

Photoswitchable Peptides as Tools for Light-controlled Modulation of Epigenetic Protein Complexes

Kumulative Dissertationsschrift

Zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Chemie der Philipps-Universität Marburg
(Hochschulkenziffer: 1180)
vorgelegt von

Lea Albert, M. Sc.

aus Ranstadt

Marburg an der Lahn, 2020

Die vorliegende Arbeit wurde in der Zeit von Februar 2016 bis November 2020 am Fachbereich Chemie der Philipps-Universität Marburg unter der Leitung von Frau Jun.-Prof. Dr. Olalla Vázquez angefertigt.

Vom Fachbereich Chemie der Philipps-Universität Marburg als Dissertationsarbeit angenommen am:

11.01.2021

Erstgutachterin: Frau Jun.-Prof. Dr. Olalla Vázquez

Zweitgutachter: Herr Prof. Dr. Alexander Brehm

Tag der mündlichen Prüfung: 15.01.2021

Für meine Familie

“Climb the mountain not to plant your flag, but to embrace the challenge. Climb it not so the world can see you, but so you can see the world.”

David McCullough Jr.

Declaration

Ich erkläre, dass eine Promotion an noch keiner anderen Hochschule als der Philipps-Universität Marburg, Fachbereich Chemie, versucht wurde.

Ich versichere, dass ich meine vorgelegte Dissertation „**Photoswitchable Peptides as Tools for Light-controlled Modulation of Epigenetic Protein Complexes**“ selbst und ohne fremde Hilfe verfasst, keine anderen als die in den angegebenen Quellen oder Hilfsmittel benutzt, alle vollständig oder sinngemäß übernommenen Zitate als solche gekennzeichnet, sowie die Dissertation in der vorliegenden oder einer ähnlichen Form noch bei keiner anderen in- oder ausländischen Hochschule anlässlich eines Promotionsgesuchs oder zu anderen Prüfungszwecken eingereicht habe.

Marburg, den

Unterschrift: _____

Lea Albert

Acknowledgements

Many people have helped and supported me during the time of my thesis and, at this point, I would like to thank ...

... Jun. Prof. Dr. Olalla Vázquez, who has been a great supervisor, not only for the possibility of working on such an interesting project but especially for the outstanding encouragement, motivation and inspiration you gave me during this time and for all the scientific discussions and support to bring my project forward. I, in particular, enjoyed the open and honest working atmosphere, appreciated your commitment to your students, the teaching of not only scientific knowledge but also team-spirit and I will always be thankful for your engagement in my life and for what we have shared in the last years.

... Prof. Dr. Alexander Brehm for being the second corrector of this thesis. It was great to work together on different collaboration projects as well as enjoying time in the TRR81 retreat 2019.

... Prof. Dr. Eric Meggers, for being part of the thesis committee, as a third referee and for reviewing several of our manuscripts before publication, giving valuable feedback and suggestions.

... Prof. Dr. Lars-Oliver Essen, for being part of the thesis committee, as a fourth referee and for the great collaborative work during my thesis.

... Dr. Uwe Linne and his team of the mass department for measuring various of my samples and always having an open ear for all kinds of questions, as well as Dr. Xiulan and her team of the NMR department for measuring some of my NMR probes and the great support. Also to Norbert Frommknecht and his team of the Feinmechanik for constructing custom made LED-array plates for us.

... Bene, Greta, Lei for the great time. It was a great pleasure to have you as my lab colleagues and friends during the last years. I could not have wished for better people to share my PhD time with. I especially appreciated our great team spirit and all your support in hard times as well as unforgettable funny moments, that brought us closer together. Also, thanks to Bene and Greta for proof-reading my thesis. As well, a big thanks to all other former and present members of the Vázquez group for the nice time and working atmosphere, as well as to all my internship, Bachelor- and Master students for their contribution to my work, especially Malte Hoffarth and Alberto Peñalver.

... Sabrina, Nathalie and Sabine for introducing me into basic biochemical working routines, for discussing the projects and the nice joint group retreats. Also to Frank for his help and support in discussing several aspects of my project.

... Dr. Ying Wang and Prema Erramsetti (Zentrum für Tumor- und Immunbiologie, Marburg) for introducing me into the working routine with human suspension cell lines.

... Ina and Andrea for the organizational help during this time.

... Prof. Dr. Yali Dou, our collaboration partner, for the possibility to spend 4 months in her lab at the University of Michigan, as well as all lab members, especially Jing Xu, for everything I learned there.

... all other collaborators, for contributing to my work, all the fruitful discussions about different projects and the nice collaborative atmosphere: Dr. Wieland Steinchen, Laura Werel, Prof. Dr. Lars-Oliver Essen, Dr. Jatin Nagpal, Prof. Dr. Ryu Soojin, Nemanja Djokovic, Dusan Ruzik, Prof. Dr. Katarina Nolic, Simone Prinz, Janina Schwenty-Lara, the members of Prof. Brehm lab especially Jonathan Lenz, Dzung Nguyen (AG Klebe), and Prof. Philipsen (student: Anna Hsieh).

... all my friends for the life and great activities besides the lab during the last years.

... meiner gesamten Familie, meinen Großeltern Anita, Rudi, Helga sowie auch den leider bereits verstorbenen Horst und Lindi. Ich weiß, wie sehr ihr euch mit mir gefreut hättet, den Moment des Doktorabschlusses zu feiern; außerdem natürlich den weltbesten Geschwistern Felix und Franzi und ein ganz besonderer Dank gilt natürlich meinen Eltern, Doris und Alexander, für eure bedingungslose Unterstützung und Liebe. Ich weiß, dass ich in jeder Lebenslage immer auf euch zählen kann. Für all eure Hilfe und aufbauenden Worte und die schönen Zeiten, die wir gemeinsam als Familie verbracht haben bin ich unglaublich froh und dankbar.

... last but most definitely not least: my husband Ahmed, for all your love and support throughout the years of our studies. For making my life so much more joyful and richer in any possible aspect, for being there for me, when I need you, for building me up, when I was struggling, and for everything beyond. And not to forget, for giving me the most valuable gift I could ever imagine. I am looking forward to our future together as a family.

Table of Contents

Abbreviations.....	III
General Remarks on Thesis Structure and Nomenclature	VII
1 Introduction	1
1.1 Epigenetics	1
1.2 Post-Translational Modifications of Histones	2
1.3 Mixed Lineage Leukemia 1.....	9
1.4 Mixed Lineage Leukemia 1 Complex.....	11
1.5 Inhibitors of the MLL1-WDR5 Interaction	15
1.6 Light as a Trigger to Modulate Biological Functions	21
1.7 References.....	22
2 Aim.....	32
3 Summaries	35
3.1 Summary in English	35
3.2 Zusammenfassung in Deutscher Sprache.....	38
4 Cumulative Part	41
4.1 Photoswitchable peptides for spatiotemporal control of biological functions	41
4.1.1 Author Contribution	64
4.2 Controlled inhibition of methyltransferases using photoswitchable peptidomimetics: towards an epigenetic regulation of leukemia.....	65
4.2.1 Author Contribution	133
4.3 Modulating Protein–Protein Interactions with Visible-Light-Responsive Peptide Backbone Photoswitches.....	134
4.3.1 Author Contribution	189
4.4 In Vivo Optochemical Control of Hematopoiesis by MLL1 Complex Disruption.....	190

4.4.1 Author Contribution	248
5 Conclusions	249
6 Further Developments	251
6.1 Additional Cell Viability Assays	251
6.2 Synthesis of Additional Cyclic CS-VIP	253
6.3 Classical Azobenzene AMPB Introduced into Cyclic Peptide	255
6.4 Synthesis of Fmoc-Hydrazine- <i>o</i> F ₄ Azobenzene	257
6.5 Cryo-Electron Microscopy (Cryo-EM) of the MLL1 Complex.....	259
6.6 Synthesis of WDR5 Targeting PROTACS	263
6.7 Synthesis of Additional Peptides to Target Protein-Protein Interactions	278
6.7.1 MLL2 Complex	278
6.7.2 NuRD Complex	281
6.7.3 tRNA-Guanine Transglycosylase Enzyme	297
7 References	303
8 Supplementary.....	309
8.1 List of Publications	309
8.2 Permissions for the Reprint of Publications	310

Abbreviations

$^1\text{H-NMR}$	$^1\text{Hydrogen-nuclear resonance spectroscopy}$
$^{13}\text{C-NMR}$	$^{13}\text{Carbon-nuclear resonance spectroscopy}$
$^{13}\text{C-APT NMR}$	$^{13}\text{Carbon-attached proton test nuclear resonance spectroscopy}$
λ	Wavelength
A	Anisotropy
Å	Angstrom
aa	Amino acid
AdoHcy	S-Adenosylhomocysteine
AdoMet	S-Adenosylmethionine
ADP	Adenosine diphosphate
Ahx	Aminohexanoic acid
ALL	Acute lymphoblastic leukemia
Alloc	Allyloxycarbonyl
AML	Acute myeloid leukemia
AMPB	(4-Aminomethyl)phenyl-azobenzoic acid
Ash2L	Absent, small, homeotic discs-2-like
ATP	Adenosine triphosphate
Azo	Azobenzene
BAH	Bromo-adjacent homology
BCL11A	B-cell lymphoma/leukemia 11A
Boc	<i>tert</i> -Butyloxycarbonyl
calcd	Calculated
CD	Chromodomain
CHD3/4	Chromodomain-helicase DNA-binding protein 3/4
Cl-Trt	Chlorotrityl
cryo-EM	Cryo-electron microscopy
CS-VIP	Conformationally strained visible-light photoswitch
CW	Zinc binding domain with conserved Cys and Trp residues
DCD	Double chromodomain
DCM	Dichloromethane
DIC	Diisopropylcarbodiimide

ABBREVIATIONS

DIPEA	Diisopropylethylamine
DMF	<i>N,N</i> -Dimethylformamide
DNA	Deoxyribonucleic acid
DODT	3,6-Dioxa-1,8-octanedithiol
DOT1	Disruptor of telomeric silencing-1
EBFAZ	Early B-cell factor-associated zinc finger protein
EtOAc	Ethylacetate
eq.	Equivalents
Evi3	Ecotropic viral integration site 3 protein
FAM	Fluorescein
FITC	Fluorescein isothiocyanate
Fmoc	Fluorenylmethoxycarbonyl
FOG-1	Friend of GATA (zinc finger protein)
FP	Fluorescence polarization
GATA	Transcription factor GATA
GSH	Glutathione
H	Histone
H3K4	Lysine 4 of histone 3
HAT	Histone acetyltransferase
HATU	<i>o</i> -(7-Azabenzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate
HDAC	Histone deacetylase
HDM	Histone demethylase
HMT	Histone methyltransferase
<i>hox</i>	<i>Homeobox</i> genes
HPLC	High performance liquid chromatography
HRMS-ESI	High resolution mass spectrometry – electron spray ionization
IC ₅₀	Half inhibitory concentration
<i>K</i> _d	Dissociation constant
<i>K</i> _i	Inhibition constant
KMT2A	Lysine [K] specific methyltransferase
MBD	Methyl-CpG-binding domain protein 2/3
MBT	Malignant brain tumor

