

# **Cytosolic delivery and characterization of monobodies interfering with SH2 domain-phosphotyrosine interactions**

THÈSE N° 8953 (2018)

PRÉSENTÉE LE 19 OCTOBRE 2018  
À LA FACULTÉ DES SCIENCES DE LA VIE  
UNITÉ DU PROF. HANTSCHHEL  
PROGRAMME DOCTORAL EN APPROCHES MOLÉCULAIRES DU VIVANT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

**Nadine Eliane SCHMIT**

acceptée sur proposition du jury:

Prof. J. Lingner, président du jury  
Prof. O. Hantschel, directeur de thèse  
Prof. M. Thomé-Miazza, rapporteuse  
Prof. W. Römer, rapporteur  
Prof. G. van der Goot, rapporteuse



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

Suisse  
2018



# Acknowledgements

This thesis would not have been possible without the support and friendship of numerous people around me.

First of all, I would like to thank my supervisor, Prof. Oliver Hantschel, for giving me the opportunity to work on this exciting project, for his motivation and guidance.

I also thank my mentor, Prof. Etienne Meylan, as well as my thesis committee members Prof. Margot Thomé Miazza, Prof. Gisou van der Goot, Prof. Winfried Römer and Prof. Joachim Lingner for their constructive feedback and discussions.

I also acknowledge the technical support of the EPFL Proteomics, Flow Cytometry and Bioimaging and Optics core facilities, especially Dr. Romain Guet and Olivier Burri for their help with image processing and analysis, R scripting and some pleasant moments at the BIOP.

I had the pleasure to spend my day-to-day life in the lab with many great colleagues. Basak Gencer, Allan Lamontanara and Barbara Gerig were the first PhD students in the lab, giving me a very warm welcome and sharing many joyful moments of the PhD-life. Sina Reckel was here almost throughout my whole PhD, and taught me many lab techniques and also how to stay strong, I thank you for that.

Tim Kükenshöner first of all brought me some amazing monobodies when he came back from Chicago, without which I could not have done this project. I am thankful for many very fruitful scientific discussions throughout our collaboration, for sharing all the ups and downs of the SFK project, and finally celebrating our publication together. Thank you also for our late-evening-conversations, for cheering me up and for many fun moments.

I was very happy, when two new PhD students, Greg La Sala and Daniel Duarte joined the lab. Daniel, I thank you for your thoughtful advice on many topics. Greg, thanks for joining Tim and me in the Monobody-team, for sharing the good and the bad times. Thanks to Greg Mann and Tim Reichart, for bringing some fun moments to the lab, with their British versus American humor during lunch and coffee breaks.

Katyayane Neopane joined the lab as a master's student, and she stayed for a total of 15 months, finishing with a short internship. She joined me on the delivery project using bacterial toxins and was super enthusiastic and motivated from the beginning. You brought the project forward immensely, with your talent and hard work. It was a true pleasure to work with you and I thank you most of all for being such a happy person, always cheering me up.

Finally, I would like to thank Sandrine Georgeon for constantly being there for me, whenever I needed something and for showing me all the techniques in the lab. You are the perfect office neighbor, and we could talk about everything. Lastly, I thank you for some really fun skitouring weekends in the Alps.

I would like to thank all my colleagues for many inspiring discussions, for all their help and support in the lab and for making our lunch and coffee breaks and our lab outings so enjoyable.

I am also thankful to all my friends I met in Lausanne, at Club Montagne and through climbing or skiing trips in the Alps. During the weekends, you provided me with lots of energy, fun moments and long-lasting memories. Thank you all for that!

A big thank you to all my friends, who continue to support and inspire me: Luisa Möhle, Yannis Bastian, Tania Hochberg, Sarah Villringer, Véra Putz, Annika Grosser, Jessica Günzle, Alex Niederbühl, and Christoph Mixtacki. Thank you for keeping in touch and for making it possible to see each other from time to time, despite living in different places.

I have the immense chance to have such an amazing family. My parents, Christiane and Nico, you are always there for me, supporting me where you can. My sister, Catherine, thank you for always being there to cheer me up via our never-ending Skype calls. My aunt, Eliane, you are like a second mother to me, thank you for your love and support.

Finally, I would like to thank my boyfriend, Sylvain. Only you know how tough it sometimes was, and you always did what you could to cheer me up and to support me. Thank you for constantly reminding me about the important things in life.

*Lausanne, September 13<sup>th</sup>, 2018*

## Abstract

Targeted cancer therapy is a very promising concept; however, small molecule inhibitors only exist for a limited number of oncoproteins and resistance development limits their clinical use. Different engineered non-antibody scaffolds are being explored to develop potent protein-based inhibitors of oncoproteins, which can overcome these hurdles.

The first goal of this thesis was to characterize monobodies, engineered binders based on the fibronectin type 3 domain, which were selected to bind the SH2 domains of 6 members of the Src kinase family that play a role in different types of cancer. I have shown that the monobodies bind specifically to their target in cells by performing tandem affinity purifications with subsequent mass spectrometry analysis. When expressed in T cells, the Lck-targeting monobodies ML1 and ML3 inhibit the phosphorylation of Zap70, a kinase which is activated downstream of T cell activation. Moreover, the fusion of ML1 or ML3 to VHL, the substrate receptor of an E3 ubiquitin ligase, resulted in the targeted degradation of Lck. The selected monobodies therefore have great potential to interfere with cancer cell signaling and to be used in therapeutic approaches.

The development of protein-based inhibitors is currently limited to extracellular targets due to the inability of proteins to cross the cell membrane. Many approaches to achieve cellular delivery are hampered by non-selective cytotoxicity or endosomal entrapment of the protein cargo. The second aim of my thesis was to explore different methods to deliver functional monobodies into cancer cells.

Firstly, I tested cell penetrating poly-disulfides (CPDs). CPD-monobody adducts showed toxic effects on the tested cell line, and mainly followed an endocytic uptake route leading to lysosomal degradation, when used at subtoxic concentrations.

Secondly, bacterial toxins have naturally evolved to deliver their payload into the cytoplasm of host cells. Both Shiga-like toxins and Exotoxin A from *Pseudomonas aeruginosa* are taken up into the cytosol via a retrograde trafficking route, avoiding endosomal entrapment. To deliver monobodies to the cytosol, we have tested a combination of their non-toxic subunits: The B subunit of Shiga-like toxin (Stx2B), binding to the sphingolipid Gb3 in the cell membrane, and the translocation domain of Exotoxin A (TDP). We could show that recombinant Stx2B-TDP-Monobody fusion proteins are taken up in HeLa cells, which naturally express Gb3. Colocalization analyses with markers for endocytic compartments demonstrated that the major part of the monobodies escape early endosomes and are not degraded in lysosomes. A recombinant fusion protein of the Stx2B-TDP construct with the monobody AS25, which allosterically inhibits Bcr-Abl kinase, leads to apoptosis in Bcr-Abl-dependent chronic myeloid leukemia cells. Furthermore, the delivery of the VHL-ML3 fusion protein into Jurkat T cells results in decreased Lck levels. The delivery of both constructs is dependent on Gb3 on the cell surface as well as the Stx2B-TDP subunits.

These results demonstrate that functional monobodies can be delivered to the cytosol and bind to their target protein in cells. This method could facilitate the development of protein-based inhibitors of intracellular proteins for cancer treatment.

**Keywords:** Monobody, SH2 domain, Src family kinases, Lck, Tyrosine phosphorylation, Protein delivery, Bacterial toxin, Targeted degradation.

## Résumé

La thérapie ciblée du cancer est prometteuse, mais les petites molécules inhibitrices n'existent que pour un nombre limité d'oncoprotéines et le développement de résistances limite leur utilisation clinique. Différentes matrices protéiques non-anticorps ont été explorées pour développer des inhibiteurs d'oncoprotéines surmontant ces limitations.

Le premier objectif de ma thèse était de caractériser des monobodies, liants synthétiques sélectionnés pour se lier aux domaines SH2 de 6 membres de la famille des kinases Src impliquées dans certains cancers. J'ai montré que les monobodies se lient spécifiquement à leur cible dans des cellules en faisant une purification d'affinité tandem suivie d'une analyse de spectrométrie de masse. Exprimés dans des cellules T, les monobodies ML1 et ML3 (se liant à Lck) inhibent la phosphorylation de Zap70, une kinase activée après activation des cellules T. En outre, la fusion de ML1 ou ML3 à VHL, le récepteur d'une ligase d'ubiquitine E3, mène à la dégradation ciblée de Lck, ce qui indique leur potentiel thérapeutique.

Le développement d'inhibiteurs protéiniques est actuellement limité aux cibles extracellulaires, car les protéines ne peuvent pas franchir la membrane cellulaire. Beaucoup d'approches pour délivrer des protéines dans les cellules se heurtent à une toxicité non-spécifique ou à un confinement endosomique. Le deuxième objectif de ma thèse était d'explorer des méthodes pour délivrer des monobodies fonctionnels dans les cellules cancéreuses.

J'ai d'abord testé des polydisulfides pénétrant les cellules (CPDs). Les adduits entre CPD et monobody se révèlent toxiques pour la lignée cellulaire testée et à des concentrations non toxiques ils entrent dans la cellule principalement par une route endocytyque menant à la dégradation lysosomale.

Puis j'ai testé les toxines bactérielles, organismes ayant évolué pour faire rentrer leur charge dans les cellules. Les toxines du type Shiga (Stx) ainsi que l'exotoxine A de *Pseudomonas aeruginosa* (TDP) suivent une voie rétrograde pour entrer dans les cellules, tout en évitant d'être confinées dans des endosomes. Afin de mener des monobodies au cytosole, nous avons testé une combinaison de domaines non-toxiques : la sous-unité B de Stx (Stx2B), qui se lie au sphingolipide Gb3 dans la membrane cellulaire, et le domaine de translocation de TDP. Nous avons démontré que la fusion entre Stx2B, TDP et un monobody entre dans des cellules HeLa, qui naturellement expriment Gb3. Des analyses de colocalisation avec des marqueurs de compartiments endocytiques démontrent que la majorité des monobodies échappent aux endosomes précoces et ne sont pas dégradés dans des lysosomes. Une fusion entre Stx2B-TDP et le monobody AS25, qui inhibe la kinase Bcr-Abl de manière allostérique, induit l'apoptose de cellules de leucémie myéloïde chronique qui dépendent de Bcr-Abl. Par ailleurs, le transfert de la fusion VHL-ML3 dans des cellules T Jurkat réduit la teneur en Lck. Le transfert des 2 protéines est dépendant de Gb3 à la surface cellulaire et de Stx2B-TDP.

Ces résultats montrent que des monobodies fonctionnels peuvent être délivrés dans le cytosole et se lient à leur protéine cible dans les cellules, et pourraient faciliter le développement d'inhibiteurs protéiniques pour le traitement de cancers.

**Mots clés:** Monobody, domaine protéinique SH2, Famille de kinases Src, Lck, phosphorylation de tyrosine, transfert cellulaire de protéines, toxine bactérielle, dégradation ciblée.

# Table of Contents

<i>Acknowledgements</i> .....	<i>iii</i>
<i>Abstract</i> .....	<i>v</i>
<i>Résumé</i> .....	<i>vii</i>
<i>List of figures</i> .....	<i>xii</i>
<i>List of tables</i> .....	<i>xiv</i>
<b>1 Introduction</b> .....	<b>1</b>
1.1 Targeted cancer therapies and their limitations .....	1
1.2 Protein tyrosine kinases .....	1
1.2.1 The SH2 domain .....	2
1.3 The Src kinase family .....	3
1.3.1 Lck: A Src kinase family member essential for T-cell receptor activation .....	4
1.3.2 The Abl kinase family .....	7
1.3.3 Abl and Src tyrosine kinase regulation .....	7
1.3.4 The Bcr-Abl fusion protein is the hallmark of chronic myeloid leukemia .....	9
1.4 Monobodies .....	11
1.4.1 Monobodies interfering with Bcr-Abl signaling .....	11
1.5 Intracellular protein delivery .....	12
1.5.1 Why is cellular protein delivery desirable? .....	12
1.5.2 Cell penetrating peptides .....	12
1.5.3 The endosomal escape problem .....	13
1.5.4 Cell penetrating polydisulfides (CPDs) .....	14
1.5.5 Bacterial toxins .....	15
1.6 Targeted degradation .....	21
<b>2 Results</b> .....	<b>25</b>
2.1 Monobodies targeting Src-family kinases (SFKs) .....	25
2.1.1 Selection of monobodies binding the SH2 domains of Src-family kinases .....	25
2.1.2 Tandem affinity purification to determine monobody interactors .....	28
2.1.3 Inhibition of SH2-pY interactions by the selected monobodies .....	33
2.1.4 The monobody ML3 inhibits Lck kinase activity in cells .....	35
2.1.5 VHL-ML3 mediates Lck degradation upon expression in Jurkat cells .....	38
2.2 Monobody delivery using CPPs .....	41
2.3 Monobody delivery using Cell Penetrating poly-Disulfides (CPDs) .....	45
2.3.1 Adduct formation and delivery of HPDP-biotin-AS25 .....	45
2.3.2 Buffer and conditions optimization to yield soluble adducts .....	48
2.3.3 Adduct formation and delivery of Avi-tagged and in vivo biotinylated monobody .....	49
2.3.4 CPD-AS25 adducts are taken up into HeLa cells without cytotoxic effects at a lower concentration .....	53
2.3.5 Immunofluorescence and colocalization analyses reveal an endocytic uptake route for CPD-monobody adducts .....	55
2.4 Monobody delivery using Bacterial Toxins .....	59
2.4.1 Cloning and expression of Toxin-Monobody fusion proteins .....	59
2.4.2 Uptake of fluorescently labelled Toxin-monobody fusion proteins in HeLa cells .....	62

2.4.3	Uptake and labeling of SNAP-tagged Toxin-monobody fusion proteins in HeLa cells .....	64
2.4.4	Little colocalization between the delivered protein and Lysotracker.....	67
2.4.5	Toxin-AS25 partially colocalizes with ER-tracker .....	68
2.4.6	Immunofluorescence imaging and quantitative colocalization analysis between the delivered monobody and endocytic compartments .....	70
2.4.7	Delivery of NLS-tagged fusion proteins into HeLa cells.....	76
2.4.8	Gb3 expression in K562 and Jurkat cells.....	78
2.4.9	AS25 inhibits the growth of K562 cells upon delivery.....	80
2.4.10	The delivery of AS25 into K562 cells leads to apoptosis .....	83
2.4.11	The delivery of a VHL-ML3 fusion protein in Jurkat cells targets Lck for degradation .....	86
<b>3</b>	<b><i>Discussion</i></b> .....	<b>91</b>
3.1	Monobodies targeting SFKs .....	91
3.2	Monobody delivery using CPPs.....	93
3.3	Monobody delivery using CPDs.....	94
3.4	Monobody delivery using bacterial toxins.....	96
<b>4</b>	<b><i>Materials and methods</i></b> .....	<b>103</b>
4.1	Cloning.....	103
4.1.1	Gateway cloning .....	103
4.1.2	Site-directed mutagenesis.....	103
4.1.3	Conventional cloning.....	103
4.1.4	InFusion cloning.....	103
4.2	Expression and purification of recombinant proteins.....	105
4.3	Ni-NTA gravity flow purification.....	105
4.4	Size exclusion chromatography.....	105
4.5	SDS-PAGE and Coomassie staining.....	106
4.6	Monobody labelling .....	106
4.6.1	Maleimide labelling.....	106
4.6.2	SNAP tag labelling.....	106
4.6.3	Purification of the labelled protein .....	106
4.7	Mammalian cell culture.....	107
4.8	Transfection .....	107
4.9	Cell lysate preparation.....	107
4.10	Immunoblotting.....	108
4.11	Retroviral transduction.....	109
4.12	Lentiviral transduction .....	110
4.13	T cell stimulation .....	110
4.14	Recombinant SH2 domain pull-down.....	110
4.15	Tandem affinity purification .....	111
4.16	Flow cytometry .....	112
4.16.1	GFP expression analysis .....	112
4.16.2	Cell sorting.....	112
4.16.3	Gb3 expression analysis.....	112
4.16.4	Protein delivery .....	113
4.16.5	AnnexinV/7AAD staining.....	113

4.17	CellTiter-Glo luminescent cell viability assay .....	114
4.18	Realtime-Glo luminescent cell viability assay.....	114
4.19	Microscopy.....	114
4.19.1	Live cell microscopy .....	114
4.19.2	Immunofluorescence .....	115
4.20	Image processing and analysis.....	116
4.20.1	Colocalization analysis .....	116
4.20.2	Nuclear localization analysis .....	116
5	<i>References</i> .....	121
	<i>Curriculum Vitae</i> .....	137

## List of figures

Figure 1.1 Lck SH2 domain bound to the phosphorylated EPQpYEEIPIYL peptide.	3
Figure 1.2 Early signaling events following T cell activation.....	5
Figure 1.3 Overview of TCR signaling. ....	7
Figure 1.4 Domain structures of the Src and Abl tyrosin kinase families as well as the oncogenic fusion kinase Bcr-Abl. ....	8
Figure 1.5 Structures of regulated c-Abl in complex with the kinase inhibitor PD166326 (left, PDB entry 1OPK) and c-Src in complex with the ATP analogue AMP-PNP (right, PDB entry 2SRC). ....	9
Figure 1.6 The Bcr-Abl oncogene results from a translocation between chromosomes 9 and 22, creating the so-called Philadelphia chromosome. (Ph 22q-) .....	10
Figure 1.7 Monobody structure. ....	11
Figure 1.8 Overview of intracellular trafficking of Shiga toxins. (Figure from <sup>170</sup> .)...	19
Figure 1.9 The Endoplasmatic Reticulum-Associated Degradation (ERAD) pathway. ....	21
Figure 1.10 Methods for targeted degradation. ....	23
Figure 2.1 Yeast binding assay of the selected monobodies to the SH2 domain of the different SFKs. ....	26
Figure 2.2 ITC measurements of different monobodies with SH2 domains. ....	27
Figure 2.3 SFK SH2 monobody interactome analysis by tandem affinity purification (TAP).....	29
Figure 2.4 Inhibition of pY peptide /SH2 interaction by monobodies as measured by fluorescence polarization assay. ....	34
Figure 2.5 SH2 pull-down assay. ....	35
Figure 2.6 ML3 or ML1 expression leads to a reduction in Zap70 phosphorylation upon TCR activation.. ....	37
Figure 2.7 Calcium flux measurement in Jurkat cells expressing no monobody, ML1, ML3 or HA4_YA upon TCR stimulation. ....	38
Figure 2.8 VHL-ML1 and -ML3 mediated degradation of Lck in Jurkat cells. ....	39
Figure 2.9 VHL-monobody mediated Lck degradation leads to reduced Zap70 phosphorylation.. ....	40
Figure 2.10 Covalent binding of a monobody to a CPP by disulfide bond. ....	41
Figure 2.11 Monobody purification.....	42
Figure 2.12 HPLC chromatograms of 6 CPPs (in blue) and the reaction product of the same peptide with Ellman's reagent (DTNB) (in green). ....	43
Figure 2.13 CPP-monobody coupling.....	44
Figure 2.14 Synthesis of CPDs and formation of biotinylated and fluorophore coupled monobodies. ....	46
Figure 2.15 Schematic showing adduct formation. ....	47
Figure 2.16 Adduct delivery in HeLa cells. ....	48
Figure 2.17 Adduct delivery and comparison between in vivo biotinylated and HPDP-biotin conjugated monobodies. ....	50
Figure 2.18 CPDs have cytotoxic effects on HeLa cells.....	52

Figure 2.19 100 nM adducts are taken up in HeLa cells without having cytotoxic effects.....	54
Figure 2.20 Colocalization analysis between CPDs and the early endosomal marker EEA1.....	57
Figure 2.21 Colocalization analysis between CPDs and the lysosomal marker Lamp1. ....	58
Figure 2.22 Schematic of all the constructs generated for the project with their size as monomeric proteins .....	60
Figure 2.23 Stx2B-TDP-Monobody fusion protein purification.....	61
Figure 2.24 Toxin-monobody delivery in HeLa cells. ....	63
Figure 2.25 Delivery of SNAP-tagged Toxin-monobody constructs in HeLa cells. ...	65
Figure 2.26 SNAP-tagged Toxin-monobody constructs remain in cells for 24h after delivery. ....	67
Figure 2.27 The majority of delivered Toxin-AS25 does not colocalize with the LysoTracker Red marker. ....	68
Figure 2.28 Toxin-AS25 partially colocalizes with ER-tracker Red.....	69
Figure 2.29 Colocalization of BG-647 labelled Toxin-SNAP-AS25 with early endosomes in HeLa cells. ....	72
Figure 2.30 Colocalization of BG-647 labelled Toxin-SNAP-AS25 with lysosomes in HeLa cells. ....	73
Figure 2.31 Colocalization of BG-647 labelled Toxin-SNAP-AS25 with the TGN in HeLa cells. ....	74
Figure 2.32 Colocalization of BG-647 labelled Toxin-SNAP-AS25 with the ER in HeLa cells. ....	75
Figure 2.33 Uptake of NLS-tagged Toxin-monobody proteins in the nucleus.....	77
Figure 2.34 Quantification of the fluorescence intensity in the nucleus after the delivery of NLS-tagged Toxin-monobody proteins. ....	78
Figure 2.35 Gb3 expression in stable Jurkat and K562 cell lines and uptake of Toxin-monobody constructs. ....	80
Figure 2.36 Realtime Glo assay after AS25 delivery in K562 cells.....	82
Figure 2.37 Flow cytometry plots of annexin V-7AAD-staining after protein delivery. ....	84
Figure 2.38 Fractions of cells in early apoptosis according to the annexin V-7AAD stain after protein delivery.....	85
Figure 2.39 Fractions of cells in late apoptosis according to the annexin V-7AAD stain after protein delivery.....	86
Figure 2.40 Stx2B-TDP-VHL-Monobody fusion protein purification. ....	87
Figure 2.41 Delivery of Toxin-VHL-ML3 in Jurkat cells.....	89

## List of tables

Table 2.1 Overview of mass spectrometry results from the interactome analysis of monobodies MS2, MY1, MY3, ML1 and ML3.....	30
Table 2.2 A: Total spectrum counts of the proteins identified by mass spectrometry during the TAP of stable K562 cell lines expressing the bait proteins MS2, MY1 or MY3.....	31
Table 2.3 Monobodies expressed in the <i>E. coli</i> BL21 strain, their targets, and yields from 1 L of expression medium after purification by SEC. ....	42
Table 2.4 Cell penetrating peptides chosen for coupling to monobodies, their origin and sequence. ....	42
Table 4.1 Expression constructs used for this thesis.....	105
Table 4.2 List of antibodies used for immunoblotting.....	109
Table 4.3 List of antibodies used for immunofluorescence .....	115

# 1 Introduction

## 1.1 Targeted cancer therapies and their limitations

Most chemotherapeutic cancer drugs act non-specifically on all rapidly dividing cells in the body, leading to numerous side-effects. Targeted drugs take advantage of the scientific discoveries of specific proteins and signaling pathways, which are mutated or over-expressed in cancer cells and drive proliferation and resistance to programmed cell death among other cancer hallmarks. Two types of targeted cancer therapeutics are currently used in the clinics: Monoclonal antibodies bind to extracellular targets or cellular receptors, whereas small-molecule inhibitors are cell-permeable and act on intracellular proteins, kinase inhibitors being the prevalent class.<sup>1,2</sup> Even though targeted therapies have led to a number of advances in cancer therapy, including increased overall survival rates and improved patient conditions, several limitations have emerged since they entered the clinics.

First of all, patients often develop a resistance towards the inhibitor, that is caused by evasive mutations in the targeted protein.<sup>3</sup> Moreover, many targeted small molecule drugs, in particular kinase inhibitors, have several off-targets, which leads to side effects and limits their efficacy.<sup>2</sup> Lastly, the majority of known oncoproteins remain untargeted, due to several reasons.<sup>4</sup> Most small molecule drugs target enzymes, which have defined deep substrate-binding pockets where small molecules can bind with high affinity and selectivity and act as inhibitors. However, many untargeted oncoproteins lack such cavities, or their mechanism of action relies on protein-protein interactions which are difficult to target due to the lack of defined small-molecule binding pockets. In addition, many untargeted proteins have close relatives so that very selective binding is hard to achieve with small molecules. Monoclonal antibodies are generally more selective than small molecules, since they bind to their target involving a larger surface interaction. They can be developed to bind virtually any target and do not rely on the presence of a specific binding pocket. However, due to their large size, they are unable to cross the lipid bilayer and can therefore not be used intracellularly.<sup>5</sup>

Many alternative scaffolds are being explored to enlarge the spectrum of protein-based inhibitors.<sup>6</sup> They could overcome several of the current limitations of small molecules and are smaller in size than classical antibodies. Despite their potential to be developed as protein-based therapeutics, they are not able to cross the cell membrane, an issue which I will address in the present thesis.

## 1.2 Protein tyrosine kinases

The human protein tyrosine kinase family counts 90 members which can be divided into receptor tyrosine kinases (RTKs) and cytoplasmic tyrosine kinases (CTKs, 34 members).<sup>7</sup> Their catalytic activity is defined by the transfer of the gamma-phosphate of ATP to a tyrosine residue in the substrate protein, where it becomes covalently attached. Protein phosphorylation was first discovered by the lab of Tony Hunter in 1979, and has since then emerged as a fundamentally important mechanism of signal transduction and regulation in all metazoan cells.<sup>8,9</sup>

The large group of CTKs is further subdivided into 8 distinct families, named after their most prominent member. For this thesis, both the Abl and the Src kinase families are of particular importance. All the members of both the Abl and the Src family contain a central kinase domain, as well as an SH2 and an SH3 domain N-terminal to the kinase domain. SH2 domains play a particular role in this thesis, as I used monobodies selected to bind the SH2 domain of different Src-family kinases and of Abl.

### 1.2.1 The SH2 domain

The human genome encodes for 121 different SH2 domains which are found in 110 human proteins with different functions, such as kinases, phosphatases, small GTPase regulators and scaffolding proteins.<sup>10</sup> They interact specifically with phospho-tyrosine (pY)-containing motifs in other proteins, thus promoting protein-protein interactions important for the propagation and amplification of cellular signals.<sup>11</sup> Since the selectivity of the active site of most tyrosine kinases is limited and determined only by a small number of amino acids surrounding it, there must be other mechanisms to confer specificity and to avoid the disastrous effects of uncontrolled substrate phosphorylation.

In 1986, short 2-amino-acid insertions into a region N-terminal to the Fes kinase domain, which is not required for catalytic activity, were found to modify kinase activity and substrate recognition.<sup>12,13</sup> This finding suggested that the kinase specificity we see in cells is mediated by this non-catalytic domain adjacent to the kinase domain. The authors named it Src homology domain 2 (SH2) because of its homology in all Src and Abl kinases. Later discoveries showed that not only kinases contain SH2 domains, but also many other proteins exerting different functions, like for example the phospholipase C gamma (PLCγ), the Ras GTPase activating protein (RasGAP) and the phosphatases SHP1 and SHP2.<sup>14-19</sup> SH2 domains act as the coupling link between kinase signaling and a range of cellular functions like phospholipid metabolism, Ras activation and transcription factor activation.<sup>13</sup>

In 1992, the first crystal structure of an SH2 domain was solved, showing how this domain is folded, and giving an insight into the nature of the SH2-pY interaction.<sup>20</sup> Indeed, the negatively charged phosphate on the tyrosine fits into a pocket where it forms a hydrogen bond with a conserved arginine residue. In addition, a second pocket was discovered through this crystal structure, which confers additional selectivity to the SH2 domain. Indeed, not only the pY alone but also the sequence context is important for peptide recognition, as together they form a motif which is selectively bound by only certain SH2-containing proteins, thus making up for the lacking specificity of kinase domains.<sup>21</sup> One year later, Lewis Cantley's lab showed that residues C-terminal to the pY, and in particular the residue at position +3 to the pY, are recognized and bind to this second pocket in the SH2 domain. Selectivity is conferred by the recognition of a set of permissive and non-permissive residues, which vary from one SH2 domain to another.<sup>22</sup> (Figure 1.1)

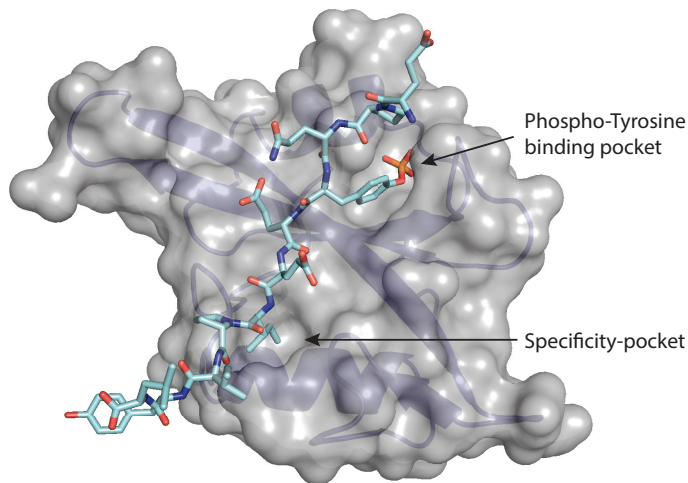


Figure 1.1 **Lck SH2 domain bound to the phosphorylated EPQpYEEIPIYL peptide.** The SH2 domain is shown as a dark blue cartoon and the peptide as sticks. The surface representation of the SH2 domain (gray) shows the pockets where the phospho-tyrosine binds (phospho-tyrosine binding pocket) and where the isoleucine at position +3, relative to the pY, binds (specificity-pocket). PDB entry: 1LCJ, figure assembly using Pymol.

However, even when this specificity-pocket is taken into account, a non-negligible degree of promiscuity remains and the highly specific pY signaling that we see taking place in cells is difficult to explain, especially as cells are crowded with different SH2 domains and pY containing proteins. Therefore, there must be other mechanisms that guarantee the specificity of protein-protein interactions. The formation of multiprotein signaling complexes, the recruitment to specific subcellular compartments or a temporal control through changes in protein abundance could all contribute to this specificity.<sup>23</sup> A recent study has shown that indeed, about 90% of SH2 domains bind plasma membrane lipids using surface cationic patches that are distinct from the pY binding pocket.<sup>24</sup> These interactions could contribute to a spatiotemporal coordination of multiple SH2 domain proteins during cellular signaling cascades.

### 1.3 The Src kinase family

The name-giving member of the Src family has first been discovered as a retroviral oncogene. In fact, v-Src was found to be the transforming gene in Rous Sarcoma Virus (RSV), which was first isolated from a chicken sarcoma by Peyton Rous in 1910.<sup>25</sup> Later studies showed that v-src is homologous to a host cellular gene (c-src) that is highly conserved in eukaryotic species.<sup>26</sup> This discovery led not only to the insight that protein kinases can have high transforming ability but also to the understanding that retroviruses can incorporate host cellular proto-oncogenes like c-Src into their genome by recombination with host DNA.

The Src-kinase family consists of 9 members: Src, Yes, Fyn, Lck, Hck, Blk, Lyn, Fgr and Yrk, which are further subdivided in the SrcA group (Yes, Src, Fyn and Fgr) and the SrcB group (Lck, Hck, Blk, Lyn and Yrk) based on sequence homology. Yes, Src and Fyn are ubiquitously expressed whereas Blk, Fgr, Hck, Lck, and Lyn are found primarily in hematopoietic cells. Lyn and Hck for example are expressed in myeloid cells, Blk in B-cells and Lck in T- and NK-cells.<sup>27,28</sup> Some of these kinases are expressed as alternatively spliced mRNAs in specific cell types.<sup>29</sup>

Src family kinases (SFKs) function as signal amplifiers downstream of receptor tyrosine kinases, G-protein-coupled receptors and immune cell receptors. They are implicated in a broad range of physiological cellular activities as well as hallmarks of tumor development, such as proliferation, differentiation, cytoskeletal modulation, metastasis and migration.<sup>30</sup> Due to their transforming ability upon deregulation, SFKs are tumor drivers in a variety of malignancies, such as leukemia, lymphoma, breast, colorectal or lung cancer and therefore have become important therapeutic targets.<sup>31</sup>

### 1.3.1 Lck: A Src kinase family member essential for T-cell receptor activation

In this thesis, I worked on targeting the SFK Lck, which is essential for T-cell development, activation and proliferation. Since aberrant Lck signaling drives T-cell leukemias, lymphomas and autoimmune diseases, there is an unmet therapeutic need to selectively inhibit Lck.

The majority of T-cell receptors are composed of an  $\alpha$ - and a  $\beta$ -chain forming a heterodimer, whereas about 1-10% of all TCRs are  $\gamma$ - $\delta$  heterodimers in humans. Each T-cell expresses several copies of the same TCR, which specifically recognizes a unique peptide-MHC complex on antigen-presenting cells. Each TCR is associated with one CD3 $\epsilon$  $\delta$ , one CD3 $\epsilon$  $\gamma$ , and one  $\zeta$ -chain homodimer, together forming the TCR complex, which constitutes one signaling unit.<sup>32</sup>

Since the TCR complex itself has no enzymatic activity, it relies on the kinase activity of associated Src-family kinases, in particular on Lck, as well as Zap70 and ITK. Depending on the T cell subtype, Lck is bound to either of the co-receptors CD4 or CD8.<sup>27</sup> When TCR:pMHC binding occurs, CD4 or CD8 brings Lck in close proximity to the TCR complex. pY residues in the immunoreceptor tyrosine-based activation motifs (ITAMs) in the cytoplasmic tails of the CD3 and  $\zeta$ -chains, which are partially phosphorylated in resting T cells, are recognized by Lck via its SH2 domain.<sup>33</sup> Lck then phosphorylates further tyrosines in the ITAMs, creating doubly phosphorylated ITAMs which serve as docking sites for the tandem-SH2 domain of Zap70 ( $\zeta$ -chain associated protein kinase of 70 kDa).<sup>34,35</sup> (Figure 1.2) Zap70 in turn is phosphorylated and activated by Lck and then phosphorylates the adaptor protein linker for activation of T-cells (LAT), which recruits additional signaling proteins to form a multiprotein complex called the LAT signalosome.<sup>36,37</sup>

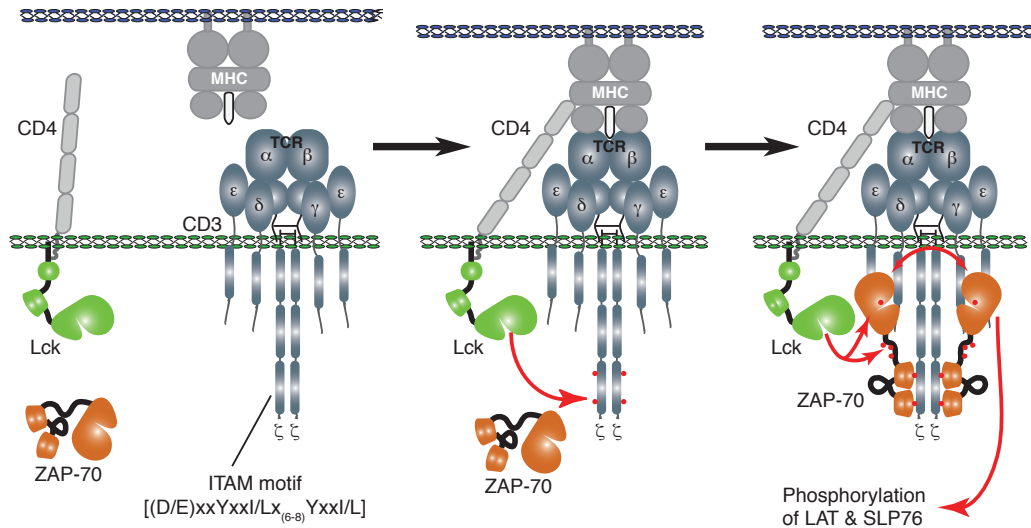


Figure 1.2 **Early signaling events following T cell activation.** Following TCR engagement, CD4-associated Lck is brought into proximity of the CD3 complex and phosphorylates ITAMs in the  $\zeta$ -chain (phosphorylation depicted as red dots). Doubly phosphorylated ITAMs then interact with the tandem SH2 domains of Zap70. After ITAM binding, Zap70 can be phosphorylated by Lck, which results in activation of Zap70 catalytic activity and its autophosphorylation. (Figure taken from <sup>36</sup>)

The threshold for T cell activation needs to be precisely set and the downstream signaling pathways accurately coordinated in order to ensure tolerance towards self-antigens and to limit effector responses. Among the proteins negatively regulating T cell activation are Csk, phosphorylating the inhibitory tyrosine in the C-terminal tail of Lck, as well as the phosphatases CD45 and SHP1, which dephosphorylate the activation loop tyrosine in Lck.<sup>38</sup>

Moreover, in resting T cells, the intracellular domains of the CD3 $\epsilon$  and  $\zeta$  chains are associated with the plasma membrane, preventing phosphorylation of the ITAM tyrosine residues by Lck.<sup>39</sup> The interaction of these domains with the membrane are primarily mediated by clusters of basic residues that interact with negatively charged lipids, such as phosphatidylserine (PS) and phosphatidylinositol-4,5-bisphosphate (PIP<sub>2</sub>).<sup>39</sup> TCR activation by binding to an MHC:peptide complex on an antigen presenting cell induces local changes in the lipid composition of these TCR microclusters.<sup>40</sup> Importantly, the density of PS in the inner leaflet of the plasma membrane is reduced and thus the negative charge, inducing the dissociation of the CD3 $\epsilon$  and  $\zeta$  cytoplasmic domains from the plasma membrane.<sup>40</sup>

There is still some debate about these early steps of T cell receptor signaling and it is not fully understood how these events allow the T cell to discriminate between self and foreign. The nowadays most widely accepted model is the kinetic segregation model, which proposes a segregation of phosphatases like CD45 out of the close-contact zones formed by the interaction of the TCR and CD4 or CD8 with a peptide-MHC. This is explained by the size of CD45 which is too bulky to remain in the close contact zone with an APC. This segregation leads to a net increase of ITAM phosphorylation by the spatial separation from CD45. For downstream phosphorylation to be initiated, the interaction between TCR and peptide-MHC needs to be stable enough so that enough ITAMs are phosphorylated before dissociation occurs, which is only the case with high

affinity interactions. Therefore, the increase in phosphorylation at the TCR occurs in an antigen-specific way and allows for a discrimination between foreign and self-antigens.<sup>41</sup> However, only a small proportion of coreceptors are associated with Lck. Moreover, the associated Lck kinases can be in either an active or an inactive conformation.<sup>32</sup> A recent study proposes a “coreceptor scanning” model explaining that of multiple coreceptors which rapidly bind and unbind the TCR-peptide-MHC complex, only the ones that are bound to active Lck will promote TCR complex phosphorylation. The authors propose this as the rate-limiting step in the kinetic proofreading of the events leading to signal propagation.<sup>42</sup>

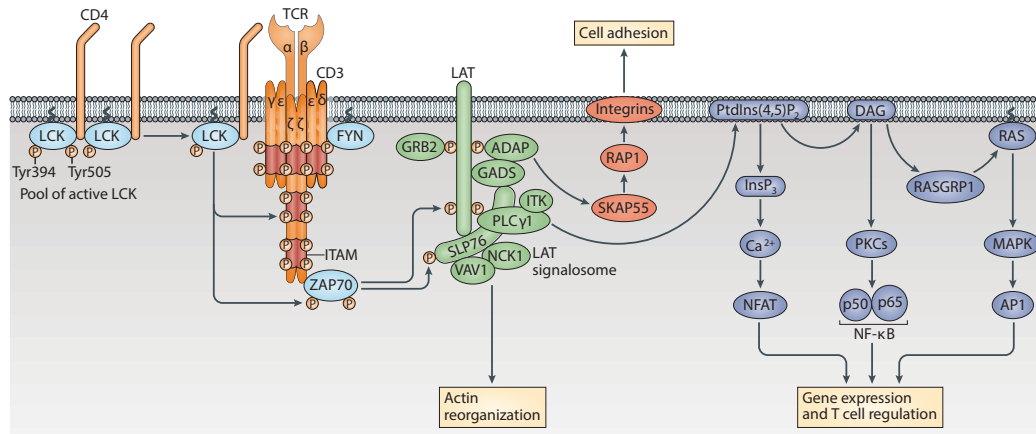
Once the TCR signaling is successfully initiated, different proteins are recruited to the LAT signalosome, leading to a number of downstream signaling events:

(a) The recruitment of the guanine nucleotide exchange factor (GEF) Vav1 activates the small RHO GTPases CDC42 and RAC1, driving actin nucleation and the reorganization of the actin cytoskeleton, which is essential for T cell activation, proliferation, migration and adhesion.<sup>43</sup>

(b) The upregulation of integrin affinity promotes cell adhesion.<sup>38</sup>

(c) The activation of phospholipase C $\gamma$ 1 (PLC $\gamma$ 1) leads to the propagation of the signal into 3 major signaling pathways, all activating different transcription factors: (i) PLC $\gamma$ 1 hydrolyses PIP<sub>2</sub> to generate the secondary messengers inositol-1,4,5-triphosphate (InsP<sub>3</sub>) and diacylglycerol (DAG). InsP<sub>3</sub> causes the opening of Ca<sup>2+</sup> channels in the ER membrane and subsequently in the plasma membrane.<sup>44</sup> The resulting calcium influx activates the transcription factor NFAT and leads to actin remodeling.<sup>45</sup> (ii) DAG activates protein kinase C (PKC) which, together with the recruitment of Grb2 and SOS to the LAT signalosome leads to the activation of the Ras-MAP kinase pathway and AP-1 transcription factors.<sup>32</sup> (iii) The activation of PKC also triggers the assembly of the CARMA1-BCL10-MALT1 complex which leads to the activation and translocation of NF- $\kappa$ B into the nucleus.<sup>46</sup>

The activation of all these transcription factors leads to changes in gene expression that are necessary for T cell growth and differentiation as well as cytokine release. In addition to the assembly of the LAT signalosome, the TCR and the costimulatory molecule CD28 activate PI3K. PI3K phosphorylates phosphatidylinositol 4,5-bisphosphate (PIP<sub>2</sub>) in the membrane to yield phosphatidylinositol (3,4,5)-trisphosphate (PIP<sub>3</sub>), recruiting ITK kinase to the membrane. ITK phosphorylates PLC $\gamma$ 1, thereby contributing to its activation.<sup>47</sup> This requirement for a costimulatory molecule to elicit a full T cell activation prevents self-reactive T cells from being activated by the engagement of their TCR alone. All these processes contribute to the differentiation of the responding T cell into an effector T cell and to its proliferation. Figure 1.3 provides an overview of these different signaling pathways.



**Figure 1.3 Overview of TCR signaling.** The first molecule to be recruited to the TCR–CD3 complex upon binding of a peptide:MHC is Lck, which phosphorylates ITAMs of the CD3  $\gamma$ ,  $\delta$  and  $\epsilon$  chains and the  $\zeta$  chains. Phosphorylation of the  $\zeta$  chain ITAMs enables the recruitment of Zap70, its phosphorylation by Lck and its activation. Activated Zap70 phosphorylates four key tyrosine residues on LAT, which recruits numerous signaling molecules to form a multiprotein complex, termed the LAT signalosome. Important molecules that constitute this complex include PLC $\gamma$ 1, GRB2, GADS, SLP76, ADAP, ITK, NCK1 and VAV1. The LAT signalosome propagates signal branching to three major signaling pathways, the  $\text{Ca}^{2+}$ , the MAP kinase and the NF- $\kappa$ B signaling pathways, leading to the mobilization of transcription factors that are critical for gene expression and essential for T cell growth and differentiation. Signals initiated from the TCR also result in actin reorganization and the activation of integrins by inside-out signaling. (Figure from <sup>38</sup>)

### 1.3.2 The Abl kinase family

The Abl family of cytoplasmic tyrosine kinases counts only two members: c-Abl (ABL1, Abelson tyrosine kinase) and Arg (ABL2, Abl related gene). Both c-Abl and Arg are highly conserved and ubiquitously expressed. The localization of c-Abl in the cell is mostly cytoplasmic, but can also be found in the nucleus, at different intracellular membranes, bound to the actin cytoskeleton or in the mitochondria depending on the tissue and on external stimuli.<sup>48,49</sup> c-Abl is implicated in a whole set of cellular activities like cell cycle regulation, actin dynamics, migration and oxidative stress and DNA damage response.<sup>49,50</sup>

### 1.3.3 Abl and Src tyrosine kinase regulation

Members of both the Abl and Src families of kinases share a similar domain organization. (Figure 1.4) The kinase domain is preceded by a Src-homology-2 (SH2) and a Src-homology-3 (SH3) domain, which both play an important role in target recognition and binding as well as in protein regulation.<sup>50</sup> The N-terminal region of both kinases contains post-translational modifications, in Abl kinases it is myristoylated and in Src-family kinases both myristoylated and palmitoylated.

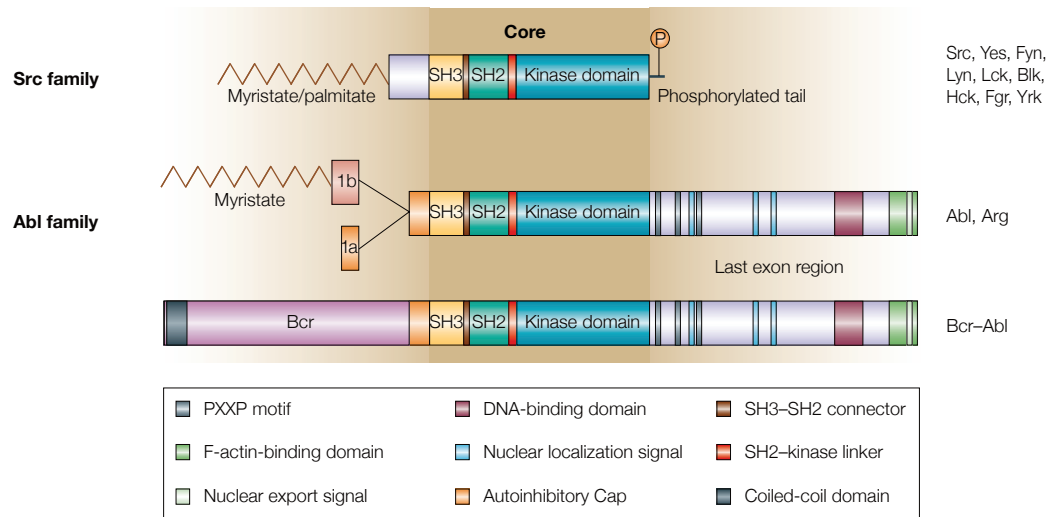


Figure 1.4 **Domain structures of the Src and Abl tyrosine kinase families as well as the oncogenic fusion kinase Bcr-Abl.** (Figure modified from <sup>50</sup>)

Since members of the Src and Abl families control many aspects of cell growth and differentiation, their catalytic activity must be under tight control. The deregulation of protein kinases through mutations or other structural changes can lead to the activation of the kinase, uncontrolled cell growth and cancer. In healthy cells, kinases are therefore mostly found in an auto-inhibited conformation, which relies on a range of intramolecular interactions. The overall structures of auto-inhibited c-Abl and Src resemble each other but differ in some aspects between the Abl and the Src family. (Figure 1.5)

Even before the first crystal structures were available, several groups have found that a tyrosine residue in the activation loop of the kinase domain (Tyr-416 in Src) needs to become phosphorylated for full activity of the enzyme.<sup>51,52</sup> Another tyrosine residue, Tyr-527, which is located in the C-terminal tail of Src family kinases, was found to render the enzyme less active when phosphorylated.<sup>53</sup> *In vivo*, Tyr-527 gets phosphorylated by the Src-specific kinase Csk<sup>54</sup> or its homolog Chk.<sup>55,56</sup>

The first crystal structures of nearly full-length c-Src<sup>57,58</sup> have shown that the SH3 and the SH2 domains pack against the N and C lobes, respectively, of the kinase domain, opposing the active site.<sup>59</sup> Together, they form a clamp, which keeps the kinase domain in an inactive conformation. The SH3 domain binds to a proline-rich region in the segment that links the SH2 domain and the N lobe of the kinase (SH2-linker) and to the N-lobe itself.<sup>60,61</sup> In Src-family kinases, the SH2 domain binds to a phosphorylated tyrosine residue (Tyr 527 in Src) in the C-terminal tail of the protein.<sup>62</sup> In c-Abl, the much larger last exon region replaces this C-terminal tail, and the clamp is held in place by a tight protein-protein interface between the SH2 domain and the C-lobe of the kinase domain.

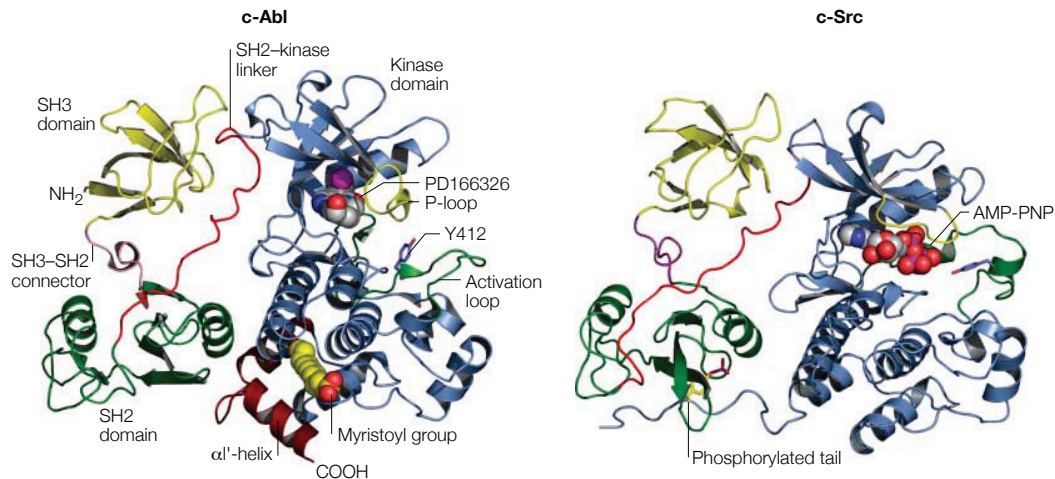


Figure 1.5 Structures of regulated c-Abl in complex with the kinase inhibitor PD166326 (left, PDB entry 1OPK) and c-Src in complex with the ATP analogue AMP-PNP (right, PDB entry 2SRC). (Figure from <sup>50</sup>.)

Another difference between Abl and Src auto-inhibition lies in the conformation of the kinase domain itself. In Src-family kinases, the SH2-SH3 clamp imposes a closed conformation on the kinase domain. In this closed conformation, the activation loop forms a helix situated between the N- and C-lobes, burying and thereby protecting Tyr-416 from phosphorylation and preventing the access of the substrate and ATP to the active site.<sup>59</sup> Moreover, Glu-310 in the  $\alpha$ C-helix of Src is turned away from the active site, which disrupts a critical salt bridge with the catalytic lysine (Lys-295 in Src).<sup>59,61</sup>

In contrast to Src kinases, the Abl kinase domain must be in an open conformation to allow the tight binding of the C-lobe with the SH2 domain.

Another particularity of Abl kinase is the myristoylated N-terminal peptide of c-Abl, which binds to a hydrophobic pocket in the C-lobe, causing a conformational change that breaks a C-terminal  $\alpha$ -helix and renders the C-lobe accessible for docking of the SH2 domain.<sup>63</sup> The N-terminal domain harboring this myristoylated tail is called “cap” region. In Src-family kinases, it is replaced by the SH4 domain, which can also contain myristoyl and/or palmitoyl groups. However, they are not involved in kinase inhibition, but in anchoring the protein to membranes.<sup>64</sup>

#### 1.3.4 The Bcr-Abl fusion protein is the hallmark of chronic myeloid leukemia

Much like Src, the c-Abl gene was discovered as the cellular homologue of a transforming gene from a virus. Abelson murine leukemia virus (A-MuLV) was first isolated in 1970 by Abelson and Rabstein from a mouse infected with Moloney murine leukemia virus (Mo-MuLV), which developed unusual tumors of the lymph node. This virus causes an oncogenic transformation of B-cells, and its transforming gene was named v-Abl.<sup>65</sup>

Chronic Myeloid Leukemia (CML) is a cancer of the blood characterized by the expansion of myeloid cells. The disease can be divided in three phases: The chronic phase during which myeloid progenitor cells expand from stem cells in the bone marrow but differentiate in a normal way. This phase is largely asymptomatic. Several years later, the disease progresses into the second phase called accelerated phase, marked by a substantial increase in myeloid cell expansion and thereafter reaches blast crisis. This final phase is

marked by a block of the hematopoietic differentiation and the blast cells accumulate in both the blood and bone marrow, causing symptoms like fatigue, weight loss, increased infection rate, and upon disease progression also anemia, splenomegaly and thrombocytopenia.<sup>66</sup>

Whereas the disease has been described as early as in 1845 by the pathologist John Hughes Bennet, its underlying cause was only defined in 1960 with the discovery of the Philadelphia chromosome by Nowell and Hungerford, the first discovered genetic aberration in cancer.<sup>67</sup> The breakthrough then came when in 1973, when Janet Rowley discovered that the Philadelphia chromosome results from a translocation between chromosomes 9 and 22, creating a shorter version of chromosome 22, called Philadelphia chromosome.<sup>68</sup> This translocation event fuses the gene encoding c-Abl kinase on chromosome 9 with the breakpoint cluster region (BCR) gene on chromosome 22. (Figure 1.6) The product of this fusion gene is a constitutively active tyrosine kinase called Bcr-Abl, which contains most part of the ABL1 kinase - only lacking the first few amino acids - and drives the oncogenic transformation of the myeloid cells giving rise to CML. Depending on the position of the break in the BCR gene, three main isoforms of Bcr-Abl can occur, which are called p190, p210 and p230, based on their molecular weight. Although the enzymatically active part of the encoded fusion protein, which lies in the Abl kinase part, is the same in all 3 proteins, they are associated with different types of leukemia.<sup>69</sup>

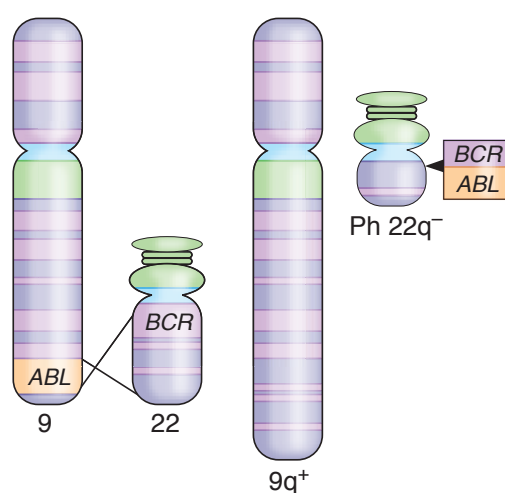


Figure 1.6 **The Bcr-Abl oncogene results from a translocation between chromosomes 9 and 22, creating the so-called Philadelphia chromosome. (Ph 22q-)** (Figure from <sup>68</sup>)

The p210 isoform is the hallmark of CML and a subset of B-cell acute lymphoblastic leukemia (B-ALL) patients, whereas the p190 isoform is only associated with the much more aggressive B-ALL type of leukemia. P230 occurs less often and is associated with chronic neutrophilic leukemia.<sup>71</sup> In all 3 isoforms, a coiled-coil domain located in the amino-terminal region of Bcr is essential for the transforming function of the fusion protein, driving constitutive tyrosine kinase activation.<sup>72</sup> Bcr-Abl activation then results in activation of major pro-proliferative and anti-apoptotic pathways, including the Ras-MAP-Kinase, the PI3K-Akt and the STAT3/5 signaling pathways.

## 1.4 Monobodies

In addition to antibodies, several other protein scaffolds are being explored to create specific protein binders, able to disrupt protein-protein interactions, block active sites of enzymes or bind to allosteric regulatory sites in target proteins. Monobodies constitutes one example of such an alternative protein scaffold.

Monobodies are small engineered proteins of 10 kDa that are based on the molecular scaffold of the 10<sup>th</sup> fibronectin type 3 (FN3) domain of human fibronectin. The FN3 domain is part of the immunoglobulin superfamily and its structure resembles that of an antibody VH domain, except that it is composed of only 7  $\beta$ -strands, connected by 3 flexible loops on each side.<sup>73</sup> Figure 1.7 A shows a comparison between an immunoglobulin domain and the FN3 domain.

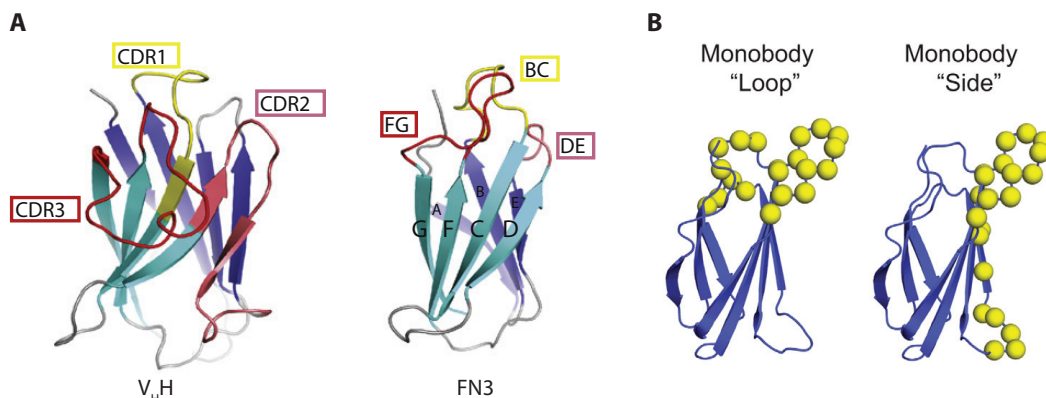


Figure 1.7 **Monobody structure.** **A:** Comparison of the structures of a variable domain of an antibody heavy chain ( $V_H$ , left) and the 10<sup>th</sup> fibronectin type 3 domain (FN3, right). **B:** Schematic showing the diversified amino acids (yellow spheres) in the monobody “loop” library (left) and “side and loop” library (right). (Figure modified from <sup>74,75</sup>.)

