added, plate poured and stored at 4 °C.
1 liter SOC	20 g Bacto-Tryptone, 5 g yeast extract, 0.5 g NaCl, 2.5 ml 1 M KCl 10 ml 1 M MgCl <sub>2</sub> , 200 μl 5 M NaOH, 980 ml ddH <sub>2</sub> O, sterilized by autoclaving, addition of 20 ml sterile 1 M glucose
Ampicillin	stock 100 mg/ml in ddH <sub>2</sub> O (complete solubilization achieved by adding NaOH), used for cultures at 100 μg/ml
Kanamycin	stock 25 mg/ml in ddH <sub>2</sub> O, used for cultures at 25 μg/ml

### Mammalian cell culture media and supplements

Appropriate cell line: HeLa and HEK293 cell line ATCC	
Dulbecco's MEM (DMEM)	(Sigma-Aldrich, cat. nr. 1-26F01)
Dulbecco's MEM (DMEM) without phenol red	(Sigma-Aldrich, cat. nr. D 5921)
Fetal calf serum (FCS)	(Sigma-Aldrich, cat. nr. 10270)
100 x L-Glutamine (200 mM)	(Sigma-Aldrich, cat. nr. G7513-100ML)
100 x Penicillin-Streptomycin solution	(Sigma-Aldrich, cat. nr. 15140-122)
10 x Trypsin-EDTA solution	(Sigma-Aldrich, cat. nr. T4174-100ml)

### Equipment

Live cell microscope	Leica Live Imaging Microscope fitted with a HCX Plan-Fluotar 63x/1.4 oil, objective and a Photometrics CCD Camera CoolSnap HQ2 with Metamorph 7.1 software (Molecular devices)
Confocal microscope	Zeiss Axiovert 200M fitted with a Plan- 63x/1.4 oil objective and SM 510 Meta (Scanning Head) with LSM software

Software	CellDesigner4.3 Metamorph 7.1 LSM Fiji 1.44b Prims 6.0f
----------	---

## Procedure

### Step-by-step protocol of HaXS method (STEP 1 to STEP 4)

STEP 1. Creation of SNAP- and HaloTag expression constructs ●TIMING up to several weeks

▲ **CRITICAL STEP** The engineering of constructs is a sustainable challenge for CIDs and extensive testing and optimization of the constructs for every single application is required, which is the most time consuming part of CID experiments

1I Create expression constructs with SNAP- and HaloTag dimerizing domains

(i) Compile DNA sequences of expression constructs containing SNAP- and HaloTag domain:

“effector translocation” approach:

anchor unit: anchoring motif, SNAP- or HaloTag, fluorophore  
effector unit: POI, SNAP- or HaloTag, fluorophore

“mutual dimerization” approach:

effector unit 1: POI1, SNAP- or HaloTag, fluorophore  
effector unit 2: POI2, SNAP- or HaloTag, fluorophore

#### Design of anchor unit

-if available use a ready-to-use plasmid from anchor unit library (see Fig 7 and Table 2), if desired available for your anchor of interest (Golgi, endosomes, lysosomes, plasma membrane, mitochondrion, actin skeleton as well as the subcellular compartments outside and inside the nucleus)

-if other organelles / subcellular compartments are targeted: design new anchor units by inserting a SNAP- resp. HaloTag dimerizing domain and fluorescent reporter into a functional plasmid containing the desired anchoring motif, while retaining the context of the targeting motif as good as possible (N- versus C-terminal-fusion, linkers between domain and targeting motif)

#### Design of effector unit

-design plasmid containing the POI, SNAP- resp. HaloTag dimerizing domain and fluorescent reporter, if available take construct containing the POI, insert SNAP- resp. HaloTag domain, keep original context of POI as good as possible possible (N- versus C-terminal-fusion, linkers between domain and targeting motif)

(ii) Amplify DNA fragments using PCR primers that contain 18 – 25 bases annealing to the 5' resp. 3' end of the gene fragment, add 15 bases homologous to sequences flanking the insertion site in the backbone to facilitate In-Fusion recombination. *Alternatively, amplify DNA fragments with conventional cloning using restriction digest and T4 DNA Ligase*

(iii) Set up 50  $\mu$ l reaction for each DNA fragment to be amplified by PCR

5x Phusion HF or GC buffer	10 $\mu$ l
Template DNA (10 ng/ $\mu$ l)	1 $\mu$ l
dNTPs (10 mM)	1 $\mu$ l
Forward Primer (50 $\mu$ M)	0.5 $\mu$ l
Reverse Primer (50 $\mu$ M)	0.5 $\mu$ l
DMSO (optional, if GC > 60%)	(1.5 $\mu$ l)
Phusion DNA Polymerase	0.5 $\mu$ l (final concentration of 1 Unit/50 $\mu$ l PCR)
Nuclease-free water	fill up to 50 $\mu$ l

(iv) Amplify DNA fragments using following PCR program on Trio Thermocycler (Biometra):

Initial denaturation of DNA	98°C	120 sec
Denaturation of DNA	98°C	10 sec

Annealing	45-72°C	20 sec
Elongation	72°C	15 sec per 1 kb
30 cycles		
Final elongation	72°C	10 min
Hold	4°C	

- (v) Digest 2  $\mu\text{g}$  of plasmid DNA with amount of Units of restriction enzymes and buffer (according to manufacturer's protocol) for 1 h at the appropriate temperature in a total volume of 20  $\mu\text{l}$ .
- (vi) Analyze undigested PCR and digested backbone by agarose gel electrophoresis, gel purify with GenElute (Gel Extraction Kit Sigma Aldrich)
- (vii) Set up In-Fusion reaction in 10  $\mu\text{l}$  according to manufacturer's protocol:
- x  $\mu\text{l}$  digested backbone (10 – 200 ng)\*
  - x  $\mu\text{l}$  DNA fragment (50 – 200 ng)\*\*
  - 0.5  $\mu\text{l}$  In-fusion reaction mix
  - fill up to 10  $\mu\text{l}$  water
- \*<0.5 kb: 10-50 ng, 0.5 to 10 kb: 50-100 ng, >10 kb: 50-200 ng  
 \*\*<10 kb: 50-100 ng, >10 kb: 50-200 ng  
 Incubate the reaction mix for 15 min at 50°C.
- (viii) Use 5  $\mu\text{l}$  of In-Fusion reaction to transform  $\text{CaCl}_2$ -competent E.coli cells (XL1 blue) and select for growth by plating on LB agar plates containing the appropriate antibiotics for selection (depending on the backbone plasmid). Incubate ON at 37°C.  
**■ PAUSE POINT:** Plate can be stored at 4°C for several weeks, In-Fusion reaction mix can be stored indefinitely at -20°C.
- (ix) Pick several clones, isolate DNA using GenElute Plasmid Miniprep Kit (Sigma Aldrich) and verify clones by analytical restriction digestion and sequencing.
- Perform **TEST EXPERIMENT 1** to verify localization and translocation of anchor and effector unit.

## STEP 2. Creation of cells co-expressing SNAP- and HaloTag expression constructs

### ● TIMING ~ 24 h

▲ **CRITICAL STEP** Find the optimal ratio of expression levels of the anchor unit to effector unit, at which the effector unit can be completely integrated in the induced dimer, but at which its expression level still sufficient high to allow induction of its correlated cellular effect.

**2I** Seed and transfect HeLa or HEK293 cells with SNAP- or HaloTag expression constructs, for microscopic analysis follow STEP 2A, for analysis in cell lysates follow STEP 2B.

#### (A) Preparing cells for analysis in cell lysates

- (i) *Day 1.* Trypsinize growing HEK293 or HeLa cells using trypsin-EDTA and count the cells using a hemocytometer
- (ii) Resuspend  $8 \times 10^6$  HEK293 resp.  $1.5 \times 10^6$  HeLa cells in a total volume of 2 ml DMEM medium
- (iii) Manually pipette cell suspension in each well of a 6-well plate (Falcon)  
 ▲ **CRITICAL STEP** The number of plated cells is crucial that plated cells be attached, forming a monolayer and transfection efficiency is high (most efficient if 60-80% confluent at day of transfection)
- (ii) *Day 2.* Transient transfection of SNAP- and HaloTag expression constructs with JetPei according to manufacturers protocol: 0.5  $\mu\text{g}$  total DNA / well in 24-well plate, 2  $\mu\text{g}$  total DNA / well in 6-well plate, use ratio of amount of DNA of both constructs according to desired expression levels

#### (B) Preparing cells for microscopic analysis

- (i) *Day 1.* Trypsinize HeLa cells using trypsin-EDTA and count the cells using a hemocytometer
- (ii) Resuspend  $5 \times 10^5$  HeLa cells in a total volume of 300  $\mu\text{l}$  pre-warmed Dulbeccos's modified Eagle medium (DMEM) with 10% heat-inactivated fetal calf serum (HIFCS), 2 mM L-glutamine (Gln), 1% penicillinstreptomycin solution (PEST)
- (iii) Place 12 mm coverslips (Menzel) in 24-well plate (Falcon), to sterilize plate and coverslip irradiate plate with coverslips with UV light
- (iv) manually pipette cell suspension on coverslips in a 24-well plate and incubate the plate at 37°C, 5%  $\text{CO}_2$

▲ **CRITICAL STEP** The number of plated cells is crucial and the plated cells should be attached and to form a monolayer. Transfection efficiency is most efficient if cells are 60-80% confluent at day of transfection.

- (ii) *Day 2.* Transient transfection of SNAP- and HaloTag expression constructs with JetPei according to manufacturer's protocol: 0.5  $\mu\text{g}$  total DNA per well in 24-well plate, use ratio of amount of DNA of both constructs according to desired expression levels

→ Perform **TEST EXPERIMENT 2** to verify localization and translocation of anchor and effector unit

*Alternatively, use another cell line with suitable cell culture medium at appropriate temperature and suitable amount of cells.*

### **STEP 3. HaXS treatment of cells co-expressing SNAP- and Halo expression constructs**

**3I** Induce dimerization by incubating cells with HaXS dimerizer either following protocol for cell lysates (STEP 3A) or protocol for live cells (STEP 3B)

#### **(A) Perform experiment in cell lysates ●TIMING ~ 60 min**

- (i) *Day 3.* Dilute HaXS8 dimerizer (from 10 mM stocks in DMSO) in pre-warmed cell culture medium to an end concentration of 5  $\mu\text{M}$
- (ii) Vortex medium with HaXS dimerizer
- (iii) Add 300  $\mu\text{l}$  medium per well in 24-well and 900  $\mu\text{l}$  per well in 6-well medium containing HaXS to the cells expressing the anchor and effector unit
- (iv) Swirl plate to mix
- (v) After suitable time of HaXS incubation at 37°C perform analysis or cleavage (STEP 4A)

*Alternatively, directly pipette DMSO stock solution of HaXS into wells with transfected cells.*

#### **(B) Perform experiment in living cells ●TIMING ~ 2- 3 h**

- (i) *Day 3.* Transfer coverslips with cells co-expressing SNAP- and HaloTag expression constructs to pre-warmed Ludin chamber (Life Imaging Services), closed configuration
- (ii) Carefully add 300  $\mu\text{l}$  pre-warmed imaging medium (phenol red free DMEM cell culture medium) in Ludin chamber
- (iii) Attach tube and syringe at Ludin chamber and transfer Ludin chamber into Ludin chamber holder in microscope in the temperature incubation box
- (iv) Start imaging by live cell microscopy
- (v) Carefully add 300  $\mu\text{l}$  of phenol red free cell culture medium containing 2x HaXS dimerizer concentration
- (vi) Image dimerization until complete, take picture every 1 to 15 sec depending on effect under investigation

▲ **CRITICAL STEP** While working with a new cell line, the different cell penetration behavior of HaXS can have an effect on the optimal concentration and thus it is important to determine the optimal concentration at which HaXS induce most efficient dimerization (perform analog to TEST EXPERIMENT 3, but with concentration gradient (0.5  $\mu\text{M}$  to 10  $\mu\text{M}$ ) instead of time course for HaXS incubation)

### **STEP 4. Cleavage of MeNV-HaXS-induced dimers**

**4I** Induce cleavage of MeNV-HaXS dimers by illuminating cells or subcellular region of cells either following protocol for cell lysates (STEP 4 A) or protocol for live cells (STEP 4B)

#### **(A) Perform experiment in cell lysates ●TIMING ~ 30 min**

- (i) *Day 3.* Remove cell culture medium, wash cells twice with 1x PBS and submerge cells with 1x PBS
- (ii) Illuminate cells with UV lamp (Blak-Ray B-100A high intensity UV lamp; 100 Watt, 365 nm) for 5 min, distance of lamp around 5 cm, keep cells on ice to prevent warming of cells

#### **(B) Perform experiment in living cells ●TIMING ~ 2- 3 h**

- (i) *Day 3.* Start imaging of cells prepared until STEP 3B:  
for global photocleavage of whole cell population:
- (ii) Select cells (field of view, suited to final experiment)
- (iii) Illuminate for appropriate time with light from mercury metal halide and standard DAPI excitation filters (ex: 377 $\pm$ 25 nm, em: 447 $\pm$ 30 nm)
- for high spatiotemporal precision:
- (ii) Select single cell or subcellular region of a cell (region of interest, ROI)

(iii) Illuminate cells for appropriate time with:

-355 nm FRAP laser (diode laser, 50 mW, XY scanning excitation laser)

-405 nm laser (laser strength as medium intensity) (UV diode)

▲ **CRITICAL STEP** To exclude unbeneficial effects exerted by UV light during cleavage process, perform control experiment with light-insensitive dimerizer HaXS8 in parallel.

▲ **CRITICAL STEP** Illumination time depends on laser strength, fluorescent lamp strength and the size of the illuminated region (ROI resp. field of view).

## Protocols for Analysis (A1 to A3)

### A1. Analysis of dimerization in fixed cells ● **TIMING** ~ 2-3 h

(i) Wash cells two times with 1 x PBS

(ii) Pipette 300  $\mu$ l of 3% PFA on cells per well

(iii) Incubate for 10 minutes at 37°C

(iv) Wash cells two times with 1x PBS

■ **PAUSE POINT** Fixed cells can be stored in 1 × PBS at 4 ° C for up to few weeks, seal plate to prevent 1 x PBS to evaporate.

(v) Pipette 5  $\mu$ l of pre-warmed mounting medium (Mowiol) on the top of a microscope slide, use forceps to wash coverslips in ddH<sub>2</sub>O and to place the coverslips on top of the mounting medium in a way that the cells are in contact with the mounting medium and the microscope slide, allow slides to dry (for minimum 6 h at RT) before microscopic analysis

■ **PAUSE POINT** Fixed cells mounted on glass slides can be stores up to several months

(vi) Microscopic analysis of localization or anchor and effector unit before and after dimerizer treatment.

### A2. Analysis of dimerization / cleavage in cell lysates ● **TIMING** ~ 24 h

(i) Wash cells with ice cold 1x PBS

(ii) Lyse cells with 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  $\mu$ M Leupeptin, 18  $\mu$ M Pepstatin, 5  $\mu$ M Aprotinin, 1 mM PMSF, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 5 mM EDTA]

(iii) Clear cell lysates by centrifugation at 13,000 rpm for 15 min at 4°C

(iv) Denature proteins with 5x sample buffer [312.5 mM Tris-HCl (pH 6.8), 10% SDS, 25%  $\beta$ -mercaptoethanol, 50% glycerol, bromphenol blue] and cooking for 6 min

(v) Separate proteins by SDS-PAGE and transfer them on Immobilon PVDF membranes (Millipore)

(vi) Use appropriate primary antibodies and secondary antibodies labeled with horseradish peroxidase (HRP-conjugated to visualize proteins using enhanced chemiluminescence (Millipore) and a CCD camera (Fusion Fx7, Vilber)

(vii) For quantification: measure intensities of monomer and dimer bands using ImageJ/Fiji and calculate the dimerization efficiency with following formula, then compare dimerization efficiency before and after dimerizer treatment / cleavage of MeNV-HaXS-induced dimers.

$$\% \text{ dimerized} = \frac{\frac{\text{dimer}}{2}}{\text{monomer}^1 + \frac{\text{dimer}}{2}} * 100$$

<sup>1</sup> = less expressed monomer

### A3. Analysis of dimerization / cleavage in live cells ● **TIMING** ~ 2-3 h

(i) Assemble and analyze movies with (ImageJA, 1.44b)

(ii) Calculate fluorescence of effector unit at selected ROIs (depending on application you plan to analyze):

$$F_{\text{effector}} = F_{\text{effector, ROI1}}$$

$$F_{\text{anchor, whole cell, n=1}} - F_{\text{effector, ROI2}} / F_{\text{anchor, whole cell}} \text{ for every frame}$$

(iii) Plot increase of effector fluorescence over time to analyze change of fluorescence at the selected ROIs.



## Protocol for TEST EXPERIMENTS (1 to 4)

**TEST EXPERIMENT 1. Verify localization and translocation of effector and anchor unit**

▲ **CRITICAL STEP** Test newly designed constructs in combination with a functional, characterized construct and optimized until they perform as expected, before two new designed constructs are used together.

▲ **CRITICAL STEP** If a new anchor unit was designed perform double labeling experiment with confirmed markers for the targeted organelle / subcellular compartment to confirm correct localization of new anchor unit.

- (i) *Day 1.* Seed cells as described in BASIC PROTOCOL STEP 2B
- (ii) *Day 2.* Transient transfection as described in BASIC PROTOCOL STEP 2B
- (iii) *Day 3.* HaXS8 treatment as described in BASIC PROTOCOL STEP 3
- (iv) **Analysis A1**

→ If anchor unit localizes properly and if translocation of effector unit towards anchor unit is efficient: go on with STEP 2, otherwise → **TROUBLESHOOTING (1)**

**TEST EXPERIMENT 2. Verify and optimize expression levels and integration of monomers**

- (i) *Day 1.* Seed cells as described in BASIC PROTOCOL STEP 1
- (ii) *Day 2.* Transient transfection as described in BASIC PROTOCOL STEP 2B, use ratio of amount of DNA as desired for final experiment
- (iii) *Day 3.* HaXS8 treatment as described in BASIC PROTOCOL STEP 3
- (iv) Perform **Analysis A1 or A3**

→ If expression levels are as desired, if dimerization is efficient and if integration of monomers is efficient, go on with STEP 3, otherwise → **TROUBLESHOOTING (2)**

**TEST EXPERIMENT 3. Analyze and optimize MeNV-HaXS treatment conditions**

- (i) *Day 1.* Seed cells as described in BASIC PROTOCOL STEP 2A
- (ii) *Day 2.* Transient transfection as described in BASIC PROTOCOL STEP 2A, use ratio of DNA suited to optimized expression levels
- (iii) *Day 3.* HaXS8 treatment as described in BASIC PROTOCOL STEP 3: perform time course of HaXS8 incubation: 5 min, 15 min, 30 min, 60 min
- (iv) perform **Analysis A2.** Determine shortest HaXS incubation time at which dimer formation is efficient

→ If shortest HaXS incubation time at which dimerization is efficient suits to the requirements of final experiment, go on with Step 4, otherwise → **TROUBLESHOOTING (3)**

**TEST EXPERIMENT 4. Analyze and optimize MeNV-HaXS cleavage conditions****(A) Optimize cleavage conditions in cell lysates**

- (i) Prepare cells until STEP 3A from BASIC PROTOCOL, induce photocleavage with UV lamp, test time course of different illumination times (5 min, 10 min, 20 min)
- (ii) Perform **Analysis A2** Determine shortest illumination time, at which MeNV-HaXS-induced dimers are efficiently cleaved, while HaXS8-induced dimers are not cleaved

→ If shortest illumination time at which dimers are efficiently cleaved suits to the requirements of final experiment, perform Final experiment, otherwise → **TROUBLESHOOTING (3)**

**(B) Optimize cleavage conditions in live cells**

- (i) Prepare cells until STEP 3B from BASIC PROTOCOL, induce photocleavage with UV lamp
- (ii) Test time course of different illumination times with laser (0.5 sec, 1 sec, 2 sec) resp. light from mercury metal halide and standard DAPI excitation filters (ex: 377±25 nm, em: 447±30 nm) (5 sec, 10 sec, 20 sec) at medium laser strength or fluorescence lamp strength
- (iii) Perform **Analysis A3** Determine shortest illumination time, at which MeNV-HaXS-induced dimers are efficiently cleaved, while HaXS8-induced dimers are not cleaved

→ If MeNV-HaXS-induced dimers are efficiently cleaved, whereas HaXS8-induced dimers are not cleaved and if cleavage conditions are compatible with final experiment requirements, go on with Final experiment, otherwise → **TROUBLESHOOTING (4)**

## Troubleshooting

Troubleshooting advice can be found in Table 3.

### TROUBLESHOOTING TABLE

STEP	PROBLEM	CAUSE	SOLUTION
1	Incorrect localization of anchor unit	Suboptimal construct design of anchor unit	<ul style="list-style-type: none"> <li>-check context of anchoring motif (has to be retained as good as possible as in original plasmid containing the functional anchor motif)</li> <li>-change order of domains</li> <li>-modulate linker between anchoring motif and next domain (fluorophore or Halo- or SNAP-Tag) (see Box 3)</li> <li>-exchange of Halo- and SNAP-tag domain</li> <li>-if a GFP or any structurally related fluorophore is integrated in the middle of the construct, consider sterical constraints of domains fused simultaneously to the N- and C-terminus of GFP due to the very close localization of the N- and C-terminus of GFP in 3D space → include long linkers to provide enough flexibility (see Box 1)</li> </ul>
	Inefficient translocation of effector unit towards anchor unit	Suboptimal onstruct design of effector unit and anchor unit	<ul style="list-style-type: none"> <li>-change order of domains</li> <li>-modulate linker between all domains (fluorophore or Halo- or SNAP-Tag) (see Box 3)</li> <li>-exchange of Halo- and SNAP-tag domain in anchor and effector unit</li> <li>-if a GFP or any structurally related fluorophore is integrated in the middle of the construct, consider sterical constraints of domains fused simultaneously to the N- and C-terminus of GFP due to the very close localization of the N- and C-terminus of GFP in 3D space → include long linkers to provide enough flexibility (see Box 1)</li> </ul>
	Inefficient translocation of effector unit towards anchor unit	Effector unit has inflexible localization before HaXS treatment	-analysis of POI integrated in effector unit (if it has a strong intracellular membrane localization modulate its targeting signal in order to get an effector unit that is translocatable upon HaXS treatment)
2	Ratio of expression levels of monomers before HaXS treatment not as expected	Expression of a suboptimal ratio of anchor and effector unit	<ul style="list-style-type: none"> <li>-vary the amount of DNA used for the transient transfection</li> <li>-use of bi-cistronic backbones based on internal ribosome entry site (IRES) <sup>[51]</sup> in which the translation initiation efficiency of the two ORFs differs significantly: initiation region of the upstream gene is very efficient, whereas IRES sequence results in a very inefficient ribosome binding and translation initiation for the second ORF → if the anchor unit is put upstream of the IRES and the effector unit downstream of the IRES, a high ratio of anchor unit to effector unit expression level is expected</li> <li>-use strong resp. weak promoters, which lead to a lower or higher expression level (weak promoter: human ubiquitin C gene (UBC) promoter, weakest promoter in various cell types (e.g. around 50% weaker than CMV and EF-1a) <sup>[52]</sup>)</li> <li>- modify cis regulatroy elements (CATA, TATA box) <sup>[53]</sup>, which affect the frequency of transcription initiation and consequently the level of gene expression</li> <li>-modify Kozak sequence flanking the AUG initiator codon, <sup>[54]</sup></li> </ul>
	Inefficient integration of monomers in dimer complex	Ratio of expressed effector unit to anchor unit is too high	-express more anchor unit and less effector unit (vary amount of DNA or follow advice in previous box)
3	Inefficient dimerization under conditions suited to final experiment (fixed incubation time)	Too low concentration of dimerizer to induce dimerization	<ul style="list-style-type: none"> <li>-use higher dimerizer concentrations to induce faster formation of dimers, consider that too high concentration result in inefficient dimer formation → compromise between the need for fast versus efficient dimerization has to be found</li> <li>-if the adaption of HaXS concentration and incubation time suited to final experiments does not result in sufficient dimer formation or integration of monomers in the dimer complex, optimization of the construct design (Troubleshooting STEP 1)</li> <li>-if only the speed but not efficiency of dimer formation is important use high dimerizer concentrations, whereas lower concentrations are recommended if efficiency of dimer formation is important, whereas time for dimer formation is less important</li> </ul>
	Inefficient dimerization under conditions suited to final experiment	Too little intracellular compound concentration (due to limited cell penetration capacity or too high activity of multidrug-resistance pumps (MDR))	<ul style="list-style-type: none"> <li>-if possible use another cell line, which display high cell permeability of the HaXS dimerizers (as e.g. HEK293, HeLa, NIH3T3, MDCK, ...)</li> <li>-investigate effects of multidrug-resistance (MDR) pump inhibitors, since it is known that some cells pump out compounds and thus dramatically reduce intracellular dimerizer concentration</li> </ul>
3	Inefficient dimerization under conditions suited to final experiment	Suboptimal dimerization efficiency	<ul style="list-style-type: none"> <li>-use higher dimerizer concentrations to induce faster formation of dimers, consider that too high concentration result in inefficient dimer formation (Fig 2) → compromise between the need for fast versus efficient dimerization has to be found</li> <li>-if the adaption of HaXS concentration and incubation time suited to</li> </ul>

			final experiments does not result in sufficient dimer formation or integration of monomers in the dimer complex, optimization of the construct design (Troubleshooting STEP 1) -if only the speed but not efficiency of dimer formation is important use high dimerizer concentrations, whereas lower concentrations are recommended if efficiency of dimer formation is important, whereas time for dimer formation is less important
4	Inefficient cleavage of MeNV-HaXS-induced dimers	Too little energy used to induce cleavage	-increase intensity of light used to induce photo-cleavage (laser strength, intensity of fluorescent lamp) -increase illumination time
	Loss of fluorescence in control cells (HaXS8-induced dimerization)	Too much energy used to induce cleavage	-decrease intensity of light used to induce photo-cleavage (laser strength, intensity of fluorescent lamp) -decrease illumination time -
	Cells dies	Too much energy used to induce cleavage	-decrease intensity of light used to induce photocleavage (laser strength, intensity of fluorescent lamp) -decrease illumination time -make less images -prepare new imaging medium (increase amount of FCS from 2% up to 10%, if compatible with amount of autofluorescence)
Final experiment	Not expected output	Suboptimal construct design,	if HaXS-induced dimerization under these optimized conditions (construct design, expression levels of anchor and effector unit and HaXS-incubation) cannot induce expected cellular effect, go back to TROUBLESHOOTING STEP1 (balance between shorter and less flexible linkers, which allow dimerization between the two units and possible functional interaction of the POI with targeted membrane or another POI has to be found)

## Timing

### Perform experiment in cell lysates: in total 4 days (without construct design)

STEP 1 Create SNAP- and HaloTag expression constructs: up to several weeks

STEP 2 Create cells co-expressing SNAP- and HaloTag expression constructs: 24 h

STEP 3 HaXS treatment: 60 min

STEP 4 Cleavage of dimers: 30 min

Analysis of dimerization and cleavage efficiency by immune blotting: 1 h

### Perform experiment in live cells: in total 4 days (without construct design)

STEP 1 Create SNAP- and HaloTag expression constructs: up to several weeks

STEP 2 Create cells co-expressing SNAP- and HaloTag expression constructs: 24 h

STEP 3 HaXS treatment: 2-3 h

STEP 4 Cleavage of dimers: 2-3 h

Analysis of dimerization and cleavage efficiency by live cell microscopy and image analysis: 5-6 h

## Anticipated results

By using this protocol, the localization of proteins can be manipulated with high spatiotemporal precision (see Fig 3), a branch of a signal network can be activate or inactivated (see Fig 4) or MeNV-HaXS anchored POIs can be released instantaneously upon UV illumination of the dimer complex, which can be exploited in an experimental setup that allow to study intracellular transport kinetics of released proteins (see Fig 5).