MeCN	Acetonitril
MLL1 / 2	Mixed-Lineage-Leukemia 1 / 2
mP	Milipolarization
MTA1/2/3	Metastasis-associated protein 1/2/3
Mtt	Methyltrityl
mut	Mutated
ncRNA	Non coding ribonucleic acid
NC	Neural crest
NSC	Neural stem cell
NuRD	Nucleosome remodeling complex
<i>o</i> F ₄ Azo	Tetra- <i>ortho</i> -fluoroazobenzene
p55	Chromatin assembly factor p55 subunit (<i>Drosophila</i> analogue to RbBP4/7)
p66	Transcriptional repressor p66
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
PHD	Plant homeodomain
pip	Piperidine
POI	Protein of interest
PPI	Protein-protein interaction
preQ1	7-Cyano-7-deaza-guanine
PROTAC	Proteolysis targeting chimera
PSS	Photostationary state
PTM	Post translational modification
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium-hexafluorophosphate
RbBP5	Retinoblastoma-binding protein-5
RbBP4/7	Histone-binding protein RbBP4/7
R _f	Retention factor
RNA	Ribonucleic acid
Ro5	Rule of five
SAH	S-Adenosylhomocysteine
SALL4	Sal-like protein 4
SAM	S-Adenosylmethionine
scr	Scrambled

ABBREVIATIONS

SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SET	Su(var), E(z) trithorax-domain
SPPS	Solid phase peptide synthesis
TCEP	Tris(2-chlorethyl)phosphate
TFA	Trifluoroacetic acid
TGT	tRNA-guanine transglycosylase
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TLC	Thin layer chromatography
t _R	Retention time
TRR81	Transregional collaborative research center 81
TTD	Tandem tudor domain
UPS	Ubiquitin-proteasome system
Ush	U-shaped protein (zinc finger protein Ush)
VHL	Von Hippel-Lindau
WD40	WD-repeats of 40-60 amino acids that end with Trp-Asp
WDR5	WD-repeat protein-5
<i>Win</i>	WDR5 interacting
wt	Wild type
Zeb2	Zinc-finger E-box-binding homeobox 2

Abbreviations for all sections of the thesis, except for the cumulative part: section 4.

General Remarks on Thesis Structure and Nomenclature

The presented thesis “Photoswitchable Peptides as Tools for Light-controlled Modulation of Epigenetic Protein Complexes” is structured into 6 main sections:

1 Introduction, providing a theoretical background on thesis relevant topics.

2 Aim, explaining the goals and motivation behind this thesis.

3 Summaries, recapitulating the results obtained during this work in English and in German.

4 Cumulative Part, consisting of 4 published / submitted manuscripts (manuscript A, B C and D).

5 Conclusions and Outlook, highlighting the contribution of this thesis to the scientific world and providing future perspectives.

6 Further Developments, discussing additionally performed experiments, which have not been included into any of the 4 above displayed manuscripts.

The Figures, Schemes and Tables displayed in the respective sections, are labelled individually starting with number “1”. In order to differentiate between the individual sections, a prefix, matching the specific section will be applied before the respective number. Prefixes utilized for the sections are:

1: **I-X**

2: **A-X**

3: no Figures, Schemes or Tables displayed

4: the manuscripts are displayed as published / submitted

5: no Figures, Schemes or Tables displayed

6: **F-X**

“**X**” stands for the respective number. As an example, Figure 1 of the Introduction would be labelled as: “**Figure I-1.**”.

The nomenclature of the molecules is addressed alike, as labelling of Figures, Schemes and Tables. E.g. the first molecule presented in the introduction would be labelled with “**I-1**”.

Exceptions are molecules, which are displayed and labelled individually inside a manuscript of the cumulative part 4 of this thesis. If such a molecule will be mentioned / displayed in any other thesis section, it will be labelled with the assigned number inside the respective manuscript. An additional prefix will be added in front of the number, describing the corresponding manuscript. The 4 different manuscripts are referred to manuscripts A, B, C and D throughout the thesis. Thus, the corresponding prefix utilized for labelling molecules of a specific manuscript will be:

manuscript A: **MA-X**

manuscript B: **MB-X**

manuscript C: **MC-X**

manuscript D: **MD-X**

“**X**” stands for the respective molecule number applied inside the specific manuscript.

1 Introduction

1.1 Epigenetics

The term “epigenetics” coined by Waddington in 1942 literally means “in addition to/ over genetics”. At that moment epigenetics was the attempt to unite genetics and developmental biology as “a branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being”.¹ This early description of epigenetic phenomena includes the whole set of developmental processes, linking the genotype with the phenotype. It was proposed that cellular differentiation can be considered an epigenetic phenomenon, governed by changes in the so-called epigenetic landscape (Figure I-1) rather than alterations in genetic inheritance.^{2,3}

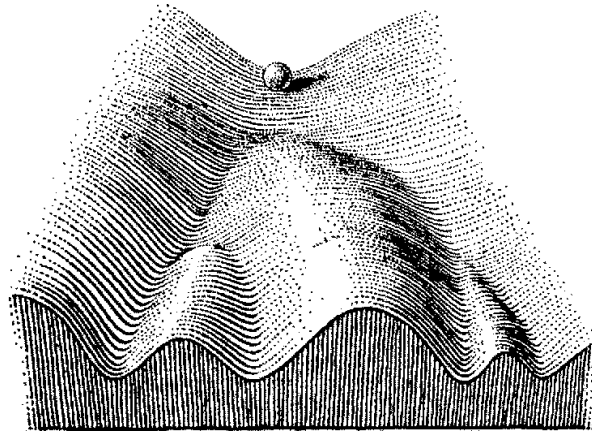


Figure I-1. Waddington proposed that the epigenetic landscape represents the process of cellular decision-making during development. The cell, depicted as a ball, can take specific permitted pathways leading to different cellular fates.³

Nowadays some scientists are still using the expression epigenetics only for mechanisms that are erased and re-established in every developmental cycle.⁴ Albeit currently, epigenetics is considered as the study of potentially heritable and environmental influenced modifications to the genome that can alter gene expression without changing the underlying DNA sequence.^{4,5} While for plants, fruit flies and nematodes, transgenerational epigenetic inheritance has been observed, whether or not epigenetic information is also passed on through the germline in mammals, is the matter of controversial debate.^{5,6} Anyhow, though genetic information is required to reveal which proteins can be expressed by a gene, epigenetic information is crucial for the determination of when, where and to what extent the genetic information will be utilized. Thus, epigenetics provides the instructions for gene expression, and hence, regulates the whole genome. Among others, the main predominant epigenetic mechanisms are histone modifications and DNA methylation, that can be influenced by non-coding RNAs (ncRNAs),^{7,8} as well as ncRNAs are functioning in modulation of chromatin, transcriptional regulation and destabilization of the nucleosome.⁹ Aligning with the initial suggestion that heritable epigenetic changes in gene expression are important for tumor progression,^{10–13} those epigenetic modifications can have vast and long-term effects on health via aberrations that can result in disease^{14–18} like Alzheimer’s,¹⁹ pediatric syndromes and autoimmune disease.²⁰ Epigenetic regulation is a dynamic and reversible process, which is controlled through a complex signaling network of protein-protein interactions (PPIs). The reversible nature of epigenetic modifications offers promising opportunities to tackle diseases by developing epigenetic-based therapeutics,²¹ wherefore it is crucial to understand the patterns of those epigenetic modifications.⁷

1.2 Post-Translational Modifications of Histones

This thesis is focused on the modulation of the histone-modifying enzyme Mixed-Lineage-Leukemia 1 (MLL1), also known as lysine [K]-specific Methyl Transferase 2A (KMT2A).^{22#} Histone modifications emerged as essential epigenetic processes to alter gene expression.²³ Histones are highly basic proteins, that are involved in the packaging of DNA into the cell nucleus (Figure I-2). The DNA is wrapped in 1.7 left-handed superhelical turns (147 base pairs) around an octamer of four different histone proteins (H3, H4, H2A, H2B). These, so-called nucleosomes, are the fundamental building blocks of chromatin, in which small sections of dsDNA serve to join the nucleosomes together, which are associated with a fifth linker histone H1.²⁴ The N-terminal 15-30 amino acid residues of the histone proteins — “histone tails” — are unstructured, highly basic and protrude from the nucleosome core particle (NCP). In general, these histone tails are subjected to a large series of post-translational modifications (PTMs), such as acetylation, methylation, phosphorylation, ubiquitination, sumoylation, biotinylation, citrullination (deimination) and ADP-ribosylation.

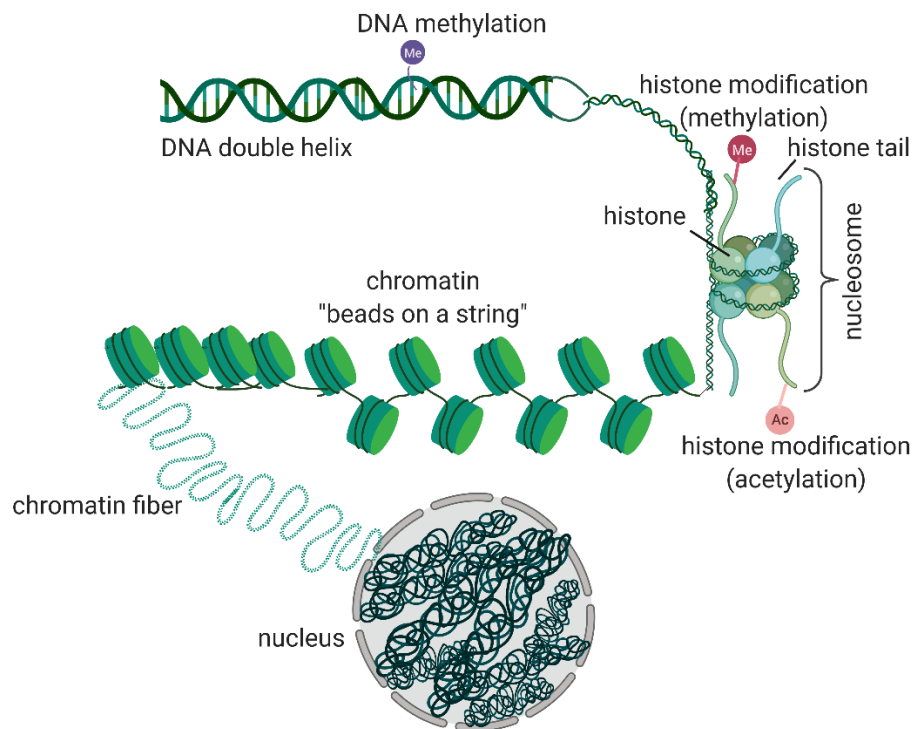


Figure I-2. Packing of DNA into the nucleus. The dsDNA helix (~2 nm), is wrapped around nucleosomes that are connected with each other by a short linker DNA building the 11 nm “beads on a string” structure of chromatin. These are further packaged into 30 nm chromatin fiber, which is condensed into the nucleus.