Large combinatorial libraries have been created by diversifying residues either in the loop regions only (loop-library) or in the  $\beta$ -strands as well as loop regions (side-and-loop library).<sup>74</sup> (Figure 1.7 B) Potent binders are selected from these libraries by phage- and yeast-display. In multiple selection rounds, high affinity and specificity to their targets can be achieved, making monobodies a robust alternative to antibodies.<sup>75</sup> A major advantage of monobodies in contrast to antibodies is that they do not contain disulfide bonds, making them functional in the reducing environment of the cytosol. They can therefore be used as genetically encoded inhibitors in mammalian cells and their expression in *Escherichia coli* is straightforward.<sup>73</sup>

### 1.4.1 Monobodies interfering with Bcr-Abl signaling

Several monobodies have already been developed in the lab of Shohei Koide as well as in our lab. HA4 was the first monobody selected to bind the SH2 domain of Abl and was shown to block Bcr-Abl-mediated Stat5 phosphorylation in cells, also demonstrating the utility of monobodies to dissect signaling networks.<sup>76</sup> A tandem monobody between HA4 and 7c12, which binds to an allosteric site in the SH2 domain of Abl, opposite to the HA4 binding site was later developed. This tandem monobody potently blocks the formation of the SH2-kinase interface necessary for Bcr-Abl activation and induces apoptosis in Bcr-Abl dependent cell lines.<sup>77</sup> A second set of monobodies termed AS25

and AS27, targeting the SH2-kinase interface with higher affinity than 7c12 were able to inhibit Bcr-Abl signaling and induce apoptosis in Bcr-Abl dependent cell lines, without the need of a tandem to increase the local concentration.<sup>78</sup>

## 1.5 Intracellular protein delivery

### 1.5.1 Why is cellular protein delivery desirable?

Since the introduction of the first therapeutic antibodies in the clinic – Rituximab (Rituxan/Mabthera) targeting CD20 and the HER2 antagonist Trastuzumab (Herceptin) – we see a rapid growth of protein-based therapeutics in the pharmaceutical industry as well as in academic research.<sup>79</sup> Protein-based inhibitors are a promising alternative to small molecules. Protein-protein interactions mostly involve a larger surface interaction than protein-small molecule interactions, including regions outside of the active site of the enzyme, which can promote a higher selectivity.<sup>80</sup> Moreover, many non-enzymatic proteins lack defined substrate binding pockets which could be used as binding sites for small molecules.<sup>81</sup>

Currently, antibodies are the best represented class of protein-based therapeutics on the market, and many alternative scaffolds are being explored to enlarge the spectrum of possible protein-based therapeutics.<sup>75,79,82</sup>

However, most proteins are not readily able to cross the cell membrane to reach potential intracellular targets and are therefore limited to extracellular or receptor targets.<sup>81</sup> As a result, different approaches for the delivery of proteins to the cytosol of mammalian cells are being explored. Additionally, the possibility to use viral carriers to deliver exogenous protein-encoding nuclear acids has been extensively studied, yet this approach is hampered by the safety issues related to genetic modifications.<sup>83,84</sup>

Direct protein delivery approaches are ranging from the use of cell-penetrating peptides, over numerous different nanocarriers to bacterial toxin domains as carriers to transport the payload across cell membranes. I will present a few approaches here as well as their advantages and limitations.

### 1.5.2 Cell penetrating peptides

The first protein which was discovered to enable cell membrane permeability was the trans-activating (Tat) protein on the surface of the human immunodeficiency virus (HIV).<sup>85,86</sup> The smallest peptide sequence endowing the Tat protein with this property was soon identified to be a 13 amino acid sequence rich in positively charged amino acids and named Tat peptide (GRKKRRQRRPPQ).<sup>87</sup> Other peptides able to deliver cargo like fluorescent dyes or RNA into cells were discovered, most of which also have a very high positive net charge. The most widely used ones include the third helix of the homeodomain of the *Drosophila* protein antennapedia which is called pentratin (RQIKIWFAQNRRMKWKK), and a sequence of 9 arginine residues called polyarginine (R9).<sup>88,89</sup> These positively charged peptides are believed to interact with the outer leaflet of mammalian cell membranes by engaging negatively charged glycosaminoglycans like heparin sulfate on the surface.<sup>90,91</sup> Especially arginine is believed to be essential, since its guanidine head can form bidentate hydrogen bonds with negatively charged groups. The

mechanism by which the peptide-cargo is then translocated into the cytoplasm after this first binding step is still under debate and is likely to differ between CPPs and cargos. In addition to the electrostatic interactions that can induce the direct uptake of small peptides through the lipid bilayer,<sup>92</sup> different endocytic mechanisms are thought to play a major role. These include macropinocytosis, clathrin-mediated endocytosis, caveolae-mediated endocytosis and clathrin/caveolae-independent endocytosis, the preferred mechanism depending on the size and charge of the attached cargo, the peptide sequence as well as the cell type specific membrane composition.<sup>93,94</sup> Local membrane destabilization and pore formation have also been proposed to contribute to peptide-cargo uptake.<sup>95</sup>

Another type of CPPs are amphipathic peptides, presenting a helical fold of which one side is hydrophilic and the other side is hydrophobic, favoring interactions with the hydrophobic tails of membrane lipids. The model amphipathic peptide (MAP, KLALKLALKALKALKLA)<sup>96</sup> is one example, as well as Pep-1 (KETWWETWWTEWSQPKKKRKV)<sup>95,97</sup> and transportan.<sup>98</sup> Some amphipathic CPPs like buforin-II have been derived from antimicrobial peptides, which often act by damaging bacterial membranes.<sup>99</sup>

Despite the multitude of different CPPs that have been described, they all share a number of drawbacks. Most CPPs unspecifically bind to and are taken up in any cell types. This lack of specificity has been partially overcome by conjugation with receptor-binding moieties or the selection of CPPs to specifically bind a certain cell type, although the molecular mechanism for CPP-cell selectivity remains unclear.<sup>100</sup> Moreover, many groups use enzymes like Cre-recombinase or fluorescent dyes or proteins as cargo, to provide a proof-of-concept of the studied CPP. However, an extremely low concentration of cargo reaching the cytosol is often sufficient to see a qualitative effect. Most studies lack a precise quantification of the amount of the delivered cargo, which makes it impossible to assess and compare the efficiency of these systems.

### 1.5.3 The endosomal escape problem

Another well-known limitation of CPPs and of other approaches is the endosomal entrapment and subsequent lysosomal degradation of CPP-cargo conjugates. The amount of protein reaching the cytosol is limited and varies with the CPP itself, the cargo and the cell line used. 36-aa miniature proteins have been optimized to escape endosomes by clustering arginine residues on the same face of an  $\alpha$ -helix.<sup>101</sup> So-called proton-sponge polymers like polyethylenimine (PEI) are believed to mediate endosome bursting by making use of the pH change occurring during endosome maturation - the pH in early endosomes is 6.5, it then drops as they mature into late endosomes while in lysosomes the pH is about 4.5. These polymers contain many secondary and tertiary amines and have pKa values between the physiological and lysosomal pH. During the acidification of endosomes, they get protonated and thereby act as a buffer, causing the ATPase at the membrane to transport more protons in an attempt to reach the desired pH. The accumulation of protons is counterbalanced by the enhanced influx of chloride ions, leading to osmotic swelling and lysis of the endosome.<sup>102-104</sup> Some naturally occurring peptides rich in glutamate, aspartate, and histidine show similar properties. These amino acids can be protonated at low pH and act as a proton sponge.<sup>105, 106</sup> Examples include the

major component of bee venom, melittin, and H2A, a peptide rich in glutamate derived from the hemagglutinin H2A subunit of the influenza virus.<sup>107,108</sup>

Aromatic ring containing residues like tryptophan and phenylalanine have also been shown to favor endosomal membrane destabilization by insertion into the lipid bilayer.<sup>109</sup> Kim et al. described an antibody that can induce membrane destabilization through structural changes exposing a so-called endosomal escape motif, rich in aromatic residues. After internalization, the acidic pH in the endosomes protonates aspartate and glutamate residues in the antibody, which thereby become more hydrophobic, facilitating hydrophobic interactions with other residues. This leads to structural changes in the antibody, the endosomal escape motif YYH becomes exposed and the clustered aromatic side chains intercalate into the endosomal membrane. When a critical concentration is reached, the antibody escapes into the cytosol by inducing membrane lipid flip-flop and pore formation.<sup>110</sup>

The group of Shiroh Futaki introduced a novel endosomolytic peptide derived from the cationic and membrane-lytic spider venom peptide M-lycotoxin. By simple co-delivery, they showed cellular uptake of Cre recombinase, Saporin, which is a protein that is cytotoxic at low concentrations, and an IgG antibody.<sup>111</sup>

#### 1.5.4 Cell penetrating polydisulfides (CPDs)

Cell penetrating polydisulfides (CPDs) are poly(amino amine)s containing disulfide linkages in their amino units.<sup>112</sup> Gasparini and colleagues introduced a method to grow CPDs directly on the cargo to deliver, using it as a substrate for polymerization initiation.<sup>113</sup> A thiol group on the substrate acts as an initiator of the polymerization. A monomer composed of a lipoic acid bound to an arginine readily forms a covalent disulfide bond with the initiator, by nucleophilic disulfide exchange. This reaction also generates a reactive thiol to attack the next monomer, starting the polymerization. The polymerization is terminated by addition of iodoacetamide, which allows for some control over the polymer length by the timepoint of termination.

The authors propose a combination of counter-ion mediated translocation, which is similar to what has been proposed for CPPs, and thiol-mediated translocation as uptake route.<sup>113</sup> They hypothesize that CPDs bind covalently to the membrane surface by disulfide exchange with exofacial thiols, and cross the membrane along transient micellar defects, upon which they detach into the cytosol by disulfide exchange with intracellular glutathione.<sup>114</sup> They claim endocytosis to be almost irrelevant, and thereby also endosomal entrapment. The self-inactivation by glutathione immediately after reaching the cytosol would explain the non-toxicity of CPDs. In the initial study, fluorophores have been used as polymerization initiators, which allowed the visualization of the uptake in HeLa cells. However, proteins cannot be used as substrates directly, due to the denaturing polymerization conditions. Two following studies simultaneously published a strategy linking a biotin to a thiol group in the initiator. The resulting biotinylated CPD is then used to form complexes with tetrameric streptavidin and biotinylated proteins.<sup>115,116</sup> The successful uptake of these complexes, which were visualized by adding fluorescently labeled biotin, into HeLa cells was shown. Fu et al. also used NTA as an initiator and coupled the resulting CPD to a His-tagged caspase-3, which induced apoptosis in the delivered cell line upon delivery, showing that functional caspase-3 had reached the cytosol. Furthermore, they functionalized a fluorescently labelled antibody

with a bis-sulfone reagent to covalently attach it to a tetrazine-CPD and demonstrated its delivery based on fluorescent imaging.<sup>115</sup>

### 1.5.5 Bacterial toxins

Bacterial toxins could overcome the limitations of CPPs and other protein delivery agents in terms of cytotoxicity and endosomal entrapment, as they have naturally evolved to deliver their payload to host cells and to escape endosomes. Moreover, they require binding to a specific host cell receptor for their entry, making them cell-selective.

Of particular interest for protein delivery are the so-called AB toxins, which are composed of two subunits, A (for activity) and B (for binding). The A subunit is the toxic protein while the B subunit binds to a specific receptor on the cell membrane and facilitates the internalization.<sup>117</sup> AB-toxins typically need to be proteolytically cleaved in order to release their active A-subunit.<sup>118</sup> In some cases, the cleavage occurs in the bacteria, like for cholera toxin, other toxins, like Diphtheria toxin, Shiga toxin and Pseudomonas Exotoxin A, contain a furin cleavage site which is processed in the target cell.<sup>118,119</sup>

For Anthrax and Clostridium botulinum C2 toxins, the B-subunit first needs to be processed and to oligomerize for the A-subunit to bind non-covalently.<sup>120,121</sup> Most of these toxins are internalized via endocytosis, but the mechanisms of translocation to the cytosol vary. Shiga toxin, Cholera toxin and Pseudomonas Exotoxin A utilize the host cell's retrograde trafficking machinery to reach the cytosol via the trans-Golgi-network and endoplasmic reticulum (ER)<sup>122</sup>, while Diphtheria and Anthrax toxins rely on the acidification of the endosome to translocate directly to the cytosol.<sup>123</sup> In the case of Anthrax toxin, the low pH triggers a conformational change in the protective antigen (PA) or B subunit heptamer to form a pore through which the effector proteins are transported to the cytosol.<sup>124</sup>

Many bacterial AB-toxins have been explored for the delivery of exogenous proteins, by either replacing the A subunit, or by fusion to a mutated A-subunit. Here I will describe a few of them more in detail and point out the advances made for the delivery of exogenous proteins.

#### 1.5.5.1 Anthrax toxin

Anthrax toxin is a virulence factor secreted by *Bacillus anthracis* that is composed of 3 distinct proteins: the receptor targeting and pore forming protective antigen (PA), the lethal factor (LF), which can bind cleaved PA by its N-terminal domain and edema factor (EF).<sup>125</sup> PA mediates the transport of the enzymatically active LF and EF into the cytosol of mammalian cells, where they exert their cytotoxic activity. Full length PA (83 kDa) binds to the host cell receptors Capillary Morphogenesis protein 2 (CMG2) and Tumor Endothelial Marker 8 (TEM8) and gets cleaved by a protease of the furin-family to yield a 63-kDa isoform that forms ring-shaped heptamers on the cell surface.<sup>121,126-129</sup> LF and EF can bind to these heptamers and the whole complex is taken up via clathrin-mediated endocytosis.<sup>130,131</sup> Early studies have shown that even exogenous proteins bound to the truncated and inactive N-terminal domain of LF are taken up into the cytosol when administered together with PA.<sup>132</sup> Following studies have focused on the delivery of peptides for vaccine development<sup>133</sup>, enzymes<sup>134</sup> or toxins<sup>135,136</sup>, of which a few copies are sufficient to see the desired effect. However, it is difficult to measure the exact amount of

protein reaching the cytosol and to assess whether exogenous proteins also reach the cytosol with a similar efficacy as EF and LF or whether the majority stays in endosomes. In 2014, Liao et al. used this system to deliver diverse antibody mimics including monobodies, aiming at perturbing intracellular protein-protein interactions. They used the tandem monobody HA4-7c12, designed to bind Bcr-Abl and interfere with downstream signaling<sup>77</sup>, and could indeed induce apoptosis in a fraction of cells.<sup>137</sup> However, they only observed a modest decrease in activation loop phosphorylation of Bcr-Abl. They reasoned that this modest effect, which is much lower than that achieved by the small molecule Abl-inhibitor Imatinib or transfection of HA4-7c12, is due to the low concentration of the monobody reaching the cytosol.<sup>137</sup> Even though the cytosolic delivery was rather inefficient, this study serves as a proof-of concept, that intracellularly delivered monobodies can perturb the activity of endogenous oncoproteins. In order to pass through the PA pore in the endosomal membrane, which can only accommodate structures of up to 15 Å in diameter, protein cargo must be at least partially unfolded by the PA's intrinsic machinery.<sup>138,139</sup> Verdurmen et al. found that exogenous proteins with a high thermodynamic stability are less efficiently translocated to the cytosol through the PA pore, and that the amount of protein reaching the cytosol can be increased by destabilizing the protein cargo.<sup>140</sup>

#### 1.5.5.2 *Pseudomonas* exotoxin A

Exotoxin A (ETA) is a virulence factor secreted by the opportunistic bacterium *Pseudomonas aeruginosa* and consists of a single polypeptide chain composed of 3 functional domains. Domain Ia binds to the cognate receptor of the toxin,  $\alpha$ 2 macroglobulin receptor/low density lipoprotein receptor on target cells.<sup>141</sup> Domain II is also called translocation domain because it enables the toxin to translocate across cell membranes and domain Ib serves as a linker between domains II and III.<sup>142</sup> Domain III is the catalytically active part of ETA, which acts as ADP-ribosyltransferase, potently inhibiting elongation factor 2, (eEF2) and thereby blocking protein synthesis in target cells.<sup>143</sup>

After binding to the receptor, the toxin is internalized by endocytosis either via clathrin-coated pits or by association with lipid rafts and caveolin-mediated endocytosis.<sup>144</sup>

In the early endosome, the toxin undergoes a conformational change, exposing a furin-cleavage site in the translocation (II) domain. After cleavage by a cellular furin protease, the disulfide bond still connecting the two fragments is reduced by protein-disulfide isomerases, and a 37-kDa fragment is detached, containing domain III and part of domain II.<sup>145</sup> After transport to late endosomes, this fragment reaches the ER via two different pathways. It can exploit a pathway regulated by the GTPase Rab9 to traffic to the trans-Golgi-network (TGN) and then bind to KDEL receptors cycling between the TGN and ER via its C-terminal REDL motif.<sup>146</sup> It can alternatively reach the TGN by a Rab-9 independent pathway and then utilize a Rab6-controlled lipid-dependent sorting pathway to reach the ER.<sup>147,148</sup>

The 37-kDa fragment comprising the cytotoxic domain III is then shuttled from the ER to the cytosol by the cellular ER-associated protein degradation pathway (ERAD).<sup>149</sup> Sequences inside the PE-domain II are recognized by the Sec61 translocon, which normally exports unfolded or misfolded proteins from the ER to the cytosol for subsequent proteasomal degradation.<sup>150,151</sup> How the toxin then escapes the proteasome

once it reaches the cytosol is not completely understood, it might simply avoid recognition by ubiquitin ligases because of the lack of any lysine residues to covalently attach ubiquitin.

In order to create targeted chimeric toxins, Domain I has been replaced by different receptor-binding ligands or antibody domains to redirect the toxin to specific cell types, creating a whole set of different immunotoxins.<sup>152,153</sup>

Mohammed et al. for the first time used the translocation domain II only, to deliver exogenous molecules into cells. They fused it to polyarginine (R10), a cationic CPP, which, used on its own, leads to internalization and endosomal entrapment of the cargo. The combination of polyarginine with *Pseudomonas* exotoxin domain II enabled the cytosolic delivery of the green fluorescent protein eGFP into HeLa cells.<sup>154</sup> It is presumably taken up by the same entry route as the full-length toxin, as it co-stains with ER markers after 4-8 hours. Unfortunately, the authors did not demonstrate the delivery of any functional protein by this method.

Verdurmen et al. have used the translocation domain of Exotoxin A with the same domain boundaries and expressed it as a fusion protein with a designed ankyrin repeat protein (DARPin) for EGF receptor targeting and another DARPin as cargo.<sup>140</sup> They have developed a biotin ligase assay that allows to precisely quantify the amount of protein reaching the cytosol after delivery. It is based on the prokaryotic biotin ligase BirA, which specifically recognizes and biotinylates a 15-amino acid tag called Avi-tag fused to the delivered cargo protein.<sup>140</sup> Since BirA is only expressed in the cytosol of reporter cell lines, only the cytosolically delivered cargo is detected by streptavidin in western blots of cell lysates.<sup>140</sup> Using this assay, they showed that their cargo DARPins were efficiently delivered to the cytosol, and that even thermodynamically very stable DARPins were successfully translocated from the ER to the cytosol, in contrast to proteins delivered by Anthrax protective antigen pores.<sup>140</sup> This is presumably due to the more potent cellular machinery used for unfolding prior to translocation by the ER associated degradation pathway.<sup>140</sup> However, this study is lacking a functional assay, showing that the delivered proteins are well folded and still able to bind their target once they reach the cytosol.

A slightly shorter version of the translocation domain was used by the group of Hak-Sung Kim, who fused it to an EGFR binding rebody, another alternative scaffold binder, to achieve targeting to EGFR expressing cancer cell lines.<sup>155</sup> They used this construct to deliver the ribosome-inactivating plant toxin Gelonin, which induced tumor regression in a xenograft mouse model.<sup>155</sup> Since this is a very potent toxin, the delivery of extremely low concentrations to the cytosol is probably sufficient to reach a cytotoxic concentration.

### 1.5.5.3 *Diphtheria toxin*

Diphtheria toxin (DT) is secreted by *Corynebacterium diphtheria* and is in many ways very similar to exotoxin A. It also consists of a single polypeptide chain, which is proteolytically cleaved into fragments A and B which remain linked by a disulfide bond.<sup>156</sup> Its catalytic domain residing in fragment A follows the same host cell killing strategy as Exotoxin A, by inhibition of elongation factor 2 (eEF2).<sup>156-158</sup> The B domain can be separated into the alpha-helical transmembrane (T) domain and the receptor

binding (R) domain with a beta-barrel structure.<sup>156-158</sup> The R domain binds to the membrane-anchored form of heparin-binding EGF-like growth factor (pro-HB-EGF).<sup>159</sup> Receptor bound toxin is internalized by clathrin-dependent endocytosis and the acidic environment in the endosome then triggers the refolding of the transmembrane domain (T), leading to its insertion into the endosomal membrane to form a 18–22 Å pore.<sup>160,161</sup> The catalytic domain is at least partially unfolded to enable its translocation through the pore, and is released into the cytosol upon reduction of the disulfide bond between fragments A and B.<sup>162,163</sup>

Auger et al. have shown that both the green fluorescent protein eGFP and the thermodynamically more stable red fluorescent mCherry proteins can be translocated as passenger proteins into the cytosol when fused to wild-type or catalytically inactive diphtheria toxin. These results suggest that cargo proteins do not need to unfold in order to be translocated through the pore formed by the transmembrane domain. They also showed the delivery of functional  $\alpha$ -amylase into cells, even though it only showed a limited effect on the cellular glycogen content.<sup>164</sup> However, the question remained, whether is it feasible to stoichiometrically target highly abundant intracellular proteins.<sup>165</sup> In order to answer this question, a precise method to quantify the amount of the delivered protein in the cytosol is first needed. Attempts based on fluorescence or cellular fractionation are prone to artifacts and are rather unprecise.<sup>166,167</sup>

Verdurmen et al. were the first to conduct a comparative study to quantified the cytosolic uptake of cargo proteins delivered to four different cell lines using different bacterial toxins, supercharged proteins and cell penetrating peptides, using their previously introduced biotin ligase assay.<sup>168</sup>

The bacterial toxins studied were Anthrax protective antigen, catalytically inactive Diphtheria toxin and the translocation domain of *Pseudomonas* exotoxin A.<sup>168</sup> Whereas cargo delivered by Anthrax or *Pseudomonas* toxin domains reached the cytosol in high concentrations, cytosolic delivery by diphtheria toxin was only achieved in one out of four cell lines.<sup>168</sup>

#### 1.5.5.4 *Shiga toxin and Shiga-like toxins*

Shiga toxin was discovered in 1897 by Kiyoshi Shiga to be secreted by the pathogenic *Shigella dysenteriae* and is the prototype of the Shiga-toxin family.<sup>169</sup> This family also includes its very close homologue Shiga-like toxin 1 (Stx1), which differs only in one amino acid from Shiga toxin and is secreted by Shiga-like toxin producing *Escherichia coli* (STEC).<sup>170,171</sup> This pathogenic *E. coli* strain can produce two variants of Stx1 (Stx1 and Stx1c) as well as five different variants of Shiga-like toxin 2 (Stx2, Stx2c, Stx2d, Stx2e, Stx2f), which cause more severe disease. Stx2 is only 56% identical to Stx1 at the amino acid level, and the Stx2 variants are 84-99% homologous to Stx2.<sup>172</sup> Shiga-like toxins are also called verotoxins, based on the discovery in 1977 that they kill Vero cells in culture.<sup>173</sup> STEC serotype O157:H7, also called enterohaemorrhagic *Escherichia coli* (EHEC), causes the most severe forms of hemorrhagic colitis in humans, which progresses to the life-threatening haemolytic uraemic syndrome (HUS) in about 15% of patients.<sup>174-176</sup>

Both Shiga toxin as well as all Shiga-like toxins share the same cellular receptor, the neutral glycosphingolipid globotriaosylceramide (Gb3; also known as CD77). All Shiga toxin family members have a very similar structure, where the B subunit forms a pentamer to which one A-subunit is non-covalently bound.<sup>177,178</sup> Each B-subunit monomer can bind up to three Gb3 molecules on the cell surface, so that each AB<sub>5</sub> unit

binds up to 15 Gb3 units.<sup>179-182</sup> This way, B-subunit binding to Gb3 induces receptor clustering on the cell membrane and membrane invaginations. The toxin is then taken up by clathrin-independent and clathrin-dependent endocytosis. Which of these mechanisms is predominant, is still a subject of debate.<sup>183</sup> From early and recycling endosomes, Shiga toxin is transferred by the retrograde transport route to the trans-Golgi network (TGN), in a process dependent on the recruitment of clathrin to the endosome by EpsinR and on the retromer complex.<sup>184</sup> This pathway is different from the recycling pathway of the mannose-6-phosphate receptor between late endosomes and the TGN.<sup>172</sup> The toxin is then transported from the TGN through the Golgi apparatus to the ER in a coat protein complex I (COPI)-independent manner.<sup>185</sup>

In the early endosome, the A subunit of the toxin gets cleaved by the endoprotease furin at a consensus cleavage site close to the B-subunit binding site, dividing it into the catalytically active A1 fragment and the short A2 fragment which is bound to the B subunit pentamer.<sup>186,187</sup> The 2 fragments stay linked by a disulfide bond, which is only cleaved once the toxin reaches the ER.<sup>172, 188</sup> The A1 fragment then uses the cellular ER-associated protein degradation (ERAD) pathway by binding to human ER-associated DNA-J protein 9 (HEDJ) and the Sec61 translocon to unfold and translocate to the cytosol, in a similar fashion as the catalytically active fragment of *Pseudomonas* exotoxin A.<sup>189</sup> (Figure 1.8 and Figure 1.9)

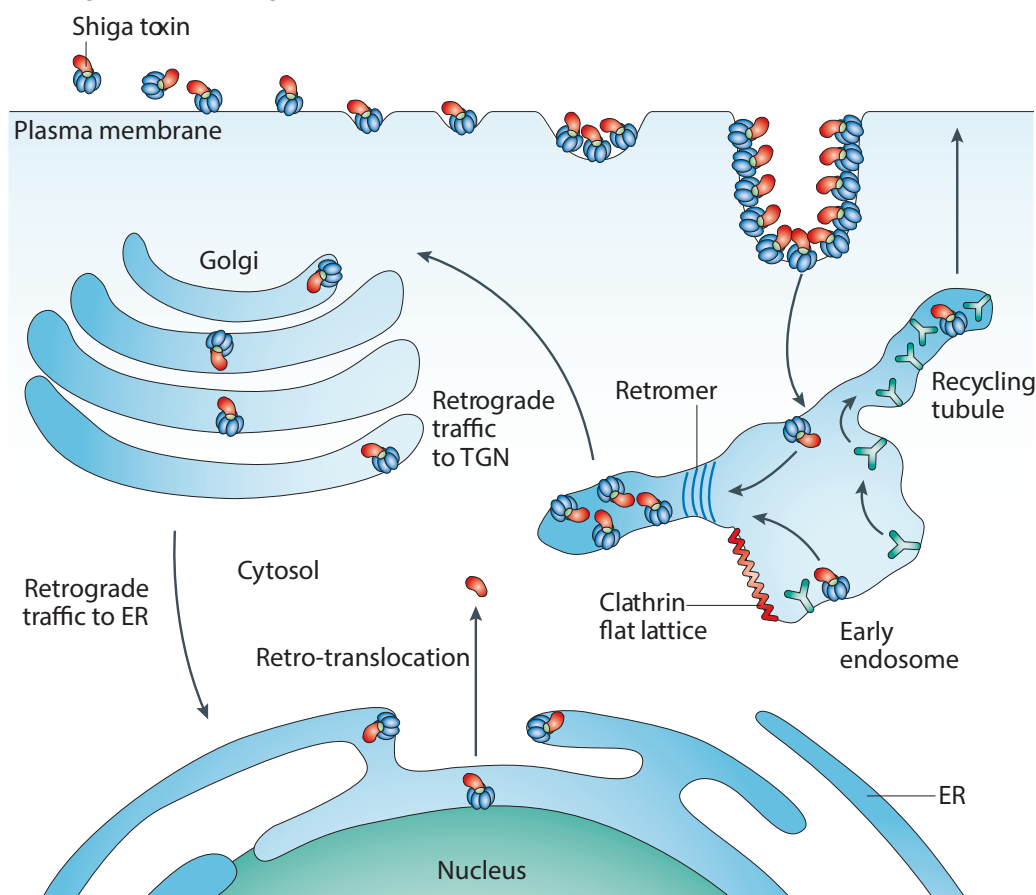


Figure 1.8 Overview of intracellular trafficking of Shiga toxins. (Figure from <sup>172</sup>.)

The Shiga toxin B subunit has been used for immunotherapy applications, to deliver viral or cancer antigens to dendritic cells, where they are presented on major histocompatibility complex class I (MHC I) molecules and can induce a cytotoxic T lymphocyte response.<sup>190-193</sup>

Moreover, since Gb3 is upregulated on a number of cancers, including colorectal, breast and ovarian carcinomas and Burkitt's lymphomas<sup>172</sup>, diverse cytotoxins and contrast agents have been coupled to the Shiga toxin B for tumor targeting and imaging.<sup>194</sup> El Alaoui et al. have coupled a prodrug based on SN-38, a camptothecin derivative that inhibits topoisomerase I and is used to treat colorectal cancer, to the Shiga toxin B subunit to achieve better cancer cell selectivity.<sup>195</sup> Photosensitizing compounds have also been bound to Shiga toxin B subunit to improve specific cellular delivery and cell killing by exposure to light.<sup>196,197</sup>

The group of Hak-Sung Kim has fused the green fluorescent protein EGFP to a truncated, inactive A subunit, which still contains the furin-sensitive loop as well as the B subunit binding part.<sup>198</sup> They expressed and purified it as a fusion construct with the B subunit, with a ribosome binding site 5' to the EGFP-A-subunit and 5' to the B subunit, so that both the whole protein as well as B subunit alone are expressed.<sup>198</sup> They showed that the truncated A subunit fused to EGFP bound to the pentameric B subunit in the same way as wild type Shiga toxin and that it was able to deliver EGFP to Vero cells.<sup>198</sup> They also demonstrated that the receptor targeting of StxB can be changed, by fusing an integrin-binding protein to the B subunit that they mutated to abolish Gb3 binding.<sup>198</sup>

In a second study, they combined the receptor binding domain of Shiga-like toxin (B subunit) and the translocation domain of *Pseudomonas aeruginosa* exotoxin A to create a more potent delivery system. They argue that this system is more versatile than Shiga-toxin alone, since it allows the fusion of diverse cargos, which was not possible with the previous construct.<sup>199</sup> They then used it to deliver different enzymes, including luciferase and adenylate cyclase, as well as the 12-amino acid peptide N8A, a mouse double minute 2 (MDM2) inhibitor.<sup>199</sup> For all these cargos, a small number of molecules reaching the cytosol are likely to be sufficient to elicit a cellular response, and a quantification of the delivery efficiency was not shown.

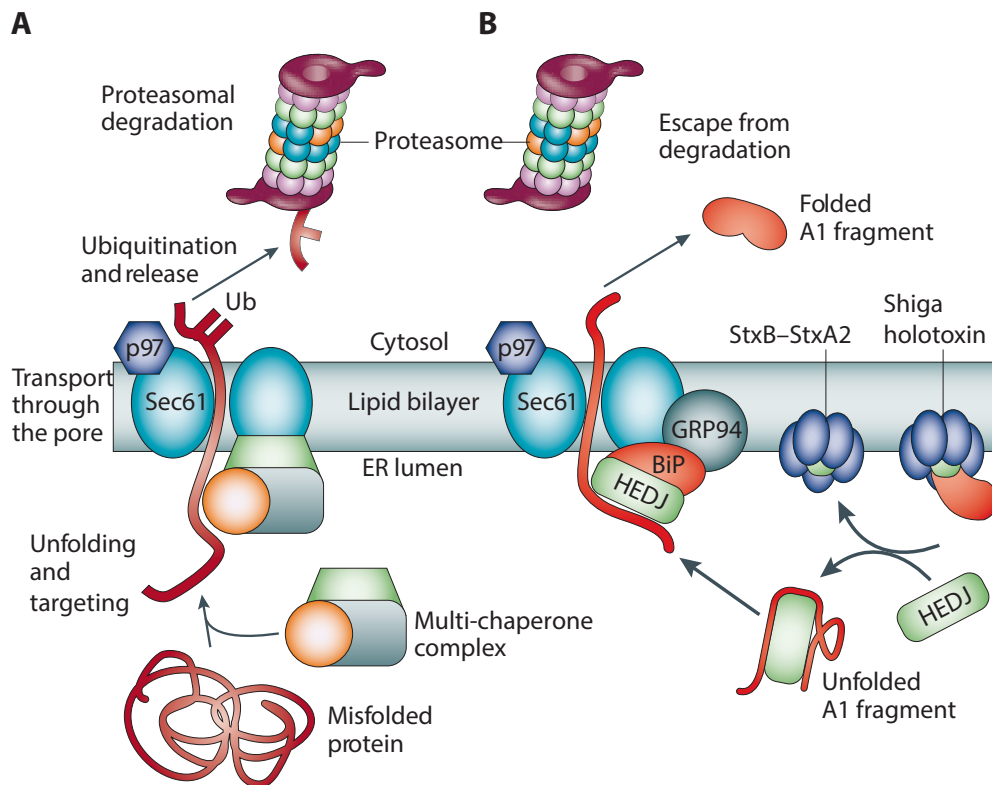


Figure 1.9 **The Endoplasmic Reticulum-Associated Degradation (ERAD) pathway.** **A:** misfolded proteins are unfolded by a multi-chaperone complex and translocated to the cytosol through the Sec61-associated translocon channel. In the cytosol, they are ubiquitinated and degraded. **B:** Following cleavage by furin, the Shiga toxin A subunit is released from the B subunit following disulfide bond reduction. The A subunit associates with host ER chaperones and is translocated into the cytosol through the Sec61-associated translocon channel. In the cytosol, it avoids proteasomal degradation, presumably due to a lack of lysine residues. (Figure from <sup>172</sup>)

## 1.6 Targeted degradation

Methods to remove distinct endogenous proteins from the proteome in a targeted way are of great utility for research and therapy. Currently, this is mostly achieved through modifications of the genome or transcriptome, either by using DNA knockout, DNA editing or RNA interference. Recently, endogenous proteins have been directly targeted for degradation by making use of the cellular ubiquitination and proteasomal degradation machineries.

Fulcher et al have developed the so-called affinity-directed protein missile system (AdPROM), which directs an E3 ubiquitin ligase to the protein of interest by means of alternative scaffold protein binders, like monobodies.<sup>200,201</sup> The monobody, or other protein binder of choice, is genetically fused to the Von-Hippel-Lindau (VHL) protein, and the construct is introduced into the target cells by transfection or viral transduction. VHL is the substrate receptor for the E3 ubiquitin ligase complex composed of Cullin 2, the RING-box protein 1 (RBX1) E3 ubiquitin ligase and the adaptor proteins Elongin B and C.<sup>202</sup> Upon expression of the VHL-monobody fusion, this complex will be recruited to the monobody target, which induces its ubiquitination by the RBX1 E3 ubiquitin ligase and subsequent degradation by the proteasome.<sup>200</sup> (Figure 1.10 A)

The authors have demonstrated the efficacy of this system for the degradation of the phosphatase SHP-2 by monobodies as well as ASC (Apoptosis-associated speck-like protein containing a CARD), a common adaptor protein of inflammasomes, using a nanobody.<sup>201</sup>

The PROTAC (proteolysis targeting chimera) approach is also harnessing the VHL-ElonginB/C-Cullin2 complex for targeted protein degradation, however, PROTACS are composed of two small molecules covalently linked via a PEG linker. One small molecule is a VHL-ligand and the other can in principle be any drug or other small molecule binder to a target protein of choice.<sup>203</sup> By binding to VHL and the target protein at the same time, the PROTAC brings them in close proximity and triggers the recruitment of the ElonginB/C-Cullin2 E3 ligase complex, inducing the ubiquitination and degradation of the target protein. Various PROTACs were constructed using thalidomide and its derivatives that bind and activate the cereblon substrate adaptor of the DDB1-Cullin4-X-box E3 ligase complex.<sup>204</sup> These PROTACs were used for the targeted degradation of a variety of substrates, such as Bcr-Abl kinase and the epigenetic regulator BRD4, which led to a suppression of the transcription factor myc.<sup>204</sup> In contrast to the AdPROM approach, PROTACs are not degraded and can subsequently engage in another round of target binding. The PROTAC approach has been applied to a variety of target proteins, including receptor tyrosine kinases.<sup>205</sup> (Figure 1.10 B)

Even though the AdPROM approach is targeting endogenous proteins for degradation, there remains the need to genetically modify the cells and the method cannot be applied to non-dividing primary cells where genome-editing is limited.

In order to overcome this limitation, Clift et al. have developed a method to degrade endogenous proteins by microinjecting antibodies against the target protein.<sup>206</sup> They make use of TRIM21, an E3 ubiquitin ligase expressed in diverse cell types that binds with high affinity to the Fc domain of antibodies.<sup>207</sup> The targeted protein is degraded within minutes following antibody injection into cells.<sup>206</sup> While in few cells, endogenous TRIM21 levels are sufficient to induce the ubiquitination and degradation of the target proteins, the authors had to overexpress TRIM21 in many cell lines in order to degrade the target protein efficiently. They reasoned that the antibody and TRIM21 protein become degraded at the same time as the target protein and therefore need to be present in excess, which can be achieved either by overexpressing TRIM21 genetically or by microinjecting recombinant TRIM21 protein along with the antibody.<sup>206</sup> (Figure 1.10 C)

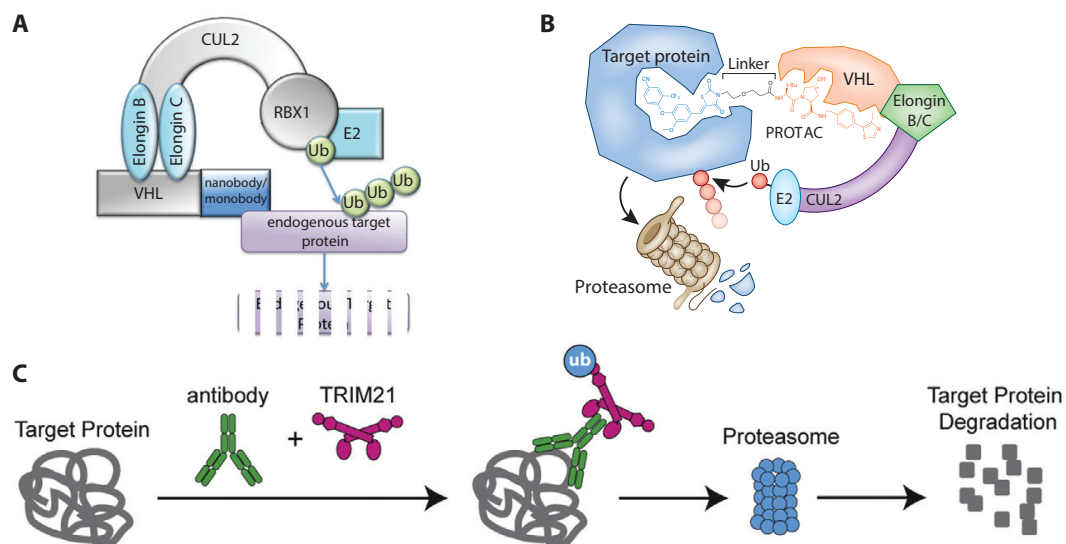


Figure 1.10 **Methods for targeted degradation.** **A:** AdPROM approach: protein binder(monobody/nanobody) fused to VHL recruits the ElonginB/C-CUL2 E3 ubiquitin complex to the target protein, which gets ubiquitinated and degraded. **B:** PROTAC approach working in the same way as the AdPROM approach, except that a fusion small molecule (PROTAC) composed of a VHL-binder and a target protein-binder brings the target protein in close proximity to VHL, leading to its ubiquitination and degradation. **C:** TRIM-away approach: An antibody binding a target protein is delivered by electroporation and recruits TRIM21, inducing the ubiquitination and degradation of the target protein. (Figures modified from <sup>201,206,208</sup>)



## 2 Results

### 2.1 Monobodies targeting Src-family kinases (SFKs)

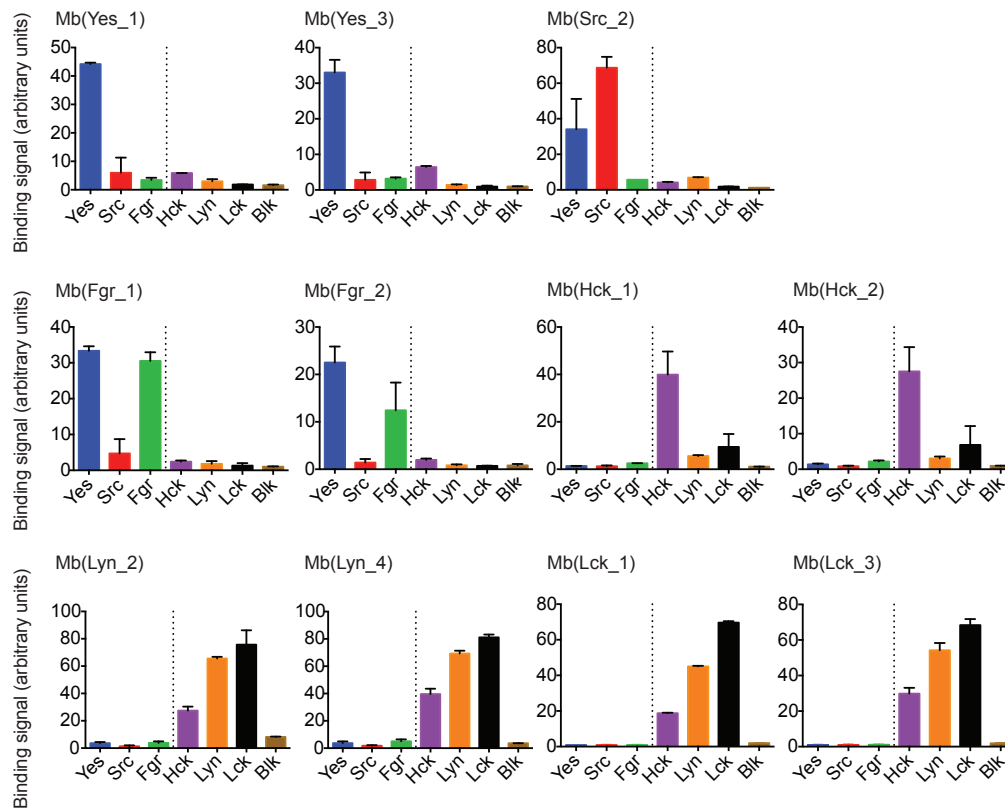
The aim of this project was to develop and characterize monobodies that disrupt the protein-protein interaction between SFKs and their substrates in a kinase-specific manner. Active SFKs bind to existing phospho-tyrosine residues in many of their substrates via the SH2 domain, enabling the phosphorylation of other critical tyrosine residues of the substrate. Small molecule ATP-competitive kinase inhibitors to SFKs are poorly selective and suffer from the rapid development of resistance through mutations in the binding site.<sup>2</sup> Several SH2-domain targeting pY-peptides, peptidomimetics and small molecules have been developed, but they are unable to discriminate between the individual SFK SH2 domains or have not been tested against all family members.<sup>209</sup> Therefore, selective inhibitors are needed to study the regulation and activity of individual SFKs in cellular signaling and to inhibit SFKs involved in oncogenic signaling pathways driving cancer cells. Monobodies were developed to bind the SH2 domain of different SFKs, due to the central role of this pY-binding domain in substrate recognition. Parts of the results of this project have been published in the *Journal of Molecular Biology*, to which I contributed as a co-first author.<sup>210</sup>

#### 2.1.1 Selection of monobodies binding the SH2 domains of Src-family kinases

The expression and purification of recombinant SFK SH2 domains and the selection of monobody binders was done by Dr. Tim Kükenshöner, a postdoc in our lab, during his visit in the laboratory of Prof. Shohei Koide, who has invented monobodies and has developed specific binders to a large variety of targets. Monobodies were selected from the “loop-only” and “side-and-loop” libraries against the 6 different SH2 domains of Yes, Src, Fgr, Hck, Lyn, and Lck, using phage and yeast display. Of each enriched pool of binders from the selection, 2 clones were characterized. Monobody clones were named Mb(SFK\_x) throughout the publication, where x is the clone number. I keep this terminology only for this section (2.1.1) and the Figure 2.1 and Figure 2.2. For simplicity reasons, Mb(Lck\_1) and Mb(Lck\_3) will be named ML1 and ML3, respectively, the monobodies targeting Src will be named MS2 and MS8 and the monobodies targeting Yes will be named MY1 and MY3 in the following sections.

The binding affinities of the different monobodies were determined using isothermal titration calorimetry (ITC) (Figure 2.1) as well as a yeast binding assay (Figure 2.2). Each monobody showed a higher affinity for the SH2 domain used in the selection than for the other SH2 domains. Moreover, in the yeast binding assay, monobodies selected against Yes, Src or Fgr (SrcA family) showed weak or no binding to the members of the SrcB family (Hck, Lyn, Lck and Blk) and the monobodies selected against Lyn, Lck or Hck did not bind to members of the SrcA family (Yes, Src, Fyn and Fgr). (Figure 2.1) For the monobodies selected against Lck, no binding could be measured to Src or Yes by ITC either, whereas Mb(Yes\_1) is binding to Lck, although with a much weaker affinity

than to Yes. (Figure 2.2) These results showed a high sub-family specificity, which had previously not been achieved with other SFK inhibitors.



**Figure 2.1 Yeast binding assay of the selected monobodies to the SH2 domain of the different SFKs.** Each panel shows the mean fluorescence signal of one monobody toward all tested targets at a concentration of 250 nM. For clarification, SrcA (Yes, Src, Fgr) and SrcB (Hck, Lyn, Lck, Blk) family members are subdivided by a dotted line. Each data point corresponds to the average of two repeats  $\pm$  SD. The data and graph were generated by Dr. Tim Kükenshörer and are taken from <sup>208</sup>.

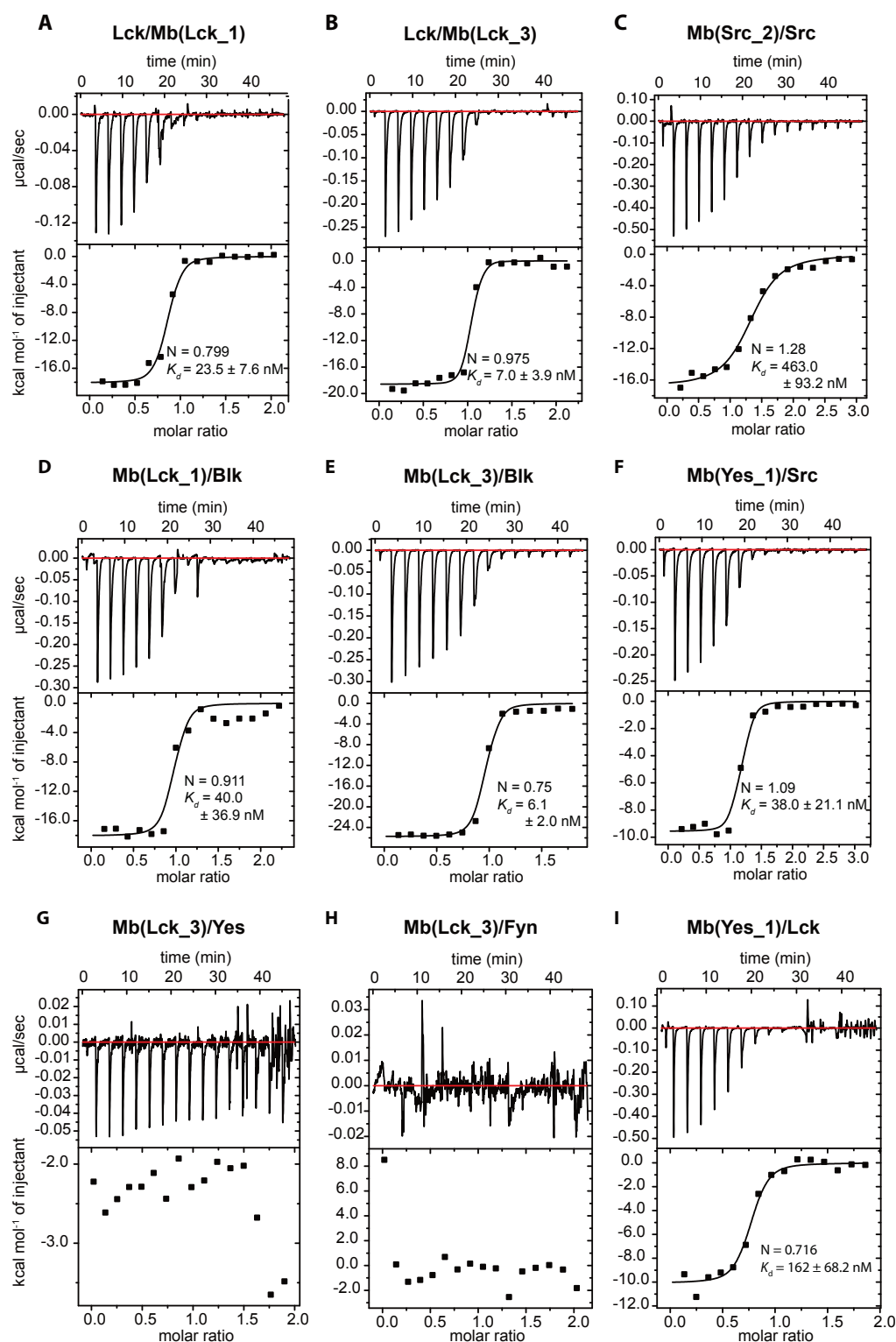


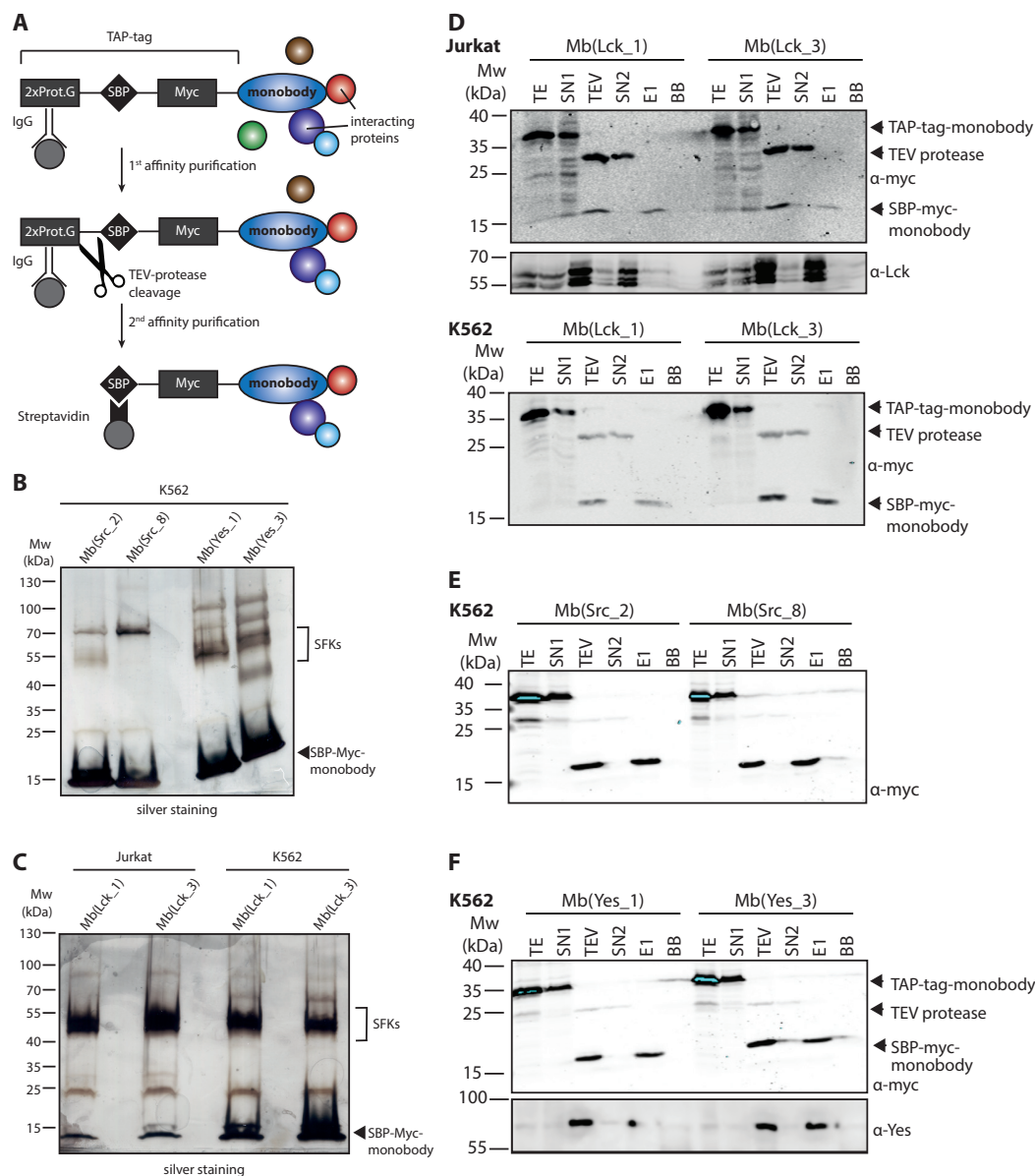
Figure 2.2 ITC measurements of different monobodies with SH2 domains. All calorimetric titrations of the monobodies with SH2 domain were performed at 25 °C. Each panel shows (at the top) the raw heat signal of an ITC experiment. The bottom panel shows the integrated calorimetric data of the area of each peak. The continuous line represents the best fit of the data based on a 1:1 binding model computed from the MicroCal software. A representative measurement is shown for each example with K<sub>D</sub> value and stoichiometry (N) calculated from the fit. **A:** Mb(Lck\_1) (70 µM) titrated to Lck (7 µM), **B:** Lck (115 µM) titrated to Mb(Lck\_3) (11 µM), **C:** Mb(Src\_2) (180 µM) titrated to Src (12 µM), **D:** Blk (142 µM) titrated to Mb(Lck\_1) (13 µM), **E:** Mb(Lck\_3) (150 µM) titrated to Blk (15 µM), **F:** Mb(Yes\_1) (215

$\mu\text{M}$ ) titrated to Lck (15  $\mu\text{M}$ ), **G:** Mb(Lck\_3) (140  $\mu\text{M}$ ) titrated to Yes SH2 domain (15  $\mu\text{M}$ ), **H:** Mb(Lck\_3) (150  $\mu\text{M}$ ) titrated to Fyn SH2 domain (15  $\mu\text{M}$ ) and **I:** Mb(Yes\_1) (400  $\mu\text{M}$ ) titrated to Lck SH2 domain (40  $\mu\text{M}$ ). The data was generated by Dr. Tim Kükenshöner and the figure is modified from <sup>210</sup>.

## 2.1.2 Tandem affinity purification to determine monobody interactors

In order to assess the selectivity of the monobodies in a complex cellular environment containing other SH2-domain containing proteins and other possible unrelated off-target proteins, I performed a tandem affinity purification (TAP) followed by tandem mass spectrometry of monobodies stably expressed in cell lines.<sup>211</sup>

I cloned the cDNAs of monobodies MY1, MY3, MS2, MS8, ML1, ML3 into a vector containing the TAP tag, which consists of two copies of the B1 domain of staphylococcal protein G (2xProtein G), a TEV protease recognition site, a streptavidin-binding peptide (SBP), and a Myc-tag N-terminal to the monobody.<sup>211</sup> (Figure 2.3 A) The chronic myeloid leukemia cell line K562 as well as the human embryonic kidney (HEK293) cell line were chosen to generate cell lines stably expressing the monobodies MY1, MY3, MS2 and MS8, and K562 cells as well as the T-cell acute lymphoblastic leukemia cell line Jurkat were transduced with the TAP-tagged ML1 and ML3 monobodies. The stable cell lines were generated by retroviral transduction and then sorted based on their expression of GFP, which is encoded on the same vector separated by an internal ribosome entry site (IRES). After confirming the expression of the monobody proteins, I expanded each cell line to a total number of  $3 \times 10^9$  cells, of which half were taken for each TAP experiment. A TAP was then performed with the lysates of the K562 cell lines expressing the monobodies MY1, MY3, MS2, MS8, ML1 and ML3 and with the Jurkat cell lines expressing ML1 and ML3. The HEK293 cell pellets were frozen and stored for a back-up analysis if necessary. An immunoblot analysis against the myc-tag confirmed that the bait monobody can be followed throughout the different steps of the TAP and can be retrieved after the second purification step. (Figure 2.3 D, E and F). A silver gel analysis of the whole elution fraction after the second purification showed only few predominant bands, indicating that only the main interactors were pulled down together with the monobodies. (Figure 2.3 B and C) The strongest bands correspond to proteins of ~58 kDa, which is the size of SFKs, as well as a band of 15 kDa, corresponding to the bait monobody. Additionally, an immunoblot analysis was performed to confirm the presence of the monobody targets in the pull-down. (Figure 2.3 D and F) The total elution fraction after the second purification was then separated on an SDS-PAGE gel and in-gel digestion with Trypsin was performed prior to analysis by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) at the EPFL proteomics core facility.



**Figure 2.3 SFK SH2 monobody interactome analysis by tandem affinity purification (TAP).** **A:** Schematic representation of a TAP-tagged monobody and the purification method including two copies of the B1 domain of staphylococcal protein G (2xProt.G), TEV protease recognition site, streptavidin-binding peptide (SBP), and Myc-tag N-terminal to the monobody. **B:** MS2, MS8, MY1 and MY3 monobody complexes after TAP (10% of the E1 fractions) from K562 cells, separated by SDS-PAGE and visualized with silver staining. **C:** ML1 and ML3 monobody complexes after TAP (10% of the E1 fractions) from K562 and Jurkat cells, separated by SDS-PAGE and visualized with silver staining. **D:** Immunoblot analysis of TAP of ML1 and ML3 monobody complexes from Jurkat cells (top) and K562 cells (bottom). The bait protein and the main target were identified by immunoblotting using an anti-Myc or anti-Lck antibody, respectively. **E:** Immunoblot analysis of TAP of MS2 and MS8 monobody complexes from K562 cells. The bait protein was identified by immunoblotting using an anti-Myc antibody. **F:** Immunoblot analysis of TAP of MY1 and MY3 monobody complexes from K562 cells. The bait protein and the main target were identified by immunoblotting using an anti-Myc or anti-Yes antibody, respectively. TE, total extract; SN1, supernatant IgG beads; TEV, eluate after TEV cleavage; SN2, supernatant streptavidin beads; E1, eluate from streptavidin beads; BB, boiled streptavidin beads to control the efficiency of elution. TAPs with monobodies ML1 and ML3 in K562 and Jurkat cells were done twice, and TAPs with monobodies Mb(Src\_2), Mb(Src\_8), Mb(Yes\_1) and Mb(Yes\_3) were done once. The figure was modified from <sup>210</sup>.