## References

- [1] C. H. Heldin, A. Ostman, A. Eriksson, A. Siegbahn, L. Claesson-Welsh, B. Westermark, *Kidney Int* **1992**, *41*, 571.
- [2] M. Putyrski, C. Schultz, *Chem Biol* **2011**, *18*, 1126.
- [3] K. W. Freeman, R. D. Gangula, B. E. Welm, M. Ozen, B. A. Foster, J. M. Rosen, M. Ittmann, N. M. Greenberg, D. M. Spencer, *Cancer Res* **2003**, *63*, 6237.
- [4] M. N. Pruschy, D. M. Spencer, T. M. Kapoor, H. Miyake, G. R. Crabtree, S. L. Schreiber, *Chem Biol* **1994**, *1*, 163; D. M. Spencer, T. J. Wandless, S. L. Schreiber, G. R. Crabtree, *Science* **1993**, *262*, 1019.
- [5] B. C. Suh, T. Inoue, T. Meyer, B. Hille, *Science* **2006**, *314*, 1454.
- [6] T. Ueno, B. H. Falkenburger, C. Pohlmeier, T. Inoue, *Sci Signal* **2011**, *4*, ra87.
- [7] S. Terrillon, M. Bouvier, *EMBO J* **2004**, *23*, 3950.
- [8] T. W. Corson, N. Aberle, C. M. Crews, *ACS Chem Biol* **2008**, *3*, 677.
- [9] R. DeRose, T. Miyamoto, T. Inoue, *Pflugers Arch* **2013**, *465*, 409.
- [10] D. Erhart, M. Zimmermann, O. Jacques, M. B. Wittwer, B. Ernst, E. Constable, M. Zvelebil, F. Beaufils, M. P. Wymann, *Chem Biol* **2013**, *20*, 549.
- [11] A. Gautier, E. Nakata, G. Lukinavicius, K. T. Tan, K. Johnsson, *J Am Chem Soc* **2009**, *131*, 17954.
- [12] G. V. Los *et al.*, *ACS Chem Biol* **2008**, *3*, 373.
- [13] M. Zimmermann, R. Cal, E. Janett, V. Hoffmann, C. G. Bochet, E. Constable, F. Beaufils, M. P. Wymann, *Angew Chem Int Ed Engl* **2014**, *53*, 4717.
- [14] X. Sun *et al.*, *Chembiochem* **2011**, *12*, 2217.
- [15] L. P. Encell *et al.*, *Curr Chem Genomics* **2012**, *6*, 55.
- [16] M. J. Hinner, K. Johnsson, *Curr Opin Biotechnol* **2010**, *21*, 766.
- [17] T. Komatsu, I. Kukelyansky, J. M. McCaffery, T. Ueno, L. C. Varela, T. Inoue, *Nat Methods* **2010**, *7*, 206.
- [18] O. Rocks, A. Peyker, P. I. Bastiaens, *Curr Opin Cell Biol* **2006**, *18*, 351.
- [19] G. Lemerrier, S. Gendreizig, M. Kindermann, K. Johnsson, *Angew Chem Int Ed Engl* **2007**, *46*, 4281.
- [20] A. Rutkowska, C. H. Haering, C. Schultz, *Angew Chem Int Ed Engl* **2011**, *50*, 12655.
- [21] F. S. Liang, W. Q. Ho, G. R. Crabtree, *Sci Signal* **2011**, *4*, rs2.
- [22] T. Miyamoto, R. DeRose, A. Suarez, T. Ueno, M. Chen, T. P. Sun, M. J. Wolfgang, C. Mukherjee, D. J. Meyers, T. Inoue, *Nat Chem Biol* **2012**, *8*, 465.
- [23] Y. C. Lin, Y. Nihongaki, T. Y. Liu, S. Razavi, M. Sato, T. Inoue, *Angew Chem Int Ed Engl* **2013**, *52*, 6450.
- [24] S. Feng, V. Laketa, F. Stein, A. Rutkowska, A. MacNamara, S. Depner, U. Klingmuller, J. Saez-Rodriguez, C. Schultz, *Angew Chem Int Ed Engl* **2014**, *53*, 6720.
- [25] P. Liu *et al.*, *Angew Chem Int Ed Engl* **2014**, *53*, 10049.
- [26] K. E. O'Reilly *et al.*, *Cancer Res* **2006**, *66*, 1500.
- [27] J. I. Luengo *et al.*, *Chem Biol* **1995**, *2*, 471.
- [28] M. S. Robinson, J. Hirst, *Curr Protoc Cell Biol* **2013**, *61*, 15.20.1.
- [29] L. A. Banaszynski, C. W. Liu, T. J. Wandless, *J Am Chem Soc* **2005**, *127*, 4715.
- [30] S. Deswal, A. K. Schulze, T. Hofer, W. W. Schamel, *PLoS One* **2011**, *6*, e22928.
- [31] H. Haruki, J. Nishikawa, U. K. Laemmli, *Mol Cell* **2008**, *31*, 925.
- [32] S. C. Phua, C. Pohlmeier, T. Inoue, *ACS Chem Biol* **2012**, *7*, 1950.
- [33] C. Dingwall, J. Robbins, S. M. Dilworth, B. Roberts, W. D. Richardson, *J Cell Biol* **1988**, *107*, 841.
- [34] A. D. Linstedt, H. P. Hauri, *Mol Biol Cell* **1993**, *4*, 679.
- [35] Y. Sancak, L. Bar-Peled, R. Zoncu, A. L. Markhard, S. Nada, D. M. Sabatini, *Cell* **2010**, *141*, 290.
- [36] J. Riedl *et al.*, *Nat Methods* **2008**, *5*, 605.
- [37] D. Kalderon, B. L. Roberts, W. D. Richardson, A. E. Smith, *Cell* **1984**, *39*, 499.
- [38] A. Busch, T. Kiel, S. Hubner, *Traffic* **2009**, *10*, 1221.
- [39] T. la Cour, L. Kiemer, A. Molgaard, R. Gupta, K. Skriver, S. Brunak, *Protein Eng Des Sel* **2004**, *17*, 527.
- [40] M. M. Kessels, B. Qualmann, *EMBO J* **2002**, *21*, 6083.
- [41] H. L. Zhao, X. Q. Yao, C. Xue, Y. Wang, X. H. Xiong, Z. M. Liu, *Protein Expr Purif* **2008**, *61*, 73.
- [42] N. Amet, H. F. Lee, W. C. Shen, *Pharm Res* **2009**, *26*, 523; Y. Bai, W. C. Shen, *Pharm Res* **2006**, *23*, 2116.
- [43] N. Amet, H. F. Lee, W. C. Shen, *Pharm Res* **2009**, *26*, 523.
- [44] J. S. Huston *et al.*, *Proc Natl Acad Sci U S A* **1988**, *85*, 5879; Y. Bai, W. C. Shen, *Pharm Res* **2006**, *23*, 2116.
- [45] M. Sabourin, C. T. Tuzon, T. S. Fisher, V. A. Zakian, *Yeast* **2007**, *24*, 39.
- [46] M. D. Ryan, A. M. King, G. P. Thomas, *J Gen Virol* **1991**, *72*, 2727; A. L. Szymczak, C. J. Workman, Y. Wang, K. M. Vignali, S. Dilioglou, E. F. Vanin, D. A. Vignali, *Nat Biotechnol* **2004**,

- 22, 589.
- [47] M. J. Kennedy, R. M. Hughes, L. A. Peteya, J. W. Schwartz, M. D. Ehlers, C. L. Tucker, *Nat Methods* **2010**, *7*, 973.
- [48] D. Strickland, Y. Lin, E. Wagner, C. M. Hope, J. Zayner, C. Antoniou, T. R. Sosnick, E. L. Weiss, M. Glotzer, *Nat Methods* **2012**, *9*, 379.
- [49] M. Yazawa, A. M. Sadaghiani, B. Hsueh, R. E. Dolmetsch, *Nat Biotechnol* **2009**, *27*, 941.
- [50] Y. I. Wu, X. Wang, L. He, D. Montell, K. M. Hahn, *Methods Enzymol* **2011**, *497*, 393.
- [51] J. Pelletier, N. Sonenberg, *Nature* **1988**, *334*, 320.
- [52] J. Y. Qin, L. Zhang, K. L. Clift, I. Hular, A. P. Xiang, B. Z. Ren, B. T. Lahn, *PLoS One* **2010**, *5*, e10611.
- [53] J. P. Ferreira, R. W. Peacock, I. E. Lawhorn, C. L. Wang, *Syst Synth Biol* **2011**, *5*, 131.
- [54] M. Kozak, *Nature* **1984**, *308*, 241.

## Author information

### Contributions

M. P. W. and F. B. supervised the project. F. B. and R. C. synthesized and chemically characterized the dimerizers. M.Z. developed the workflow of the protocol, validated the dimerizers in cells and performed experiments. M. ZV. performed modeling. M. Z. performed modeling analysis (CellDesigner). M. Z. and M. P. W wrote manuscript. F. B., R. C, M. Z and M. P. W discussed results during development of project.

### Corresponding author

Matthias P. Wymann, Institute of Biochemistry & Genetics, Department of Biomedicine, University of Basel, Switzerland  
matthias.wymann@unibas.ch

Florent Beaufils, Institute of Biochemistry & Genetics, Department of Biomedicine, University of Basel, Switzerland  
florent.beaufils@unibas.ch

## Supplementary Information

### Generation of plasmids encoding anchor and effector unit

Generation of anchor units Halo-RFP-Giantin, Halo-RFP-Rheb15, Mito-GFP-Halo, Halo-RFP-LAMP1, LifeAct-mTFP1-SNAP, SNAP-GFP-CAAX and NLS-CFP-SNAP is described in<sup>[13]</sup>. Generation of HT7-GFP and SNAP26m-GFP is described in<sup>[10]</sup>, generation of HT7(L273Y)-GFP and SNAPf-GFP is described in<sup>[13]</sup>.

For NES-DsRed-Halo: FRB from NES-DsRed-FRB expression vector (kind gift from S. Hübner, Würzburg) was exchanged by HT7(L273Y) sequence. Maps and expression vector sequences can be obtained from the authors upon request.

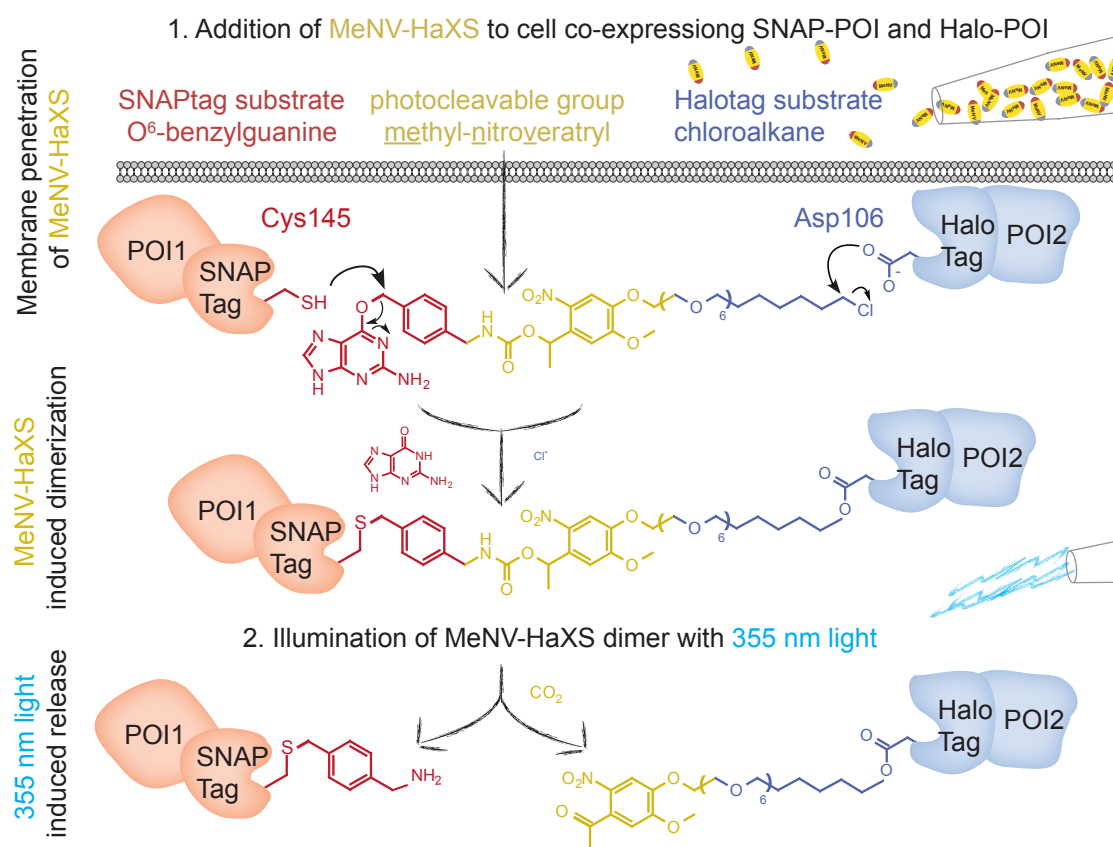
### Statistical Analysis

Statistical analysis was performed with GraphPad Prism v6. For Student's t test (two sided, non-paired with Welch correction,  $p < 0.05$ )  $\geq$  two independent experiments were compared.

### Molecular modeling

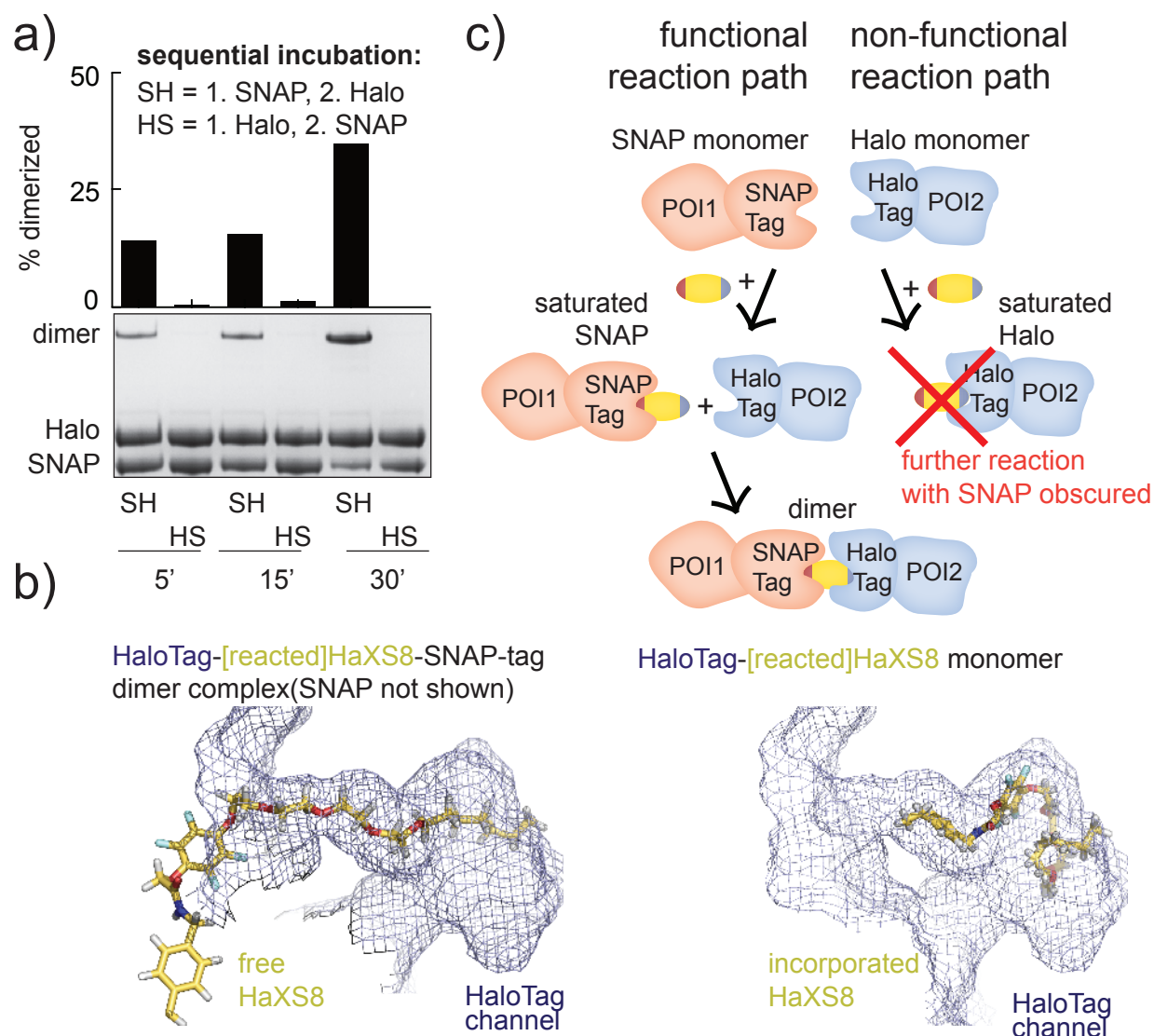
Starting values for the modeling are the crystal structures of the Halotag (Newman et al., 1999) and the SNAP-tag (Mollwitz et al., 2012). A first manual docking of HaXS8 dimerizer was refined using the genetic algorithm application GOLD (Jones et al., 1997), before the linker HaXS dimerizer and protein structures were subjected to energy minimization using the program Yasara (Krieger et al., 2004). Modeling was performed by Marketa Zvelebil, London. Pymol software package was used to substitute amino acid leucine with tyrosine (L273Y) and to create illustration of the models (for Fig 1 and 2a). VMD was used to create illustration of the models software (Supplementary Figure S1b).

## Scheme 1\_Principle of MeNV-HaXS CID



**Scheme 1. Principle of MeNV-HaXS induced dimerization of SNAP- and HaloTag fusion proteins followed by subsequent UV-induced reversion of dimer complex.** A photocleavable, cell-permeable SNAP- and HaloTag reactive chemical inducer of dimerization (CID) with a methyl-6-nitroveratryl (MeNV) core (MeNV-HaXS) enters the cell. The O<sup>6</sup>-benzylguanine substrate (red) reacts with Cys145 in the SNAP-tag, while the chloroalkane (blue) reacts with Asp106 in the HaloTag, which results in dimerization of HaloTag- and SNAP-tag-fused proteins of interest (POI). Illumination of MeNV-HaXS (360 nm;  $\epsilon = 4058 \text{ M}^{-1}\text{cm}^{-1}$ ; quantum yield = 0.075) cleaves the link between the POIs, and releases them from the covalent complex.

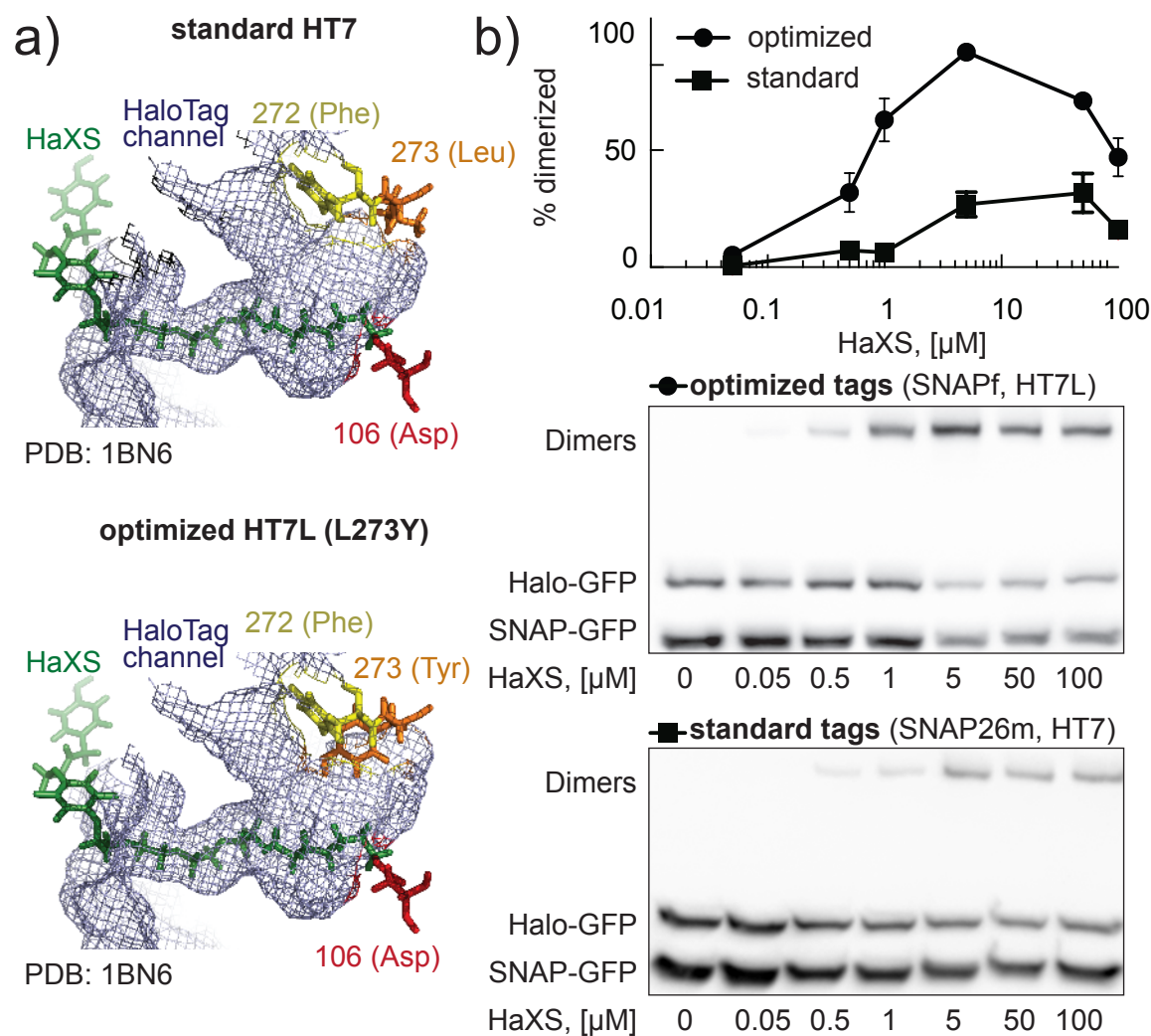
## Figure 1\_Reaction path of HaXS dimerizer



**Figure 1. Reaction path of HaXS dimerizer.** a) 5  $\mu$ M HaXS was sequentially incubated with recombinant SNAP-GFP and Halo-GFP fusion proteins (5  $\mu$ M each) at 37°C (SH: HaXS is first incubated with SNAP-GFP, then with Halo-GFP, for time indicated, HS: vice versa) before proteins were denatured and reduced (in sample buffer containing SDS and  $\beta$ -mercaptoethanol), separated by SDS-PAGE and detected by Coomassie-Blue staining. b) Left: Modeled structure of a linked HaloTag-[reacted]HaXS8-SNAP-tag dimer complex. [reacted]HaXS8 molecule was linked to Asp106 in the HaloTag (blue) and to Cys145 in the SNAP-tag (SNAP-tag not shown) crystal structure, before randomized structural starting points were put through energy minimization (see Supporting Information). Right: Modeled structure of HaloTag-[reacted]HaXS8 monomer, same procedure but [reacted]HaXS8 molecule was linked to HaloTag only (blue) before energy minimization (modeling performed by Marketa Zvelebil, London). c) Scheme showing the two possible reaction paths of the HaXS dimerizer. Only the reaction path in which HaXS first reacts with the SNAP-tag leads to efficient dimer formation (functional reaction path).

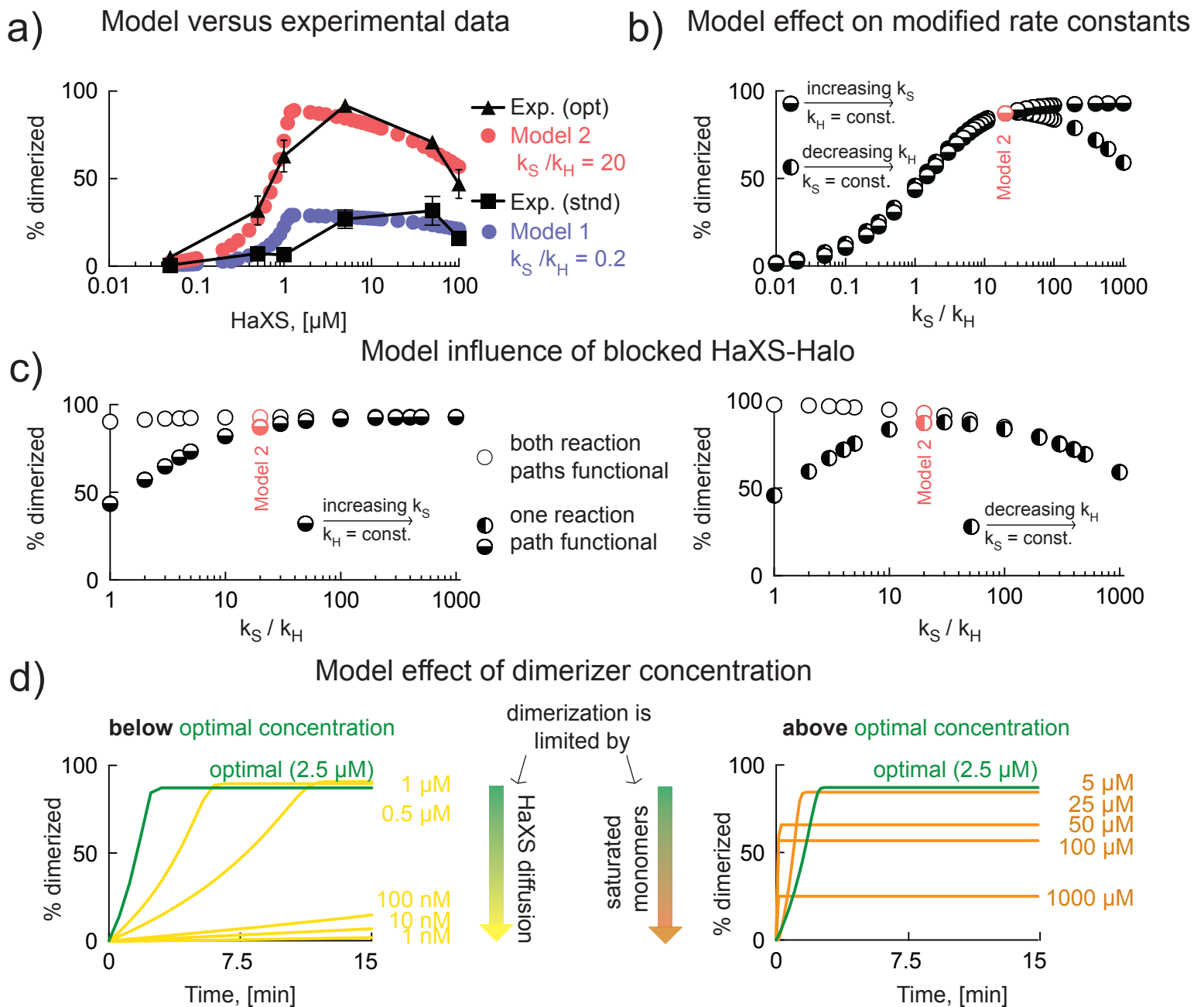


## Figure 2\_Improve Dimerization Efficiency of HaXS CID



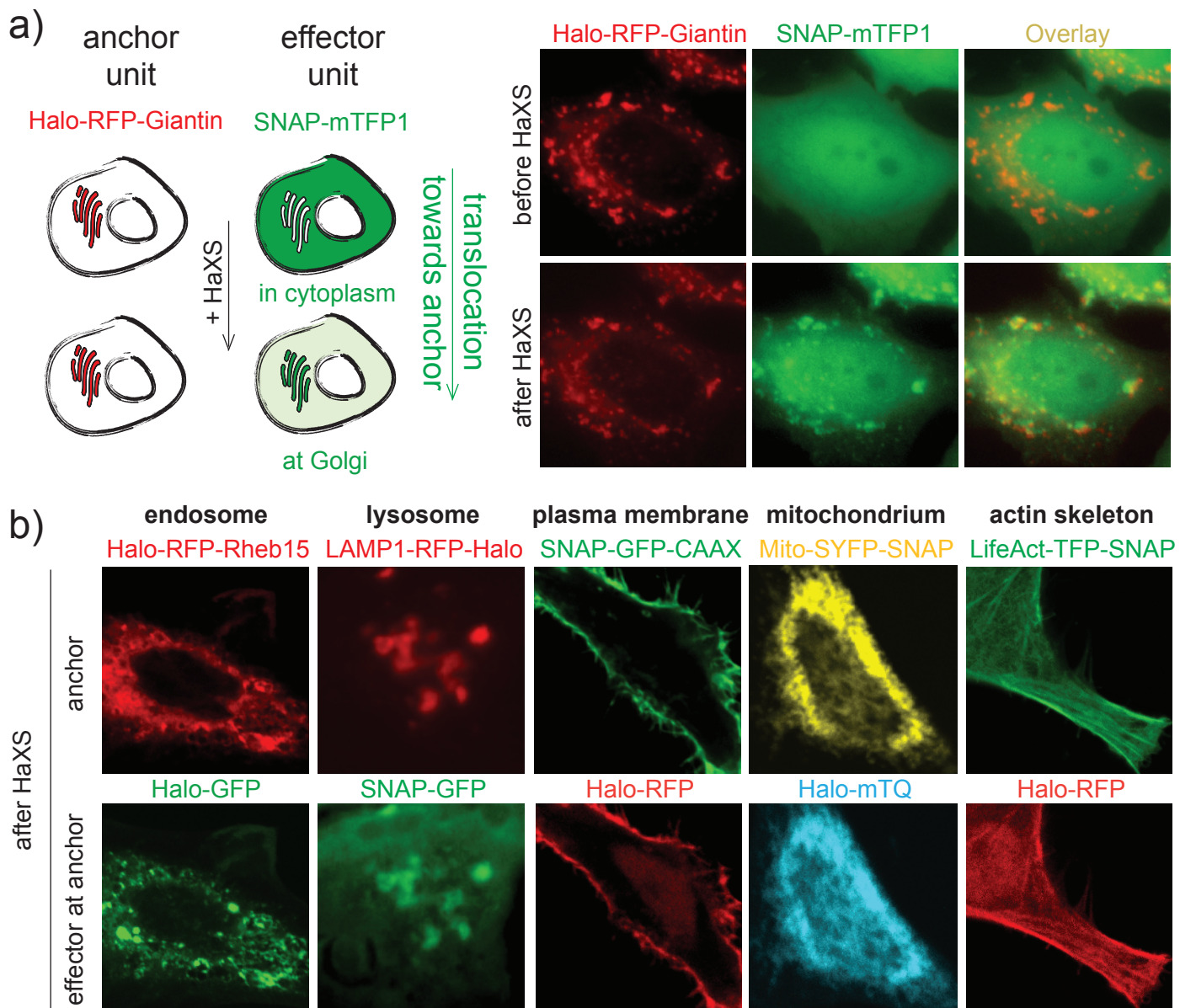
**Figure 2. Adaption of rate constants of SNAP- and Halo-Tag to improve dimerization efficiency.** a) Illustration of crystallographic structure of reaction cavity of standard (HT7) and optimized (HT7L) HaloTag, (Dha, PDB 1BN6). In HT7L leucine at position 273 was substituted by a tyrosine (L273Y, shown orange). The amino acids 272 and 106 involved in the reaction with the HaloTag substrate are shown yellow and red. b) HeLa cells co-expressing the optimized (SNAPf/HT7L273Y) or standard (SNAP26m/HT7) tag combination of SNAP-GFP and Halo-GFP were exposed to an increasing amount of HaXS (from 50 nM to 100 μM, as indicated) in complete cell culture medium at 37°C. After 5 min cells were lysed and proteins were subjected to SDS-PAGE and immune-blotting. Tagged proteins were detected using anti-GFP (primary) and horseradish peroxidase labeled (secondary) antibodies, and chemiluminescence. Quantifications represent mean of ± SEM of at least three independent experiments.

### Figure 3\_Model dimerization reactions of HaXS CID



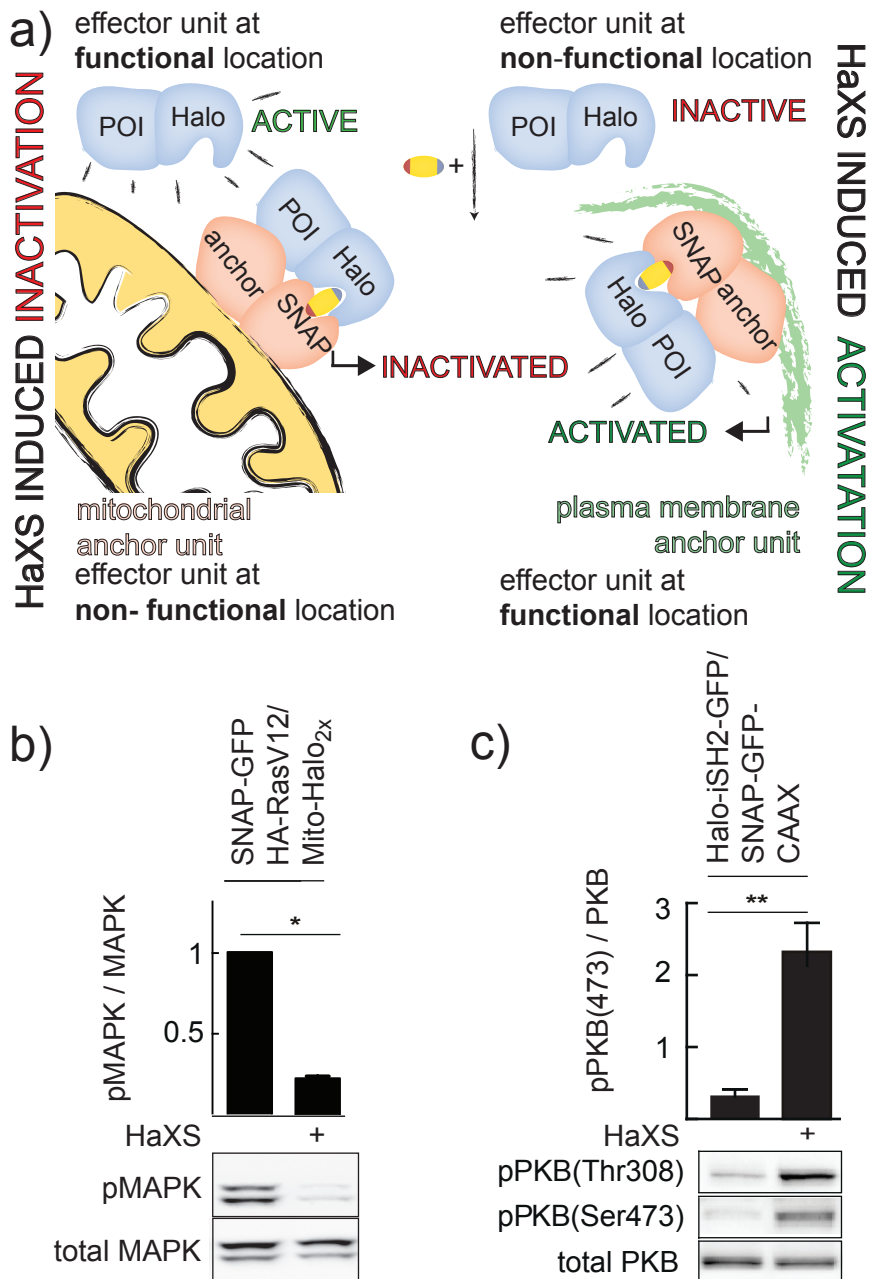
**Figure 3. Model dimerization reactions of HaXS CID.** a) Comparison of data obtained from model and intracellular dimerization experiment from Fig 2b). b) Effect of increasing ratios of rate constants  $k_S$  to  $k_H$  is simulated. Graph showing correlation of dimerization efficiencies (5 min, 2.5  $\mu\text{M}$  HaXS) with increasing ratios of rate constant  $k_S$  to  $k_H$ . Rate constant of Halo is kept constant ( $k_{\text{Halo}} = 900$ ), while rate constant of SNAP was increased (vertical half filled circles) or rate constant of SNAP was kept constant ( $k_{\text{SNAP}} = 18000$ ), while rate constant of Halo was decreased (horizontal half filled circles). Dimerization efficiency at which ratio of rate constants  $k_S$  to  $k_H$  is 20 as defined in Model 2 is labelled in pink. c) Effect of blocked HaXS-Halo on dimerization efficiencies at different ratios  $k_S$  to  $k_H$  is simulated. Graph showing correlation of dimerization efficiencies (5 min, 2.5  $\mu\text{M}$  HaXS) with increasing ratios of rate constant  $k_S$  to  $k_H$ . Rate constant of Halo is kept constant ( $k_{\text{Halo}} = 900$ ), while rate constant of SNAP was increased (left) with one (horizontal half filled circles) or with two (empty circles) functional reaction paths. Rate constant of SNAP was kept constant ( $k_{\text{SNAP}} = 18000$ ), while rate constant of Halo was decreased (right) with one (vertical half filled circles) or with two (empty circles) functional reaction paths. Dimerization efficiency at which ratio of rate constants  $k_S$  to  $k_H$  is 20 (is labelled in pink) d) Scheme describing influence of HaXS concentration on dimerization efficiency. Dimerization efficiency of concentrations below the optimal concentration (2.5  $\mu\text{M}$  green curve) are limited through diffusion of HaXS into the cell (left, yellow curves). Dimerization efficiency of concentrations above the optimal concentration (2.5  $\mu\text{M}$ , green curve) are limited through the amount of saturated monomers (right, orange curves).