PTMs can have high influence on the chromatin structure, which can lead to exposure of DNA regions to various transcription factors to bind to the specific DNA sequences. This makes them influential for many cellular events, including gene expression, DNA replication, cell cycle control and chromatin compaction.²⁵ The chromatin is further organized into distinct domains that differentiate regions of transcriptional activity from transcriptionally silent ones. Thus, regions that are highly condensed, termed heterochromatin, are transcriptionally repressed. Contrarily, the loose euchromatin contains most of the transcriptionally active genes. The degree of chromatin packaging is influenced by various factors including histone chaperons that facilitate the deposition and eviction of histones,^{26,27} chromatin remodelers that use the power of ATP hydrolysis to reposition nucleosomes,²⁸ the actions

² # The nomenclature of the H3K4 methyltransferases has been proposed to be changed from MLL to KMT2 in 2013 due to increasing misuse of *mll2* for two different genetic loci. Still, to date, many scientists refer to the human *mll1* gene on chromosome 11q23.3 as *mll1* instead of *kmt2a*²²

of linker histone H1 that regulate the formation of higher-order chromatin states²⁹ and chemical modifications to the DNA^{30,31} and histones.^{32,33} In terms of PTMs, the histone tails of the euchromatic regions are highly acetylated, which weakens the interaction of the normally basic, positively charged histones to the negatively charged phosphate backbone of the DNA by changing the net charge of the histones. Thus, transcription factors are more likely to access the DNA compared to the condensed heterochromatin (Figure I-3).^{34,35}

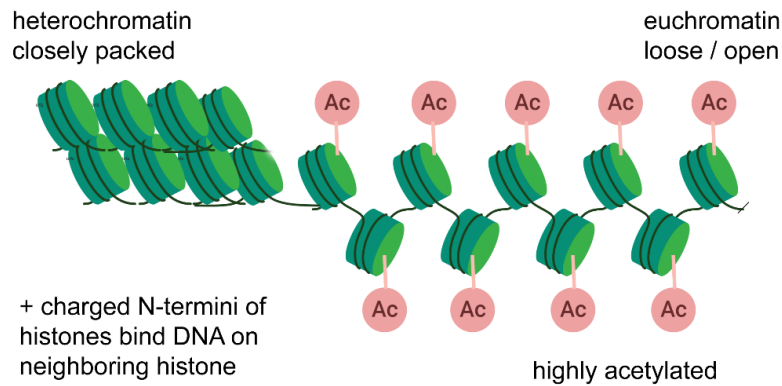


Figure I-3. The closely packed, condensed heterochromatin with positively charged N-termini of lysine residues that interact with the phosphate backbone of the DNA (left) is transcriptionally inactive. The euchromatin (right) with highly acetylated lysine residues is transcriptionally active.

This would suggest that acetylation of histones generally leads to gene activation. However, it is known that distinct PTMs can work in combination to determine downstream events. The remarkable diversity and specificity of biological actions that are associated with these PTM patterns have led to the proposal that PTMs are forming a “histone language” on the histone tails. Initially, Strahl and Allis referred to this theory as a “histone code”.³⁶ This initial histone code hypothesis was based on findings of an interplay between different histone modifications, yielding specific functions (Figure I-4).

	Residue:	Modification state	Function
H3	N 14 9 10 14 28	Unmodified	Gene silencing
	N	Acetylated	Gene activation
	N	Acetylated	Histone deposition?
	N	Phosphorylated	Mitosis/meiosis
	N	Phosphorylation/ acetylation	Gene activation
	N	Methylated	Gene activation
	N	Higher-order combinations	?
	H4	N 5 12	Acetylated
N 8 16		Acetylated	Gene activation

Figure I-4. Initial histone code hypothesis. Distinct PTMs work sequentially or in combination to determine downstream events. The displayed patterns have been demonstrated to being closely linked to the indicated biological events.

Since the original histone code hypothesis, there have been many advances to understand chromatin regulation. Genomic wide profiling of histone modifications has provided novel insights of crosstalk between various modifications.^{37,38} Thus, nowadays the histone code is appreciated as a whole new rich and vibrant “histone language” due to its complexity and the various combinatorial modification patterns.^{39–41} In order to further understand the language of histones, the identification of the protein machineries that incorporate (write), remove (erase) and recognize (read) specific PTMs have to be identified and studied. Thus, PTMs can function through the selective recruitment of specific effector proteins (type of a reader protein) that interact with distinct PTMs to direct downstream events in chromatin.⁴⁰ The first identified reader protein, a bromodomain (BD) for acetyl-lysine, was identified in 1999,⁴² followed by the discovery of a chromodomain (CD)^{43–45} and a plant homeodomain (PHD),^{46–49} both recognizing methyl-lysine. Growing identification of previously unknown types and sites of PTMs as well as characterization of novel chromatin modifying enzymes is continuously raising the complexity of the already complicated histone language phenomenon.^{50,51} Yet, in 2018, there were over 100 “reader” proteins, around 50 “writers” and about 12 “erasers” known for the different PTMs.⁵²

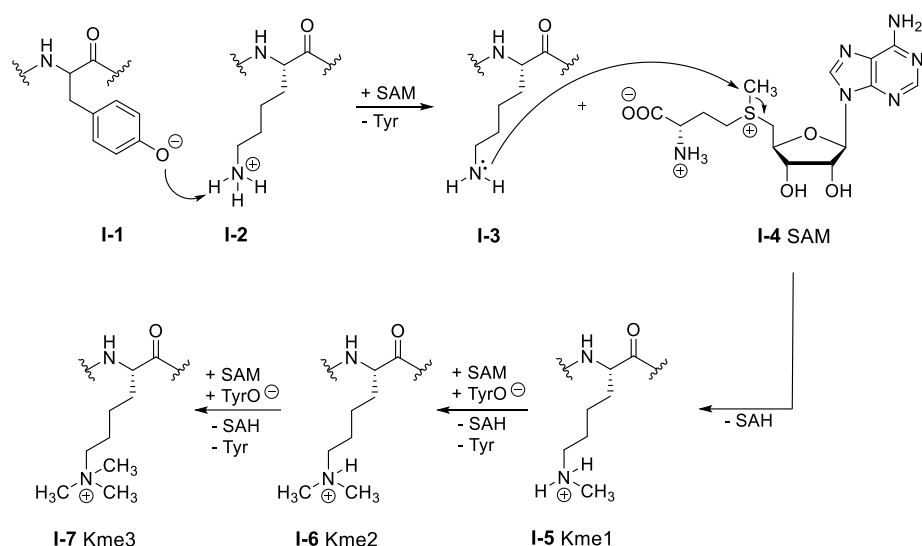
This thesis is centered on the histone lysine methylation of H3K4, wherefore the focus here is set on the writers, erasers and readers of methyl-lysine as an epigenetic mark. Since histone lysine methylation is not affecting the net electrostatic charge of chromatin, its implication in gene expression is more complicated. Gene activation or repression signals depend on the methylation state and its precise position.⁵³ For example, the mono- and dimethylation mark on H3K9 stands for transcriptionally silent genes, whereas the trimethylation of H3K9 and H3K4 activates them.^{53,54} Thus, H3K4me3 is strongly enhanced at promoter regions, where chromatin exists in an open conformation.^{54,55} Likewise, H3K4me1 is highly enriched at enhancers and H3K4me2 towards the 5’ end of actively transcribing genes.^{54,56–58} Table I-1 summarizes identified lysine [K] methyltransferases (KMTs), lysine [K] demethyltransferases (KDMTs) and the known readers of histone lysine methylation marks.

Table I-1. Summary of identified writers, erasers and readers of histone lysine methylation marks. DOT1 = disruptor of telomeric silencing-1; BAH = bromo-adjacent homology; CD = chromodomain, CW = zinc binding domain with conserved Cys and Trp residues; DCD = double chromodomain; MBT = malignant brain tumor; PHD = plant homeodomain; PWWP = single letter amino acid sequence; SAWWADEE = single letter amino acid sequence; TTD = tandem-tudor domain; WD40 = WD-repeats of 40-60 amino acids that end with Trp-Asp.

Histone PTM	Writers ²⁵	Erasers ²⁵	Reader domains ⁵⁰
H3Kme1/ me2/me3	See below	See below	MBT
H3K4me1	SET1A/1B, MLL1/2/3/4, SMYD1/2, SET7/9, PRDM9	LSD1/2, NO66, JARID1B	Chromobarrel, DCD, WD40
H3K4me2	SET1A/1B, MLL1/2/3/4, PRDM9	LSD1/2, JARID1A JARID1B, JARID1C, JARID1D	DCD, PHD, WD40
H3K4me3	SET1A/1B, MLL1/2, PRDM9	NO66, JARID1A, JARID1B, JARID1C, JARID1D	DCD, PHD, Spindlin (in combination with H3R8me2a), TTD, CW, WD40

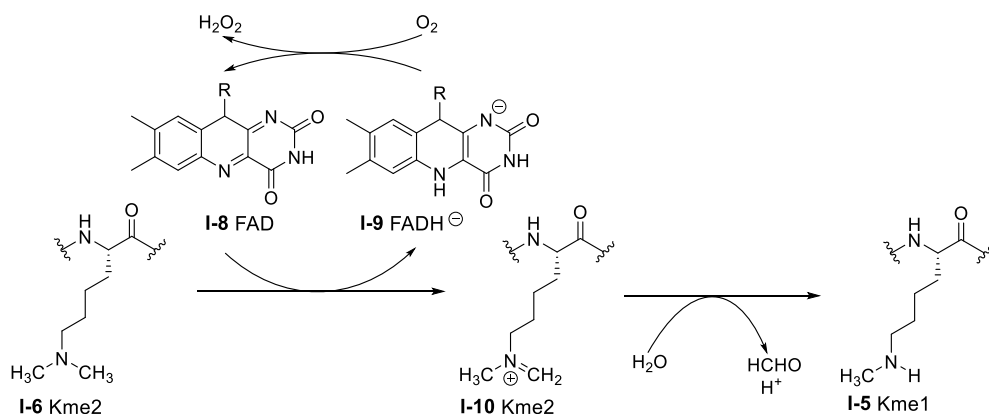
H3K9me1	G9a, SETDB1	JHDM2, PHF8	Ankyrin, SAWADEE
H3K9me2	SUV39H1/2, G9a, GLP, SETDB1	JHDM2, JHDM3, PHF8	Ankyrin, CD, SAWADEE
H3K9me3	SUV39H1/2, GLP	JHDM3	ADD, CD, PHD, SAWADEE, TTD, WD40
H3K27me2	EZH1/2	UTX, UTY, JMJD3, KIAA1718, PHF8	CD
H3K27me3	EZH1/2	UTX, UTY, JMJD3	BAH, CD, WD40
H3K36me1	NSD1/2/3, SETD3	JHDMI1	Not known
H3K36me2	NSD1/2/3, SETD3, SMYD2, ASH1L, SETMAR	JHDMI1/3	Chromobarrel
H3K36me3	SETD2	JHDMI3	Chromobarrel, PWWP, Tudor
H3K79me1	DOT1L	Not known	Tudor
H3K79me2	DOT1L	FBXL10 ⁵⁹	Not known
H3K79me3	DOT1L	FBXL10 ⁵⁹	PWWP
H4K20me1	SET8	PHF8, LSD1n	Chromobarrel, PWWP
H4K20me2	SUV4-20H1/2	PHF2	BAH, TTD
H4K20me3	SUV4-20H1/2	PHF2	PWWP