We later found that MS2 and MS8 have the same sequence, as all the 8 clones selected among the pool of binders to the Src SH2 domain. We therefore continued the analysis with only the MS2 sample. Table 2.1 shows a summary of the results of the mass spectrometry analysis. The SrcA subgroup members Src, Yes and Fyn were identified among the most abundant proteins bound by the monobodies MY1, MY3 and MS2 in K562 cell lines. In the cells expressing the Yes-binding monobodies, Lyn, Btk and Tec were the only other SH2-domain containing proteins detected to bind to MY1 or MY3, although at much lower total spectrum counts than Src, Yes and Fyn. ML1 and ML3 bound predominantly to Lck in Jurkat cells and to Lyn in K562 cells. In the first replicate of the experiment, Lck and Lyn were the only SFKs detected. In the second replicate of this experiment, the total number of detected proteins as well as their total spectrum counts were much higher. The SFKs Hck and Src were detected along with Lck and Lyn, although with total spectrum counts which were more than one order of magnitude lower than for Lck in Jurkat and Lyn in K562 cells. For Lck and Src targeting monobodies, no other SH2-domain containing proteins were detected which do not belong to the Src kinase family. (Table 2.1)

belong to the Src kinase family). (Table 2.1)

Cell line	Bait	total number of proteins identified*		number of proteins meeting selection criteria**		All identified SH2 domain-containing proteins in each sample (in brackets: total spectrum counts, rank***)	
		1 <sup>st</sup> TAP	2 <sup>nd</sup> TAP	1 <sup>st</sup> TAP	2 <sup>nd</sup> TAP	1 <sup>st</sup> TAP	2 <sup>nd</sup> TAP
K562	ML1	68	252	16	50	Lyn (112, 2); Lck (7, 34)	Lyn (1523, 1); Hck (141, 13); Lck (133, 15); Src (44, 38)
	ML3	31	335	6	31	Lyn (53, 1)	Lyn (1614, 1); Hck (172, 9); Lck (130, 15); Src (45, 32)
Jurkat	ML1	33	285	3	23	Lck (114, 1); Lyn (1, 22)	Lck (1811, 1); Src (210, 6); Lyn (138, 13); Hck (131, 16)
	ML3	25	275	6	19	Lck (109, 1); Lyn (1, 19)	Lck (1930, 1); Lyn (177, 6); Hck (162, 7); Src (159, 9)
K562	MS2	39		15		Src (35, 7); Yes1 (31, 8); Fyn (31, 8)	
	MY1	46		13		Yes1 (101, 4); Fyn (76, 6); Src (61, 7); Lyn (12, 20); Tec (6, 27); Btk (1, 47)	
	MY3	259		75		Yes1 (295, 1); Fyn (139, 6); Src (107, 11); Lyn (32, 28); Btk (4, 105)	

**Table 2.1 Overview of mass spectrometry results from the interactome analysis of monobodies MS2, MY1, MY3, ML1 and ML3** \* Proteins were scored as identified if at least 2 unique peptides were detected. The complete list of identified proteins is provided in Supplementary Tables 1-3. \*\* Selection criteria: Proteins with a selectivity score of >0.75 (present in <25% of samples analyzed), with a total spectrum count >1% of the most abundant protein in the sample, excluding different isoforms of keratin and the bait protein itself. \*\*\* All identified proteins, including common contaminants, such as Keratin, were sorted by the number of assigned spectra from highest to lowest.

Together with the very few SH2-domain containing proteins, many other proteins were identified by the mass spectrometry analysis, the total number of proteins detected ranging from 25 to 335. Most of these proteins were detected in the majority of all the

TAP experiments performed in our lab by me or other lab members. These commonly found proteins can be contaminants such as keratins or heat-shock proteins, highly abundant proteins that are difficult to eliminate completely during the purifications or proteins which bind unspecifically to all monobodies, affinity beads or parts of the TAP-tag. By comparing all the protein lists from our lab database of TAP experiments, I calculated for each protein the complement of the percentage of TAP experiments where it was detected. We called this number selectivity score because it is a measure of the uniqueness of a protein in a given TAP experiment and indicates whether it is likely to be a specific interactor of the monobody or its target protein. The following formula was used:  $S=1-(1/T \times P)$ , where T is the total number of TAP experiments and P is the number of experiments where the protein is detected. For the analysis, I then only kept proteins with a selectivity score >0.75, meaning that it is present in <25% of all TAP experiments and with a total spectrum count of ≥10% of the most abundant proteins in the sample in order to exclude very low abundant proteins and excluded common contaminants such as keratin. Table 2.2 A, B and C contain the complete lists of proteins meeting these criteria that were detected and their total spectrum counts in each TAP.

Identified Proteins	Gene Name	K562 MS2	K562 MY1	K562 MY3	Selectivity score
<b>Proto-oncogene tyrosine-protein kinase Src</b>	<b>SRC</b>	<b>40</b>	<b>66</b>	<b>114</b>	<b>0.89</b>
<b>Non-specific protein-tyrosine kinase Yes</b>	<b>YES1</b>	<b>36</b>	<b>111</b>	<b>311</b>	<b>0.96</b>
<b>Tyrosine-protein kinase Fyn</b>	<b>FYN</b>	<b>36</b>	<b>82</b>	<b>147</b>	<b>0.78</b>
Cilia- and flagella-associated protein 54	CFAP54	21	13	8	0.91
Monobody		20	0	11	
40S ribosomal protein S3	RPS3	8	16	16	0.88
Polyubiquitin-B	UBB	6	5	75	0.91
Collagen alpha-1(III) chain	COL3A1	6	0	0	0.96
Collagen alpha-1(I) chain	COL1A1	5	1	2	0.84
Tubulin alpha-1A chain	TUBA1A	3	10	38	0.76
Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2	0	0	52	0.99
Heat shock 70 kDa protein 1-like	HSPA1L	0	0	39	0.78
Hsp90 co-chaperone Cdc37	CDC37	0	11	32	0.89
<b>Isoform 2 of Tyrosine-protein kinase Lyn</b>	<b>LYN</b>	<b>0</b>	<b>13</b>	<b>32</b>	<b>0.8</b>

Table 2.2 A: Total spectrum counts of the proteins identified by mass spectrometry during the TAP of stable K562 cell lines expressing the bait proteins MS2, MY1 or MY3. Only the proteins that met our criteria described in the text are shown as well as their selectivity score. Proteins containing an SH2 domain are highlighted in bold.

Identified Proteins	Gene Name	Jurkat ML1	Jurkat ML3	K562 ML1	K562 ML3	Selectivity score
<b>Tyrosine-protein kinase Lck</b>	<b>LCK</b>	<b>112</b>	<b>109</b>	<b>7</b>	<b>0</b>	<b>0.89</b>
<b>Tyrosine-protein kinase Lyn</b>	<b>LYN</b>	<b>1</b>	<b>1</b>	<b>107</b>	<b>44</b>	<b>0.8</b>
Monobody		11	11	97	59	
CAD protein	CAD	2	0	15	0	0.75
Heat shock 70 kDa protein 1-like	HSPA1L	0	0	14	0	0.78
CUB domain-containing	CDCP1	0	0	27	6	0.94

protein 1						
Talin-1	TLN1	0	0	9	6	0.89
Diacylglycerol kinase theta	DGKQ	0	0	0	4	0.97
Heat shock 70 kDa protein 1-like	HSPA1L	0	0	14	0	0.78

**B: Total spectrum counts of the proteins identified by mass spectrometry during the 1<sup>st</sup> TAP of stable Jurkat and K562 cell lines expressing the bait proteins ML1 or ML3.** Only the proteins that met our criteria described in the text are shown as well as their selectivity score. Proteins containing an SH2 domain are highlighted in bold.

Identified Proteins	Gene name	Jurkat ML1	Jurkat ML3	K562 ML1	K562 ML3	Selectivity score
<b>Tyrosine-protein kinase Lck</b>	<b>LCK</b>	<b>1811</b>	<b>1930</b>	<b>133</b>	<b>130</b>	<b>0.89</b>
<b>Tyrosine-protein kinase Lyn</b>	<b>LYN</b>	<b>138</b>	<b>180</b>	<b>1523</b>	<b>1614</b>	<b>0.8</b>
<b>Non-specific protein-tyrosine kinase Hck</b>	<b>HCK</b>	<b>131</b>	<b>162</b>	<b>141</b>	<b>172</b>	<b>0.94</b>
<b>Proto-oncogene tyrosine-protein kinase Src</b>	<b>SRC</b>	<b>210</b>	<b>159</b>	<b>44</b>	<b>45</b>	<b>0.89</b>
Monobody		485	225	1245	1780	0.94
Tubulin alpha-1B chain	TUBA1B	168	122	117	150	0.78
Tubulin beta-2A chain	TUBB2A	0	115	143	152	0.96
Protein unc-119 homolog A	UNC119	114	102	0	0	0.94
Protein unc-119 homolog B	UNC119B	64	69	0	0	0.96
Hsp90 co-chaperone Cdc37	CDC37	66	52	23	22	0.89
Sodium/potassium-transporting ATPase subunit $\alpha$ -1	ATP1A1	41	49	4	2	0.94
Translational activator GCN1	GCN1L1	44	42	62	26	0.81
Ubiquitin carboxyl-terminal hydrolase 22	USP22	42	38	0	0	0.96
Tubulin beta-6 chain	TUBB6	0	36	51	44	0.92
4F2 cell-surface antigen heavy chain	SLC3A2	20	25	14	11	0.92
Retinol dehydrogenase 11	RDH11	22	21	0	0	0.97
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	RPN1	21	21	2	1	0.92
Isoform 2 of Extended synaptotagmin-2	ESYT2	17	21	1	0	0.96
Transferrin receptor protein 1	TFRC	13	21	15	7	0.94
Heat shock 70 kDa protein 1A	HSPA1A	30	18	113	119	0.88
Very-long-chain enoyl-CoA reductase	TECR	26	18	0	0	0.97
Exportin-2	CSE1L	17	17	20	14	0.8
Bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase, mitochondrial	MTHFD2	28	15	19	16	0.94
Exportin-1	XPO1	17	14	51	50	0.89
RNA-binding protein FUS	FUS	18	12	21	10	0.84
Polyadenylate-binding protein 1	PABPC1	21	11	26	29	0.86
60S acidic ribosomal protein P0	RPLP0	13	11	33	22	0.88
Leucine-rich PPR motif-containing protein, mitochondrial	LRPPRC	11	11	34	20	0.8
40S ribosomal protein S8	RPS8	11	11	18	10	0.83
Serine palmitoyltransferase 1	SPTLC1	10	10	19	28	0.94
60S ribosomal protein L4	RPL4	8	10	30	17	0.94
60S ribosomal protein L18 (Fragment)	RPL18	7	9	25	20	0.94

Isoleucine--tRNA ligase, cytoplasmic	IARS	1	9	28	11	0.84
T-complex protein 1 subunit epsilon	CCT5	8	7	17	6	0.76
Trifunctional enzyme subunit alpha, mitochondrial	HADHA	7	7	26	21	0.86
Transportin-1	TNPO1	13	6	37	32	0.94
Isoform B of Phosphate carrier protein, mitochondrial	SLC25A3	26	5	9	3	0.89
60S ribosomal protein L6	RPL6	16	4	27	17	0.86
60S ribosomal protein L3	RPL3	10	4	22	4	0.86
60S ribosomal protein L7a	RPL7A	9	4	13	17	0.83
Vimentin	VIM	2	4	28	42	0.84
60S ribosomal protein L15	RPL15	6	3	18	11	0.94
Importin subunit beta-1	KPNB1	4	3	19	14	0.91
Importin-7	IPO7	0	1	16	14	0.94
Tubulin beta-3 chain	TUBB3	89	0	104	107	0.91
Heat shock-related 70 kDa protein 2	HSPA2	47	0	42	50	0.88
ATP-dependent RNA helicase A	DHX9	5	0	22	13	0.88
X-ray repair cross-complementing protein 6	XRCC6	5	0	16	9	0.96
Clathrin heavy chain	CLTC	1	0	34	26	0.92
Large proline-rich protein BAG6	BAG6	1	0	35	22	0.88
Arginine--tRNA ligase, cytoplasmic	RARS	1	0	16	11	0.96
Talin-1	TLN1	0	0	22	62	0.89
CUB domain-containing protein 1	CDCP1	0	0	106	61	0.94
Peroxisomal multifunctional enzyme type 2	HSD17B4	0	0	25	18	0.78
Annexin A1	ANXA1	0	0	17	9	0.83
Leucine--tRNA ligase, cytoplasmic	LARS	0	0	16	6	0.96
Kinesin-like protein KIF11	KIF11	0	0	16	6	0.92
Serine/threonine-protein kinase mTOR	MTOR	0	0	26	5	0.86
RNA-binding protein 4	RBM4	0	0	29	3	0.97

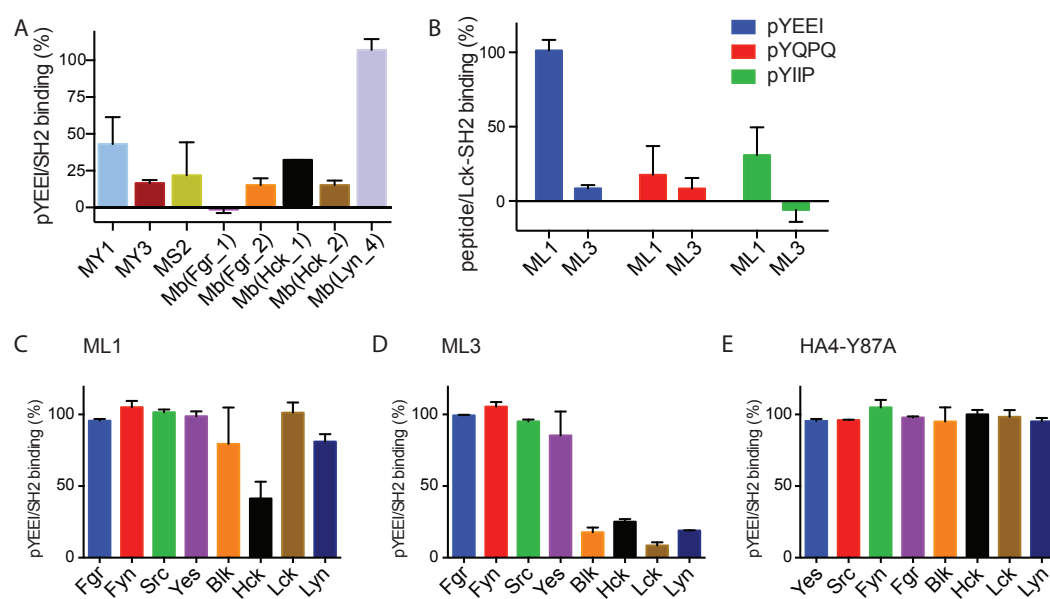
**C: Total spectrum counts of the proteins identified by mass spectrometry during the 2<sup>nd</sup> TAP of stable Jurkat and K562 cell lines expressing the bait proteins ML1 or ML3.** Only the proteins that met our criteria described in the text are shown as well as their selectivity score. Proteins containing an SH2 domain are highlighted in bold.

Altogether, the results demonstrate the ability of these monobodies to bind with high selectivity to their full-length target protein in cells. They only bind to few other SFK SH2 domains and 2 closely related kinases, but no other major off-targets.

### 2.1.3 Inhibition of SH2-pY interactions by the selected monobodies

We next wanted to assess if the monobodies not only bind SH2 domains but also inhibit their interaction with pY-containing substrates. We first determined whether they could block the binding to a peptide containing the pYEEI motif, a known binding motif of SFK SH2 domains. The pY peptide competition assays were performed by Tim Kükenshöner using recombinant SH2 domains and monobodies previously purified. All the tested monobodies competed with the pYEEI peptide for binding to their on-target SH2 domain, except for monobodies ML1, Mb(Lyn\_4) and a control non-binding

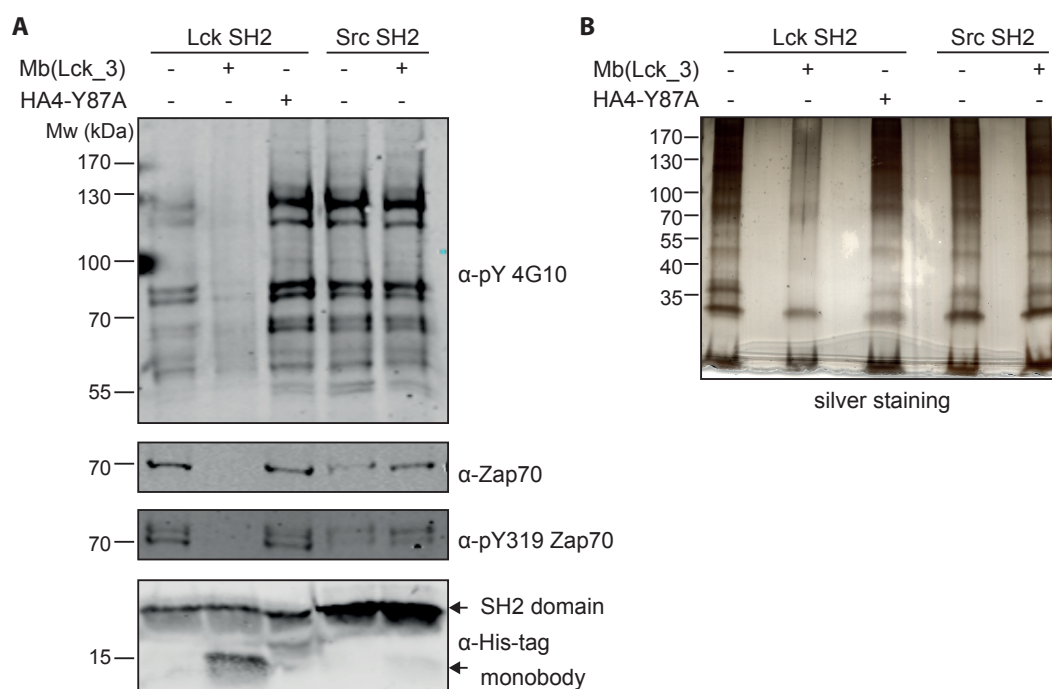
monobody HA4-Y87A (HA4\_YA)<sup>76</sup>. (Figure 2.4 A and E) To further study this discrepancy between ML3 and ML1, their competition against two additional pY peptides with lower affinities for the Lck SH2 domain, pYIIP and pYQPQ, was tested. Remarkably, both Lck monobodies displaced these peptides, suggesting that they are both able to inhibit protein-protein interactions of the Lck SH2 domain. (Figure 2.4 B) ML1 was also not able to compete with the pYEEI peptide binding to any of the other purified SH2 domains, while ML3 displaced the pYEEI peptide from all the SrcB family SH2 domains but none of the SrcA family SH2 domains. (Figure 2.4 C and D) This result is in line with the previous results from the ITC and yeast binding assay and underlines the subgroup selectivity of ML3.



**Figure 2.4 Inhibition of pY peptide /SH2 interaction by monobodies as measured by fluorescence polarization assay.** **A:** Relative pYEEI peptide binding (in %) to the main target SH2 domain in the presence of the indicated monobody. All eight SH2 domains have been measured without and in presence of the monobody selected for the respective on-target. The pYEEI peptide in isolation and the SH2/pYEEI complex were set to 0% and 100% binding, respectively. Each data point corresponds to the average of at least two repeats  $\pm$  SD. **B:** Relative peptide binding (in %) of three different peptides (pYEEI, pYQPQ, pYIIP) to the Lck SH2 domain in the presence of each of the two Lck monobodies ML1 or ML3. **C:** Relative pYEEI peptide binding (in %) to the indicated SH2 domains in the presence of ML1. **D:** Relative pYEEI peptide binding (in %) to the indicated SH2 domains in the presence of ML3. **E:** Relative pYEEI peptide binding (in %) to the indicated SH2 domains in the presence of HA4\_Y87A. The data was generated by Dr. Tim Kükenshöner and the figure is modified from <sup>210</sup>.

I then sought to study the potential of the developed monobodies as functional antagonists of SH2 domain mediated protein-protein interactions with full-length proteins rather than isolated pY peptides. I therefore did a pull-down assay on T-cell receptor-stimulated Jurkat cells with recombinant Lck SH2 domain. Jurkat cells were stimulated with a T-cell receptor binding antibody prior to lysis. I then immobilized recombinant biotinylated SH2 domain of Lck or Src on Streptavidin-coated beads and added either ML3 or the control monobody HA4-Y87A<sup>76</sup> prior to incubation with the cell lysate. Interacting proteins were analyzed on a silver stained SDS-PAGE gel as well as by immunoblotting against tyrosine phosphorylated proteins. (Figure 2.5) Both analyses showed a substantially reduced number of interactors when ML3 was added as compared

to the SH2 domain alone or after addition of the control monocbody HA4\_YA. In contrast, the Src-SH2 domain bound to the same amount of tyrosine-phosphorylated proteins, whether recombinant ML3 was added or not. In particular, the interaction with a known interactor of Lck, Zap70, was disrupted by ML3 but not the control monocbody. ML3 specifically blocks the interaction of the Lck SH2 domain with Zap70 phosphorylated on Y319, which is known to be responsible for the binding to Lck. These results show that ML3 is a selective competitor of SH2 domain mediated protein-protein interactions of Lck in a complex cellular lysate.



**Figure 2.5 SH2 pull-down assay. A:** The monocbodies ML3 or HA4-Y87A or buffer alone were added to biotinylated Lck or Src SH2 domains immobilized on Streptavidin-coated beads prior to incubation with lysate from anti-TCR stimulated Jurkat T cells. Immunoblot analysis after pull-down using an anti-pY antibody for a 40% fraction of beads (upper blot), using an anti-Zap70 antibody and an anti-phospho-Zap70 (pY319) antibody for a 40% fraction of beads (second and third blot from top), and an anti-His-tag antibody to detect the recombinant SH2 domains and monocbodies for a 10% fraction of beads (lower blot). One representative immunoblot is shown out of 5 repeats. **B:** The remaining 10% fraction of beads was separated on an SDS-Page gel and visualized by silver staining. One representative silver stained gel is shown out of 5 biological repeats.

#### 2.1.4 The monocbody ML3 inhibits Lck kinase activity in cells

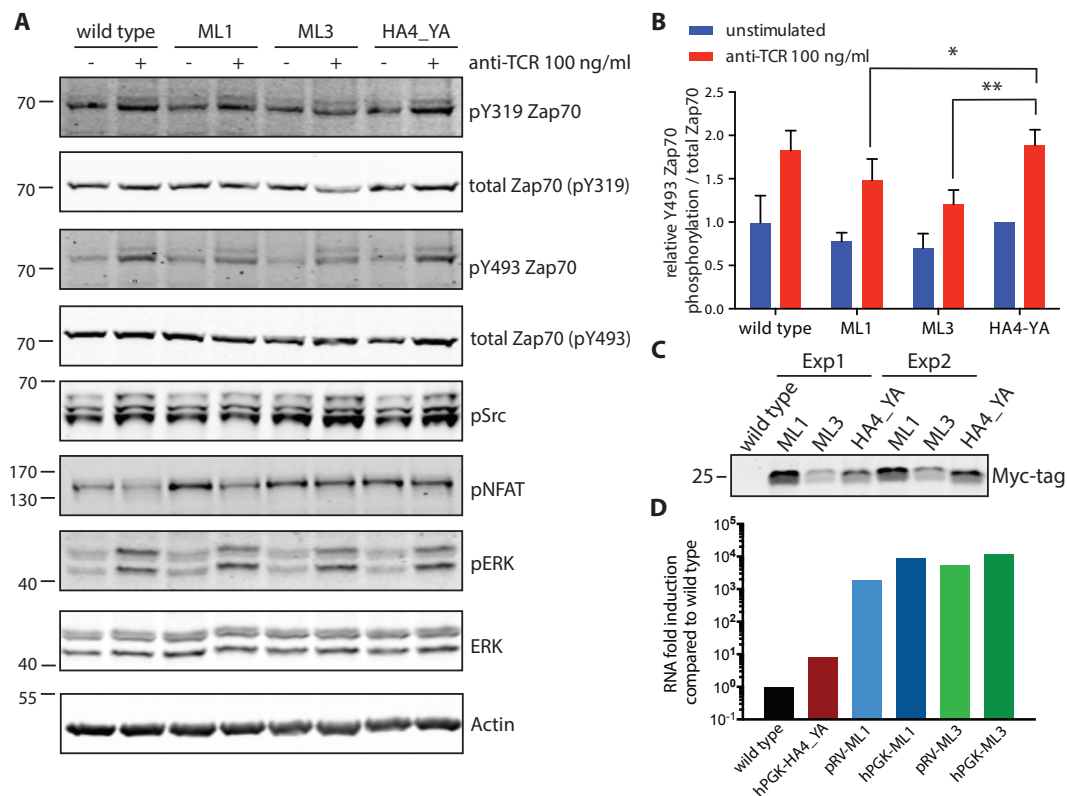
Finally, our aim was to test if the developed SFK SH2 monocbodies can interfere with SFK dependent signaling pathways in cells. We focused on the Lck-binding monocbodies since they showed the most remarkable ability to discriminate between the SrcA and SrcB subgroups. Since the Lck SH2 domain is indispensable for the activation of Zap70 kinase in the T-cell receptor complex and for T-cell activation, I assessed the effects of ML1 and ML3 on Zap70 phosphorylation.

In a first attempt, I used the same Jurkat cell lines stably expressing the monocbodies as for the TAP. However, I did not observe any effect on Zap70 phosphorylation in these cell lines, which is probably due to the comparatively low expression of the monocbodies

using the TAP-tagged vector. I then generated stable Jurkat cell lines expressing Myc-tagged ML1, ML3, or the control monobody HA4-Y87A using a lentiviral hPGK vector obtained from the laboratory of Prof. Jörg Huelsken at EPFL. This vector also contains an IRES followed by a GFP and following infection with the lentivirus generated by transfection of HEK293T cells with the constructs, the cells were sorted by FACS. I could confirm a higher expression of the monobodies using this vector by western blot against the Myc-tag (data not shown).

These cells were then stimulated with an anti-T-cell receptor (TCR) antibody to activate the T cell receptor pathway and the phosphorylation of the activation loop of Zap70 (Y493) as well as of Y319, which is in the Zap70 kinase-SH2 linker, were monitored by immunoblotting. (Figure 2.6 A and B) In non-transduced Jurkat cells as well as the cells expressing HA4-YA, I observed an increase in Zap70 phosphorylation on both tyrosine residues upon stimulation with the TCR antibody. (Figure 2.6 A and B) In contrast, in cells expressing ML3, TCR stimulation resulted in much weaker Zap70 phosphorylation, which was decreased as compared to the control cells even in non-stimulated cells. ML1 also reduced Zap70 phosphorylation, but showed milder effects, which is in line with its weaker pY competition activity. (Figure 2.6 A and B) This result shows that ML3 and ML1 can interfere with Lck-mediated signaling in T-cells and inhibit the activation of Zap70.

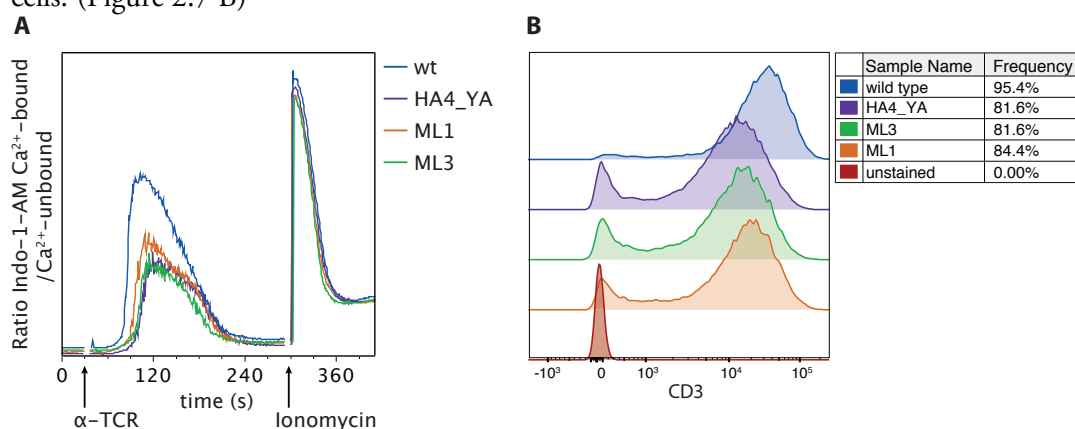
Surprisingly, I observed that albeit the more pronounced Lck inhibition by ML3 than by ML1, ML3 protein levels were much lower than ML1 or HA4\_YA protein levels. I observed this effect in several independent experiments where cell lysates prepared on different days were immunoblotted with an anti-myc antibody. (Figure 2.6 C) Interestingly, the mRNA levels of both ML1 and ML3 were similar, and 5 to 10-fold higher in cells transduced with the hPGK vector than with the pRV vector (TAP-tagged monobodies). This result suggests that the Jurkat cells may down-regulate protein levels of the ML3 monobody that inhibits TCR activation, by an unknown post-translational mechanism.



**Figure 2.6 ML3 or ML1 expression leads to a reduction in Zap70 phosphorylation upon TCR activation.** **A:** Jurkat cells expressing no monobody, ML1, ML3 or the control monobody HA4\_YA were stimulated with an anti-TCR antibody for 5 minutes and lysed. Immunoblot analysis of the cell lysate with antibodies against phosphorylated Y319 residue of Zap70 and total Zap70 on the same blot, against phosphorylated Y493 residue of Zap70 and total Zap70 on the same blot, against activation loop phosphorylated Src-family kinases, against phosphorylated NFAT, phosphorylated ERK and total ERK on the same blot and against Actin. (order as displayed, from top to bottom) **B:** Quantification of the ratio between the Y493-phosphorylated Zap70 and total Zap70 signals. The data was transformed by setting the pZap70:Zap70 ratio of the unstimulated HA4-Y87A control cells to 1.0 for each experiment. Average values and standard deviations from two independent experiments done in two technical replicates were used, and *p* values were calculated using a ratio paired t-test on the untransformed data. \**p* < 0.05, \*\**p* < 0.01. **C:** Immunoblot analysis of the expression levels of the myc-tagged monobodies from 2 different lysates of the same stable cell lines expressing the hPGK vector constructs. **D:** RT-PCR analysis of mRNA levels of expressed monobodies. The RT-PCR was performed with primers binding to ML1 and ML3 but not to HA4\_YA. The data was normalized to a GAPDH control run in parallel and the experiment was done in triplicates, of which the mean values are shown.

I also assessed the phosphorylation of proteins that are activated further downstream in the TCR activation pathway. The MAP kinase ERK was similarly phosphorylated upon TCR stimulation in cells expressing ML1 or ML3 or HA4\_YA, and the dephosphorylation of the transcription factor NFAT was similar in all cell lines as well. (Figure 2.6 A) I then set out to measure the calcium flux in the cells after stimulation with the anti-TCR antibody. The cells were loaded with the Ca<sup>2+</sup> sensitive Indo-1-AM dye, which upon excitation with UV-light displays a different emission spectrum when bound to Ca<sup>2+</sup> than in the unbound state. The ratio of the two emission maxima of Indo-1-AM-loaded cells was recorded on a flow cytometer and then the antibody was added to elicit the calcium flux response. Ionomycin, a TCR independent stimulant was added after the ratio had dropped back to the baseline level as a positive control. No difference was observed between the cells expressing ML1 or ML3 or the control monobody.

(Figure 2.7 A) Only wild type cells showed a more pronounced calcium flux response than the cells expressing a monobody, which is most likely due to the amount of TCR expressed on the cell surface, being lower in monobody expressing cells than in wild type cells. (Figure 2.7 B)



**Figure 2.7 Calcium flux measurement in Jurkat cells expressing no monobody, ML1, ML3 or HA4\_YA upon TCR stimulation.** **A:** Jurkat cells were loaded with Indo-1-AM, and the ratio between the 475 nm ( $\text{Ca}^{2+}$ -bound) and the 400 nm (not bound to  $\text{Ca}^{2+}$ ) signal was measured over time on a flow cytometer. The signal was measured for 30 seconds to set the baseline, then an anti-TCR antibody was added to the tube and the measurement was continued for 5 minutes, after which Ionomycin was added to the tube. One representative plot is shown out of 3 repeats. **B:** CD3 expression in the same Jurkat cells measured by flow cytometry. One representative plot is shown out of 2 repeats.

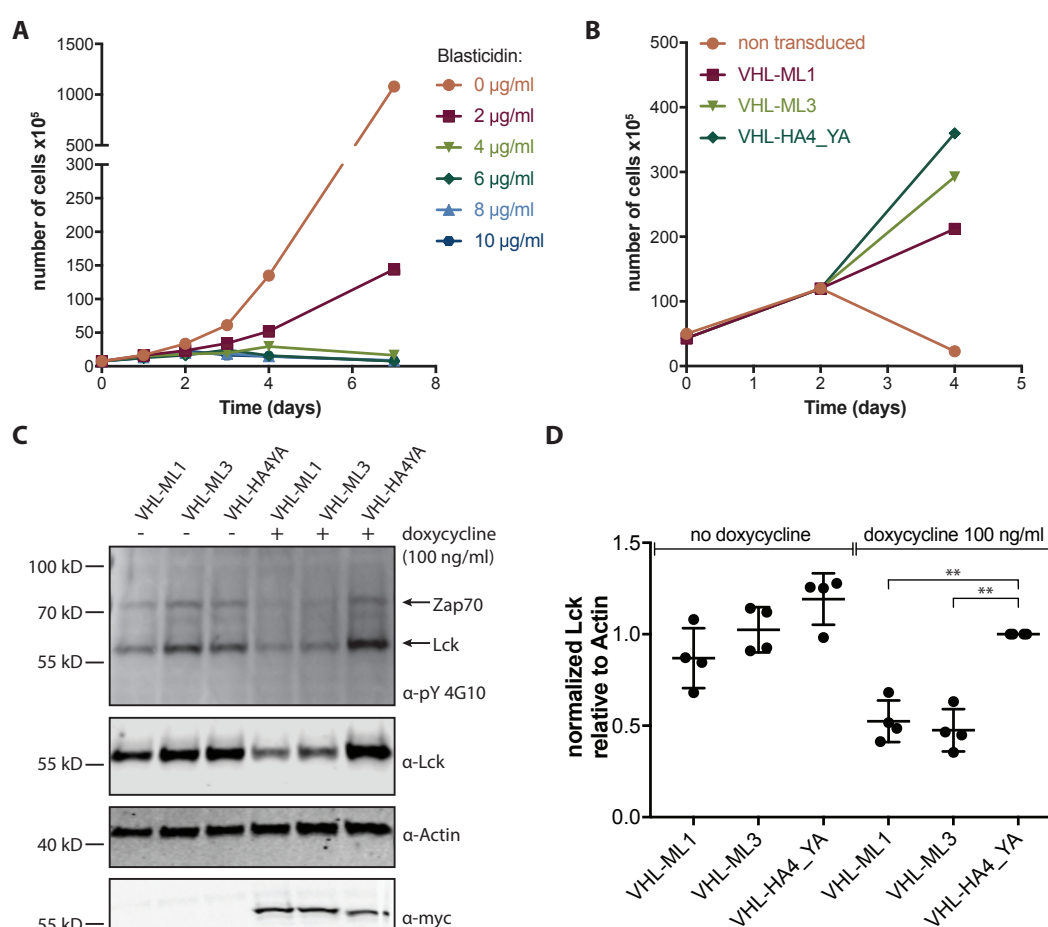
### 2.1.5 VHL-ML3 mediates Lck degradation upon expression in Jurkat cells

We reasoned that we could use the ML1 and ML3 monobodies for the targeted degradation of Lck, and that by this, we could potentially further decrease Lck-mediated signaling in T cells. I therefore fused the ML1 and ML3 monobodies as well as the negative control monobody HA4-YA to the E3 ubiquitin ligase recruiting Von-Hippel-Lindau (VHL) protein. I hypothesized that upon binding of the monobody to Lck, the fused VHL would recruit the Cullin2 E3 ubiquitin ligase complex, leading to the polyubiquitination and proteasomal degradation of Lck.

The VHL cDNA was fused to the ML3 cDNA and cloned into a lentiviral vector containing a Tet-on system for doxycycline dependent expression and a blasticidin resistance gene for selection. (pEM24) The Tet-on system in the vector allows to turn on the expression and thereby degradation of Lck by addition of doxycycline to the growth medium.

I first determined the minimal lethal concentration of blasticidin to use for the selection of the positive cells after transduction. (Figure 2.8 A) Following infection with the lentivirus generated by transfection of HEK293T cells with the constructs, the cells which had integrated the construct were selected by addition of 5  $\mu\text{g}/\text{ml}$  of blasticidin and after 4 days of selection, the non-transduced cells were dead, meaning that 100% of the remaining transduced cells had integrated the construct. (Figure 2.8 B) The expression of the VHL-monobody constructs was induced by addition of doxycycline to the medium and the expression of the myc-tagged constructs was tested by immunoblotting with an anti-myc antibody. The expression was effectively induced by the addition of doxycycline and no leakiness of the constructs could be detected. (Figure 2.8 C, bottom blot)

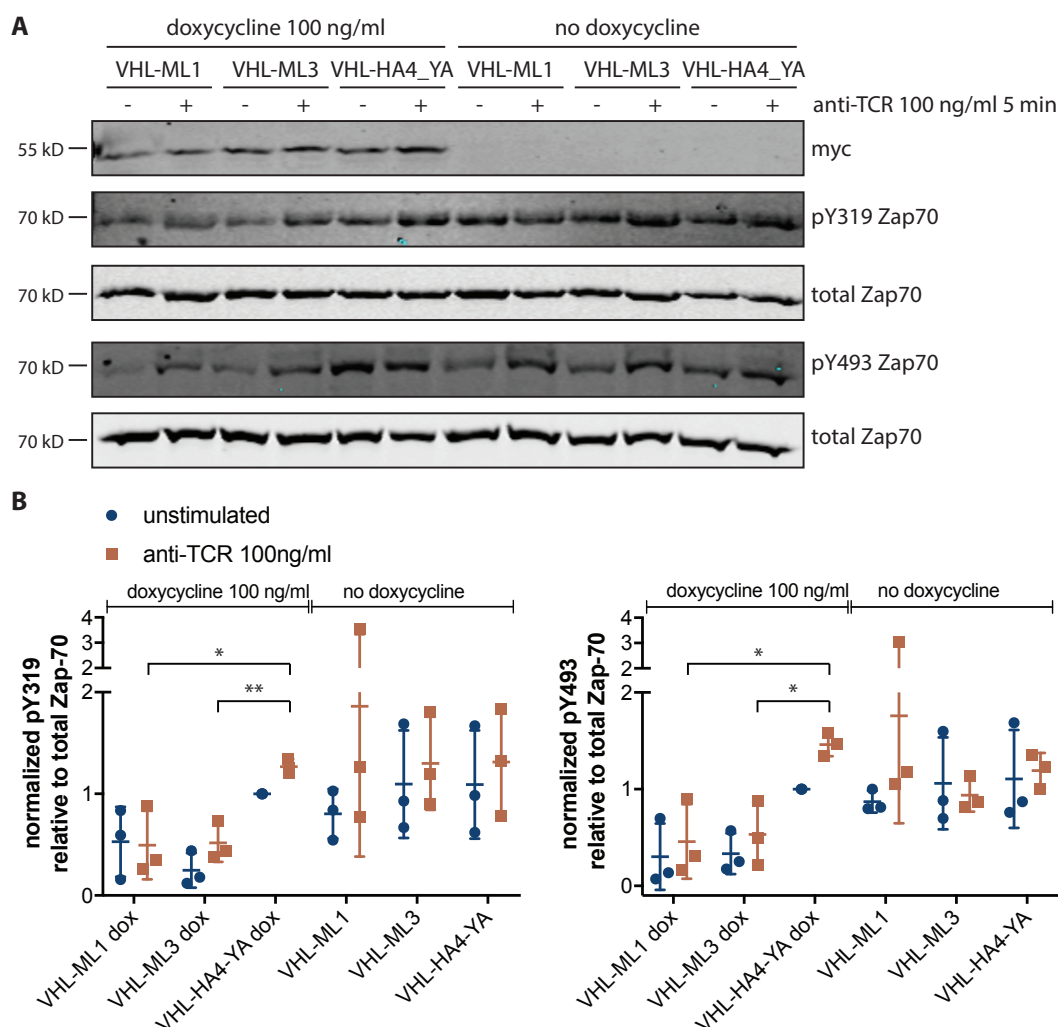
In order to assess whether the expression of VHL-monobody fusion proteins indeed leads to the degradation of the monobody target protein, cell lysates of cells grown in presence and absence of doxycycline were immunoblotted with an anti-Lck antibody. I observed that indeed, in cells expressing VHL-ML1 or VHL-ML3 fusion proteins, but not in cells expressing the control monobody HA4-YA fused to VHL, the protein levels of Lck are reduced by 50%, whereas in the same cell lines without addition of doxycycline, the level of Lck protein was similar across all the cell lines. (Figure 2.8 C and D) Moreover, when looking at total phospho-tyrosine levels, I observed that in doxycycline induced cells expressing VHL-ML1 or VHL-ML3, the bands at approximately 55 kD, probably corresponding to phospho-Lck and at about 70 kD, which corresponds to Zap70, were less strong than in uninduced cells or cells expressing VHL-HA4\_YA. This observation indicates that the loss of Lck leads to a reduced Zap70 phosphorylation already in unstimulated cells.



**Figure 2.8 VHL-ML1 and -ML3 mediated degradation of Lck in Jurkat cells.** **A:** Wild type Jurkat cells were supplemented with blasticidin at the indicated concentration and the cell number was measured for 7 days. This measurement was done once. **B:** Jurkat cells were transduced with the PEM24 vector containing the indicated VHL-monobody constructs or left untransduced and 5  $\mu\text{g/ml}$  blasticidin was added to the growth medium 1 day after the second transduction (day 0). The cell numbers were measured on day 2 and day 4. **C:** Immunoblot analysis of the lysate of VHL-monobody transduced Jurkat cells grown in presence and absence of doxycycline with antibodies against tyrosine-phosphorylated proteins (4G10), Lck, Actin and the myc-tag. (Order as displayed from top to bottom) **D:** Quantification of the ratio between the Lck and Actin signals. The data was transformed by setting the Lck:Actin ratio of doxycycline induced VHL-HA4\_YA transduced cells to 1.0 for each experiment. Average values and standard deviations from 4

independent experiments were used, and  $p$  values were calculated from the untransformed data using a ratio paired t-test.  $**p < 0.01$

I next stimulated these cells with the anti-TCR antibody, as previously, and assessed whether the phosphorylation of Zap70 is reduced in cells where Lck is degraded. I observed that indeed, both in unstimulated cells as well as in cells stimulated with an anti-TCR antibody for 5 minutes, the phosphorylation of Zap70 is reduced, and that this effect is stronger than in cells expressing the monobodies alone without VHL. (Figure 2.9)



**Figure 2.9 VHL-monobody mediated Lck degradation leads to reduced Zap70 phosphorylation. A:** Jurkat cells expressing VHL-ML1, -ML3 or -HA4\_YA and induced with doxycycline or grown in doxycycline-free medium were stimulated with an anti-TCR antibody for 5 minutes and lysed. Immunoblot analysis of the cell lysate with antibodies against the myc-tag (upper blot), phosphorylated Y319 residue of Zap70 and total Zap70 on the same blot (blots 2 and 3 from top) and against phosphorylated Y493 residue of Zap70 and total Zap70 on the same blot (blots 4 and 5 from top). **B:** Quantification of the ratio between the Y319-phosphorylated Zap70 and total Zap70 signals (left) and of the ratio between the Y319-phosphorylated Zap70 and total Zap70 signals (right). The data was transformed by setting the pZap70:Zap70 ratio of the unstimulated (doxycycline induced) HA4-Y87A control cells to 1.0 for each experiment. Average values and standard deviations from 3 independent experiments were used, and  $p$  values were calculated using a paired t-test.  $*p < 0.05$ ,  $**p < 0.01$ .

Collectively, these results demonstrate that the selected monobodies are SFK subgroup specific and can interfere with the TCR signaling pathway in cells, leading to the degradation of their target protein when coupled to the E3 ubiquitinating ligase recruiting protein VHL.

## 2.2 Monobody delivery using CPPs

My first attempt at delivering monobodies into cells was to use cell penetrating peptides (CPPs). In order to couple monobodies to CPPs, a previous master student in our lab had attempted to express and purify recombinant monobody-CPP fusion proteins in *E. coli*. Despite optimization attempts, the resulting proteins were very unstable and prone to precipitation, so that no soluble protein could be purified. My attempt at purifying monobody-CPP fusion proteins failed for the same reason, and I therefore set out to establish a method to chemically ligate monobodies to CPPs. The selected method consists of using DTNB (5,5'-Dithiobis(2-nitrobenzoic acid)) or Ellman's reagent) in order to covalently link a monobody to a CPP via a disulfide bond. (Figure 2.10) DTNB is normally used as a colorimetric reagent to quantify the number of free thiols in a protein, as the reaction product yields TNB, which is yellow.

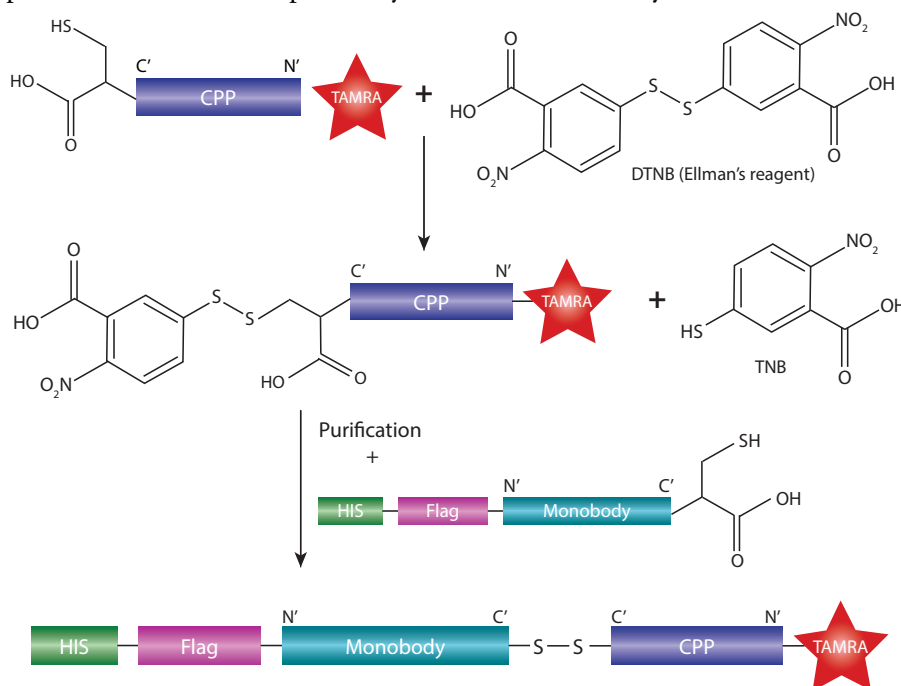


Figure 2.10 **Covalent binding of a monobody to a CPP by disulfide bond.** The CPP is first activated with Ellman's reagent and after purification, reacts with the monobody to form a disulfide bond.

I added a cysteine residue to the C-termini of 5 monobodies that had previously been characterized in our lab. Since monobodies do not contain cysteines, the C-terminal cysteine provides the only free thiol in the monobodies to react with DTNB-activated CPPs. I then expressed the monobodies, containing a Flag-tag as well as a 10x-His-tag at the N-terminus, in the *E. coli* BL21 strain, allowing for purification on a Ni-NTA column. Table 2.3 gives an overview of the 5 monobodies chosen, as well as their target and the yield of expression in *E. coli*.

Monobody	Target	Yield / l of expression medium
HA4	Abl SH2 domain <sup>76</sup>	11 mg
AS25	Abl SH2-kinase interface <sup>78</sup>	17 mg
AS27	Abl SH2-kinase interface <sup>78</sup>	12 mg
Nsa5	SHP2 N-terminal SH2 domain <sup>212</sup>	10 mg
Nsa5 S63K	Mutated non-binding form of Nsa5 <sup>75</sup>	8 mg

Table 2.3 **Monobodies expressed in the *E. coli* BL21 strain, their targets, and yields from 1 L of expression medium after purification by SEC.** Autoinduction medium was used as expression medium, cultures were incubated for 20h at 20°C, shaking, prior to lysis.

Three different CPPs were selected and synthesized with a cysteine in the C-terminus and both unlabeled as well as with a tetramethyl-6-carboxyrhodamine (TAMRA) dye label in the N-terminus (Table 2.4).

Peptide	Origin	Sequence	Length
Tat	HIV	GYGRKKRRQRRPPQC	16 aa
CPP44	synthetic, CML- and AML-specific	KRPTMRFRYTWNPMKC	16 aa
Polyarginine	synthetic	RRRRRRRRRC	10 aa

Table 2.4 **Cell penetrating peptides chosen for coupling to monobodies, their origin and sequence.**

Figure 2.11 A shows a Coomassie-stained gel of all purification steps of AS27 as a representative example for all the monobodies. In Figure 2.11 B, the size exclusion chromatogram of AS25 is representative for all the monobodies purified. A good yield and high purity were obtained for all five monobodies, as shown on a Coomassie-stained SDS-PAGE gel of the purified monobodies. (Figure 2.11 C and Table 2.3)

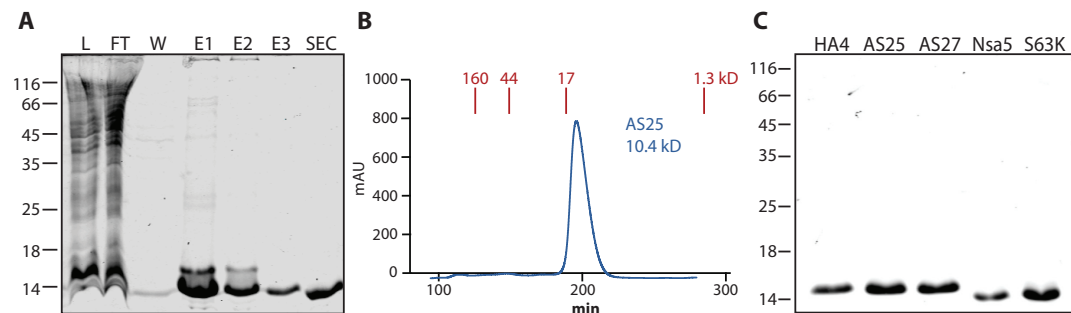


Figure 2.11 **Monobody purification.** The crude bacteria lysate was first purified on a Ni-NTA column and then by SEC. **A:** Coomassie stained SDS PAGE gel with the fractions from the Ni-NTA purification of monobody AS27 and the main peak of the SEC after concentration (similar purity was obtained from the other monobodies) (L=crude lysate, FT=flow-through, W=wash, E1-3=Ni-NTA elution fractions 1-3) **B:** size exclusion chromatogram of monobody AS25 as an example. (All other monobodies gave similar chromatograms). **C:** Coomassie stained SDS PAGE gel of all the purified monobodies after purification.

DTNB (or Ellman's reagent) was then used to covalently link the single cysteine residues in the CPPs and the monobodies via a disulfide bridge, as depicted in the scheme in Figure 2.11.

For the first step, activating the CPP with DTNB, different non-denaturing phosphate- and TRIS-based buffers as well as pH-values distinct from the pI of the peptides were tested. Both Ellman's reagent and the peptides were soluble in all buffers, but upon

mixing, a yellow precipitate formed. The reaction product was contained in the precipitate and not the supernatant, as seen by HPLC analysis (data not shown). By using a denaturing phosphate buffer containing 6 M guanidinium, a yellow solution was obtained, and no precipitation was observed. Ellman's reagent was used in 5-fold molar excess and the reaction was immediate. To monitor the reaction, the initial CPP as well as the final product were analyzed by HPLC and mass spectrometry. Figure 2.12 shows the HPLC analysis of each peptide and of the reaction product with Ellman's reagent as an overlay. Many of the peptides showed major impurities either in HPLC or in mass spectrometry analysis and did not match the data on the quality control sheet of the supplier. Even after re-synthesis by the supplier, the quality did not improve and new peptides were finally ordered from a different company.

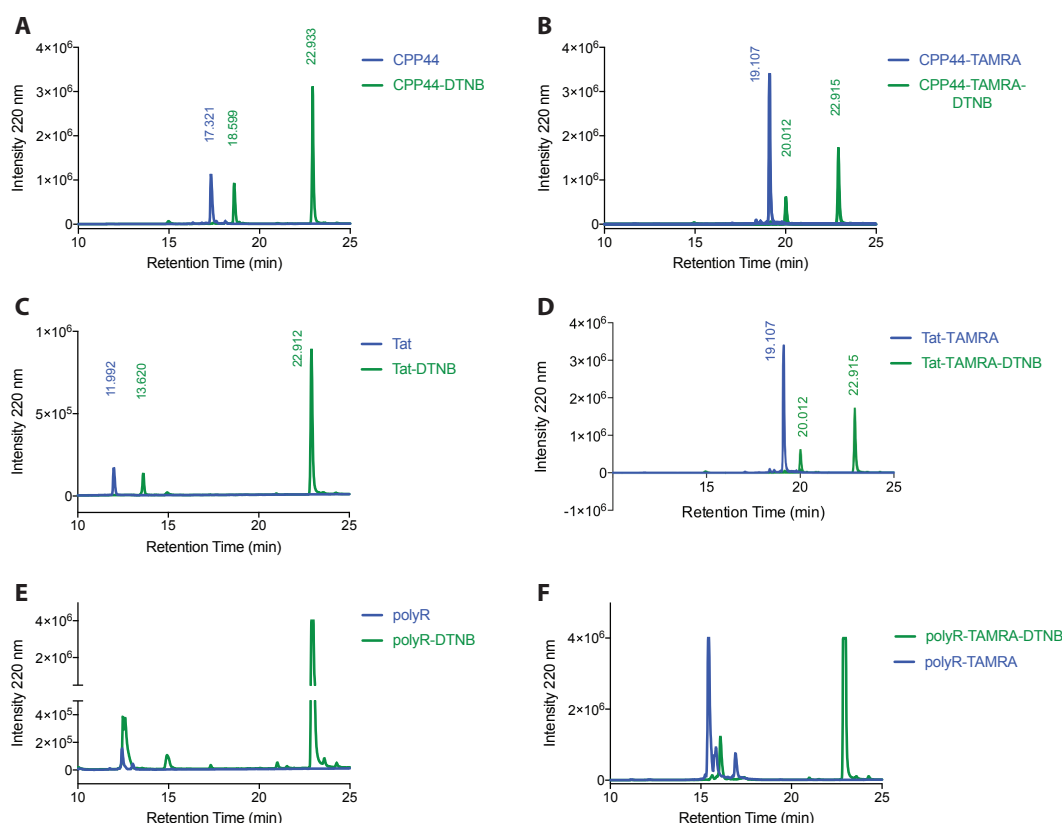


Figure 2.12 HPLC chromatograms of 6 CPPs (in blue) and the reaction product of the same peptide with Ellman's reagent (DTNB) (in green). A: CPP44, B: TAMRA-labeled CPP44, C: Tat, D: TAMRA-labeled Tat, E: polyarginine, F: TAMRA-labeled polyarginine. In each of the chromatograms, the largest peak of the green curves corresponds to excess Ellman's reagent. An analytical scale C-18 column was used and the flow rate was constant at 1ml/min.

In order to establish the coupling of the activated CPPs to monobodies, unlabeled CPP44 and Tat were chosen to test the next reaction step. The reaction with Ellman's reagent was repeated at a larger scale and the activated CPP was purified by HPLC on a semi-preparative C-18 column, to remove the excess DTNB and the TNB reaction product. After lyophilization of the eluate, it was re-dissolved in a sodium phosphate buffer at pH 6.5 and the activated peptide remained soluble. To assess the purity of the product, an analytical HPLC was run, which gave a single major peak of the expected mass for both CPP44 and Tat. (Figure 2.13 A) The DTNB-activated and purified CPPs were then mixed with monobody protein. After 30 minutes of incubation on ice, a

yellow color formation due to TNB production indicated that the reaction had taken place. The analytical HPLC showed that the product eluted in one major peak, which was shifted compared to the monobody alone. (Figure 2.13 B) Unfortunately, the resulting monobody-CPP product was very unstable and prone to precipitation, similar to the recombinant fusion protein. When I replaced the reaction buffer by a sodium phosphate buffer containing 6 M guanidinium, the product remained soluble, since the protein is denatured. Different attempts to refold the protein by dialysis into a guanidinium-free buffer have failed so far, and another coupling method might be needed to ligate CPPs to monobodies.

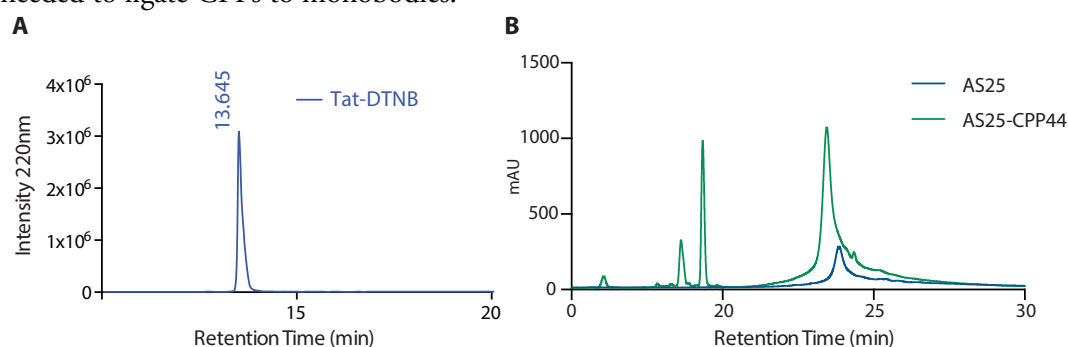


Figure 2.13 **CPP-monobody coupling**. A: Analytical HPLC chromatogram of purified Tat after activation with DTNB. CPP44 resulted in a similarly pure single peak. B: Analytical HPLC chromatogram of AS25 after coupling to DTNB-activated and purified CPP44 (in green) and of AS25 alone (in blue).

## 2.3 Monobody delivery using Cell Penetrating poly-Disulfides (CPDs)

Cell penetrating poly-disulfides (CPDs) have been developed by the group of Prof. Matile at the University of Geneva, with whom we collaborated on this project. The Matile group has shown that CPDs can be used to deliver fluorophores and different macromolecules, including peptides, into cells, but protein delivery had not been assessed.<sup>113,116</sup> CPDs are synthetic mimics of poly-arginine in which the polypeptide backbone is replaced with poly-disulfides. Upon cellular uptake, CPDs are rapidly degraded in the cytosol by glutathione (GSH)-mediated reduction of the disulfide bonds.