## Figure 4\_Protein targeting to selected intracellular organelles.



**Figure 4. Protein targeting to selected intracellular organelles.** a) Scheme describing HaXS induced translocation of a cytosolic effector unit (SNAPf-mTFP1, green) to the Golgi anchor unit (Halo-RFP-Giantin, red). HeLa cells grown on 12 mm coverslips (Menzel) expressing an organelle anchor unit (first row) and a cytosolic effector unit (second row) were exposed to HaXS dimerizer, before cells were washed twice with PBS, fixed with 4% pformaldehyde (PFA, in PBS), and mounted in Mowiol (Plüss-Stauffer) containing 1% propyl gallate (Sigma-Aldrich). Translocation of cytosolic effector unit toward the Golgi is documented. b) Library of anchor units targeting endosomes, lysosomes, Golgi, plasma membrane, mitochondrion and the actin skeleton (first row). Pictures in second row show effector unit translocated toward the anchor unit upon addition of HaXS.

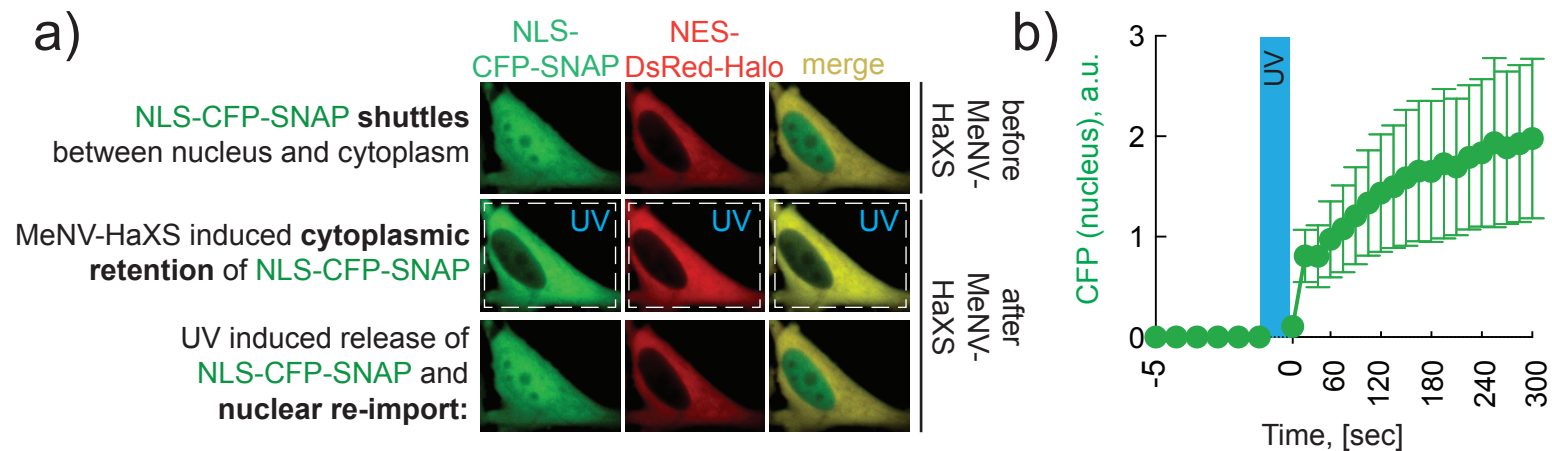
## Figure 5\_Modulate signaling pathways



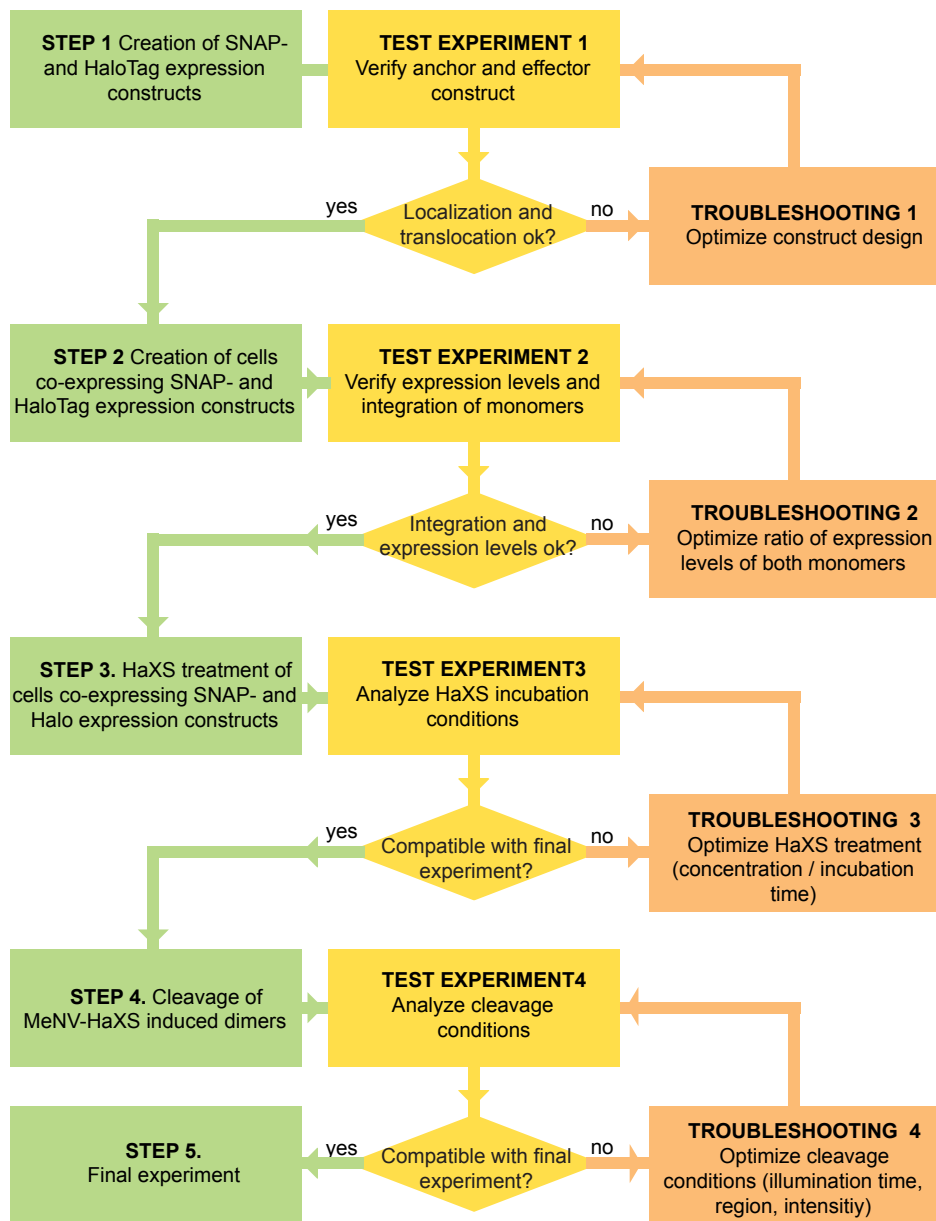
**Figure 5. Control of protein localization as strategy to interfere with the activation of signaling pathways.** a) Scheme describing principle of HaXS-induced inactivation (left) and HaXS-induced activation (right) of a signaling protein through the translocation of the effector unit containing the signaling protein towards or away from its functional localization via dimerization with the anchor unit. b) HEK293 cells co-expressing the effector unit SNAP-GFP-HA-RasV12 and the anchor unit Mito-Halo<sub>2x</sub> were starved overnight (ON) and exposed to 0.5 μM HaXS for 30 min, before cells were lysed and proteins were subjected to SDS-PAGE and immunoblotting. Tagged proteins were detected using the indicated primary antibodies and horseradish peroxidase labeled secondary antibodies, and chemiluminescence. Quantifications of signal intensities represent the mean of ± SEM of two independent experiments. c) HEK293 cells co-expressing a translocatable effector unit GFP-Halo-iSH2, with the inter-Src homology 2 (iSH2) domain from p85, which complexes in cells with endogenous PI 3-kinase p110α and a plasma membrane anchor unit SNAP-GFP-CAAX, were starved ON and exposed to 0.5 μM HaXS in complete medium for 40 min, before cells were lysed and PI3K / PKB pathway was analyzed using the indicated phospho-specific antibodies. Quantifications of signal intensities represent the mean of ± SEM of two independent experiments.



## Figure 6\_Intracellular Transport Kinetics

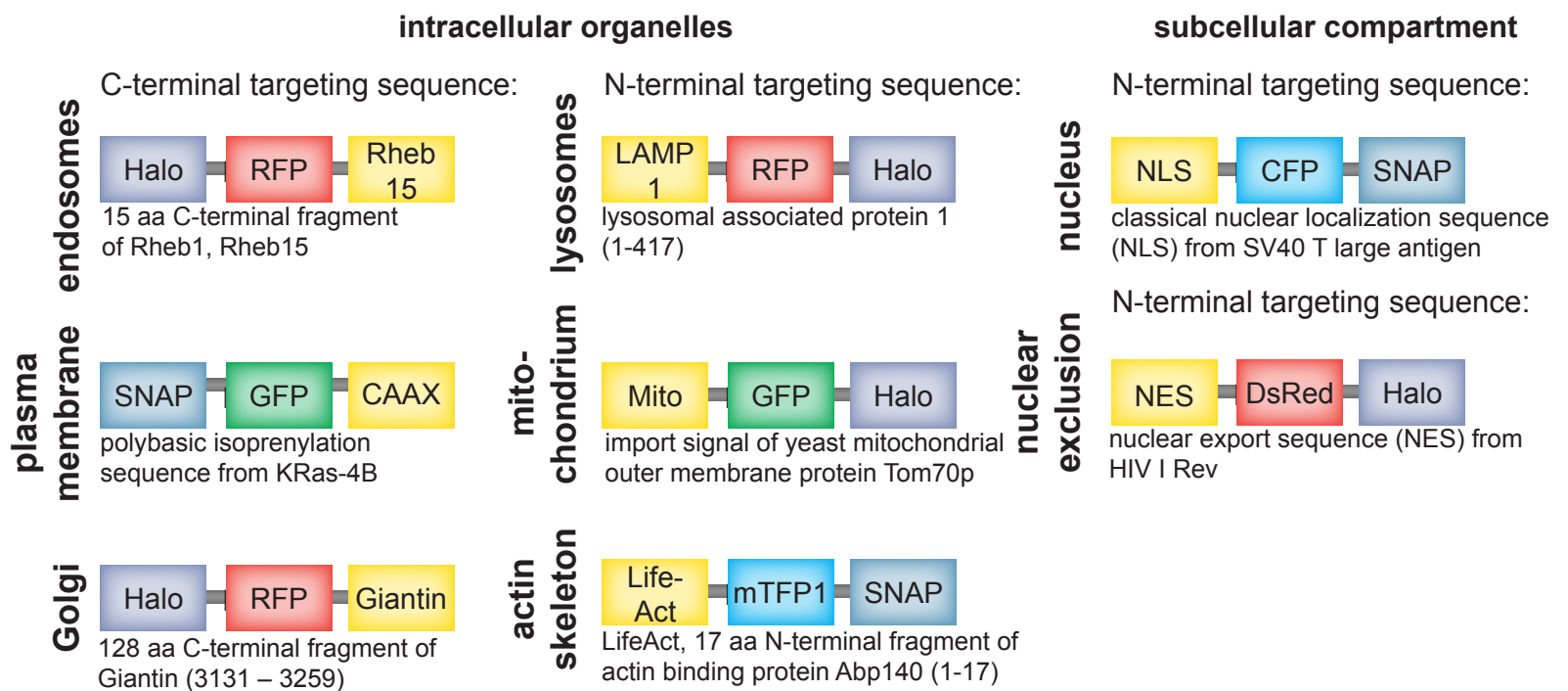


**Figure 6. Control translocation of effector units from one subcellular compartment to another subcellular compartment.** a) HeLa cells co-expressing NES-DsRed-Halo and NLS-CFP-SNAP were exposed to 5  $\mu$ M MeNV-HaXS in cell-culture medium for 15 min at 37°C, which induced the cytoplasmic retention of NLS-CFP-SNAP. Upon illumination of the MeNV-HaXS dimer complex, NLS-CFP-SNAP is released and can translocate back into the nucleus. CFP fluorescence intensity was monitored in the indicated circular regions of interest (ROI) by live-cell microscopy, before and after illumination of the cell (white, dashed rectangle) with a scanning FRAP laser (355 nm). b) Quantification of CFP fluorescence intensity in selected ROIs (labeled circles) monitoring nuclear NLS-CFP-SNAP concentrations (green curve) before and after illumination of the cells are shown; values are means  $\pm$  SEM, n=6 cells, error bars not shown where smaller than symbols used.



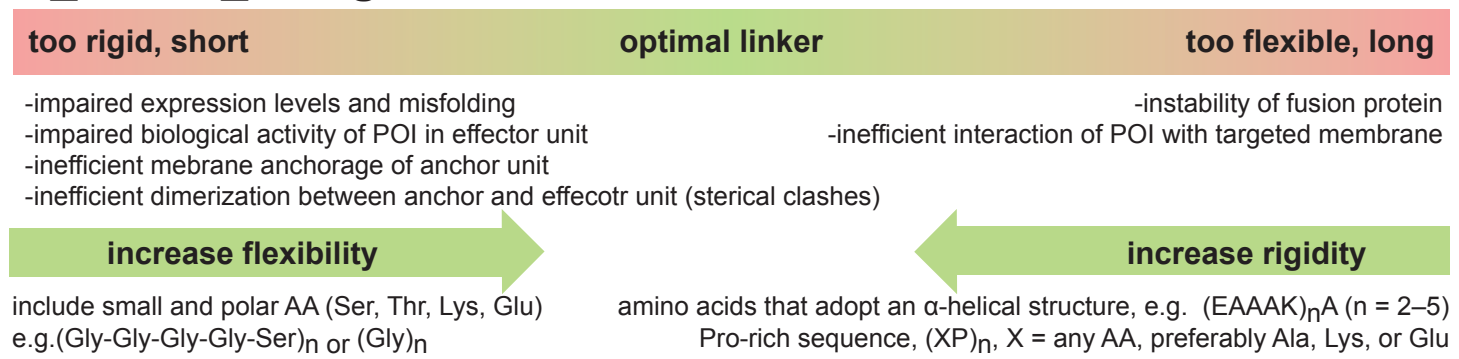
**Flowchart 1. Basic protocol of HaXS CID.** Flowchart describing the strategic planning and optimization steps of the MeNV-HaXS CID.

## Figure 7\_Construct Design of Anchor and Effector unit



**Figure 7. Design of anchor unit and optimal choice of linker.** Scheme describing plasmids encoding anchor units targeting various organelles such as the endosomes, lysosomes, plasma membrane, mitochondrion, actin skeleton and the Golgi as well as subcellular compartments as the nucleus and space outside the nucleus. Targeting sequence of each construct is shown in yellow and described in the text below.

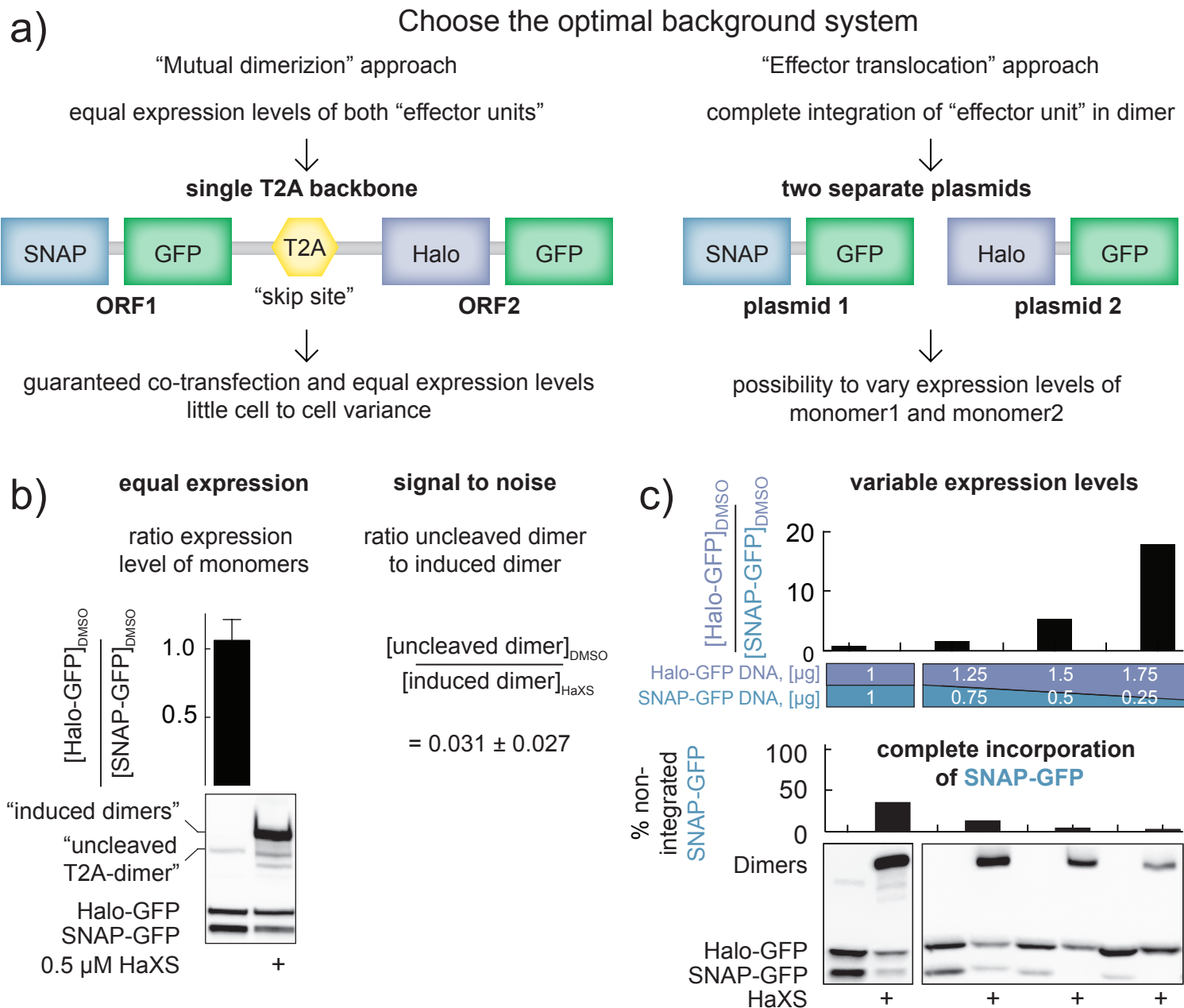
## Box1\_Linkers\_Design



**Box 1. Choice of optimal linker.** Potential effects of too long or flexible and too short or rigid linkers on construct performances are listed. Advice on how to increase flexibility resp. rigidity of linkers is indicated.

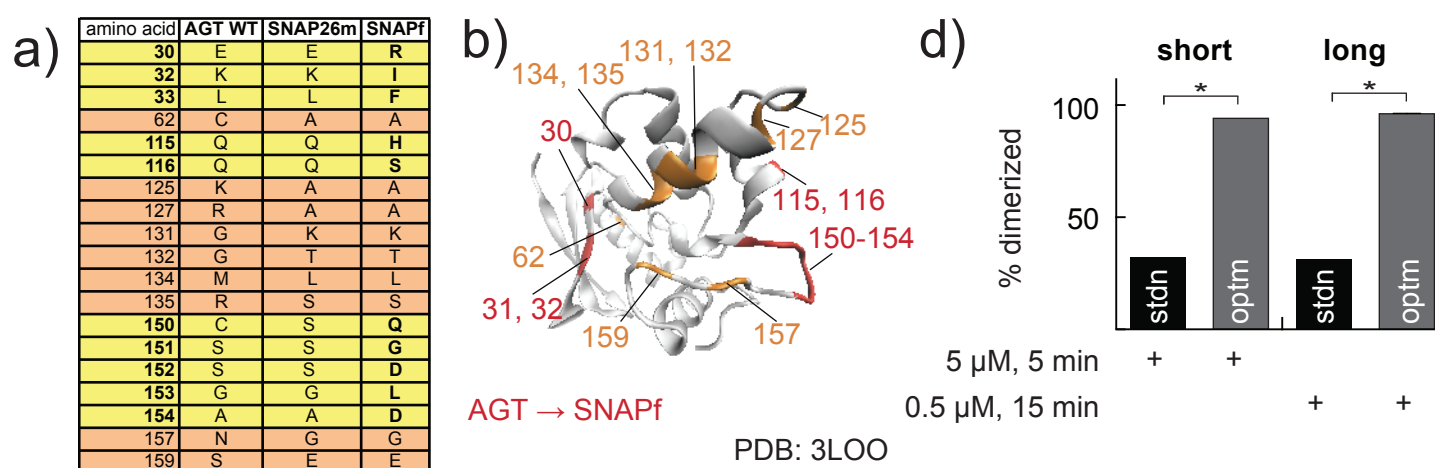


## Figure 8\_Construct Design of Anchor and Effector unit



**Figure 8. Optimal expression levels of anchor and effector unit.** a) Scheme describing two different backbone systems. Left: both ORFs are encoded from a single T2A-based plasmid. Right: Both ORF are encoded on separate plasmids. b) To determine the ratio of monomers HeLa cells expressing SNAP-GFP-T2A-Halo-GFP were exposed to 0.5 μM HaXS in complete cell culture medium at 37 °C. After 60 min cells were lysed and proteins were subjected to SDS-PAGE and immune blotting. Tagged proteins were detected using primary anti-GFP and horseradish peroxidase labeled secondary antibodies, and chemiluminescence. Ratio of monomers was determined by dividing amount of Halo-GFP and SNAP-GFP in DMSO treated cells. Quantification represent mean ± SEM of three independent experiments. To determine ratio of uncleaved and induced dimers, the amount of uncleaved T2A-dimers in DMSO treated cells was divided by amount of induced dimer in HaXS treated cells. Mean ± SEM of three independent experiments was determined. c) HeLa cells co-expressing equal amount of DNA (1 μg each, lane 1 and 2) or different ratios of amount of DNA (1.25 μg Halo-GFP and 0.75 μg SNAP-GFP lane 3 and 4, 1.5 μg Halo-GFP and 0.5 μg SNAP-GFP lane 5 and 6, 1.75 μg Halo-GFP and 0.25 μg SNAP-GFP lane 7 and 8) were exposed to 5 μM HaXS in complete cell culture medium at 37°C. After 15 min cells were lysed and analyzed by immune blot analysis as described previously. Ratio of monomers was determined as described previously. To determine the percentage of non-incorporated SNAP-GFP monomers, the amount of monomeric SNAP-GFP in HaXS treated cells was divided by the amount of SNAP-GFP in DMSO treated cells.

## Supplementary Figure 1\_Optimized SNAP-tag (SNAPf)

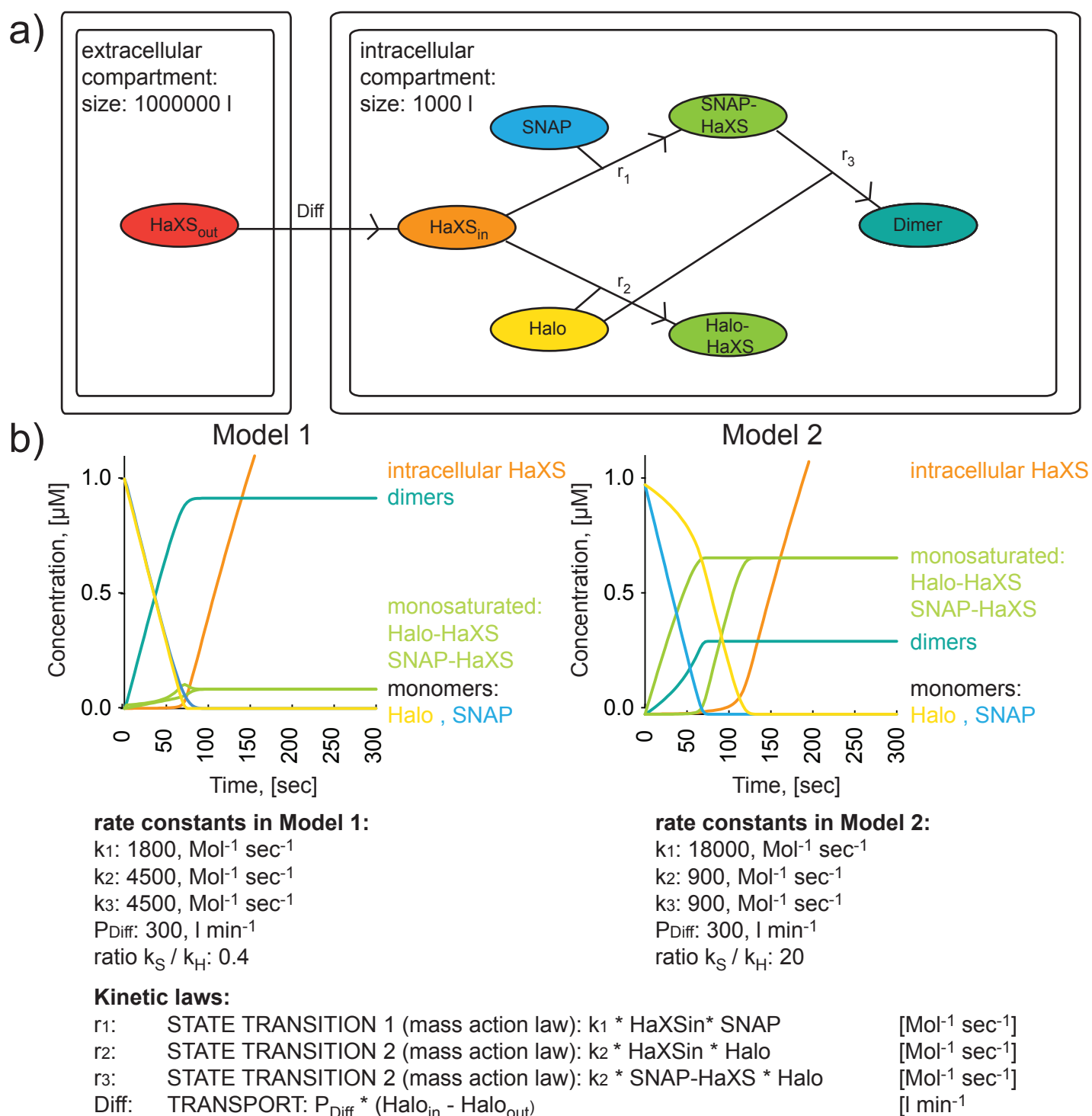


c)

AGT	MDKDCEMKRRTTLDSP LGKLELSGCEQGLHEIKLLGKGTSA	40
SNAP26m	MDKDCEMKRRTTLDSP LGKLELSGCEQGLHEIKLLGKGTSA	40
SNAP <sub>f</sub>	MDKDCEMKRRTTLDSP LGKLELSGCEQGLHRIIFLGKGTSA	40
AGT	ADAVEVPAPAAVLGGPEPLMQCTAWLNAYFHQPEAIEEFP	80
SNAP26m	ADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFP	80
SNAP <sub>f</sub>	ADAVEVPAPAAVLGGPEPLMQATAWLNAYFHQPEAIEEFP	80
AGT	VPALHHPVFQQESFTRQVLWKLKLVVVKFGEVISYQQLAAL	120
SNAP26m	VPALHHPVFQQESFTRQVLWKLKLVVVKFGEVISYQQLAAL	120
SNAP <sub>f</sub>	VPALHHPVFQQESFTRQVLWKLKLVVVKFGEVISYSHLAAL	120
AGT	AGNPKATRAVGGAMRGNPVPILIPCHR VVCS SSGAVGNYSG	160
SNAP26m	AGNPAATAAVKTALS GN PVPILIPCHR VVSS SSGAVGGYEG	160
SNAP <sub>f</sub>	AGNPAATAAVKTALS GN PVPILIPCHR VVQGDLDVGGYEG	160
AGT	GLAVKEWLLAHEGHRLGK PGLGGSSGLAGAWLKGAGATSG	200
SNAP26m	GLAVKEWLLAHEGHRLGK PGLG	182
SNAP <sub>f</sub>	GLAVKEWLLAHEGHRLGK PGLG	182
AGT	SPPAGR N	207
SNAP26m		182
SNAP <sub>f</sub>		182

**Supplementary Figure 1. Optimized SNAP-tag (SNAPf)** a) Amino acids substitutions from wild type human repair protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase (AGT) to SNAPf (orange) and from SNAP26m to SNAPf (yellow) In addition to SNAP26m, which carries 19 amino acids substitutions and a C-terminal deletion compared to the wild type human DNA repair protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase (hAGT), SNAPf carries ten extra mutations compared to SNAP26m, Sun et al., 2011 . b) Model of crystal structure of SNAPf, indicated mutations introduced to get SNAPf out of AGT (red) and SNAPf out of SNAP26m (orange), PDB: 3L00. c) Sequence alignment of AGT, SNAP26m, and SNAPf. The specific amino acid changes of the AGT mutants SNAP26m and SNAPf are depicted in boxes (light green from AGT to SNAPf, dark green from SNAP26m to SNAPf). d) HeLa cells co-expressing the standard or optimized SNAP- and Halo-tag combination were exposed to 0.5 resp. 5 μM HaXS in complete cell culture medium for the time indicated at 37°C before cells were lysed and analyzed by Western Blot analysis as described previously. Quantifications represent the mean ± SEM of two independent experiments.