The predominant class of KMTs, including our target MLL1, contains the catalytically active SET (Suppressor of variegation (Su(var)) Enhancer of Zeste (*E(z)*) Trithorax) domain. It is a widely conserved ~130 amino acid motif with a β -architecture embedded within a larger α -helical fold.⁶⁰ The SET domain utilizes the cofactor S-adenosylmethionine (SAM or AdoMet) to facilitate substrate methylation. Notably, the binding sites for the lysine substrate and the SAM are located on opposite sites and are connected by a deep hydrophobic channel, running through the core of the SET domain.⁶⁰ This open hydrophobic channel is a key feature of the SET domain, because it connects the substrate- and cofactor binding faces.⁶¹ Such binding cleft is allowing the transfer of the methyl group from the SAM to the amine group of the lysine side chain. The unusual arrangement of the substrate binding site was proposed to allow specific protein recognition and addition of multiple methyl groups.^{60,62,63} The methyl transfer was further demonstrated to be catalyzed by the hydroxyl group of a conserved Tyr residue, located at a narrow pore connecting the two sites.⁶⁰ The methylation via a S_N2 nucleophilic substitution reaction, leaving the demethylated cofactor side product S-adenosylhomo-cysteine (SAH or AdoHyc), is displayed in Scheme I-1. The specific structural characteristics of the SET1 domain of our target enzyme MLL1 are specified in the following section 1.3.



Scheme I-1. Mechanism of the lysine methylation catalyzed by a conserved Tyr residue **I-1** of the SET domain. The hydroxyl group of the Tyr deprotonates the ϵ -amine group of the substrate lysine **I-2**. The methyl group is transferred from the SAM **I-4** to the lysine side chain via a S_N2 nucleophilic reaction to yield mono- (**I-5**), bi- (**I-6**) or trimethylated (**I-7**) lysine.

Initially, lysine methylation was believed to be an irreversible process until the first lysine demethylase (KDMT) of H3K4, the FAD (Flavin adenine dinucleotide)-dependent nuclear amine oxidase LSD1 (also known as KDM1A), has been identified.⁶⁴ The LSD1/2 KDMTs are able to demethylate H3K4me1/me2^{64,65} where they are recruited to their target genes by members of the zinc-finger transcription factor family Snail.⁶⁶ The demethylation mechanism (Scheme I-2) of the LSD family was proposed to work via a redox process where a conserved Tyr and a Lys-water-flavin motif play a crucial role to properly orient the FAD with respect to the substrate. The reaction was postulated to work through a direct hydride transfer mechanism.⁶⁷ It was revealed that the pK_a value of H3K4 in complex with LSD1 was remarkably lower than in its free state (pK_a of H3K4 substrate alone: 10.4; in LSD1 complex: < 0), which was ascribed to the microenvironment. This suggests that the predominant form of H3K4 under physiological conditions is preferentially bound to LSD1 in an uncharged dimethylated form.⁶⁷



Scheme I-2. Proposed catalytic mechanism of lysine demethylation of LSD1.⁶⁷ Lys and Tyr of LSD are positioning the FAD **I-8** cofactor in the correct position for hydride transfer. FAD **I-8** is acting as an oxidant being itself reduced to FADH⁻ **I-9** via a direct hydride transfer from the lysine from the lysine α -carbon to flavin, whereby the methylated lysine is oxidized.

Subsequently, another class of lysine demethylases was identified, the JmjC (Jumonji C) domain-containing proteins. NO66 belongs to this protein family and is able to demethylate all three states of H3K4 methylation via an oxidative reaction using Fe(II) and α -ketoglutarate as cofactors.⁶⁸ The JARID1 (jumonji domain ARID-containing protein) family proteins were demonstrated to specifically demethylate H3K4me2/ me3.^{69–72}

Because the methylation state of lysines is not affecting the electronic charge of the chromatin, its function is considered to be mainly exerted by effector proteins that specifically recognize the methylated site.²⁵ Such “reader” proteins exert the methyl-lysine binding domains, summarized in Table I-1. This includes the H3K4 recognizing Tudor, tandem-Tudor (TTD) PHD (plant-homeodomain), chromo, CW (zinc binding domain with conserved Cys and Trp residues), MBT (malignant brain tumor) and WD40 (WD-repeats of 40-60 amino acids that end with Trp-Asp) domains, which have the ability of distinguishing target lysines by their methylation state and amino-acid surrounding.^{25,50} Although different binding domains are folded differently, the domains that recognize the same epigenetic mark, display remarkable resemblance between each other. Thus, most histone lysine readers form an aromatic cage to accommodate the long hydrophobic methylated lysine side chain. Readers of mono- and dimethylation generally exert small keyhole-like cavities, limiting the access of trimethylated lysines. Contrarily, readers of di- and trimethylation marks usually utilize a wider, more accessible binding groove.⁷³ Structures of the respective reader domains are exemplified in Figure I-5.

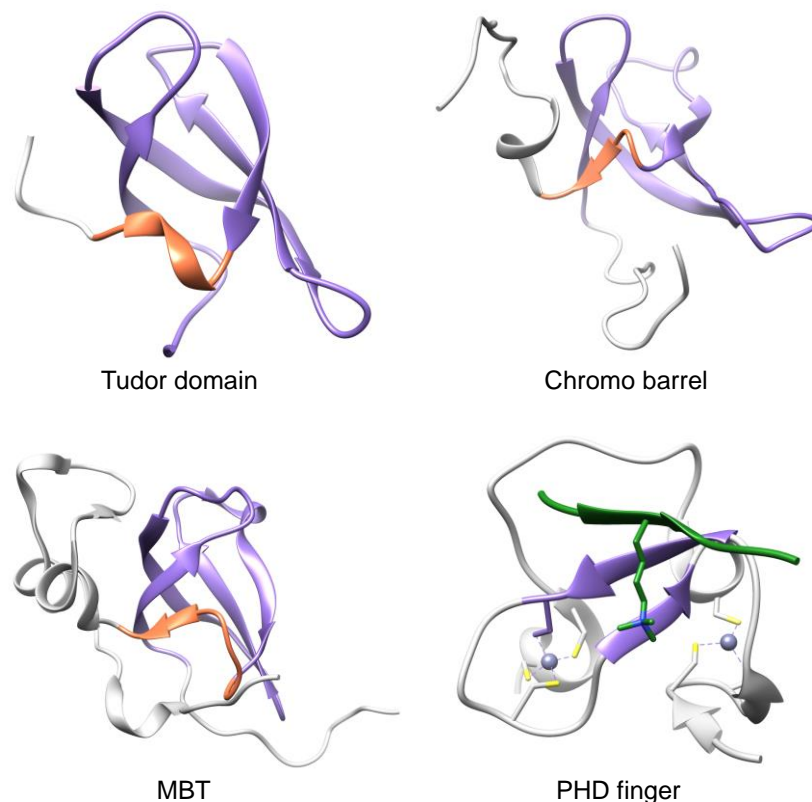


Figure I-5. Structural features of H3K4 methylation effector proteins of Tudor domain, chromo barrel, MBT and PHD finger. β -barrel motifs = blue; protein = grey; H3K4me3 peptide = green; additional β -strand or helix, occupying one face of the usually four stranded core = orange. PDBs: Tudor = 1MHN;⁷⁴ chromobarrel = 2BUD;⁷⁵ MBT = 1OZ2;⁷⁶ PHD = 2PNX.⁷⁷

The Tudor, as well as the spindlin domains form tudor like β -barrels, which are packed against each other forming a triangular assembly, where the H3K4me3 mark is recognized by a characteristic binding site consisting of four aromatic residues.⁷⁸ The PHD fingers comprise one of the largest reader

domains to date, with over 100 proteins containing such model.⁷⁹ They are 50-80 amino acid long, Zn-coordinating domains consisting of a two-strand anti-parallel β -sheet and a C-terminal α -helix. They recognize unmodified as well as methylated lysine side chains, the majority binding the methylation mark at H3K4me3.⁸⁰ For example, the plant specific histone reader SHL was demonstrated in being capable of recognizing H3K4me3 via its BAH (bromo-adjacent homology) and PHD domains.⁸¹ Binding of the H3 tail to the PHD finger is mediated by formation of an anti-parallel β -strand of the N-terminal H3 residues to the two-stranded β -sheet of the PHD domain, while recognition of the H3K4me3 mark is as well facilitated through an aromatic cage.⁸² The trimethyl ammonium group is stabilized by van der Waals and cation- π interactions, similar to those observed for the MBT (malignant brain tumor) and chromo-domains.

Another evolving class of histone readers are the WD40 protein family, which is one of the most abundant protein domains in the human proteome with more than 360 domains currently annotated.⁸³ The WD40 repeats are comprised of 40-60 amino acid motifs that preferentially end with a Trp and Asp (WD). Typically, WD40 proteins consist of a 7-bladed β -propeller fold, where each blade comprises a four-stranded anti-parallel β -sheet, and a central cavity.⁸⁴⁻⁸⁶ The crystal structure of an exemplified WD40 protein, is displayed in Figure I-6.