### 2.3.1 Adduct formation and delivery of HPDP-biotin-AS25

Since the polymerization of CPDs is carried out in organic solvents, proteins cannot be used as initiators of the reaction directly. We therefore decided to adopt a different strategy to conjugate monobodies to CPDs. We reasoned that since the binding of biotin to streptavidin is one of the strongest naturally occurring non-covalent conjugations and the reagents are commercially available, we could use streptavidin to couple biotinylated monobodies to CPDs grown from a biotinylated initiator. (Figure 2.14 A) Tetramethylrhodamine (TMR) coupled to biotin and bearing a free thiol group is used as an initiator (I) to form biotinylated and fluorophore-labelled CPDs. The CPD building block consists of lipoic acid coupled to arginine (M), as the strained disulfide offers a target for nucleophilic disulfide exchange polymerization. The thiol group of the initiator reacts with this disulfide, forms a new covalent disulfide bond with the monomer and generates a free reactive thiol to react with the next monomer. The polymerization reaction can be stopped at different time points by addition of the terminator (T) in excess, to yield products with different polymer lengths. (Figure 2.14 A)

In order to make a biotinylated and fluorophore-labelled monobody, our initial approach was to add a cysteine residue to the C-terminus of the monobody by site-directed mutagenesis, which could be used to react with the pyridyldithiol-activated, sulfhydryl-reactive biotinylation reagent Biotin-HPDP, as shown in Figure 2.14 B. We decided to use AS25 as a first monobody to test this approach. A bacterial expression plasmid with the monobody AS25 with the C-terminal cysteine was expressed in the *E. coli* BL21 strain and purified on a Ni-NTA column and by size exclusion. NHS-ester coupled Dylight488 was then used to label the free amines in the protein. Since there are 3 lysine residues in the monobody in addition to the N-terminus, this would result in up to 4 Dylight488 molecules per monobody protein. Dylight488 was added to the monobody at a 7.5 fold molar excess and incubated for 1 hour at room temperature. After purification of the labelled protein from excess dye by dialysis, the ratio of dye molecules per monobody was determined to be 3.27 by spectrophotometric absorbance measurement. This labelled monobody was then incubated with Biotin-HPDP for 1 hour at room temperature. (Figure 2.14 B)

Streptavidin could then be used to form adducts with this biotinylated and fluorophore labelled CPD as well as biotinylated monobodies. (Figure 2.15) Streptavidin is a homotetramer with 4 biotin binding sites, so different stoichiometries could be used to make adducts. We reasoned that forming adducts at a 1:2.5:1 (Monobody:CPD:Streptavidin) ratio, we would maximize the chance of uptake due to an average of 2.5 molecules of

CPD on each adduct. We did not choose a 1:3:1 ratio to avoid having residual free CPD in the solution.

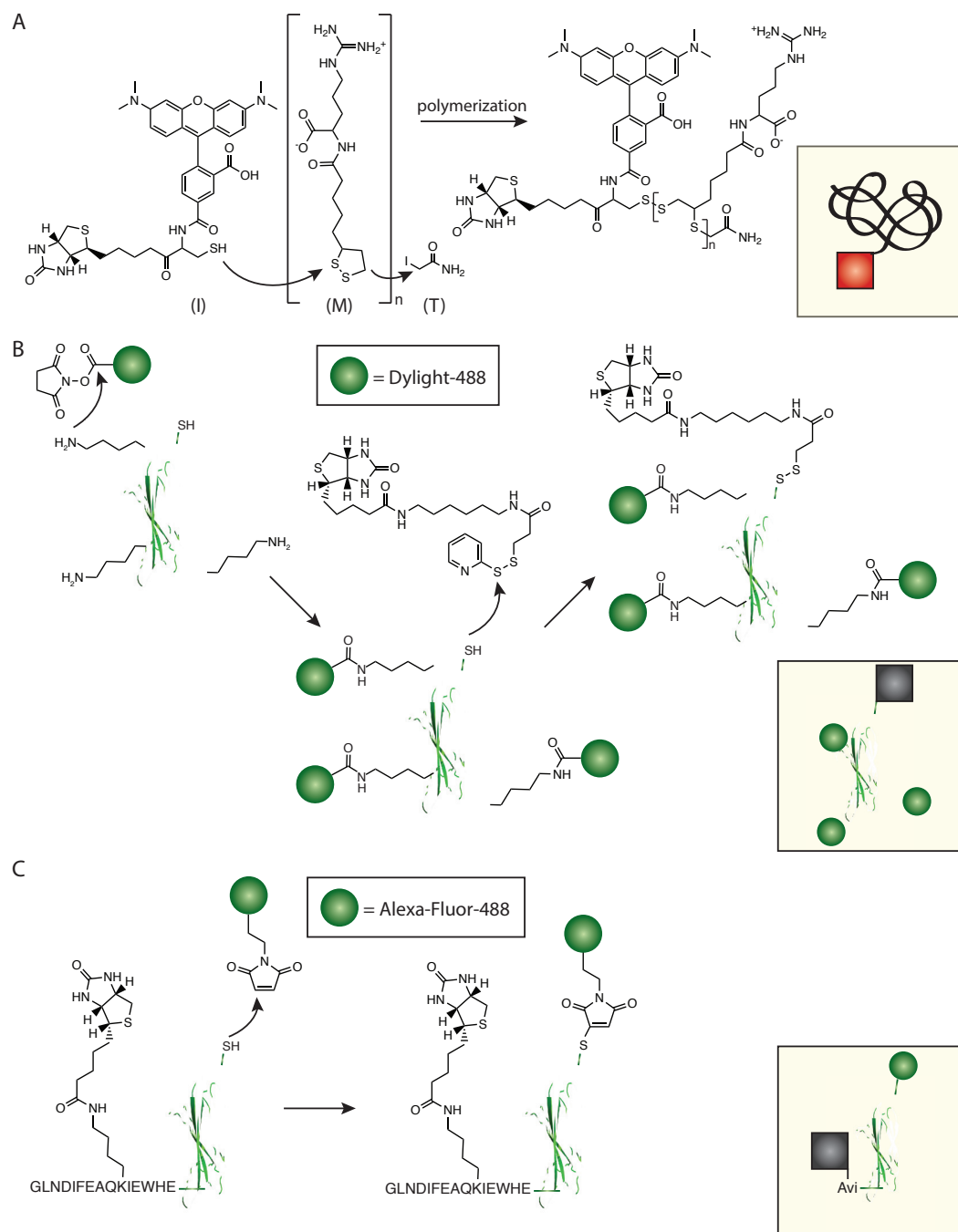


Figure 2.14 **Synthesis of CPDs and formation of biotinylated and fluorophore coupled monobodies.**

**A:** TMR-coupled and biotinylated initiator (I) is used to start a nucleophilic disulfide exchange reaction with the strained disulfide in the monomer (M), consisting of a lipoic acid coupled to arginine, the reaction is stopped by addition of the terminator (T). A schematic representation of the resulting CPD is shown in the colored box and will be used throughout the chapter. **B:** Free amines in a monobody react with NHS-ester coupled Dylight-488 and after purification of the fluorophore-coupled monobody, reacts with biotin-HPDP to yield a biotinylated and fluorophore-coupled monobody, as shown schematically in the colored box. **C:** Avi-tagged monobody has been biotinylated by co-expression with BirA enzyme in *E. coli* and its free thiol of the cysteine at the C-terminus reacts with maleimide-coupled Alexa Fluor 488. The resulting biotinylated and fluorophore-coupled monobody is shown as a schematic in the colored box.

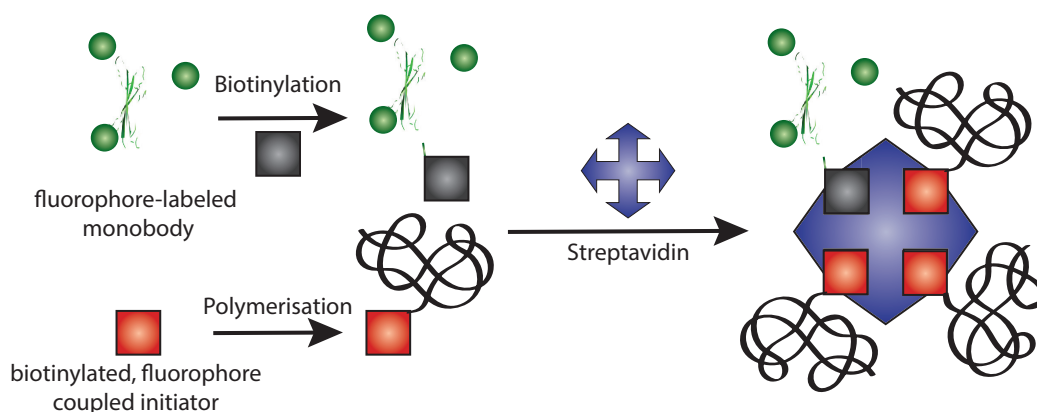


Figure 2.15 **Schematic showing adduct formation.** Both monobodies and CPDs are biotinylated and fluorophore coupled and mixed with streptavidin at a 1:2.5:1 (monobody:CPD:streptavidin) ratio. For simplicity, the graph shows resulting adducts at 1:3:1 ratio.

For the initial testing, we decided to form the adducts by simple mixing at the intended ratio, without purifying adducts. This would result in a polydispersed product, with a majority of adducts at the intended stoichiometry of 1 monobody and 2.5 CPDs per streptavidin and a fraction of adducts with no monobody or more than one monobody per streptavidin. Streptavidin was first mixed with CPDs at a 2.5:1 ratio at room temperature for 2 hours, and then the biotinylated and labelled monobody was added and the mixture was incubated for additional 2 hours at room temperature. As a control, a biotin-fluorescein conjugate was added instead of the monobody. Initially, with the reaction performed in PBS at pH 7.5, most of the monobody, but not the biotin-fluorescein, precipitated upon mixing with the streptavidin-CPD. The resulting adducts were centrifuged to remove the precipitate and diluted in Leibovitz's medium to a calculated concentration of 500 nM. When considering the precipitation, the actual concentration was likely to be lower than 500 nM, however the precise concentration was difficult to determine, since the CPDs do not show prominent UV absorption in contrast to proteins. Leibovitz's medium is commonly used for live cell microscopy as it allows cells to grow at atmospheric CO<sub>2</sub> concentration. The adducts diluted Leibovitz's medium were added to HeLa cells grown in microscopy dishes and incubated at 37°C for 1 hour.

Figure 2.16 shows images of live HeLa cells after incubation with the adducts for 1 hour and a subsequent washing step. Both the adducts with labelled monobody or fluorescein seem to be taken up in HeLa cells, since signal from both the TMR on the CPDs and the Dylight488 or fluorescein (from the monobody) are detected inside the cells. The TMR signal is much stronger in the cells incubated with fluorescein-adducts than the ones with monobody-adducts, even though the adducts were prepared in the same way and the same number of CPDs are bound to streptavidin on average. This difference is most likely due to the precipitation of the monobody-adducts and the resulting lower concentration. When the cells were incubated with monobody-adducts at a calculated concentration of 1 µM, the intensity of TMR was comparable to the cells incubated with fluorescein-adducts. The signal from the fluorescein or Dylight-488 on the monobody is largely colocalized with the TMR signal and both signals are mostly coming from distinct speckles throughout the cytosol. (Figure 2.16) In some of the cells that were incubated with the fluorescein adducts or with the higher concentration of the monobody-adducts, signs of apoptosis, such as membrane blebbing, irregular nuclear shapes and cell debris were observed.

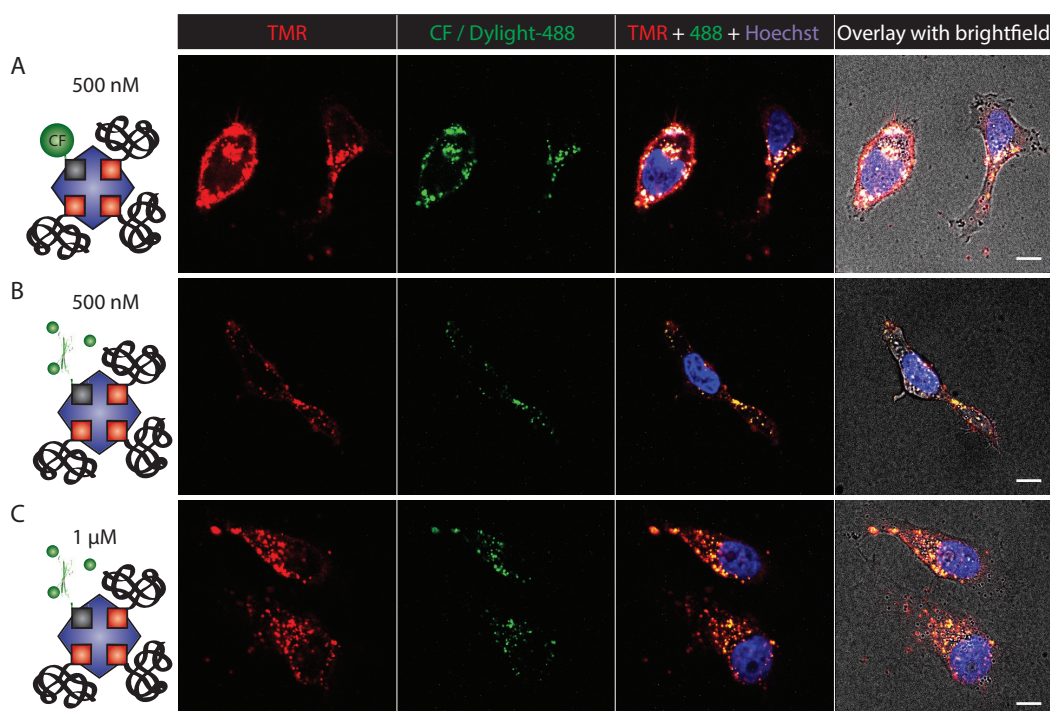


Figure 2.16 **Adduct delivery in HeLa cells.** HeLa cells were incubated with 500 nM CPD-Streptavidin-CF (A) or 500 nM (B) or 1  $\mu$ M (C) CPD-Streptavidin-AS25 1h, washed and imaged on a confocal microscope. TMR signal is shown in red, CF and Dylight-488 signal in green and nuclei are stained with Hoechst and shown in blue. Representative images are shown out of 3 biological experiments where 5 images were taken from each condition randomly. Scalebars represent 10  $\mu$ m.

### 2.3.2 Buffer and conditions optimization to yield soluble adducts

Since the adduct formation lead to extensive precipitation of the monobody, I decided to optimize the buffer conditions to yield a more soluble product. I tested Tris-HCl, PBS and MES buffers at a pH values ranging from 5.5 to 8.5, different salt and glycerol concentrations, as well as the addition of polyethyleneglycol (PEG) of different sizes. The best condition was found to be MES buffer at pH 5.5 with 150 mM NaCl and 10% glycerol, as it yielded less precipitation than all the other conditions. The addition of PEG did not result in any change as compared to the same buffers without PEG. In addition, by reducing the concentration of both the monobodies and the CPDs to 15  $\mu$ M prior to mixing, I was able to further improve the solubility. The conditions for the formation of adducts were also slightly changed to favor the protein stability and solubility. The reaction of Biotin-HPDP with monobody was performed at 4°C and the biotinylated monobody was added to the streptavidin-CPD adducts at 4°C drop by drop and with gentle mixing, then left to complete the reaction for 1 hour at 4°C. Using this new buffer and reaction conditions, the final product appeared to be soluble with no visible precipitate. Any invisible precipitated material was removed by centrifugation and the resulting soluble adducts were diluted in Leibovitz's medium to the required concentration for the respective experiment. The incubation with HeLa cells was performed on the same day of adduct formation to avoid any precipitation due to freeze-thaw cycles or prolonged storage at 4°C.

### 2.3.3 Adduct formation and delivery of Avi-tagged and in vivo biotinylated monobody

Since the addition of several Dylight488 molecules on the monobody could interfere with its binding to the target protein, we also decided to test another biotinylation and labeling approach. By adding an Avi-tag to the N-terminus of the monobody, we reasoned that we could use this to attach a biotin, thereby freeing the C-terminal cysteine for reaction with a maleimide coupled dye. This method has the advantage that the reaction with the maleimide coupled dye is site-specific, since there is no other cysteine residue in the monobody. I co-expressed Avi-tagged monobody with BirA enzyme in presence of biotin in the *E. coli* BL21 strain. BirA is a bacterial enzyme specifically recognizing the Avi-tag sequence (GLNDIFEAQKIEWHE) and attaching one biotin molecule to the lysine residue. The co-expression of both proteins yielded a monobody that was biotinylated to about 90% according to a streptavidin bead pulldown assay. After purification by Ni-NTA and SEC, the monobody was labeled on the single C-terminal cysteine using maleimide-AlexaFluor488 and purified from excess dye using a PD-10 sephadex G25 column. (Figure 2.14 C) The formation of adducts was done in the same way as for the HPDP-biotinylated and NHS-ester labelled monobody with the improved buffer and reaction conditions.

I incubated HeLa cells for 1 hour with adducts with either HPDP-biotinylated, NHS-ester labelled (AS25-Dylight488) or Avi-tag-biotinylated, maleimide labelled (Avi-AS25-AF488) monobody and observed their uptake in live cells. As shown in Figure 2.17, both the TMR signal from the CPDs and the 488 nm signal from the monobody can be seen in speckles in the cytosol, but not all the cells take up the same amount of adducts, as already observed before. The 488 nm signal partially colocalizes with the TMR signal, and the Avi-AS25-AF488 adducts seem to be localized closer to the cell membrane after uptake than the AS25-Dylight488 adducts. No 488 nm signal can be observed when the cells are incubated with adducts containing an unlabeled monobody, or with CPD alone. Since no major differences were observed between the two labelling and biotinylation protocols of the monobodies, I used the *in vivo* biotinylated AS25 with only one fluorescent dye on the C-terminus for all the subsequent experiments. In all the conditions, except for the control cells which were incubated with labelled monobody only, signs of apoptosis, such as membrane blebbing, irregular nuclear shapes and cell debris were observed. These effects were most clearly observed in cells incubated with 500 nM CPD only, indicating that the observed toxicity stemmed from the CPD and not the monobody. The experiment was repeated with different batches and sizes of CPDs with the same result. The observed toxicity was surprising, since results from of similar adducts containing the same CPDs at concentrations up to 1  $\mu$ M have been published and shown to have no toxic effects on HeLa cells in an MTT assay by the Matile group.<sup>113</sup>

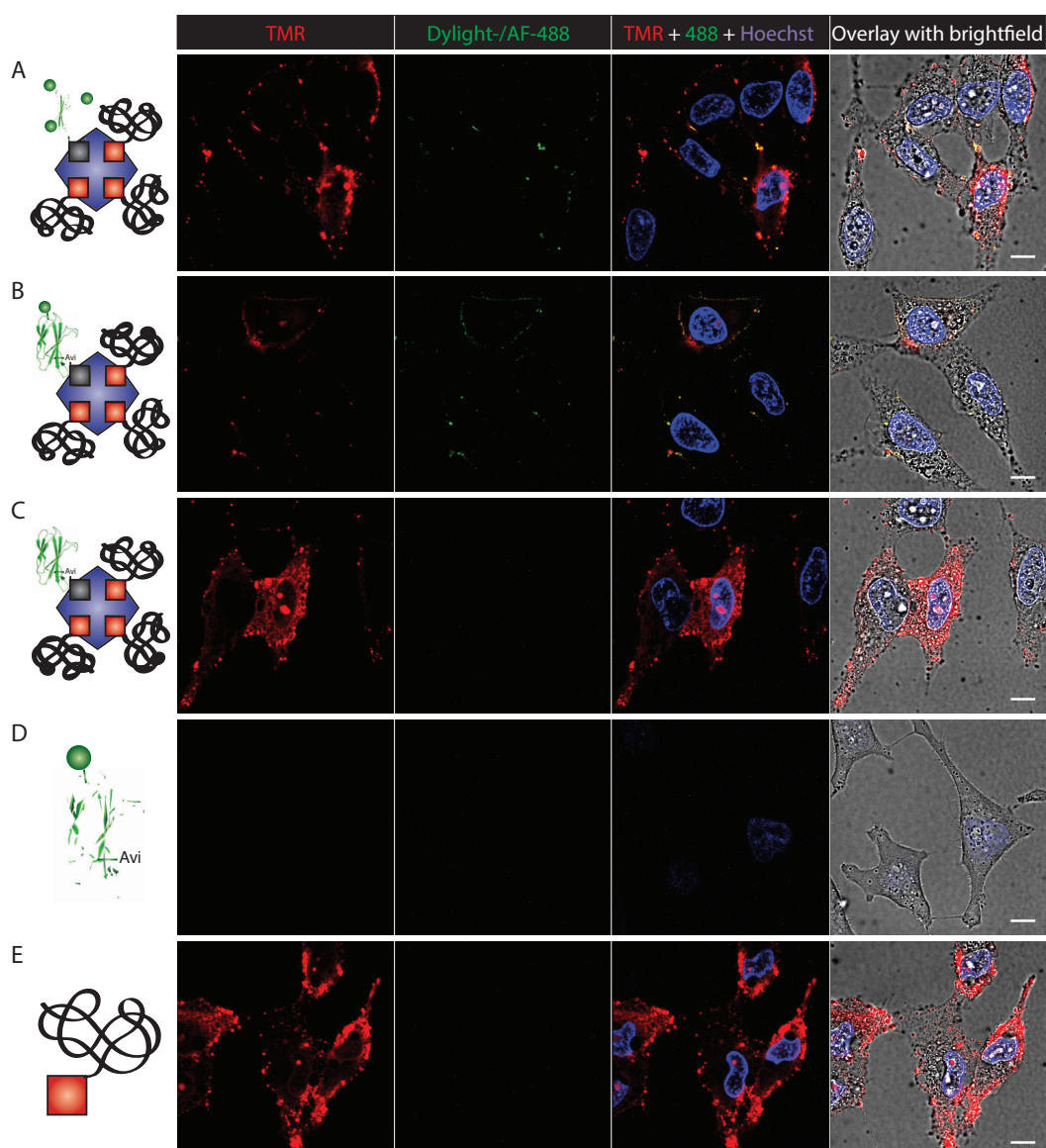


Figure 2.17 **Adduct delivery and comparison between in vivo biotinylated and HPDP-biotin conjugated monobodies.** HeLa cells were incubated with adducts or controls at 500nM for 1h, washed and imaged using a confocal microscope. A: Adduct with HPDP-biotinylated, NHS-ester labelled AS25 (AS25-Dylight488); B: Adduct with Avi-tag-biotinylated, maleimide labelled AS25 (Avi-AS25-AF488); C: Adduct with Avi-tag biotinylated, unlabeled AS25; D: Avi-AS25-AF488 alone; E: TMR-labelled, biotinylated CPDs alone. TMR signal is shown in red, CF and Dylight-488 signal in green and nuclei are stained with Hoechst and shown in blue. Representative images are shown out of 2 biological experiments where 5 images were taken from each condition randomly. Scalebars represent 10  $\mu$ m.

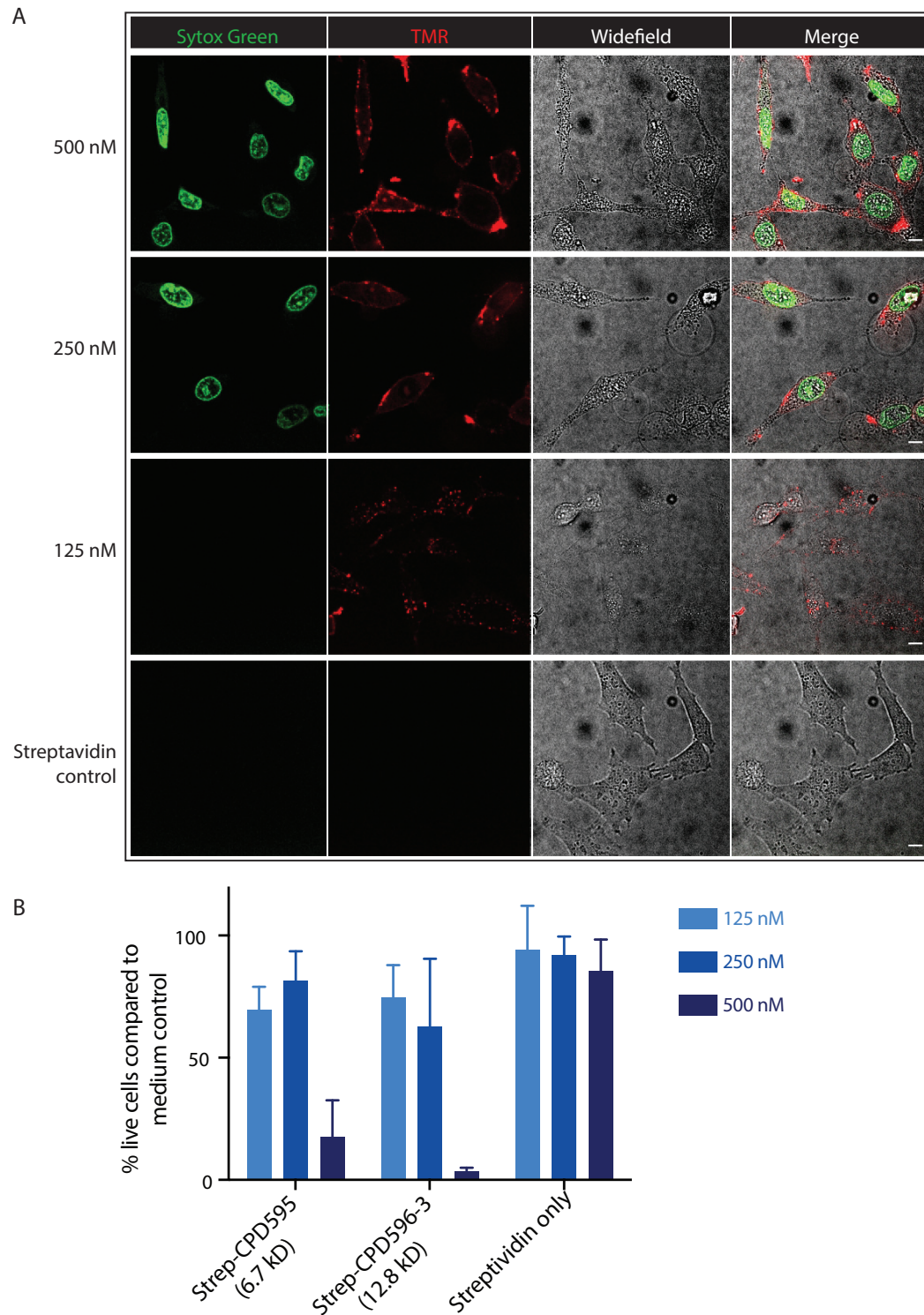
#### 2.3.4 CPD containing adducts have cytotoxic effects on HeLa cells

In order to test whether the adducts were indeed inducing cytotoxic effects, I stained the cells with Sytox Green after the incubation with adducts containing only streptavidin and CPDs at an average 1:2.5 ratio. Sytox Green is a DNA intercalating dye, which is impermeable in healthy cells but can penetrate the cell membrane as well as the nuclear envelope in cells undergoing apoptosis due to membrane permeabilization. As shown in Figure 2.18 A, the nuclei of cells incubated with adducts for 1 hour indeed become positive for Sytox Green staining, at adduct concentrations of 250 nM and 500 nM but

not at 125 nM or lower concentrations. When Streptavidin alone is added to the cells at 500 nM, the cells remain negative for Sytox Green. This result clearly indicates that concentrations of adducts of 250 nM or above cause membrane permeabilization in HeLa cells. We reasoned that this permeabilization could also be due to the cell penetrating nature of the CPDs and a byproduct of the uptake of the adducts into the cytosol. I therefore repeated the same experiment including a 24-hour incubation time in adduct free medium after the incubation with the adducts as well as several washing steps, in order to allow the cells to recover from the membrane permeabilization. When subsequently staining the cells with Sytox Green, I also observed the Sytox Green signal in the nuclei, showing that even when the cells are left to recover from the delivery, their membranes remain permeabilized which is a clear sign of occurring cell death. (data not shown)

I additionally performed a CellTiter Glo assay to test if the number of viable cells is decreasing in response to the incubation with CPD adducts. CellTiterGlo is a cell viability assay measuring the amount of ATP in a well. Since only healthy cells produce and contain high amounts of ATP, this is a direct measure of the number of viable cells. In the wells with 500 nM adduct concentration, the luminescent signal and thereby the number of viable cells was significantly decreased as compared to the control with streptavidin only. (Figure 2.18 B) The cytotoxic effects observed were independent of the CPD length, as two different CPD sizes as well as preparation batches were tested and gave similar results. Incubation of the cells with adducts at 250 nM or 125 nM had much less effect on cell viability in the CellTiter Glo assay. (Figure 2.18 B)

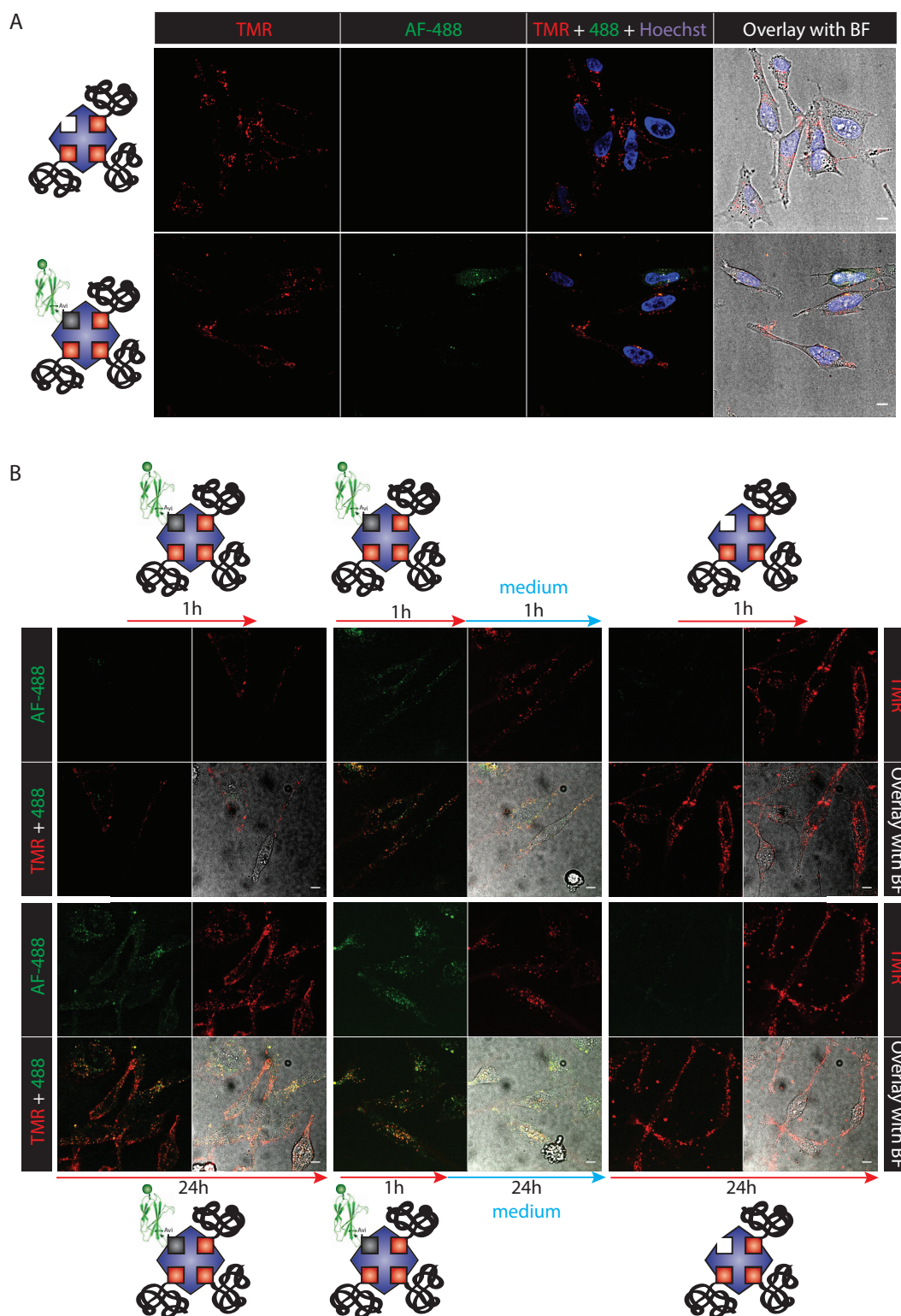
Based on these results, I decided to work with adducts at 100 nM concentration for subsequent experiments, as this concentration did not induce any cytotoxic effect in either the CellTiter Glo or the Sytox Green assay.



**Figure 2.18 CPDs have cytotoxic effects on HeLa cells.** **A:** Sytox Green stain of live HeLa cells after 1h incubation with Streptavidin-CPD adducts at indicated concentrations. TMR signal is shown in red and Sytox Green signal in green. Scalebars represent 10  $\mu$ m. Representative images are shown out of 2 biological repeats. **B:** CellTiter Glo assay on HeLa cells which were incubated with Streptavidin-CPD adducts of different CPD lengths for 1 h, then washed and incubated in growth medium for 24 h before adding the CellTiter Glo reagent. This experiment was done once in triplicates.

#### 2.3.4 CPD-AS25 adducts are taken up into HeLa cells without cytotoxic effects at a lower concentration

In order to assess if the monobodies can also be delivered to the cytoplasm after incubation at a lower concentration, I performed live cell imaging on HeLa cells after incubation with adducts at 100nM and subsequent washing. After 1 hour of incubation with the adducts at 100 nM, a TMR signal can be observed in the cytosol of the cells, and a very weak 488 nm signal from the monobody is detectable, which is absent in cells incubated with CPD adducts without monobody as a control. (Figure 2.19 A) Longer incubation times with the adducts, as well as in medium after a shorter incubation with adducts were tested in a subsequent experiment to assess whether more adducts are taken up over time. (Figure 2.19 B) When the cells are incubated for 1 hour in medium after the incubation with adducts, the 488 nm signal becomes more visible in the cytosol. This suggest that the fraction of adducts not containing any monobody might be taken up faster than the monobody-adducts. Both the 488 nm and the TMR signal are mostly localized in distinct speckles throughout the cytosol. After incubating the cells for 24 hours with the adducts, both the TMR and the 488 nm signals become much stronger. When incubating the cells for 1 hour with the adducts and then for 24 hours in medium, the signal was weaker than in cells incubated with adducts for 24 hours, but still remained. However, even after prolonged incubation, the signal remains mainly in distinct speckles, which possibly represent vesicles, in most of the cells. Cells incubated with control adducts containing only CPDs but no monobody do not display any 488 nm signal, showing that the observed signal in the monobody-adduct incubated cells is indeed coming from the AF488 label on the monobody. No cytotoxic effects could be observed in these cells, even after prolonged incubation with the adducts and a Sytox Green stain remained negative. (data not shown)



**Figure 2.19 100 nM adducts are taken up in HeLa cells without having cytotoxic effects. A:** HeLa cells were incubated with 100 nM AS25-streptavidin-CPD adducts or streptavidin-CPD adducts for 1h, washed and imaged on a confocal microscope. Representative images are shown out of one experiment where 5 images were taken from each condition randomly. **B:** HeLa cells were incubated with 100 nM AS25-streptavidin-CPD adducts for 1h (upper left) or for 24h (lower left), or with streptavidin-CPD control adducts for 1h (upper right) or for 24h (lower right) washed, then imaged or incubated for 1h with AS25-streptavidin-CPD adducts, washed and incubated for 1h (upper middle) or for 24h (lower middle) in growth medium before imaging. Live cell images were taken on a confocal microscope, TMR signal is

shown in red, AF-488 signal in green and nuclei are stained with Hoechst and shown in blue (only in A). Representative images are shown out of one experiment where 5 images were taken from each condition randomly. Scalebars represent 10  $\mu$ m.

### 2.3.5 Immunofluorescence and colocalization analyses reveal an endocytic uptake route for CPD-monobody adducts

Since the observed speckles were moving when observed in live HeLa cells, we hypothesized that they could be endosomes, suggesting an endocytic uptake route for CPD adducts at this concentration. We also wanted to assess whether the endocytosed adducts are translocated to the cytosol or degraded in lysosomes. I therefore performed colocalization experiments with antibodies against the early endosomal marker EEA1 as well as the lysosomal marker Lamp1. HeLa cells grown on coverslips were incubated with adducts for different durations and subsequently washed, fixed and permeabilized. The coverslips were then incubated with an antibody either against EEA1 or Lamp1, a secondary anti-mouse antibody coupled to Cy5 and the nuclear stain Hoechst. The coverslips were mounted on glass slides and analyzed using a confocal microscope. Images of 5 sections at the mid-nuclear plane were taken and analyzed using ImageJ and the JaCOP plugin developed by the BIOP at EPFL. I chose to measure the colocalization between CPDs and the antibody stain, since the 488 nm signal from the monobodies was not detectable at incubation times below 1 hour and fainter than the TMR signal at later timepoints. The 488 nm signal was much fainter in the fixed cells than in the previous experiments where live cells were observed.

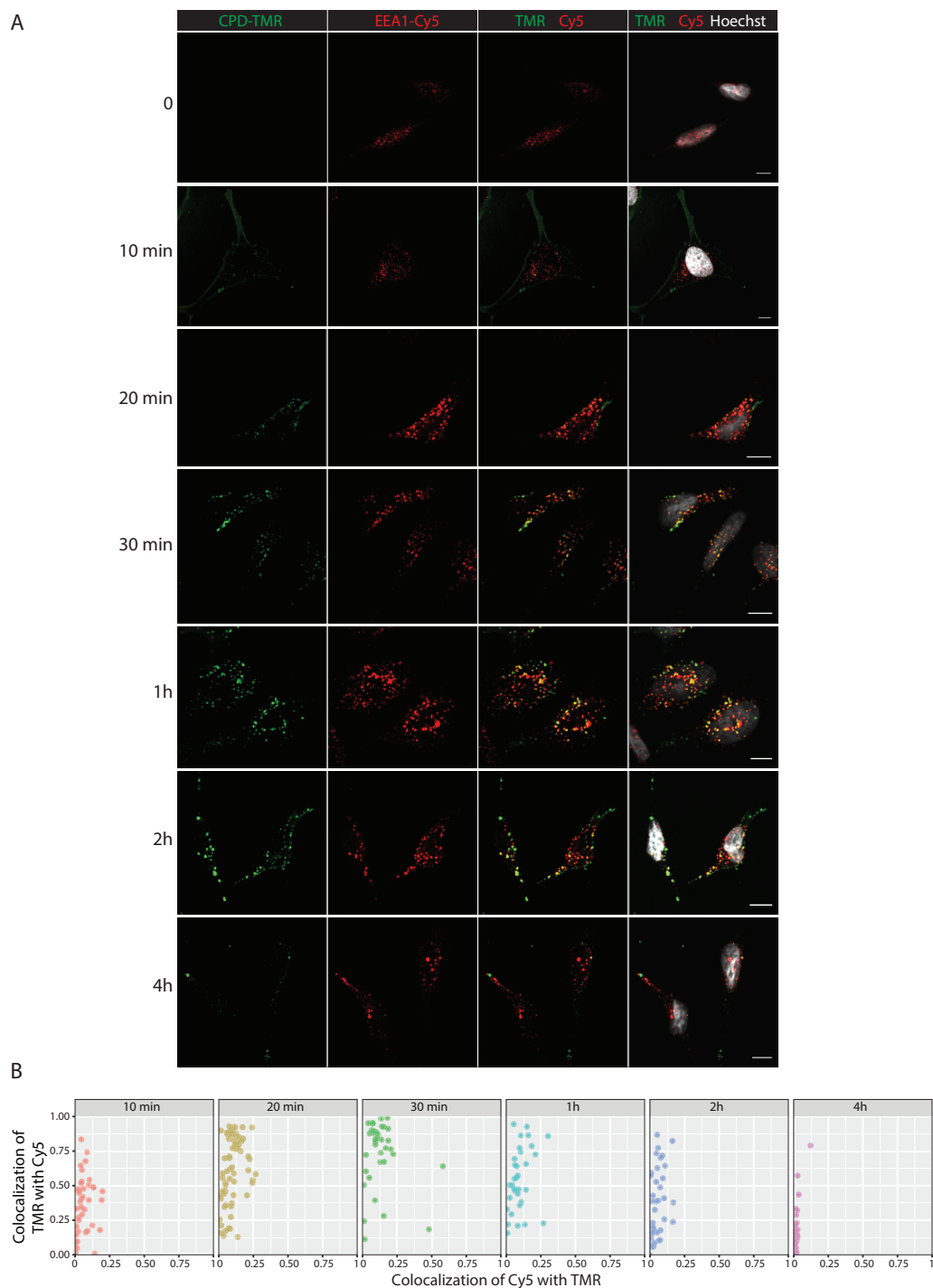
Briefly, cells were defined as regions of interest (ROIs) and throughout the 5 z-slices of each individual ROI, the Mander's overlap coefficient was calculated between the TMR and the Cy5 channel. The results were plotted as single dots, each representing a single cell and showing the incidence of the antibody colocalized with the CPDs (Mander's overlap coefficient 1) on the x-axis and the incidence of the CPDs colocalized with the antibody (Mander's overlap coefficient 2) on the y-axis. Both coefficients range from 0 to 1, 1 on the y-axis meaning that all the pixels of the TMR channel colocalize with pixels from the Cy5 channel.

As shown in Figure 2.20, the speckles corresponding to the TMR signal partially colocalize with the EEA1 signal after 20 minutes and up to 2 hours of incubation with the adducts, reaching its maximum after 30 minutes. This suggests that CPD adducts are taken up in early endosomes where they can be seen from 20 minutes to 2 hours after addition of the adducts to the cells. From there, they then translocate to a different compartment as the colocalization decreases between 30 minutes and 4 hours after addition of the adducts. (Figure 2.20)

In cells stained with the antibody against Lamp1, we observed an increasing colocalization between the TMR and the Cy5 channels with a maximum at 4h after addition of the adducts, which is the longest time-point analyzed. (Figure 2.21) This result indicates that the adducts are following a lysosomal degradation route, as the localization in lysosomes increases over time.

Collectively, these results suggest that CPDs coupled to monobodies mediate an endocytic uptake route, meaning that CPD-coupled monobodies are first taken up in endosomes, which fuse with lysosomes to promote their degradation. These results are in

contradiction to previously published data arguing that CPDs induce a direct translocation over the membrane, called thiol-mediated uptake. The idea is that CPDs bind to thiols on the cell surface by dynamic covalent disulfide exchange, upon which they translocate through micellar pores and get released into the cytosol by disulfide exchange with glutathione.<sup>114</sup>



**Figure 2.20 Colocalization analysis between CPDs and the early endosomal marker EEA1. A:** Confocal images of fixed HeLa cells after incubation with adducts for the indicated time-points, and staining with a primary antibody against EEA1 and a secondary Cy5-coupled antibody. TMR signal is shown in green, Cy5 signal in red and Hoechst staining in gray. Scalebars represent 10  $\mu$ m. Representative images are shown out of one biological experiment where 15 images were taken from each condition randomly. **B.:** Colocalization analysis using JaCOP tools and Fiji ImageJ software. Each dot represents a single cell and the Manders coefficient 1 is plotted on the x axis and 2 on the y axis.

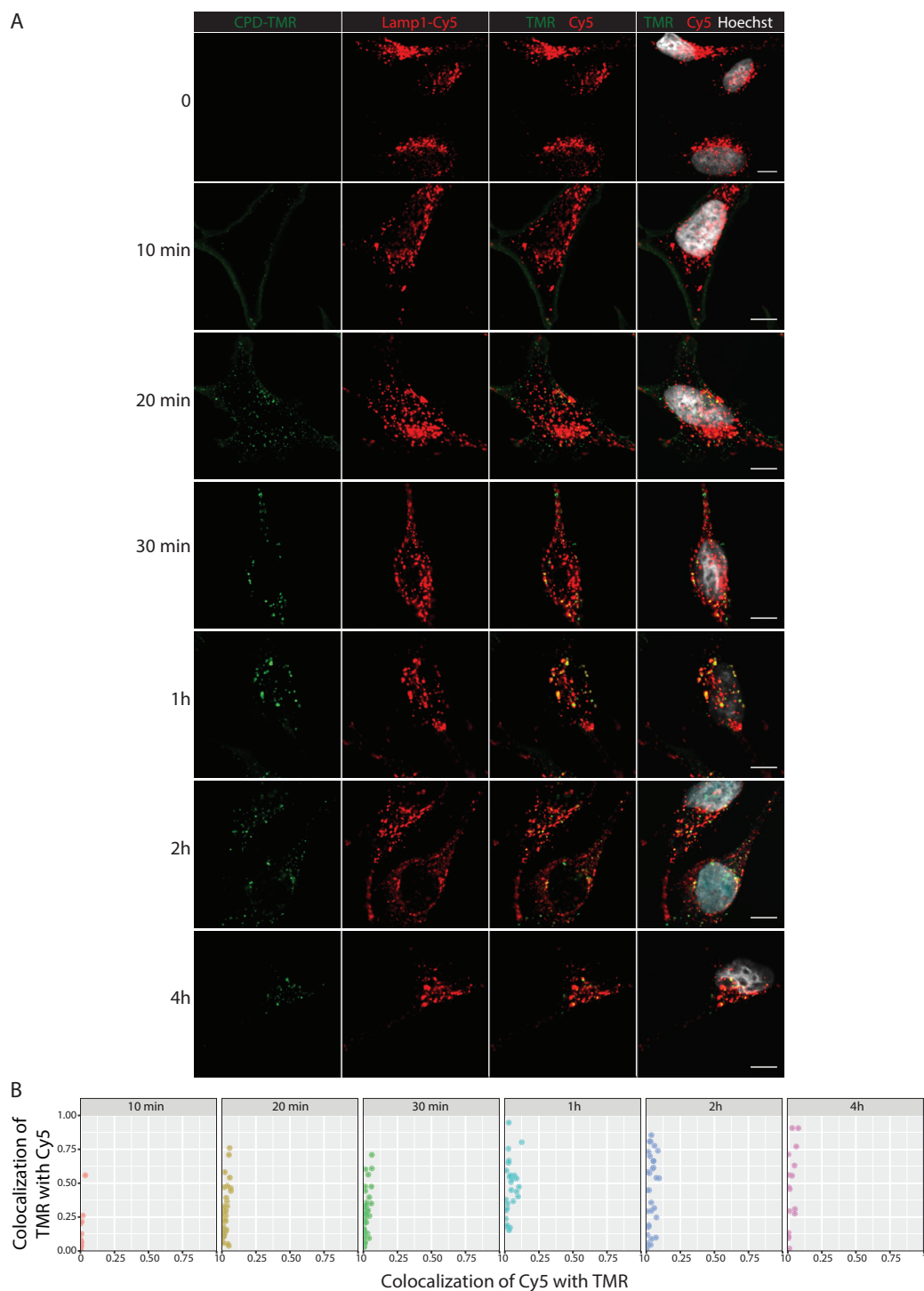


Figure 2.21 **Colocalization analysis between CPDs and the lysosomal marker Lamp1.** **A:** Confocal images of fixed HeLa cells after incubation with adducts for the indicated time-points, and staining with a primary antibody against Lamp1 and a secondary Cy5-coupled antibody. TMR signal is shown in green, Cy5 signal in red and Hoechst staining in gray. Representative images are shown out of one biological experiment where 15 images were taken from each condition randomly. Scalebars represent 10  $\mu\text{m}$ . **B.:** Colocalization analysis using JaCOP tools and Fiji ImageJ software. Each dot represents a single cell and the Mander's coefficient 1 is plotted on the x axis and 2 on the y axis.

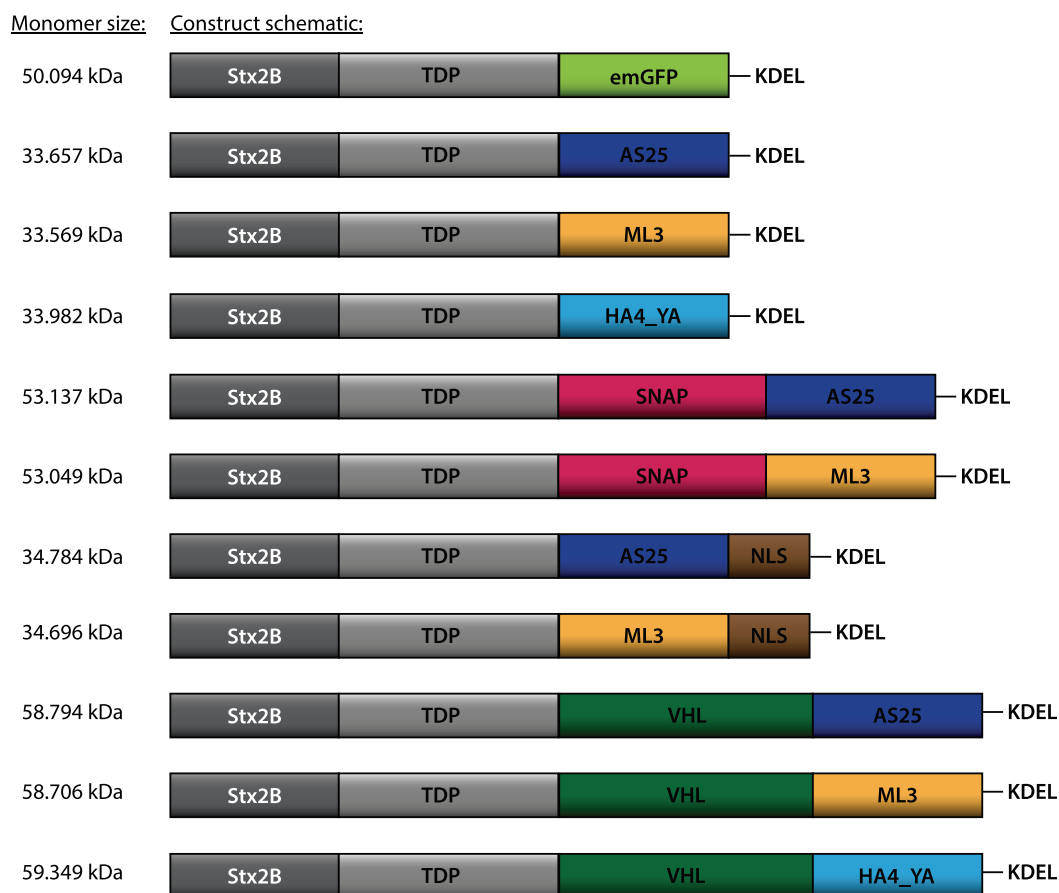
## 2.4 Monobody delivery using Bacterial Toxins

The lab of Prof. Hak-Sung Kim has developed an effective and receptor-specific protein delivery system by fusing the B-subunit of Shiga-like toxin (Stx2B) from *E. coli* to the translocation domain or domain II of Exotoxin A (TDP) from *P. aeruginosa*.<sup>191,199,213</sup> They have shown that functional enzymes and short peptides can be delivered into the cytosol using this bacterial toxin fusion construct. However, a small number of the delivered molecules are likely to be sufficient to elicit a cellular response when reaching the cytosol. The delivery of functional stoichiometric binders to endogenous proteins into the cytosol has not yet been demonstrated with this toxin construct. Moreover, the precise uptake mechanism and intracellular trafficking route of proteins delivered using the Stx2B-TDP fusion construct still remains to be fully elucidated. The authors hypothesized that they follow the same retrograde uptake route after endocytosis as Shiga toxin on its own, a quantitative analysis however remains to be done to fully answer this question. We therefore wanted to assess if this system could be used to deliver monobodies to the cytosol of cells, in sufficient amounts to have an inhibitory effect on their target protein. This project was done with the help of Katyayane Neopane, who did her master's thesis in our lab and whom I supervised. For simplicity reasons, the STx2B-TDP fusion protein is named Toxin in all the fusion constructs throughout this thesis.

### 2.4.1 Cloning and expression of Toxin-Monobody fusion proteins

In this project, we used the monobodies AS25, targeting the Bcr-Abl kinase, and ML3, targeting the Lck kinase, to study the efficiency of the delivery system as well as the consequences of the monobody delivery on signaling. In addition, we used the green fluorescent protein emGFP and the non-binding HA4\_Y87A monobody<sup>76</sup> as control proteins.

In a first step, we tested if recombinant fusion proteins composed of the two toxin subunits and a monobody can be expressed and purified. We genetically fused a monobody followed by a KDEL sequence to the C-terminus of the TDP. KDEL is an ER retention motif which is thought to facilitate the retrograde transport since it is required for the cytotoxicity of Exotoxin A and constructs comprising a C-terminal KDEL motif have shown enhanced cytosolic uptake.<sup>191,199,213</sup> In order to visualize the uptake of the monobody proteins, we used two different labelling techniques. A cysteine residue was introduced C-terminal to the monobody and N-terminal to the KDEL motif by site-directed mutagenesis, to allow the site-specific labelling of the recombinant protein using thiol-reactive maleimide-coupled fluorescent dyes. In addition, we generated a construct containing a SNAP-tag N-terminal to the monobody, allowing the labelling with benzylguanine (BG) as a substrate for the SNAP-tag. We also introduced a 6x-Histidine-tag (6xHis) N-terminal to the KDEL motif to allow the purification of the protein using Ni-NTA beads. In addition to these constructs, we generated constructs bearing a nuclear localization sequence C-terminal to the monobody and constructs with an the E3 ubiquitin ligase recruiting protein VHL in addition N-terminal to the monobody. The purification of these constructs as well as the experiments they were used for will be discussed in section 2.4.7 and 2.4.11, respectively. Figure 2.22 shows an overview of all the constructs generated.



**Figure 2.22 Schematic of all the constructs generated for the project with their size as monomeric proteins**

The resulting constructs were transformed into the *E. coli* strain Origami, which bears a mutation in a thiol reductase, helping to conserve the disulfide bonds necessary for the correct folding of the toxin subunits. Upon expression, the proteins were purified using a Ni-NTA gravity flow column and size exclusion chromatography (SEC). Since the Stx2B is forming pentamers, we expected our recombinant proteins to be pentameric in native conditions. Indeed, we could confirm the pentameric nature of the native proteins by SEC, where the largest peak corresponds to the molecular weight of the pentamer, according to the column calibration. (Figure 2.23 A-E) The purified fractions corresponding to the largest peak were analyzed on a Coomassie stained SDS-PAGE gel, where the denatured proteins appeared as a single band corresponding to the size of the monomeric fusion protein. (Figure 2.23 F-J) The integrity of the obtained proteins was further confirmed by an immunoblot using an anti-penta-His antibody. (Figure 2.23 K-O) All constructs were solubly expressed with reasonable yield (between 2 and 10 mg per liter of culture) and could be purified to good homogeneity (>90%).

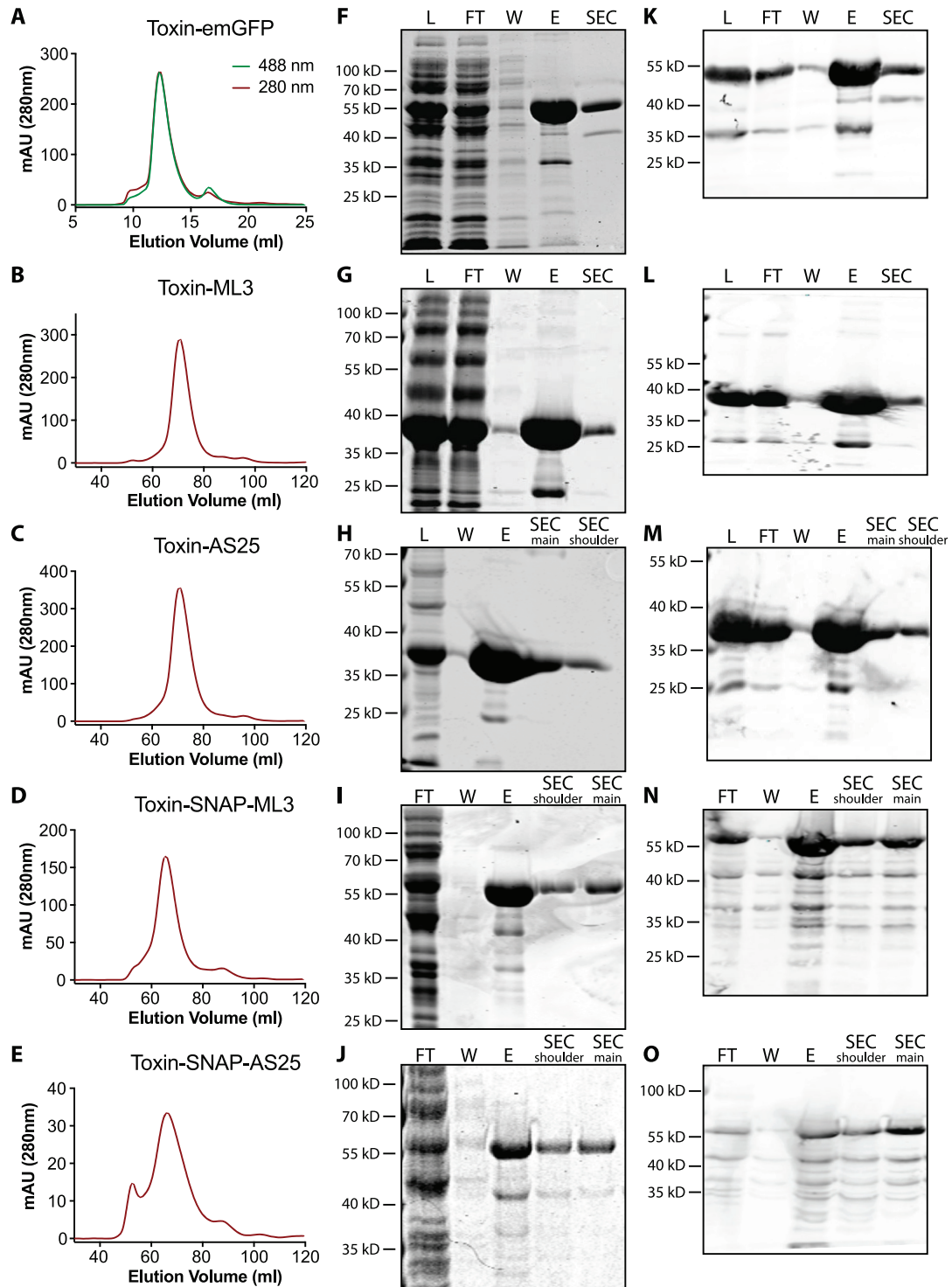


Figure 2.23 **Stx2B-TDP-Monobody fusion protein purification.** The crude bacteria lysate was first purified on a Ni-NTA column and then by SEC. **A-E:** Size exclusion chromatograms, **F-J:** Coomassie stained SDS-PAGE gels with the fractions from the Ni-NTA purification (L=crude lysate, FT=flow-through, W=wash, E=elution) and the main peak of the SEC after concentration, as well as in some gels the shoulder fraction from the SEC. **K-O:** Corresponding immunoblots with an antibody recognizing penta-His. Purified proteins: Stx2B-TDP-emGFP (**A, F, K**), Stx2B-TDP-ML3 (**B, G, L**), Stx2B-TDP-AS25 (**C, H, M**), Stx2B-TDP-SNAP-ML3 (**D, I, N**), Stx2B-TDP-SNAP-AS25 (**E, J, O**).