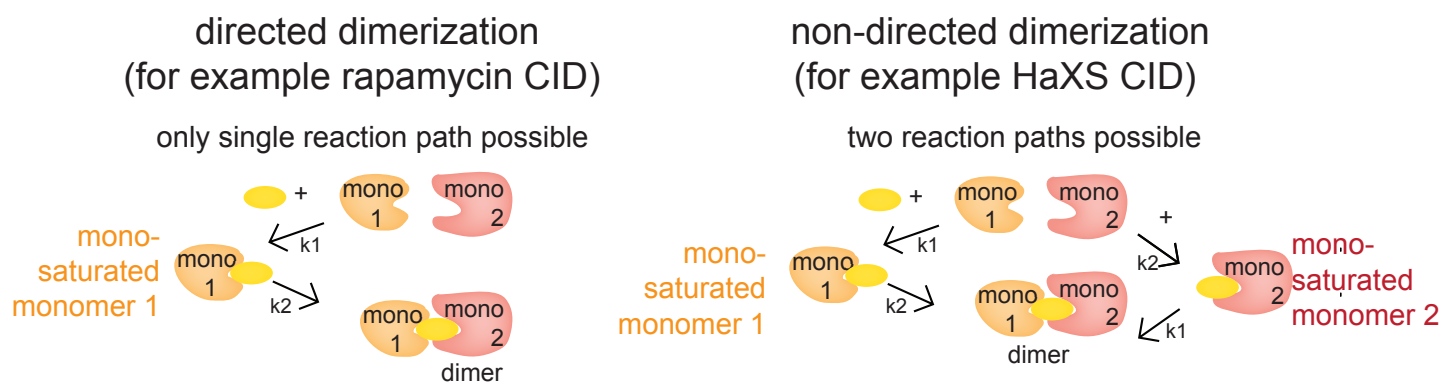
## Supplementary Figure 2\_Simulation of HaXS CID



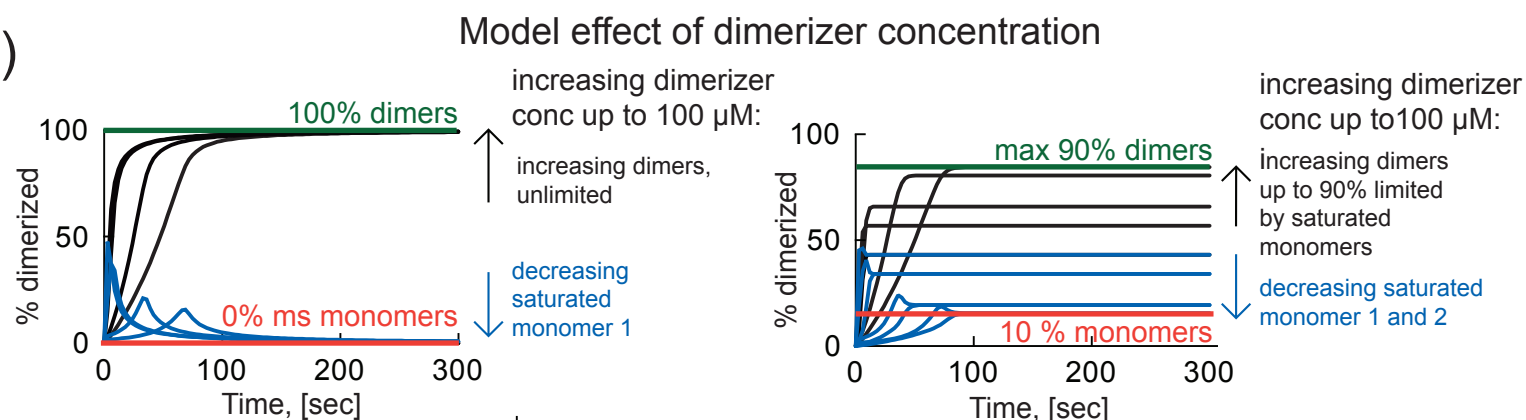
**Supplementary Figure 2.** a) Model of HaXS dimerization was created using CellDesigner 4.3, a software package that enables modeling of molecular interactions. Graphical representation of the model editor is shown. b) Simulation control interface of Model 1 and Model 2 with start values Halo = 1 µM, SNAP = 1 µM and HaXS<sub>out</sub> = 5 µM is shown. Kinetic laws as well as rate constants and diffusion coefficient for all reactions are indicated.

## Supplementary Figure 3\_Directed versus Non-directed CID

a)

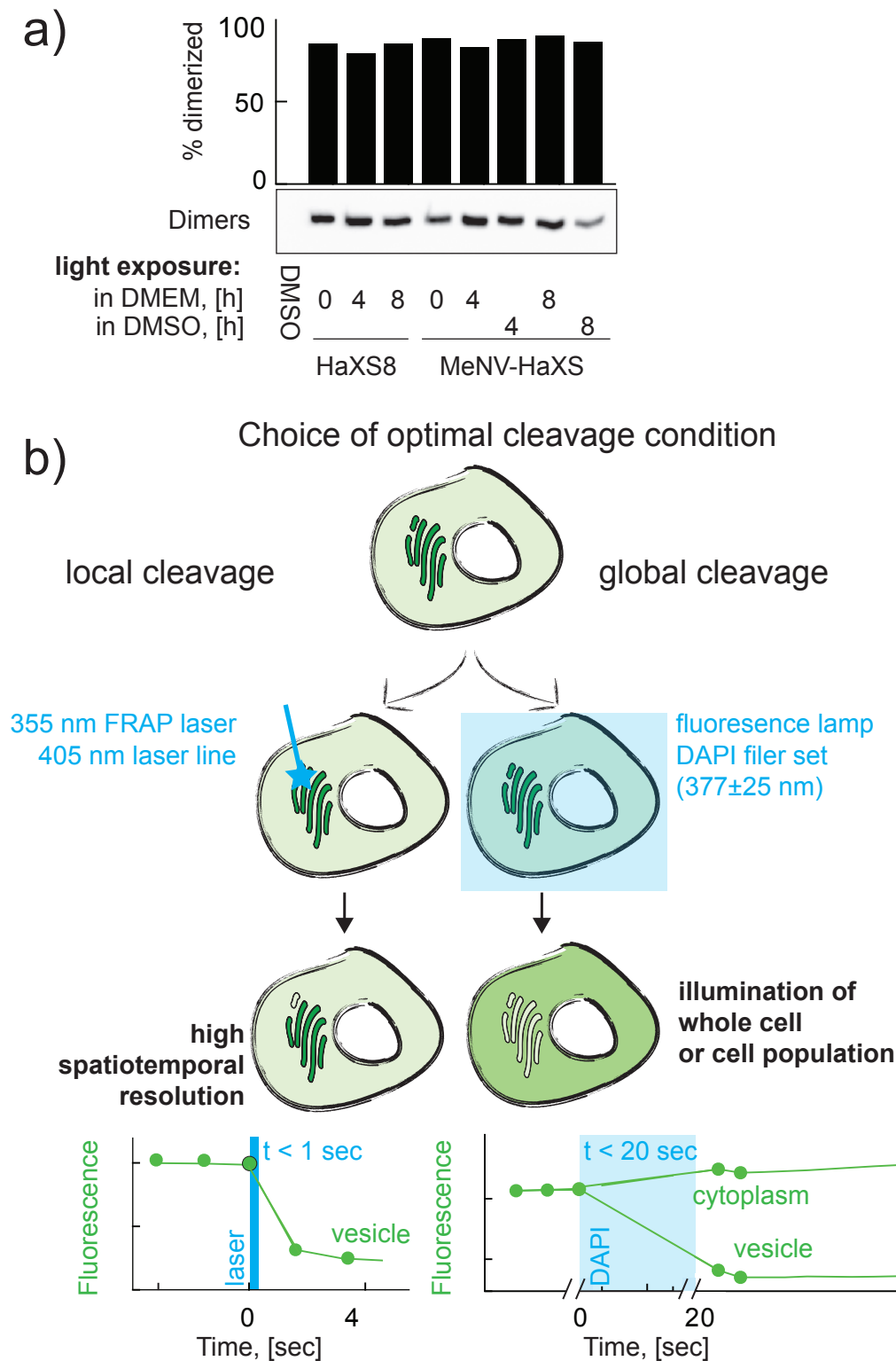


b)



**Supplementary Figure 3. Directed versus non-directed reactino paths of CID.** a) Scheme describing a dimerizer that follows a “directed” and a “non-directed” pathway. If the dimerizer can follow a “directed” dimerization path (as for example rapamycin), the dimerizer can only bind to monomer 2 (dark red) if it has first reacted with monomer 1, whereas dimerizer that follows a “non-directed” pathway can either first react with monomer 1 (orange) or monomer 2 (dark red). b) In a “directed” CID (left): high dimerizer concentration result in up to 100% dimer formation, whereas in a “non-directed” CID (right), high concentration of the dimerizer never result in 100% of dimer formation, as the amount of saturated monomers is increasing and limiting the dimer formation.

## Supplementary Figure 4\_Intracellular cleavage efficiency



**Supplementary Figure 4. Photocleavage conditions of MeNV-HaXS.** a) HeLa cells co-expressing SNAP-GFP and Halo-GFP were exposed to 5  $\mu$ M MeNV-HaXS or 5  $\mu$ M HaXS8 which were previously exposed to environmental light for the indicated times. After 15 min cells were lysed and analyzed for induced SNAP/HaloTag dimers as described previously. b) Scheme showing two possibilities to cleave MeNV-HaXS induced dimers: local illumination with a laser pulse ( $t < 1$  sec) to achieve high spatiotemporal precise cleavage or illumination with light from a fluorescence lamp and DAPI filter set ( $t < 20$  sec) for global cleavage of MeNV-HaXS of whole cell population with less temporal resolution.



# SHORT COMMUNICATION

DOI: 10.1002/ejoc.200((will be filled in by the editorial staff))

## Efficient Synthesis of C16-carbamyl Rapalogs via Lewis Acid Decomplexation

Ruben Cal,<sup>[a]</sup> Mirjam Zimmermann,<sup>[a]</sup> Kaspar Zimmermann,<sup>[b]</sup> Vladimir Cmiljanovic,<sup>[a]</sup> Olivier Jacques,<sup>[a]</sup> Edwin Constable,<sup>[b]</sup> Daniel Häussinger,<sup>[b]</sup> Florent Beaufils<sup>[a]\*</sup> and Matthias P. Wymann<sup>[a]\*</sup>

**Keywords:** Carbamylation/Natural products/Lewis Acids/Reaction Mechanisms/ Protein-protein interactions/chemically-induced dimerization/CID

C16-derivatized rapamycin has been utilized in chemically-induced dimerization (CID). Minor rapamycin contaminations can interfere with endogenous mammalian target of rapamycin (mTOR) complex 1 (TORC1) activity. Here, we describe the C16-Lewis acid-mediated carbamylation of rapamycin. The reported method is efficient and proceeds under mild reactions conditions to give excellent yields of differently substituted rapalogs. Furthermore, an NMR-

based mechanistic analysis of the Lewis acid-mediated carbamylation is described. The resulting (*R*)-C16-phenylcarbamate-rapamycin ((*R*)-pcRap) rapidly dimerizes mutated FRB<sub>T2098L</sub>- and FKBP12 fusion proteins, but spares endogenous TORC1 signaling. Pure (*R*)-pcRap can therefore be used in CID approaches investigating phosphoinositide 3-kinase (PI3K)/mTOR pathways.

[a] University of Basel, Department of Biomedicine, Mattenstrasse 28, Basel, Switzerland  
[b] University of Basel, Department of Chemistry, Spitalstrasse 51, Basel, Switzerland  
Fax: +41 61 267 3566  
E-mail: Florent.Beaufils@Unibas.CH,  
E-mail: Matthias.Wymann@Unibas.CH  
Homepage: <http://biomedizin.unibas.ch>  
Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/ejoc.xxxxxxxx>.

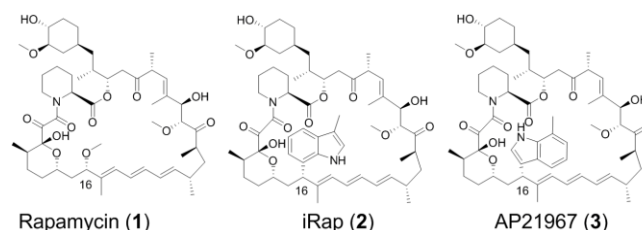


Figure 1. Structure of Rapamycin (1), C16-(*S*)-3-methylindole-rapamycin (iRap; 2) and C16-(*S*)-7-methylindole-rapamycin (AP21967; 3).

### Introduction

Rapamycin (1) binds with high affinity to the 12 kDa FK506 binding protein (FKBP12). Once formed, this FKBP-rapamycin complex binds the FRB (FKBP12-rapamycin-binding) domain of mTOR (mammalian target of rapamycin), which is a major signaling hub controlling cell growth and proliferation.<sup>[1]</sup> Fusion proteins of FKBP12 and FRB can be conditionally dimerized by the addition of rapamycin, which makes this system a valuable tool to chemically dimerize proteins of interest.<sup>[2,3]</sup> As interference with endogenous mTOR complex 1 (TORC1) signaling leads to feedback loops affecting upstream events, such as phosphoinositide 3-kinase activation,<sup>[1,4,5]</sup> rapamycin analogs (so called rapalogs<sup>[6]</sup>) and alternative covalent dimerizers<sup>[7]</sup> have been designed.

Chemical modifications at C16-methoxy group of rapamycin were reported to prevent cross-reaction with endogenous TORC1, 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.<sup>[6]</sup> Rapalogs devoid of wild type FRB domain-binding such as iRap (2) and AP21967 (3) bear bulky substituents at the C16 position, and can be used in combination with mutated FRB proteins even in studies of growth control, immunity or metabolism.<sup>[8]</sup> Although various nucleophiles were introduced at the C16 position, no general method to introduce carbamates was published.

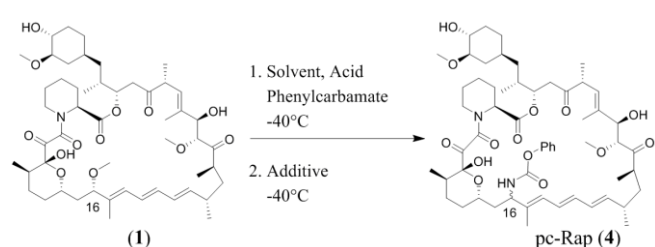
Here, we report the synthesis of C16-phenylcarbamate-rapamycin (pcRap) (4) using a modified Lewis acid-mediated substitution of the C16 position of rapamycin. A detailed optimization of the reaction procedure, as well as the investigation of the reaction mechanism is described: the scope of this reaction was also explored by the synthesis of C16-benzyl-, *n*-butyl- and *tert*-butyl carbamate-rapalogs, which were produced with excellent yields. Finally, protein translocation with the obtained (*R*)-pcRap and a FRB<sub>T2098L</sub>/FKBP12 system has been successfully achieved without any interference with endogenous mTOR signaling.

### Results and Discussion

Under acidic conditions, the C16-methoxy group of rapamycin undergoes a heterolytic cleavage, and the stabilized carbocation can be trapped with various nucleophiles.<sup>[6,8,9]</sup> Attempts to carry out the nucleophilic addition of phenyl carbamate (pc) to C16-rapamycin carbocation by using TFA (Table 1, entry 1) and *p*-Toluenesulfonic acid (Table 1, entry 2), in CH<sub>2</sub>Cl<sub>2</sub> at -40°C, gave an inseparable mixture of products containing only traces of pcRap (4). 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 targeted the use of a Lewis acid to generate the carbocation at C16.<sup>[6]</sup> Addition of BF<sub>3</sub>-Et<sub>2</sub>O to a solution of Rap and pc, in CH<sub>2</sub>Cl<sub>2</sub> at -40°C, produced the desired carbocation at C16, which was indicated by a dark-red color change. Although the active intermediate was observed, a new

mixture of inseparable products containing starting material, the product of methoxy-elimination<sup>[6]</sup> and traces of pcRap was formed (products identified by mass spectroscopy; Table 1, entry 3). The formation of the transient carbocation complex was not accomplished by using THF as a solvent (no color change was seen), but no degradation products were generated, and only starting material (rapamycin and pc) was recovered (Table 1, entry 4). Addition of BF<sub>3</sub>-Et<sub>2</sub>O, at -40°C, to a solution of pc and Rap in CH<sub>2</sub>Cl<sub>2</sub> (generation of the C16 carbocation), followed by the addition of an excess of THF, 5 minutes later, gave pcRap with a yield of 83% (diastereoselective ratio [d.r.] 52:48; Table 1, entry 5). Polarity effects were investigated for this sequential transformation, and similar results were obtained using Et<sub>2</sub>O or dioxane as additives (Table 1, entry 6 and 7).

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



Entry	Solvent	Additive	Acid	Yield [%]	C16 (R)/(S) d.r. [f]
1	CH <sub>2</sub> Cl <sub>2</sub>	-	TFA <sup>[a]</sup>	Traces	-
2	CH <sub>2</sub> Cl <sub>2</sub>	-	p-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 <sup>[e]</sup>	BF <sub>3</sub> -Et <sub>2</sub> O <sup>[c]</sup>	83	52:48
6	CH <sub>2</sub> Cl <sub>2</sub> <sup>[d]</sup>	Dioxane <sup>[e]</sup>	BF <sub>3</sub> -Et <sub>2</sub> O <sup>[c]</sup>	76	50:50
7	CH <sub>2</sub> Cl <sub>2</sub> <sup>[d]</sup>	Et <sub>2</sub> O <sup>[e]</sup>	BF <sub>3</sub> -Et <sub>2</sub> O <sup>[c]</sup>	79	50:50

[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.; [f] d.r. diastereoselective ratio.

The fact that the product is obtained as a racemate with respect to the configuration at C16, implies a S<sub>N</sub>1 reaction mechanism mediated by BF<sub>3</sub>. To support this assumption, low temperature NMR studies of the reaction mixture before adding the THF were performed. <sup>1</sup>H-, <sup>11</sup>B- and <sup>19</sup>F-NMR spectra show the presence of an equilibrium between rapamycin and the BF<sub>3</sub>-coordinated rapamycin, as well as further BF<sub>3</sub>-containing species in slow exchange with each other. Monitoring the reaction between BF<sub>3</sub>-Et<sub>2</sub>O and rapamycin in CD<sub>2</sub>Cl<sub>2</sub> at -30°C by <sup>1</sup>H-NMR (Supporting information see Figure S1), yielded very broad spectra, indicating a dynamic equilibrium. The disappearance of the correlation peak for the methoxy group at C16 in the HMQC spectra (Supporting information Figure S1) clearly demonstrates that C16 is the most favorable coordination site for the reversible formation of a tight ion pair, while the other methoxy groups are less affected. An irreversible formation of the free or solvent separated carbocation

can be ruled out, as no product is formed under these conditions. On the other hand, using THF as the sole solvent, this ion pair is also not formed, because the large excess of THF prevents the BF<sub>3</sub> from interacting with the methoxide at C16. In contrast, addition of THF as co-solvent leads to sufficient stabilisation of the carbocation, subsequent formation of a solvent separated ion pair and finally to a fast reaction with the nucleophile. The fact that CH<sub>2</sub>Cl<sub>2</sub> and THF have similar dielectric constants indicates that the stabilisation of the carbocation is an effect of the coordinating ability of the co-solvent, rather than its polarity. This is further supported by the observation that the reaction is also possible using less polar co-solvents, such as diethyl ether. Kinetic data of the reaction has not been determined because the reaction mechanism was not the main target of this project.

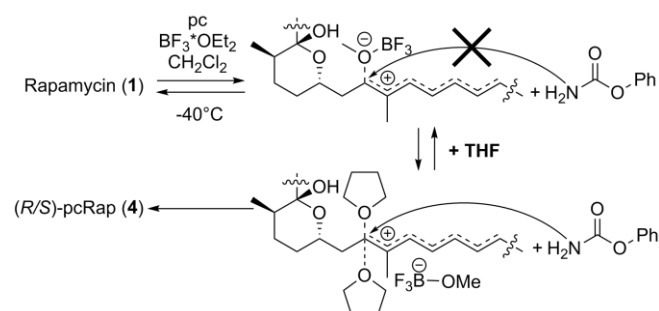
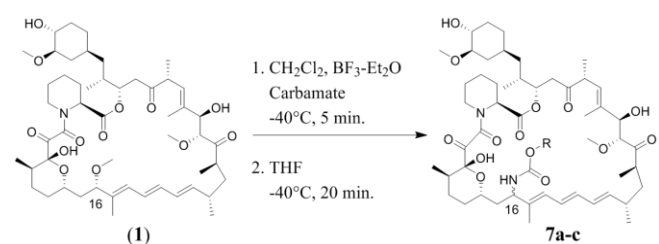


Figure 2. Proposed mechanism for the Lewis acid-mediated formation of pcRap.

Once the optimal experimental conditions had been determined, the scope and generality of the reaction was next investigated with the introduction of representative carbamates. Using benzyl carbamate, C16-benzylcarbamate-Rap (**7a**) was synthesized in excellent yield (88%, Table 2, entry 1). Reaction with the lower electron-donating, and unhindered aliphatic n-butyl carbamate gave C16-n-butylcarbamate-Rap (**7b**) with a yield of 73% (Table 2, entry 2). Interestingly, comparable results were still achieved with the acid sensitive, and sterically demanding tert-butyl carbamate (compound **7c**, 78% yield, Table 2, entry 3). We assume that compound **7c** could be applied in extended selective modification of the rapamycin core via the C16-Boc-protected amine.

Table 2. BF<sub>3</sub>-Et<sub>2</sub>O/THF-Mediated Synthesis of C16-carbamate-Rapalogs.



Entry	Carbamate	R	Product	Yield [%]	C16 (R)/(S) ratio
1	benzylcarbamate	Bn	7a	88	51:49
2	n-butylcarbamate	nBu	7b	73	51:49
3	tert-butylcarbamate	tBu	7c	78	51:49

Indicated carbamate (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.

On the basis of previous studies focusing on the development of rapalogs sparing endogenous mTOR,<sup>[6]</sup> each pcRap diastereoisomer might possess a stereospecific profile towards endogenous wild type and mutated FRB<sub>T2098L</sub> domains. The stereochemistry of (*R*)-pcRap and (*S*)-pcRap was thus confirmed by ROESY experiments (Figure 3). Correlations between the protons at the cyclic hemiacetal, at C15 and C16 as well as 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 3. Only the (*S*)-configuration at C16 could explain the observed correlation between H16 and H13a, and without violating any of the other correlations. The latter correlation was not observed for the (*R*)- diastereoisomer. A detailed description of the assignment strategy of both diastereoisomers is given in the supporting information.

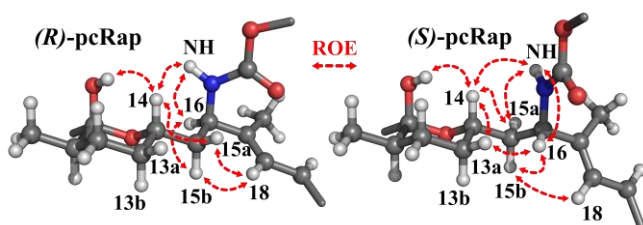


Figure 3. (*R*)-pcRap (d.r. 51%), left, and (*S*)-pcRap (d.r. 49%), right. Key ROESY correlations are indicated in red.

Finally, a potential interference of pcRap isomers with endogenous TORC1, and their validation as tools for chemically-induced dimerization (CID), was investigated. To test for a potential interference of pcRap isomers with TORC1, the activation status of a kinase substrate of TORC1, the ribosomal protein S6 kinase (p70<sup>S6K</sup>) was monitored by analyzing the phosphorylation of Thr389 on p70<sup>S6K</sup>. Figure 4A demonstrates that exposure of HEK cells to rapamycin or (*S*)-pcRap leads to the inhibition of the phosphorylation of S6K1 on Thr389, whereas (*R*)-pcRap does not interfere with TORC1, as S6K1 phosphorylation is maintained (Figure 4A). Luengo et al.<sup>[6]</sup> have investigated the activity of selected (*R*)- and (*S*)-rapalog stereoisomers on mTOR-dependent T-cell responses. Due to co-eluting contaminations of rapamycin and rapamycin by-products along with the (*S*)-pcRap, it is presently not entirely clear if contaminants or (*S*)-pcRap impair mTOR activity. Using the biologically validated (*R*)-pcRap, we next targeted the translocation of a cytosolic fluorescent protein to the cell's plasma membrane: a red fluorescent protein (monomeric RFP; TagRFP) fused to mutated FRB (FRB<sub>T2098L</sub>), was co-expressed with FKBP12-GFP with a C-terminal isoprenylation sequence from KRas-4B (the KRas-4B CAAX box), which targets the resulting FKBP12-GFP-CAAX fusion protein exclusively to the plasma membrane. Upon addition of (*R*)-pcRap, the cytosolic mutated FRB<sub>T2098L</sub>-RFP was translocated to the plasma membrane after addition of (*R*)-pcRap (see Figure 4B).

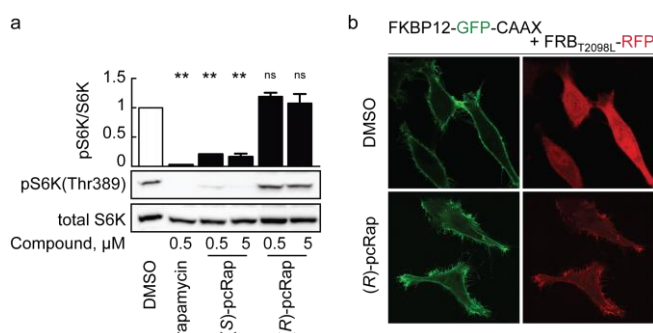


Figure 4. (A) HEK cells grown in fetal calf serum supplemented media were exposed to DMSO, rapamycin, (*S*)-pcRap and (*R*)-pcRap for 15 min at 37°C, before cells were lysed and proteins were subjected to SDS-PAGE,

and immune-blotting using antibodies against total p70<sup>S6K</sup> (total S6K), and phosphorylated p70<sup>S6K</sup> (pS6K, Thr389). Data represent means  $\pm$  SEM, n=3; difference from DMSO control: \*\* indicates p<0.01; ns, not significant. (B) HeLa cells expressing the FKBP12-GFP-CAAX membrane anchor [green (-CAAX is the polybasic isoprenylation sequence from KRas-4B)] and FRB<sub>T2098L</sub>-RFP (red) fusion proteins were exposed to DMSO or 0.5  $\mu$ M (*R*)-pcRap for 15 min. at 37°C. Translocation of FRB<sub>T2098L</sub>-RFP to the plasma membrane was imaged by confocal microscopy in fixed cells.

## Conclusions

We have successfully developed a facilitated access to C16-carbamyl-rapalogs using mild reaction conditions. The mechanism of the addition of carbamates onto the C16-carbocation appears to involve the prior establishment of a tight rapamycin-BF3 ion pair (in a dynamic equilibrium). The addition of THF allows subsequently an irreversible formation of the desired rapalogs. The structures of the novel C16-rapamycin diastereomers were determined through NMR spectroscopic analysis. Additionally, we have demonstrated that phenyl carbamate substitution can abolish potential interferences with endogenous mTOR, and high-lighted (*R*)-pcRap as a powerful, non-toxic tool to chemically induce protein-protein interactions.

## Experimental Section

**General methods:** Materials and reagents were of the highest commercially available grade and used without further purification. Reactions were monitored by thin layer chromatography using Merck silica gel 60 F254 plates. Compounds were visualized by UV and KMnO<sub>4</sub>. Flash chromatography was performed using Merck silica gel 60, particle size 40 - 63  $\mu$ m and DCM/MeOH 40:1). All NMR experiments were performed on a Bruker Avance III NMR spectrometer operating at 600.13 MHz, equipped with a z-axis pulsed field gradient broadband direct detection probe-head. Chemical shifts were referenced to residual solvent peaks and the temperature was calibrated using a methanol sample. NMR experiments for the assignment of the Rapalogs were performed at 298 K and for the mechanistic study at 243 K. 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. All solvents used for reactions were purchased as anhydrous grade from Fluka. Solvents for extractions, column chromatography and TLC were commercial grade. Mass spectra were recorded with a VG70-250 (FAB), Finnigan MAT MS 312 (EI) or Finnigan MAT LCQ (ESI) spectrometer. High-resolution mass spectra were recorded with a thermo Fisher Scientific LTQ Orbitrap XL, nanoelectrospray ion source.

**Analytical HPLC:** HPLC analyses were carried out using a Dr. Maisch Reprosil 100 Si, 5  $\mu$ m 100 x 10 mm column. The instrument was equipped with an SPD-M20A diode array detector and an RF-10AXL fluorescence detector. 1 mg of the crude mixture was dissolved in 1 ml CH<sub>2</sub>Cl<sub>2</sub>. 50  $\mu$ l of this standard solution was injected in the HPLC system. Samples were eluted using a 60-min linear gradient of 100:0 to 96:4 (CH<sub>2</sub>Cl<sub>2</sub>:isopropanol) and a flow rate of 2 ml·min<sup>-1</sup>.

**Preparative HPLC:** Preparative HPLC was performed with a Dr. Maisch Reprosil 100 Si, 5  $\mu$ m 250 x 40 mm column. 30 mg of the crude mixture was dissolved in 2 ml CH<sub>2</sub>Cl<sub>2</sub> and injected into the HPLC system. Samples were eluted using a 100-min linear



gradient of 100:0 to 92:8 (CH<sub>2</sub>Cl<sub>2</sub>:isopropanol) and a flow rate of 30 ml·min<sup>-1</sup>.

**General Procedure for the synthesis of C16-carbamyl rapalogs:** Rapamycin (100 mg, 0.109 mmol) and the corresponding carbamate (0.656 mmol) were dissolved in 20 ml of CH<sub>2</sub>Cl<sub>2</sub>. The reaction mixture was cooled to -40°C and BF<sub>3</sub>·Et<sub>2</sub>O (57 µl, 0.438 mmol) was added. This solution was stirred for 5 min before 10 ml of THF was added. After 30 min stirring at -40°C, the reaction was quenched with a solution of sat. NaHCO<sub>3</sub>. The aqueous layers were extracted with ethyl acetate three times. The combined organic layers were dried over sodium sulfate and concentrated under reduced pressure. The crude mixture was analyzed by analytical HPLC. The crude mixture was redissolved in CH<sub>2</sub>Cl<sub>2</sub> and cooled down to -40°C. BF<sub>3</sub>·Et<sub>2</sub>O (57 µl, 0.438 mmol) were added and the mixture was stirred for 5 min, prior to THF addition. The mixture was stirred for 30 min at -40°C. The reaction was quenched with a solution of sat. NaHCO<sub>3</sub> and the aqueous layer were extracted with ethyl acetate three times. The combined organic layers were dried over sodium sulfate and concentrated under reduced pressure. This reaction procedure was repeated 3-4 times until complete consumption of rapamycin (HPLC monitoring). Purification of the crude mixture was done by Flash chromatography and separation of the diastereomers was done by preparative HPLC.

**C16-phenyl carbamyl rapamycin, pcRap (4):** Flash chromatography gave 92 mg as a mixture of two diastereomers (83% yield; 52:48 ratio). Diastereomers were separated with preparative HPLC. 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 C<sub>57</sub>H<sub>82</sub>O<sub>14</sub>N<sub>2</sub> [M-Na<sup>+</sup>]: calcd. 1041.5658, found 1041.5664.