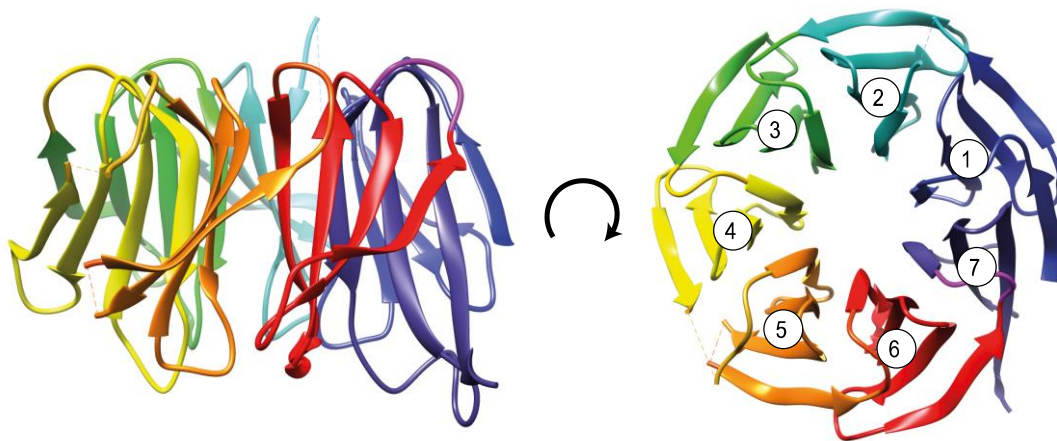


Figure I-6. Structural features of WD40 proteins. They often fold into seven-bladed β -propellers with a funnel-like shape. Displayed protein is the WD40 protein Tup1. PDB: 1ERJ.⁸⁷

The WD40 protein class is highly conserved throughout bacteria and mammals and are often part of multiprotein complexes, where they are mediating protein-protein interactions (PPIs).⁸⁸ Due to the fact that WD40 proteins exert various surfaces, enabling interaction to multiple binding partners, it comes as no surprise that they are crucial for maintaining the integrity of those complexes. Although, to date, no WD40 protein has been demonstrated to display intrinsic enzymatic activity, they rather act as presenter modules to promote further modifications.⁸⁹ As chromatin readers, WD40 proteins are able to bind to the H3 tail at the center of their β -propeller structure. WD40 domains are able to read various post-translational methylation marks. In regard to H3K4 methylation, it has been demonstrated that a WD40 protein, WDR5 (WD-repeat protein-5), is directly associating with the di- and trimethylation mark at that position.⁹⁰⁻⁹⁴ WDR5 is, thereby, facilitating nucleosome binding of the full SET1 methyltransferase complex and, thus, enabling global H3K4 trimethylation and *hox* (*homeobox*) gene activation in mammalian cells.⁹⁰ *Hox* genes, in general, encode for homodomain transcriptional factors, which, in turn, regulate the organ formation during embryo development as well as proper hematopoiesis in adults.⁹⁵ Analysis of WDR5-H3 structure, revealed that Ala1, Arg2 and Thr3 of the H3

tail are making important contacts to WDR5, where the Arg2 side chain inserts into the central channel of the β -propeller, while Lys4 does not exert contacts to WDR5.^{91–93} Binding affinity of H3 to WDR5 was demonstrated to be in the μM range ($K_d = 35 \mu\text{M}^{93}$ and $6.6 \mu\text{M}^{91}$). It may be possible that the Lys recognition mode is yet to be discovered, or that other proteins in complex with WDR5 in lysates mediate the observed methylated lysine specificity. Although the methylation status of H3K4 might not directly influence H3 binding to WDR5, the methylation status of H3R2 does. Thus, binding affinity of WDR5 to H3 is increased upon symmetrical demethylation of H3R2 (K_d : H3R2me2 = $0.1 \mu\text{M}$; H3R2 = $5.6 \mu\text{M}$).⁹⁶ Notably, the two WD40 proteins WDR5 and RbBP5 (retinoblastoma-binding protein-5) are both components of the histone modifying protein complex of MLL1, whose structural characteristics and function will be detailed in section 1.4.

1.3 Mixed Lineage Leukemia 1

The Mixed Lineage Leukemia 1 (MLL1) enzyme (or KMT2A, see section 1.2) belongs to the SET1 family of histone 3 lysine 4 (H3K4) methyltransferases. Upon translation, wild-type MLL1 is proteolytically cleaved by caspase-1⁹⁷ leading to a 320 kDa N-terminal fragment, containing all domains, except the transcriptional activation and SET domain, which are located in the 180 kDa C-terminal fragment (Figure I-7). The SET1 enzyme family (other family members: MLL2, MLL3, MLL4, SET-1A, SET-1B) engages a pivotal role in the transcriptional regulation of genes involved in hematopoiesis and development. Thus, the trimethylation of H3K4 together with an H3K9/14 acetylation in conjunction with the H3K36 trimethyl-mark are characteristic for transcriptionally active genes.⁹⁸ Allis et al.⁹⁹ postulated that the H3K4me3 is a recognition site for additional factors that are required for transcriptional activation and that it facilitates the recruitment of transcriptional complexes leading to gene activation.^{32,55} Additionally, it is known that the *mll1* gene has over 100 fusion partners, which are involved in cancers. For example, MLL1-ZCH13 fusions were found in solid prostate tumors, and they induced chromosomal instability, affected mitotic progression and enhanced tumor formation.¹⁰⁰ The fusion of MLL1 to AF4 protein (*all1*-fused gene from chromosome 4) is responsible for ~ 50 % of infant acute lymphoblastic leukemia (ALL) and 75 % of adult MLL-rearranged (*MLL-r*) ALLs.¹⁰¹ AF9 is the second most common fusion partner for MLL1, which is usually associated with acute myeloid leukemias (AMLs).¹⁰² Target genes of MLL1 are *homeobox (hox)* genes like *hox-a9* and *hox-c8*, as well as *meis* and *depror* genes. Overexpression of the *depror* gene was found in multiple myeloma cells¹⁰³ as well as T-cell leukemia.¹⁰⁴ Concerning the *hox* genes, it was demonstrated that *hox-c8* is expressed at 24000 times higher levels in the presence of MLL1, which is able to bind directly to the promoter regions of the *hox-c8* locus and trimethylates H3K4.¹⁰⁵ *Hox* and *meis* (HOX cofactor) are the most frequently over-expressed genes in *MLL-r* leukemias.^{106,107} Although, in the hematopoietic system *hox* and *meis* genes are normally expressed at highest levels in stem cells and early lineage progenitor cells, upon differentiation expression levels are downregulated.^{108,109} Persistent overexpression of both genes has been observed in various leukemias.^{106,110} Aberrations of MLL1 and increased levels of MLL1 target genes were observed in AML and ALL,^{111,112} which turns the inhibition of MLL1 methylation activity into an attractive therapeutic target.

The writer domain for H3K4 methylation of the MLL1 enzyme, is the SET1 domain, whose general features have been discussed in section 1.2. In the case of the SET1 domain of MLL1, the essential active site residues of MLL1 are Phe3884, Tyr3942, Tyr3944, Phe3946, Cys3882 to Phe3885, forming the hydrophobic channel located between the SET-I and the postSET loop (Figure I-7).⁶¹

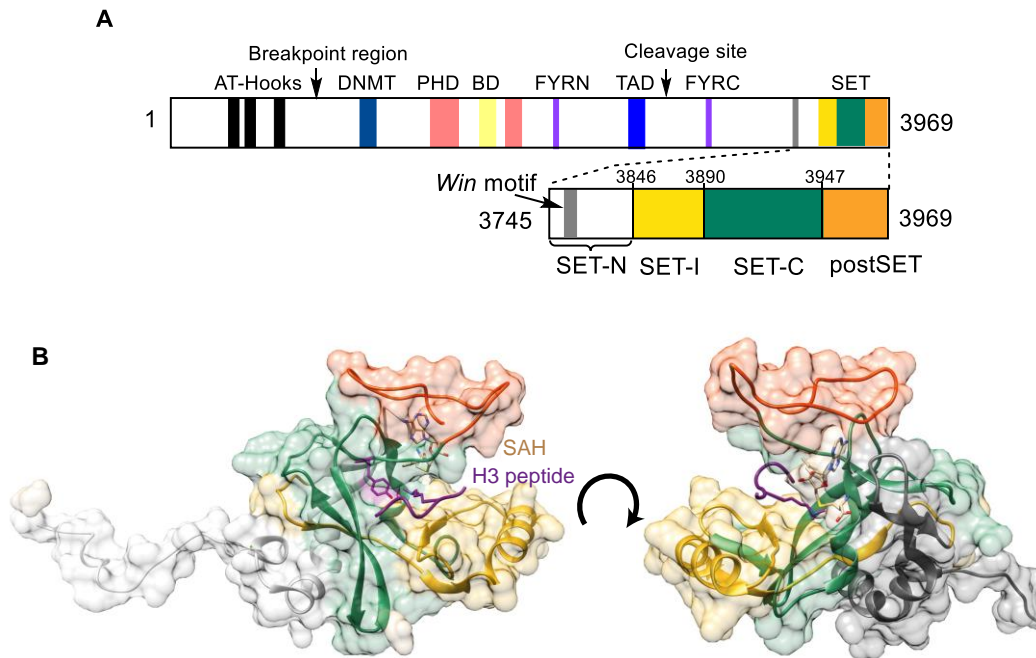


Figure I-7. A) Schematic illustration of the MLL1 enzyme with the C-terminal SET-domain indicating the subdomains. **B)** Surface representation of the MLL1 SET domain with the bound cofactor product SAH (tan) and the histone H3 peptide substrate with the targeted lysine side chain (purple).⁶¹ (PDB MLL1: 2W5Y⁶¹).

In 2015, Zhang et al.¹¹³ identified a previously unknown function of the postSET loop of MLL1 for its activation mechanism. In comparison of the crystal structures of MLL1 and MLL4 (Figure I-8) the SET-I region in MLL1 is located further away from the active site.^{114,115} It was demonstrated that in MLL1, the postSET loop engages a more open conformation that is more distant from the SET-I region than in MLL4. This leads to a suboptimal positioning of the target lysine, which is proposed to contribute to the weaker intrinsic activity of the isolated SET domain of the MLL1 enzyme.⁶¹ The MLL1 core complex, which is enhancing the methylation activity, is discussed in the following section 1.4.

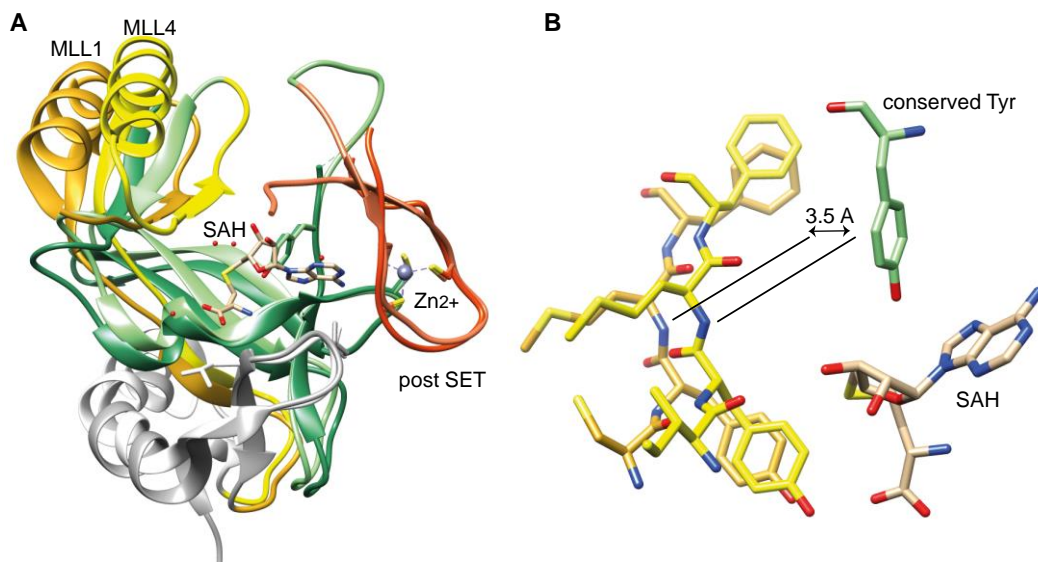


Figure I-8. A) Overlay of MLL1- (dark colors) and MLL4-SET (bright colors) domains. **B)** Positioning details of the channel-tetrapeptide, demonstrating suboptimal positioning of MLL1 domain for lysine methylation, due to broader channel opening (3.5 Å difference to MLL4; PDB MLL1: 2W5Y⁶¹, PDB MLL4: 4Z4P¹¹³).