### 2.4.2 Uptake of fluorescently labelled Toxin-monobody fusion proteins in HeLa cells

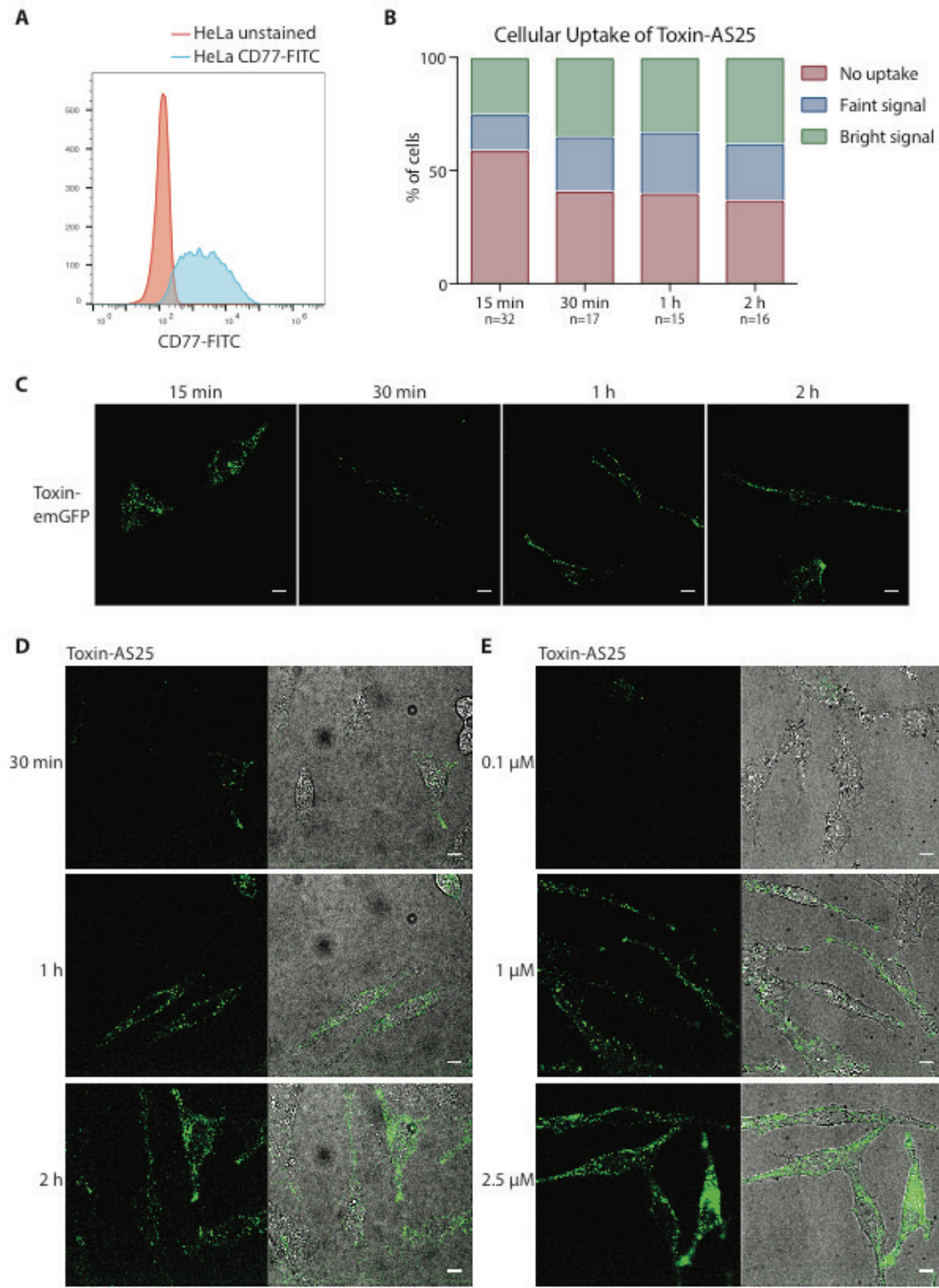
In order to assess the uptake efficiency of the purified proteins, we chose HeLa cells as a model cell type since they grow as adherent cells, have a large cytoplasm compared to the nucleus, and subcellular compartments can be distinguished. We determined whether they express the receptor for Stx2B - Gb3 - on their surface by staining them with a FITC coupled anti-Gb3 (CD77) antibody for flow cytometry analysis. Most of the HeLa cells are CD77 positive, but the amount of CD77 expression is broadly distributed. (Figure 2.24 A)

We first tested the uptake of Toxin-emGFP in HeLa cells, to verify that the recombinant toxin subunits are well folded and functional. Upon incubation with HeLa cells, fluorescent signal from the emGFP could be observed in the cytoplasm already after 15 minutes of incubation. (Figure 2.24 C) However, independently of the time of incubation, some cells did not display any measurable fluorescence, and some cells were brighter than others. This is in line with the broad distribution of the Gb3 expression level on the surface within a population of HeLa cells observed by flow cytometry. (Figure 2.24 A) In order to quantify the number of cells taking up protein over time, individual cells from each timepoint were categorized into either “no fluorescence”, “faint” or “bright”, according to the fluorescence intensity measured across the cell. (Figure 2.24 B) This categorization shows that the percentage of cells taking up protein increases by 20% between 15 and 30 minutes of incubation, and marginally between 30 minutes and 1 or 2 hours of incubation.

Recombinant Toxin-AS25 with a cysteine at the C-terminus of the monobody was labelled using maleimide-coupled Alexa488 fluorescent dye and purified using a desalting column to remove excess dye. We then incubated the labelled proteins with HeLa cells at different concentrations and for different time periods to assess the efficiency of uptake by confocal microscopy.

When incubating the cells with 1  $\mu$ M protein for different times, we observed an increase in the mean fluorescence over time. (Figure 2.24 D) Similarly, when increasing the concentration of the protein, the fluorescence intensity increased in most cells. While 0.1  $\mu$ M protein was hardly detectable in the cytosol, in cells incubated with 2.5  $\mu$ M protein, the fluorescence intensity was on average higher when compared to cells incubated with 1  $\mu$ M protein. (Figure 2.24 E) However, in line with the heterogenous uptake efficiency of Toxin-emGFP, also with Toxin-AS25, we observed differences in fluorescence intensity within a population treated for the same time and with the same concentration of protein.

Collectively, these results show the uptake of both emGFP and monobodies in cells which express the Shiga toxin receptor Gb3, and that the uptake is both concentration and time dependent.



**Figure 2.24 Toxin-monomobody delivery in HeLa cells.** **A:** Flow cytometry plot of HeLa cells stained with CD77-FITC. **B:** HeLa cells were incubated for different durations with AF488-labelled Toxin-AS25 and categorized based on the average fluorescence throughout the cell. The number of cells analyzed is indicated below each bar. **C:** HeLa cells were incubated with 1  $\mu$ M Toxin-emGFP for different times. **D:** HeLa cells were incubated with 1  $\mu$ M AF488-labelled Toxin-AS25 for different durations. **E:** HeLa cells were incubated with 0.1  $\mu$ M or 1  $\mu$ M or 2.5  $\mu$ M Toxin-AS25 for 1 h. Scale bars correspond to 10  $\mu$ M. Representative images are shown out of one experiment where 10 images were taken from each condition randomly.

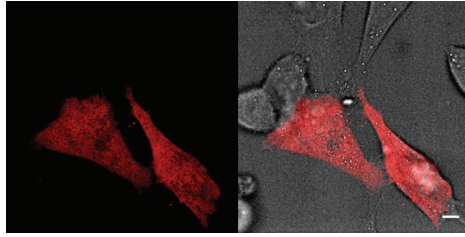
### 2.4.3 Uptake and labeling of SNAP-tagged Toxin-monobody fusion proteins in HeLa cells

In order to test the labeling of SNAP-tagged proteins inside cells using BG-coupled fluorophores, HeLa cells were transiently transfected with a plasmid driving the expression of SNAP-AS25 under a CMV promoter and incubated with 500 nM (cell-permeable) BG-Silicorhodamine (BG-SiR) 2 days after the transfection. The cells were observed by confocal microscopy and the transfected cells showed an evenly distributed fluorescence, already after 30 minutes of incubation with BG-SiR, whereas cells transfected with AS25 alone and incubated with BG-SiR did not show any fluorescence. (Figure 2.25 A) We next labelled recombinant Toxin-SNAP-AS25 with the cell-impermeable SNAP substrate BG-Alexa-Fluor-647 (BG-647) and purified the labelled protein using a desalting column. We then incubated HeLa cells with this labelled protein for different durations and observed them by confocal microscopy. In some cells we already observed a weak fluorescent signal after only 15 or 30 minutes of incubation, which became stronger after 1 hour of incubation with the protein. (Figure 2.25 B) The staining pattern and signal intensities at different incubation times are similar as with the AF-488-labelled Toxin-monobody and the Toxin-emGFP constructs, suggesting that the uptake mechanism and duration are similar.

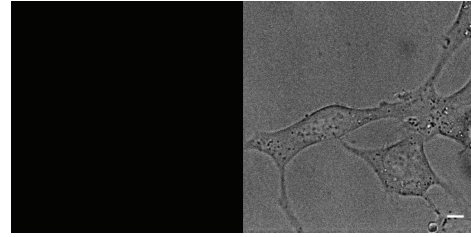
We next incubated the cells with unlabeled Toxin-SNAP-monobody constructs and added the cell-permeable BG-SiR after a washing step. The BG-SiR effectively stained the monobody inside the cells, resulting in a similar staining pattern as with the pre-labeled SNAP-tagged constructs. Control experiments with cells incubated with Toxin-SNAP-monobody and the cell-impermeable SNAP substrate BG-647 did not show any staining, and cells incubated only with BG-SiR but no proteins showed a very weak signal. (Figure 2.25 C) These experiments show that we can use the SNAP-tag as an alternative means to label recombinant toxin-monobody fusions.

**A** Transfection and staining with Bg-SiR

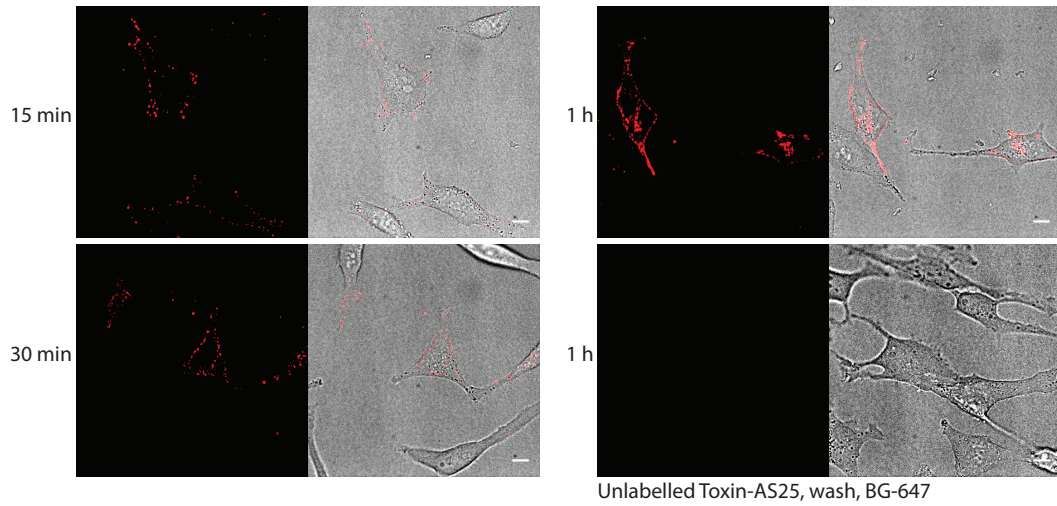
SNAP-AS25



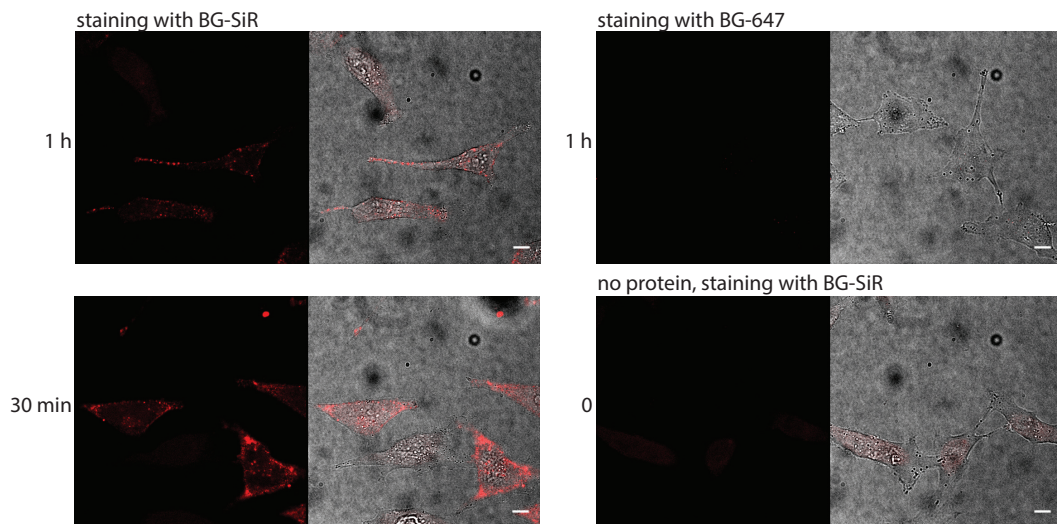
AS25



**B** Incubation with BG-647-labelled Toxin-SNAP-AS25



**C** Incubation with unlabelled Toxin-SNAP-ML3



**Figure 2.25 Delivery of SNAP-tagged Toxin-monobody constructs in HeLa cells.** **A:** HeLa cells were transfected with SNAP-AS25 (left) or AS25 (right) and incubated with BG-SiR for 30 minutes 2 days post transfection. **B:** HeLa cells were incubated with Toxin-SNAP-AS25 prelabelled with BG-647 for the indicated durations. The lower right image shows HeLa cells incubated with the same unlabeled construct for 1h, washed and subsequently incubated with BG-647 for 1h. **C:** HeLa cells were incubated with unlabelled Toxin-SNAP-ML3 for the indicated durations, washed and incubated with BG-SiR for 30 minutes. The upper right image shows HeLa cells incubated with the same construct followed by an incubation with BG-647. The lower right image shows HeLa cells incubated only with BG-SiR, but no protein. All images were taken of live cells on a confocal microscope. Scalebars correspond to 10  $\mu$ M.

Representative images are shown out of one (A) or two (other panels) biological repeats where 10 images were taken from each condition randomly.

We next wanted to assess whether the staining pattern we observed in the previous experiments changes with longer incubation times. Since most of the signal was seen in distinct speckles and not evenly distributed throughout the cytosol when we incubated cells with Toxin-monobody constructs for up to 4 hours, we wanted to test if the signal becomes more distributed when the cells are incubated with the protein for longer durations. We therefore labelled Toxin-SNAP-AS25 with BG-647 and incubated it with HeLa cells for 1, 2.5, 9 or 24 hours or we incubated the cells with the same protein for 1 hour, followed by a washing step and further incubation in growth medium for 1, 2.5, 9 or 24 hours. (Figure 2.26)

We observed the same staining pattern for each duration of incubation, and even after prolonged incubation, the strongest fluorescent signal came from distinct speckles, with only a very faint signal which was dispersed throughout the cells. (Figure 2.26) The signal intensity increased with longer incubation times with the labelled protein. However, when the cells were washed and then further incubated in growth medium, the signal strength stayed the same for up to 24 hours. (Figure 2.26) This result indicates that the protein stays within the cells for at least 24 hours after it has been removed from the medium and is not degraded or exported out of the cells.

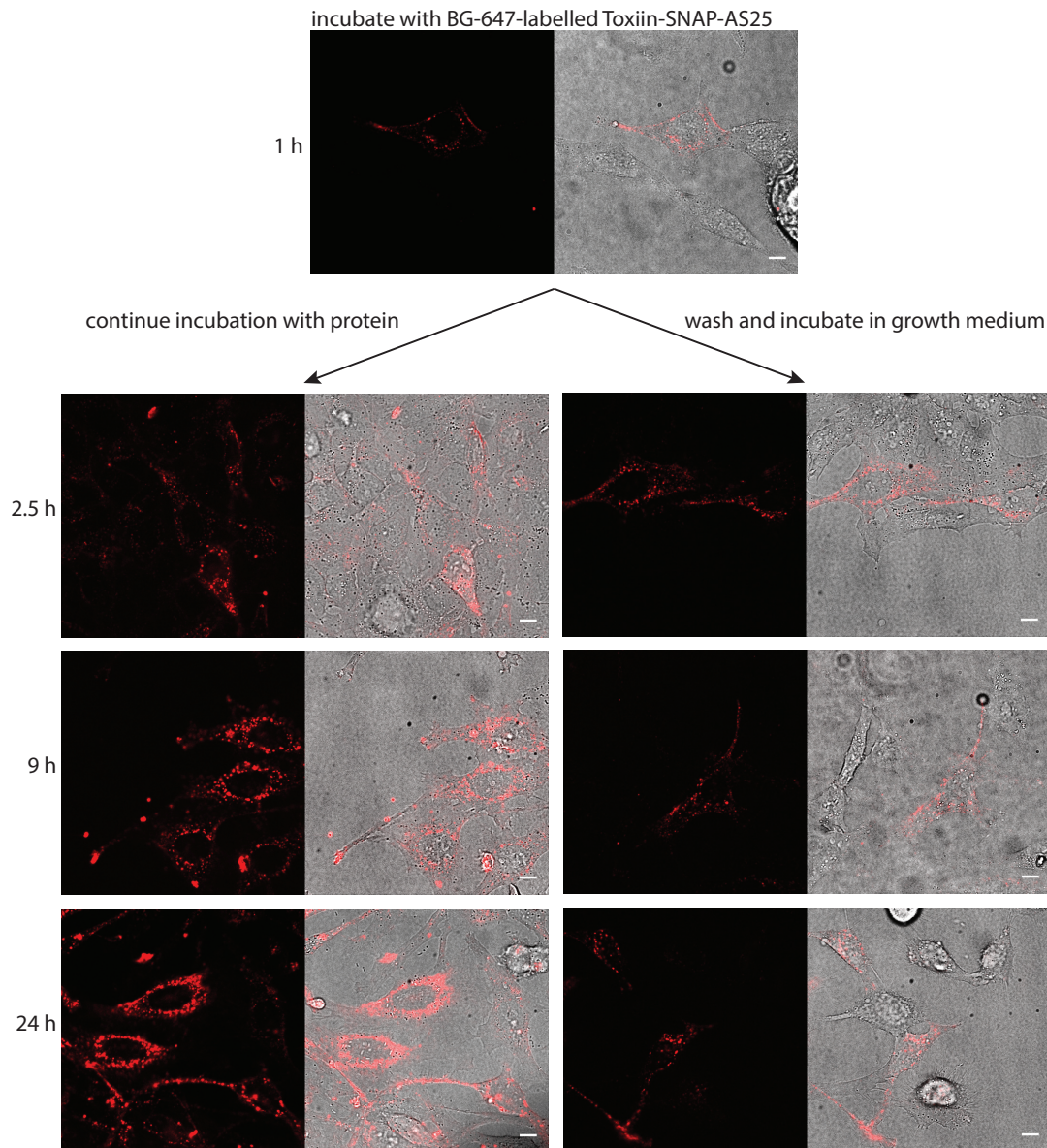


Figure 2.26 **SNAP-tagged Toxin-monobody constructs remain in cells for 24h after delivery.** HeLa cells were incubated for the indicated durations with BG-647-labelled Toxin-SNAP-AS25 (left column) or washed after 1h of incubation with the protein and incubated in medium for the indicated duration (right column). Live cells were imaged on a confocal microscope. Scale bars correspond to 10  $\mu$ m. Representative images are shown out of one experiment where 10 images were taken from each condition randomly.

#### 2.4.4 Little colocalization between the delivered protein and Lysotracker

In order to test if the delivered proteins employ a retrograde trafficking route to reach the cytosol, like native Shiga toxin and Exotoxin A, we set out to do a colocalization analysis with markers for different endocytic compartments in the cell.

Since we observed most of the fluorescent signal in distinct speckles, we first tested if these speckles could correspond to lysosomes, meaning that the protein would be

endocytosed and degraded in lysosomes instead of translocating to the TGN and ER via the retrograde trafficking route.

We stained HeLa cells with the live cell marker Lysotracker Red for 30 minutes, after having incubated them for different periods with AF488-labelled Toxin-AS25. Figure 2.27 shows that at the analyzed time-points, little of the AF488 signal colocalizes with the lysotracker signal, indicating that most of the lysosomes do not contain the delivered protein. (Figure 2.27)

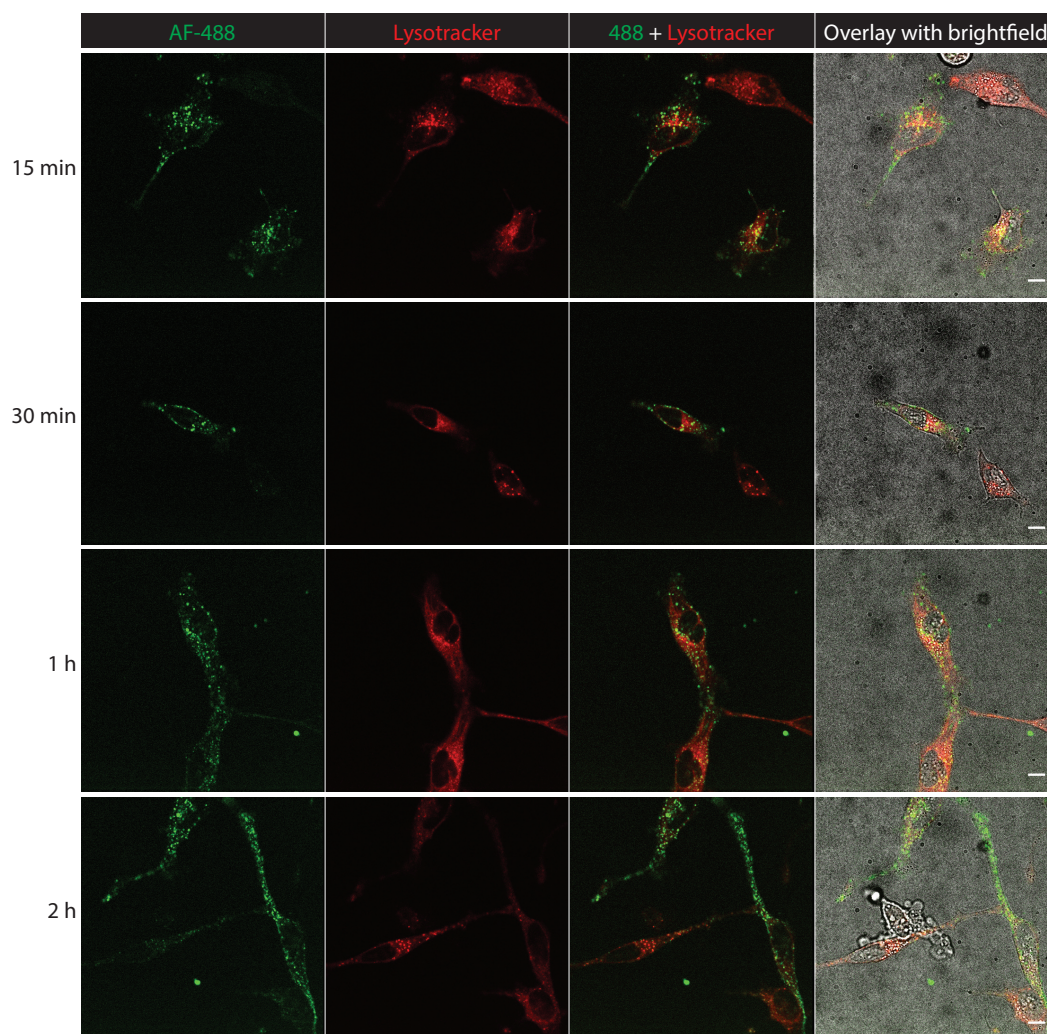


Figure 2.27 **The majority of delivered Toxin-AS25 does not colocalize with the Lysotracker Red marker.** HeLa cells were incubated with AF-488-labelled Toxin-AS25 for the indicated time periods and subsequently with Lysotracker red for 30 minutes. Images were taken with a confocal microscope. Scalebars correspond to 10  $\mu$ m. Representative images are shown out of one experiment where 15 images were taken from each condition randomly.

#### 2.4.5 Toxin-AS25 partially colocalizes with ER-tracker

We next wanted to test whether the delivered monobodies are taken up into the ER via the retrograde trafficking pathway. We therefore incubated HeLa cells with the live cell marker ER-tracker Red after incubation with AF-488-labelled Toxin-AS25 for 30 minutes or for 1, 2 or 4 hours.

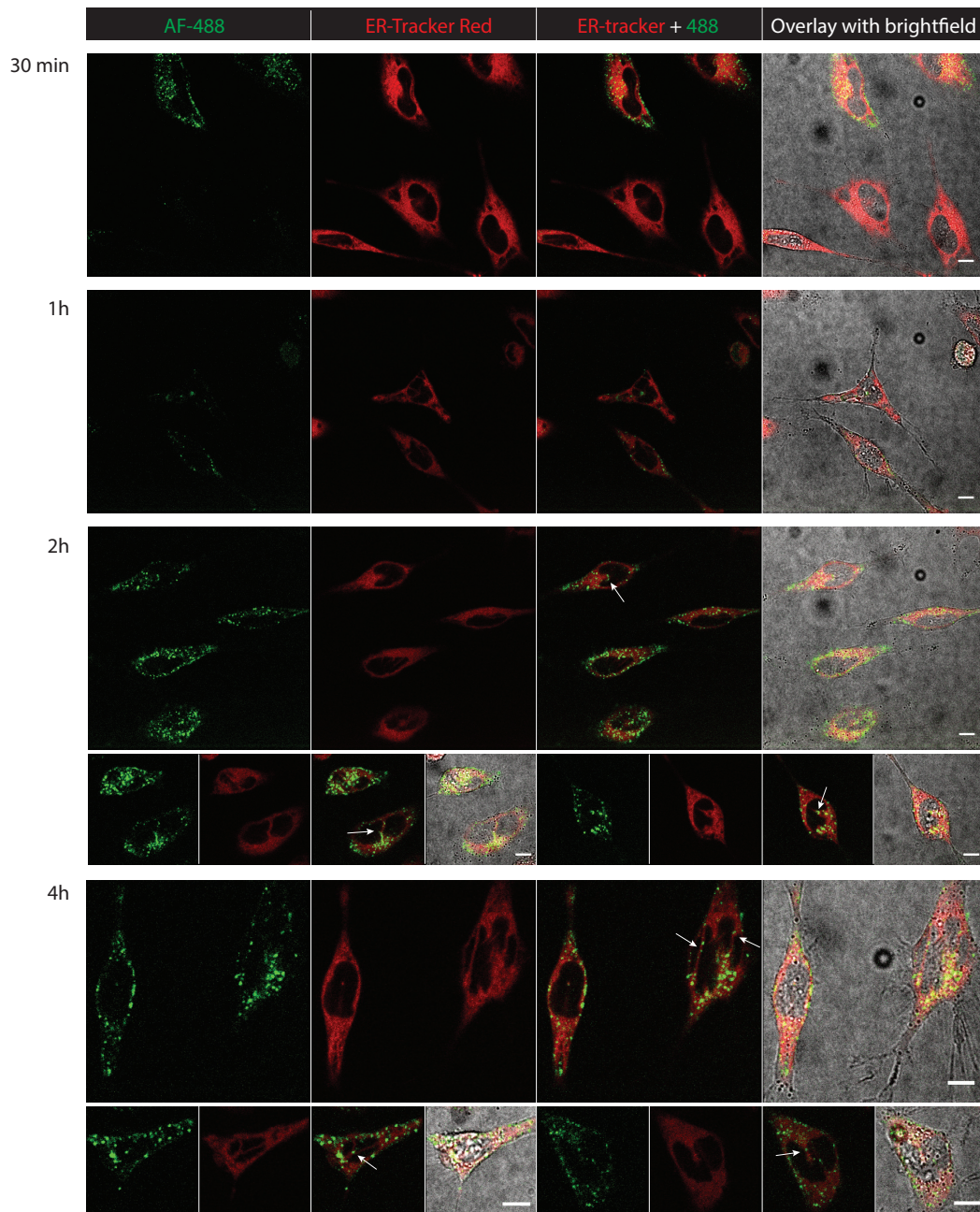


Figure 2.28 **Toxin-AS25 partially colocalizes with ER-tracker Red.** HeLa cells were incubated for the indicated durations with 1  $\mu$ M AF488-labelled Toxin-AS25 and subsequently washed and incubated with ER-tracker for 20 minutes. Arrows indicate particularly visible areas where the AF488 and ER-tracker labels colocalize. Scale bars correspond to 10  $\mu$ m. Representative images are shown out of one experiment where 15 images were taken from each condition randomly.

Figure 2.28 shows that after 30 minutes as well as 1 hour of incubation with the protein, most of the AF-488 signal does not colocalize with the ER-tracker signal. However, after 2 and 4 hours of incubation, some of the green signal colocalizes with ER-tracker, which is particularly visible where the ER structures are thin. (Figure 2.28, indicated by arrows) The ER-tracker stained almost the whole cytoplasm in most cells, making it difficult to assess whether the AS25 really colocalizes with the ER or is in close proximity. Nevertheless, the images taken after 2 and 4 hours indicate that the delivered protein

could indeed be partially localized in the ER, showing that it is translocated via the retrograde trafficking route.

#### 2.4.6 Immunofluorescence imaging and quantitative colocalization analysis between the delivered monobody and endocytic compartments

Both the live cell markers ER-tracker and LysoTracker resulted in a high background signal, which made it difficult to quantify the colocalization between the respective live cell marker and the signal of the delivered protein. We additionally wanted to test whether the protein is taken up into early endosomes, for which no live cell marker exists. We therefore decided to do immunofluorescence experiments with fixed cells after protein delivery, to be able to perform a quantitative colocalization analysis thereafter. We set out to do a pulse chase experiment, where HeLa cells were incubated at 37°C with BG-647-labelled Toxin-SNAP-AS25 for 10 minutes, after which they were washed to stop any further uptake and incubated in growth medium. The cells were fixed at different time points thereafter, with one hour being the longest time point, based on the finding that Toxin-emGFP can be observed in the ER after 30 minutes by the group of professor Hak-Sung Kim.<sup>199</sup>

The cells were subsequently fixed and permeabilized to allow the penetration of specific antibodies against EEA1 to stain early endosomes (Figure 2.29 A), Lamp1 for lysosomes (Figure 2.30 A), Golgi97 to stain the Trans-Golgi-network (TGN) (Figure 2.31 A) and the ER resident protein Calreticulin (Figure 2.32 A). All antibodies used were first tested using different fixation and permeabilization reagents and then titrated to determine the optimal concentration for the staining. Controls with the FITC or AF-488-coupled secondary antibodies alone were performed to determine the background staining from the secondary antibodies, which was negligible in all cases.

The samples were imaged using a confocal microscope and 5 sections were taken at the nuclear plane. Images were thereafter analyzed using ImageJ and the JaCOP plugin developed by the BIOP at EPFL. Briefly, cells were defined as regions of interest (ROIs) and a threshold algorithm was applied to distinguish true signal from background noise for both the delivered protein signal and the subcellular compartment marker antibody signal. The Mander's overlap coefficient was then calculated between the signal of the 647 nm channel corresponding to the delivered protein and the signal of the 488 nm channel corresponding to the compartment marker stain after application of the threshold algorithm. Representative images showing the result of each threshold algorithm are shown in the panel B of the Figure 2.29 to Figure 2.32. The results were plotted as single dots, each representing a single cell and showing the incidence of the delivered protein colocalized with the marker (Mander's overlap coefficient 1) on the x-axis and the incidence of the marker colocalized with the delivered protein (Mander's overlap coefficient 2) on the y-axis. Both coefficients range from 0 to 1, 1 on the x-axis meaning that all the pixels of the 647 nm channel colocalize with pixels from the 488 nm channel. (Figure 2.29 to Figure 2.32, panel C)

Figure 2.29 A shows that there is a colocalization of part of the signal from the antibody against EEA1 (green) with the AF-647 signal from the delivered protein (red). The quantitative analysis indicates that the colocalization increases after 10 and 30 minutes of incubation after the protein delivery and decreases again after 1 hour of incubation in growth medium. (Figure 2.29 C) This result indicates that the delivered proteins are

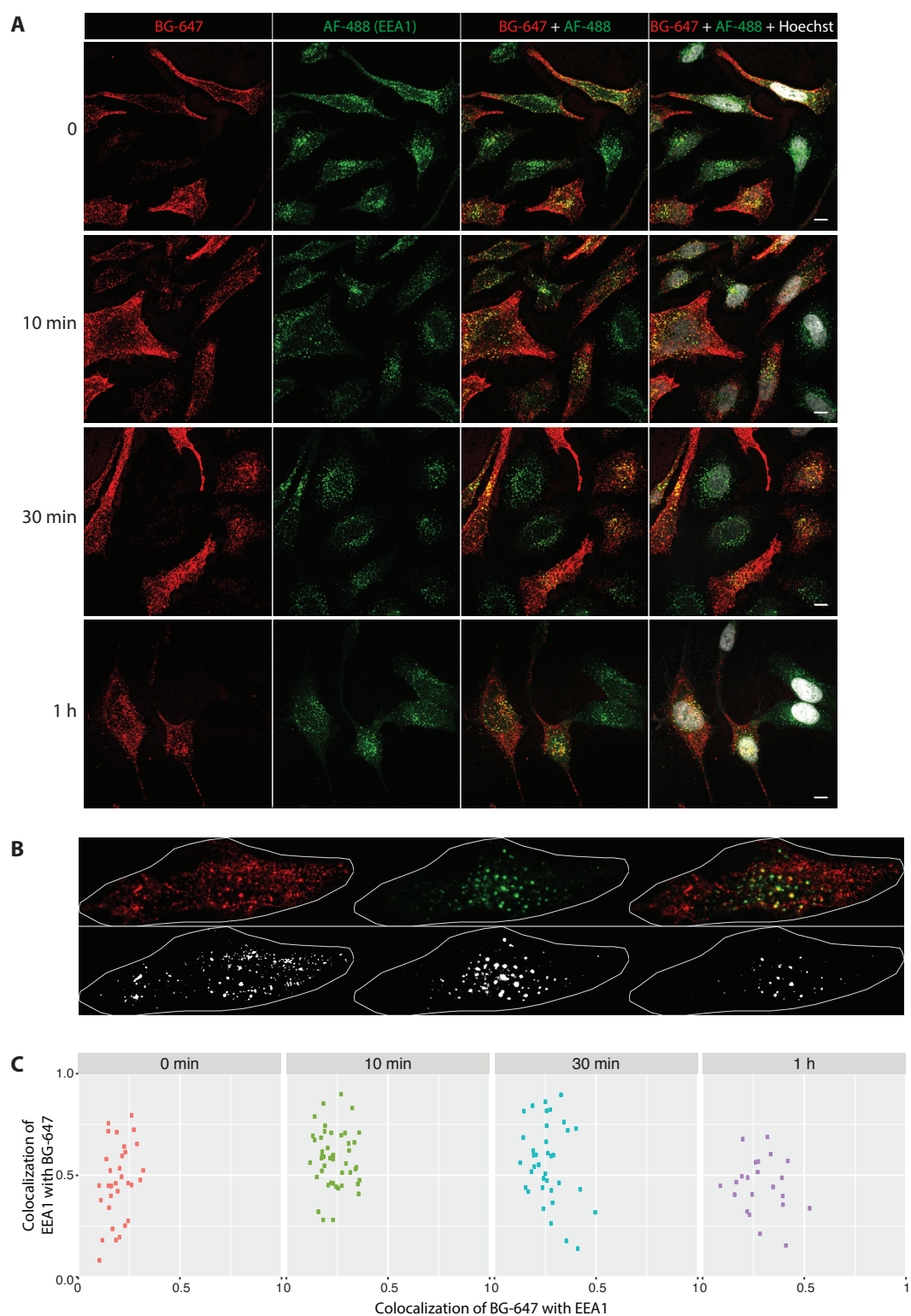
indeed endocytosed and leave the early endosomes again between 30 minutes and 1 hour after their internalization.

There is very little colocalization of the protein with the signal from the Lamp1 antibody, suggesting that the early endosomes containing the proteins do not fuse with lysosomes and that the delivered monobodies are thus not degraded by lysosomal proteases. (Figure 2.30)

The marker for the TGN, Golgin97, produces a high background signal, despite optimization attempts, and a careful evaluation of the threshold algorithm was needed to remove all the cells from the analysis, where part of the background signal was selected by the algorithm. In the remaining cells, only the strongest signal was selected as true signal. (Figure 2.31 B) A colocalization of the Golgin97 signal with the protein was observed in part of the cells 10 minutes after the protein uptake, whereas 30 minutes or 1 hour after the delivery, very little colocalization between the two signals was observed, suggesting that the protein did not reach the TGN at the studied timepoints. (Figure 2.31 C)

The staining with the antibody against Calreticulin resulted in a strong signal throughout almost the whole cytoplasm, which made it difficult to find an algorithm capable of distinguishing between the ER and the rest of the cytosol. (Figure 2.32 Figure 2.19 A) The selected threshold algorithm only selected the pixels with the strongest signal (Figure 2.32 figure 2.19 B) and almost no colocalization was measured between the antibody and protein signals at any of the timepoints. (Figure 2.32 Figure 2.19 C) Another antibody against Calnexin was also tested and resulted in a very similar staining pattern than the Calreticulin antibody. This result indicates that the protein does not reach the ER at the studied timepoints, which is in line with the previous analysis using ER-tracker, with which the delivered protein only partially colocalized after 2 to 4 hours of incubation. However, the results need to be evaluated carefully, since the staining pattern of the antibody is not sufficiently distinct to allow a precise colocalization analysis.

Collectively, the results of this colocalization analysis show that the delivered protein is taken up in early endosomes, after which it is not degraded in lysosomes but presumably translocated to a different endocytic compartment. The analysis does not allow any conclusions about the question, if the delivered proteins follow a retrograde trafficking route after their uptake, as the protein is not observed to be localized in the TGN or ER for up to one hour after the delivery, and longer timepoints might be needed to answer this question.



**Figure 2.29 Colocalization of BG-647 labelled Toxin-SNAP-AS25 with early endosomes in HeLa cells.** **A:** HeLa cells were incubated with the protein for 10 minutes, washed, incubated in growth medium and fixed after 0, 10 or 30 min or 1h. The cells were permeabilized and stained with a primary antibody against EEA1, a secondary anti-mouse antibody coupled to FITC and the nuclear stain Hoechst. Scale bars represent 10 $\mu$ m. Representative images are shown out of one experiment where 15 images were taken from each condition randomly. **B:** Result of the applied threshold algorithm, showing the fluorophore signal in the upper panel and the corresponding pixel selection by the threshold algorithm in the lower panel for one

representative ROI. **C:** Result of the colocalization analysis showing the average of 5 slices for each ROI, represented as a single dot.

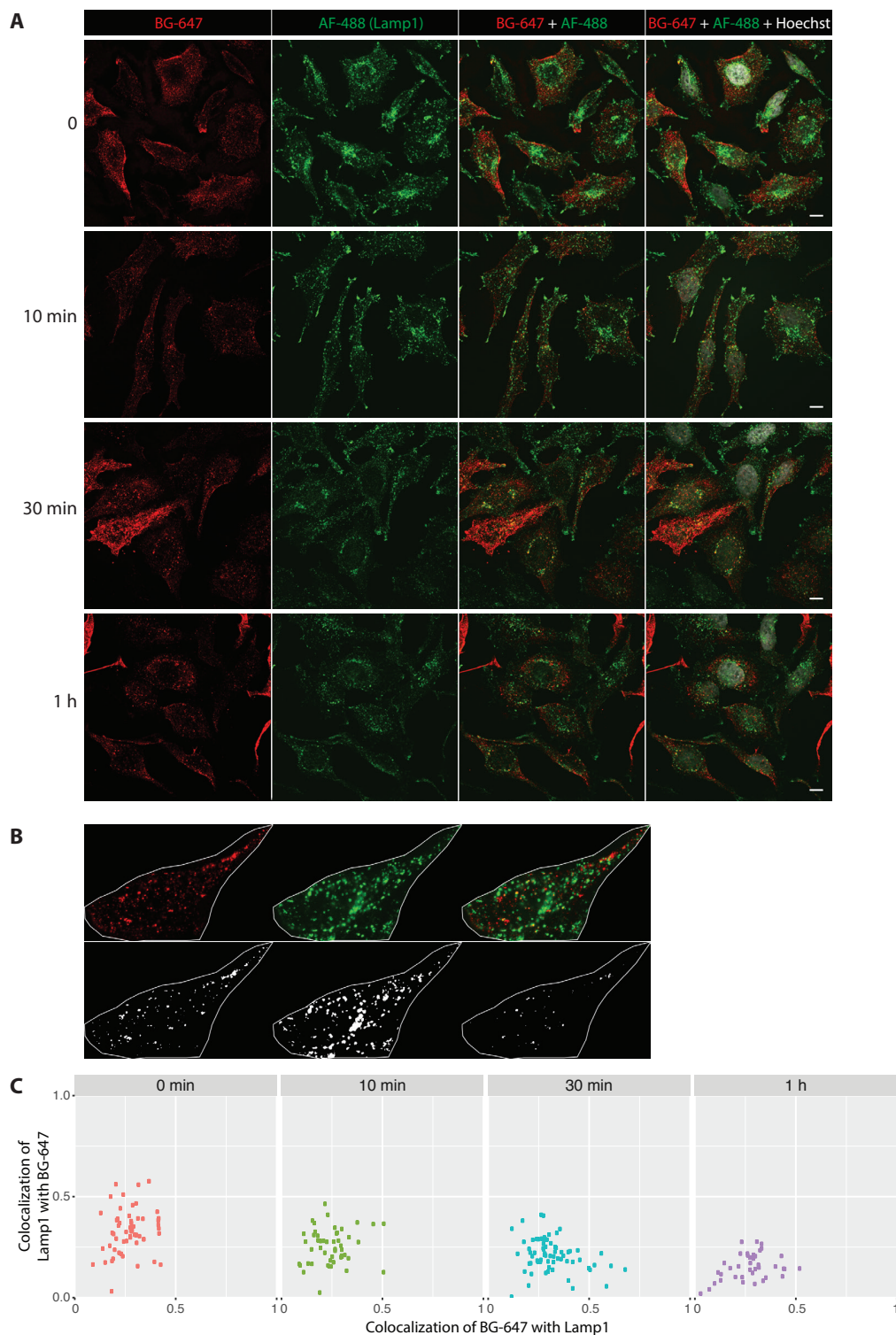


Figure 2.30 **Colocalization of BG-647 labelled Toxin-SNAP-AS25 with lysosomes in HeLa cells.** **A:** HeLa cells were incubated with the protein for 10 minutes, washed, incubated in growth medium and fixed after 0, 10 or 30 min or 1h. The cells were permeabilized and stained with a primary antibody against Lamp1, a secondary anti-mouse antibody coupled to FITC and the nuclear stain Hoechst. Scale bars represent 10 $\mu$ m. Representative images are shown out of one experiment where 15 images were taken from

each condition randomly. **B:** Result of the applied threshold algorithm, showing the fluorophore signal in the upper panel and the corresponding pixel selection by the threshold algorithm in the lower panel for one representative ROI. **C:** Result of the colocalization analysis showing the average of 5 slices for each ROI, represented as a single dot.

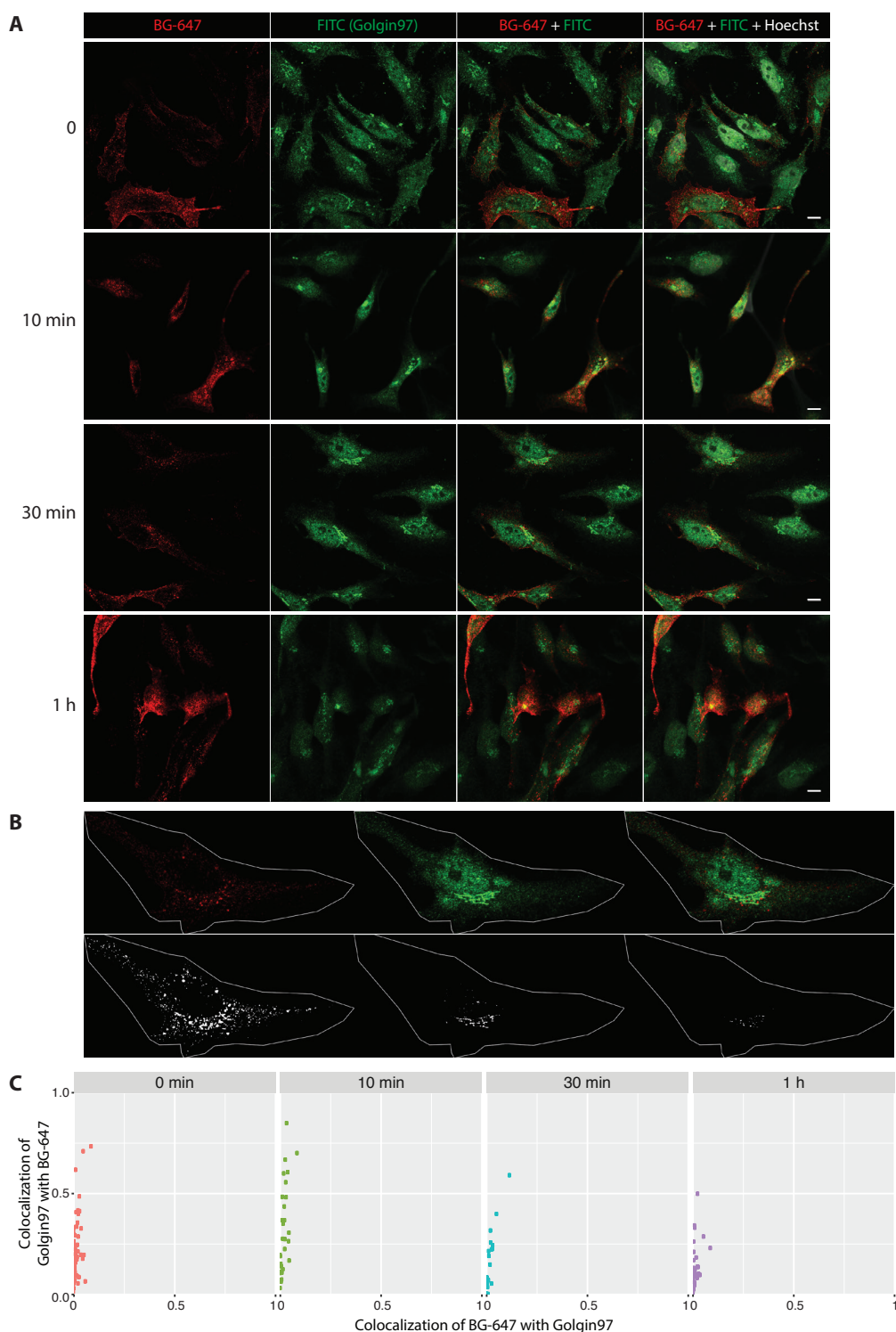


Figure 2.31 **Colocalization of BG-647 labelled Toxin-SNAP-AS25 with the TGN in HeLa cells.** **A:** HeLa cells were incubated with the protein for 10 minutes, washed, incubated in growth medium and fixed after 0, 10 or 30 min or 1h. The cells were permeabilized and stained with a primary antibody against Golgin97, a secondary anti-rabbit antibody coupled to Alexa Fluor 488 (AF-488) and the nuclear stain

Hoechst. Scale bars represent 10 $\mu$ m. Representative images are shown out of one experiment where 15 images were taken from each condition randomly. **B:** Result of the applied threshold algorithm, showing the fluorophore signal in the upper panel and the corresponding pixel selection by the threshold algorithm in the lower panel for one representative ROI. **C:** Result of the colocalization analysis showing the average of 5 slices for each ROI, represented as a single dot.

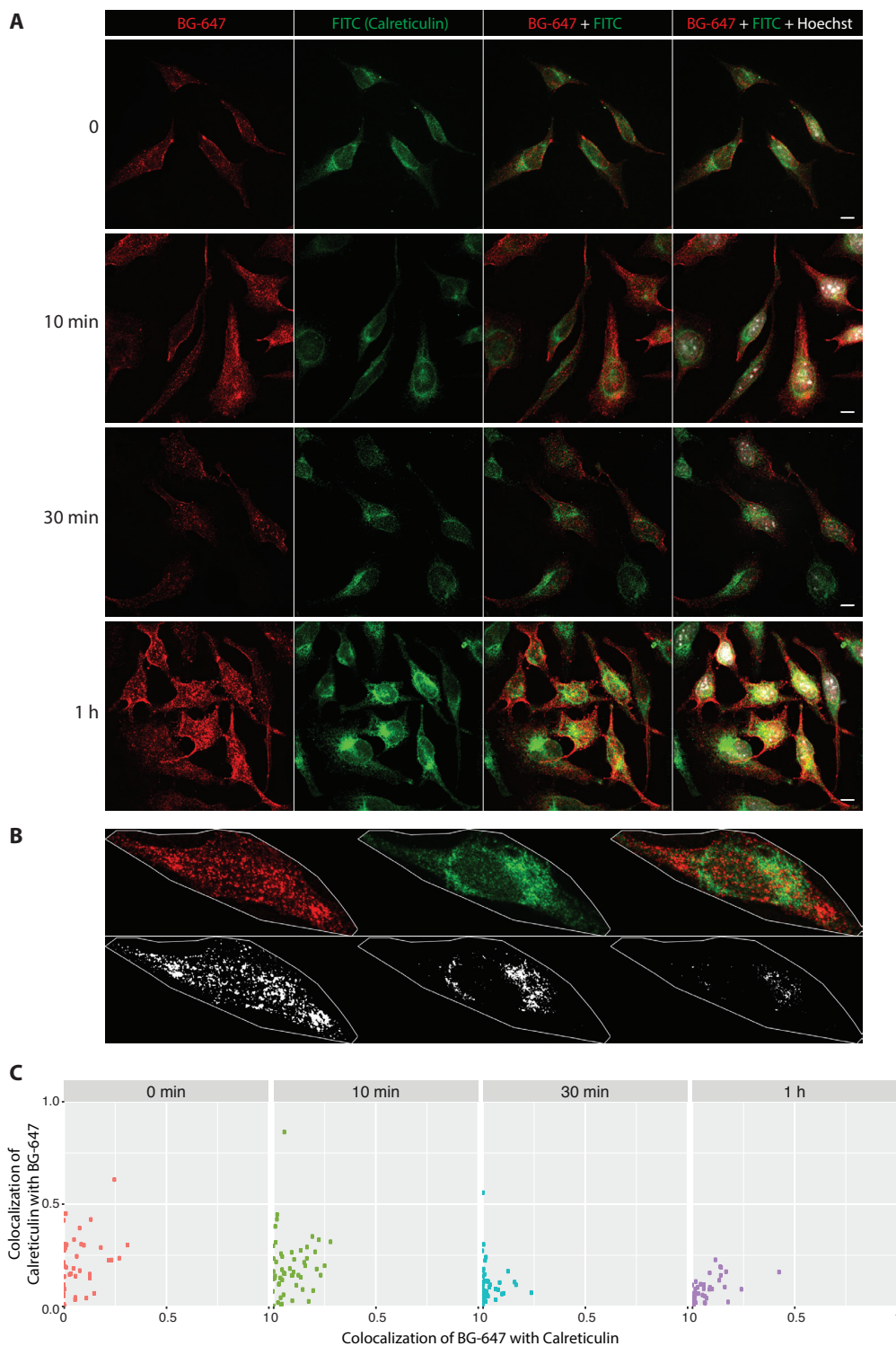


Figure 2.32 **Colocalization of BG-647 labelled Toxin-SNAP-AS25 with the ER in HeLa cells.** **A:** HeLa cells were incubated with the protein for 10 minutes, washed, incubated in growth medium and fixed after 0, 10 or 30 min or 1h. The cells were permeabilized and stained with a primary antibody against

Calreticulin, a secondary anti-rabbit antibody coupled to Alexa Fluor 488 (AF-488) and the nuclear stain Hoechst. Scale bars represent 10 $\mu$ m. Representative images are shown out of one experiment where 15 images were taken from each condition randomly. **B:** Result of the applied threshold algorithm, showing the fluorophore signal in the upper panel and the corresponding pixel selection by the threshold algorithm in the lower panel for one representative ROI. **C:** Result of the colocalization analysis showing the average of 5 slices for each ROI, represented as a single dot.

#### 2.4.7 Delivery of NLS-tagged fusion proteins into HeLa cells

We next wanted to assess if part of the delivered monobody protein reaches the cytosol. We hypothesized that the fluorescent signal from proteins located in the cytosol could be masked by the much stronger signal from vesicles and therefore difficult to detect. The same amount of protein in vesicles than in the cytosol would also result in a weaker signal from the cytosol as the volume is bigger and the protein more dilute. We therefore added a nuclear localization signal (NLS) to our proteins, so that after reaching the cytosol, they would be transported to the nucleus, where we could easily distinguish them from vesicles in the cytoplasm. Since NLS-containing proteins are recognized in the cytosol by importins, to be translocated through nuclear pores, the delivery of protein to the nucleus means that this protein must have reached the cytosol after delivery.

We cloned the NLS sequence PKKKRKVED to the C-terminal of the monobodies, N-terminal of the KDEL sequence and expressed and purified the encoded protein in *E. coli* Origami cells as for the other toxin-monobody fusions, obtaining similar yields and purity. We then labelled the resulting proteins using the cysteine-reactive maleimide coupled Alexa Fluor 488 and purified the labelled protein as described previously. The protein was incubated with HeLa cells for 2.5 hours, then the cells were washed to remove any protein that was not internalized, and live HeLa cells were observed using the confocal microscope. The cells were further incubated in growth medium thereafter and observed again after a total of 7 hours and 24 hours of incubation. Images of 5 nuclear plane sections were taken for each position and processed using Fiji ImageJ. Figure 2.33 shows a representative image for each condition.

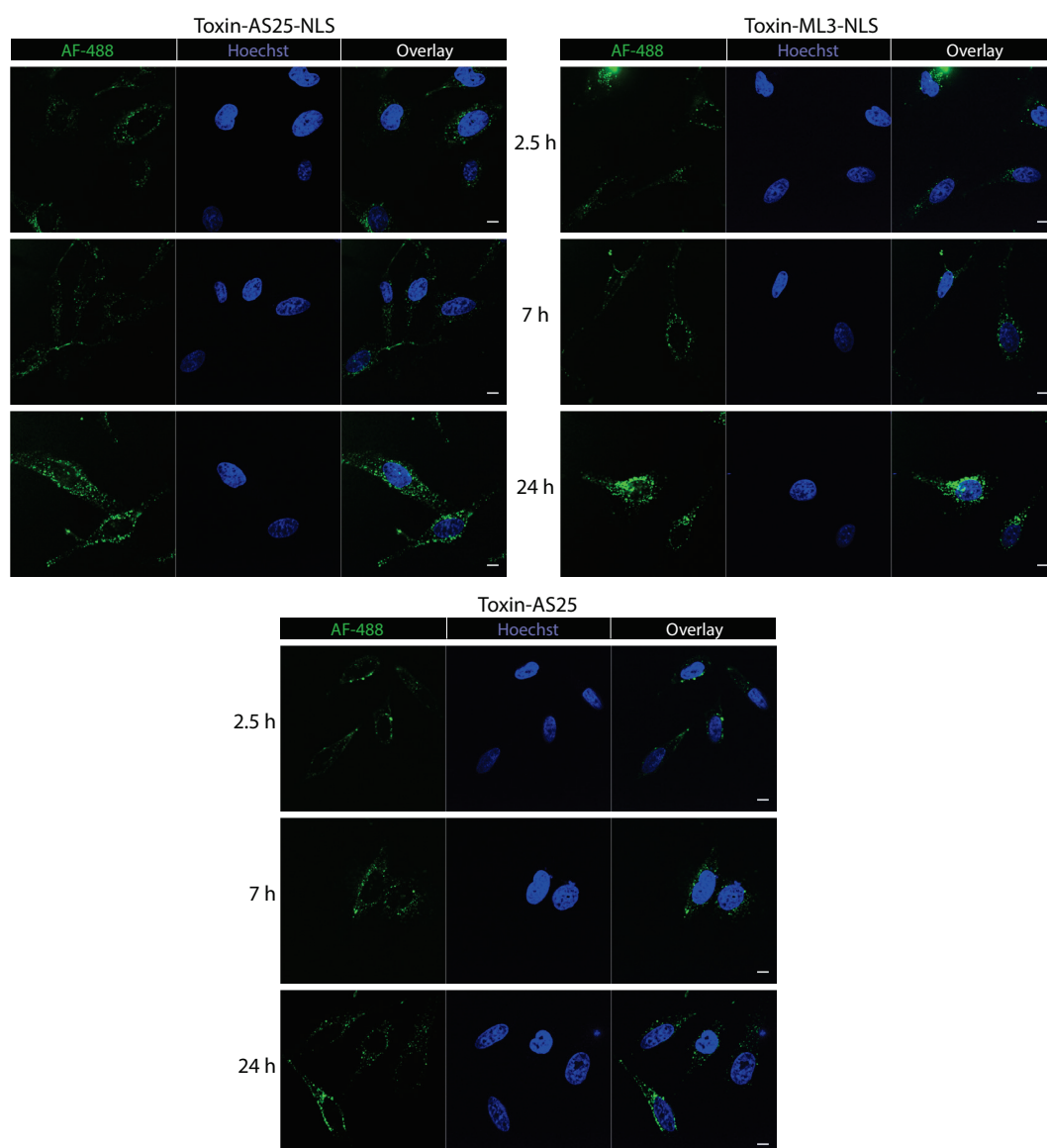


Figure 2.33 **Uptake of NLS-tagged Toxin-monobody proteins in the nucleus.** HeLa cells were incubated with AF-488 labelled NLS-Toxin-AS25 or NLS-Toxin-ML3 or Toxin-AS25 without NLS for 2.5h, washed and incubated in growth medium, so that the total incubation time was 7h or 24h. The cells were then stained with Hoechst and imaged using a confocal microscope. Scalebars correspond to 10 μm. Representative images are shown out of one experiment where 15 images were taken from each condition randomly.

In order to quantify the fluorescence intensity in the nucleus of each cell, I wrote a macro script in Image J, measuring the intensity of the 488 nm signal in each nucleus defined by the Hoechst signal. A description as well as the entire macro code can be found in section 4.20.2. The results of this analysis are represented in Figure 2.34.

The fluorescent signal of the labelled protein in the nucleus after 24 hours of incubation is significantly higher in the cells incubated with Toxin-ML3-NLS or Toxin-AS25-NLS than with Toxin-AS25. (Figure 2.34) This result demonstrates that part of the delivered proteins must have reached the cytosol. After 2.5 or 7 hours of incubation, no difference can be seen between the three different proteins, suggesting that the translocation to the

cytosol followed by nuclear import of the proteins occurs between 7 and 24 hours after the initial uptake in endosomes.