**(R)-pcRap:** <sup>1</sup>H NMR (600MHz, DMSO) δ: 7.87 (d, <sup>3</sup>J = 7.8 Hz, 1H, H<sub>52</sub>), 7.37 (dd, <sup>3</sup>J = 6.8 Hz, 8.0 Hz, 2H, H<sub>56</sub>), 7.19 (t, <sup>3</sup>J = 6.8 Hz, 1H, H<sub>57</sub>), 7.09 (d, <sup>3</sup>J = 8.0 Hz, 2H, H<sub>55</sub>), 6.40 (dd, <sup>3</sup>J = 10.8 Hz, 11.4 Hz, 1H, H<sub>19</sub>), 6.39 (s, 1H, OH<sub>10</sub>), 6.21 (dd, <sup>3</sup>J = 11.4, 12.4 Hz, 1H, H<sub>20</sub>), 6.13 (d, <sup>3</sup>J = 10.6 Hz, 12.4 Hz, 1H, H<sub>21</sub>), 6.08 (d, <sup>3</sup>J = 10.8 Hz, 1H, H<sub>18</sub>), 5.52 (dd, <sup>3</sup>J = 9.4 Hz, 12.4 Hz, 1H, H<sub>22</sub>), 5.24 (d, <sup>3</sup>J = 4.4 Hz, 1H, OH<sub>28</sub>), 5.22 (d, <sup>3</sup>J = 9.8 Hz, 1H, H<sub>30</sub>), 5.10 (m, 1H, H<sub>34</sub>), 4.96 (d, <sup>3</sup>J = 5.3 Hz, 1H, H<sub>2</sub>), 4.59 (d, <sup>3</sup>J = 4.3 Hz, 1H, OH<sub>40</sub>), 4.07 (m, 1H, H<sub>28</sub>), 4.02 (m, 1H, H<sub>16</sub>), 4.01 (m, 1H, H<sub>27</sub>), 3.95 (m, 1H, H<sub>14</sub>), 3.57 (m, 1H, H<sub>6</sub>), 3.31 (m, 3H, H<sub>51</sub>), 3.27 (m, 1H, H<sub>31</sub>), 3.18 (m, 3H, H<sub>50</sub>), 3.16 (m, 1H, H<sub>40</sub>), 3.06 (m, 1H, H<sub>6</sub>), 2.81 (m, 1H, H<sub>39</sub>), 2.81 (m, 1H, H<sub>33</sub>), 2.44 (m, 1H, H<sub>33</sub>), 2.39 (m, 1H, H<sub>25</sub>), 2.25 (m, 1H, H<sub>23</sub>), 2.16 (m, 1H, H<sub>3</sub>), 2.04 (m, 1H, H<sub>11</sub>), 1.91 (m, 1H, H<sub>38</sub>), 1.84 (m, 1H, H<sub>15</sub>), 1.75 (m, 1H, H<sub>41</sub>), 1.74 (m, 3H, H<sub>47</sub>), 1.72 (m, 3H, H<sub>44</sub>), 1.70 (m, 1H, H<sub>35</sub>), 1.67 (m, 1H, H<sub>4</sub>), 1.64 (m, 1H, H<sub>13</sub>), 1.61 (m, 1H, H<sub>12</sub>), 1.60 (m, 1H, H<sub>3</sub>), 1.55 (m, 1H, H<sub>12</sub>), 1.52 (m, 3H, H<sub>5</sub>, H<sub>15</sub>, H<sub>42</sub>), 1.40 (m, 1H, H<sub>24</sub>), 1.36 (m, 2H, H<sub>4</sub>, H<sub>5</sub>), 1.31 (m, 1H, H<sub>13</sub>), 1.27 (m, 1H, H<sub>37</sub>), 1.17 (m, 1H, H<sub>41</sub>), 1.10 (m, 1H, H<sub>36</sub>), 1.04 (m, 1H, H<sub>24</sub>), 0.99 (m, 3H, H<sub>45</sub>), 0.96 (m, 1H, H<sub>36</sub>), 0.91 (m, 3H, H<sub>48</sub>), 0.85 (m, 1H, H<sub>42</sub>), 0.81 (m, 3H, H<sub>46</sub>), 0.80 (m, 3H, H<sub>49</sub>), 0.76 (m, 3H, H<sub>43</sub>), 0.58 (m, 1H, H<sub>38</sub>). <sup>13</sup>C NMR: 210.2 (C<sub>26</sub>), 208.0 (C<sub>32</sub>), 199.0 (C<sub>9</sub>), 169.4 (C<sub>1</sub>), 167.0 (C<sub>8</sub>), 153.8 (C<sub>53</sub>), 151.1 (C<sub>54</sub>), 140.0 (C<sub>17</sub>), 138.4 (C<sub>22</sub>), 136.9 (C<sub>29</sub>), 131.8 (C<sub>20</sub>), 130.4 (C<sub>21</sub>), 129.2 (C<sub>56</sub>), 127.6 (C<sub>19</sub>), 124.8 (C<sub>57</sub>), 124.6 (C<sub>30</sub>), 123.6 (C<sub>18</sub>), 121.6 (C<sub>55</sub>), 99.0 (C<sub>10</sub>), 85.5 (C<sub>27</sub>), 83.8 (C<sub>39</sub>), 75.7 (C<sub>28</sub>), 73.4 (C<sub>34</sub>), 73.2 (C<sub>40</sub>), 66.6 (C<sub>14</sub>), 56.9 (C<sub>50</sub>), 56.8 (C<sub>51</sub>), 53.7 (C<sub>16</sub>), 51.2 (C<sub>2</sub>), 44.8 (C<sub>31</sub>), 43.9 (C<sub>6</sub>), 40.4 (C<sub>33</sub>), 39.9 (C<sub>25</sub>), 39.6 (C<sub>24</sub>), 39.4 (C<sub>15</sub>), 38.3 (C<sub>36</sub>), 35.2 (C<sub>38</sub>), 34.7 (C<sub>11</sub>), 34.4 (C<sub>23</sub>), 33.5 (C<sub>35</sub>), 32.9 (C<sub>41</sub>), 32.6 (C<sub>37</sub>), 31.3 (C<sub>13</sub>, C<sub>42</sub>), 26.6 (C<sub>3</sub>), 26.4 (C<sub>12</sub>), 24.4 (C<sub>5</sub>), 21.5 (C<sub>45</sub>), 20.5 (C<sub>4</sub>), 16.0 (C<sub>48</sub>), 15.6 (C<sub>43</sub>), 15.2 (C<sub>44</sub>), 15.0 (C<sub>49</sub>), 13.8 (C<sub>47</sub>), 12.9 (C<sub>46</sub>).

**(S)-pcRap:** <sup>1</sup>H NMR (600MHz, DMSO) δ: 8.07 (d, <sup>3</sup>J = 8.8 Hz, 1H, H<sub>52</sub>), 7.37 (dd, <sup>3</sup>J = 7.38 Hz, 7.68 Hz, 2H, H<sub>56</sub>), 7.19 (t, <sup>3</sup>J = 7.38 Hz, 1H, H<sub>57</sub>), 7.07 (d, <sup>3</sup>J = 7.68 Hz, 2H, H<sub>55</sub>), 6.51 (s, 1H, OH<sub>10</sub>), 6.36 (dd, <sup>3</sup>J = 11.3 Hz, 13.1 Hz, 1H, H<sub>19</sub>), 6.20 (dd, <sup>3</sup>J = 10.8 Hz, 13.1 Hz, 1H, H<sub>20</sub>), 6.12 (dd, <sup>3</sup>J = 10.8 Hz, 13.1 Hz, 1H, H<sub>21</sub>), 6.05 (d, <sup>3</sup>J = 10.9 Hz, 1H, H<sub>18</sub>), 5.48 (dd, <sup>3</sup>J = 9.2 Hz, 13.1 Hz, 1H, H<sub>22</sub>), 5.27 (d, <sup>3</sup>J = 4.4 Hz, 1H, OH<sub>28</sub>), 5.10 (d, <sup>3</sup>J = 10.0 Hz, 1H,

H<sub>30</sub>), 4.99 (m, 1H, H<sub>34</sub>), 4.96 (d, <sup>3</sup>J = 6.3 Hz, 1H, H<sub>2</sub>), 4.60 (m, 1H, OH<sub>40</sub>), 4.13 (m, 1H, H<sub>16</sub>), 4.01 (m, 1H, H<sub>28</sub>), 3.94 (m, 1H, H<sub>14</sub>), 3.87 (d, <sup>3</sup>J = 5.0 Hz, 1H, H<sub>27</sub>), 3.45 (m, 1H, H<sub>6</sub>), 3.33 (m, 1H, H<sub>31</sub>), 3.32 (m, 3H, H<sub>51</sub>), 3.25 (m, 1H, H<sub>6</sub>), 3.17 (m, 1H, H<sub>40</sub>), 3.16 (m, 3H, H<sub>50</sub>), 2.83 (m, 1H, H<sub>39</sub>), 2.76 (m, 1H, H<sub>33</sub>), 2.49 (m, 1H, H<sub>25</sub>), 2.44 (m, 1H, H<sub>33</sub>), 2.22 (m, 1H, H<sub>23</sub>), 2.08 (m, 1H, H<sub>3</sub>), 2.05 (m, 1H, H<sub>11</sub>), 1.96 (m, 1H, H<sub>15</sub>), 1.92 (m, 1H, H<sub>38</sub>), 1.82 (m, 1H, H<sub>13</sub>), 1.75 (m, 6H, H<sub>41</sub>, H<sub>42</sub>, H<sub>44</sub>), 1.74 (m, 3H, H<sub>47</sub>), 1.70 (m, 1H, H<sub>35</sub>), 1.67 (m, 1H, H<sub>4</sub>), 1.63 (m, 1H, H<sub>3</sub>), 1.58 (m, 1H, H<sub>5</sub>), 1.54 (m, 1H, H<sub>12</sub>), 1.52 (m, 1H, H<sub>12</sub>), 1.43 (m, 1H, H<sub>15</sub>), 1.41 (m, 2H, H<sub>4</sub>, H<sub>24</sub>), 1.31 (m, 1H, H<sub>5</sub>), 1.28 (m, 1H, H<sub>37</sub>), 1.25 (m, 1H, H<sub>13</sub>), 1.18 (m, 1H, H<sub>41</sub>), 1.07 (m, 1H, H<sub>36</sub>), 1.06 (m, 1H, H<sub>24</sub>), 0.98 (m, 3H, H<sub>45</sub>), 0.97 (m, 1H, H<sub>36</sub>), 0.89 (m, 3H, H<sub>48</sub>), 0.86 (m, 1H, H<sub>42</sub>), 0.85 (m, 3H, H<sub>46</sub>), 0.79 (m, 3H, H<sub>49</sub>), 0.76 (m, 3H, H<sub>43</sub>), 0.60 (m, 1H, H<sub>38</sub>). <sup>13</sup>C NMR: 210.9 (C<sub>26</sub>), 207.7 (C<sub>32</sub>), 198.9 (C<sub>9</sub>), 169.2 (C<sub>1</sub>), 166.8 (C<sub>8</sub>), 154.2 (C<sub>53</sub>), 151.1 (C<sub>54</sub>), 138.9 (C<sub>22</sub>), 138.2 (C<sub>17</sub>), 132.0 (C<sub>20</sub>), 130.4 (C<sub>21</sub>), 129.3 (C<sub>56</sub>), 127.3 (C<sub>19</sub>), 125.5 (C<sub>18</sub>), 124.9 (C<sub>57</sub>), 125.3 (C<sub>30</sub>), 121.7 (C<sub>55</sub>), 99.0 (C<sub>10</sub>), 85.9 (C<sub>27</sub>), 83.8 (C<sub>39</sub>), 75.8 (C<sub>28</sub>), 73.7 (C<sub>34</sub>), 73.2 (C<sub>40</sub>), 66.8 (C<sub>14</sub>), 57.1 (C<sub>50</sub>), 56.7 (C<sub>51</sub>), 54.9 (C<sub>16</sub>), 50.6 (C<sub>2</sub>), 45.3 (C<sub>31</sub>), 43.4 (C<sub>6</sub>), 39.7 (C<sub>33</sub>), 39.6 (C<sub>24</sub>), 39.2 (C<sub>25</sub>), 38.4 (C<sub>15</sub>, C<sub>36</sub>), 37.4 (C<sub>29</sub>), 35.4 (C<sub>38</sub>), 34.9 (C<sub>11</sub>, C<sub>23</sub>), 33.3 (C<sub>35</sub>), 32.9 (C<sub>41</sub>), 32.6 (C<sub>37</sub>), 31.2 (C<sub>42</sub>), 29.6 (C<sub>13</sub>), 26.4 (C<sub>3</sub>), 26.2 (C<sub>12</sub>), 24.4 (C<sub>5</sub>), 21.6 (C<sub>45</sub>), 20.4 (C<sub>4</sub>), 15.6 (C<sub>43</sub>), 15.5 (C<sub>48</sub>), 14.7 (C<sub>49</sub>), 13.7 (C<sub>46</sub>), 13.2 (C<sub>47</sub>), 11.9 (C<sub>44</sub>).

**C16-benzyl carbamyl rapamycin (7a):** Flash chromatography gave 99 mg 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 C<sub>57</sub>H<sub>82</sub>O<sub>14</sub>N<sub>2</sub> [M]: calcd 1033.5995, found 1033.5992.

**(R)-benzyl carbamyl rapamycin:** <sup>1</sup>H NMR (600MHz, DMSO) δ: <sup>1</sup>H NMR (600MHz, DMSO) δ: 7.45 (d, <sup>3</sup>J = 7.7 Hz, 1H, H<sub>52</sub>), 7.38 (m, 2H, H<sub>57</sub>), 7.35 (m, 1H, H<sub>58</sub>), 7.08 (m, 2H, H<sub>56</sub>), 6.37 (dd, <sup>3</sup>J = 10.8 Hz, 12.4 Hz, 1H, H<sub>19</sub>), 6.34 (s, 1H, OH<sub>10</sub>), 6.17 (dd, <sup>3</sup>J = 10.8 Hz, 12.4 Hz, 1H, H<sub>20</sub>), 6.13 (dd, <sup>3</sup>J = 10.8 Hz, 12.0 Hz, 1H, H<sub>21</sub>), 6.02 (d, <sup>3</sup>J = 10.8 Hz, 1H, H<sub>18</sub>), 5.50 (dd, <sup>3</sup>J = 8.8 Hz, 12.0 Hz, 1H, H<sub>22</sub>), 5.23 (m, 1H, OH<sub>28</sub>), 5.21 (m, 1H, H<sub>30</sub>), 5.08 (m, 1H, H<sub>34</sub>), 5.01 (m, 2H, H<sub>54</sub>), 4.95 (d, <sup>3</sup>J = 5.3 Hz, 1H, H<sub>2</sub>), 4.59 (d, <sup>3</sup>J = 4.65 Hz, 1H, OH<sub>40</sub>), 4.07 (m, 1H, H<sub>28</sub>), 3.95 (m, 1H, H<sub>16</sub>), 3.88 (m, 1H, H<sub>27</sub>), 3.85 (m, 1H, H<sub>14</sub>), 3.57 (m, 1H, H<sub>6</sub>), 3.30 (m, 3H, H<sub>51</sub>), 3.27 (m, 1H, H<sub>31</sub>), 3.18 (m, 3H, H<sub>50</sub>), 3.15 (m, 1H, H<sub>40</sub>), 3.03 (m, 1H, H<sub>6</sub>), 2.82 (m, 1H, H<sub>39</sub>), 2.79 (m, 1H, H<sub>33</sub>), 2.50 (m, 2H, H<sub>33</sub>, H<sub>25</sub>), 2.42 (m, 1H, H<sub>23</sub>), 2.15 (m, 1H, H<sub>3</sub>), 2.02 (m, 1H, H<sub>11</sub>), 1.90 (m, 1H, H<sub>38</sub>), 1.76 (m, 1H, H<sub>41</sub>), 1.73 (m, 3H, H<sub>47</sub>), 1.72 (m, 1H, H<sub>15</sub>), 1.69 (m, 1H, H<sub>35</sub>), 1.68 (m, 1H, H<sub>4</sub>), 1.65 (m, 3H, H<sub>44</sub>), 1.59 (m, 1H, H<sub>3</sub>), 1.58 (m, 1H, H<sub>12</sub>), 1.56 (m, 2H, H<sub>13</sub>, H<sub>42</sub>), 1.52 (m, 1H, H<sub>12</sub>), 1.51 (m, 1H, H<sub>5</sub>), 1.40 (m, 2H, H<sub>15</sub>, H<sub>24</sub>), 1.39 (m, 1H, H<sub>5</sub>), 1.38 (m, 1H, H<sub>4</sub>), 1.29 (m, 1H, H<sub>13</sub>), 1.27 (m, 1H, H<sub>41</sub>), 1.26 (m, 1H, H<sub>37</sub>), 1.09 (m, 1H, H<sub>36</sub>), 1.03 (m, 1H, H<sub>24</sub>), 0.99 (m, 3H, H<sub>45</sub>), 0.95 (m, 1H, H<sub>36</sub>), 0.90 (m, 3H, H<sub>48</sub>), 0.86 (m, 3H, H<sub>46</sub>), 0.85 (m, 1H, H<sub>42</sub>), 0.79 (m, 3H, H<sub>49</sub>), 0.73 (m, 3H, H<sub>43</sub>), 0.57 (m, 1H, H<sub>38</sub>). <sup>13</sup>C NMR (400 MHz, DMSO) δ: 210.2 (C<sub>26</sub>), 207.9 (C<sub>32</sub>), 199.1 (C<sub>9</sub>), 169.4 (C<sub>1</sub>), 167.0 (C<sub>8</sub>), 156.3 (C<sub>55</sub>), 155.7 (C<sub>53</sub>), 137.2 (C<sub>17</sub>), 138.3 (C<sub>22</sub>), 137.2 (C<sub>29</sub>), 131.5 (C<sub>20</sub>), 130.4 (C<sub>21</sub>), 128.3 (C<sub>57</sub>), 127.7 (C<sub>58</sub>), 127.6 (C<sub>19</sub>), 127.4 (C<sub>56</sub>), 124.4 (C<sub>30</sub>), 123.3 (C<sub>18</sub>), 99.0 (C<sub>10</sub>), 85.9 (C<sub>27</sub>), 83.8 (C<sub>39</sub>), 75.8 (C<sub>28</sub>), 73.4 (C<sub>34</sub>), 73.2 (C<sub>40</sub>), 67.0 (C<sub>14</sub>), 65.2 (C<sub>54</sub>), 56.9 (C<sub>50</sub>), 56.8 (C<sub>51</sub>), 53.7 (C<sub>16</sub>), 51.3 (C<sub>2</sub>), 44.9 (C<sub>31</sub>), 43.8 (C<sub>6</sub>), 40.4 (C<sub>33</sub>), 39.9 (C<sub>25</sub>), 39.6 (C<sub>24</sub>), 39.5 (C<sub>15</sub>), 38.4 (C<sub>36</sub>), 35.4 (C<sub>38</sub>), 34.7 (C<sub>11</sub>), 34.4 (C<sub>23</sub>), 33.5 (C<sub>35</sub>), 33.0 (C<sub>41</sub>), 32.6 (C<sub>37</sub>), 31.8 (C<sub>13</sub>), 31.2 (C<sub>42</sub>), 26.4 (C<sub>3</sub>, C<sub>12</sub>), 24.5 (C<sub>5</sub>), 21.5 (C<sub>45</sub>), 20.5 (C<sub>4</sub>), 16.0 (C<sub>48</sub>), 15.6 (C<sub>43</sub>), 15.0 (C<sub>49</sub>), 14.1 (C<sub>46</sub>), 14.0 (C<sub>47</sub>), 12.0 (C<sub>44</sub>).

**(S)-benzyl carbamyl rapamycin:** <sup>1</sup>H NMR (600MHz, DMSO) δ: 7.58 (d, <sup>3</sup>J = 9.0 Hz, H<sub>52</sub>, 1H), 7.34 (m, 2H, H<sub>57</sub>), 7.32 (m, 1H, H<sub>58</sub>) 7.07 (m, 2H, H<sub>56</sub>), 6.45 (s, 1H, OH<sub>10</sub>), 6.33 (dd, <sup>3</sup>J = 10.8 Hz, 12.4 Hz, 1H, H<sub>19</sub>), 6.19 (dd, <sup>3</sup>J = 10.2 Hz, 12.4 Hz, 1H, H<sub>20</sub>), 6.12 (dd, <sup>3</sup>J = 10.2 Hz, 12.8 Hz, 1H, H<sub>21</sub>), 6.01 (d, <sup>3</sup>J = 10.8 Hz, 1H, H<sub>18</sub>), 5.49 (dd, <sup>3</sup>J = 8.8 Hz, 12.8 Hz, 1H, H<sub>22</sub>), 5.28 (m, 1H, OH<sub>28</sub>), 5.10 (m, 1H, H<sub>30</sub>), 5.03 (m, 1H, H<sub>34</sub>), 4.98 (m, 1H, H<sub>34</sub>), 4.95 (d, <sup>3</sup>J = 4.9 Hz, 1H, H<sub>2</sub>), 4.59 (d, <sup>3</sup>J = 4.38 Hz, 1H, OH<sub>40</sub>), 4.11 (m, 1H, H<sub>16</sub>), 4.01

(m, 1H, H<sub>28</sub>), 3.87 (m, 1H, H<sub>27</sub>), 3.86 (m, 1H, H<sub>14</sub>), 3.42 (m, 1H, H<sub>6</sub>), 3.32 (m, 4H, H<sub>31</sub>, H<sub>51</sub>), 3.21 (m, 1H, H<sub>6</sub>), 3.18 (m, 1H, H<sub>40</sub>), 3.16 (m, 3H, H<sub>50</sub>), 2.83 (m, 1H, H<sub>39</sub>), 2.76 (m, 1H, H<sub>33</sub>), 2.50 (m, 2H, H<sub>33</sub>, H<sub>25</sub>), 2.23 (m, 1H, H<sub>23</sub>), 2.08 (m, 1H, H<sub>3</sub>), 2.02 (m, 1H, H<sub>11</sub>), 1.92 (m, 1H, H<sub>38</sub>), 1.88 (m, 1H, H<sub>15</sub>), 1.76 (m, 1H, H<sub>41</sub>), 1.74 (m, 3H, H<sub>47</sub>), 1.69 (m, 1H, H<sub>35</sub>), 1.68 (m, 3H, H<sub>44</sub>), 1.65 (m, 1H, H<sub>4</sub>), 1.62 (m, 1H, H<sub>3</sub>), 1.54 (m, 3H, H<sub>5</sub>, H<sub>12</sub>, H<sub>13</sub>), 1.53 (m, 1H, H<sub>42</sub>), 1.45 (m, 1H, H<sub>12</sub>), 1.39 (m, 1H, H<sub>4</sub>), 1.37 (m, 1H, H<sub>15</sub>), 1.36 (m, 1H, H<sub>24</sub>), 1.31 (m, 1H, H<sub>5</sub>), 1.30 (m, 1H, H<sub>37</sub>), 1.26 (m, 2H, H<sub>13</sub>, H<sub>41</sub>), 1.12 (m, 1H, H<sub>36</sub>), 1.08 (m, 1H, H<sub>24</sub>), 0.99 (m, 3H, H<sub>45</sub>), 0.96 (m, 1H, H<sub>36</sub>), 0.89 (m, 3H, H<sub>48</sub>), 0.86 (m, 1H, H<sub>42</sub>), 0.85 (m, 3H, H<sub>46</sub>), 0.79 (m, 3H, H<sub>49</sub>), 0.75 (m, 3H, H<sub>43</sub>), 0.60 (m, 1H, H<sub>38</sub>). <sup>13</sup>C-NMR: 210.8 (C<sub>26</sub>), 207.7 (C<sub>32</sub>), 198.8 (C<sub>9</sub>), 169.2 (C<sub>1</sub>), 166.9 (C<sub>8</sub>), 156.4 (C<sub>55</sub>), 155.6 (C<sub>53</sub>), 139.1 (C<sub>17</sub>), 138.7 (C<sub>22</sub>), 137.2 (C<sub>29</sub>), 131.8 (C<sub>20</sub>), 130.5 (C<sub>21</sub>), 128.3 (C<sub>57</sub>), 127.7 (C<sub>58</sub>), 127.6 (C<sub>19</sub>), 127.4 (C<sub>56</sub>), 125.1 (C<sub>30</sub>), 125.0 (C<sub>18</sub>), 98.8 (C<sub>10</sub>), 85.9 (C<sub>27</sub>), 83.9 (C<sub>39</sub>), 75.9 (C<sub>28</sub>), 73.8 (C<sub>34</sub>), 73.3 (C<sub>40</sub>), 67.0 (C<sub>14</sub>), 65.2 (C<sub>54</sub>), 57.2 (C<sub>50</sub>), 56.8 (C<sub>51</sub>), 53.4 (C<sub>16</sub>), 50.8 (C<sub>2</sub>), 45.4 (C<sub>31</sub>), 43.6 (C<sub>6</sub>), 39.7 (C<sub>24</sub>), 39.6 (C<sub>33</sub>), 39.3 (C<sub>25</sub>), 38.7 (C<sub>15</sub>), 38.3 (C<sub>36</sub>), 35.5 (C<sub>38</sub>), 34.9 (C<sub>11</sub>, C<sub>23</sub>), 33.4 (C<sub>35</sub>), 33.0 (C<sub>41</sub>), 32.8 (C<sub>37</sub>), 31.3 (C<sub>13</sub>, C<sub>42</sub>), 26.4 (C<sub>3</sub>, C<sub>12</sub>), 24.5 (C<sub>5</sub>), 21.7 (C<sub>45</sub>), 20.5 (C<sub>4</sub>), 15.7 (C<sub>43</sub>), 15.5 (C<sub>48</sub>), 14.9 (C<sub>49</sub>), 13.7 (C<sub>46</sub>), 13.4 (C<sub>47</sub>), 12.1 (C<sub>44</sub>).

**C16-butyl carbamyl rapamycin (7b):** Flash chromatography gave 79 mg 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 C<sub>55</sub>H<sub>86</sub>O<sub>14</sub>N<sub>2</sub> [M-Na<sup>+</sup>]: calcd. 1021.5971, found 1021.5945.

**(R)-butyl carbamyl rapamycin:** <sup>1</sup>H NMR (600MHz, DMSO) δ: 7.43 (d, <sup>3</sup>J = 7.8 Hz, 1H, H<sub>52</sub>), 6.36 (dd, <sup>3</sup>J = 10.7 Hz, 12.3 Hz, 1H, H<sub>19</sub>), 6.34 (s, 1H, OH<sub>10</sub>), 6.17 (dd, <sup>3</sup>J = 11.0 Hz, 12.3 Hz, 1H, H<sub>20</sub>), 6.13 (dd, <sup>3</sup>J = 11.0 Hz, 12.3 Hz, 1H, H<sub>21</sub>), 6.01 (d, <sup>3</sup>J = 10.7 Hz, 1H, H<sub>18</sub>), 5.50 (dd, <sup>3</sup>J = 9.8 Hz, 12.3 Hz, 1H, H<sub>22</sub>), 5.23 (d, <sup>3</sup>J = 4.1 Hz, 1H, OH<sub>28</sub>), 5.22 (d, <sup>3</sup>J = 11.1 Hz, 1H, H<sub>30</sub>), 5.09 (m, 1H, H<sub>34</sub>), 4.95 (d, <sup>3</sup>J = 5.1 Hz, 1H, H<sub>2</sub>), 4.59 (d, <sup>3</sup>J = 4.53 Hz, 1H, OH<sub>40</sub>), 4.07 (m, 1H, H<sub>28</sub>), 4.03 (m, 1H, H<sub>27</sub>), 3.95 (m, 1H, H<sub>16</sub>), 3.93 (m, 2H, H<sub>54</sub>), 3.86 (m, 1H, H<sub>14</sub>), 3.57 (m, 1H, H<sub>6</sub>), 3.30 (m, 3H, H<sub>51</sub>), 3.26 (m, 1H, H<sub>31</sub>), 3.18 (m, 3H, H<sub>50</sub>), 3.16 (m, 1H, H<sub>40</sub>), 3.03 (m, 1H, H<sub>6</sub>), 2.81 (m, 1H, H<sub>39</sub>), 2.78 (m, 1H, H<sub>33</sub>), 2.50 (m, 1H, H<sub>33</sub>), 2.35 (m, 1H, H<sub>25</sub>), 2.25 (m, 1H, H<sub>23</sub>), 2.16 (m, 1H, H<sub>3</sub>), 2.02 (m, 1H, H<sub>11</sub>), 1.91 (m, 1H, H<sub>38</sub>), 1.76 (m, 1H, H<sub>41</sub>), 1.73 (m, 3H, H<sub>47</sub>), 1.72 (m, 3H, H<sub>15</sub>, H<sub>55</sub>), 1.69 (m, 1H, H<sub>35</sub>), 1.68 (m, 1H, H<sub>4</sub>), 1.67 (m, 3H, H<sub>44</sub>), 1.61 (m, 1H, H<sub>3</sub>), 1.57 (m, 1H, H<sub>13</sub>), 1.53 (m, 3H, H<sub>5</sub>, H<sub>12</sub>), 1.52 (m, 3H, H<sub>42</sub>, H<sub>55</sub>), 1.45 (m, 1H, H<sub>15</sub>), 1.39 (m, 1H, H<sub>24</sub>), 1.38 (m, 1H, H<sub>4</sub>), 1.37 (m, 1H, H<sub>5</sub>), 1.29 (m, 2H, H<sub>56</sub>), 1.26 (m, 1H, H<sub>37</sub>), 1.25 (m, 1H, H<sub>13</sub>), 1.17 (m, 1H, H<sub>41</sub>), 1.09 (m, 1H, H<sub>36</sub>), 1.03 (m, 1H, H<sub>24</sub>), 0.99 (m, 3H, H<sub>45</sub>), 0.95 (m, 1H, H<sub>36</sub>), 0.90 (m, 3H, H<sub>48</sub>), 0.85 (m, 4H, H<sub>42</sub>, H<sub>57</sub>), 0.81 (m, 3H, H<sub>46</sub>), 0.80 (m, 3H, H<sub>49</sub>), 0.74 (m, 3H, H<sub>43</sub>), 0.57 (m, 1H, H<sub>38</sub>). <sup>13</sup>C-NMR: 209.9 (C<sub>26</sub>), 207.8 (C<sub>32</sub>), 198.8 (C<sub>9</sub>), 169.3 (C<sub>1</sub>), 167.0 (C<sub>8</sub>), 155.6 (C<sub>53</sub>), 140.4 (C<sub>17</sub>), 138.0 (C<sub>22</sub>), 136.7 (C<sub>29</sub>), 130.2 (C<sub>21</sub>), 127.4 (C<sub>19</sub>), 124.1 (C<sub>30</sub>), 123.0 (C<sub>18</sub>), 99.0 (C<sub>10</sub>), 85.2 (C<sub>27</sub>), 83.6 (C<sub>39</sub>), 75.5 (C<sub>28</sub>), 73.2 (C<sub>34</sub>), 73.0 (C<sub>40</sub>), 66.8 (C<sub>14</sub>), 65.4 (C<sub>54</sub>), 56.6 (C<sub>50</sub>), 56.5 (C<sub>51</sub>), 53.5 (C<sub>16</sub>), 50.8 (C<sub>2</sub>), 44.6 (C<sub>31</sub>), 43.6 (C<sub>6</sub>), 39.8 (C<sub>25</sub>), 39.6 (C<sub>33</sub>), 39.3 (C<sub>15</sub>, C<sub>24</sub>), 38.1 (C<sub>36</sub>), 35.0 (C<sub>38</sub>), 34.4 (C<sub>11</sub>), 34.2 (C<sub>23</sub>), 33.2 (C<sub>35</sub>), 32.8 (C<sub>41</sub>), 32.7 (C<sub>55</sub>), 32.4 (C<sub>37</sub>), 31.0 (C<sub>13</sub>, C<sub>42</sub>), 26.2 (C<sub>3</sub>, C<sub>12</sub>), 24.2 (C<sub>5</sub>), 22.3 (C<sub>56</sub>), 21.3 (C<sub>45</sub>), 20.3 (C<sub>4</sub>), 15.8 (C<sub>48</sub>), 15.4 (C<sub>43</sub>), 15.0 (C<sub>44</sub>), 14.8 (C<sub>49</sub>), 14.7 (C<sub>57</sub>), 13.7 (C<sub>47</sub>), 12.6 (C<sub>46</sub>).