1.4 Mixed Lineage Leukemia 1 Complex

The low enzymatic activity of the isolated SET domain of MLL1 indicates that the methylation activity is influenced by additional factors.⁶¹ Indeed, the catalytic activity of MLL1 depends on the formation of a protein core complex, which consists of MLL1, WDR5, RbBP5, Ash2L (absent, small, homeotic discs-2-like oncoprotein) and DPY30 (Protein dpy-30 homologue) (Figure I-9). The 4-mer complex (without DPY30) was firstly reconstituted by Dou et al.,¹¹⁶ who revealed that the interaction between MLL1 and WDR5 especially influences the H3K4 methyltransferase activity of MLL1. Thus, WDR5 is bridging MLL1 to the other units of the core complex, maintaining complex integrity, as well as it mediates interactions between MLL1 and the histone substrate.¹¹⁶ Additionally, it was suggested that WDR5 recognizes mono- and dimethylated H3K4, facilitating its trimethylation through MLL1.¹¹⁷

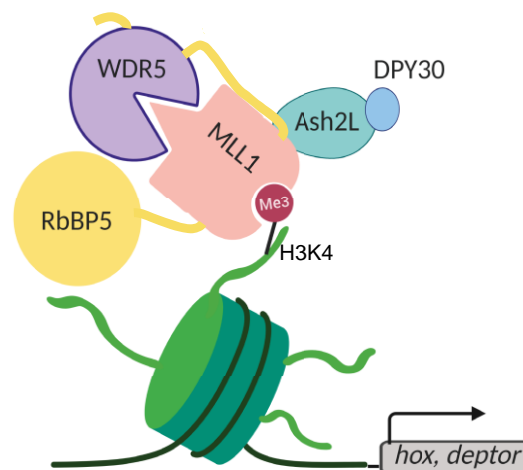


Figure I-9. Model of the MLL1-core complex enabling methylation activity of MLL1 SET domain, activating gene expression.

Just recently, the overall structure of the MLL1 complex bound to the nucleosome has been revealed using cryo-electron microscopy (cryo-EM) to a resolution of 3.2 Å, by Xue et al.¹¹⁸ The overall architecture of the complete MLL1 core complex was demonstrated to be Y-shaped. This Y-shape is in agreement with the previously reported structures obtained from the low resolution cryo-EM map of the human MLL1 complex, as well as of the analogue yeast Set1 complex (sometimes also called COMPASS (COMplex of Proteins Associated with Set1) complex).^{119–121} It was revealed that the MLL1 enzyme engages a position in the center of the complex, while WDR5 and RbBP5 form two arms and the Ash2L SPRY domain the bottom of the Y-shape (Figure I-10). A few weeks later, the groups of Dou and Cho also published the cryo-EM map (6.2 Å) of the MLL1 complex bound to the NCP.¹²² Both studies shed light on the mechanism by which the MLL1 complex recognizes H3K4, enabling its methylation. They corroborated the importance of the Ash2L-RbBP5 dimer that displays extensive interactions with the nucleosomal DNA as well as with the surface of the H4 tail. It is demonstrated that those interactions anchor the MLL1 complex to the NCP, thereby aligning the SET domain in a favored position to allow H3K4 substrate proximity, which enables methylation activity.^{118,122} Park et al.¹²² highlighted that especially the Ash2L-NCP interaction of the IDR (intrinsically disordered regions, residues 178-277) of Ash2L to nucleosomal DNA is critical for H3K4 trimethylation. Thus, utilizing immunoblot assays, it was validated that deletion of the Ash2L linker IDR completely abolishes MLL1 activity on the NCP. In addition, Ala mutation of Ash2L residues K205/R206/K207 significantly reduced

MLL1 trimethylation activity on the NCP, but not on the free H3 tail, indicating that these residues are crucial to promote trimethylation on the NCP.¹²² Xue et al.¹¹⁸ likewise point out the importance of Ash2L IDR-NCP interactions, although they neither did perform in depth analysis of the exact Ash2L residues involved, nor studied the methylation state affected. Instead their publication focuses on different cryo-EM maps of MLL1 compared to MLL3 (comparison of MLL1 and MLL3 will be discussed later on in this section) bound to the unmodified as well as to an ubiquitinated (H2BK120ub) NCP. Ubiquitination of Lys120 of histone H2B (H2BK120ub1) is known to stimulate methylation activity.^{123,124} This PTM mark was, now, found to be recognized by RbBP5. Acting as a reader module, RbBP5 is arranging the remaining MLL1 complex subunits in an optimal position to enable H3K4 methylation.¹²² In addition, kinetic analysis of MLL1 activity revealed that it displays a two-fold higher catalytic efficiency on ubiquitinated nucleosome core particles (ubNCP) than on unmodified nucleosomes, suggesting that histone H2B, ubiquitinated on Lys120 (H2BK120ub1) is further stabilizing the MLL1 interaction with the histone core.¹¹⁸

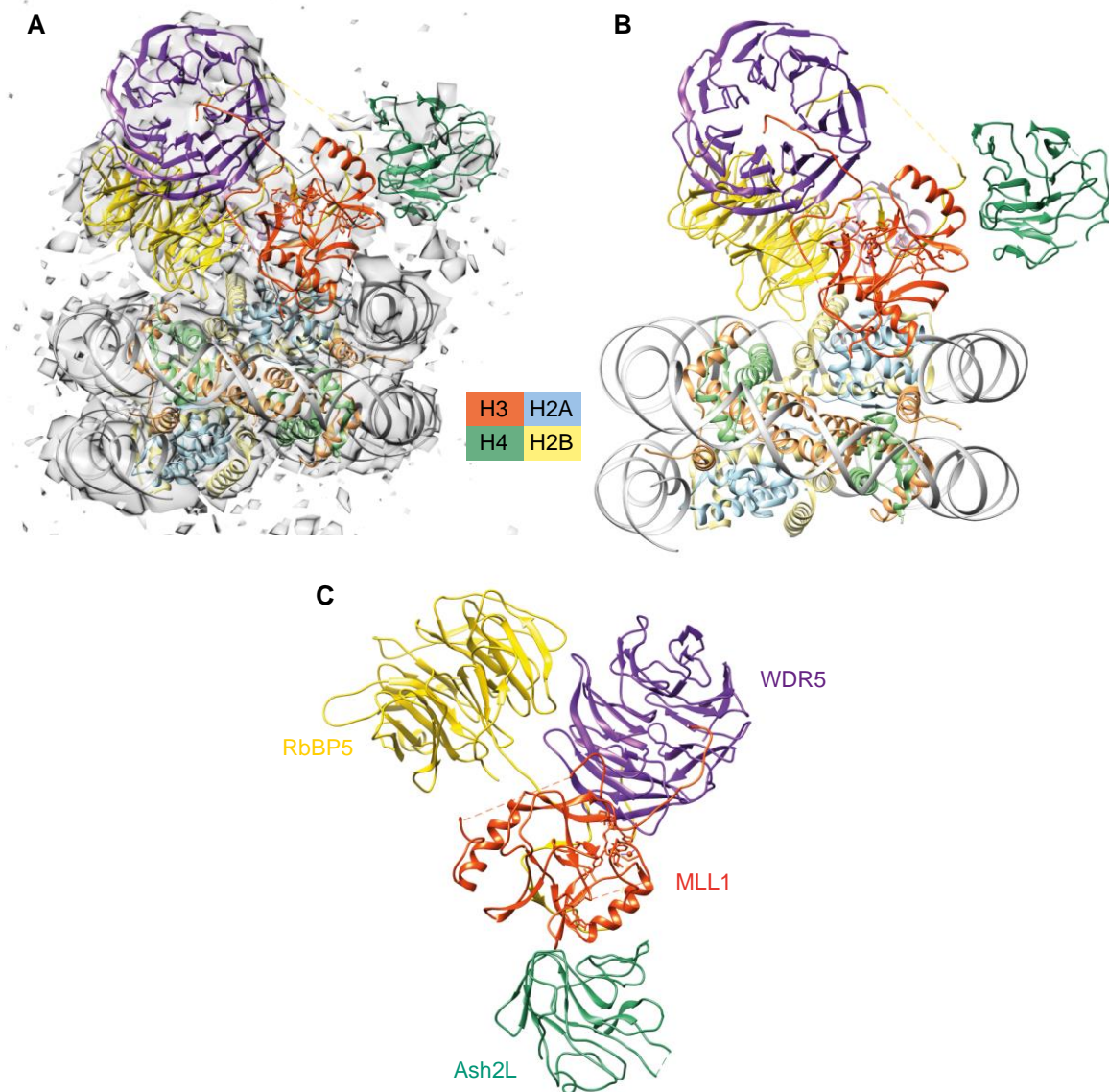


Figure I-10. Overview of the structure of the human MLL1 complex associated to the H2BK120ub1 nucleosome (ubNCP). **A)** Atomic model of human MLL1 complex mapped upon the cryo-EM density map. **B)** Atomic model of human MLL1 complex interacting with ubNCP. **C)** Atomic model of MLL1 complex highlighting Y-shaped conformation. (PDB: 6KIU; EMD: 9998).¹¹⁸