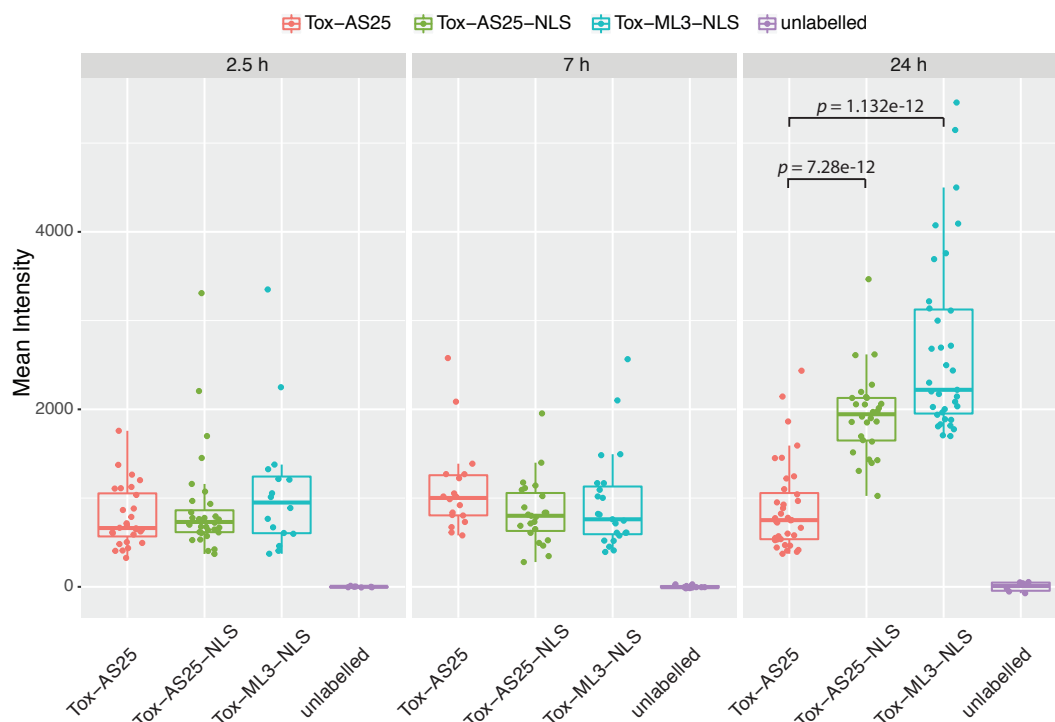


Figure 2.34 **Quantification of the fluorescence intensity in the nucleus after the delivery of NLS-tagged Toxin-monobody proteins.** HeLa cells were incubated with the indicated AF-488-labelled proteins for 2.5h, washed, and incubated in medium for the indicated durations, prior to nuclear staining with Hoechst and imaging on a confocal microscope. Each dot represents the mean 488 nm fluorescence in the nucleus of a single cell, as measured in 5 z-planes through the nucleus and normalized to the mean of the control cells incubated with unlabelled Tox-AS25-NLS. *P*-values were calculated using a Welch two-sample t-test.

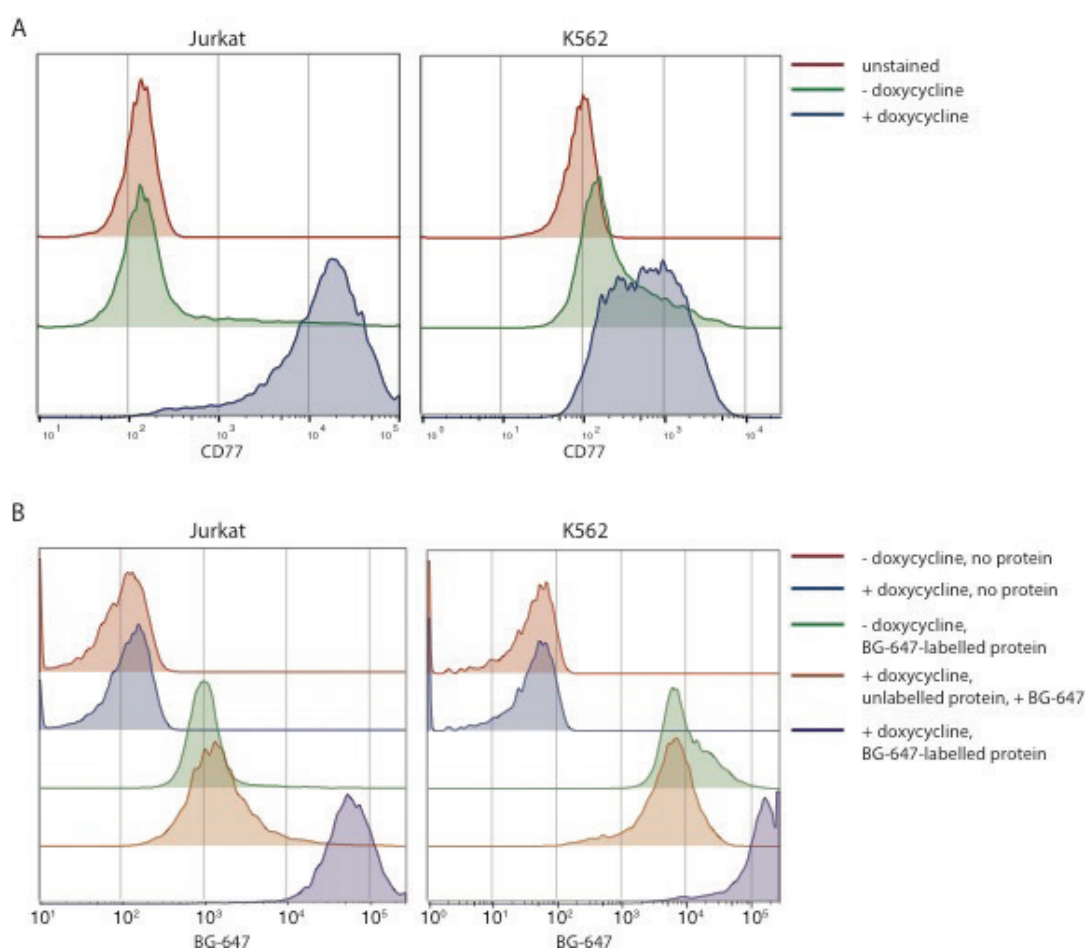
## 2.4.8 Gb3 expression in K562 and Jurkat cells

We next set out to assess if the delivered monobodies remain correctly folded and functional during the uptake, and whether the amount reaching the cytosol is sufficient to inhibit endogenous proteins and signaling pathways.

Since we wanted to analyze the effect of the delivery of AS25 into K562 cells and of ML3 into Jurkat cells, we set out to measure the expression of the Shiga-toxin receptor Gb3 in these cell lines. K562 and Jurkat cells were stained using an antibody against Gb3 (CD77) and analyzed by flow cytometry. The analysis revealed that both cell lines do not have Gb3 on their surface and thus we would not expect any uptake of the Toxin-Monobody fusion proteins in these cell lines.

The formation of Gb3 and its presence in the cell membrane is dependent on the expression of Lactosylceramide 4- $\alpha$ -galactosyltransferase (A4GALT) or Gb3 synthase, which catalyzes the transfer of galactose to lactosylceramide to form Gb3. We therefore reasoned that expressing this enzyme could lead to the presence of Gb3 on the surface and render the cells susceptible for delivery using our toxin construct. The A4GALT cDNA encoding for Gb3 synthase was cloned into a lentiviral vector containing a blasticidin resistance gene for selection as well as a Tet-on system for doxycycline

dependent expression of Gb3 synthase. The dependence on doxycycline for the expression would provide us with an internal control cell line, which in absence of doxycycline would carry the same transgene but not express Gb3 on the cell surface. Stable Jurkat and K562 cells were generated by lentiviral transduction. We determined a minimal lethal dose of blasticidin in K562 cells of 7.5 µg/ml, in the same way as described in section 2.1.5 for Jurkat cells. We then supplemented the growth media of the K562 cells with 7.5 µg/ml and that of the Jurkat cells with 5 µg/ml blasticidin to select for the transduced cells. After one week of selection, all the non-transduced cells in a control dish were dead, meaning that the selection has resulted in a pure population where 100% of the cells carry the transgene. We treated the surviving cells with doxycycline for 24 hours to test for the presence of Gb3 on the cell surface by immunostaining and flow cytometry. Both cell lines expressed high levels of Gb3 compared to HeLa cells when doxycycline was added to the growth medium but not when left untreated. (Figure 2.35 A) In K562 cells, a small amount of Gb3 positive cells were also observed in the untreated cells, showing some degree of leakiness of the plasmid in this cell line. We next assessed the ability of these cell lines to take up the monobodies fused to the toxin construct. We induced the expression of Gb3 synthase with doxycycline in both the transduced Jurkat and K562 cell lines and incubated the K562 cells with 1 µM of recombinant BG-AF647 labelled Toxin-SNAP-AS25 and the Jurkat cells with the respective concentration of labelled Toxin-SNAP-ML3 for 1 hour. After washing the cells, they were observed using flow cytometry to detect the BG-647 signal. 97% of both the doxycycline induced Jurkat and K562 cells which had been incubated with the labelled proteins were BG-647 positive, whereas the intensity of the signal from uninduced cells incubated with the same proteins was much lower. (Figure 2.35 B) To confirm that the proteins are indeed internalized into the cells and not only binding to the cell surface, doxycycline treated cells were incubated with unlabeled Toxin-SNAP-ML3 or Toxin-SNAP-AS25, washed with PBS to remove surface-bound proteins and then incubated with the cell impermeable SNAP substrate BG-AF647. A fluorescent signal of much lower intensity than in cells incubated with labelled proteins was observed. (Figure 2.35 B) These results suggest that the cells have internalized the proteins and that the signal does not come from surface-bound proteins.



**Figure 2.35 Gb3 expression in stable Jurkat and K562 cell lines and uptake of Toxin-monombody constructs.** **A:** Jurkat (left) and K562 (right) cells were transduced with a doxycycline inducible A4GALT gene, to express Gb3 on their surface. The cells were treated with doxycycline for 24h, stained with an anti-Gb3 (CD77) antibody and analyzed by flow cytometry. **B:** Transduced Jurkat cells (left panel) were treated with doxycycline for 24h (blue, orange and purple lines) or left untreated (red and green lines) and then incubated with BG-647 labelled Toxin-SNAP-ML3 (green and purple lines) or in medium alone (red and blue lines) or with unlabeled Toxin-SNAP-ML3, washed and subsequently incubated with BG-647 (orange line). K562 cells (right panel) were treated in the same way, but with Toxin-SNAP-AS25 instead of Toxin-SNAP-ML3. One representative plot is shown out of 2 biological repeats.

#### 2.4.9 AS25 inhibits the growth of K562 cells upon delivery

We next wanted to assess whether the amount of AS25 monobody delivered to the cytosol of K562 cells is sufficient to have an inhibitory effect on Bcr-Abl signaling. AS25 has been shown to induce apoptosis upon retroviral transduction in K562 cells, which are dependent on Bcr-Abl signaling for their growth and survival.<sup>78</sup>

We set out to measure cell survival of K562 cells after the delivery of Toxin-AS25 protein using the RealTime Glo MT cell viability assay, which is composed of the NanoLuc luciferase and a cell permeable substrate, which, when reduced in the cytosol, reacts with the enzyme to produce a luminescent signal. Since only metabolically active cells have a reducing potential, the signal intensity is directly proportional to the number of viable cells present.

Doxycycline induced K562 cells were seeded in a 96-well assay plate, and recombinant Toxin-AS25, or control proteins were added. Uninduced K562 cells which do not express Gb3 on the surface were used as a control. The cells were incubated for 4 days and luminescence signal was measured every 24 hours. Figure 2.36 A shows a representative time-course and Figure 2.36 B shows the luminescence signal at the 48-hour and 72-hour timepoint comparing all the proteins added. All the experiments were done in triplicates and 6 biological repeats were done of the same experiment, whereas not all the control proteins were included in each repeat.

Figure 2.36 A shows that, whereas the cell number increases with time when doxycycline induced cells were treated with AS25 alone or left untreated, but it decreases when the cells are treated with Toxin-AS25 or Toxin-SNAP-AS25. In uninduced cells not expressing Gb3, the addition of Toxin-AS25 or the SNAP-tagged version have no effect on cell growth, showing that the growth inhibitory effect of the Toxin-conjugated AS25 is dependent on uptake via the receptor Gb3. Nilotinib is a small molecule Bcr-Abl kinase inhibitor and is used as a positive control. Figure 2.36 B shows that in uninduced cells which were not treated with doxycycline, none of the added proteins has an effect on cell growth. However, the incubation of doxycycline induced K562 cells with Toxin-AS25 or Toxin-SNAP-AS25 leads to a reduction of viable cells as compared to uninduced cells. When 2  $\mu$ M Toxin-AS25 or Toxin-SNAP-AS25 are added, the number of viable cells is further decreased than with 1  $\mu$ M of the same proteins, indicating that the effect is concentration-dependent. AS25 alone has no effect on the growth of K562 cells, showing that the inhibitory effect of AS25 is dependent on its uptake. The delivery of an unrelated protein, such as emGFP has no effect either, which indicates that protein delivery on its own has no inhibitory effect on cell growth. However, when adding Toxin-ML3 or Toxin-SNAP-ML3, the cells are also growing less, even though the effect is less pronounced than with AS25. This could be explained by the binding of ML3 to the Lyn kinase, which is expressed in K562 cells and the inhibition of which could have an effect on cell growth, in line with reports in the literature.<sup>214-219</sup> The delivery of the non-binding control monobody Toxin-HA4\_YA also mildly reduces cell growth as compared to the control. This might be due to a weak residual binding of the HA4\_YA mutant monobody to Abl. On the other hand, the effect of the other monobodies on cell growth could also be due to a general toxicity of monobodies when delivered to K562 cells using the bacterial toxin construct.

Taken together, these results demonstrate that the uptake of AS25 in K562 cells has an inhibitory effect on cell growth and is dependent on the toxin construct and Gb3 on the surface of the target cells.

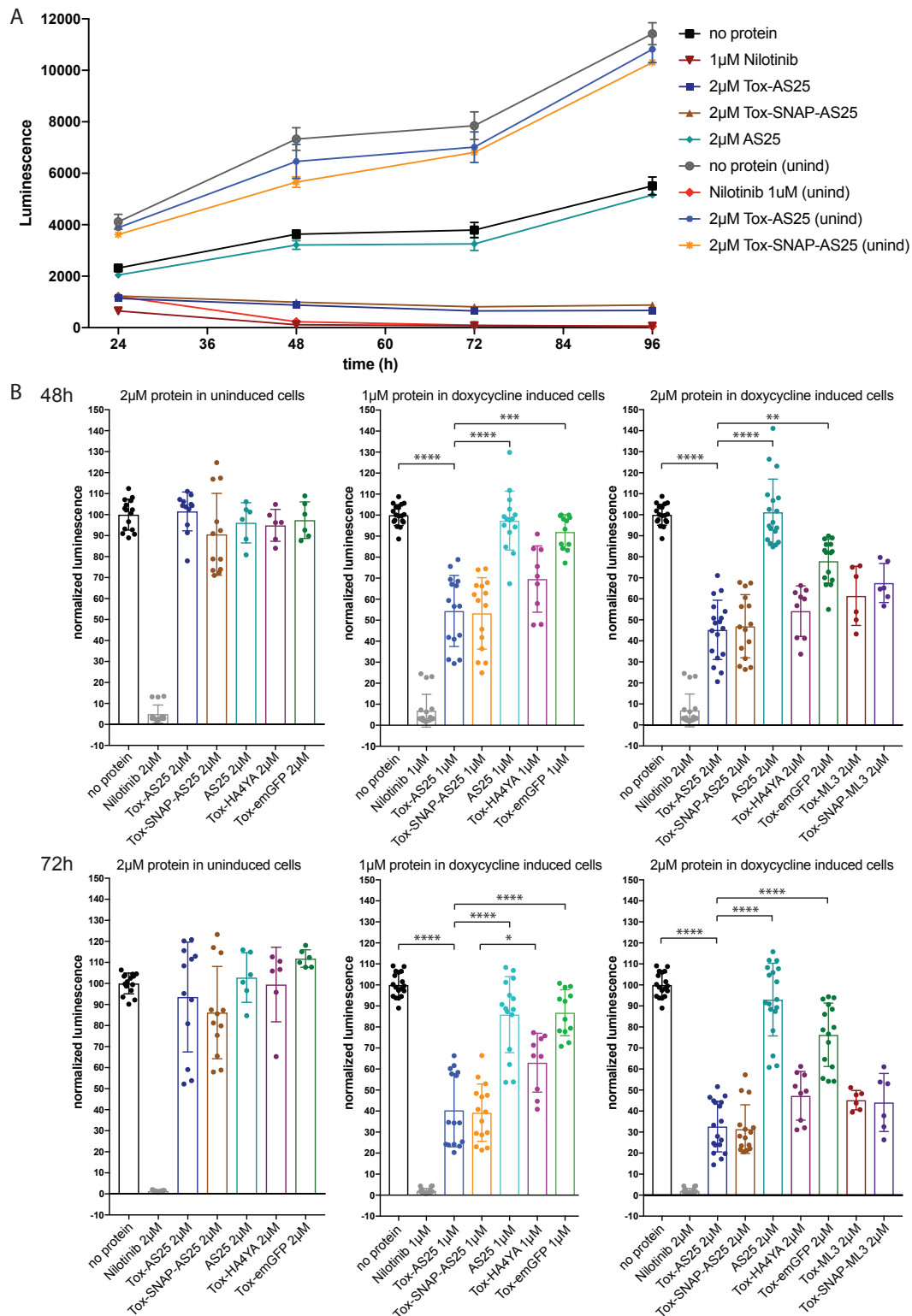


Figure 2.36 **Realtime Glo assay after AS25 delivery in K562 cells.** K562 cells were incubated with the indicated amounts of proteins and Relatime Glo reagents for 96h. **A:** Time course of one representative experiment. Error bars indicate the SD of 3 technical replicates. **B:** 48h (upper panels) and 72h (lower panels) timepoints comparing the different conditions with each dot representing one assay point. The experiment was done in technical triplicates with between 2 and 6 biological repeats, depending on the protein and concentration. All the technical and biological repeats are shown and *p*-values are calculated by averaging the technical replicates of each biological repeat, using a one-way ANOVA without correcting for

multiple comparisons. (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ ) Error bars indicate the SD of all the repeats.

#### 2.4.10 The delivery of AS25 into K562 cells leads to apoptosis

Small molecule kinase inhibitors targeting Bcr-Abl lead to apoptosis in CML cells by inhibiting Bcr-Abl mediated signaling pathways which are necessary for their growth and survival. We thus analyzed whether the effect on cell growth of the delivered AS25 is also due to apoptosis. We stained the cells with fluorophore coupled annexin V and the nuclear stain 7AAD. Annexin V binds to phosphatidylserine, which, in healthy cells is in the inner leaflet of the plasma membrane but flips to the outer leaflet when cells undergo apoptosis. The cells thus become positive for annexin V staining as soon as they enter apoptosis. 7AAD is a cell impermeable DNA binding dye, which can only enter cells that have a permeable cell membrane because they are apoptotic. As membrane permeabilization occurs only in late apoptosis, the double staining allows to distinguish between early (only annexin V positive) and late (double positive) apoptotic cells.

The cells were stained and analyzed 24, 48 or 72 hours after incubation with Toxin-AS25 or control proteins. Figure 2.37 shows the flow cytometry plots from one representative experiment with the gating strategy to separate early and late apoptotic cells. Figure 2.38 and Figure 2.39 show the fractions of cells in early and late apoptosis, respectively, from all the biological replicate experiments. About 3 to 5% of the cells are in early apoptosis and 1-2% in late apoptosis, when no protein or treatment is added, which is expected for a healthy cell population. When treated with nilotinib, about 30% of the cells are found in early apoptosis after 24 hours of treatment, and after 48 or 72 hours, this percentage increases to about 40%. About 10% of the nilotinib treated cells are in late apoptosis after 24 hours of treatment, 38% after 48 hours and more than 50% after 72 hours of treatment. In the Gb3 expressing cells incubated with Toxin-AS25 or Toxin-SNAP-AS25, 20 to 25% are in early apoptosis after 24 hours of incubation, and this fraction stays the same also after longer incubation. (Figure 2.38) 10% of these cells are in late apoptosis after 24 hours of treatment, similar to the nilotinib treated cells, and after 48 or 72 hours, this percentage increases to about 20% for the Toxin-AS25 treated cells. (Figure 2.39) Cells treated with AS25 alone, and uninduced cells treated with Toxin-AS25 or Toxin-SNAP-AS25 do not undergo apoptosis, demonstrating that the effect is dependent on the expression of Gb3 on the surface as well as the presence of the toxin construct, leading to the uptake of AS25. The delivery of Toxin-ML3 and Toxin-SNAP-ML3 also leads to apoptosis in Gb3 expressing K562 cells, indicating that ML3 might indeed inhibit Lyn in these cells lines and thereby induce apoptosis. When delivering emGFP, no increase in apoptosis is observed as compared to untreated cells and with the Toxin-conjugated non-binding monobody HA4\_YA, also only a small fraction of cells undergoes apoptosis, which is marginally higher than in untreated cells after 48 and 72 hours of incubation.

These results collectively show that AS25 specifically induces apoptosis in K562 cells after toxin mediated delivery, as in absence of Gb3 or the toxin construct, no increased apoptosis is observed.

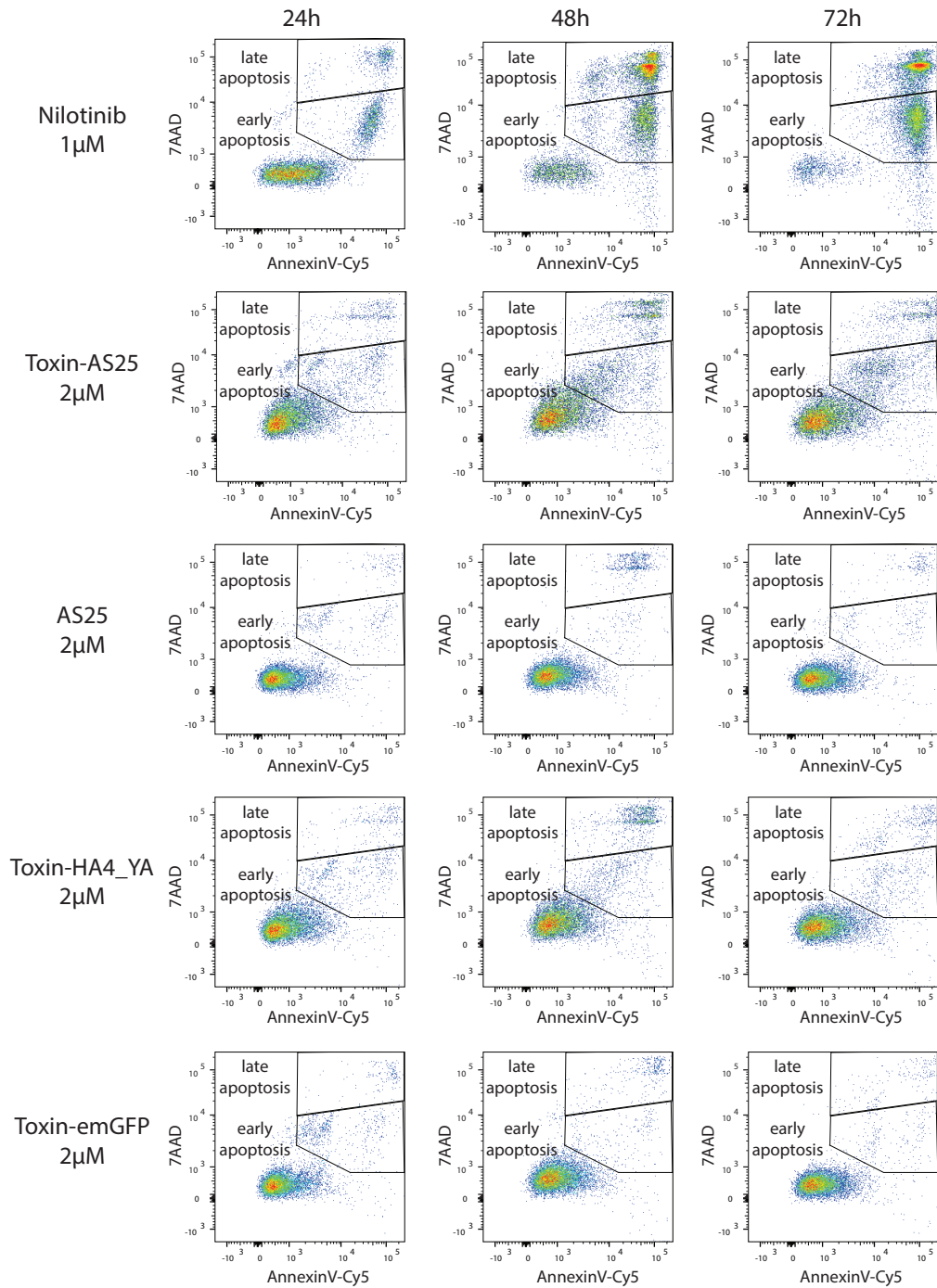
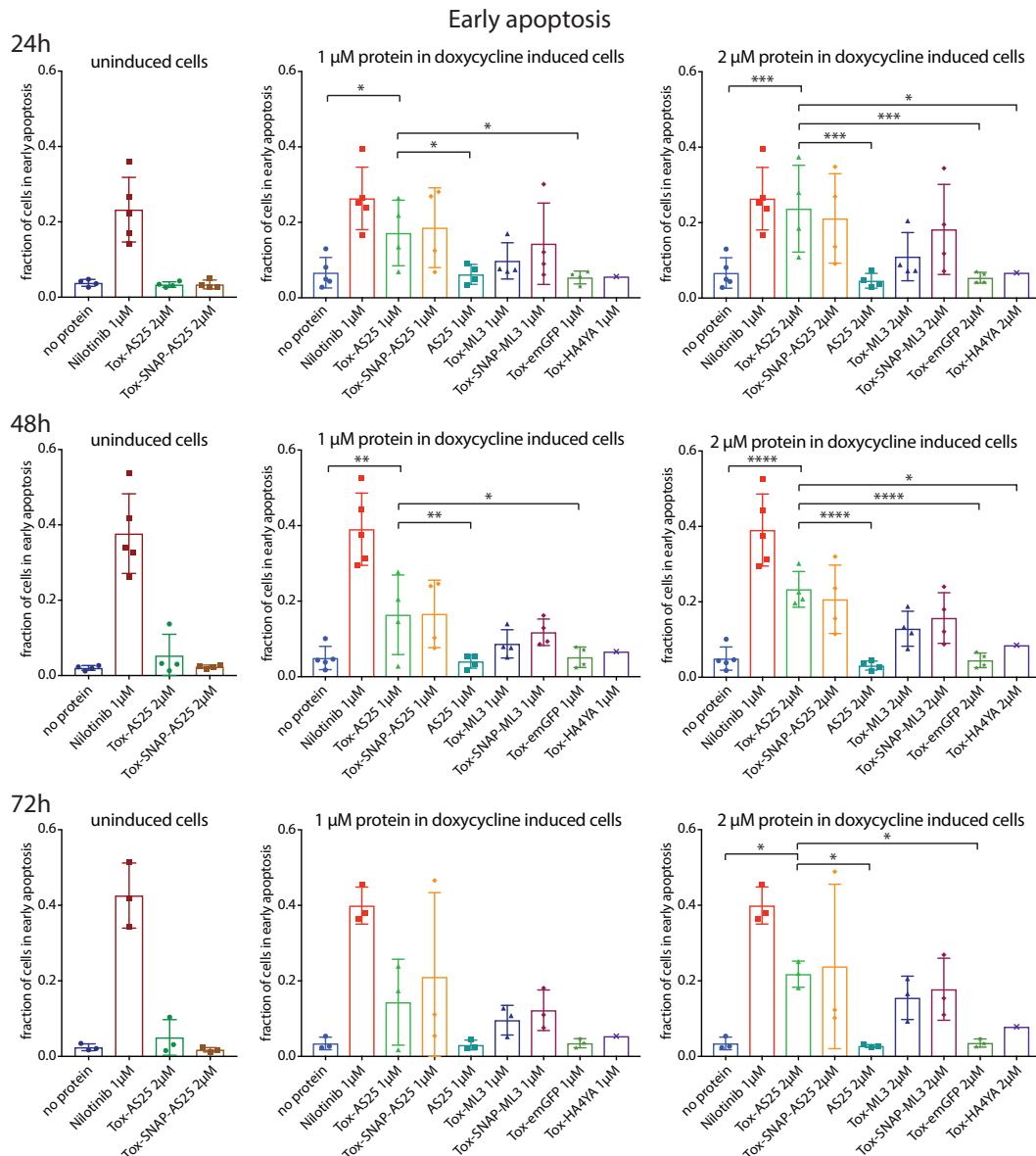


Figure 2.37 **Flow cytometry plots of annexin V-7AAD-staining after protein delivery.** K562 cells were incubated with the indicated proteins for 24, 48 or 72 hours, washed and stained with a Cy5-coupled antibody against annexin V and with 7AAD, which are plotted on the x and y axis, respectively. The gates indicate the early and late apoptotic populations. Representative plots are shown out of 1 (Toxin-HA4\_YA) to 4 (other samples) biological repeats.



**Figure 2.38 Fractions of cells in early apoptosis according to the annexin V-7AAD stain after protein delivery.** K562 cells were incubated with the indicated proteins, washed and stained with a Cy5-coupled antibody against annexin V and with 7AAD and analyzed by flow cytometry. The percentages of early apoptotic cells correspond to the gates shown in Figure 2.37 figure 2.24. Each dot corresponds to a biological repeat of the experiment. *P*-values are calculated using a one-way ANOVA without correcting for multiple comparisons. (\**p*<0.05, \*\**p*<0.01, \*\*\**p*<0.001, \*\*\*\**p*<0.0001) Error bars indicate the SD of all the repeats.

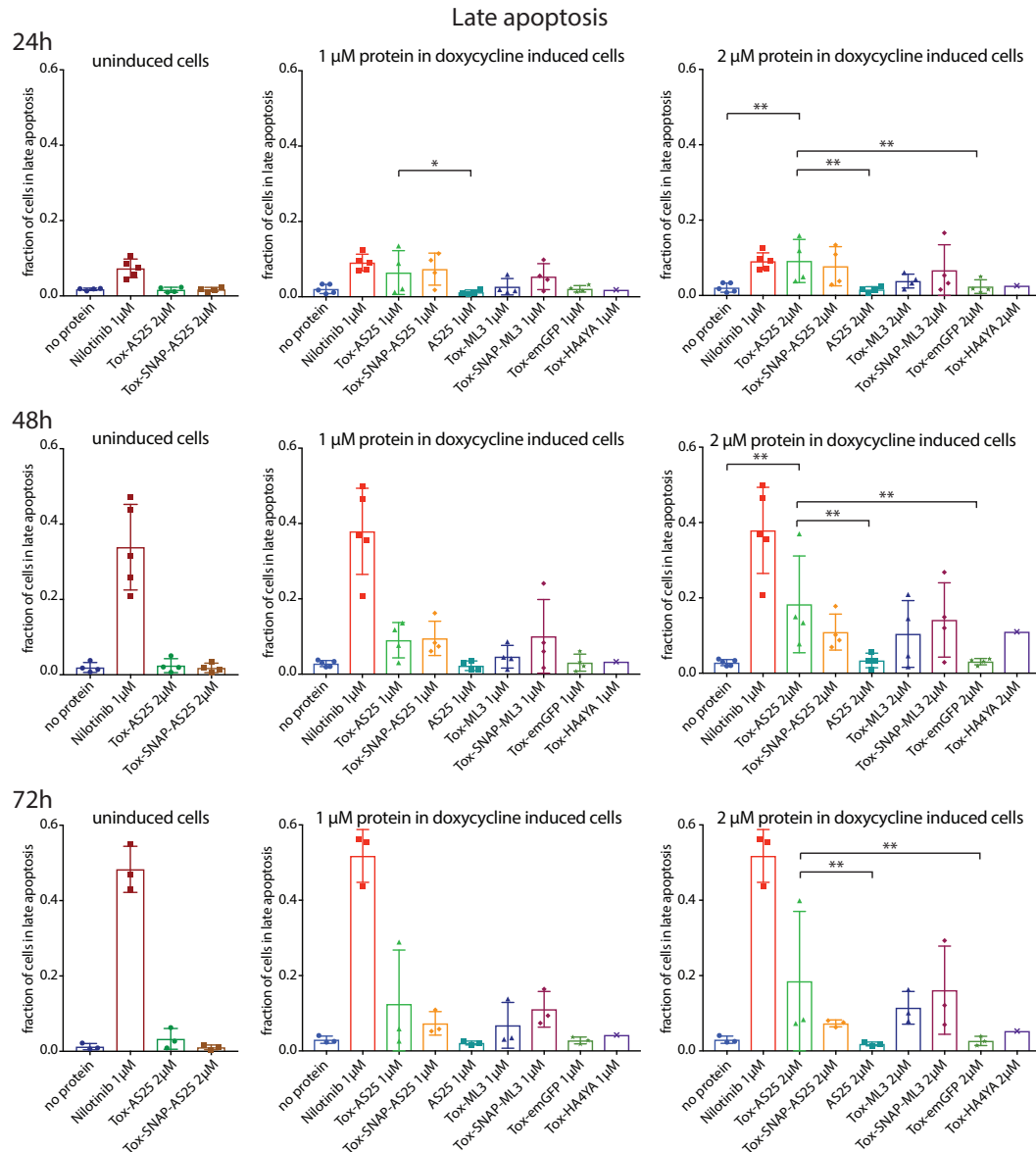


Figure 2.39 **Fractions of cells in late apoptosis according to the annexin V-7AAD stain after protein delivery.** K562 cells were incubated with the indicated proteins, washed and stained with a Cy5-coupled antibody against annexin V and with 7AAD and analyzed by flow cytometry. The percentages of late apoptotic cells correspond to the gates shown in Figure 2.37 figure 2.24. Each dot corresponds to a biological repeat of the experiment. *P*-values are calculated using a one-way ANOVA without correcting for multiple comparisons. (\* $p < 0.05$ , \*\* $p < 0.01$ ) Error bars indicate the SD of all the repeats.

#### 2.4.11 The delivery of a VHL-ML3 fusion protein in Jurkat cells targets Lck for degradation

I chose the VHL-ML3 fusion protein studied in section 2.1.5 to study protein delivery using the bacterial toxin construct in Jurkat cells. As the degradation of Lck is easy to assess by immunoblotting, without the need of T cell receptor stimulation, it is a suitable readout to investigate whether functional VHL fusion proteins can be delivered using the bacterial toxin construct.

We first cloned the VHL-ML3 and VHL-HA4\_YA constructs shown in Figure 2.22 into the expression vector containing the bacterial toxin and expressed and purified the resulting fusion proteins in *E. coli* Origami cells as described previously for the other proteins. The expression of these proteins containing the VHL gave much lower yields and purity after the Ni-NTA column purification than the previously purified proteins. We therefore performed a series of small-scale expression trials to test different expression media as well as lysis buffers in order to find conditions better suited for these proteins. According to the results of this screen we decided to use TB medium for expression and HEPES buffer supplemented with Triton-X-100 as a lysis buffer. With these conditions, we still obtained a low, but acceptable yield of about 2 mg from 1 liter of bacteria culture.

Figure 2.40 A and B show size exclusion chromatograms of Toxin-VHL-ML3 and Toxin-VHL-HA4\_YA, respectively. The shoulder eluting just after the main peak was collected after the size exclusion, corresponding to the correct size of the pentamer according to the column calibration. Figure 2.40 C shows a Coomassie-stained gel of the different fractions from the Ni-NTA column and SEC purification of Toxin-VHL-ML3 and Toxin-SNAP-VHL-ML3. The SNAP-tagged version was purified at the same time because we reasoned that the SNAP-tag might stabilize the protein and lead to a better yield, which was not the case.

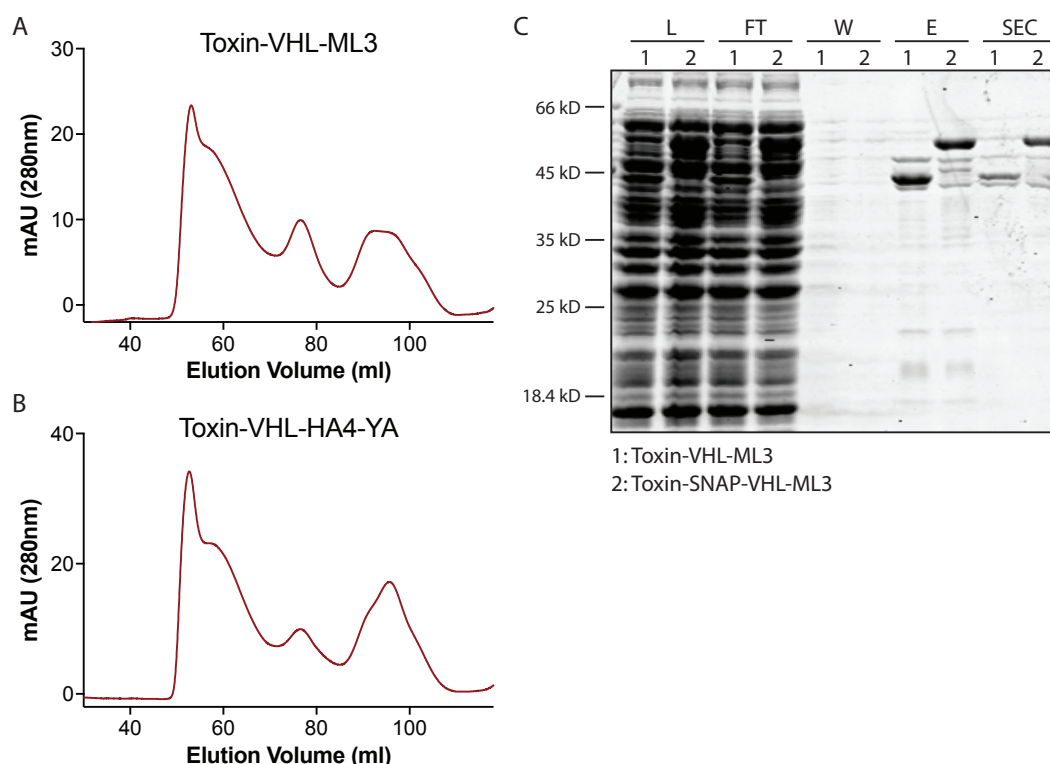
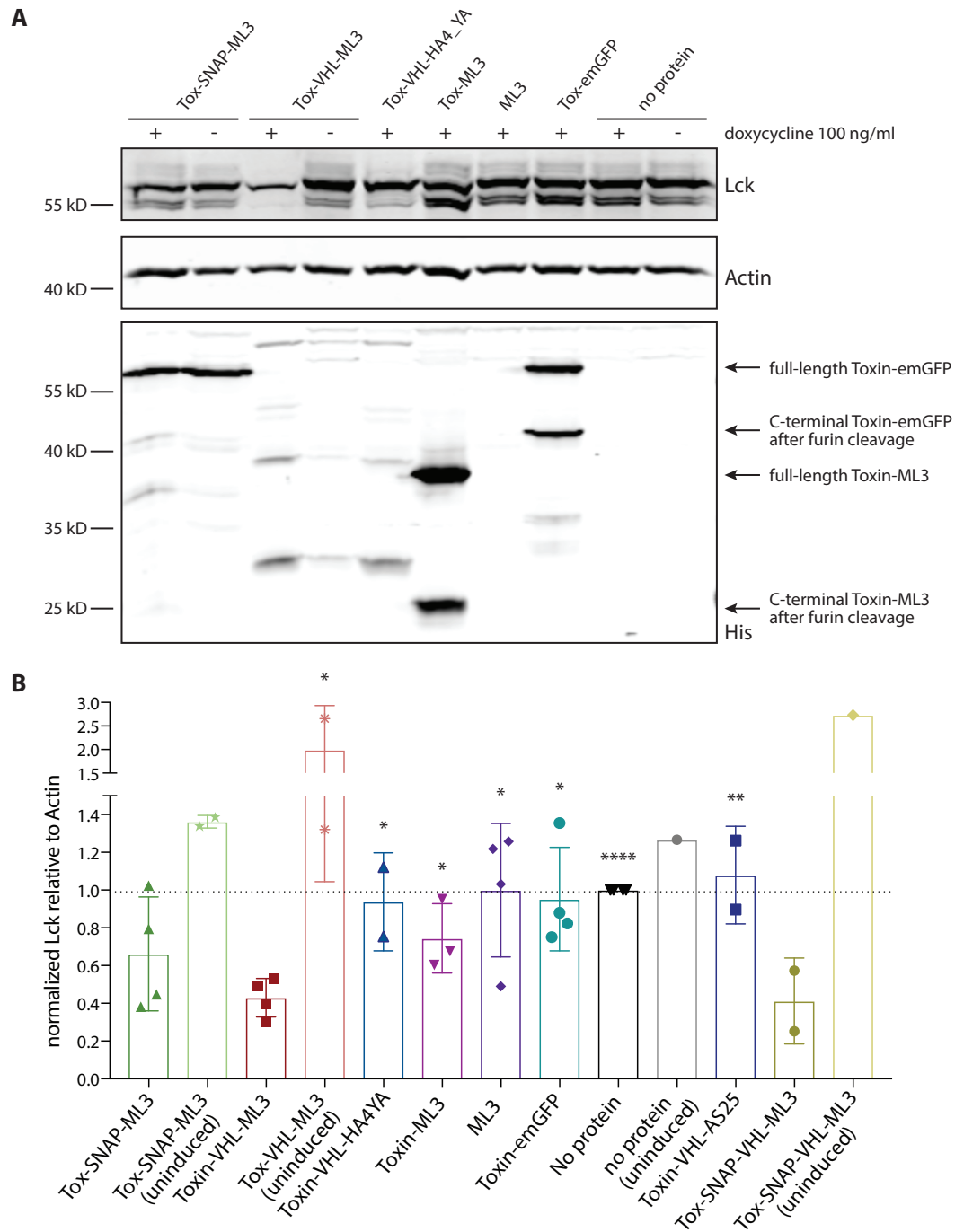


Figure 2.40 **Stx2B-TDP-VHL-Monobody fusion protein purification.** The crude bacteria lysate was first purified on a Ni-NTA column and then by SEC. **A and B:** Size exclusion chromatograms, the shoulder of the first peak corresponds to the right size of the pentamer and was purified. **C:** Coomassie stained SDS PAGE gel with the fractions from the Ni-NTA purification (L=crude lysate, FT=flow-through, W=wash, E=elution) and the shoulder of the main peak of the SEC after concentration.

We next incubated the purified proteins with Jurkat cells for 48 hours before washing and lysing the cells to study their effect on Lck protein levels. We observed a decrease of the Lck protein level in the Toxin-VHL-ML3-treated cells of more than 50%, as compared to cells where no protein was added. (Figure 2.41 A and B) When the cells were incubated with Toxin-emGFP, Toxin-VHL-HA4\_YA or Toxin-VHL-AS25, we did not observe such a drop in Lck protein levels. The incubation with Toxin-ML3, without the VHL protein resulted in a very mild reduction of Lck levels in 2 out of 3 repeats. In the cells where the expression of Gb3 had not been induced with doxycycline, we did not observe any decrease of Lck protein levels after adding Toxin-VHL-ML3. On the contrary, Lck levels even increased slightly. On the anti-His immunoblots, we observed that the bands corresponding to the delivered VHL-fusion constructs were much fainter than of the proteins not bearing a VHL sequence. (Figure 2.41, lower blot) This indicates that the VHL-tagged proteins are indeed recruited to the CUL2-E3 ubiquitin ligase complex and degraded. Moreover, we observed an additional smaller band in the lanes corresponding to Toxin-ML3 and Toxin-emGFP, which correspond to the size of the C-terminal, His-tagged parts of the proteins after cleavage by the protease furin. (Figure 2.41 A, lower blot) This indicates that the protein is indeed internalized and cleaved by furin along the endocytic or retrograde pathway. In the lane corresponding to ML3, no band can be seen, because the protein without the toxin-construct is not taken up into the cells. In the lane corresponding to uninduced cells treated with Toxin-VHL-ML3, no band can be seen either, showing that Gb3 needs to be present on the cell surface for protein uptake. When uninduced cells were incubated with Toxin-SNAP-ML3, a band of the same strength than with doxycycline induced cells can be seen, which is surprising, as the cells should not be able to take up this protein without the expression of Gb3 either. (Figure 2.41 A, lower blot) The same observation was made in the other repeats of the experiment, and only Toxin-SNAP-ML3 but no other proteins resulted in a band in the anti-His immunoblot when incubated with uninduced cells, indicating that this protein might bind to the cell membrane through a protein-protein interaction which is independent of Gb3.

Collectively, these results demonstrate that the delivery of VHL-ML3 into Jurkat cells is dependent on the binding of the toxin to Gb3 on the cell surface. Furthermore, the VHL protein is necessary to induce Lck degradation, indicating that the VHL-ML3 protein recruits Lck to the CUL2-E3 ubiquitin ligase complex after its delivery to the cytosol in a correctly folded state.



**Figure 2.41 Delivery of Toxin-VHL-ML3 in Jurkat cells.** Jurkat cells were incubated for 48 hours with the indicated proteins, washed and lysed. The cell lysate was immunoblotted with antibodies against Lck, Actin and the His-tag. **A:** One representative set of western blots from the same experiment. **B:** Quantification of the Lck immunoblot normalized to Actin and to the control where no protein was added. Each dot represents an individual biological repeat of the experiment. *P*-values are calculated using a two-tailed t-test, comparing the values for Toxin-VHL-ML3 with the values where \* are shown. (\**p*<0.05, \*\**p*<0.01, \*\*\**p*<0.001, \*\*\*\**p*<0.0001) Error bars indicate the SD of all the repeats.



## 3 Discussion

### 3.1 Monobodies targeting SFKs

The aim of this project was to develop monobodies that are able to discriminate between different SFKs and to inhibit their function in cells. Since off-target binding is a major challenge when developing selective protein binders, I first set out to characterize the selectivity of the SFK SH2 domain binding monobodies in cells. The TAP coupled to LC-MS/MS analysis is a rigorous way of determining monobody selectivity, as it assays the selectivity of monobodies on a proteome-wide scale instead of concentrating on a particular protein family. The TAP-MS protocol has been developed with the goal to keep protein complexes intact during the pull-down and to obtain a comprehensive list of proteins which might directly or indirectly interact with the monobodies.<sup>211</sup>

We observed large differences in the total number of identified proteins and in their total spectrum counts between the first and the second repeat of the analysis of Lck binding monobodies. (Table 2.1 and Table 2.2) This is most probably due to differences in TAP efficiency, sample handling and possibly the analysis itself carried out by the proteomics core facility. Still, the ranking and list of the most abundant monobody binders was very similar between both repeats, which demonstrates the accuracy and reproducibility of the method. One major drawback of this type of mass spectrometry analysis is the large number of contaminant proteins that are typically detected, despite the two consecutive affinity purification steps, which makes it difficult to distinguish them from true interactors. The online contaminant repository for affinity purifications (CRAPome) database has been created in an attempt to list these common contaminants from affinity-purification experiments and to facilitate their identification. (<http://crapome.org>) Although very useful when doing analyses with similar purification tags than the ones listed in the database, a comparison with data obtained by the same protocol is in our case certainly more reliable. I therefore compared my results with all the protein lists already generated by our lab using exactly the same protocol. This approach allowed me to calculate a selectivity score for each protein, and to distinguish between true interactors and contaminants.

When analyzing the resulting protein list, one has to consider, that the amount of a given protein seen by mass spectrometry analysis is determined by a combination of its cellular abundance and its affinity for the bait. We therefore have to take into account the cell-type specific expression of potential monobody off-targets, especially for the different SFKs. Even though the SrcA family members Yes, Src and Fyn are constitutively expressed across cell types, Lck might for example be much more abundant in T cells, and therefore lead to higher spectral counts, independently of monobody binding affinity. However, the fact that Yes, Src and Fyn were almost completely lacking in the MS analysis strongly supports the findings of the yeast binding assay (Figure 2.1), ITC (Figure 2.2) and pY peptide competition assays (Figure 2.4), all demonstrating the subgroup specificity of the Lck binding monobodies.

Most of the selected monobodies bind with nanomolar affinity to their respective targets, and especially ML3 is one of the strongest binders and most potent pY peptide displacing

monobodies. (Figure 2.4) However, in a functional assay, it only showed a relatively mild effect at inhibiting Zap70 phosphorylation upon T cell receptor activation. Despite the decreased Zap70 phosphorylation, there was no effect on downstream signaling processes, such as calcium flux and NFAT dephosphorylation. (Figure 2.6 and Figure 2.7) This lack of signal propagation could be due to redundancies in the signaling pathway that are not entirely dependent on Lck or may still be functional despite blockade of the SH2 domain by a monobody. Fyn kinase has been reported to contribute to proximal TCR signaling, albeit it is less important than Lck.<sup>220</sup> Some reports indicate that Fyn can directly bind to the T cell receptor<sup>221,222</sup> and that it can partially compensate for the loss of Lck, especially during T cell development.<sup>223</sup> Experiments with the Lck-deficient Jurkat cell line JCaM1 suggest that Fyn alone is sufficient to activate the MAPK-ERK signaling pathway, via a PLC $\gamma$ 1 independent pathway by recruiting Grb2-SOS complexes directly to phosphorylated  $\zeta$ -chains in the TCR.<sup>224,225</sup> Since ML1 and ML3 are subgroup specific and only bind Lck but not Fyn, this might be the reason for the relatively mild reduction in Zap70 phosphorylation. Moreover, a process known as kinome reprogramming, which consists in the activation of alternative signaling pathways in response to kinase inhibition has been frequently observed in different types of cancer<sup>2,226</sup> and could also occur in response to Lck inhibition. Unlike classical kinase inhibitors, ML1 and ML3 do not inhibit the kinase domain of Lck, but the SH2 domain. Albeit inhibiting certain protein-protein-interactions of Lck with its substrates, the activity of the kinase domain is not altered by the monobodies, and Lck might still be able to bind substrates via its SH3 domain and/or phosphorylate substrates independently of the SH2 domain. An alternative explanation for the reduced Zap70 phosphorylation in ML1 or ML3 expressing cells could be that the monobodies are binding directly to Zap70 as an off-target. However, we detected only very low spectral counts for Zap70 in the second repeat of the TAP experiment in Jurkat cells expressing ML1 and ML3, with total spectrum counts (TSC) of less than 1% of the TSC for Lck. Moreover, the addition of ML3 to lysates from activated T cells blocked the binding of Zap70 to recombinant Lck SH2 domain, (Figure 2.5) providing further evidence that ML1 and ML3 inhibit the binding of Zap70 to Lck and do not bind Zap70 directly.

A puzzling observation was that, despite the very similar binding mode to the Lck SH2 domain, ML1 is not able to compete with a pYEEI peptide while ML3 is, whereas both monobodies competed with two other pY peptides having a lower affinity for the Lck SH2 domain. (Figure 2.4) In line with this, the inhibition of Zap-70 phosphorylation by ML1 was less pronounced than by ML3. The crystal structures of ML1 or ML3 in complex with the Lck SH2 domain, that were obtained during the study,<sup>210</sup> show that both monobodies bind to the +3-specificity pocket, which is most likely responsible for their capacity to distinguish between the SrcA and SrcB subgroups. In contrast, the pY binding pocket of the SH2 domain is less occluded by the CD loop of ML1 than ML3, which could be the reason why ML3 shows more potent binding inhibition in several assays.

We repeatedly observed lower protein levels of ML3 than ML1 in Jurkat cells transduced with the respective plasmids, even when using a different expression vector. (Figure 2.6) However, we observed the same mRNA levels in both cell lines suggesting that protein levels are regulated post translationally, and that the cells might be able to adapt to the

inhibitory effects of the monobodies by reducing their protein levels, or that ML3 is a less stable protein as ML1 in cells and therefore has a shorter half-life.

The degradation of Lck by ML1 or ML3 fused to VHL resulted in a more pronounced reduction in Zap70 phosphorylation upon TCR activation, as compared to the expression of ML1 or ML3 alone. (Figure 2.9) This observation is consistent with reports from the group of Prof. Craig Crews, that lower IC<sub>50</sub> values can be reached when conjugating kinase inhibiting drugs to a small molecule VHL-ligand (PROTAC concept) than with the drug on its own.<sup>227,228</sup> They rationalize this finding with the fact that degradation of the target protein inhibits downstream signaling in a more sustained way than the binding of an inhibitor which needs to continuously bind at saturating concentrations.<sup>205</sup>

Experiments measuring Ca<sup>2+</sup>-flux and phosphorylation of downstream proteins in these cells would clarify whether this more pronounced reduction in Zap70 phosphorylation could have consequences on downstream signaling in contrast to cells expressing ML1 or ML3 alone. Another approach would be to use a reporter cell line for one of the downstream transcription factors, for example NFAT.

Conclusively, we showed that the monobodies ML1 and ML3 inhibit proximal T-cell signaling and that this effect can be increased by coupling them to the E3-ubiquitin ligase VHL, inducing Lck degradation. They could be further used to study T cell receptor signaling or other signaling pathways involving Lck, Lyn, Blk or Hck, and to inhibit deregulated signaling in cancer cells or autoimmune disease.

### 3.2 Monobody delivery using CPPs

Since CPPs have been the first method of delivering proteins into cells, and many of them are of natural origin, we first chose to test their applicability to monobodies. We chose Tat because it is a naturally occurring peptide and the first discovered CPP,<sup>87</sup> and CPP44 because it has been proposed to be homing specifically to AML and CML cells.<sup>100</sup> Polyarginine has been proposed to be the simplest sequence still able to deliver proteins, which is why we included it into our study. We ordered all peptides with and without a red fluorescent TAMRA, because this fluorophore is pH-independent and thus still visible in the acidic compartment of lysosomes.

I established a protocol to activate these cysteine-modified CPPs using DTNB and thereby making them reactive towards free thiols, allowing to covalently ligate them to cysteine-containing proteins by a disulfide bond. However, even though the ligation of the activated CPPs to monobodies appeared to be effective, as demonstrated by the colorimetric reaction as well as analytical HPLC, the resulting protein was almost entirely lost through precipitation. Different attempts to dilute the protein in guanidinium-containing buffers and to then refold it have all failed. This apparent tendency of CPP-coupled monobodies to aggregate or precipitate is in line with our previous difficulties to express and purify them as fusion proteins in *E. coli*. Highly positively charged CPPs might form electrostatic interactions with negatively charged amino acids in the monobodies that could have a destabilizing effect. Therefore, an unfolded linker between the CPP and the monobody might help improving the solubility of the fusion proteins and could be assessed in a future attempt. It would be very interesting to test the cellular effects of monobody delivery using CPPs, if a stable construct can be generated.

On the other hand, CPPs have been shown to lead to endosomal entrapment of the majority of the protein cargo by an increasing number of groups.<sup>94,229,230</sup> In order to overcome this limitation, different strategies to enhance the endosomal escape have been proposed.<sup>109,231</sup> The combination of endosomal-escape enhancing mechanisms with CPPs might be useful also for the delivery of monobodies and would be interesting to assess. However, by comparing CPPs to the other two protein delivery approaches that I have worked on during my PhD, I think that CPPs alone are less promising for the delivery of functional proteins to the cytosol, even though they might be used in combination with other methods.

### 3.3 Monobody delivery using CPDs

As CPDs have been proposed as cellular delivery agents for certain macromolecules, which can circumvent the endosomal escape problem frequently encountered with CPPs and other methods, I chose to test if I could use them to deliver functional monobodies into cells. The first attempts of coupling biotinylated monobodies to CPDs via streptavidin were hampered by extensive precipitation of the adducts. However, rigorous and systematic testing of different buffer conditions led to an optimized protocol, which allowed me to avoid major precipitation issues. When performing the first tests in HeLa cells, I could observe a fluorescent signal inside the cells, demonstrating that the adducts were indeed taken up. (Figure 2.16 and Figure 2.17) However, I also observed signs of cytotoxicity, such as membrane blebbing, irregular nucleus shapes and cell debris in a large fraction of cells. When I counterstained the cells with Sytox Green, a DNA-intercalating dye which does not penetrate healthy cells, the nuclei became green fluorescent, demonstrating that the cells were permeabilized. Even when I washed the cells after having incubated them with adducts for one hour, and then incubated them in growth medium for another 24 hours, the cells remained permeabilized, as indicated by the Sytox Green staining. (Figure 2.18 A) I additionally performed a Cell Titer Glo assay, which also showed that the adducts at a concentration of 500 nM are toxic to HeLa cells. (Figure 2.18 B) Since I also did the same tests with adducts which did not contain any monobodies, the cytotoxic effects must stem from the CPDs, and not from the monobodies.

These results clearly demonstrate that CPDs are toxic to HeLa cells when added at concentrations of 500 nM. Cells incubated with 250 nM adducts were also positive for the Sytox Green staining but did not show cytotoxicity in the CellTiter Glo assay. These results are contradicting previous studies from the group of Prof. Stefan Matile, claiming no cytotoxicity of the CPDs for concentrations up to 2.5  $\mu$ M and negligible cytotoxic effects at 10  $\mu$ M, as measured by an MTT assay.<sup>113</sup> The reasons for these observed differences are unclear to us.

I nevertheless decided to continue the study with a lower concentration of 100 nM of the adducts as no cytotoxicity was observed with this concentration. Since endosomal entrapment is a common challenge for protein delivery, I set out to assess whether CPDs could overcome this limitation and deliver monobodies to the cytosol of HeLa cells.

A fluorescent signal from the labelled monobodies was observed in the cells after incubation, showing that even at lower concentrations, CPDs were effective at promoting the cellular uptake of the adducts. (Figure 2.19) However, most of the fluorescent signal originated from distinct speckles within the cell and no dispersed signal was observed,

also in contrast to delivery experiments with different cargoes, which have been published by the group of Prof. Matile.<sup>113,116</sup> When observed in live cells, these speckles were in part moving in the cell, indicating that they could be intracellular vesicles.

I hypothesized that these speckles could be endosomes and therefore performed immunofluorescence experiments with antibodies against the early endosomal marker EEA1 as well as the lysosomal marker Lamp1. (Figure 2.20 and Figure 2.21) First of all, I observed a much less intense fluorescent signal from the labelled monobodies in these fixed cells as when observing live cells. This could be due to a quenching of the signal by the fixation and permeabilization method, or to a less efficient monobody delivery in this set of experiments, as compared with the live cell experiments. Since a new batch of CPDs was used, there could also be a variation in the coupling or delivery efficiency between different batches of CPDs. Because the intensity of the monobody signal was very low, especially at early time points of the delivery, I chose to analyze the colocalization between the signal from the CPDs with the signal from the antibodies. I observed an increasing colocalization between early endosomes and the CPDs for up to 30 minutes, after which it decreased again. (Figure 2.20) This observation suggests that CPDs coupled to proteins are at least in part taken up via endocytosis. I also observed an increasing colocalization of the CPDs with lysosomes, starting as early as 20 minutes after addition of the adducts. (Figure 2.21)

These results collectively show that CPD-monobody adducts are at least partially taken up by endocytosis, which fuse with lysosomes along the endocytic route, leading to a degradation of the adducts. I did not address whether the adducts are degraded over time, and further experiments including longer time points as well as a quantitative analysis would be needed to answer this question. The fact that the fluorescent signal from the CPDs was lower after incubation with adducts for one hour followed by medium for 24 hours, than after incubation with adducts for 24 hours indicates that the adducts might be degraded when they are removed from the medium and are not taken up any longer. Further experiments would be needed to test this hypothesis.