**(S)-butyl carbamyl rapamycin:** <sup>1</sup>H NMR (600MHz, DMSO) δ: 7.25 (d, <sup>3</sup>J = 8.8 Hz, 1H, H<sub>52</sub>), 6.44 (s, 1H, OH<sub>10</sub>), 6.33 (dd, <sup>3</sup>J = 10.9 Hz, 12.4 Hz, 1H, H<sub>19</sub>), 6.19 (dd, <sup>3</sup>J = 10.4 Hz, 12.4 Hz, 1H, H<sub>20</sub>), 6.12 (dd, <sup>3</sup>J = 10.4 Hz, 11.3 Hz, 1H, H<sub>21</sub>), 6.02 (d, <sup>3</sup>J = 11.3 Hz, 1H, H<sub>18</sub>), 5.48 (dd, <sup>3</sup>J = 9.5 Hz, 12.4 Hz, 1H, H<sub>22</sub>), 5.10 (s, 1H, OH<sub>28</sub>), 5.09 (m, 1H, H<sub>30</sub>), 4.99 (m, 1H, H<sub>34</sub>), 4.96 (m, 1H, H<sub>2</sub>), 4.59 (d, <sup>3</sup>J = 4.37 Hz, 1H, OH<sub>40</sub>), 4.10 (m, 1H, H<sub>16</sub>), 4.01 (d, <sup>3</sup>J = 5.2 Hz, 1H, H<sub>28</sub>), 3.93 (m, 2H, H<sub>54</sub>), 3.88 (d, <sup>3</sup>J = 5.2 Hz, 1H, H<sub>27</sub>), 3.83 (m, 1H, H<sub>14</sub>), 3.44 (m, 1H, H<sub>6</sub>), 3.32 (m, 3H, H<sub>51</sub>), 3.31 (m, 1H, H<sub>31</sub>), 3.23 (m, 1H, H<sub>6</sub>), 3.17 (m, 1H, H<sub>40</sub>), 3.14 (s, 3H, H<sub>50</sub>), 2.81 (m, 1H, H<sub>39</sub>), 2.75 (m, 1H, H<sub>33</sub>), 2.49 (m, 1H, H<sub>25</sub>), 2.41 (m, 1H, H<sub>33</sub>), 2.22 (m, 1H, H<sub>23</sub>), 2.09 (m, 1H, H<sub>3</sub>), 2.02 (m, 1H, H<sub>11</sub>), 1.91 (m, 1H, H<sub>38</sub>), 1.88 (m, 1H, H<sub>15</sub>), 1.79 (m, 1H, H<sub>13</sub>), 1.76 (m, 2H, H<sub>4</sub>, H<sub>41</sub>), 1.74

(m, 4H, H<sub>35</sub>, H<sub>47</sub>), 1.73 (m, 3H, H<sub>44</sub>), 1.55 (m, 1H, H<sub>5</sub>), 1.53 (m, 2H, H<sub>42</sub>, H<sub>3</sub>), 1.52 (m, 2H, H<sub>55</sub>), 1.47 (m, 1H, H<sub>12</sub>), 1.44 (m, 1H, H<sub>12</sub>), 1.41 (m, 1H, H<sub>24</sub>), 1.34 (m, 1H, H<sub>15</sub>), 1.30 (m, 2H, H<sub>4</sub>, H<sub>5</sub>), 1.28 (m, 1H, H<sub>37</sub>), 1.26 (m, 2H, H<sub>56</sub>), 1.24 (m, 1H, H<sub>13</sub>), 1.16 (m, 1H, H<sub>41</sub>), 1.09 (m, 1H, H<sub>36</sub>), 1.06 (m, 1H, H<sub>24</sub>), 0.99 (m, 3H, H<sub>45</sub>), 0.96 (m, 1H, H<sub>36</sub>), 0.89 (m, 3H, H<sub>48</sub>), 0.86 (m, 3H, H<sub>57</sub>), 0.85 (m, 1H, H<sub>42</sub>), 0.84 (m, 3H, H<sub>46</sub>), 0.79 (m, 3H, H<sub>49</sub>), 0.75 (m, 3H, H<sub>43</sub>), 0.60 (m, 1H, H<sub>38</sub>). <sup>13</sup>C-NMR: 210.8 (C<sub>26</sub>), 207.9 (C<sub>32</sub>), 198.8 (C<sub>9</sub>), 169.1 (C<sub>1</sub>), 166.4 (C<sub>8</sub>), 155.5 (C<sub>53</sub>), 138.5 (C<sub>22</sub>), 137.5 (C<sub>17</sub>), 131.7 (C<sub>20</sub>), 130.2 (C<sub>21</sub>), 127.1 (C<sub>19</sub>), 125.0 (C<sub>30</sub>), 124.9 (C<sub>18</sub>), 99.0 (C<sub>10</sub>), 85.5 (C<sub>27</sub>), 83.6 (C<sub>39</sub>), 75.5 (C<sub>28</sub>), 73.5 (C<sub>34</sub>), 72.9 (C<sub>40</sub>), 67.0 (C<sub>14</sub>), 65.4 (C<sub>54</sub>), 57.3 (C<sub>50</sub>), 56.8 (C<sub>16</sub>), 56.5 (C<sub>51</sub>), 50.4 (C<sub>2</sub>), 45.0 (C<sub>31</sub>), 43.2 (C<sub>6</sub>), 39.5 (C<sub>25</sub>, C<sub>33</sub>), 39.3 (C<sub>24</sub>), 38.2 (C<sub>15</sub>, C<sub>36</sub>), 35.1 (C<sub>38</sub>), 34.6 (C<sub>11</sub>, C<sub>23</sub>), 32.7 (C<sub>55</sub>), 32.6 (C<sub>35</sub>, C<sub>41</sub>), 31.2 (C<sub>42</sub>), 30.9 (C<sub>37</sub>), 29.6 (C<sub>5</sub>), 28.7 (C<sub>13</sub>), 26.0 (C<sub>3</sub>, C<sub>55</sub>), 25.9 (C<sub>12</sub>), 22.3 (C<sub>56</sub>), 21.4 (C<sub>45</sub>), 20.3 (C<sub>4</sub>), 15.4 (C<sub>43</sub>), 15.3 (C<sub>48</sub>), 14.7 (C<sub>57</sub>), 14.5 (C<sub>49</sub>), 13.5 (C<sub>46</sub>), 13.0 (C<sub>44</sub>), 12.9 (C<sub>47</sub>).

**C16-butyl carbamyl rapamycin (7c):** Flash chromatography gave 85 mg 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-butyl carbamyl rapamycin:** <sup>1</sup>H NMR (600MHz, DMSO) δ: 6.98 (d, <sup>3</sup>J = 9.0 Hz, 1H, H<sub>52</sub>), 6.42 (s, 1H, OH<sub>10</sub>), 6.33 (dd, <sup>3</sup>J = 11.1 Hz, 12.4 Hz, 1H, H<sub>19</sub>), 6.18 (dd, <sup>3</sup>J = 11.9 Hz, 12.4 Hz, 1H, H<sub>20</sub>), 6.11 (dd, <sup>3</sup>J = 11.9 Hz, 12.4 Hz, 1H, H<sub>21</sub>), 5.97 (d, <sup>3</sup>J = 11.1 Hz, 1H, H<sub>18</sub>), 5.49 (dd, <sup>3</sup>J = 9.1 Hz, 12.4 Hz, 1H, H<sub>22</sub>), 5.11 (m, 2H, OH<sub>28</sub>, H<sub>30</sub>), 5.08 (m, 1H, H<sub>34</sub>), 4.95 (d, <sup>3</sup>J = 4.9 Hz, 1H, H<sub>2</sub>), 4.52 (m, 1H, OH<sub>40</sub>), 4.08 (m, 1H, H<sub>16</sub>), 4.07 (m, 1H, H<sub>28</sub>), 4.05 (m, 1H, H<sub>27</sub>), 3.83 (m, 1H, H<sub>14</sub>), 3.55 (m, 1H, H<sub>6</sub>), 3.31 (m, 3H, H<sub>51</sub>), 3.26 (m, 1H, H<sub>31</sub>), 3.18 (m, 3H, H<sub>50</sub>), 3.17 (m, 1H, H<sub>40</sub>), 3.00 (m, 1H, H<sub>6</sub>), 2.81 (m, 1H, H<sub>39</sub>), 2.77 (m, 1H, H<sub>33</sub>), 2.50 (m, 1H, H<sub>25</sub>), 2.43 (m, 1H, H<sub>33</sub>), 2.23 (m, 1H, H<sub>23</sub>), 2.11 (m, 1H, H<sub>3</sub>), 1.92 (m, 1H, H<sub>11</sub>), 1.91 (m, 1H, H<sub>38</sub>), 1.75 (m, 1H, H<sub>41</sub>), 1.74 (m, 3H, H<sub>47</sub>), 1.70 (m, 1H, H<sub>15</sub>), 1.69 (m, 1H, H<sub>35</sub>), 1.67 (m, 1H, H<sub>4</sub>), 1.66 (m, 3H, H<sub>44</sub>), 1.60 (m, 1H, H<sub>3</sub>), 1.57 (m, 1H, H<sub>42</sub>), 1.52 (m, 1H, H<sub>12</sub>), 1.50 (m, 2H, H<sub>13</sub>, H<sub>5</sub>), 1.42 (m, 1H, H<sub>15</sub>), 1.40 (m, 2H, H<sub>24</sub>, H<sub>12</sub>), 1.37 (m, 5H, H<sub>5</sub>, H<sub>4</sub>, H<sub>55</sub>), 1.26 (m, 2H, H<sub>37</sub>, H<sub>13</sub>), 1.25 (m, 1H, H<sub>41</sub>), 1.09 (m, 1H, H<sub>36</sub>), 1.03 (m, 1H, H<sub>24</sub>), 0.98 (m, 3H, H<sub>45</sub>), 0.95 (m, 1H, H<sub>36</sub>), 0.90 (m, 3H, H<sub>48</sub>), 0.86 (m, 4H, H<sub>46</sub>, H<sub>42</sub>), 0.79 (m, 3H, H<sub>49</sub>), 0.73 (m, 3H, H<sub>43</sub>), 0.59 (m, 1H, H<sub>38</sub>). <sup>13</sup>C-NMR: 210.1 (C<sub>26</sub>), 207.7 (C<sub>32</sub>), 198.8 (C<sub>9</sub>), 169.2 (C<sub>1</sub>), 166.9 (C<sub>8</sub>), 155.0 (C<sub>53</sub>), 139.1 (C<sub>17</sub>), 138.7 (C<sub>22</sub>), 137.2 (C<sub>29</sub>), 131.8 (C<sub>20</sub>), 130.5 (C<sub>21</sub>), 127.6 (C<sub>19</sub>), 125.1 (C<sub>30</sub>), 125.0 (C<sub>18</sub>), 98.8 (C<sub>10</sub>), 85.4 (C<sub>27</sub>), 83.9 (C<sub>39</sub>), 77.6 (C<sub>54</sub>), 75.9 (C<sub>28</sub>), 73.8 (C<sub>34</sub>), 73.3 (C<sub>40</sub>), 67.0 (C<sub>14</sub>), 57.2 (C<sub>50</sub>), 56.8 (C<sub>51</sub>), 53.4 (C<sub>16</sub>), 50.8 (C<sub>2</sub>), 45.4 (C<sub>31</sub>), 43.6 (C<sub>6</sub>), 39.7 (C<sub>24</sub>), 39.6 (C<sub>33</sub>), 39.3 (C<sub>25</sub>), 38.7 (C<sub>15</sub>), 38.4 (C<sub>36</sub>), 35.5 (C<sub>38</sub>), 34.9 (C<sub>11</sub>, C<sub>23</sub>), 33.4 (C<sub>35</sub>), 33.0 (C<sub>41</sub>), 32.8 (C<sub>37</sub>), 31.3 (C<sub>13</sub>, C<sub>42</sub>), 28.3 (C<sub>55</sub>), 26.4 (C<sub>3</sub>, C<sub>12</sub>), 24.5 (C<sub>5</sub>), 21.7 (C<sub>45</sub>), 20.5 (C<sub>4</sub>), 15.7 (C<sub>43</sub>), 15.5 (C<sub>48</sub>), 14.9 (C<sub>49</sub>), 13.7 (C<sub>46</sub>), 13.4 (C<sub>47</sub>), 12.1 (C<sub>44</sub>).

**(S)-C16-butyl carbamyl rapamycin:** <sup>1</sup>H NMR (600MHz, DMSO) δ: 7.07 (d, <sup>3</sup>J = 9.0 Hz, 1H, H<sub>52</sub>), 6.39 (s, 1H, OH<sub>10</sub>), 6.36 (m, 1H, H<sub>19</sub>), 6.15 (m, 1H, H<sub>20</sub>), 6.12 (m, 12.4 Hz, 1H, H<sub>21</sub>), 5.99 (m, 1H, H<sub>18</sub>), 5.51 (dd, <sup>3</sup>J = 9.1 Hz, 12.4 Hz, 1H, H<sub>22</sub>), 5.22 (s, 1H, OH<sub>28</sub>), 5.20 (m, 1H, H<sub>30</sub>), 4.98 (m, 1H, H<sub>34</sub>), 4.95 (d, <sup>3</sup>J = 5.60 Hz, 1H, H<sub>2</sub>), 4.04 (m, 1H, H<sub>16</sub>), 4.02 (m, 1H, H<sub>27</sub>), 3.89 (m, 1H, H<sub>28</sub>), 3.84 (m, 1H, H<sub>14</sub>), 3.44 (m, 1H, H<sub>6</sub>), 3.32 (m, 3H, H<sub>51</sub>), 3.30 (m, 1H, H<sub>31</sub>), 3.20 (m, 1H, H<sub>6</sub>), 3.17 (m, 1H, H<sub>40</sub>), 3.15 (m, 3H, H<sub>50</sub>), 2.83 (m, 1H, H<sub>39</sub>), 2.75 (m, 1H, H<sub>33</sub>), 2.50 (m, 1H, H<sub>25</sub>), 2.43 (m, 1H, H<sub>33</sub>), 2.23 (m, 1H, H<sub>23</sub>), 2.11 (m, 1H, H<sub>3</sub>), 2.01 (m, 1H, H<sub>11</sub>), 1.91 (m, 1H, H<sub>38</sub>), 1.83 (m, 1H, H<sub>15</sub>), 1.76 (m, 1H, H<sub>35</sub>), 1.74 (m, 1H, H<sub>41</sub>), 1.73 (m, 3H, H<sub>47</sub>), 1.67 (m, 1H, H<sub>4</sub>), 1.66 (m, 3H, H<sub>44</sub>), 1.55 (m, 2H, H<sub>5</sub>, H<sub>13</sub>), 1.54 (m, 1H, H<sub>12</sub>), 1.53 (m, 1H, H<sub>42</sub>), 1.52 (m, 2H, H<sub>12</sub>), 1.51 (m, 1H, H<sub>3</sub>), 1.40 (m, 1H, H<sub>24</sub>), 1.39 (m, 1H, H<sub>4</sub>), 1.37 (m, 9H, H<sub>55</sub>), 1.32 (m, 2H, H<sub>5</sub>, H<sub>15</sub>), 1.31 (m, 1H, H<sub>13</sub>), 1.28 (m, 1H, H<sub>37</sub>), 1.18 (m, 1H, H<sub>41</sub>), 1.17 (m, 1H, H<sub>36</sub>), 1.06 (m, 1H, H<sub>24</sub>), 0.97 (m, 3H, H<sub>45</sub>), 0.89 (m, 3H, H<sub>48</sub>), 0.86 (m, 1H, H<sub>36</sub>), 0.85 (m, 1H, H<sub>42</sub>), 0.83 (m, 3H, H<sub>46</sub>), 0.78 (m, 3H, H<sub>49</sub>), 0.74 (m, 3H, H<sub>43</sub>), 0.60 (m, 1H, H<sub>38</sub>). <sup>13</sup>C-NMR: 209.9 (C<sub>26</sub>), 207.9 (C<sub>32</sub>), 198.8

(C<sub>9</sub>), 169.5 (C<sub>1</sub>), 167.0 (C<sub>8</sub>), 155.2 (C<sub>53</sub>), 138.2 (C<sub>22</sub>), 137.8 (C<sub>29</sub>), 136.8 (C<sub>17</sub>), 131.4 (C<sub>20</sub>), 130.5 (C<sub>21</sub>), 127.8 (C<sub>19</sub>), 124.4 (C<sub>30</sub>), 123.1 (C<sub>18</sub>), 99.0 (C<sub>10</sub>), 85.4 (C<sub>27</sub>), 83.9 (C<sub>39</sub>), 77.7 (C<sub>54</sub>), 75.8 (C<sub>28</sub>), 73.4 (C<sub>34</sub>), 73.2 (C<sub>40</sub>), 67.3 (C<sub>14</sub>), 56.9 (C<sub>50</sub>), 56.8 (C<sub>51</sub>), 53.7 (C<sub>16</sub>), 51.3 (C<sub>2</sub>), 44.9 (C<sub>31</sub>), 43.7 (C<sub>6</sub>), 40.0 (C<sub>25</sub>), 39.8 (C<sub>15</sub>), 39.6 (C<sub>33</sub>), 39.5 (C<sub>24</sub>), 38.7 (C<sub>36</sub>), 35.5 (C<sub>38</sub>), 35.4 (C<sub>11</sub>), 34.5 (C<sub>23</sub>), 33.5 (C<sub>35</sub>), 33.0 (C<sub>41</sub>), 31.8 (C<sub>37</sub>), 31.3 (C<sub>42</sub>), 31.2 (C<sub>13</sub>), 28.3 (C<sub>55</sub>), 26.5 (C<sub>3</sub>, C<sub>12</sub>), 24.5 (C<sub>5</sub>), 21.6 (C<sub>45</sub>), 20.6 (C<sub>4</sub>), 16.0 (C<sub>48</sub>), 15.6 (C<sub>43</sub>), 15.1 (C<sub>49</sub>), 14.1 (C<sub>47</sub>), 14.0 (C<sub>46</sub>), 12.1 (C<sub>44</sub>).

### Cell Culture and Transfection

HeLa and HEK293 were cultured in complete Dulbecco'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™ (Polyplus-transfection) using 2.5 µg of total plasmid DNA, according to manufacturer's guidelines. Materials and media for cell biology were from Sigma-Aldrich, where not indicated otherwise.

### Immune-blot Analysis and Cellular heterodimerization

For immuno-blot analysis, 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. The phosphorylation of p70<sup>S6K</sup> was analyzed using primary antibodies against total p70<sup>S6K</sup> (total S6K; New England Biolabs [NEB]) and phosphorylated p70S6K (pS6K, Thr389, NEB) and horseradish peroxidase (HRP)-coupled secondary goat anti-mouse or anti-rabbit antibodies. Protein bands were visualized by enhanced chemiluminescence (ECL; Millipore). Statistical analysis was performed using GraphPad Prism 6.0c (Student's t test; two sided, non-paired). For microscopy, HeLa cells grown on 12 mm cover slips (Menzel) were transfected with expression constructs for FKBP12-GFP-CAAX and mutated FRB<sub>T2098L</sub>-RFP fusion proteins. Cells were exposed to DMSO or rapalogs 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 (Plüss-Stauffer) containing 1% Propyl gallate.

**Supporting Information** (see footnote on the first page of this article): Copies of the <sup>1</sup>H, <sup>13</sup>C, COSY, HMBC and HMQC NMR spectra and HRMS for all final products.

### Acknowledgments

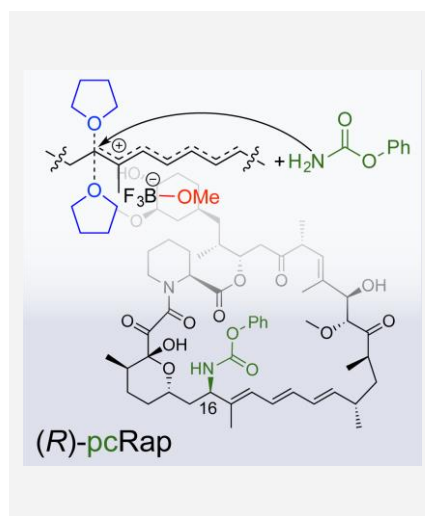
We thank B. Giese for valuable advice and discussions, T. Inoue and K. Johansson for FRB- and FKBP12 domain-containing plasmids, M. Mayor for technical support. This work was supported by Swiss National Science Foundation grants 205320-138302 and 31EM30-126143, the ESF EuroMEMBRANE programme grant FP-018, and the Novartis (Jubilée) Foundation.

- [1] a) R. Zoncu, D. M. Sabatini, A. Efeyan *Nat. Rev. Mol. Cell. Biol.* **2011**, *12*, 21–35.
- [2] G. R. Crabtree, S. L. Schreiber, *Trends Biochem. Sci.*, **1996**, *21*, 418–422.
- [3] M. Putyrski, C. Schultz, *FEBS Lett.*, **2012**, *585*, 2097–2105.
- [4] M. P. Wymann, G. Schneider, *Nat. Rev. Mol. Cell Biol.*, **2008**, *9*, 162–176.
- [5] M. P. Wymann, C. Schultz, *Chembiochem.*, **2012**, *13*, 2022–2035.
- [6] J. I. Luengo, D. S. Yamashita, D. Dunnington, A. K. Beck, L. W. Rozamus, H.-K. Yen, M. J. Bossard, M. A. Levy, A. Hand, T. Newman-Tarr, A. Badger, L. Faucette, R. K. Johnson, K. D'Alessio, T. Porter, A. Y.L. Shu, R. Heys, J. Choi, P. Kongsaree, J. Clardy, D. A. Holt, *Chem. Biol.*, **1995**, *7*, 471–481.

- [7] D. Erhart, M. Zimmermann, O. Jacques, M. B. Wittwer, B. Ernst, E. Constable, M. Zvelebil, F. Beaufils, M. P. Wymann, *Chem. Biol.*, **2013**, *20*, 549–557.
- [8] J. H. Bayle, J. S. Grimley, K. Stankunas, J. E. Gestwicki, T. J. Wandless, G. R. Crabtree, *Chem. Biol.*, **2006**, *13*, 99–107.
- [9] S. D. Liberles, S. T. Diver, D. J. Austin, S. L. Schreiber, *Proc., Natl. Acad. Sci.*, **1997**, *94*, 7825–7830

Received: ((will be filled in by the editorial staff))  
Published online: ((will be filled in by the editorial staff))

Rapalogs are valuable tools to chemically induce dimerization of engineered proteins. A convenient synthesis of C16-carbamyl substituted rapalogs using sequential Lewis Acid activation-decomplexation of the C16-methoxide provides a path to (*R*)-pcRap, a rapamycin derivative that does not interfere with endogenous target of rapamycin (TOR) signaling.



#### Chemical Inducer of Dimerization

Ruben Cal, Mirjam Zimmermann, Kaspar Zimmermann, Vladimir Cmiljanovic, Olivier Jacques, Edwin Constable, Daniel Häussinger, Florent Beaufils\* and Matthias P. Wymann\*

Efficient Synthesis of C16-carbamyl Rapalogs via Lewis Acid Decomplexation

**Keywords:** Carbamylation/Natural products/Lewis Acids/Reaction Mechanisms/ Protein-protein interactions/chemically-induced dimerization/CID/mammalian target of rapamycin/mTOR

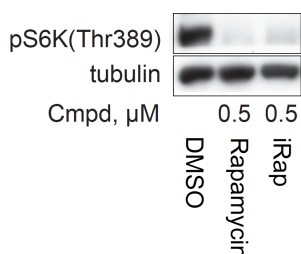
## Comments to pcRap manuscript:

The high cell permeability of rapamycin as well as the high rate constants of rapamycin with FKBP12<sup>[31]</sup> ( $k_a = 5.8 \times 10^6 \text{ M}^{-1} \cdot \text{sec}^{-1}$ ) and FKBP12–rapamycin with the FRB domain of mammalian target of rapamycin<sup>[31]</sup> (mTORC1) ( $k_a = 1.7 \times 10^6 \text{ M}^{-1} \cdot \text{sec}^{-1}$ ) allow rapamycin to induce dimerization between FKBP12 and FRB with excellent kinetics and efficiency.

Speed as well as the efficiency of the HaXS-induced dimerization is lower compared to the rapamycin CID. Even the amount of HaXS-induced dimerization exceeds values required for many applications such as manipulation of surface receptor, which recruits cytosolic signaling molecules<sup>[67]</sup>; some applications require a faster and more efficient dimerization. However, as rapamycin interferes with endogenous signaling pathways, the utility of the rapamycin CID is limited for some applications.

To overcome the effect of the rapamycin-induced mTORC1 inhibition, rapamycin analogs (so-called rapalogs) were developed, in which the interface of rapamycin that interacts with the endogenous FRB domain is synthetically equipped with a bulky substituent in order to abolish binding to the wild type FRB domain. To restore the dimerizing potency of the rapalogs, a mutated FRB domain (Thr2098Leu), which displays a compensatory cavity that accommodates the binding of the bulky substituent is used in combination with these rapalogs. Overall, these rapalogs should be devoid of inhibitory effects towards mTORC1, but still be able to dimerize fusions of FKBP12 and the mutated FRB domain (Thr2098Leu)<sup>[38]</sup>.

The successful use of several rapalogs such as iRap and AP21967 were reported<sup>[39]</sup>. However, we analyzed the activation status of mTORC1, through analyzing the phosphorylation status of an mTORC1 downstream substrate, the ribosomal protein S6 kinase (S6K). Exposure of HEK293 cells with 0.5  $\mu\text{M}$  iRap for 60 min, resulted in a clear downregulation of pS6K1 at Thr389 (as compared to the DMSO control), revealing that iRap interferes with mTORC1 (Fig 1).



**Figure 1.** HEK293 cells were exposed to DMSO, 0.5  $\mu\text{M}$  rapamycin resp. 0.5  $\mu\text{M}$  iRap in complete medium for 60 min at 37°C before cells were lysed and proteins were subjected to SDS-PAGE and immune blotting using antibodies against tubulin and phosphorylated S6K (pS6k1 Thr389). Experiment performed by Dominik Erhart. iRap kindly received from Tom Wandless.

Furthermore, the effect on mTORC1 kinase activity by iRap and another commonly used rapalog AP29167 was determined through the determination of the  $\text{IC}_{50}$  values in HEK293 cells. Surprisingly, the results revealed that both rapalogs inhibit mTORC1 ( $\text{IC}_{50} = 0.1 \text{ nM}$  for rapamycin,  $\text{IC}_{50} = 5 \text{ nM}$  for iRap and  $\text{IC}_{50} = 10 \text{ nM}$  for AP29167)<sup>[40]</sup>, either due to impurities of rapamycin or rapamycin-byproducts, or an undefined spatial orientation of the bulky group at the C16 position.

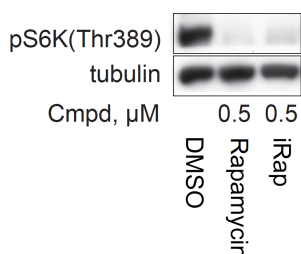
## Comments to pcRap manuscript:

The high cell permeability of rapamycin as well as the high rate constants of rapamycin with FKBP12<sup>[31]</sup> ( $k_a = 5.8 \times 10^6 \text{ M}^{-1} \cdot \text{sec}^{-1}$ ) and FKBP12–rapamycin with the FRB domain of mammalian target of rapamycin<sup>[31]</sup> (mTORC1) ( $k_a = 1.7 \times 10^6 \text{ M}^{-1} \cdot \text{sec}^{-1}$ ) allow rapamycin to induce dimerization between FKBP12 and FRB with excellent kinetics and efficiency.

Speed as well as the efficiency of the HaXS-induced dimerization is lower compared to the rapamycin CID. Even the amount of HaXS-induced dimerization exceeds values required for many applications such as manipulation of surface receptor, which recruits cytosolic signaling molecules<sup>[67]</sup>; some applications require a faster and more efficient dimerization. However, as rapamycin interferes with endogenous signaling pathways, the utility of the rapamycin CID is limited for some applications.

To overcome the effect of the rapamycin-induced mTORC1 inhibition, rapamycin analogs (so-called rapalogs) were developed, in which the interface of rapamycin that interacts with the endogenous FRB domain is synthetically equipped with a bulky substituent in order to abolish binding to the wild type FRB domain. To restore the dimerizing potency of the rapalogs, a mutated FRB domain (Thr2098Leu), which displays a compensatory cavity that accommodates the binding of the bulky substituent is used in combination with these rapalogs. Overall, these rapalogs should be devoid of inhibitory effects towards mTORC1, but still be able to dimerize fusions of FKBP12 and the mutated FRB domain (Thr2098Leu)<sup>[38]</sup>.

The successful use of several rapalogs such as iRap and AP21967 were reported<sup>[39]</sup>. However, we analyzed the activation status of mTORC1, through analyzing the phosphorylation status of an mTORC1 downstream substrate, the ribosomal protein S6 kinase (S6K). Exposure of HEK293 cells with 0.5  $\mu\text{M}$  iRap for 60 min, resulted in a clear downregulation of pS6K1 at Thr389 (as compared to the DMSO control), revealing that iRap interferes with mTORC1 (Fig 1).



**Figure 1.** HEK293 cells were exposed to DMSO, 0.5  $\mu\text{M}$  rapamycin resp. 0.5  $\mu\text{M}$  iRap in complete medium for 60 min at 37°C before cells were lysed and proteins were subjected to SDS-PAGE and immune blotting using antibodies against tubulin and phosphorylated S6K (pS6k1 Thr389). Experiment performed by Dominik Erhart. iRap kindly received from Tom Wandless.

Furthermore, the effect on mTORC1 kinase activity by iRap and another commonly used rapalog AP29167 was determined through the determination of the  $\text{IC}_{50}$  values in HEK293 cells. Surprisingly, the results revealed that both rapalogs inhibit mTORC1 ( $\text{IC}_{50} = 0.1 \text{ nM}$  for rapamycin,  $\text{IC}_{50} = 5 \text{ nM}$  for iRap and  $\text{IC}_{50} = 10 \text{ nM}$  for AP29167)<sup>[40]</sup>, either due to impurities of rapamycin or rapamycin-byproducts, or an undefined spatial orientation of the bulky group at the C16 position.

Overall, these data revealed once more that the synthesis and purification of rapalogs is very challenging. Already minor contaminations of rapamycin and rapamycin-byproducts (<<1%) can inhibit mTORC1 signaling.

#### Aim of this project

In collaboration with Ruben Cal we set out to generate a rapalog CID, which does not interfere with mTORC1 signaling and which dimerizes FKBP12 and FRB (Thr2098Leu) fusion proteins with excellent kinetics, comparable to the one of the rapamycin CID. Furthermore, the synthesized rapalog should be able to be derivatized with functional groups, such as with a photocleavable group resulting in a photoactivable rapalog or with a SNAP-tag resp. Halo-Tag substrate resulting in a trimerizer molecule.

#### Adapted protocol

While repeating the synthesis and purification of (R)-pcRap (performed by Ruben Cal), the protocol was slightly changed than the one described in the manuscript. Purification by the preparative HPLC was performed with a 90-min 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> and not with a 100-min linear gradient of 100:0 to 92:8 (CH<sub>2</sub>Cl<sub>2</sub>:isopropanol) and a flow rate of 30 ml·min<sup>-1</sup> as indicated in the manuscript. Further analysis by the analytical HPLC was performed with a linear gradient of 100:0 to 96:4 (CH<sub>2</sub>Cl<sub>2</sub> : methanol) and not with a linear gradient of 100:0 to 96:4 (CH<sub>2</sub>Cl<sub>2</sub> : isopropanol) as indicated in the manuscript. Additionally, the established protocol for the purification of (R)-pcRap is inefficient and need to be further optimized.

#### Conclusion

We successfully develop a novel synthetic route for a rapalog (pcRap), which efficiently induces dimerization between FKBP12 and FRB (Thr2098Leu)-fusion proteins while not interfering with mTORC1 signaling. Additionally, the (R)-pcRap CID can be used orthogonal with the HaXS CID, which enables to control two cellular events in parallel and thus to study more complex signaling pathways.

However, a detailed characterization of (R)-pcRap is missing and further experiments for validation of (R)-pcRap are required, such as the determination of the (R)-pcRap stability in aqueous solution or a side-by-side comparison of dimerization kinetics of rapamycin and (R)-pcRap. To determine *in vitro* rate constants of (R)-pcRap with FKBP and (R)-pcRap-FKBP with FRB, Biacore experiments could be performed, which allow measuring protein-protein and protein-small molecule interactions in real time. Intracellular dimerization kinetics should be compared with the one from rapamycin to get indications on how the introduction of the phenyl-carbamate on the rapamycin core affected the cell permeability and the dimerization behavior. Alternatively, a PAMPA (parallel artificial membrane permeability assay) could be performed in order to determine cell permeability constants of rapamycin and (R)-pcRap.



Furthermore, the described protocol for the synthesis and purification of (R)-pcRap is inefficient and needs to be optimized.

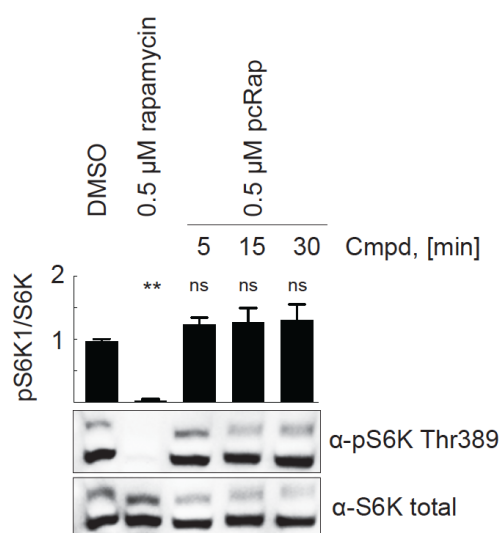
The further derivatization of (R)-pcRap in order to develop novel variants of this rapalog such as fluorescent or photoactivable variants was not successful. The functionalization either at the C40 or C16 position of the pcRap core dramatically lowered the cell penetration abilities (data now shown).