The specific interactions between the distinct components of the MLL1 core complex have already been characterized previously. Thus, in 2008, Patel et al.¹²⁵ identified a WDR5 interacting motif (*Win* motif: ³⁷⁶²GSARAE), located in the N-SET region of MLL1, whose Arg3765 residue is inserting into the central cavity of WDR5. This was validated by sedimentation velocity experiments using Ala point mutations of *Win* amino acids (S3763, R3765, E3767) of the SET1 domain. The interaction between MLL1 and WDR5, was further analyzed by the two groups of Cosgrove¹²⁶ and Kingston¹²⁷ who both determined the crystal structures of WDR5 bound to the MLL1 *Win* peptide (Ac-GSARAEVHLRKS-NH₂; affinity to WDR5: $K_i = 160$ nM), independently. The structure of WDR5 consists of seven WD40 β -propeller blades that are organized around a central cavity. The two channel openings at the top and the bottom are connected by a water-filled tunnel. The N-terminus of MLL1 *Win* peptide forms a 3_{10} -helix (three amino acids per 360°-turn enclosing ten atoms in a ring by formation of $i + 3 \rightarrow i$ hydrogen bonds) that fits perfectly into the opening of the central cavity, similar to the H3 peptide. This helix is stabilized by intramolecular main-chain hydrogen bonds between Ser3763 and Ala3766 as well as between Ala3764 and Glu3767. Furthermore, the hydrogen bond between the carboxylate side chain of Glu3767 and the amide nitrogen of Gly3762 gives the sequence a cyclical structure that entails the 3_{10} -helix to fit into the outer opening of the central cavity (Figure I-11 A). Intermolecular contacts of MLL1 to WDR5 are facilitated by sandwiching of the Arg3763 of MLL1 between two Phe13 and 263 residues of WDR5. In addition, an extensive network of (water mediated) hydrogen bonds and van der Waals interaction between the *Win* peptide and WDR5 is stabilizing the binding of the 3_{10} -helix to the WDR5 cavity. Thus, MLL1 Arg3763 displays multiple contacts to residues Ser91, Phe133, Ser175, Ser218, Cys261, Phe263 and Ile305 of WDR5 as well as Gly3762 and Ser3763 of MLL1 interact with WDR5 residues Ala47, Ala65, Gly89, Ile90, Ser91, Asp107.^{126,127}

In contrast, the binding site for RbBP5 is located on the bottom of WDR5, in proximity of a V-shaped cleft between blade 5 and 6. The crystal structure of a RbBP5 peptide (³⁷²H₂N-DEEVDVTSV-NH₂; $K_i = 2.2$ μ M) bound to WDR5 revealed that RbBP5 residues responsible for RbBP5-WDR5 affinity are Glu374, Val375, Asp376, Thr378 and Ser379, while WDR5 residues involved in RbBP5 recognition are Ser223, Pro224, Asn225, Tyr228, Leu240, Leu249, Lys250, Phe266, Leu288, Gln289 (Figure I-11 B).^{128,129}

The interaction of RbBP5 with the SPRY domain (the primary moiety in Ash2L that recognizes RbBP5, name based on sequence repeat of SP1A kinase and RYanodine receptor) of Ash2L is based on RbBP5 residues 344-363. Crystal structure analysis of a RbBP5 peptide (³⁴⁷H₂N-ERESEFDIE-NH₂; $K_d = 750$ nM, determined by sedimentation experiments using full length RbBP5-Ash2L proteins¹³⁰) to the Ash2L SPRY domain revealed that Glu349 and Asp353 of RbBP5 are sitting on two arginine residues Arg343 and Arg367 at the center of the basic pocket of Ash2L SPRY. The SPRY domain, thus, adopts a twisted β -sandwich composed of two antiparallel β -sheets. Additional Ash2L residues involved in RbBP5 contacts are Tyr313, Ala355, Pro356, Tyr359, Ile374, Ser377, Ile378, and Tyr475 (Figure I-11 B).^{131,132}

Binding interfaces between the RbBP5-Ash2L dimer and MLL1 were found to be located in the region between 3883-3904 of MLL1. Specifically, interactions of RbBP5 residue Phe336 to MLL1 residues Tyr3883, Arg3903 and Phe3904 were identified. Here, Arg3903 sticks towards an acidic pocket formed by RbBP5-Ash2L SPYR (residues RbBP5: Glu347; Ash2L: Tyr313, Gln354. (Figure I-11 B).¹³² In 2019, Kaustov et al.¹³³ identified a previously unknown binding region between MLL1 and RbBP5 using a combination of X-ray scattering (SAXS), cross-linking mass spectrometry (XL-MS), NMR spectroscopy and computational modelling. This new RbBP5-binding region of MLL1 is not conserved in the other MLL family members and unique for MLL1. The novel interaction site to MLL1 was proposed to be located in the N-terminal region of the WD40 repeat domain of RbBP5 (via hydrophobic RbBP5

residues: Val249, Ile283, Leu286, Val287, Ile289, Gln273, Tyr277, Pro253), making several contacts to MLL1 residues, located between the *Win* motif and the SET-I domain (³⁷⁸⁶RQPPEYN; called RBS (RbBP5 Binding Sequence) motif). The interaction of the MLL1_{RBS-SET1} construct (3785-3969) and RbBP5_{NTD} construct (residues 10-340) was estimated to display affinities in the μM region ($K_d = \sim 8 \mu\text{M}$) using BLI (biolayer interferometry).¹³³

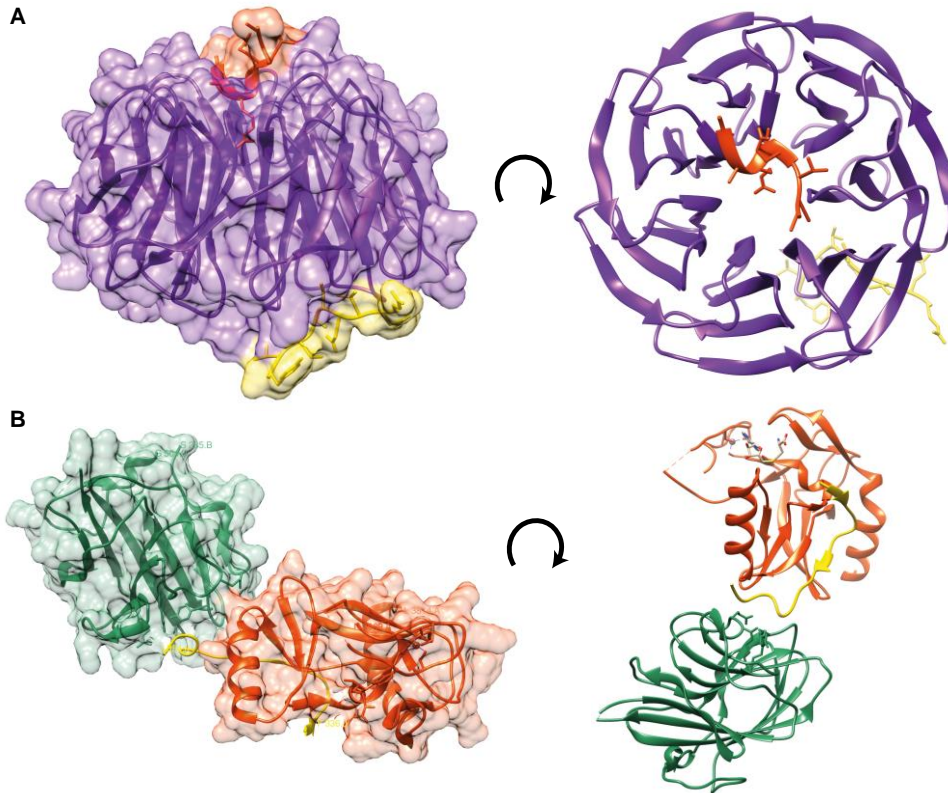


Figure I-11. A) Structure of WDR5 (purple) bound to MLL1 *Win* (orange) and RbBP5 peptide (yellow). MLL1 is bound to a pocket, located at center of WD40 β -propeller, while RbBP5 peptide is bound on opposite site between blade 5 and 6; (PDB: 3P4F¹²⁸). **B)** Structure of MLL1-Ash2L-RbBP5 interface displaying ternary complex; orange = MLL1, yellow = RbBP5, green = Ash2L (PDB: 5F6L¹³²).

Comparison of MLL1 to other MLL family members, revealed that MLL1 displays a unique dependency on WDR5 binding, in order to trigger its methylation activity.¹³² Thus, it has been demonstrated that MLL1 alone, only exhibits $\sim 1\%$ activity, while the MLL1-RbBP5-Ash2L 3mer displays $\sim 20\%$ and addition of WDR5 increases activity of the 4mer complex to $\sim 100\%$. In contrast, the activity of other MLL family members is only dependent on the formation of the 3mer complex via MLL-interaction with RbBP5-Ash2L. For MLL3 it was even revealed that addition of WDR5 is hampering methylation activity (MLL3 alone $< 5\%$; MLL3-RbBP5-Ash2L $\sim 88\%$, MLL3-RbBP5-Ash2L-WDR5 $\sim 80\%$). This hindered methylation activity of MLL3 upon WDR5 addition could be structurally interpreted by comparison of the two cryo-EM structures of MLL1 and MLL3. Thus, different orientations of the β -propellers of WDR5 and RbBP5, accompanied by changes in the organization of interfaces between WDR5-RbBP5-MLL components were revealed. For MLL1, it was validated that the, so called, activation segments of RbBP5 (residues 329-336) and MLL1 (residues 3777-3780) are brought together via their respective interaction with WDR5. Contrarily, MLL3 does not display a counterpart of this activation motif, where WDR5 engages direct interactions to the N-terminal residues of MLL3 (Figure I-12). These differences between binding of MLL1 and MLL3 to WDR5 and RbBP5 might be an explanation for the hampered methylation activity of MLL3 in presence of WDR5, while for MLL1 the interaction to WDR5 is essential.¹¹⁸

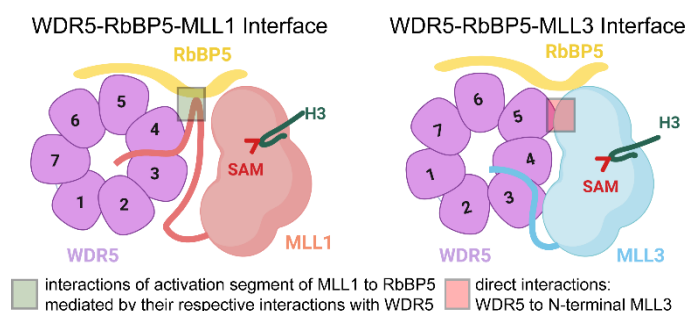


Figure I-12. Schematic representation of different interacting modes of MLL1 and MLL3 with WDR5 and RbBP5, giving structural explanations for the different outputs between MLL1 and MLL3 upon WDR5 binding. Fig. inspired by Xue et al.¹¹⁸

However, Li et al.¹³² demonstrated that WDR5 is not directly involved in the activation of MLL1 but only needed to facilitate binding of MLL1 to the RbBP5-Ash2L dimer. They proved this by fusing MLL1 to RbBP5, which led to high methylation activity of the MLL1 3mer complex where further stimulation via WDR5 was only marginal (activity of MLL1-RbBP5 fusion <5%, MLL1-RbBP5 fusion + Ash2L ~88%, MLL1-RbBP5 fusion + Ash2L + WDR5 ~100%). The addition of DPY30 to the 4mer-complex did not increase methylation activity any further.¹³⁴ Although it was demonstrated that DPY30 did not further enhance activity of MLL1 utilizing this in vitro assay and the H3-substrate, it is known that deletion of DPY30 leads to a reduction of global H3K4 methylation,¹³⁵ to defects in lineage specification,¹³⁵ severe effects on the proliferation and differentiation of hematopoietic progenitor cells¹³⁶ as well as it alters identity and function of adult hematopoietic stem cells.¹³⁷ To find an explanation for the difference between MLL1 and the other family members, studies on the example of MLL1 vs. MLL3 were performed. Structural investigations revealed that the two residues of MLL1 Gln3867 and Asn3861, which are differing from the residues of MLL3, are responsible for the lower binding affinity of MLL1 to RbBP5, because of their bulkiness and hydrophilicity. Therefore, WDR5 is required to facilitate binding of MLL1 to RbBP5. Additionally, normal mode analysis was utilized to suggest a highly dynamic motion of the SET-I motif of isolated MLL enzymes, which is suppressed upon RbBP5-Ash2L binding. This reduced flexibility might contribute to cofactor binding and substrate recognition of MLL. Since for MLL1, WDR5 is essential to bring the RbBP5-Ash2L dimer into the correct position to achieve such SET-I stabilization, the MLL1-WDR5 interaction is of such importance to maintain MLL1 enzymatic activity. This unique dependence of MLL1 activation on WDR5, compared to the other MLL family members, opens the avenue for specific inhibition of MLL1 methyl-transferase activity by blocking its interaction to WDR5 using small WDR5 high affinity binders.