These results do not agree with the model proposed by the group of Prof. Stefan Matile, by which CPDs are delivered into cells through a thiol-mediated uptake mechanism, and not through endocytosis.<sup>113,114</sup> They observe very strong fluorescent signals from cells incubated with fluorescent CPDs, including fluorescence in the nucleoli already after a few minutes of incubation.<sup>113</sup> However, the thiol-mediated uptake model is a theoretical hypothesis and remains to be confirmed experimentally. On the other hand, CPDs could behave differently and follow a different cellular uptake route when coupled to proteins than when coupled to fluorophores or other small molecules, which were used in the studies of the Matile group. It has been shown that CPPs can be taken up into cells by different mechanisms depending on the cargo that they are coupled to, as well as the cell type.<sup>93,94</sup> Therefore, and because streptavidin together with the monobody is approximately 65 kD in size, it is very likely that the uptake is not following the same route as when a smaller cargo is attached to the CPDs. CPDs have also been shown to at least partially colocalize with the live cell marker for lysosomes, LysoTracker, by Fu et al., who also studied protein delivery using CPDs.<sup>115</sup> They have shown PARP1 cleavage upon the delivery of CPD-conjugated Caspase 3 as well as a decrease in cell viability upon delivery of the small molecule doxycycline using CPDs. However, they incubate cells with adducts at concentrations below 100 nM, which is probably sufficient since low amounts of caspase 3 as well as of doxycycline already elicit a biological response.

In a later publication, the Matile group reported the delivery of streptavidin reaching the nucleoli after delivery.<sup>116</sup> They used the same protocol for the formation of streptavidin-CPD adducts and HeLa cells were incubated for 1 hour with 1  $\mu$ M of the resulting adducts of a size of approximately 77kD. The cells show a strong fluorescent signal from both the CPDs as well as a fluorophore coupled to streptavidin, mostly in large speckles in the nuclei. However, even though the cells are rounded up and show signs of cytotoxicity, no toxicity test was done, and no explanation is given, on the mechanism, by which these large adducts can rapidly enter cells and even accumulate in nucleoli.

In order to reduce the cargo size, a different way of coupling monobodies to CPDs without the need of streptavidin could be envisaged, such as native chemical ligation, which is commonly used to ligate peptides. On the other hand, a combination of CPDs with peptides promoting endosomal escape could be used to avoid the endosomal entrapment observed after the delivery of monobody-CPD-adducts. Several endosomal escape promoting peptides have been described and are listed in the introduction.

In conclusion, I demonstrated that low, subtoxic concentrations of CPDs can be used to deliver monobodies into cells. However, this method might benefit from a combination with peptides promoting endosomal escape, or other methods to avoid endosomal entrapment.

### 3.4 Monobody delivery using bacterial toxins

To enable the development of monobodies and potentially other protein binders as protein-based therapeutics, we were able to validate the use of bacterial toxins for the cellular delivery of monobodies. Different bacterial toxins have been proposed and tested for protein delivery, and differences in their efficacy, which might be dependent on both the cargo and the cell type have been identified. The anthrax toxin pore has for example been shown to be effective at delivering proteins which are able to unfold in order to translocate through the pore into the cytosol, whereas more stable proteins will remain trapped in endosomes.<sup>140</sup> Endosomal entrapment is a common drawback for many protein delivery approaches, including for bacterial toxin-mediated delivery as it has been observed with Diphtheria toxin and in some cases Anthrax toxin.<sup>168</sup> Toxins which employ the host cell's retrograde trafficking route to reach the cytosol could overcome this limitation. We therefore chose to study the delivery of monobodies by a fusion construct between Shiga-like toxin and *Pseudomonas* Exotoxin A, which has been introduced by the group of Prof. Hak-Sung Kim.<sup>199</sup>

For monitoring the uptake of the proteins into cells, we used two different labelling strategies. In the first approach, a maleimide coupled dye was covalently attached to free thiol groups on cysteine residues. Since the fusion protein is purified and kept in a non-reducing buffer, we assumed that the four cysteines present in the toxin part would remain in their disulfide state, which is essential for their native structure, and that the dye would only be coupled to the C-terminal cysteine. The fluorescent signal observed within the cells after incubation with labelled Toxin-monobody fusion proteins demonstrates that the toxin part remains correctly folded and functional after labelling. (Figure 2.24) The second labeling strategy, using the SNAP-tag has several advantages over labelling with cysteine-reactive dyes. First, it allows the site-specific labelling of a specific tag which is far away from the toxin subunits essential for the uptake. Second,

since cell-permeable as well as -impermeable SNAP substrates are commercially available, it enabled us to distinguish between surface-bound and internalized proteins, providing an additional control for the cellular uptake of the delivered proteins. Third, the SNAP-tag helped stabilizing the recombinant toxin-monobody fusion proteins, leading to higher protein purification yields, as well as less precipitation upon labelling as compared to labelling with the maleimide-coupled dye. The staining patterns we observed after incubating HeLa cells with the labelled proteins were similar with cysteine-labelled proteins as with SNAP-tag-labelled proteins, suggesting that Stx2B is in both cases able to bind Gb3 and that neither of the two strategies hinder the protein uptake. (Figure 2.25)

Live cell imaging demonstrated the uptake of the fusion proteins into cells and no reduction in fluorescence intensity was observed for up to 24 hours after the uptake. (Figure 2.26) The strongest fluorescent signal however came from distinct speckles of which some were moving in live cells, suggesting that the delivered proteins were at least partially located in vesicles. A strong fluorescent signal coming from vesicles could also mask a fainter signal from proteins which are dispersed in the cytosol, as the much bigger volume of the cytosol would lead to a dilution of the protein and a less intense signal than in vesicles. In order to determine the nature of these vesicles and the uptake route followed by the monobody, we performed a colocalization analysis with markers for different endocytic compartments. Our first attempt using live cell imaging and markers proved challenging due to the lack of specific live cell markers for different types of endosomes and the high background signal of the existing markers for lysosomes (LysoTracker) and ER (ERTracker). (Figure 2.27 and Figure 2.28) We therefore set out to do immunofluorescence experiments, allowing the use of antibodies against proteins specifically expressed in different subcellular compartments. The colocalization analysis with EEA1 suggested that the proteins first enter early endosomes, from which they appear to escape between 30 minutes and 1 hour after the uptake. (Figure 2.29) Immunofluorescence with an anti-Lamp1 antibody showed only little colocalization between the delivered proteins and the antibody, suggesting that the proteins do not undergo lysosomal degradation, or only to a small extent. (Figure 2.30) The presence of fluorescent protein in live cells for up to 24 hours further supported this hypothesis. (Figure 2.26) The antibodies used as markers for the TGN and for the ER both resulted in high background signal despite efforts of optimization and testing a second antibody (anti-Calnexin) as a marker for the ER, which hampered the application of a threshold algorithm to distinguish true signal from background. Very little colocalization between the delivered proteins and either the markers of the TGN or the ER was detected at any of the analyzed time points, suggesting that the proteins do not reach these compartments within one hour after their uptake. (Figure 2.31 and Figure 2.32) The studied time points were chosen based on the publication from the group of Prof. Hak-Sung Kim, showing that EGFP delivered using this toxin fusion construct already reached the ER after 30 minutes of incubation.<sup>199</sup> However, this hypothesis is based on a qualitative evaluation of cells incubated with the live cell marker ER-tracker, which gave the same diffuse background signal as in our experiments and therefore might be partly unspecific. Early studies of the retrograde transport of Shiga-toxin report the presence of Shiga-toxin B subunit in the Golgi apparatus as early as 20 minutes after the internalization, and its colocalization with a KDEL-tagged fluorescent protein only 2 to 5 hours after the uptake.<sup>147</sup> When coupled to the TDP subunit and a protein to be

delivered, we would therefore maybe expect even longer trafficking times as for the B-subunit alone. Further experiments including longer time points might therefore be needed to show a localization of the delivered proteins in the TGN or ER, demonstrating their uptake via the retrograde trafficking route.

In order to determine the nature of the vesicles in which at least part of the delivered proteins remain for up to 24 hours, further experiments will be needed. Retrograde trafficking can occur at different stages along the endocytic route and originate from different types of endosomes, early, late or recycling endosomes.<sup>184,232</sup> Which of these compartments are involved in the trafficking of Stx2B-TDP-monobody fusion proteins remains to be elucidated. This question could be answered by immunofluorescence experiments with antibodies against proteins which localize to these compartments, for example Rab9 for late endosomes and Rab11 for recycling endosomes.

In order to elucidate whether the delivered proteins are at least in part reaching the cytosol, we added an NLS sequence to their C-terminus. Nuclear localization is easier to measure than cytoplasmic localization, due to the smaller volume of the nucleus as compared to the cytosol and the absence of vesicles which might mask the fainter fluorescent signal in the cytosol. The NLS sequence is bound by importins in the cytosol, which enable its translocation through the nuclear pore. In the nucleus, GTP bound form of the small GTPase Ran binds to this complex, reducing the affinity of importin to the protein and thereby releasing the protein into the nucleus.<sup>233</sup> Thus, if proteins are transported to the nucleus, they must have reached the cytosol, as they can only be recognized by importins in the cytosol. We have found an increase in fluorescence in the nucleus after 24 hours of incubation with two different monobodies, as compared to a control construct of the same sequence which is lacking the NLS. (Figure 2.33 and Figure 2.34) This result demonstrates that the delivered proteins reach the cytosol at least partially, as the fluorescent signal from the nuclei from cell incubated with the NLS-tagged constructs was about 2.5 to 3 times higher than with the control construct. However, this experiment does not allow a quantitative analysis to determine the exact amount of protein taken up in the cytosol.

The lack of precise quantification methods for protein delivery is a common problem in the field, which has been addressed by subcellular fractionation methods. There are numerous different protocols for this technique, which is however prone to artifacts. Our attempts to do a subcellular fractionation of HeLa cells have failed, because we were unable to completely avoid any contamination between the cytosolic and the different endosomal compartments. Besides, this technique requires a large number of cells, as some of the fractions are of very small volumes, which would require large amounts of recombinant proteins to study their subcellular localization using this technique.

In order to overcome these limitations, the group of Prof. Andreas Plückthun has proposed a biotin ligase-based assay to objectively quantify cytosolic delivery, which is independent of protein function.<sup>140</sup> The bacterial BirA enzyme is expressed in the studied cell line and biotinylates delivered proteins carrying an Avi-tag. Since the enzyme is only expressed in the cytosol, the amount of biotinylated versus non-biotinylated protein then gives a measure of the fraction of the protein which has reached the cytosol.

It would be interesting to assess the amount of monobodies delivered using the bacterial toxin construct using this approach, in order to test how effectively the delivered monobodies are translocated to the cytosol.

Another way of showing that the monobodies are reaching the cytosol is by testing their function. Since monobodies interfere with their target in a 1:1 stoichiometry, monobody concentrations above the  $K_D$  of the target binding are required in order to effectively inhibit the target protein.

We have demonstrated that the growth of K562 cells, which are dependent on Bcr-Abl activity for their survival, is inhibited after the delivery of AS25. Moreover, the delivery and growth inhibiting effect of AS25 is dependent on both the presence of Gb3 on the cell surface and the bacterial toxin subunits fused to the monobody. (Figure 2.36) In a second step, we could also demonstrate that K562 cells undergo apoptosis upon delivery of AS25, again in a Gb3- and toxin-dependent manner. (Figure 2.37) We did not observe any growth arrest or apoptosis when replacing the monobody with emGFP, whereas ML3 or the non-binding control monobody HA4\_YA also affected the cell growth and apoptosis, albeit the percentage of cells undergoing apoptosis was lower than with AS25. (Figure 2.38 and Figure 2.39)

The growth-inhibiting effect of ML3 could be explained by the high affinity of ML3 for Lyn, which has been shown to be required for the proliferation of Bcr-Abl dependent cells.<sup>214</sup> In addition, Lyn has been implicated in the survival of drug resistant CML cells.<sup>216,217,219</sup> HA4\_YA however, which also showed mild inhibition, is not binding to either Lyn<sup>210</sup> or Abl kinase<sup>76</sup> and should thus not influence the survival of K562 cells. We cannot exclude that HA4\_YA might bind to another protein involved in K562 cell survival, which we have missed so far. On the other hand, these results could also be explained by a general toxicity of monobodies when delivered into cells. Testing Bcr-Abl negative cell lines could help to resolve this issue.

Shiga toxin and *Pseudomonas* Exotoxin A both hijack the cellular ER-associated protein degradation pathway (ERAD), responsible for the translocation of misfolded proteins from the ER to the cytosol.<sup>144,150,172</sup> Following furin cleavage, the A subunit of Shiga toxin is released from the B-subunit by disulfide bond reduction and then binds to chaperones, which unfold and direct the protein to the translocon channel in the ER membrane for retrotranslocation to the cytosol.<sup>172</sup> The PE domain-II is also recognized by the translocon, leading to the export of the PE domain-II and -III from the ER to the cytosol via the same mechanism.<sup>144</sup> Proteins thus need to unfold to pass through the translocon channel, and the failure of proteins to unfold and exit the ER leads to their accumulation. It has been shown that the accumulation of unfolded or misfolded proteins can cause ER-stress and the activation of the unfolded protein response (UPR) pathway leading to a more rapid protein clearance or apoptosis of the cell.<sup>234,235</sup> This mechanism has been identified as one of the reasons for neuron degeneration in Alzheimer's disease in response to amyloid- $\beta$  accumulation<sup>236</sup> and can also be induced by the expression of recombinant proteins.<sup>237</sup> Monobodies are thermodynamically very stable<sup>238</sup> and might unfold less readily than the effector domains of the bacterial toxins. After furin cleavage and disulfide reduction of the TDP, part of the TDP protein remains covalently attached to the monobody. A stable monobody could possibly prevent the complete unfolding of the cleaved TDP-monobody fusion protein, leading to its accumulation in the ER. This would explain the apoptosis observed to a varying extend

in the cells incubated with different Toxin-monobody constructs. To test this hypothesis, we incubated HeLa cells with the same proteins and measured their growth with the Realtime Glo assay. Although the number of HeLa cells was slightly reduced (15% as compared to the control) after 72 hours of incubation with Toxin-AS25, this effect was much less pronounced than with K562 cells, and we did not observe any decrease of cell numbers upon incubation of the cells with Toxin-SNAP-AS25, Toxin-HA4\_YA or Toxin-emGFP. (data not shown) These results indicate that monobodies do not cause any cytotoxicity related to their accumulation in the ER upon delivery, but rather that AS25 is specifically cytotoxic to K562 cells. On the other hand, HeLa cells could also be less susceptible to ER stress caused by unfolded proteins, or they could take up less protein than K562 cells. Given that the expression of Gb3 in the cell membrane is similar in HeLa cells as in A4GALT-expressing K562 cells, I think that this latter explanation is unlikely. (Compare Figure 2.24 A and Figure 2.35 A)

Kinase inhibitors against Bcr-Abl, such as Imatinib and Nilotinib, bind the kinase domain and thereby inhibit the phosphorylation of downstream targets, including of the transcription factor Stat5, leading to apoptosis.<sup>239-241</sup> When measuring Stat5 phosphorylation in the K562 cells which we incubated with the toxin-AS25 constructs, I could not detect a reduction in Stat5 phosphorylation. (data not shown) In the original publication describing AS25, the authors observed that the percentage of GFP-positive cells after transduction with AS25 decreased over time, indicating that they undergo apoptosis due to Bcr-Abl inhibition by AS25.<sup>78</sup> Even though a reduction of phospho-Stat5 levels has been shown in response to HA4, another Bcr-Abl targeting monobody, that blocks the pY pocket of the SH2 domain, it has not yet been demonstrated that Bcr-Abl inhibition by AS25 also has an effect on Stat5 phosphorylation.<sup>76</sup> A strong effect of the Bcr-Abl SH2-kinase interface on STAT5 activation was demonstrated in Bcr-Abl transduced Ba/F3 and UT-7 cells<sup>77</sup>, but dependence of K562 cell survival on STAT5 activation has not been demonstrated. It therefore remains possible, that Bcr-Abl inhibition by AS25 in K562 cells leads to apoptosis via a different mechanism that is independent of STAT5, or that AS25 is killing K562 cells by an off-target effect. The rapid cell death of K562 cells transduced with AS25 did not allow a selectivity study by affinity purification and it thus remains an open question, whether AS25 has off-targets in K562 cells. Another way of testing if AS25 inhibits Bcr-Abl would be to directly measure the activation loop phosphorylation of Bcr-Abl, which has been shown to be reduced in cells transduced with the tandem monobody HA4-7c12.<sup>77</sup>

Further evidence, that proteins coupled to the Stx2B-TDP toxin construct can reach the cytosol following uptake, is provided by the delivery of VHL-ML3 into Jurkat cells.

In the previous section, I have discussed the degradation of Lck induced by the expression of VHL-ML3 in Jurkat cells. VHL can only actively recruit the Cullin2 E3 ubiquitin ligase complex when in the cytosol, providing a direct way of testing whether the amount of delivered monobodies reaching the cytosol is sufficient to induce the degradation of a target protein. We could demonstrate that indeed, Lck protein levels are reduced in cells expressing Gb3 that were incubated for 48 hours with the Toxin-VHL-ML3 fusion protein. (Figure 2.41) We did not observe this effect in cells lacking Gb3 or when incubating them with VHL-ML3 alone, demonstrating that the degradation of Lck is dependent on the delivery of VHL-ML3 by the bacterial toxin construct. We moreover did not observe the effect when replacing ML3 with HA4\_YA or with AS25, which

shows that ML3 must be folded and fully functional to be able to bind Lck and to lead to its degradation. In order to demonstrate that Lck is degraded following its ubiquitination by the Cullin2 ubiquitin ligase complex, it would be interesting to measure ubiquitinated Lck upon delivery of VHL-ML3.

In summary, we have demonstrated that proteins can be delivered into the cytosol of different cell types by the Stx2B-TDP bacterial toxin fusion construct. The lack of Gb3 expression on Jurkat and K562 cells was used to demonstrate the dependence of the delivery on the binding of Stx2B to Gb3. This provides the delivery system with cell-selectivity that could be used to specifically target certain cancer cell types that overexpress Gb3, such as Burkitt's lymphoma cells, gastric adenocarcinoma, colorectal cancer cells and others.<sup>194,242-244</sup> On the other hand, the receptor specificity might also be altered in order to target other cell types. Efforts to replace Stx2B by a, EGFR targeting rebody or by human EGF have been made by the group of Hak-Sung Kim.<sup>155,199</sup> Albeit less efficient for cellular delivery than the constructs with Stx2B, these approaches are promising and further development could lead to an even more versatile bacterial toxin based delivery system.



## 4 Materials and methods

### 4.1 Cloning

#### 4.1.1 Gateway cloning

The Gateway cloning system was used to clone all the monobodies into the pRV expression vector for the TAP experiments and to clone the SNAP-AS25 and Toxin-monobody-NLS constructs into the pCS2 vector for transient transfection. The monobody genes were first amplified by PCR from the pHFT vector using primers compatible with gateway cloning. The amplified DNA was then introduced into the pDONR201 vector using the BP clonase mix (Invitrogen) and then from there into the destination vector pRVNTAP-GS-Gw using LR clonase mix (Invitrogen), prior to transformation in the *E. coli* strain DH5 $\alpha$ . The hPGK vector was obtained from the laboratory of Prof. Jörg Huelsken (EPFL), and the monobodies ML1, ML3 and HA4\_YA were introduced using Gateway cloning, using the same protocol as for the pRV vector.

All DNA preparations were made using the QIAGEN Miniprep kit and all the resulting constructs were confirmed by sequencing.

#### 4.1.2 Site-directed mutagenesis

A C-terminal cysteine was introduced by site directed mutagenesis using the Quik-change site directed mutagenesis kit (Stratagene) at the C-terminus of the monobodies AS25, AS27, HA4, Nsa5, and Nsa5\_S63K for the CPP project as well as the monobodies AS25 and ML3 for the bacterial toxin project.

#### 4.1.3 Conventional cloning

The toxin-monobody, Toxin-SNAP-monobody constructs and the constructs with nuclear localization sequence (NLS) were generated by conventional cloning. The Stx2B-TDP construct was obtained from the laboratory of Prof. Hak-Sung Kim (KAIST, Korea), and the original insert was replaced by the monobodies using HindIII and NheI restriction enzymes.

The monobodies AS25, AS27, HA4, Nsa5, and Nsa5\_S63K had already previously been cloned in the pHFT2 vector containing a 10x-His-tag and a Flag-tag and were obtained from the laboratory of Prof. Shohei Koide (University of New York, USA).

#### 4.1.4 InFusion cloning

The A4GALT gene was obtained from the Gene expression core facility at EPFL and amplified using primers with 15bp overhangs complementary to the PEM24 vector and compatible with InFusion recombinase. The PEM24 vector was obtained from the laboratory of Prof. Etienne Meylan (EPFL). The vector was linearized using HpaI and PacI restriction enzymes and the A4GALT sequence was introduced using the InFusion recombinase kit according to the manufacturer's instructions (In-Fusion<sup>®</sup> HD Cloning Kit User Manual, Clontech. (2012)). The VHL sequence was obtained from the laboratory of Prof. Sapkota (Dundee University, UK) and was introduced into the

PEM24 vector containing ML3 using InFusion cloning after linearizing the vector using the HpaI restriction enzyme. All constructs in the PEM24 vector were transformed in the *E. coli* strain HB101.

All constructs were confirmed by sequencing and are shown in Table 4.1.

Cloning method	Constructs	Restriction enzymes	N-terminal Tags	C-terminal Tags
Conventional cloning	pHFT2-AS25		10x-His, Flag	
	pHFT2-AS27		10x-His, Flag	
	pHFT2-HA4		10x-His, Flag	
	pHFT2-Nsa5		10x-His, Flag	
	pHFT2-Nsa5_S63K		10x-His, Flag	
	pET21a-Stx2b-TDP-AS25/ML3	HindIII, NheI		6X His, KDEL
	pET21a-Stx2b-TDP-SNAP-AS25/ML3	NotI, NheI	SNAP	6X His, KDEL,
	pET21a-Stx2b-TDP-AS25/ML3-NLS	HindIII, XhoI		6X His, KDEL, NLS
	pET21a-Stx2b-TDP-SNAP-AS25/ML3-NLS	NotI, XhoI	SNAP	6X His, KDEL, NLS
InFusion cloning	pEM24-A4GALT	PacI, HpaI		
	pEM24-VHL-ML3	NdeI, NheI	6X Myc, Flag	
	pEM24-VHL-HA4_Y87A	NdeI, NheI	6X Myc, Flag	
	pET21a-Stx2b-TDP-VHL-ML3	HindIII		6X His, KDEL
	pET21a-Stx2b-TDP-VHL- HA4_Y87A	HindIII		6X His, KDEL
	pET21a-Stx2b-TDP-SNAP-VHL-ML3	NotI	SNAP	6X His, KDEL
Gateway cloning	pRV-N-TAP(GS 2xT)_MS2		2x-Protein G, TEV site, SBP, Myc	
	pRV-N-TAP(GS 2xT)_MS8		2x-Protein G, TEV site, SBP, Myc	
	pRV-N-TAP(GS 2xT)_MY1		2x-Protein G, TEV site, SBP, Myc	
	pRV-N-TAP(GS 2xT)_MY3		2x-Protein G, TEV site, SBP, Myc	
	pRV-N-TAP(GS 2xT)_ML1		2x-Protein G, TEV site, SBP, Myc	
	pRV-N-TAP(GS 2xT)_ML3		2x-Protein G, TEV site, SBP, Myc	
	hPGK-ML1		6x-myc, Flag	
	hPGK-ML3		6x-myc, Flag	
	hPGK-HA4_Y87A		6x-myc, Flag	

	pCS2-SNAP-AS25		6X Myc	
	pCS2-Stx2b-TDP-AS25/ML3-NLS		6X Myc	
	pCS2-Stx2b-TDP-SNAP-AS25/ML3-NLS		6X Myc, SNAP	

Table 4.1 **Expression constructs used for this thesis.**

## 4.2 Expression and purification of recombinant proteins

*E. coli* cells of the BL21(DE3) strain were transformed with the plasmids of all constructs except the toxin constructs, which were transformed in the *E. coli* strain Origami. A preculture was grown over night in 5 ml Luria Bertani (LB) medium containing the appropriate antibiotics Ampicillin for all constructs and additionally Tetracycline for the expression in the Origami strain. The preculture was used to inoculate 1 l Auto Induction Medium (AIM) with the same antibiotics which was incubated for 72 hours at 20°C shaking at 200 rpm. For the constructs bearing a VHL, Terrific Broth (TB) was used instead of AIM, the cells were grown at 37°C until their OD<sub>595 nm</sub> had reached between 0.8 and 1, 1 mM IPTG was added and the culture was further incubated at 18°C shaking at 200 rpm for 15 to 18 hours.

## 4.3 Ni-NTA gravity flow purification

The cells were harvested by centrifugation at 6000g for 15 minutes and the pellet was resuspended using Buffer A (50mM Tris-HCl pH 7.5, 500mM NaCl, 10% glycerol, 20mM Imidazole). 1mM DTT was added to the Buffer A for all the constructs except the ones bearing the Stx2b-TDP construct. For the constructs bearing a VHL, Hepes-based Buffer A was used (50mM Hepes pH 7.5, 500mM NaCl, 10% glycerol, 0.1% Triton-X-100, 20mM Imidazole). DNase was added before the cells were lysed using the Avestin Emulsifex C3. The lysates were centrifuged at 14000g for 45 minutes. The supernatant containing protein was filtered through 0.45µm filter. The protein samples were incubated with 1.5 ml Nickel agarose resin equilibrated with Buffer A for 1 hour at 4°C on a rotor. The lysates were then transferred into a 20mL plastic body column and the flow-through was allowed to pass completely. The column was washed with 50 column volumes of Buffer A prior to elution with 6mL of Buffer B (the same as the buffer A but with 400 mM Imidazole). 50ul each of the lysate, flowthrough, wash and elution fractions were kept for SDS-PAGE analysis later. The concentration of the eluted protein was measured using the Nanodrop spectrophotometer.

## 4.4 Size exclusion chromatography

Ni-NTA purification was followed by size-exclusion chromatography (SEC). The samples were first filtered through 0.22µm syringe filters before loading onto a Superdex size exclusion column attached to the Akta Avant system (GE) and pre-equilibrated with Gel filtration (GF) buffer (25mM Tris-HCL, pH 7.5, 150mM NaCl, and 5% Glycerol, (+1mM DTT, only for monobodies, not for Stx2B-TDP constructs)). A Superdex200 16-600 column was used for purification of Stx2B-TDP constructs, and a Superdex75 26-600 was used for purification of monobodies. The elution fractions corresponding to the molecular weight of the protein were collected and concentrated using Amicon

centrifugal filters. The concentration of the protein was measured using the Nanodrop spectrophotometer.

#### **4.5 SDS-PAGE and Coomassie staining**

The expression of the protein was analyzed by Sodium dodecyl sulphate Polyacrylamide Gel electrophoresis (SDS-PAGE). The flow-through, wash and elution fractions from Ni-NTA purification and elution fraction from SEC were mixed with (4X) sample buffer (0.2M Tris-HCL, 8% SDS, 400mM DTT, 40% glycerol and 0.02% bromophenol blue), boiled at 95°C and loaded in the pockets of a 12% polyacrylamide gel. The gel was run at 100V constant in SDS running buffer (0.25M Tris, 1.92M glycine, 1% (w/v) SDS). SDS-PAGE gels were either used for western blotting or stained in coomassie staining solution (10% acetic acid, 50% absolute ethanol and 0.025% coomassie brilliant blue G-250) for 1-2 hours at room temperature after boiling for few seconds, and destained using destaining solution (10% acetic acid and 20% ethanol). The membranes were imaged using the Odyssey infrared imaging system (Li-Cor) and analysed using ImageStudio software (Li-Cor).

#### **4.6 Monobody labelling**

Two strategies were applied to label the Toxin-monobody constructs. The first strategy was to label the free cysteine in the monobody with maleimide coupled to AlexaFluor488 (ThermoFisher) and the second strategy was to use the fluorogenic benzylguanine (BG) substrates to label the SNAP tag.

##### **4.6.1 Maleimide labelling**

The concentration of the protein was measured using the Nanodrop spectrophotometer. A ten fold molar excess of the required protein amount was used for labelling. Maleimide was added in a 1:10 molar ratio (Protein:dye) and the labelling reaction was left to proceed in dark overnight at 4°C with mild shaking.

##### **4.6.2 SNAP tag labelling**

The concentration of the protein was measured using the Nanodrop spectrophotometer. A ten-fold molar excess of the required protein amount was mixed with the cell impermeable fluorogenic SNAP substrate (BG-Cy5 or BG-547) in a 1:5 molar ratio (Protein:dye) and the reaction was left to proceed in dark for 2 hours at room temperature with mild shaking. For in-cell labelling, 500nM permeable SNAP substrate (BG-Sir) was added to the cells previously incubated with the SNAP-protein for 30 minutes. The cells were then washed with PBS and imaged.

##### **4.6.3 Purification of the labelled protein**

For the recovery and purification of the labelled protein, PD SpinTrap G-25 or PD MidiTrap G-25 (GE Healthcare) columns prepacked with Sephadex G-25 resin were used depending on the sample size. The columns were first equilibrated with the

equilibration buffer (25mM Tris-HCl pH7.5, 150mM NaCl, 5% glycerol). The sample was slowly allowed to pass through the equilibrated resin and the purified labelled protein was eluted in the same buffer.

#### **4.7 Mammalian cell culture**

The adherent cell lines HeLa (human cervical carcinoma), HEK293 (Human Embryonic Kidney), HEK293gp and HEK293T were cultured in Dulbecco's Modified Eagle Medium, DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco) and 1% penicillin/streptomycin (P/S, Amimed). K562 cells and Jurkat cells, which grow in suspension, were cultured in RPMI (Gibco), supplemented with 10% FBS and 1% P/S. All the cells were cultured in flat, round cell culture dishes of 10 or 15 cm diameter at 37°C in 5% CO<sub>2</sub>. Adherent cells were split 3 times per week by washing with PBS, followed by incubation with 1-2 ml Trypsin-EDTA (Gibco) at 37°C, and resuspension by pipetting in growth medium. 10-20% of the cell suspension was added to a new dish containing growth medium. Suspension cells were split 3 times per week. Cells were first centrifuged at 400g for 3 minutes, resuspended in fresh growth medium and counted using the CASYton cell counter. The concentration was adjusted to 1x10<sup>5</sup> cells/ml and the cells were transferred to a new culture dish. Blasticidin at the experimentally determined concentrations of 5 µg/ml for Jurkat (acute T cell leukemia) cells and 7.5 µg/ml for K562 (chronic myeloid leukemia) cells were continuously added to the growth medium of cells after lentiviral transduction with a PEM24 vector containing a blasticidin-resistance.

#### **4.8 Transfection**

Cells were transiently transfected with plasmids using the Polyfect transfection reagent (QIAGEN), according to the manufacturer's protocol. Briefly 1.5ug plasmid DNA was diluted in serum-free DMEM to a total volume of 100µl per dish with 12µl polyfect transfection reagent per dish. The mixture was then vortexed for 10 seconds and incubated for 10 minutes at room temperature. While complex formation occurred, the medium was removed from HeLa cells in the dishes and replaced with 7ml warm complete DMEM. 600µl of complete DMEM was added to the transfection complexes. 712µl of the mixture was added slowly and dropwise to the cells. The plates were swirled and incubated at 37°C in 5% CO<sub>2</sub>.

#### **4.9 Cell lysate preparation**

Adherent cells were washed twice with cold PBS and then lysed by addition of cold lysis buffer (50mM Tris-HCl, pH7.5, 150mM NaCl, 5mM EDTA, 5mM EGTA, 1% NP-40, 50mM NaF, 1mM orthovanadate, 1 mM PMSF (Sigma), 1mg/ml TPCK (Applichem), 10µg/ml protease inhibitor cocktail (Roche)). For 10 cm dishes, 400 µl lysis buffer (recipe below) was used. The cells were detached using a rubber scraper and transferred into an Eppendorf tube by pipetting.

Suspension cells were centrifuged for 5 minutes at 500g at 4°C. The pellet was resuspended in 10 ml cold PBS. In case of the delivery experiment, the cells incubated with protein in Leibovitz medium were directly transferred into 10 ml cold PBS. After

centrifugation for 5 minutes at 500g at 4°C, the pellet was washed once with 1 ml PBS, transferred into a cold Eppendorf tube and resuspended in 50 µl lysis buffer. After 10 minutes incubation on ice, the lysate was cleared by centrifugation for 10 minutes at 20000g at 4°C. The supernatant (lysate) was transferred into fresh cold tubes. Protein concentrations in lysates were determined using Bradford solution (Bio-Rad) diluted 1:5 in water and measuring the OD<sub>595 nm</sub> and comparing to a standard curve determined by measuring known amounts of a  $\gamma$ -globulin solution.

#### 4.10 Immunoblotting

Protein lysate samples were diluted in Laemmli buffer containing 1mM DTT to contain 100 µg protein per 20 µl and then boiled at 95°C for 5 minutes. The samples were separated by SDS-PAGE (as described in section 4.5) and transferred to a nitrocellulose membrane using a semi-dry blotting system (Bio-Rad) or transferred to a nitrocellulose membrane using the iBlot dry blotting system (ThermoFisher). For using the semi-dry blotting system, a nitrocellulose membrane (Whatman, Protran BA85, GE Healthcare) was placed on top of three Whatman cellulose chromatography papers presoaked in 1x western blot buffer (25 mM Tris-HCl, pH 7.5, 192 mM glycine, 10% v/v methanol) and the SDS-PAGE gel was placed on top, followed by another 3 presoaked Whatman papers. The transfer was done in 30 minutes using 300 mA constant current per minigel. For the iBlot system, materials provided by the manufacturer and specified for the iBlot were used and the transfer was done using program 4, according to the manufacturer's protocol.

The membrane was blocked for 1h in a blocking buffer consisting of either 5% w/v non-fat dry milk powder (Bio-Rad) or 5% w/v BSA or Li-Cor blocking buffer diluted 1:5 in either PBS-T or PBS-T (PBS or TBS supplemented with 0.1% Tween20). The buffers used for the individual antibodies are listed in table. The membrane was then incubated with the appropriate antibody (Table 4.2) diluted 1:1000 in blocking buffer at 4°C overnight with gentle shaking. After washing 3x for 5 minutes with either PBS-T or TBS-T, the membrane was incubated with a secondary IRDye680 or IRDye800 or HRP-coupled antibody diluted 1:10000 in wash buffer for 1 hour at room temperature with gentle shaking. The membrane was again washed 3x for 5 minutes and then imaged using the Odyssey fluorescent imaging system (Li-Cor) for IRDye coupled secondary antibodies and analysed using ImageStudio software (Li-Cor). For HRP-coupled secondary antibodies, Amersham ECL Prime Western Blotting Detection reagent (Ge Healthcare) was used for development according to the manufacturer's protocol. Membranes were imaged using the FluorChem IS-8900 (Alpha Innotech).

Primary antibody	Blocking buffer	Wash buffer	Secondary Antibody
Rabbit Yes, CST (3201)	5% BSA in PBS-T	TBS-T	Anti-rabbit HRP, Jackson ImmunoResearch (111-035-003)
Mouse penta-His, Qiagen (34610)	5% BSA in TBS-T	TBS-T	Anti-mouse IRDye 680, Li-Cor (926-32210)
Mouse Myc, ThermoFisher (MA1-2-1316-0800)	5% BSA in TBS-T	TBS-T	Anti-mouse IRDye 680, Li-Cor (926-32210)
Mouse Lck, CST (2657)	5% BSA in TBS-T	TBS-T	Anti-mouse IRDye 680,

			Li-Cor (926-32210)
Mouse Zap70, CST (2709)	5% BSA in TBS-T	TBS-T	Anti-mouse IRDye800, Li-Cor (925-32210)
Rabbit pZap70(Y493), CST (2704)	5% BSA in TBS-T	TBS-T	Anti-rabbit IRDye680, Li-Cor (925-68071)
Rabbit pZap70(Y319), CST (2701)	5% BSA in TBS-T	TBS-T	Anti-rabbit IRDye680, Li-Cor (925-68071)
Mouse tubulin, Sigma (T-9026)	5% BSA in PBS-T	PBS-T	Anti-mouse IRDye680, Li-Cor (926-32210)
Mouse pY 4G10, Millipore (05-1050)	1:5 Li-Cor buffer in PBS-T	PBS-T	Anti-mouse-HRP, Jackson Immunoresearch (115-035-003)
Mouse pERK (T202/Y204), CST (9106)	5% BSA in PBS-T	PBS-T	Anti-mouse IRDye680, Li-Cor (926-32210)
Rabbit ERK, CST (9102)	5% BSA in PBS-T	PBS-T	Anti-rabbit IRDye800, Rockland (611-732-127)
Rabbit pSrc-family (activation loop), CST (D49G4)	5% BSA in TBS-T	TBS-T	Anti-rabbit IRDye800, Rockland (611-732-127)
Rabbit pNFAT (S54), ThermoFisher (44-944G)	1% milk + 1% BSA in TBS-T	TBS-T	Anti-rabbit IRDye680, Li-Cor (925-68071)
Mouse beta-Actin, ThermoFisher (MA1-140)	5% BSA in PBS-T	PBS-T	Anti-mouse IRDye800, Li-Cor (925-32210) or Anti-mouse IRDye 680, Li-Cor (926-32210)

Table 4.2 List of antibodies used for immunoblotting

#### 4.11 Retroviral transduction

Stable cell lines expressing monobodies for TAP experiments were generated by retroviral transduction with the pRV vector constructs, which carry a GFP gene following an IRES site, for selection. HEK293gp cells stably expressing MLV gag and pol proteins were used to produce the virus. They were co-transfected with the VSV-G cDNA and the pRV vector plasmid DNA by calcium phosphate transfection. 4 µg VSV-G DNA and 12 µg pRV plasmid DNA were mixed with 330 µl 0.1x Tris-EDTA, 55 µl 2.5 M CaCl<sub>2</sub> and with water to a total volume of 550 µl. 550 µl 2x Hepes Buffered Saline (HBS) were added dropwise while vortexing. The transfection complexes were incubated at room temperature for 20 minutes, then added dropwise to the HEK293gp cells in 10 cm dishes.

The cells were incubated for 16 hours at 37°C in 5% CO<sub>2</sub> and then the medium was replaced with 6 ml fresh growth medium and the cells were incubated for another 24 hours. The supernatant was removed, filtered through a 0.22µm filter and added to the target cells, after having removed their medium and incubated them with 2µl Polybrene for 5 minutes. The supernatant of the HEK293gp cells was again replaced with 6 ml fresh medium and the infection of the target cells was repeated 24 hours later. After incubating again for 24 hours, the medium of the target cells was replaced, and 2 days later, they were analyzed on a flow cytometer for GFP fluorescence, indicating the integration of the gene.

## 4.12 Lentiviral transduction

Stable cell lines expressing the PEM24 vector with the A4GALT gene, with the VHL fused to the monobodies ML1, ML3 or HA4\_YA, and the hPGK vector with the monobodies ML1, ML3 or HA4\_YA were generated by lentiviral transduction. The hPGK vector carries a GFP gene following an IRES site and the PEM24 vector carries a blasticidin resistance gene for selection.

HEK293T cells were co-transfected with a plasmid encoding the viral envelope, a packaging plasmid and the construct of interest to produce lentivirus by calcium phosphate transfection. For a 10-cm dish, 11.25 µg construct DNA, 3.95 µg envelope plasmid and 7.3 µg packaging plasmid DNA were mixed with 330 µl 0.1x Tris-EDTA, 55 µl 2.5 M CaCl<sub>2</sub> and with water to a total volume of 550 µl. 550 µl 2x Hepes Buffered Saline (HBS) were added dropwise while vortexing. The transfection complexes were incubated at room temperature for 20 minutes, then added dropwise to the HEK293gp cells in 10 cm dishes.

The cells were incubated for 16 hours at 37°C in 5% CO<sub>2</sub> and then the medium was replaced with 7 ml fresh growth medium and the cells were incubated for another 8 hours. The supernatant was removed, filtered through a 0.22µm filter and transferred to a tube for overnight storage. The supernatant of the HEK293T cells was again replaced with 7 ml fresh medium and harvested again in the same way 16 hours later. The medium was removed from the target cells and replaced with half of the pooled and filtered HEK293T supernatant containing the virus. A second infection was performed 24 hours later in the same way. After incubating them for 24 hours, the medium of the target cells was replaced, and 2 days later, they were either analyzed on a flow cytometer for GFP fluorescence (hPGK constructs), or blasticidin was added to the culture medium to select for cell having integrated the gene of interest.

To induce the expression of the gene of interest of cells transduced with the pEM24 vector, 100 ng/ml Doxycycline was added to the medium in a dropwise manner and replaced every 2-3 days. The effective concentration for doxycycline was previously determined. The expression of the protein of interest was determined by lysis and western blot (previously described).

## 4.13 T cell stimulation

Jurkat cells were stimulated with anti-TCR antibody (clone C305, Millipore) at a concentration of 100 ng/ml for 5 min at 37°C. The cells were cooled to 4°C immediately after stimulation by addition of ice-cold PBS and centrifugation at 4°C at 500g for 5 minutes. Cell lysates were prepared as described in section 4.9.

## 4.14 Recombinant SH2 domain pull-down

Jurkat cells were stimulated with the anti-TCR antibody as described in section 4.13 and lysed as described in section 4.9 and the protein content was quantified using Bradford reagent. Recombinant Lck and Src SH2 domains were biotinylated in vitro using recombinant BirA enzyme. Magnetic beads coated with streptavidin were incubated with either biotinylated Lck or Src SH2 domain for 2 hours at 4°C, washed 3 times with TBS

and then incubated for 30 minutes with ML3 or the non-binding HA4\_Y87A control monobody. After washing 3 times with TBS, the beads were incubated for 2 hours with 3 mg of lysate from stimulated Jurkat cells. After another 3 washing steps, the beads were boiled for 5 minutes in SDS-PAGE loading buffer prior to separation on an SDS-PAGE gel.

#### 4.15 Tandem affinity purification

The cells stably expressing the monobodies in the pRV-vector were expanded to approximately  $3 \times 10^9$  cells in 300 ml shaker flasks for mammalian cells at 37°C in 5% CO<sub>2</sub> and while shaking at 180 rpm. The cells were first grown up to a concentration of  $10^6$  cells/ml in 20 ml growth medium in a 15-cm dish and then diluted with growth medium to 100 ml and transferred into a 300 ml shaker flask with a conical bottom after addition of 1x Pluronic-F-68 (Gibco) (100x stock diluted 1:100) to reduce the shear stress during shaking. The cells were then grown while shaking at 180 rpm for 2 days, before diluting the culture medium 1:3 with growth medium supplemented with Pluronic-F-68. The cells were diluted 1:3 every 2 days until reaching a total volume of 1800 ml of cell culture. The cells were then pelleted by centrifugation at 500g for 5 minutes in mammalian cell centrifugation bottles of 150 ml and the pellets were resuspended in 10 ml PBS each, pooled and centrifuged again. The pellet was resuspended in 50 ml PBS and the cells were counted. Between  $2.2$  and  $3.4 \times 10^9$  cells were obtained from 1800 ml cell culture medium. The 50 ml cell suspension in PBS was divided into 2 tubes, centrifuged again, the supernatant was removed and the pellets were frozen in liquid N<sub>2</sub> and stored in the -80°C freezer.

During the TAP, all the steps were performed with cold buffers, at 4°C while working in the cold room. The cell pellets were thawed by resuspension in 25 ml cold TAP buffer (50mM Tris-HCl, pH 7.5, 100 nM NaCl, 5% glycerol, 0.2% NP-40, 1.5 mM MgCl<sub>2</sub> and 25 mM NaF, 1mM orthovanadate, 1 mM PMSF (Sigma), 1mg/ml TPCK (Applichem), 10µg/ml protease inhibitor cocktail (Roche)) and incubated on ice for 15 minutes for lysis to proceed. The lysate was first centrifuged at 4500g for 15 minutes and the supernatant was transferred into ultracentrifuge tubes (polycarbonate thick walls, 38 ml, Beckman-Coulter). The lysates were centrifuged for 1 hour at 4°C at 100,000g in the ultracentrifuge using the rotor SW32Ti (Beckman Coulter).

The supernatant was transferred to a 50-ml tube while avoiding the lipid layer (total extract, TE) and an aliquot of 50µl was saved. IgG beads (Sigma, A2909) were washed and equilibrated with TAP buffer and then 200µl of the beads diluted 1:1 in TAP buffer were added to the lysate and incubated on a rotating wheel at 4°C for 2 hours.

The 50µl TE sample was used to calculate the protein concentration using a Bradford assay. The beads were recovered by centrifugation at 600 rpm for 3 minutes at 4°C, an aliquot of the supernatant was saved (SN1) and the rest discarded and the beads were transferred into a Bio-Spin column (Bio-Rad, 732-6008). The beads were washed with 10 ml TAP buffer and then with 5 ml TEV-NP-40 buffer (10 mM Tris-HCl, pH7.5, 100 mM NaCl, 0.5 M EDTA, pH8.0, 0.2% NP-40). 50µg of TEV protease were diluted to a final volume of 400µl in TEV buffer (10 mM Tris-HCl, pH7.5, 100 mM NaCl, 0.5 M EDTA, pH8.0) and added to the beads after closing the bottom of the column. The beads were incubated for 1h at 16°C while shaking at 800rpm. The TEV eluate was collected from the column into an Eppendorf tube and another 400µl of TEV-NP-40

buffer were added to the beads and eluted by gravity flow into the same tube. 30µl were saved (TEV).

Streptavidin UltraLink Resin beads (PierceThermo, 53114) were washed 3 times with TEV buffer and 150µl of the beads diluted 1:1 in TEV buffer were added per sample and incubated for 1h on a rotating wheel at 4°C. The Streptavidin beads were transferred into a Bio-Spin column, the eluate was discarded except for a 50µl sample which was saved (SN2) and the beads were washed with 10 ml TEV-NP-40 buffer. The proteins were eluted from the column with 2x 250 µl 150 mM HCl into tubes containing 125 µl TEAB. A 30 µl sample was saved (E1). The beads were boiled with 75µl 4x sample buffer for 5 minutes at 95°C and a sample was saved (BB). The eluate was frozen in liquid N<sub>2</sub>, lyophilized overnight and resuspended in 50 µl 4x Laemmli sample buffer (Bio-Rad) containing 0.5% β-mercapto-ethanol. A 10% fraction of the sample was separated on an SDS-PAGE gel and protein bands were visualized by silver staining. The saved samples (TE, SN1, TEV, SN2, E1 and BB) were separated by SDS-PAGE and immunoblotted with the indicated antibodies. The 90% fraction of the final eluate was separated by SDS-PAGE (any kD gel, Bio-Rad, 456-9033), stained with R-250 Coomassie Blue solution and submitted for mass spectrometry analysis by the EPFL proteomics core facility. Gel lanes were cut into pieces and subjected to in-gel digestion with trypsin. Extracted peptides were separated by reversed-phase chromatography on a Dionex Ultimate 3000 RSLC nano UPLC system (Dionex) on-line connected in-line either with an Orbitrap Elite or an Orbitrap Q-Exactive Mass Spectrometer (Thermo Fischer Scientific). Raw data were processed with Proteome Discoverer software and searched with Mascot against a human database (UniProt). The data was obtained from the facility in a Scaffold3 software file format and was analyzed as described in the results section.

## 4.16 Flow cytometry

### 4.16.1 GFP expression analysis

Flow cytometry was used to assess the DNA integration of the pRV and hPGK vectors by measuring the 488 nm signal from the GFP. 1 ml of cell suspension with 10<sup>5</sup> to 10<sup>6</sup> cells were taken, washed 2 times with PBS, resuspended in PBS supplemented with 0.5% BSA (FACS buffer) and then analyzed on the BD LSRII flow cytometer (Beckman-Dickinson) using the 488nm laser and a 525/40nm bandpass filter.

### 4.16.2 Cell sorting

Cell sorting was performed on a FACS AriaII device at the EPFL Flow cytometry core facility. 10<sup>7</sup> cells washed with PBS and resuspended in FACS buffer, before sorting them according to their 488 nm fluorescence into tubes containing FBS. Afterwards, the cells were centrifuged at 500g for 5 minutes and resuspended in growth medium.

### 4.16.3 Gb3 expression analysis

Flow cytometry was used to assess the expression of Gb3 receptor (CD77) on the surface of different cells. HeLa cells were first trypsinized and resuspended in 5mL fresh DMEM.

5x10<sup>5</sup> cells per sample were centrifuged for 3 minutes at 500g and washed once with PBS. The cell pellet was resuspended in 100 µl FACS buffer containing human FITC-conjugated anti-CD77 antibody (Biolegend, 357103) diluted 1:20 and the tube was incubated at dark on ice for 20 minutes. The cells were washed with 3 ml PBS and the pellet was resuspended in 400µl FACS buffer for analysis on the BD LSRII flow cytometer (Beckman-Dickinson) using the 488nm laser and a 525/40nm bandpass filter.

#### 4.16.4 Protein delivery

Flow cytometry was used to assess the delivery of the proteins in suspension cells. 5x10<sup>5</sup> K562 and Jurkat cells transduced with the A4GALT gene and induced with 100 ng/ml doxycycline or uninduced were collected after centrifugation at 500g for 3 minutes and resuspension in fresh buffer. After one washing step with PBS, the supernatant was aspirated and pellet was resuspended in Leibovitz medium containing the AlexaFluor488 or SNAP-substrate or unlabeled protein to be delivered. The cells were incubated in a 48-well plate for 30 minutes at 37°C. The cells were collected in 15ml tubes, centrifuged and washed three times with PBS. The cell pellets were then resuspended in 400 µl FACS buffer for analysis. The samples were analyzed on an LSRII flow cytometer (Beckman-Dickinson) using the 488nm laser and a 525/40nm bandpass filter for the delivery of AF488-labelled proteins or using the 640nm laser and a 670/14nm bandpass filter for the delivery of SNAP-tagged proteins.

#### 4.16.5 AnnexinV/7AAD staining

K562 cells either induced with 100 ng/ml doxycycline for 24h or uninduced were pelleted by centrifugation at 400g for 3 minutes and resuspended in Leibovitz's medium supplemented with 10% FBS, 1% P/S and 7.5 µg/ml blasticidin. The cells were counted on the CASYton cell counter and their concentration was adjusted to 4x10<sup>5</sup> cells/ml for the 24-hour timepoint, 2x10<sup>5</sup> cells/ml for the 48-hour timepoint and 1x10<sup>5</sup> cells/ml for the 72-hour timepoint. 100ng/ml doxycycline was added to the cells previously induced with doxycycline. 250 µl of cell suspension were transferred in each well of a 24-well plate and 250 µl of recombinant protein diluted in Leibovitz's medium and sterile filtered were added. The cells were incubated for 24, 48 or 72 hours at 37°C in 5% CO<sub>2</sub>.

AnnexinV binding buffer was prepared by diluting the 10x AnnexinV binding buffer (BD) 1:10 in sterile water. The AnnexinV/7AAD staining solution was prepared by diluting Cy5-coupled AnnexinV (BD) 1:50 and 7AAD (BD) 1:20 in 1x AnnexinV binding buffer. The cells were transferred from the wells of the 24-well plate into 15-ml tubes with 10 ml PBS and centrifuged at 500g for 5 minutes. The pellets were resuspended in 100µl AnnexinV staining solution and transferred to FACS tubes. After 15 minutes of incubation at room temperature in the dark, 200µl of AnnexinV binding buffer was added without washing the cells and the cells were analyzed on the LSRII flow cytometer (Beckman-Dickinson) using the 640nm laser and a 670/14nm bandpass filter for detection of Cy5 and the 561nm laser and a 670/30nm bandpass filter for detection of 7AAD. Unstained and single stained controls were included in each experiment and used for compensation between the Cy5 and 7AAD signals. Analysis and compensation were done using FlowJo software.

#### 4.17 CellTiter-Glo luminescent cell viability assay

The cells were seeded in a 96-well plate and incubated with adducts for 1h, and then in adduct-free medium for 24h after 3 washing steps with PBS. The cells were then lysed by addition of the CellTiter Glo reagent (Promega), according to the manufacturer's protocol. The luminescent signal was measured using the M5 plate reader (Molecular devices).

#### 4.18 Realtime-Glo luminescent cell viability assay

The Realtime-Glo assay kit (Promega) was used to monitor cell viability of K562 cells after protein delivery. For this, K562 cells transduced with A4GALT were induced with 100 ng/ml of doxycycline, which was added to the medium 24 hours before the experiment, or were left uninduced. The cells were pelleted by centrifugation at 400g for 3 minutes and the pellets were resuspended in Leibovitz's medium supplemented with 10% FBS, 1% P/S and 7.5 µg/ml blasticidin. After counting the cells using the CASYton cell counter, they were diluted to  $1 \times 10^5$  cells/ml and the NanoLuc enzyme (Promega) as well as the MT cell viability substrate were added to the cell suspension at a 1:1000 dilution as indicated in the manufacturer's protocol. The cells were then seeded in an opaque white 96-well plate at a density of  $5 \times 10^3$  cells in 50µl per well. The proteins diluted at 2x the indicated end concentration in Leibovitz's medium supplemented with 10% FBS, 1% P/S and 7.5 µg/ml blasticidin were sterile filtered and 50µl were added per well. The plates were incubated at 37°C in 5% CO<sub>2</sub>. After 24, 48, 72 and 96 hours of incubation, the plates were inserted into the M5 plate reader (Molecular devices) and luminescence was measured.

#### 4.19 Microscopy

##### 4.19.1 Live cell microscopy

$8 \times 10^4$  cells were seeded 24 hours before incubation with the protein in 35 mm glass-bottom dishes. Before protein incubation, the medium was aspirated and the cells were washed once with 1 ml sterile PBS. The protein was diluted in Leibovitz medium supplemented with 10% FBS and 1% P/S and the cells were incubated in 1 ml of the protein dilution at the indicated concentration for the indicated time. For incubation times longer than 1 hour, the protein dilutions were sterile filtered prior to addition to the cells. Before imaging, the cells were washed twice with PBS and fresh Leibovitz's medium was added. The dishes were imaged using a spinning disk confocal microscope (Yokogawa Spinning-Disk "CSU-W1") at the BIOP, EPFL with the 60x oil objective using appropriate laser settings. For AF488 labelled proteins, the 488nm laser was used. For BG-Cy5 and BG-Sir labelled SNAP tagged proteins, 640nm laser was used. For each sample, individual z-slices were imaged at the nuclear plane. The same laser settings were used for each sample of the same experiment.

#### 4.19.2 Immunofluorescence

8 x10<sup>4</sup> cells were seeded 24 hours before incubation with the protein in a six-well plate with 18mm sterile round coverslips. Before protein incubation, the medium was aspirated and the cells were washed once with 1 ml PBS. The protein was diluted in Leibovitz medium supplemented with 10% FBS and 1% P/S and the cells were incubated in 1 ml of the protein dilution at the indicated concentration for the indicated time. The cells were washed once with PBS and then fixed in 4% Paraformaldehyde (PFA) at room temperature for 15 minutes (all antibodies except for anti-Calreticulin) or in ice-cold 100% Methanol for 5 minutes at -20°C. (anti-Calreticulin antibody) The cells were rehydrated three times in PBS supplemented with 0.01% Triton-X-100 (PBS-TX). The coverslips were incubated for 1h in permeabilization solution (PBS, Saponin 0.1%, BSA 5%) at room temperature and then washed 3 times with PBS-TX. The coverslips were placed on the parafilm-coated and labelled lid of a 6-well plate in a humidity chamber. PBS-TX was added on top to prevent drying. PBS-TX was aspirated and 50µl of primary antibody solution diluted 1:500 in PBS-TX supplemented with 1% BSA was added. The coverslips were incubated overnight at 4°C in the humidity chamber. The coverslips were washed with PBS-TX once, then 3 times with 5 minutes incubation between washes, each time by aspirating the solution from the coverslips and pipetting fresh solution immediately to the rim of the coverslip to prevent coverslips from drying. 50µl of secondary antibody diluted 1:1000 in PBS-TX supplemented with 1% BSA was added on each coverslip and incubated for 45 minutes in the humidity chamber at room temperature. The cells were washed once with PBS-TX, then 3 times with 5 minutes incubation between washes. 100µl of 1:1000 dilution of Hoechst in PBS was added on each coverslip and incubated for 10 minutes at room temperature in the humidity chamber. The cells were washed 3 times with PBS. For each coverslip, 2µl Fluoromount-G mounting medium was put on a microscopy slide just before mounting. Each coverslip was taken up with tweezers, dipped into a becher with water and the excess water was removed with a paper tissue. It was then placed upside down on the slide on top of the mounting media. Excess media was removed with a paper tissue and the slides were allowed to dry in the dark. After drying, the cells were imaged on the spinning disk confocal microscopy (Yokogawa Spinning-Disk “CSU-W1”) at the BIOP, EPFL with the 60x oil objective using appropriate laser settings.

Primary antibodies	Organelle stained	Secondary antibodies
Mouse EEA1 (Sigma BD610547)	Early Endosomes	Anti-Mouse FITC (Sigma F0257) and anti-Mouse 647 (Biolegend 405322)
Mouse Lamp1 (Sigma BD555798)	Lysosomes	
Rabbit Golgin-97 (Abcam Ab84340)	Golgi bodies	Anti-Rabbit 488 (Abcam Ab15007) and anti-Rabbit Cy5
Rabbit Calnexin (Abcam Ab22595)	Endoplasmatic Reticulum	
Rabbit Calreticulin (Abcam Ab92516)		

Table 4.3 List of antibodies used for immunofluorescence

## 4.20 Image processing and analysis

After image acquisition, the images were processed and analysed in Fiji/ImageJ software. The Common Tools Plugin (written by the BIOP, EPFL) was used to process all the images. First, Look Up Tables (LUTs) were selected for each channel in the image. Then, brightness and contrast (B/C) for each channel was adjusted. The image without protein labelling was used as a reference to subtract background signal from the true signal. The same settings were applied to all the images from one experiment.

### 4.20.1 Colocalization analysis

For colocalization analysis, Regions of Interest (ROIs) were selected from the processed images with correct LUTs and B/C. This was done using the MultiManualSelect tool (written by the BIOP, EPFL). ROIs were drawn for each cell in each image using polygon selection tool. At least 15-20 images per experimental condition were analyzed with 1-4 cells per image. After selecting ROIs, a threshold algorithm was selected for each channel, which was appropriate to select true signal and omit the background noise for each image in a similar way. In contrast to using a fixed threshold value, this method allowed to correct for overall intensity differences between the images. The images were then all processed in the same way using the selected threshold for the individual channels. Using the JaCOP tools (written by PTBIOP, EPFL) in ImageJ, the Mander's overlap coefficients between the channels of the protein signal and the antibody signal were calculated for each cell (ROI) individually by averaging over the different z-stacks.