In summary, we developed an alternative CID with fast dimerization kinetics and most importantly, which does not interfere with endogenous mTORC1.

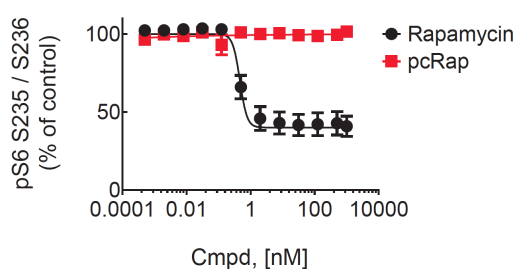
### Supplementary Information

We performed the same experiment as described in the following manuscript (Fig 4a) but with increased incubation time of (R)-pcRap (30 min) to analyze whether a longer incubation of (R)-pcRap will affect mTORC1 in HEK293 cells. The analysis revealed that also under these tested conditions (30 min, 0.5  $\mu$ M), (R)-pcRap does not interfere with mTORC1 (see Fig S1).

Furthermore, we analyzed the effect of (R)-pcRap on mTORC1 in A2058 cells. We added (R)-pcRap for 60 min with an increasing concentration (up to 5  $\mu$ M) on A2058 cells. The phosphorylation status of the ribosomal S6 on Ser235/236 (pS6 S235/236) was analyzed by In-Cell Western experiments and revealed that (R)-pcRap does not interfere with mTORC1 even after 60 min incubation with up to 5  $\mu$ M (Fig S2). The IC<sub>50</sub> value of (R)-pcRap was determined to be > 1  $\mu$ M and the one of rapamycin 0.47 nM, a value comparable to values reported in the literature<sup>[68]</sup>.



**Figure S1.** HEK293 cells grown in fetal calf serum supplemented media were exposed to DMSO, 0.5  $\mu$ M rapamycin, 0.5  $\mu$ M (R)-pcRap for times indicated 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  $\pm$  SEM, n=3; difference from DMSO control: \*\* indicates p<0.01; ns, not significant.



**Figure S2.** A2058 cells were grown in 96-well in fetal calf serum supplemented media. 24 h later cells were exposed for 60 min to rapamycin resp. (R)-pcRap at concentrations indicated. Ribosomal S6 phosphorylation on Ser235/236 (pS6 S235/236) was analyzed by In-Cell Western experiments. Data were analysed with GraphPad Prism, and represent means  $\pm$  SEM, n=3. Experiment performed by Anna Melone.



## 4. Materials & Methods

### Materials

AccuGel 29:1, ultra pure (40% acrylamide stock)	National Diagnostics
Acetic acid	Scharlau
Agarose	Sigma-Aldrich
Albumin from bovine serum (BSA)	New England Biolabs
Ammonium persulfate (APS)	Bio-Rad
Ampicillin trihydrate	Fluka
Aprotinin	Fluka
Bromophenole blue sodium salt	Riedel-de Haen
Calcium chloride (CaCl <sub>2</sub> x 2H <sub>2</sub> O)	Fluka
Deoxynucleotides (dNTPs)	Sigma-Aldrich
Dimethylsulfoxide (DMSO)	Sigma-Aldrich
DNA AluI marker (low range)	Fermentas
DNA Lambda molecular weight marker (high range)	LabForce
T4 DNA ligase	NEB
Ethanol, absolute, HPLC grade	Scharlau
Ethidium bromide	Sigma-Aldrich
Ethylene diamine tetraacetate (EDTA)	Fluka
Fetal calf serum	Sigma-Aldrich
D (+)-glucose-monohydrate	Fluka
L-Glutamine, 200mM	Sigma-Aldrich
Glycerol, anhydrous	Fluka
Hoechst 33342	Juro Supply
Hydrochloric acid (HCl)	Merck
In-Fusion recombinase	Clonetech
Isopropanol, analytical grade	Scharlau
Kanamycin sulfate	Fluka
Leupeptin	Alexis Corporation
Magnesium chloride hexahydrate (MgCl x 6H <sub>2</sub> O)	Fluka
β-mercaptoethanol	Sigma-Aldrich
Methanol, HPLC grade	Scharlau
Milk powder (non-fat)	Migros
Molecular weight marker (14.4 – 167 kDa)	Sigma-Adrich
Nonident P40 (NP-40)	Fluka
Paraformaldehyde, powder, 95% pure	Sigma-Aldrich
Penicillin-Streptomycin solution	Sigma-Aldrich

Pepstatin A	Alexis Corporation
Phenylmethylsulfonylfluoride (PMSF)	Fluka
Potassium chloride (KCl)	Fluka
Potassium dihydrogen phosphate (KH <sub>2</sub> PO <sub>4</sub> )	Fluka
Phusion DNA Polymerase	New England Biolabs
Rapamycin	LC Laboratories
Restriction enzymes	New England Biolabs
Sodium azide (NaN <sub>3</sub> )	Sigma Aldrich
Sodium chloride (NaCl)	Fluka
Sodium dihydrogen phosphate dihydrate (Na <sub>2</sub> HPO <sub>4</sub> x 2H <sub>2</sub> O)	Fluka
Sodium dodecyl sulfate (SDS)	Fluka
Sodium fluoride (NaF)	Fluka
Sodium hydroxide (NaOH)	Merck
Sodium orthovanadate (Na <sub>3</sub> VO <sub>3</sub> )	Fluka
Sucrose	Sigma-Aldrich
N,N,N',N'tetramethylethylenediamine (TEMED)	Bio-Rad
Trizma Base (Tris)	Sigma-Aldrich
Trypsin-EDTA solution (10 x)	Sigma-Aldrich
Tween 20	Fluka
Water, endotoxin-free	Sigma-Aldrich
Dimerizer molecules	
HaXS8	synthesized by Olivier Jacques and Florent Beaufigli, stock solution in DMSO, chemical synthesis and characterization described in Erhart et al., 2013
MeNV-HaXS	synthesized by Viktor Hofmann, Ruben Cal and Florent Beaufigli, stock solution in DMSO, chemical synthesis and characterization described in Zimmermann et al., 2014
Rapamycin	LC Laboratories
pcRap	synthesized by Ruben Cal, stock solution in DMSO, chemical synthesis and characterization described in pcRap manuscript (in preparation)

## Kits

Bio-Rad Protein Assay	Bio-Rad
GenElute Gel Extraction Kit	Sigma Aldrich
GenElute HP Endotoxin-Free Plasmid Maxiprep Kit	Sigma-Aldrich
GenElute PCR Clean Up Kit	Sigma-Aldrich
GenElute Plasmid Miniprep Kit	Sigma-Aldrich
Immobilion Western, Chemiluminescent HRP	Substrate Millipore
In-fusion Cloning Kit	Clontech
JetPEI DNA transfection agent	Polyplus
Phusion DNA Polymerase and PCR buffer	New England Biolabs
Restriction endonucleases and NEB buffers	New England Biolabs

## Buffers and solutions

Blue juice (10x)	80% glycerol, 20% EDTA of 0.5 M pH 8, 1 spatula Bromophenole blue, 1 spatula Xylene Cyanol FF
Coomassie blue staining solution	0.1% Coomassie blue R-250, 50% methanol, 7% acetic acid, 43% ddH <sub>2</sub> O
Coomassie blue destaining solution	90% ethanol, 2% acetic acid, 8% ddH <sub>2</sub> O
10 x PBS (phosphate buffer saline)	28.4 g Na <sub>2</sub> HPO <sub>4</sub> x 2H <sub>2</sub> O, 3.8 g KH <sub>2</sub> PO <sub>4</sub> , 3.8 g KCl, 159 g NaCl, pH adjusted to 7.4, filled up to 2 liter with ddH <sub>2</sub> O, sterilized by autoclaving
10 x TBS	24.4 g Trizma Base, 80 g NaCl, pH adjusted to 7.6 with HCl, filled up to 1 liter with ddH <sub>2</sub> O
1 x TBS-T	1 x TBS + 0.1% (v/v) Tween 20, 5 x 2.5 ml of 1.25 M Tris-HCl pH 6.8, 1 g SDS, 2.5 ml β-mercaptoethanol, 5.8 ml of 87% glycerol, 5 mg bromophenol blue, 35 ml ddH <sub>2</sub> O, stirred until SDS and bromophenol blue had dissolved and filled up to 50 ml with ddH <sub>2</sub> O
Sample buffer	
1M Tris-HCl	121.14 g Tris, dissolved in 800 ml ddH <sub>2</sub> O and pH adjusted to 6.8 resp. 8.8 with HCl. Volume adjusted to 1

---

	liter with ddH <sub>2</sub> O. Sterilized by autoclaving.
Standard lysis buffer	20 mM Tris-HCl pH 8, 138 mM NaCl, 2.7 mM KCl, 5% glycerol, 1% NP-40, 10 x Leupeptin (2mM), 100 x Pepstatin (1.8 mM), 100 x PMSF (100 mM), 100 x Aprotinin (1mg/ml), 500 mM NaF, 100 mM Na <sub>3</sub> VO <sub>4</sub>
10 x electrode buffer (Tris-Glycine)	30.3 g Tris-HCl (250 mM Tris-HCl), 144.2 g glycine (1.92 M), 10 g SDS (1% SDS), filled up to 1 liter with ddH <sub>2</sub> O (pH approx. 8.3)
10 x transfer buffer	250 mM Tris-HCl, 1.92 M glycine, to prepare 1 x transfer buffer, methanol is added to a final concentration of 20% (v/v)
50 x TAE (5 liters)	1.21 kg Trizma Base dissolved in H <sub>2</sub> O, 500 ml 500 mM EDTA, pH 8 and 285.5 ml glacial acetic acid, volume adjusted to 5 liters with ddH <sub>2</sub> O
10 x TE pH 7.5	100 mM Tris-HCl pH 7.5, 10 mM EDTA pH 8
3% paraformaldehyde (PFA) in PBS (20 ml)	18 ml ddH <sub>2</sub> O and 7.5 $\mu$ l 1 M NaOH added to 600 mg PFA, stirring on a hot plate until PFA is dissolved, 2 ml 10 x PBS pH 6.55 added, mixed, cooled to 37°C and pH adjusted to pH 7.2
Media and supplements	
LB Miller	10 g NaCl, 5 g yeast extract, 10 g Bacto-Tryptone, 5 ml 1 M NaOH, filled up to 1 liter with ddH <sub>2</sub> O and sterilized by autoclaving.
LB Miller agar	LB Miller and 12.5 g agar, after autoclaving, medium cooled down to 60°C, appropriate antibiotics added, plate poured and stored at 4°C

---

1 liter SOC	20 g Bacto-Tryptone, 5 g yeast extract, 0.5 g NaCl, 2.5 ml 1 M KCl, 10 ml 1 M MgCl <sub>2</sub> , 200 $\mu$ l 5 M NaOH, 980 ml ddH <sub>2</sub> O, sterilized by autoclaving, 20 ml sterile 1 M glucose added
Ampicillin (1000 x)	dissolved 100 mg/ml in ddH <sub>2</sub> O (1000x stock)
Kanamycin (1000 x)	dissolved 25 mg/ml in ddH <sub>2</sub> O (1000x stock)

## Mammalian cell culture media and supplements

Dulbecco's MEM (DMEM)	Sigma-Aldrich
Dulbecco's MEM (DMEM) without phenol red	Sigma-Aldrich
Fetal calf serum (FCS)	Sigma-Aldrich
100 x L-Glutamine (200 mM)	Sigma-Aldrich
100 x Penicillin-Streptomycin solution	Sigma-Aldrich
10 x Trypsin-EDTA solution	Sigma-Aldrich

## Plasmids

Name	Plasmid	Insert	Constructor
pcDNA3_SNAP26m-EGFP_p1417	pcDNA3	SNAP26m-GFP	Dominik Erhart
pEGFP-N3_HT7_p1445	pEGFP-N3	HT7-GFP	Dominik Erhart
pcDNA3_SNAPf-EGFP_1522	pcDNA3	SNAPf-GFP	Mirjam Zimmermann
pEGFP-N3_HT7(L273Y)-GFP_1560	pEGFP-N3	HT7(L273Y)-GFP	Mirjam Zimmermann
pEGFP-C1_GFP-HT7(L273Y)_1551	pEGFP-C1	GFP-HT7(L273Y)	Mirjam Zimmermann
pEGFP-C1_SNAPf_1525	pEGFP-C1	GFP-SNAPf	Mirjam Zimmermann
pEGFP-N3_HA-HT7(L273Y)_1566	pEGFP-N3	HA-HT7(L273Y)	Sandra Dehn
pcDNA3_HA-SNAPf_1567	pcDNA3	HA-SNAPf	Sandra Dehn
pcDNA3_SNAPf-GFP-T2A-HT7(L273Y)-GFP_1568	pcDNA3	SNAP-GFP-T2A-HT7(L273Y)-GFP	Sandra Dehn
pcDNA3_SNAPf-GFP-2xT2A-HT7(L273Y)-GFP_1587	pcDNA3	SNAP-GFP-2xT2A-HT7(L273Y)-GFP	Sandra Dehn
pcDNA3_SNAPf-GFP-T2A_1601	pcDNA3	SNAP-GFP-T2A	Sandra Dehn
pcDNA3_SNAPf-GFP-T2A-HT7(L273Y)-RFP_1604	pcDNA3	SNAP-GFP-T2A-HT7(L273Y)-RFP	Sandra Dehn
pEGFP-C1_HT7(L273Y)-RFP-Giantin_1590	pEGFP-C1	HT7(L273Y)-RFP-Giantin	Mirjam Zimmermann
pcDNA3_SNAPf-GFP-T2A-Halo-RFP-Giantin_1602	pcDNA3	SNAP-GFP-T2A-HT7(L273Y)-RFP-Giantin	Sandra Dehn
pC4EN-F1_SNAPf-GFP-NLS_1593	pC4EN-F1	SNAP-GFP-NLS	Mirjam Zimmermann
pcDNA3-NLS-GFP-SNAPf-T2A-Halo-RFP	pcDNA3	NLS-GFP-SNAP-T2A-HT7(L273Y)-RFP	Sandra Dehn
pC4-RHE_NES-DsRed-HT7(L273Y)	pC4-RHE	NES-DsRed-HT7(L273Y)	Mirjam Zimmermann
pC4EN-F1<SNAPf-GFP-NLS>_1398	pC4EN-F1	SNAP-GFP-NLS	Mirjam Zimmermann
pcDNA3<HT7(L273Y)-FRB>_1529	pcDNA3	HT7(L273Y)-FRB	Mirjam Zimmermann
pTagRFP<HT7(L273Y)-RFP-Rheb15>	pTagRFP	HT7(L273Y)-RFP-Rheb15	Mirjam Zimmermann
pEGFP-C1<LifeAct-mTFP-SNAPf>_1547	EGFP-C1	LifeAct-mTFP1-SNAPf	Mirjam Zimmermann
pcDNA3_SNAP26m-EGFP-CAAX_p1418	pcDNA3	SNAP26-GFP-CAAX	Dominik Erhart
pEGFP-C1_LAMP-RFP-HT7(L273Y)_1588	pEGFP-C1	LAMP-RFP-HT7(L273Y)	Mirjam Zimmermann
pEGFP-C1_Tom70-HT7(L273Y)2x_1717	pEGFP-C1	Tom70-HT7(L273Y)2x	Dominik Buser
pEGFP-C1_Tom70-GFP-HT7(L273Y)_1718	pEGFP-C1	Tom70-GFP-HT7(L273Y)2x	Dominik Buser
pEGFP-C2<SNAPf-GFP-HA-HRasV12_1630	pEGFP-C2	SNAPf-GFP-HA-HRasV12	Mirjam Zimmermann
pcDNA3_2xFKBP12-eGFP-CAAX_p1440	pcDNA3	2xFKBP12-eGFP-CAAX	Dominik Erhart
pTagRFP_FRB T2098L_p1442	pTagRFP	RFP-FRB T2098L	Dominik Erhart
pEGFP-N3<HT7 L273Y-TQ>_1615	pEGFP-N3	HT7 L273Y-TQ	Mirjam Zimmermann
pcDNA3_Mito-SYFP-HT7(L273Y)_1698	pcDNA3	Mito-SYFP-HT7(L273Y)	Mirjam Zimmermann

Plasmids, which are not listed here, are described in the manuscripts Erhart et al., 2013 and Zimmermann et al., 2014.

## Antibodies

Name	Isotype	Antigen	Source	Dilution
anti-phospho p70 S6 Kinase Thr389 (1A5) #9206	mouse, monoclonal IgG2a	Phospho-p70 S6 kinase Thr389	Cell Signaling Technology	Western blot (1:2000)
anti-p70 S6 Kinase #9202	rabbit, polyclonal	S6K	Cell Signaling Technology	Western blot (1:2000)
anti-phospho-Akt Ser473 (193H12) #4058	rabbit, polyclonal	Phospho-Akt/PKB Ser473	Cell Signaling Technology	Western blot (1:2000)
anti-phospho-Akt Thr308 (244F9)#4056	rabbit, polyclonal	Phospho-Akt/PKB Thr308	Cell Signaling Technology	Western blot (1:1000)

anti-PKB (19G7/C7)	mouse, monoclonal	PKB	Gift from Emilio Hirsch University of Torino (Italy)	Western blot (1:1000)
anti-GFP (mixture of clones 7.1 and 13.1)	mouse, monoclonal IgG1	GFP	Roche	Western blot (1:5000)
anti- $\alpha$ -tubulin (DM1A) #T9026	mouse, monoclonal IgG1	$\alpha$ -Tubulin	Sigma	Western blot (1:50'000)
anti-HA.11 (16B12) #MMs-101R	mouse, monoclonal IgG1	HA	LucernaChem	Western blot (1:5000)
Anti pMAPK	mouse, monoclonal	Rabbit IgG	Promega	Western blot (1:5000)
Anti MAPK total	rabbit, monoclonal	Mouse IgG	Promega	Western blot (1:5000)
anti-rabbit IgG peroxidase conjugate #A6154	goat, polyclonal	Rabbit IgG	Sigma	Western blot (1:5000)
anti-mouse IgG peroxidase conjugate #A4416	goat, polyclonal	Mouse IgG	Sigma	Western blot (1:5000)

## Molecular biological methods

### Agarose gel electrophoresis of DNA fragments

0.7 to 2% agarose in 1x TAE buffer containing ethidium bromide (0.4  $\mu\text{g/ml}$ ) was heated up in the microwave and poured on a glass plate. An appropriate comb was inserted and gel was cooled down until gel was polymerized. DNA samples were diluted in 10x blue juice and run with a molecular size marker (Lambda DNA/EcoRI+HindIII Marker 3 for 0.7% and 1% gels and pBR322 DNA/AluI Marker 20, Fermentas for 2% gels). Electrophoresis was performed in 1x TAE buffer at 70V in electrophoresis chambers (Werkstatt, Institute of Biochemistry Fribourg, Switzerland).

### Preparation of competent E.coli cells XL-1 Blue

Bacteria were grown ON at 37°C on LB-Miller agar plate from which a single colony was picked and used to inoculate a pre-culture of 100 ml liquid LB-Miller medium. After incubation ON (37°C, 300 rpm) the pre-culture was used to inoculate 1 l LB-Miller medium (1:100 dilution). The culture was incubated on the shaker until OD600 reached 0.6. The bacteria culture was placed for 20 minutes on ice and all the following steps were performed at 4°C. After centrifugation for 20 minutes at 3000 rpm (Sorvall RC6+, rotor SLA1500), the bacteria pellet was resuspended in 500 ml ice-cold 50 mM CaCl<sub>2</sub>. Solution is kept a few hours on ice while mixing occasionally. Then bacteria were pelleted for 15 minutes at 3000 rpm, and gently resuspended in 100 ml precooled 50 mM CaCl<sub>2</sub>/10% glycerol. Aliquots of 500  $\mu\text{l}$  were prepared, flash frozen in liquid nitrogen and stored at -80°C until use.

#### Transformation of competent E.coli cells XL-1 Blue

CaCl<sub>2</sub>-competent E.coli cells (XL-1 Blue) were thawed on ice for about 30 minutes. 5 to 7.5  $\mu$ l of the ligation reaction were gently mixed with 100  $\mu$ l competent bacteria in a sterile Eppendorf tube. After 20 min incubation on ice, bacteria were heat-shocked at 42°C for 45 seconds on a Thermomixer. After 5 min incubation on ice, 1 ml of 37°C prewarmed SOC medium was added to the bacteria and incubated at 37°C (250 rpm) for 30-45 min for bacteria containing ampicillin resistance or for 40-60 min for kanamycin resistance. Following centrifugation (13k rpm, 1 min, RT), the bacteria pellet was resuspended in 100  $\mu$ l LB-Miller medium and plated onto LB-Miller agar plates with the corresponding antibiotic. Plate was incubated at 37°C ON. Next day single colonies were picked from the plate and used for the inoculation of liquid cultures.

#### General cloning procedures

##### Preparation of plasmid DNA

Plasmid DNA from E.coli was prepared using commercially available kits from Sigma-Aldrich (Mini, Maxi) according to the manufacture's protocol.

##### Isopropanol precipitation of DNA

0.1 volumes of 3 M sodium acetate pH 5.2 and 0.7 volumes of isopropanol were added to plasmid DNA received from Maxi preparation in a 50 ml falcon tube. After mixing by inversion of the tube, DNA was pelleted by centrifugation (20'000 x g at 4°C for 30 min), rinsed with 1.5 ml 70% ethanol and centrifuged as before for another 30 min. Pellet was dried at the air until it turned white. Then DNA pellet was dissolved in 1x TE buffer.

##### Restriction enzyme digest

For analytical purposes 0.5  $\mu$ g plasmid DNA was digested with 5 Units of restriction enzyme for 1 hour at the appropriate temperature in a total volume of 20  $\mu$ l. For preparative purposes 2  $\mu$ g of plasmid DNA or 35  $\mu$ l of the PCR product was digested with 10 Units of restriction enzymes for 1 hour at the appropriate temperature in a total volume of 20  $\mu$ l.

##### Ligation of DNA fragments into vectors with T4 DNA ligase

Ligation of DNA fragments was performed with T4 DNA ligase in a total volume 15  $\mu$ l. The molar ratio of vector and insert was 1:2 up to 1:6 for single insertions and 1:6 for multiple insertions. Following reaction setup was used:

x  $\mu$ l vector (50-100 ng)  
x  $\mu$ l insert  
1.5  $\mu$ l 10x T4 ligation buffer  
0.5  $\mu$ l T4 DNA ligase (400 U/ $\mu$ l)

Filled up to 15  $\mu$ l with ddH<sub>2</sub>O



The ligation mixture was incubated for 1-2 hour at RT or ON at 4°C. 5 to 7.5  $\mu$ l of ligation reaction was transformed into CaCl<sub>2</sub>-competent E.coli.

#### Insertion of fragments into vector with In-fusion recombinase

Recombination of DNA fragment and backbone with homologous regions was performed with In-fusion recombinase (Clontech) according to manufacturer's protocol.

Following reaction setup was used:

2 $\mu$ l vector
2 $\mu$ l insert
1 $\mu$ l In-fusion reaction mix
5 $\mu$ l water

The recombination mixture was incubated for 15 min at 50°C. 5  $\mu$ l was transformed into CaCl<sub>2</sub>-competent E.coli.

#### Polymerase Chain Reaction (PCR)

PCR was performed with the Phusion High-Fidelity Polymerase. The following reaction setup was prepared on ice and mixed gently:

5x Phusion HF or GC buffer	10 $\mu$ l
MgCl <sub>2</sub> (20 mM)	3 $\mu$ l
Template DNA (10 ng/ $\mu$ l)	1 $\mu$ l
dNTPs (10 mM)	1 $\mu$ l
Forward Primer (50 $\mu$ M)	0.5 $\mu$ l
Reverse Primer (50 $\mu$ M)	0.5 $\mu$ l
DMSO (optional, if GC > 60%)	(1.5 $\mu$ l)
Phusion DNA Polymerase	0.5 $\mu$ l (final concentration of 1 Unit/50 $\mu$ l PCR)
Nuclease-free water	filled up to 50 $\mu$ l

PCR was performed on the Professional Trio Thermocycler (Biometra) using the following protocol:

Initial denaturation of DNA	98°C	120 sec
Denaturation of DNA	98°C	10 sec
Annealing	45-72°C	20 sec (depending on the sequence)
Elongation	72°C	15 sec per 1 kb
30 cycles		
Final elongation	72°C	10 min
Cooling	4°C	indefinitively

After performing the PCR reaction, the product was analyzed by gel electrophoresis.

### Recombinant protein production

For recombinant protein production His-eGFP-HT7, His-eGFP-SNAP, His-HT7(L273Y), His-SNAPf, Halo-eGFP and SNAP-eGFP ORFs were cloned into pTriEx-4 (Novagen) and expressed as N-terminal (His)<sub>6</sub> fusion proteins, and purified on Ni<sup>2+</sup>-NTA beads (QIAGEN) according to the manufacturer's instructions.

### Cell culture methods

#### Adherent cell culture (HeLa, HEK293)

HeLa cells resp. HEK293 cells were cultured with complete DMEM medium (DMEM medium supplemented with 10% FCS, 2 mM L-glutamine and 100 U/ml penicillin/streptomycin) at 37°C in an atmosphere of 5% CO<sub>2</sub>. The cells were splitted every 2 to 3 days when confluence reached 80-90%.

#### Thawing and freezing of cells

Frozen cells from the liquid nitrogen tank were thawed in the 37°C water bath before cells were resuspended in 10 ml pre-warmed cell culture medium and centrifuged (5 min, 900 rpm, RT). The pellet was resuspended with pre-warmed cell culture medium and transferred into a cell culture flask. Trypsinized cells were pelleted (5 min, 900 rpm, RT) and resuspended in freezing medium (10% DMSO in FCS) at a minimal density of 5-10 x10<sup>6</sup> cells/ml, transferred into 2 ml cryotubes (Nunc) and placed into a freezing container (Nalgene) for 24 hours at -80°C. The 100% isopropyl alcohol in the freezing container ensures a constant cooling rate of 1°C/min. Frozen cells are then transferred to the liquid nitrogen tank for long-term storage.

#### Transfection of cells with JetPei

Transfection of adherent cells was performed using JetPEI™ (Brunschwig) according to the manufacturer's guidelines. For optimal transfection efficiency cells should be about 50-60% confluent.

Following number of cells were seeded 24 h before transfection was performed:

Culture vessel	Number of adherent cells / well	Volume of medium per well
24-well	50 000 HeLa	1 ml
6-well	300 000 – 500 000 HEK293 100 000 – 300 000 HeLa	2 ml

#### Preparation of the transfection solutions:

Culture vessel	DNA solution		JetPEI solution	
	Amount of DNA (µg)	Volume of 150 mM NaCl (µl)	Volume of JetPEI reagent (µl)	Volume of 150mM NaCl (µl)
24-well	0.5	50	1	50
6-well	2	50	3	50

The DNA solution and the JetPEI solution were gently vortexed for 10 sec. Then JetPEI solution was added to the DNA solution, vortexed again for 10 sec. After incubation for 20-30 min at RT the mixture was added drop-wise onto the cells.

## Biochemical methods

### Intracellular heterodimerization with HaXS8 resp. MeNV-HaXS

HEK293 resp. HeLa cells co-expressing constructs containing the dimerizing domains were exposed to HaXS8 resp. MeNV-HaXS dimerizer stock solutions in DMSO (at concentrations from 0.05  $\mu$ M to 50  $\mu$ M) for 5 to 60 min at 37°C. DMSO was used as negative control.

### Protein Denaturation, Cell Lysis, and Immune-blotting

All working steps were performed with pre-cooled reagents and on ice. HEK293 resp. HeLa cells were washed with ice cold PBS and lysed in a NP-40 lysis buffer (standard lysis buffer supplemented with 1% NP-40, 40 mM NaF, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 20 mM Leupeptin, 18 mM Pepstatin, 5 mM Aprotinin, 1 mM PMSF, 1 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>). Cell lysates were cleared by centrifugation at 13,000 rpm for 15 min. Protein concentration was determined using Bradford. Then proteins were denatured by the addition of 5x sample buffer and boiling for 6 min at 96°C.

### Protein separation by SDS PAGE gels

SDS page gels were placed into a Hoefer SE250 Mighty Small II complete electrophoresis unit. 10-15  $\mu$ g proteins were loaded. Three  $\mu$ l of MWM SDS-PAGE marker (Molecular Weight Mixture, 14.4 – 167 kDa, Sigma-Aldrich) was loaded per gel. The gels were run at 220 V and 20 mA per gel for about 60 minutes.

Equal amounts of proteins were loaded on gels and separated by SDS-PAGE before they were transferred to Immobilon PVDF membranes (Millipore). Primary antibodies were used to detect proteins by immune-blotting. Secondary antibodies were labeled with horseradish peroxidase (HRP)-conjugated goat anti mouse IgG and goat anti-rabbit IgG (Sigma) for visualization using enhanced chemiluminescence (Millipore) detected by a CCD camera system (Fusion Fx7, Vilber).

### Determination of protein concentrations with Bradford

Bradford reagent (Bio-Rad) was diluted 1:5 in ddH<sub>2</sub>O. 1 ml diluted Bradford reagent was mixed with 2  $\mu$ l cell lysate, mixed and incubated for 5 minutes at RT. Absorbance was measured in semi-micro cuvettes (1.6 ml, Greiner Bio-One) at 595 nm with a spectrophotometer (BioPhotometer, Eppendorf).

## SDS polyacrylamide gel electrophoresis

The concentration of the acrylamide in the separating gel was chosen according to the size of the proteins to be analyzed. Preparation of 10 gels (8 x 10 x 0.75 cm) was performed in a multicasting cassette (Mighty Small SE2000):

	Separating gel (total 75 ml)				Stacking gel (total 30 ml)
	7.5%	10%	Gradient gel		5%
			5%	15%	
Acrylamide Stock (40%)	14.1 ml	18.8 ml	3.1 ml	9.4 ml	4.8 ml
Distilled H <sub>2</sub> O	44.9 ml	40.2 ml	16.5 ml	8.3 ml	21.6 ml
Tris-HCl pH 8.8	15 ml	15 ml	5 ml	5 ml	-
Tris-HCl pH 6.8	-	-	-	-	3 ml
Sucrose				3.7 g	
10% SDS	750 $\mu$ l	750 $\mu$ l	250 $\mu$ l	250 $\mu$ l	300 $\mu$ l
10% APS	250 $\mu$ l	250 $\mu$ l	83.3 $\mu$ l	83.3 $\mu$ l	102 $\mu$ l
TEMED	37.5 $\mu$ l	37.5 $\mu$ l	12.5 $\mu$ l	12.5 $\mu$ l	30 $\mu$ l

SDS, APS and TEMED was added directly before pouring the gels. The separating gel solution was prepared first, poured and overlaid with isopropanol. After about 30 minutes the separating gel is polymerized. Isopropanol was removed and the gel rinsed with ddH<sub>2</sub>O. Then the stacking gel was prepared, poured on top of the separating gel and 10 resp. 15-slot combs were inserted.

For the preparation of gradient gels, two different solutions for the separating gel were prepared (see Table 5). Both solutions were poured simultaneously and mixed by using a gradient forming apparatus (Pharmacia).

## Semi-dry-transfer and western blotting

A PVDF membrane (Immobilion P PVDF membrane, cut 9 x 6.5 cm, Millipore) was activated in 100% methanol before soaked together with Whatman papers (cut 9 x 6.5 cm, Macherey & Nagel) in 1x Transfer Buffer (25 mM Tris, 192 mM Glycine, 20% methanol, pH 8.3). To perform the transfer a sandwich (three Whatman papers, PVDF membrane, gel, three Whatman papers) was assembled in a semi dry blotter (Witec AG). The transfer was performed at 40 V, 30 W and 65 mA per gel for 75 minutes. After blockage of the membrane in 5% milk in TBS-Tween (TBST, 0.1% Tween) for 30 minutes at RT, the membrane was incubated with the primary antibody diluted in 5% milk in TBST for 1-2 hours at RT or ON at 4°C. After three washing steps of 5 minutes in TBST, the membrane was incubated with the secondary HRP-coupled antibody (diluted 1:5000 in 5% milk TBST) for 1 hour at RT and washed again three times for 5 min. For detection of the horseradish-labeled secondary antibodies, the membrane was incubated with 2 ml chemiluminescent HRP substrate (Horseradish peroxidase, Immobilion Western, Millipore). The detection was performed with the Fusion FX7<sup>TM</sup> imaging system (analysis software FUSION CAPT). The images were processed with Photoshop ImageJ and Canvas11. The quantification was performed using ImageJ or the FUSION CAPT software.

## Microscopy

### Plating and preparation of cells

For fluorescence microscopy using a laser scanning confocal microscope (Axiovert 200M and laser scanning module LSM510 Meta Zeiss), HeLa cells were seeded on 12 mm coverslip in 24 well plates with 1 ml complete DMEM medium. After 24 hours cells were transfected as described before.

For microscopy using the Operetta HCS microscope, HeLa cells were seeded in 6-well plates, transfected and after 24 hours transferred in 96-well plates. Twenty-four hours later cells were treated with the heterodimerizer HaXS8 (0.5-1.0  $\mu\text{g}/\mu\text{l}$ ) for 15-60 minutes at 37°C or with DMSO as negative control.

### Cell fixation

Cells were washed two times with 1 x PBS and incubated with 3% PFA for 10 minutes at 37°C. To visualize nuclei, cells were treated with 1  $\mu\text{M}$  Hoechst for 15 min at RT. Afterwards, cells were washed two times with 1x PBS and mounted with 37°C pre-warmed Mowiol on glass slides (for confocal microscopy) or stored in 1 x PBS (for Operetta microscopy).

### Fluorescence microscopy

All microscopy pictures were taken with the confocal microscope Axiovert 200M using 40 x, 63 x or 100 x objectives with oil and operated using LSM510 software (Zeiss). The images were processed with ImageJ and assembled with Canvas11.

To quantify the transfection efficiency microscopy pictures were taken with the Operetta HCS Microscope, operated using Harmony and analyzed with Columbus software.

### Microscopic analysis of HaXS-induced dimerization

See Results part, Protocol Manuscript, STEP 3A (for induction of dimerization in cell lysates) or STEP 3B (for induction of dimerization in live cells)

### Microscopic analysis of UV light-induced cleavage of MeNV-HaXS dimers

See Results part, Protocol Manuscript, STEP 4A (for induction of cleavage in cell lysates) or STEP 4B (for induction of cleavage in live cells).

## 5. Discussion

The understanding of biological systems highly depends on tools available to manipulate cellular processes and to assay phenotypic responses. Nowadays, a huge variety of molecular techniques are available and intensive research results in successive technological advances, which enables interrogation of biological systems under study with increasing precision while minimizing off-target perturbations.

A fundamental way to investigate the role of a protein is to increase or decrease its function and observe the response on the cellular system. In principle, each step in the conversion of a gene into a protein can be targeted. Genetic techniques targeting DNA (such as site-specific genome editing or site directed mutations) are robust and specific but suffer from the disadvantage of a slow onset of the induced effect as well as from potential compensation effects due to redundant pathways. Control over RNA stability through RNAi is faster and easier to implement, but the efficiency of knockdowns and the possibility of off-target effects limit the utility of this approach<sup>[69]</sup>. Direct targeting of proteins via small, cell-permeable inhibitors or activators offers many advantages. The chemical molecules act rapidly, can be applied tunably and additionally can reversibly bind a protein target. However, this approach is mainly limited by potential off-target effects<sup>[70][71]</sup>, as well as the availability of specific molecules for the protein under study as most proteins do not have high-affinity small-molecule binding partners.

A promising strategy is to implement a dimerization system with small molecules in a single approach. This combination enables to profit from the specificity and modularity of genetically encoded dimerizing domains with the speed of chemical molecule-based approaches<sup>[72]</sup>. A successful example of this strategy is based on small molecules that simultaneously bind two protein domains and thereby induces their proximity. These so-called chemical inducers of dimerization (CIDs) are powerful tools to manipulate protein localizations in living cells with high spatiotemporal precision. Over the last years many different CIDs were developed and have successfully been used to control a wide range of cellular events, including various signal transduction pathways such as the Ras signaling pathway in T cells<sup>[73]</sup> or the activation of the Raf-1 kinase<sup>[74]</sup>, gene transcription<sup>[13]</sup>, post-translational modification<sup>[16]</sup>, subcellular localization of proteins<sup>[75]</sup> and many more. Each CID system has its own limitations and the choice of the CID that suits best to your applications requirements is critical.

New class of CID based on Halo- and SNAP-tag reactive dimerizers

We developed a new class of chemical dimerizers based on Halo- and SNAP-tag reactive heterodimerizer, called HaXS. The cell-permeable HaXS molecules efficiently induce intracellular dimerization of Halo- and SNAP-tagged fusion proteins with high specificity and without interfering with endogenous signaling molecules and inducing feedback mechanisms. This enables to investigate cellular systems without considering toxic effects induced by the

heterodimerizer molecule, as for example necessary while using the rapamycin CID. Rapamycin does not only bind FKBP12 and FRB fusion proteins in the dimerizing constructs, but also endogenous FKBP12 and mTORC1. This can induce unwanted side effects on the regulation of cell growth, proliferation as well as autophagy<sup>[76]</sup>, rendering the interpretation of resulting phenotypes challenging.

Besides the fast and efficient HaXS-induced dimerization, which does not interfere with endogenous signaling pathways, the HaXS CID offers another advantage. The covalent reaction of the chemical dimerizer with Halo- and SNAP-tag fusion proteins enables the simple monitoring of the dimerization efficiency under denaturing conditions through performing immune blots. The most time consuming part while performing experiments with the HaXS CID or CIDs in general is the design of new fusion proteins bearing the dimerizing domains, which needs careful consideration and optimization for every single application. The possibility for a fast evaluation of the dimerization behavior of newly designed constructs greatly simplifies analysis and optimization of constructs and cuts down the overall experiment time of the CID experiments. In contrast, the dimerization efficiency of CIDs based on non-covalent interactions can only be analyzed by microscopic translocation experiments or through a time-consuming analysis of expected cellular outputs. To our knowledge, HaXS is the only covalent CID that is able to link freely diffusible proteins in cells, whereas another covalent CID (S-CROSS<sup>[45]</sup>), only links pre-associated dimerizing proteins in cell lysates.

Light as regulatory trigger to enable manipulation of protein dynamics with high spatiotemporal precision

Although CIDs are extensively used to study various signaling events, the cellular uptake of the chemical molecule limits the spatial and temporal resolution of these tools. In contrast, light is an excellent regulatory trigger. Its delivery is immediate and can be applied with high spatiotemporal precision. Extensive constructs testing and time-intensive case-by-case optimization is required to put single proteins under light control either via genetically or chemical incorporation of a light-sensitive module. The indirect control of protein activities through a light-induced control of protein localizations with a dimerizer system is a more modular approach and the dimerizing domains can be fused to various target proteins. Genetically-encoded light-induced (optogenetic) dimerization systems have promising features, but existing tools still have limitations such as large protein tags<sup>[54][57]</sup>, requirement of an exogenous co-factor<sup>[57]</sup>, slow kinetics<sup>[57]</sup>, sensitivity of accidental exposure to environmental light as well as the spectral overlap between activating light and fluorescent reporter<sup>[55],[56]</sup>.

Photolabile groups to investigate cellular systems

Nitrobenzyl- and coumarinyl-4-methyl derivatives represent both classes of photolabile groups suitable to be integrated in biological tools. Photolabile groups involved in studying cellular events require the ability to be excited at light at wavelengths higher than 300 nm in order to prevent cell damage, have to display a high quantum yield and undergo an efficient photolysis reaction. Additionally, the intermediates and products formed during the photolysis reactions cannot be cytotoxic or absorb the emitted light<sup>[77]</sup>. The coumarinyl-4-methyl derivatives are especially promising, as this class of photolabile groups display higher extinction coefficients and absorption maxima compared to the nitrobenzyl derivatives<sup>[77]</sup>. Furthermore, coumarinyl-4-methyl derivatives are suited for two-photon excitation (> 700 nm), and thus enables to induce excitation of this group under less toxic conditions.

Modular synthetic strategy of HaXS enable generation of HaXS derivatives with novel features

The chemical setup of the HaXS dimerizer is modular and the core module linking the Halo-tag and SNAP-tag substrate building blocks are prepared separately. This enables the relatively simple generation of HaXS derivatives with novel properties, through modifying or substituting the core module of HaXS with functional groups with novel features. Through the substitution of the core module with the photocleavable methyl-6-nitroveratryl (MeNV) group, we successfully generated a photocleavable HaXS dimerizer with excellent intracellular dimerization kinetics<sup>[2]</sup>. Furthermore, we developed an alternative photocleavable dimerizer based on a coumarin derivative through substituting the core module of HaXS with a 7-alkoxycoumarinyl-4-methylhydroxyl- group to generate HCM-HaXS. The intracellular dimerization and cleavage conditions are comparable to the MeNV-HaXS dimerizer (data now shown). Its applicability for two-photon cleavage is under investigation.

Summing up, the success of the modular synthetic strategy of the HaXS dimerizer is exemplified by the successful generation of these two photocleavable dimerizers, which both retained the excellent intracellular dimerization kinetics of the starting molecule HaXS8, demonstrating that modifications and substitution of the core module of HaXS8 does not dramatically affect the intracellular penetration capabilities of the dimerizers.

In contrast, the chemical derivatization of naturally occurring molecules such as the introduction of novel functional groups (for example fluorescent or photoactivable groups) is very challenging and already small modifications can dramatically lower their cell penetration capabilities. In the most popular CID based on rapamycin, only few positions of the rapamycin core (C16, C28, C40) can be functionalized and already additions of small groups at these positions can dramatically affect the cell permeability of rapamycin. In this context, rapamycin is a less versatile tool than the HaXS CID as the generation of rapamycin derivatives with novel functions is very challenging.



MeNV-HaXS is a valuable alternative CID, which unifies many advantageous features

With the development of a photocleavable MeNV-HaXS, we provide a novel tool that combines the advantages of a modular system of a CID with the ability to control protein dynamics with high spatiotemporal precision by light. The combination of chemical-induction and light-induced reversion of the induced dimerization offers the possibility to use MeNV-HaXS for many different reaction schemes, such as the dimerizer-induced rerouting and sequestering of proteins away from their normal site of action, which results in their inactivation, analog to the so-called “knocksideway” approach described by Robinson and colleagues<sup>[29]</sup>. In contrast to traditional gene perturbation techniques like gene knockouts and knockdowns, the effects of the “knocksideways” occur on a faster timescales. Additionally, since proteins are not inactivated through their destruction but through their removal from their place of function, these proteins can in principle be re-activated. The use of the photocleavable MeNV-HaXS in this conditional protein inactivation approach enables to reactivate the inactivated proteins through releasing of trapped proteins upon cleavage of the MeNV-HaXS-dimers as well as to observe translocation kinetics of trapped proteins back to their normal localization.

In most present available CIDs, the induced dimerization is not reversible, either due to covalent (as in the HaXS8 CID) or high affinity (as in the rapamycin CID) interactions of the dimerizer with the respective dimerizing domains. Current strategies to implement reversibility in available CIDs are not very promising. Either a sophisticated experimental setup of combining two CIDs (Dual translocation strategy<sup>[63]</sup>), which requires a high effort for the construct design or slow dimerization (> 10 min) and/or reversion of the induced dimers (> 10 min) through replacing the dimerizer molecules with competitor molecules (rCD1<sup>[65]</sup>, TMP<sup>[66]</sup>), limit the applications of these reversible CIDs.

In contrast, MeNV-HaXS-induced dimer complexes of HaloTag and SNAP-Tag fusion proteins can be efficiently cleaved by UV light, which results in the immediate reversion of induced dimers or release of anchored proteins. This enables to control a defined time-window of an activated status with a clear switch-on and switch-off point and thus to reversibly control signaling events and to adjust the duration of an induced signaling event or a cellular response, which are important prerequisites to reproduce physiological cellular events.

Furthermore, the big choice of light sources that can be used to induce photocleavage of MeNV-HaXS (fluorescent lamp, 355 nm FRAP lasers, 405 nm laser or a bulk UV lamp) makes the MeNV-HaXS system applicable to many labs. Additionally, the choice between global and local illumination opens a wide range of applications, such as the study of subcellular compartment-associated signaling and the simulation of cell-wide signaling dynamics. Furthermore, since there is no spectral overlap between the excitation wavelength (355 nm) and the imaging of commonly used fluorescent reporters, most available fluorescent proteins can be integrated in the MeNV-HaXS CID. In contrast, some optogenetic dimerizer systems require blue light to induce dimerization<sup>[55-57]</sup>, which dramatically limits the choice of

fluorescent reporters compatible with these systems. Furthermore, MeNV-HaXS-induced dimers are stable over a long time (tested up to 8 h) and not affected through exposures to ambient light, thus there is no need to perform experiments in dark, as it is the case while performing experiments with some of the optogenetic dimerizer system<sup>[57]</sup>. Due to the short illumination required to induce cleavage of MeNV-HaXS, no phototoxic effects on cells are expected. However, to exclude any off-target effects induced by UV light, the light-insensitive HaXS8, which displays comparable intracellular dimerization properties as MeNV-HaXS, can be used as internal control.

Simulation of HaXS-induced dimerization reaction reveals important features of the HaXS CID and CIDs in general

Besides these advantages of the HaXS CID, the non-directed reaction mechanism as well as the fact that only the reaction pathway in which HaXS first reacts with the HaloTag protein yields in efficient dimer formation limits the overall dimerization performance of the HaXS CID. However, the simulation of HaXS-induced dimerization reactions under various conditions, revealed important insights into the mechanism of the HaXS CID and CIDs in general. Through increasing the ratio of the SNAP-tag rate constant to the HaloTag rate constant and through defining the optimal HaXS concentration, the dimerization performance of the HaXS CID can be dramatically improved. In this context, we demonstrated that depending on the targeted application, the use of a slower variant of the commercially available HaloTag (HT7), can results in better outcomes than HT7, which was optimized for fast ligand labeling.

We investigated various strategies to prevent the integration of HaXS into the HaloTag channel, which blocks its further reaction with the SNAP-tag and thus results in a non-functional reaction pathway of the HaXS CID. None of the analyzed approaches, such as the generation of caged HaloTag substrates, increasing the length of the chloroalkyl chain of the HaloTag substrate or the integration of a more rigid core module in the dimerizer (data not shown), could increase the reaction efficiency of saturated HaloTags (HaXS-Halo) with SNAP-tags. However, simulation of the HaXS-induced dimerization reactions revealed that irrespective of the fact that only one reaction pathway results in efficient dimer formation, these optimized tag variant combination (HT7(L273Y) and SNAPf) enables a very high dimerization efficiency of the HaXS CID under our tested conditions.

Alternative CIDs to expand toolbox of existing dimerization systems

In summary, all CIDs offer their own advantages and the choice of the method that suits best to your planned application is critical. As true for all CIDs, the optimization of the systems to each single application is necessary. The application of efficient optimization strategies of the HaXS CID, such as the simple evaluation of construct performance, choice between two background systems to insert the two POIs that will yield in desired expression levels of the monomers as well as the choice of different light sources to induce cleavage, enables to improve and optimally adjust performance of the HaXS CID to various targeted application.

Summing up, the three dimerizer molecules (HaXS8, MeNV-HaXS, pcRap) definitively contribute important additions to the available toolbox of CIDs. We unified many important features, which are necessary to reproduce physiological signaling pathways in our CIDs. Due to the orthogonal dimerizing domains of the HaXS and pcRap CID, these systems can be used simultaneously to control multiple proteins in a single cell and thus greatly improves the possible complexity for cellular interrogations.

## References

- [1] D. Erhart, M. Zimmermann, O. Jacques, M. B. Wittwer, B. Ernst, E. Constable, M. Zvelebil, F. Beaufils, M. P. Wymann, *Chem Biol* **2013**, *20*, 549.
- [2] M. Zimmermann, R. Cal, E. Janett, V. Hoffmann, C. G. Bochet, E. Constable, F. Beaufils, M. P. Wymann, *Angew Chem Int Ed Engl* **2014**, *53*, 4717.
- [3] D. P. Doupe, N. Perrimon, *Sci Signal* **2014**, *7*, re1; B. Turgeon, S. Meloche, *Physiol Rev* **2009**, *89*, 1.
- [4] M. A. Farrar, J. Alberol-Ila, R. M. Perlmutter, *Nature* **1996**, *383*, 178.
- [5] M. Skwarczynska, M. Molzan, C. Ottmann, *Proc Natl Acad Sci U S A* **2013**, *110*, E377.
- [6] F. S. Liang, W. Q. Ho, G. R. Crabtree, *Sci Signal* **2011**, *4*, rs2.
- [7] T. Miyamoto, R. DeRose, A. Suarez, T. Ueno, M. Chen, T. P. Sun, M. J. Wolfgang, C. Mukherjee, D. J. Meyers, T. Inoue, *Nat Chem Biol* **2012**, *8*, 465.
- [8] K. Baker, C. Bleczynski, H. Lin, G. Salazar-Jimenez, D. Sengupta, S. Krane, V. W. Cornish, *Proc Natl Acad Sci U S A* **2002**, *99*, 16537.
- [9] S. S. Gallagher, L. W. Miller, V. W. Cornish, *Anal Biochem* **2007**, *363*, 160.
- [10] J. L. Czapinski, M. W. Schelle, L. W. Miller, S. T. Laughlin, J. J. Kohler, V. W. Cornish, C. R. Bertozzi, *J Am Chem Soc* **2008**, *130*, 13186.
- [11] S. Gendreizig, M. Kindermann, K. Johnsson, *J Am Chem Soc* **2003**, *125*, 14970.
- [12] D. M. Spencer, I. Graef, D. J. Austin, S. L. Schreiber, G. R. Crabtree, *Proc Natl Acad Sci U S A* **1995**, *92*, 9805.
- [13] V. M. Rivera, *Methods* **1998**, *14*, 421; R. Pollock, R. Issner, K. Zoller, S. Natesan, V. M. Rivera, T. Clackson, *Proc Natl Acad Sci U S A* **2000**, *97*, 13221.
- [14] B. R. Graveley, *RNA* **2005**, *11*, 355.
- [15] S. Schlatter, C. Senn, M. Fussenegger, *Biotechnol Bioeng* **2003**, *83*, 210.
- [16] J. J. Kohler, C. R. Bertozzi, *Chem Biol* **2003**, *10*, 1303.
- [17] S. N. Ho, S. R. Biggar, D. M. Spencer, S. L. Schreiber, G. R. Crabtree, *Nature* **1996**, *382*, 822.
- [18] P. J. Belshaw, D. M. Spencer, G. R. Crabtree, S. L. Schreiber, *Chem Biol* **1996**, *3*, 731.
- [19] T. Ueno, B. H. Falkenburger, C. Pohlmeier, T. Inoue, *Sci Signal* **2011**, *4*, ra87.
- [20] P. Shahi *et al.*, *PLoS One* **2012**, *7*, e30814.
- [21] V. M. Rivera *et al.*, *Nat Med* **1996**, *2*, 1028.
- [22] D. M. Janse, B. Crosas, D. Finley, G. M. Church, *J Biol Chem* **2004**, *279*, 21415.
- [23] S. Lee, H. Park, T. Kyung, N. Y. Kim, S. Kim, J. Kim, W. D. Heo, *Nat Methods* **2014**, *11*, 633.
- [24] K. W. Freeman, R. D. Gangula, B. E. Welm, M. Ozen, B. A. Foster, J. M. Rosen, M. Ittmann, N. M. Greenberg, D. M. Spencer, *Cancer Res* **2003**, *63*, 6237.
- [25] H. D. Mootz, E. S. Blum, A. B. Tyszkiewicz, T. W. Muir, *J Am Chem Soc* **2003**, *125*, 10561.
- [26] M. Putyrski, C. Schultz, *Chem Biol* **2011**, *18*, 1126.
- [27] S. Terrillon, M. Bouvier, *EMBO J* **2004**, *23*, 3950.
- [28] H. Haruki, J. Nishikawa, U. K. Laemmli, *Mol Cell* **2008**, *31*, 925.
- [29] M. S. Robinson, D. A. Sahlender, S. D. Foster, *Dev Cell* **2010**, *18*, 324.
- [30] G. R. Crabtree, S. L. Schreiber, *Trends Biochem Sci* **1996**, *21*, 418.
- [31] L. A. Banaszynski, C. W. Liu, T. J. Wandless, *J Am Chem Soc* **2005**, *127*, 4715.
- [32] B. C. Suh, T. Inoue, T. Meyer, B. Hille, *Science* **2006**, *314*, 1454; P. Varnai, B. Thyagarajan, T. Rohacs, T. Balla, *J Cell Biol* **2006**, *175*, 377.
- [33] F. Castellano, P. Montcourrier, P. Chavrier, *J Cell Sci* **2000**, *113*, 2955; T. Inoue, W. D. Heo, J. S. Grimley, T. J. Wandless, T. Meyer, *Nat Methods* **2005**, *2*, 415.
- [34] S. D. Liberles, S. T. Diver, D. J. Austin, S. L. Schreiber, *Proc Natl Acad Sci U S A* **1997**, *94*, 7825.
- [35] T. Clackson,
- [36] J. Choi, J. Chen, S. L. Schreiber, J. Clardy, *Science* **1996**, *273*, 239.
- [37] J. I. Luengo *et al.*, *Chem Biol* **1995**, *2*, 471.

- [38] J. H. Bayle, J. S. Grimley, K. Stankunas, J. E. Gestwicki, T. J. Wandless, G. R. Crabtree, *Chem Biol* **2006**, *13*, 99.
- [39] T. Inoue, T. Meyer, *PLoS One* **2008**, *3*, e3068.
- [40] S. R. Edwards, T. J. Wandless, *J Biol Chem* **2007**, *282*, 13395.
- [41] M. S. Robinson, J. Hirst, *Curr Protoc Cell Biol* **2013**, *61*, 15.20.1.
- [42] T. Ozawa, *Int J Mol Med* **2006**, *18*, 187.
- [43] J. C. Coutinho-Budd, S. B. Snider, B. J. Fitzpatrick, J. E. Rittiner, M. J. Zylka, *J Negat Results Biomed* **2013**, *12*, 13.
- [44] T. W. Corson, N. Aberle, C. M. Crews, *ACS Chem Biol* **2008**, *3*, 677.
- [45] A. Gautier, E. Nakata, G. Lukinavicius, K. T. Tan, K. Johnsson, *J Am Chem Soc* **2009**, *131*, 17954.
- [46] M. Endo, K. Nakayama, Y. Kaida, T. Majima, *Angew Chem Int Ed Engl* **2004**, *43*, 5643.
- [47] J. P. Pellois, M. E. Hahn, T. W. Muir, *J Am Chem Soc* **2004**, *126*, 7170.
- [48] G. Mayer, A. Heckel, *Angew Chem Int Ed Engl* **2006**, *45*, 4900.
- [49] S. Szobota *et al.*, *Neuron* **2007**, *54*, 535.
- [50] J. Binschik, J. Zettler, H. D. Mootz, *Angew Chem Int Ed Engl* **2011**, *50*, 3249.
- [51] M. E. Hahn, J. P. Pellois, M. Vila-Perello, T. W. Muir, *Chembiochem* **2007**, *8*, 2100.
- [52] J. P. Pellois, T. W. Muir, *Angew Chem Int Ed Engl* **2005**, *44*, 5713.
- [53] Y. I. Wu, X. Wang, L. He, D. Montell, K. M. Hahn, *Methods Enzymol* **2011**, *497*, 393.
- [54] A. Levskaya, O. D. Weiner, W. A. Lim, C. A. Voigt, *Nature* **2009**, *461*, 997.
- [55] M. J. Kennedy, R. M. Hughes, L. A. Peteya, J. W. Schwartz, M. D. Ehlers, C. L. Tucker, *Nat Methods* **2010**, *7*, 973.
- [56] D. Strickland, Y. Lin, E. Wagner, C. M. Hope, J. Zayner, C. Antoniou, T. R. Sosnick, E. L. Weiss, M. Glotzer, *Nat Methods* **2012**, *9*, 379.
- [57] M. Yazawa, A. M. Sadaghiani, B. Hsueh, R. E. Dolmetsch, *Nat Biotechnol* **2009**, *27*, 941.
- [58] N. Umeda, T. Ueno, C. Pohlmeier, T. Nagano, T. Inoue, *J Am Chem Soc* **2011**, *133*, 12.
- [59] A. V. Karginov, Y. Zou, D. Shirvanyants, P. Kota, N. V. Dokholyan, D. D. Young, K. M. Hahn, A. Deiters, *J Am Chem Soc* **2011**, *133*, 420.
- [60] S. Ahmed, J. Xie, D. Horne, J. C. Williams, *J Biol Chem* **2014**, *289*, 4546.
- [61] R. Wollman, T. Meyer, *Nat Cell Biol* **2012**, *14*, 1261; O. Dyachok, O. Idevall-Hagren, J. Sagetorp, G. Tian, A. Wuttke, C. Arrieumerlou, G. Akusjarvi, E. Gylfe, A. Tengholm, *Cell Metab* **2008**, *8*, 26.
- [62] A. Y. Karpova, D. G. Tervo, N. W. Gray, K. Svoboda, *Neuron* **2005**, *48*, 727.
- [63] Y. C. Lin, Y. Nihongaki, T. Y. Liu, S. Razavi, M. Sato, T. Inoue, *Angew Chem Int Ed Engl* **2013**, *52*, 6450.
- [64] M. Skwarczynska, M. Molzan, C. Ottmann, *Proc Natl Acad Sci U S A* **2013**, *110*, E377.
- [65] S. Feng, V. Laketa, F. Stein, A. Rutkowska, A. MacNamara, S. Depner, U. Klingmuller, J. Saez-Rodriguez, C. Schultz, *Angew Chem Int Ed Engl* **2014**, *53*, 6720.
- [66] P. Liu *et al.*, *Angew Chem Int Ed Engl* **2014**, *53*, 10049.
- [67] S. Deswal, A. K. Schulze, T. Hofer, W. W. Schamel, *PLoS One* **2011**, *6*, e22928.
- [68] C. J. Kuo, J. Chung, D. F. Fiorentino, W. M. Flanagan, J. Blenis, G. R. Crabtree, *Nature* **1992**, *358*, 70.
- [69] S. Milstein, M. Nguyen, R. Meyers, A. de Fougerolles, *Methods Enzymol* **2013**, *533*, 57.
- [70] E. Lounkine *et al.*, *Nature* **2012**, *486*, 361.
- [71] A. Gaulton *et al.*, *Nucleic Acids Res* **2012**, *40*, D1100.
- [72] L. A. Banaszynski, L. C. Chen, L. A. Maynard-Smith, A. G. Ooi, T. J. Wandless, *Cell* **2006**, *126*, 995.
- [73] L. J. Holsinger, D. M. Spencer, D. J. Austin, S. L. Schreiber, G. R. Crabtree, *Proc Natl Acad Sci U S A* **1995**, *92*, 9810.
- [74] A. Fegan, B. White, J. C. Carlson, C. R. Wagner, *Chem Rev* **2010**, *110*, 3315.
- [75] P. J. Belshaw, S. N. Ho, G. R. Crabtree, S. L. Schreiber, *Proc Natl Acad Sci U S A* **1996**, *93*, 4604.
- [76] S. Wullschlegel, R. Loewith, M. N. Hall, *Cell* **2006**, *124*, 471.

- [77] P. Klan, T. Solomek, C. G. Bochet, A. Blanc, R. Givens, M. Rubina, V. Popik, A. Kostikov, J. Wirz, *Chem Rev* **2013**, *113*, 119.

## Acknowledgments

First of all, I want to thank all people who supported me during my time as a PhD student. My greatest thanks goes to Matthias Wymann for giving his guidance, motivation and support, but also freedom to develop my own project. Furthermore, I would like to thank Prof K. Ballmer-Hofer for being a member of my PhD committee and for evaluating my thesis and being a co-examiner at my defense.

A special thank goes to chemists Ruben Cal, Viktor Hoffmann and Florent Beaufile, who synthesized and characterized all molecules used for this project. Special thanks goes also to Dominik Erhart, Florent Beaufile and Olivier Jacque, who initiated the HaXS CID project. Additionally, I would like to thank all past and present members of the Wymann lab for the friendly and supportive working atmosphere, as well as for nice lunches and coffee breaks. Finally, I would like to everyone who makes my life interesting and beautiful as it is.

# MIRJAM ZIMMERMANN

St. Galler-Ring 57, 4055 Basel, Switzerland ♦ mirjam.zimmerman@unibas.ch ♦ +41 79 272 37 58

## PERSONAL DATA

---

Born on 19th of October, 1984, in Brugg AG, Switzerland

Unmarried

Swiss nationality

## EDUCATION

---

- 05/2010 – 03/2015 PhD in Natural Sciences (Biochemistry) in the group of Prof. Dr. M. Wymann, University of Basel  
Thesis title: “Novel photocleavable intracellular heterodimerizer to manipulate protein dynamics with high spatiotemporal precision”
- 02/2008 - 08/2009 Master of Science ETH in Biology with focus in cell biology, ETH Zurich
- 10/2004 - 02/2008 Bachelor of Science ETH in Biology with focus on chemistry, ETH Zurich
- 08/2000 - 06/2004 General qualifications for university entrance, Kantonsschule Wettingen

## LANGUAGE SKILLS

---

- German: Mother tongue
- English: Excellent knowledge
- French: Good knowledge
- Italian: Basic knowledge

## PEER-REVIEWED PUBLICATIONS

---

1. [Zimmermann, M.](#), R. Cal, E. Janett, V. Hoffmann, C. G. Bochet, E. Constable, F. Beaufils, and M. P. Wymann. 2014. Cell-permeant and photocleavable chemical inducer of dimerization. *Angew Chem Int Ed Engl* 53: 4717-4720.
2. Erhart, D., [M. Zimmermann](#), O. Jacques, M. B. Wittwer, B. Ernst, E. Constable, M. Zvelebil, F. Beaufils, and M. P. Wymann. 2013. Chemical development of intracellular protein heterodimerizers. *Chem Biol* 20: 549-557.
3. Wyler, E., [M. Zimmermann](#), B. Widmann, M. Gstaiger, J. Pfannstiel, U. Kutay, and I. Zemp. 2011. Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. *RNA* 17: 189-200.



## COURSES

---

- Animal experimentation course, Module 1, Biozentrum (University of Basel), April 2011

## MEETINGS AND PRESENTATIONS

---

- Membrane Biology, Non-Funded Workshop, April 12-16 2014, Lake Como, Italy, oral presentation: "Photocleavable chemical inducer of dimerization to control protein dynamics"
- Targeting the kinome III, Conference, September 26 – 28, 2014, Basel, Switzerland, poster presentation: "Cell Permeant and Photocleavable Chemical Inducer of Dimerization"
- PhD Student Winter Retreat I, II and III, Hasliberg, Switzerland, 2012, 2013 and 2014, 32, oral presentation and poster presentation about small molecules controlling signal transduction in time and space
- Membrane Dynamics in Physiology and Disease, Conference, Basel, Switzerland, June 5 – 8, 2012, poster presentation: "Covalent tag-chemistry targeting receptor dynamics and signaling"
- Molecular Recognition, When Biology meets Chemistry, USGEB Annual Meeting, February 6 – 7, 2012, attendance
- TOR, PI3K and Akt – 20 Years On, Conference, Basel, Switzerland, September 11–13, 2011, poster presentation: "Novel synthetic route for the development of a chemically induced protein dimerizer system based on rapamycin derivatives"
- Probes for lipid systems biology, Workshop, Lake Como, Italy, October 11 to 15, 2011, oral presentation: "Chemical and dynamic control of cellular signaling"
- Lipid and fluorescent probe development, Workshop, Lake Como, Italy, October 16 to 20, 2010, oral presentation: "Novel synthetic routes to inducible rapalogs"