### 4.20.2 Nuclear localization analysis

After acquiring images as described in section 4.19.1 from cell incubated with NLS-tagged proteins, the following ImageJ script was used to process all the images and quantify the signal from the 488nm channel in the nucleus. It loops through all the images in a chosen folder, and for each image, it loops through the 5 slices taken in the nuclear plane. The Hoechst stained nuclei of each slice are first defined as regions of interest (ROIs), by setting a threshold selecting pixels with a Hoechst signal above a certain value. Each distinct area with a minimum number of pixels is then saved as a ROI, discarding small speckles that might result from noise or artifacts. The mean intensity of the 488 nm signal of each ROI is measured, corresponding to the signal from the delivered proteins. The mean intensity values of the 5 slices corresponding to the same ROI are matched and the average is calculated. This gives a measure for the average 488 nm fluorescence intensity throughout the 5 slices of each nucleus. (Text written in Calibri are comments, explaining the function of the macro and not part of the macro itself.)

```
// @File(label="Select the folder of images to process",
style="directory") myFolder
// @Integer(label="DAPI channel",value=2) chDAPI
// @Integer(label="channel to measure",value=1) chMeasure
```

// Macro to measure the intensity of a defined channel in the DAPI-stained area for each cell. The above commands must stand at the very beginning of the macro to function as a dialogue with the user

```

// START OF : Functions

// This function defines a list of accepted file formats

function isImage(filename){
    extensions= newArray("lsm", "lei", "lif", "tif", "ics", "bmp",
    "png", "TIF", "tiff", "czi", "zvi", "nd2");
    for (i=0; i<extensions.length; i++) { // for each index of the
    array extensions
        if(endsWith(filename, "."+extensions[i])) {
            return true;
        }
    }
    return false;
}

//End of Functions

//Start of Macro

// These commands are used to install the BIOP Library
call("BIOP_LibInstaller.installLibrary",
"BIOP"+File.separator+"BIOPLib.ijm");
run("Set Measurements...", "area mean standard min display
redirect=None decimal=5");
//make a new folder within the directory called "saved images"
saveDir = myFolder + File.separator + "results" + File.separator ;
File.makeDirectory(saveDir);
run("Close All");

//make text file to write results

f = File.open("");
print(f, "Image, ROI, Mean \n");

//Make an array containing all the file names in the folder
filesList = getFileList(myFolder);

//set the particle limit for the analyze particles
partLimit = "20-2000"

setBatchMode(true);

// for each file of the folder to process
for (j = 0 ; j < lengthOf(filesList) ; j++)
{
    run("Close All");
    run("Clear Results");
    roiManager("Reset");
    // check if it's an image
    if ( isImage( filesList[j] ) ) {
        open(myFolder+File.separator+filesList[j]);

        run("Set Measurements...", "area mean standard min display
        redirect=None decimal=5");

    }

}

// get some infos about the image : Name, size, pixel...
imageName = getTitle();
getDimensions(width, height, channels, slices, frames);
getVoxelSize(widthPixel, heightPixel, depthPixel, unitPixel);
print(imageName);
slicesToAnalyze=slices-2;

```

```

//Split channels
    run("Split Channels");
//select DAPI channel
    selectWindow("C"+chDAPI+"-"+imageName+"");
//split stacks to images
    run("Stack to Images");
//select channel to measure
    selectWindow("C"+chMeasure+"-"+imageName+"");
//split stacks to images
    run("Stack to Images");

/*open a loop to go through all slices of the DAPI image to add ROIs of the nuclei to the ROI
manager,
and then measure the intensity in these ROIs in the corresponding slice of the channel of
interest*/

//for each slice:
    for (i = 0 ; i < slicesToAnalyze ; i++ )
    {
//set the measurements, because they get reset each time in the writeResults function
        run("Set Measurements...", "area mean standard min
display redirect=None decimal=5");
//reset ROI manager
        roiManager("Reset");
//Set Threshold for DAPI stain
        selectWindow("c:"+chDAPI+"/3 z:"+ (i+1) + "/" + slices + " -
sdcGFP/sdcDAPI/Brightfield");
        setAutoThreshold("Li dark");
        run("Convert to Mask");
        run("Median...", "radius=3");
//Analyze the particles and add them to the ROI manager
        run("Analyze Particles...", "size="+partLimit+"
show=[Count Masks] exclude add");
//count the added ROIs
        ROIcount=roiManager("Count");
//create an array the size of the number of ROIs, ex 3
        intensity_roi = newArray(ROIcount);
        print(ROIcount+" ROIs have been counted in slice "+i+" in
image "+imageName);
//Give an error message when the ROI count is 0
        if (ROIcount == 0)
        {
            waitForUser("No ROI detected");
        }
// Measure the detected ROI on the channel of interest
        selectWindow("c:"+chMeasure+"/3 z:"+ (i+1) + "/" + slices + " -
sdcGFP/sdcDAPI/Brightfield");
        roiManager("Measure");

// and add the values for each ROI from the results table in one array per slice
        for (x=0; x< ROIcount; x++ )
        {
            intensity_roi[x]=getResult("Mean", x);
            Array.print(intensity_roi);
        }
        IJ.renameResults("Results", "temp");
        writeResults("Final_Results", "Label", "next",
imageName);
        writeResults("Final_Results", "Slice", "current",
"Slice_"+(i+1));

```

```

//add a column for the following ROI
    for (l=0; l< ROIcount; l++ )
    {
        run("Set Measurements...", "area mean standard min
        display redirect=None decimal=5");
        writeResults("Final_Results", "ROI_"+(l+1), "current",
        intensity_roi[l]);
    }

    selectWindow("temp");
    run("Close");
}

//rename the final results table to results to be able to extract values
    selectWindow("Final_Results");
    waitForUser("check final results table");
    IJ.renameResults("Final_Results", "Results");

//write the values form this table into the csv file
//make a new array containing the values from each slice for 1 ROI each
    ROI = newArray(slicesToAnalyze);
//make a new array containing the mean intensity over all slices for each ROI
    Mean_ROI = newArray(ROIcount);
    for(r=0; r<ROIcount; r++)
    {
        for (s=0; s<slicesToAnalyze; s++)
        {
            print("the intensity of ROI "+r+" in slice "+s+" is
            "      +getResult("ROI_"+(r+1), s));
            ROI[s]=getResult("ROI_"+(r+1), s);
        }
        //get the mean of intensities from each slice for this ROI
        Array.print(ROI);
        Array.getStatistics(ROI,MIN,MAX,MEAN,STDDEV);
        print("the mean of ROI "+r+1+" in image "+imageName+" is
        "+MEAN);
        //put this value as a first value in the array Mean_ROI, containing the means of intensities from
        all ROIs from this image
        Mean_ROI[r]=MEAN;
        Array.print(Mean_ROI);
        //write the Mean_ROI for this ROI in a new line in the csv file
        print(f, imageName +", "+r+1+", "+Mean_ROI[r]+"\\n");
    }
    run("Close All");
    run("Clear Results");
}else{
// print name of files that are not images
    print(filesList[j]+" is not an image");
}
}
//End of Macro

```



## 5 References

1. Sliwkowski, M. X. & Mellman, I. Antibody Therapeutics in Cancer. *Science* **341**, 1192–1198 (2013).
2. Hantschel, O. Unexpected off-targets and paradoxical pathway activation by kinase inhibitors. *ACS Chem. Biol.* **10**, 234–245 (2015).
3. Holohan, C., Van Schaeybroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: an evolving paradigm. *Nature reviews. Cancer* **13**, 714–726 (2013).
4. Fedorov, O., Müller, S. & Knapp, S. The (un)targeted cancer kinome. *Nature chemical biology* **6**, 166–169 (2010).
5. Bruce, V., Ta, A. & Mc Naughton, B. Minimalist Antibodies and Mimetics: An Update and Recent Applications. *Chembiochem : a European journal of chemical biology* 1–10 (2016).
6. Kintzing, J. R., Interrante, M. V. F. & Cochran, J. R. Emerging Strategies for Developing Next-Generation Protein Therapeutics for Cancer Treatment. *Trends in Pharmacological Sciences* 1–16 (2016).
7. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002).
8. Eckhart, W., Hutchinson, M. A. & Hunter, T. An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. *Cell* **18**, 925–933 (1979).
9. Hunter, T. Tyrosine phosphorylation: thirty years and counting. *Curr. Opin. Cell Biol.* **21**, 140–146 (2009).
10. Liu, Z. *et al.* Legumain protease-activated TAT-liposome cargo for targeting tumours and their microenvironment. *Nature communications* **5**, 4280 (2014).
11. Moran, M. F. *et al.* Src homology region 2 domains direct protein-protein interactions in signal transduction. *Proceedings of the National Academy of Sciences* **87**, 8622–8626 (1990).
12. Sadowski, I., Stone, J. C. & Pawson, T. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus P130gag-fps. *Mol. Cell. Biol.* **6**, 4396–4408 (1986).
13. Pawson, T. Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* **116**, 191–203 (2004).
14. Meisenhelder, J., Suh, P. G., Rhee, S. G. & Hunter, T. Phospholipase C-gamma is a substrate for the PDGF and EGF receptor protein-tyrosine kinases in vivo and in vitro. *Cell* **57**, 1109–1122 (1989).
15. Margolis, B. *et al.* The tyrosine phosphorylated carboxyterminus of the EGF receptor is a binding site for GAP and PLC-gamma. *The EMBO journal* **9**, 4375–4380 (1990).
16. Stahl, M. L., Ferez, C. R., Kelleher, K. L., Kriz, R. W. & Knopf, J. L. Sequence similarity of phospholipase C with the non-catalytic region of src. *Nature* **332**, 269–272 (1988).

17. Vogel, U. S. *et al.* Cloning of bovine GAP and its interaction with oncogenic ras p21. *Nature* **335**, 90–93 (1988).
18. Ellis, C., Moran, M., McCormick, F. & Pawson, T. Phosphorylation of GAP and GAP-associated proteins by transforming and mitogenic tyrosine kinases. *Nature* **343**, 377–381 (1990).
19. Neel, B. G., Gu, H. & Pao, L. The ‘Shp’ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. *Trends in Biochemical Sciences* **28**, 284–293 (2003).
20. Waksman, G. *et al.* Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature* **358**, 646–653 (1992).
21. Domchek, S. M., Auger, K. R., Chatterjee, S., Burke, T. R. & Shoelson, S. E. Inhibition of SH2 domain/phosphoprotein association by a nonhydrolyzable phosphonopeptide. *Biochemistry* **31**, 9865–9870 (1992).
22. Songyang, Z. *et al.* SH2 domains recognize specific phosphopeptide sequences. *Cell* **72**, 767–778 (1993).
23. Scott, J. D. & Pawson, T. Cell signaling in space and time: where proteins come together and when they're apart. *Science* **326**, 1220–1224 (2009).
24. Park, M.-J. *et al.* SH2 Domains Serve as Lipid-Binding Modules for pTyr-Signaling Proteins. 1–15 (2016).
25. Rous, P. A transmissible avian neoplasm. (Sarcoma of the common fowl). 1–13 (1910).
26. Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt P. K. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. 1–4 (1976).
27. Veillette, A., Bookman, M. A., Horak, E. M. & Bolen, J. B. The CD4 and CD8 T cell surface antigens are associated with the internal membrane tyrosine-protein kinase p56lck. *Cell* **55**, 301–308 (1988).
28. Yamanashi, Y. *et al.* Selective expression of a protein-tyrosine kinase, p56lyn, in hematopoietic cells and association with production of human T-cell lymphotropic virus type I. *Proceedings of the National Academy of Sciences* **86**, 6538–6542 (1989).
29. Thomas, S. M. & Brugge, J. S. Cellular functions regulated by Src family kinases. *Annu. Rev. Cell Dev. Biol.* **13**, 513–609 (1997).
30. Lowell, C. A. & Soriano, P. Knockouts of Src-family kinases: stiff bones, wimpy T cells, and bad memories. *Genes Dev.* **10**, 1845–1857 (1996).
31. Zhang, S. & Yu, D. Targeting Src family kinases in anti-cancer therapies: turning promise into triumph. *Trends in Pharmacological Sciences* **33**, 122–128 (2012).
32. Courtney, A. H., Lo, W.-L. & Weiss, A. TCR Signaling: Mechanisms of Initiation and Propagation. *Trends in Biochemical Sciences* **43**, 108–123 (2018).
33. Straus, D. B., Chan, A. C., Patai, B. & Weiss, A. SH2 domain function is essential for the role of the Lck tyrosine kinase in T cell receptor signal transduction. *Journal of biological chemistry* **271**, 9976–9981 (1996).
34. Straus, D. B. & Weiss, A. Genetic evidence for the involvement of the lck tyrosine kinase in signal transduction through the T cell antigen receptor. *Cell* **70**, 585–593 (1992).

35. Straus, D. B. & Weiss, A. The CD3 chains of the T cell antigen receptor associate with the ZAP-70 tyrosine kinase and are tyrosine phosphorylated after receptor stimulation. *J. Exp. Med.* **178**, 1523–1530 (1993).
36. Wang, H. *et al.* ZAP-70: An Essential Kinase in T-cell Signaling. *Cold Spring Harbor Perspectives in Biology* **2**, a002279–a002279 (2010).
37. Balagopalan, L., Coussens, N. P., Sherman, E., Samelson, L. E. & Sommers, C. L. The LAT story: a tale of cooperativity, coordination, and choreography. *Cold Spring Harbor Perspectives in Biology* **2**, a005512–a005512 (2010).
38. Brownlie, R. J. & Zamoyska, R. T cell receptor signalling networks: branched, diversified and bounded. *Nat Rev Immunol* **13**, 257–269 (2013).
39. Xu, C. *et al.* Regulation of T Cell Receptor Activation by Dynamic Membrane Binding of the CD3 $\epsilon$  Cytoplasmic Tyrosine-Based Motif. *Cell* **135**, 702–713 (2008).
40. Gagnon, E., Schubert, D. A., Gordo, S., Chu, H. H. & Wucherpfennig, K. W. Local changes in lipid environment of TCR microclusters regulate membrane binding by the CD3 $\epsilon$  cytoplasmic domain. *J. Exp. Med.* **209**, 2423–2439 (2012).
41. Davis, S. J. & van der Merwe, P. A. The kinetic-segregation model: TCR triggering and beyond. *Nat Immunol* **7**, 803–809 (2006).
42. Stepanek, O. *et al.* Coreceptor scanning by the T cell receptor provides a mechanism for T cell tolerance. *Cell* **159**, 333–345 (2014).
43. Billadeau, D. D., Nolz, J. C. & Gomez, T. S. Regulation of T-cell activation by the cytoskeleton. *Nat Rev Immunol* **7**, 131–143 (2007).
44. Feske, S., Skolnik, E. Y. & Prakriya, M. Ion channels and transporters in lymphocyte function and immunity. *Nat Rev Immunol* **12**, 532–547 (2012).
45. Bunnell, S. C., Kapoor, V., Tribble, R. P., Zhang, W. & Samelson, L. E. Dynamic actin polymerization drives T cell receptor-induced spreading: a role for the signal transduction adaptor LAT. *Immunity* **14**, 315–329 (2001).
46. Paul, S. & Schaefer, B. C. A new look at T cell receptor signaling to nuclear factor- $\kappa$ B. *Trends Immunol.* **34**, 269–281 (2013).
47. Andreotti, A. H., Schwartzberg, P. L., Joseph, R. E. & Berg, L. J. T-cell signaling regulated by the Tec family kinase, Itk. *Cold Spring Harbor Perspectives in Biology* **2**, a002287–a002287 (2010).
48. Wetzler, M. *et al.* Subcellular localization of Bcr, Abl, and Bcr-Abl proteins in normal and leukemic cells and correlation of expression with myeloid differentiation. *J. Clin. Invest.* **92**, 1925–1939 (1993).
49. Van Etten, R. A. Cycling, stressed-out and nervous: cellular functions of c-Abl. *Trends Cell Biol.* **9**, 179–186 (1999).
50. Hantschel, O. & Superti-Furga, G. Regulation of the c-Abl and Bcr-Abl tyrosine kinases. *Nature reviews. Molecular cell biology* **5**, 33–44 (2004).
51. Worms, H. P., Kaplan, D. R. & Roberts, T. M. Retrovirus Shuttle Vector for Study of Kinase Activities of pp60csrc Synthesized. *Mol. Cell. Biol.* 1–8 (1986).
52. Hunter, T. A Tail of Two src's: Mutatis Mutandis. 1–4 (1987).
53. Courtneidge, S. A. Activation of the pp60C-src kinase by middle T antigen binding or by dephosphorylation. *The EMBO journal* 1–7 (1985).

54. Nada, S., Okada, M., MacAuley, A., Cooper, J. A. & Nakagawa, H. Cloning of a complementary DNA for a protein-tyrosine kinase that specifically phosphorylates a negative regulatory site of p60c-src. *Nature* **351**, 69–72 (1991).
55. Hamaguchi, I. *et al.* Analysis of CSK homologous kinase (CHK/HYL) in hematopoiesis by utilizing gene knockout mice. *Biochemical and biophysical research communications* **224**, 172–179 (1996).
56. Davidson, D., Chow, L. M. & Veillette, A. Chk, a Csk family tyrosine protein kinase, exhibits Csk-like activity in fibroblasts, but not in an antigen-specific T-cell line. *Journal of biological chemistry* **272**, 1355–1362 (1997).
57. Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595–602 (1997).
58. Williams, J. C. *et al.* The 2.35 Å crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. *J. Mol. Biol.* **274**, 757–775 (1997).
59. Xu, W., Harrison, S. C., Doshi, A. & Eck, M. J. Crystal Structures of c-Src Reveal Features of Its Autoinhibitory Mechanism. *Molecular Cell* 1–10 (1999).
60. Barilá, D. & Superti-Furga, G. An intramolecular SH3-domain interaction regulates c-Abl activity. *Nat. Genet.* **18**, 280–282 (1998).
61. Sicheri, F. & Kuriyan, J. Structures of Src-family tyrosine kinases. *Curr. Opin. Struct. Biol.* **7**, 777–785 (1997).
62. Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S. A. & Draetta, G. Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. *The EMBO journal* **12**, 2625–2634 (1993).
63. Hantschel, O. *et al.* A myristoyl/phosphotyrosine switch regulates c-Abl. *Cell* **112**, 845–857 (2003).
64. Silverman, L. & Resh, M. D. Lysine residues form an integral component of a novel NH2-terminal membrane targeting motif for myristylated pp60v-src. *J. Cell Biol.* **119**, 415–425 (1992).
65. Abelson, H. T. & Rabstein, L. S. Lymphosarcoma: Virus-induced Thymic-independent Disease in Mice. 1–11 (1970).
66. Melo, J. V. & Barnes, D. J. Chronic myeloid leukaemia as a model of disease evolution in human cancer. *Nature reviews. Cancer* **7**, 441–453 (2007).
67. Nowell, P. C. & Hungerford, D. A. A minute chromosome in human chronic granulocytic leukemia. *Science* **142** (1960).
68. Rowley, J. D. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukemia identified by quinacrine fluorescence and Giemsa staining. *Nature* **243**, 290–293 (1973).
69. Deininger, M. W., Goldman, J. M. & Melo, J. V. The molecular biology of chronic myeloid leukemia. *Blood* **96**, 3343–3356 (2000).
70. Lydon, N. *Attacking cancer at its foundation. Nature medicine* **15**, 1153–1157 (2009).
71. Advani, A. S. & Pendergast, A. M. Bcr-Abl variants: biological and clinical aspects. *Leuk. Res.* **26**, 713–720 (2002).

72. McWhirter, J. R., Galasso, D. L. & Wang, J. Y. A coiled-coil oligomerization domain of Bcr is essential for the transforming function of Bcr-Abl oncoproteins. *Mol. Cell. Biol.* **13**, 7587–7595 (1993).
73. Koide, A., Bailey, C. W., Huang, X. & Koide, S. The fibronectin type III domain as a scaffold for novel binding proteins. *J. Mol. Biol.* **284**, 1141–1151 (1998).
74. Koide, A., Wojcik, J., Gilbreth, R. N., Hoey, R. J. & Koide, S. Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold. *J. Mol. Biol.* **415**, 393–405 (2012).
75. Sha, F., Salzman, G., Gupta, A. & Koide, S. Monobodies and other synthetic binding proteins for expanding protein science. *Protein Sci.* **26**, 910–924 (2017).
76. Wojcik, J. *et al.* A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain. *Nature structural & molecular biology* **17**, 519–527 (2010).
77. Grebien, F. *et al.* Targeting the SH2-kinase interface in Bcr-Abl inhibits leukemogenesis. *Cell* **147**, 306–319 (2011).
78. Wojcik, J. *et al.* Allosteric Inhibition of Bcr-Abl Kinase by High Affinity Monobody Inhibitors Directed to the Src Homology 2 (SH2)-Kinase Interface. *The Journal of biological chemistry* **291**, 8836–8847 (2016).
79. Serna, N. *et al.* Protein-Based Therapeutic Killing for Cancer Therapies. *Trends in biotechnology* **36**, 318–335 (2018).
80. Chapman, A. M. & McNaughton, B. R. Scratching the Surface: Resurfacing Proteins to Endow New Properties and Function. *Cell Chemical Biology* **23**, 543–553 (2016).
81. Bruce, V. J. & McNaughton, B. R. Inside Job: Methods for Delivering Proteins to the Interior of Mammalian Cells. *Cell Chemical Biology* **24**, 924–934 (2017).
82. Reichert, J. M., Rosensweig, C. J., Faden, L. B. & Dewitz, M. C. Monoclonal antibody successes in the clinic. *Nature biotechnology* **23**, 1073–1078 (2005).
83. El-Aneed, A. An overview of current delivery systems in cancer gene therapy. *Journal of controlled release : official journal of the Controlled Release Society* **94**, 1–14 (2004).
84. Mingozzi, F. & High, K. A. Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat. Rev. Genet.* **12**, 341–355 (2011).
85. Frankel, A. D. & Pabo, C. O. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **55**, 1189–1193 (1988).
86. Green, M. & Loewenstein, P. M. Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein. *Cell* **55**, 1179–1188 (1988).
87. Vives, E., Brodin, P. & Lebleu, B. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *The Journal of biological chemistry* **272**, 16010–16017 (1997).

88. Joliot, A., Pernelle, C., Deagostini-Bazin, H. & Prochiantz, A. Antennapedia homeobox peptide regulates neural morphogenesis. *Proceedings of the National Academy of Sciences* **88**, 1864–1868 (1991).
89. Rothbard, J. B. *et al.* Conjugation of arginine oligomers to cyclosporin A facilitates topical delivery and inhibition of inflammation. *Nat Med* **6**, 1253–1257 (2000).
90. Fuchs, S. M. & Raines, R. T. Pathway for polyarginine entry into mammalian cells. *Biochemistry* **43**, 2438–2444 (2004).
91. Poon, G. M. K. & Gariepy, J. Cell-surface proteoglycans as molecular portals for cationic peptide and polymer entry into cells. *Biochem. Soc. Trans.* **35**, 788–793 (2007).
92. Mai, J. C., Shen, H., Watkins, S. C., Cheng, T. & Robbins, P. D. Efficiency of protein transduction is cell type-dependent and is enhanced by dextran sulfate. *Journal of biological chemistry* **277**, 30208–30218 (2002).
93. Lindgren, M. & Langel, U. Classes and prediction of cell-penetrating peptides. *Methods in molecular biology* **683**, 3–19 (2011).
94. Koren, E. & Torchilin, V. P. Cell-penetrating peptides: breaking through to the other side. *Trends in molecular medicine* **18**, 385–393 (2012).
95. Henriques, S. T. & Castanho, M. A. R. B. Translocation or membrane disintegration? Implication of peptide-membrane interactions in pep-1 activity. *J. Pept. Sci.* **14**, 482–487 (2008).
96. Oehlke, J. *et al.* Cellular uptake of an alpha-helical amphipathic model peptide with the potential to deliver polar compounds into the cell interior non-endocytically. *Biochimica et biophysica acta* **1414**, 127–139 (1998).
97. Morris, M. C., Depollier, J., Mery, J., Heitz, F. & Divita, G. A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nature biotechnology* **19**, 1173–1176 (2001).
98. Pooga, M., Hällbrink, M., Zorko, M. & Langel, U. Cell penetration by transportan. *FASEB J.* **12**, 67–77 (1998).
99. Park, C. B., Yi, K. S., Matsuzaki, K., Kim, M. S. & Kim, S. C. Structure-activity analysis of buforin II, a histone H2A-derived antimicrobial peptide: the proline hinge is responsible for the cell-penetrating ability of buforin II. *Proceedings of the National Academy of Sciences* **97**, 8245–8250 (2000).
100. Kondo, E. *et al.* Tumour lineage-homing cell-penetrating peptides as anticancer molecular delivery systems. *Nature communications* **3**, 951–13 (2012).
101. Appelbaum, J. S. *et al.* Arginine Topology Controls Escape of Minimally Cationic Proteins from Early Endosomes to the Cytoplasm. *Chemistry & biology* **19**, 819–830 (2012).
102. Pack, D. W., Hoffman, A. S., Pun, S. & Stayton, P. S. Design and development of polymers for gene delivery. *Nature reviews. Drug discovery* **4**, 581–593 (2005).
103. Sonawane, N. D., Szoka, F. C. & Verkman, A. S. Chloride accumulation and swelling in endosomes enhances DNA transfer by polyamine-DNA polyplexes. *Journal of biological chemistry* **278**, 44826–44831 (2003).
104. Lin, C. & Engbersen, J. F. J. Effect of chemical functionalities in poly(amido amine)s for non-viral gene transfection. *Journal of controlled release : official journal of the Controlled Release Society* **132**, 267–272 (2008).

105. Varkouhi, A. K., Scholte, M., Storm, G. & Haisma, H. J. Endosomal escape pathways for delivery of biologicals. *Journal of controlled release* **151**, 220–228 (2011).
106. Pack, D., Putnam, D. & Langer, R. Design of Imidazole-Containing Endosomolytic Biopolymers for Gene Delivery. *Biotechnology and Bioengineering* 1–7 (1999).
107. Meyer, M., Philipp, A., Oskuee, R., Schmidt, C. & Wagner, E. Breathing life into polycations: functionalization with pH-responsive endosomolytic peptides and polyethylene glycol enables siRNA delivery. *Journal of the American Chemical Society* **130**, 3272–3273 (2008).
108. Wharton, S. A., Martin, S. R., Ruigrok, R. W., Skehel, J. J. & Wiley, D. C. Membrane fusion by peptide analogues of influenza virus haemagglutinin. *J. Gen. Virol.* **69**, 1847–1857 (1988).
109. Lönn, P. *et al.* Enhancing Endosomal Escape for Intracellular Delivery of Macromolecular Biologic Therapeutics. *Scientific reports* 1–9 (2016).
110. Kim, J.-S. *et al.* Endosomal acidic pH-induced conformational changes of a cytosol-penetrating antibody mediate endosomal escape. *Journal of controlled release* **235**, 165–175 (2016).
111. Akishiba, M. *et al.* Cytosolic antibody delivery by lipid-sensitive endosomolytic peptide. *Nature Chemistry* **9**, 751–761 (2017).
112. Piest, M. *et al.* Novel poly(amido amine)s with bio-reducible disulfide linkages in their diamino-units: structure effects and in vitro gene transfer properties. *Journal of controlled release* **130**, 38–45 (2008).
113. Gasparini, G. *et al.* Cellular uptake of substrate-initiated cell-penetrating poly(disulfide)s. *Journal of the American Chemical Society* **136**, 6069–6074 (2014).
114. Gasparini, G., Bang, E.-K., Montenegro, J. & Matile, S. Cellular uptake: lessons from supramolecular organic chemistry. *Chemical Communications* **51**, 10389–10402 (2015).
115. Fu, J., Yu, C., Li, L. & Yao, S. Q. Intracellular Delivery of Functional Proteins and Native Drugs by Cell-Penetrating Poly(disulfide)s. *Journal of the American Chemical Society* **137**, 12153–12160 (2015).
116. Gasparini, G. & Matile, S. Protein delivery with cell-penetrating poly(disulfide)s. *Chemical Communications* **51**, 17160–17162 (2015).
117. Falnes, P. O. & Sandvig, K. Penetration of protein toxins into cells. *Curr. Opin. Cell Biol.* **12**, 407–413 (2000).
118. Gordon, V. M. & Leppla, S. H. Proteolytic activation of bacterial toxins: role of bacterial and host cell proteases. *Infect. Immun.* **62**, 333–340 (1994).
119. Chiron, M. F., Fryling, C. M. & FitzGerald, D. J. Cleavage of pseudomonas exotoxin and diphtheria toxin by a furin-like enzyme prepared from beef liver. *Journal of biological chemistry* **269**, 18167–18176 (1994).
120. Molloy, S. S., Bresnahan, P. A., Leppla, S. H., Klimpel, K. R. & Thomas, G. Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen. *Journal of biological chemistry* **267**, 16396–16402 (1992).
121. Klimpel, K. R., Molloy, S. S., Thomas, G. & Leppla, S. H. Anthrax toxin protective antigen is activated by a cell surface protease with the sequence

- specificity and catalytic properties of furin. *Proceedings of the National Academy of Sciences* **89**, 10277–10281 (1992).
122. Sandvig, K. *et al.* Pathways followed by protein toxins into cells. *Int. J. Med. Microbiol.* **293**, 483–490 (2004).
  123. Lord, J. M., Smith, D. C. & Roberts, L. M. Toxin entry: how bacterial proteins get into mammalian cells. *Cell. Microbiol.* **1**, 85–91 (1999).
  124. Miller, C. J., Elliott, J. L. & Collier, R. J. Anthrax protective antigen: prepore-to-pore conversion. *Biochemistry* **38**, 10432–10441 (1999).
  125. Young, J. A. T. & Collier, R. J. Anthrax toxin: receptor binding, internalization, pore formation, and translocation. *Annu. Rev. Biochem.* **76**, 243–265 (2007).
  126. Scobie, H. M., Rainey, G. J. A., Bradley, K. A. & Young, J. A. T. Human capillary morphogenesis protein 2 functions as an anthrax toxin receptor. *Proceedings of the National Academy of Sciences* **100**, 5170–5174 (2003).
  127. Bradley, K. A., Mogridge, J., Mourez, M., Collier, R. J. & Young, J. A. Identification of the cellular receptor for anthrax toxin. *Nature* **414**, 225–229 (2001).
  128. Milne, J. C., Furlong, D., Hanna, P. C., Wall, J. S. & Collier, R. J. Anthrax protective antigen forms oligomers during intoxication of mammalian cells. *Journal of biological chemistry* **269**, 20607–20612 (1994).
  129. Petosa, C., Collier, R. J., Klimpel, K. R., Leppla, S. H. & Liddington, R. C. Crystal structure of the anthrax toxin protective antigen. *Nature* **385**, 833–838 (1997).
  130. Elliott, J. L., Mogridge, J. & Collier, R. J. A quantitative study of the interactions of *Bacillus anthracis* edema factor and lethal factor with activated protective antigen. *Biochemistry* **39**, 6706–6713 (2000).
  131. Abrami, L., Liu, S., Cosson, P., Leppla, S. H. & van der Goot, F. G. Anthrax toxin triggers endocytosis of its receptor via a lipid raft-mediated clathrin-dependent process. *J. Cell Biol.* **160**, 321–328 (2003).
  132. Arora, N. & Leppla, S. H. Residues 1-254 of anthrax toxin lethal factor are sufficient to cause cellular uptake of fused polypeptides. *Journal of biological chemistry* **268**, 3334–3341 (1993).
  133. Ballard, J. D., Collier, R. J. & Starnbach, M. N. Anthrax toxin-mediated delivery of a cytotoxic T-cell epitope in vivo. *Proceedings of the National Academy of Sciences* **93**, 12531–12534 (1996).
  134. Hobson, J. P., Liu, S., Rønø, B., Leppla, S. H. & Bugge, T. H. Imaging specific cell-surface proteolytic activity in single living cells. *Nature methods* **3**, 259–261 (2006).
  135. Cordero, C. L., Kudryashov, D. S., Reisler, E. & Satchell, K. J. F. The Actin cross-linking domain of the *Vibrio cholerae* RTX toxin directly catalyzes the covalent cross-linking of actin. *Journal of biological chemistry* **281**, 32366–32374 (2006).
  136. Zahaf, N. I. *et al.* Targeted delivery of an ADP- ribosylating bacterial toxin into cancer cells. *Scientific reports* 1–10 (2017). doi:10.1038/srep41252
  137. Liao, X., Rabideau, A. E. & Pentelute, B. L. Delivery of Antibody Mimics into Mammalian Cells via Anthrax Toxin Protective Antigen. *Chembiochem : a European journal of chemical biology* **15**, 2458–2466 (2014).

138. Feld, G. K. *et al.* Structural basis for the unfolding of anthrax lethal factor by protective antigen oligomers. *Nature structural & molecular biology* **17**, 1383–1390 (2010).
139. Thoren, K. L. & Krantz, B. A. The unfolding story of anthrax toxin translocation. *Mol. Microbiol.* **80**, 588–595 (2011).
140. Verdurmen, W. P., Luginbuhl, M., Honegger, A. & Pluckthun, A. Efficient cell-specific uptake of binding proteins into the cytoplasm through engineered modular transport systems. *Journal of controlled release : official journal of the Controlled Release Society* **200C**, 13–22 (2014).
141. Kounnas, M. Z. *et al.* The alpha 2-macroglobulin receptor/low density lipoprotein receptor-related protein binds and internalizes Pseudomonas exotoxin A. *Journal of biological chemistry* **267**, 12420–12423 (1992).
142. Jinno, Y. *et al.* Domain II mutants of Pseudomonas exotoxin deficient in translocation. *Journal of biological chemistry* **264**, 15953–15959 (1989).
143. Siegall, C. B., Chaudhary, V. K., FitzGerald, D. J. & Pastan, I. Functional analysis of domains II, Ib, and III of Pseudomonas exotoxin. *Journal of biological chemistry* **264**, 14256–14261 (1989).
144. Michalska, M. & Wolf, P. Pseudomonas Exotoxin A: optimized by evolution for effective killing. *Front Microbiol* **6**, 963 (2015).
145. McKee, M. L. & FitzGerald, D. J. Reduction of furin-nicked Pseudomonas exotoxin A: an unfolding story. *Biochemistry* **38**, 16507–16513 (1999).
146. Jackson, M. E. *et al.* The KDEL retrieval system is exploited by Pseudomonas exotoxin A, but not by Shiga-like toxin-1, during retrograde transport from the Golgi complex to the endoplasmic reticulum. *Journal of Cell Science* **112** ( Pt 4), 467–475 (1999).
147. White, J. *et al.* Rab6 coordinates a novel Golgi to ER retrograde transport pathway in live cells. *J. Cell Biol.* **147**, 743–760 (1999).
148. Smith, D. C. *et al.* Internalized Pseudomonas exotoxin A can exploit multiple pathways to reach the endoplasmic reticulum. *Traffic* **7**, 379–393 (2006).
149. Theuer, C. P., Buchner, J., FitzGerald, D. & Pastan, I. The N-terminal region of the 37-kDa translocated fragment of Pseudomonas exotoxin A aborts translocation by promoting its own export after microsomal membrane insertion. *Proceedings of the National Academy of Sciences* **90**, 7774–7778 (1993).
150. Hazes, B. & Read, R. J. Accumulating evidence suggests that several AB-toxins subvert the endoplasmic reticulum-associated protein degradation pathway to enter target cells. *Biochemistry* **36**, 11051–11054 (1997).
151. Koopmann, J. O. *et al.* Export of antigenic peptides from the endoplasmic reticulum intersects with retrograde protein translocation through the Sec61p channel. *Immunity* **13**, 117–127 (2000).
152. Wolf, P. & Elsässer-Beile, U. Pseudomonas exotoxin A: from virulence factor to anti-cancer agent. *Int. J. Med. Microbiol.* **299**, 161–176 (2009).
153. Akbari, B. *et al.* Immunotoxins in cancer therapy: Review and update. *Int. Rev. Immunol.* **36**, 207–219 (2017).
154. Mohammed, A. F. *et al.* The Pseudomonas aeruginosa exotoxin A translocation domain facilitates the routing of CPP-protein cargos to the

- cytosol of eukaryotic cells. *Journal of controlled release : official journal of the Controlled Release Society* **164**, 58–64 (2012).
155. Kim, H.-Y. *et al.* Intracellular Protein Delivery System Using a Target-Specific Repebody and Translocation Domain of Bacterial Exotoxin. *ACS Chem. Biol.* **12**, 2891–2897 (2017).
  156. Collier, R. J. & Kandel, J. Structure and activity of diphtheria toxin. I. Thiol-dependent dissociation of a fraction of toxin into enzymically active and inactive fragments. *Journal of biological chemistry* **246**, 1496–1503 (1971).
  157. Gill, D. M. & Pappenheimer, A. M. Structure-activity relationships in diphtheria toxin. *Journal of biological chemistry* **246**, 1492–1495 (1971).
  158. Choe, S. *et al.* The crystal structure of diphtheria toxin. *Nature* **357**, 216–222 (1992).
  159. Naglich, J. G., Metherall, J. E., Russell, D. W. & Eidels, L. Expression cloning of a diphtheria toxin receptor: identity with a heparin-binding EGF-like growth factor precursor. *Cell* **69**, 1051–1061 (1992).
  160. Donovan, J. J., Simon, M. I., Draper, R. K. & Montal, M. Diphtheria toxin forms transmembrane channels in planar lipid bilayers. *Proceedings of the National Academy of Sciences* **78**, 172–176 (1981).
  161. Kagan, B. L., Finkelstein, A. & Colombini, M. Diphtheria toxin fragment forms large pores in phospholipid bilayer membranes. *Proceedings of the National Academy of Sciences* **78**, 4950–4954 (1981).
  162. Oh, K. J., Senzel, L., Collier, R. J. & Finkelstein, A. Translocation of the catalytic domain of diphtheria toxin across planar phospholipid bilayers by its own T domain. *Proceedings of the National Academy of Sciences* **96**, 8467–8470 (1999).
  163. Ratts, R. *et al.* A conserved motif in transmembrane helix 1 of diphtheria toxin mediates catalytic domain delivery to the cytosol. *Proceedings of the National Academy of Sciences* **102**, 15635–15640 (2005).
  164. Auger, A. *et al.* Efficient Delivery of Structurally Diverse Protein Cargo into Mammalian Cells by a Bacterial Toxin. *Mol. Pharm.* **12**, 2962–2971 (2015).
  165. Beilhartz, G. L., Sugiman-Marangos, S. N. & Melnyk, R. A. Repurposing bacterial toxins for intracellular delivery of therapeutic proteins. *Biochemical Pharmacology* 1–8 (2017). doi:10.1016/j.bcp.2017.04.009
  166. Dunn, K. W., Mayor, S., Myers, J. N. & Maxfield, F. R. Applications of ratio fluorescence microscopy in the study of cell physiology. *FASEB J.* **8**, 573–582 (1994).
  167. Holden, P. & Horton, W. A. Crude subcellular fractionation of cultured mammalian cell lines. *BMC Res Notes* **2**, 243 (2009).
  168. Verdurmen, W. P. R., Mazlami, M. & Plückthun, A. A quantitative comparison of cytosolic delivery via different protein uptake systems. *Scientific reports* 1–13 (2017). doi:10.1038/s41598-017-13469-y
  169. Shiga, K. Ueber den erregere der dysenterie in Japan. Vorläufige mitteilung. *Zentralbl. Bakteriол. Microbiol. Hyg* 599–600 (1936).
  170. O'Brien, A. D. *et al.* Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* **226**, 694–696 (1984).

171. Jackson, M. P., Newland, J. W., Holmes, R. K. & O'Brien, A. D. Nucleotide sequence analysis of the structural genes for Shiga-like toxin I encoded by bacteriophage 933J from *Escherichia coli*. *Microb. Pathog.* **2**, 147–153 (1987).
172. Johannes, L. & Römer, W. Shiga toxins--from cell biology to biomedical applications. *Nat. Rev. Microbiol.* **8**, 105–116 (2010).
173. Konowalchuk, J., Speirs, J. I. & Stavric, S. Vero response to a cytotoxin of *Escherichia coli*. *Infect. Immun.* **18**, 775–779 (1977).
174. Karmali, M. A., Steele, B. T., Petric, M. & Lim, C. Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *Lancet* **1**, 619–620 (1983).
175. Tarr, P. I., Gordon, C. A. & Chandler, W. L. Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. *Lancet* **365**, 1073–1086 (2005).
176. Nguyen, Y. & Sperandio, V. Enterohemorrhagic *E. coli* (EHEC) pathogenesis. *Front Cell Infect Microbiol* **2**, 90 (2012).
177. Fraser, M. E. *et al.* Structure of shiga toxin type 2 (Stx2) from *Escherichia coli* O157:H7. *Journal of biological chemistry* **279**, 27511–27517 (2004).
178. Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. Crystal structure of the cell-binding B oligomer of verotoxin-1 from *E. coli*. *Nature* **355**, 748–750 (1992).
179. Lindberg, A. A. *et al.* Identification of the carbohydrate receptor for Shiga toxin produced by *Shigella dysenteriae* type 1. *Journal of biological chemistry* **262**, 1779–1785 (1987).
180. Jacewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A. & Keusch, G. T. Pathogenesis of shigella diarrhea. XI. Isolation of a shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *J. Exp. Med.* **163**, 1391–1404 (1986).
181. Waddell, T., Cohen, A. & Lingwood, C. A. Induction of verotoxin sensitivity in receptor-deficient cell lines using the receptor glycolipid globotriosylceramide. *Proceedings of the National Academy of Sciences* **87**, 7898–7901 (1990).
182. Ling, H. *et al.* Structure of the shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry* **37**, 1777–1788 (1998).
183. Sandvig, K., Bergan, J., Dyve, A.-B., Skotland, T. & Torgersen, M. L. Endocytosis and retrograde transport of Shiga toxin. *Toxicon* **56**, 1181–1185 (2010).
184. Johannes, L. & Popoff, V. Tracing the retrograde route in protein trafficking. *Cell* **135**, 1175–1187 (2008).
185. Girod, A. *et al.* Evidence for a COP-I-independent transport route from the Golgi complex to the endoplasmic reticulum. *Nat. Cell Biol.* **1**, 423–430 (1999).
186. Lea, N., Lord, J. M. & Roberts, L. M. Proteolytic cleavage of the A subunit is essential for maximal cytotoxicity of *Escherichia coli* O157:H7 Shiga-like toxin-1. *Microbiology (Reading, Engl.)* **145**, 999–1004 (1999).
187. Garred, O., van Deurs, B. & Sandvig, K. Furin-induced cleavage and activation of Shiga toxin. *Journal of biological chemistry* **270**, 10817–10821 (1995).

188. Tam, P. J. & Lingwood, C. A. Membrane cytosolic translocation of verotoxin A1 subunit in target cells. *Microbiology (Reading, Engl.)* **153**, 2700–2710 (2007).
189. Yu, M. & Haslam, D. B. Shiga toxin is transported from the endoplasmic reticulum following interaction with the luminal chaperone HEDJ/ERdj3. *Infect. Immun.* **73**, 2524–2532 (2005).
190. Haicheur, N. *et al.* The B subunit of Shiga toxin coupled to full-size antigenic protein elicits humoral and cell-mediated immune responses associated with a Th1-dominant polarization. *Int. Immunol.* **15**, 1161–1171 (2003).
191. Lee, R. S. *et al.* Major histocompatibility complex class I presentation of exogenous soluble tumor antigen fused to the B-fragment of Shiga toxin. *Eur. J. Immunol.* **28**, 2726–2737 (1998).
192. Vingert, B. *et al.* The Shiga toxin B-subunit targets antigen in vivo to dendritic cells and elicits anti-tumor immunity. *Eur. J. Immunol.* **36**, 1124–1135 (2006).
193. Adotevi, O. *et al.* B subunit of Shiga toxin-based vaccines synergize with alpha-galactosylceramide to break tolerance against self antigen and elicit antiviral immunity. *J. Immunol.* **179**, 3371–3379 (2007).
194. Janssen, K.-P. *et al.* In vivo tumor targeting using a novel intestinal pathogen-based delivery approach. *Cancer Res.* **66**, 7230–7236 (2006).
195. Alaoui, El, A. *et al.* Shiga toxin-mediated retrograde delivery of a topoisomerase I inhibitor prodrug. *Angew. Chem. Int. Ed. Engl.* **46**, 6469–6472 (2007).
196. Tarragó-Trani, M. T., Jiang, S., Harich, K. C. & Storrie, B. Shiga-like toxin subunit B (SLTB)-enhanced delivery of chlorin e6 (Ce6) improves cell killing. *Photochem. Photobiol.* **82**, 527–537 (2006).
197. Amessou, M. *et al.* Retrograde delivery of photosensitizer (TPPp-O-beta-GluOH)<sub>3</sub> selectively potentiates its photodynamic activity. *Bioconjugate chemistry* **19**, 532–538 (2008).
198. Ryou, J. H., Sohn, Y. K., Hwang, D. E. & Kim, H. S. Shiga-like toxin-based high-efficiency and receptor-specific intracellular delivery system for a protein. *Biochemical and biophysical research communications* (2015).
199. Ryou, J. H. & Kim, H. S. Engineering of bacterial exotoxins for highly efficient and receptor-specific intracellular delivery of diverse cargosEngineering of bacterial exotoxins for highly efficient and receptor-specific intracellular delivery of diverse cargos. *Biotechnology and Bioengineering* 1–8 (2016).
200. Fulcher, L. J. *et al.* An affinity-directed protein missile system for targeted proteolysis. *Open Biol.* **6**, 160255 (2016).
201. Fulcher, L. J., Hutchinson, L. D., Macartney, T. J., Turnbull, C. & Sapkota, G. P. Targeting endogenous proteins for degradation through the affinity-directed protein missile system. *Open Biol.* **7**, 170066–11 (2017).
202. Kamura, T. *et al.* VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev.* **18**, 3055–3065 (2004).

203. Buckley, D. L. *et al.* Small-molecule inhibitors of the interaction between the E3 ligase VHL and HIF1 $\alpha$ . *Angew. Chem. Int. Ed. Engl.* **51**, 11463–11467 (2012).
204. Lai, A. C. & Crews, C. M. Induced protein degradation: an emerging drug discovery paradigm. *Nature Publishing Group* **16**, 101–114 (2017).
205. Burslem, G. M. *et al.* The Advantages of Targeted Protein Degradation Over Inhibition: An RTK Case Study. *Cell Chemical Biology* **25**, 67–77.e3 (2018).
206. Clift, D. *et al.* A Method for the Acute and Rapid Degradation of Endogenous Proteins. *Cell* 1–34 (2017).
207. James, L. C. Intracellular antibody immunity and the cytosolic Fc receptor TRIM21. *Curr. Top. Microbiol. Immunol.* **382**, 51–66 (2014).
208. Deshaies, R. J. Protein degradation: Prime time for PROTACs. *Nature chemical biology* **11**, 634–635 (2015).
209. Kraskouskaya, D., Duodu, E., Arpin, C. C. & Gunning, P. T. Progress towards the development of SH2 domain inhibitors. *Chemical Society reviews* **42**, 3337–3370 (2013).
210. Kükenshöner, T. *et al.* Selective Targeting of SH2 Domain-Phosphotyrosine Interactions of Src Family Tyrosine Kinases with Monobodies. *J. Mol. Biol.* **429**, 1364–1380 (2017).
211. Burckstummer, T. *et al.* An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nature methods* **3**, 1013–1019 (2006).
212. Sha, F. *et al.* Dissection of the BCR-ABL signaling network using highly specific monoclonal inhibitors to the SHP2 SH2 domains. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 14924–14929 (2013).
213. Chaudhary, V. K., Jinno, Y., FitzGerald, D. & Pastan, I. Pseudomonas exotoxin contains a specific sequence at the carboxyl terminus that is required for cytotoxicity. *Proceedings of the National Academy of Sciences* **87**, 308–312 (1990).
214. Ku, M. *et al.* Src family kinases and their role in hematological malignancies. *Leuk. Lymphoma* **56**, 577–586 (2015).
215. Gioia, R. *et al.* Quantitative phosphoproteomics revealed interplay between Syk and Lyn in the resistance to nilotinib in chronic myeloid leukemia cells. *Blood* **118**, 2211–2221 (2011).
216. Samanta, A. K. *et al.* Jak2 inhibition deactivates Lyn kinase through the SET-PP2A-SHP1 pathway, causing apoptosis in drug-resistant cells from chronic myelogenous leukemia patients. **28**, 1669–1681 (2009).
217. Gamas, P. *et al.* Inhibition of imatinib-mediated apoptosis by the caspase-cleaved form of the tyrosine kinase Lyn in chronic myelogenous leukemia cells. **23**, 1500–1506 (2009).
218. Wilson, M. B., Schreiner, S. J., Choi, H.-J., Kamens, J. & Smithgall, T. E. Selective pyrrolo-pyrimidine inhibitors reveal a necessary role for Src family kinases in Bcr-Abl signal transduction and oncogenesis. *Oncogene* **21**, 8075–8088 (2002).
219. Ptasznik, A., Nakata, Y., Kalota, A., Emerson, S. G. & Gewirtz, A. M. Short interfering RNA (siRNA) targeting the Lyn kinase induces apoptosis in

- primary, and drug-resistant, BCR-ABL1(+) leukemia cells. *Nat Med* **10**, 1187–1189 (2004).
220. Salmond, R. J., Filby, A., Qureshi, I., Caserta, S. & Zamoyska, R. T-cell receptor proximal signaling via the Src-family kinases, Lck and Fyn, influences T-cell activation, differentiation, and tolerance. *Immunol. Rev.* **228**, 9–22 (2009).
  221. Samelson, L. E., Phillips, A. F., Luong, E. T. & Klausner, R. D. Association of the fyn protein-tyrosine kinase with the T-cell antigen receptor. *Proceedings of the National Academy of Sciences* **87**, 4358–4362 (1990).
  222. Lin, H. *et al.* Association of p59(fyn) with the T lymphocyte costimulatory receptor CD2. Binding of the Fyn Src homology (SH) 3 domain is regulated by the Fyn SH2 domain. *Journal of biological chemistry* **273**, 19914–19921 (1998).
  223. Groves, T. *et al.* Fyn can partially substitute for Lck in T lymphocyte development. *Immunity* **5**, 417–428 (1996).
  224. Denny, M. F., Patai, B. & Straus, D. B. Differential T-cell antigen receptor signaling mediated by the Src family kinases Lck and Fyn. *Mol. Cell. Biol.* **20**, 1426–1435 (2000).
  225. Chau, L. A. & Madrenas, J. Phospho-LAT-independent activation of the ras-mitogen-activated protein kinase pathway: a differential recruitment model of TCR partial agonist signaling. *J. Immunol.* **163**, 1853–1858 (1999).
  226. Graves, L. M., Duncan, J. S., Whittle, M. C. & Johnson, G. L. The dynamic nature of the kinome. *Biochem. J.* **450**, 1–8 (2013).
  227. Schneekloth, J. S. *et al.* Chemical genetic control of protein levels: selective in vivo targeted degradation. *Journal of the American Chemical Society* **126**, 3748–3754 (2004).
  228. Bondeson, D. P. *et al.* Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nature chemical biology* **11**, 611–617 (2015).
  229. Raucher, D. & Ryu, J. S. Cell-penetrating peptides: strategies for anticancer treatment. *Trends in molecular medicine* **21**, 560–570 (2015).
  230. LeCher, J. C., Nowak, S. J. & McMurry, J. L. Breaking in and busting out: cell-penetrating peptides and the endosomal escape problem. *Biomolecular Concepts* **8**, 84–19 (2017).
  231. El-Sayed, A., Futaki, S. & Harashima, H. Delivery of macromolecules using arginine-rich cell-penetrating peptides: ways to overcome endosomal entrapment. *AAPS J* **11**, 13–22 (2009).
  232. Bonifacino, J. S. & Rojas, R. Retrograde transport from endosomes to the trans-Golgi network. *Nature reviews. Molecular cell biology* **7**, 568–579 (2006).
  233. Cavazza, T. & Vernos, I. The RanGTP Pathway: From Nucleo-Cytoplasmic Transport to Spindle Assembly and Beyond. *Front Cell Dev Biol* **3**, 82 (2015).
  234. Shen, X., Zhang, K. & Kaufman, R. J. The unfolded protein response--a stress signaling pathway of the endoplasmic reticulum. *J. Chem. Neuroanat.* **28**, 79–92 (2004).
  235. Sitia, R. & Braakman, I. Quality control in the endoplasmic reticulum protein factory. *Nature* **426**, 891–894 (2003).

236. Nakagawa, T. *et al.* Caspase-12 mediates endoplasmic-reticulum-specific apoptosis and cytotoxicity by amyloid-beta. *Nature* **403**, 98–103 (2000).
237. Cudna, R. E. & Dickson, A. J. Endoplasmic reticulum signaling as a determinant of recombinant protein expression. *Biotechnology and Bioengineering* **81**, 56–65 (2003).
238. Batori, V., Koide, A. & Koide, S. Exploring the potential of the monobody scaffold: effects of loop elongation on the stability of a fibronectin type III domain. *Protein Eng.* **15**, 1015–1020 (2002).
239. Ren, R. Mechanisms of BCR-ABL in the pathogenesis of chronic myelogenous leukaemia. *Nature reviews. Cancer* **5**, 172–183 (2005).
240. Hoelbl, A. *et al.* Stat5 is indispensable for the maintenance of bcr/abl-positive leukaemia. *EMBO Mol Med* **2**, 98–110 (2010).
241. Nieborowska-Skorska, M. *et al.* Signal transducer and activator of transcription (STAT)5 activation by BCR/ABL is dependent on intact Src homology (SH)3 and SH2 domains of BCR/ABL and is required for leukemogenesis. *J. Exp. Med.* **189**, 1229–1242 (1999).
242. Samir, T. & Wiels, J. Intracellular Signaling Events in CD77-Mediated Apoptosis of Burkitt's Lymphoma Cells. *Blood* 1–12 (1997).
243. Geyer, P. E. *et al.* Gastric Adenocarcinomas Express the Glycosphingolipid Gb3/CD77: Targeting of Gastric Cancer Cells with Shiga Toxin B-Subunit. *Molecular Cancer Therapeutics* **15**, 1008–1017 (2016).
244. Kovbasnjuk, O. *et al.* The glycosphingolipid globotriaosylceramide in the metastatic transformation of colon cancer. *Proceedings of the National Academy of Sciences* **102**, 19087–19092 (2005).



# Curriculum Vitae

## Nadine Eliane Schmit

**Address** Chemin des Ramiers 16, 1022 Chavannes-près-Renens  
**Phone** +41 788 19 65 16  
**Email** nadeschmit@gmail.com  
**LinkedIn** <https://www.linkedin.com/in/nadine-schmit/>  
**Personal** Luxembourgish citizen, Swiss residence permit B, 30 years old

- Strong background in Cancer, Immunology and Protein Biochemistry
- Excellent communication and interpersonal skills
- Fluent in English, French and German

## PROFESSIONAL EXPERIENCE

---

<b>Swiss Institute for Experimental Cancer Research (ISREC), EPFL</b> <i>PhD thesis:</i> "Cytosolic delivery and characterization of monobodies interfering with SH2 domain-phosphotyrosine interactions" - Development of a new method to deliver protein therapeutics into cells, characterization of protein binders inhibiting kinase function, teaching and student supervision, management of collaborations and regular presentations	<b>2014 - present</b>
<b>Novartis Institutes for Biomedical Research (NIBR) in Basel</b> <i>Post-graduate trainee at the Autoimmunity, Transplantation and Inflammation (ATI) department.</i> Contributed to the development of novel protein-based therapeutics to treat autoimmune disease	<b>2014</b>
<b>Centre for Chronic Immunodeficiency in Freiburg, Germany</b> <i>Diploma thesis:</i> Discovered an implication of the CARD11 adaptor protein in immunodeficiency by analyzing patient samples	<b>2012-2013</b>
<b>LERU Bright Student Conference</b> Organizer: "Energy transition in the 21 <sup>st</sup> century", Freiburg, Germany	<b>2013</b>
<b>Centre for Chronic Immunodeficiency in Freiburg, Germany</b> <i>Undergraduate Research Assistant:</i> Studied innate immune signaling pathways and macrophage differentiation in response to bacterial challenges in mouse models	<b>2012</b>
<b>Uppsala Biomedical Centre, Uppsala, Sweden</b> <i>Research Internship:</i> Contributed to the development of therapeutic vaccines for allergy treatment	<b>2010</b>
<b>Hematology unit at Laboratoire national de Santé, Luxembourg</b> <i>Research Internship:</i> Analyzed patient samples for hematological disease	<b>2009</b>

## EDUCATION

---

<b>Ecole polytechnique fédérale de Lausanne (EPFL)</b> PhD in Molecular Life Sciences Courses in Project Management, Scientific writing and Image processing	<b>2014 - present</b>
<b>University of Freiburg, Germany</b> Diploma in Biology, Majors: Immunology, Cell Biology and Animal Physiology (grade: 1,3 "very good")	<b>2007 - 2013</b>

University of Uppsala, Sweden  
ERASMUS exchange year

2010-2011

Lycée Classique d' Echternach, Luxembourg  
Secondary School diploma, Natural Science and Mathematics Section

2007

## PROFESSIONAL SKILLS

---

- Cellular and molecular biology techniques
- Protein purification and biochemistry techniques
- Confocal microscopy and Image processing and analysis using Image J
- IT: ImageJ Macro language, basic R programming, basic Python, GraphPad Prism, Adobe Illustrator, PubMed (database research), MS Office (Word, Excel, Powerpoint)
- Design and execution of scientific experiments, management of research projects
- Teamwork with internal and external international collaborators
- Writing of scientific reports, articles and experimental protocols
- Effective communication with scientists, healthcare professionals and lay audiences
- Teaching and supervision of interns and students in English, French and German
- Establishment and methodical testing of new and technically challenging protocols

## LANGUAGES

---

Luxembourgish (native), German (C2), French (C2), English (C2), Swedish (B2), Spanish (A1)

## EXTRACURRICULAR ACTIVITIES

---

Sponsoring and climbing responsible at **Club Montagne, EPFL** 2014 - present

Organization of the **PhD Hiring days** of the Life Science Doctoral Schools, EPFL 2014 - 2018

Member of the **Choir "Laudate"**, Lausanne 2014 - present

## CONFERENCE PRESENTATIONS

---

International Symposium in Chemical Biology, Geneva, CH, *Speaker* January 2018

FASEB Science Research Conference "Protein Kinase Signaling Network Regulation", Snowmass, CO, USA, *poster presentation* July 2016

FEBS Advanced Lecture Course "Molecular Mechanisms in Signal Transduction and Cancer", August 16th to 24th, 2015, Spetses, Greece, *poster presentation* August 2015

## PUBLICATIONS

---

Kükenshöner T.\*, **Schmit N. E.\***, Sha F., Pojer F., Koide A., Seeliger M., Koide S., and Hantschel O., Selective Targeting of SH2 Domain-Phosphotyrosine Interactions of Src Family Tyrosine Kinases with Monobodies. JMB, 2017.

Rakhmanov M., Sic H., Kienzler A.K., Fischer B., Rizzi M., Seidl M., Melkaoui K., Unger S., Moehle L., **Schmit N. E.**, Deshmukh S.D., Ayata C. K., Schuh W., Zhang Z., Cosset F. L., Verhoeven E., Peter H.H., Voll R.E., Salzer U., Eibel H., Warnatz K., High Levels of SOX5 Decrease Proliferative Capacity of Human B Cells, but Permit Plasmablast Differentiation., PLOS One, 2014.