1.5 Inhibitors of the MLL1-WDR5 Interaction

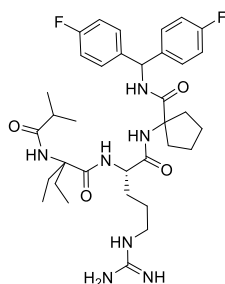
As explained in section 1.4, the inhibition of the MLL1-WDR5 interaction provides a promising therapeutic cancer target, due to the unique dependence of MLL1 catalytic activity on WDR5 binding. After the structural investigations of the H3⁹¹⁻⁹⁴ and *Win* peptide^{126,127} to WDR5, the group of Dou was evaluating the essential elements of the *Win* peptide (Ac-GSARAEVHLRKS-NH₂) that are required for high affinity WDR5 binding, in order to enable the design of small molecule inhibitors.¹³⁸ Thus, they determined the binding affinities of a battery of truncated *Win* peptide fragments to WDR5 using a fluorescence polarization (FP)-based assay (Table I-2). It was demonstrated that the two peptides: Ac-10mer and NH₂-11mer exhibit increased WDR5 affinities compared to the parental *Win* peptide. The two 3-mer peptides: Ac-ARA-NH₂ and Ac-ART-NH₂ based on the *Win* and H3 peptide, respectively,

demonstrated to be the minimal motif, required for high affinity WDR5 binding. The H3-Thr derivative displayed one order of magnitude tighter binding than its *Win*-Ala analogue. It was highlighted that the high binding affinities are based on two intramolecular hydrogen bonds formed within this motif (H-bonds between: backbone carbonyl oxygen N-terminal of Ala3 contacts with backbone NH of Ala5; backbone carbonyl oxygen of Ala3 contacts with backbone NH C-terminal of Ala5). Absence of one of these hydrogen bonds, led to decreased binding affinity. Thus, truncating the Ala from the N-terminus resulted in a decrease of 1500 times, while removal of amino acids from the C-terminus were tolerated up to the second Ala, maintaining the essential –CO-ARA-NH– sequence.¹³⁸ Previously, binding affinity of the original H3 peptide (H₂N-ARTKQTARKSTGGKAPRKQ-NH₂; $K_d = 6.6 \mu\text{M}$ (ITC)⁹¹) to WDR5 was demonstrated to be lower than of the corresponding *Win* peptide ($K_i = 160 \text{ nM}$ (FP)¹³⁸) as it is lacking one of the essential hydrogen bonds. However, Karatas et al.¹³⁸ revealed that by reestablishment of such hydrogen bond via acetylation of a 10mer- and 3mer H3 peptide, affinities are increased. Even tighter binding than the corresponding *Win* analogues was detected (Table I-2). This addition of an N-terminal carbonyl group next to the Ala1 substantiates the importance of the two intramolecular hydrogen bonds for high affinity binding. The increased affinity of acetylated H3-peptides over *Win* analogues led to the investigation of FAM-labelled *Win*-peptides, where the second Ala was replaced by a Thr in order to increase binding affinities. Such peptides have been used as tracers for the FP-based assays.¹³⁸ The developed Win-6mer peptide (Table I-2) was utilized to study its selectivity for MLL1, compared to the other MLL family members. This study revealed that WDR5 binding is only disrupting the methylation activity of MLL1 and SET-1A, while the activity of the remaining MLL family members was maintained. In addition, high WDR5-binding affinity is preserved, as such peptide displays similar binding modes than previously reported *Win* analogues, where additional hydrogen bonding is observed by its Tyr residue. Based on the previous findings, in 2013, Karatas et al.¹³⁹ pursued the design of the first set of small molecule peptidomimetics as inhibitors for the WDR5-MLL1 interaction. Initially, they performed an amino acid mutation scan of the Ac-ARA-NH₂ motif, which revealed that mutation of the first as well as the second Ala with aminobutyric acid (Abu) or Val, led to the largest affinity increases (K_i : Ac-AbuRA-NH₂ = $60 \pm 1.0 \text{ nM}$; Ac-VRA-NH₂ = $50 \pm 5.0 \text{ nM}$; Ac-ARAbu-NH₂ = $6.0 \pm 2.0 \text{ nM}$; Ac-ARV-NH₂ = $20 \pm 3.0 \text{ nM}$). Further exploration of mutating these positions with unnatural amino acids that exert small condensed hydrophobic side chains as well as of the N- and C-termini led to compound MM-102, the most promising candidate at the time ($K_i < 1 \text{ nM}$; Table I-2). It was demonstrated that MM-102 is able to inhibit methylation activity of MLL1, to downregulate transcription of MLL1 target genes *hoxA9* and *meis-1*, as well as to specifically inhibit cell viability of MV4-11 leukemia cells.¹³⁹ To improve cellular uptake and microsomal stability, conformationally constrained macrocyclic analogues were prepared.^{140,141} To date, the most promising candidates of small molecule peptidomimetics are the macrocyclic peptides MM-401 and MM-589 (Table I-2), whose design was based on the previously reported MM-102/ MM-101.¹³⁹ Further stabilization was derived from π - π stacking interactions of the phenyl residues to Tyr260 of WDR5, as well as from additional WDR5 interactions of the Abu residues of the two macrocycles. In MM-589 advanced stabilization was gained from hydrophobic interactions of the methyl group, next to the arginine residue. Thus, enhanced WDR5 affinity can be attributed to those additional hydrophobic interactions, as well as to the restricted conformation of the cyclic derivatives. Both, MM-401 and MM-589, were proven to be specific for MLL1, while they did not show any effects in the other MLL family members. Thus, they are specifically inhibiting MLL1 HMT-activity, blocking proliferation of MLL cells, inducing cell cycle arrest, apoptosis and myeloid differentiation without toxicity for non-MLL cells.¹⁴⁰ MM-401 was additionally proven to be able to reprogram epiblast stem cells to naive pluripotency¹⁴¹ as well as its application can cause NSCs (neural stem cells) to stop neuron production.¹⁴²

Table I-2. Peptide inhibitors for WDR5-MLL1 interaction, displaying their binding affinities to WDR5 and the used technique as well as further investigations applied for the respective inhibitor (FP: Fluorescence Polarization, ITC: isothermal calorimetry, HMT-activity: histone methyltransferase activity of MLL1, ^a: radiometric assay, ^b: AlphaLISA assay; Ahx: aminohexanoic acid, Abu: aminobutyric acid; ^c: 95% confidence intervals).

Peptide	Sequence/ Structure	Binding Affinity	Further Investigations
Win ¹³⁸	Ac-GSARAEVHLRKS-NH ₂	160 ± 20 nM (<i>K_i</i> , FP)	PDB: 3EG6 ¹²⁶
Ac-11mer ¹³⁸	Ac-SARAEVHLRKS-NH ₂	200 ± 30 nM (<i>K_i</i> , FP)	-
H ₂ N-11mer ¹³⁸	H ₂ N-SARAEVHLRKS-NH ₂	20 ± 2.0 nM (<i>K_i</i> , FP)	-
Ac-10mer ¹³⁸	Ac-ARAEVHLRKS-NH ₂	3.0 ± 1.0 nM (<i>K_i</i> , FP)	-
Ac-9mer ¹³⁸	Ac-RAEVHLRKS-NH ₂	6300 ± 800 nM (<i>K_i</i> , FP)	-
Ac-7mer ¹³⁸	Ac-ARAEVHL-NH ₂	30 ± 1.0 nM (<i>K_i</i> , FP)	-
Ac-3mer ¹³⁸	Ac-ARA-NH ₂	120 ± 10 nM (<i>K_i</i> , FP)	-
Ac-2mer ¹³⁸	Ac-AR-NH ₂	27000 ± 1400 nM (<i>K_i</i> , FP)	-
H ₂ N-H3-10mer ¹³⁸	H ₂ N-ARTKQTARKS-NH ₂	15100 ± 130 nM (<i>K_i</i> , FP)	-
Ac-H3-10mer ¹³⁸	Ac-ARTKQTARKS-NH ₂	< 1.0 nM (<i>K_i</i> , FP)	PDB: 2CO0 ⁹⁴
H ₂ N-H3-3mer ¹³⁸	H ₂ N-ART-NH ₂	27300 ± 2500 nM (<i>K_i</i> , FP)	-
Ac-H3-3mer ¹³⁸	Ac-ART-NH ₂	20 ± 1.0 nM (<i>K_i</i> , FP)	-
H3-peptide ¹⁴³	H-ARTKQA-NH ₂	1100 ± 100 nM (<i>K_d</i> , ITC)	-
Ac-H3 ¹⁴³	Ac-ARTKQA-NH ₂	130 ± 20 nM (<i>K_d</i> , ITC)	PDB: 3PSI HMT-activity ^a : IC ₅₀ = 4.1 ± 0.9 μM
10mer-Ala-FAM ¹³⁸	Ac-ARAEVHLRKS-Ahx-Ahx-K(FAM)-NH ₂	14 ± 2.0 nM (<i>K_d</i> , FP)	-
10mer-Thr-FAM ¹³⁸	Ac-ARTEVHLRKS-Ahx-Ahx-K(FAM)-NH ₂	1.0 ± 0.3 nM (<i>K_d</i> , FP)	-
Win6-mer ¹⁴⁴	Ac-ARTEVY-NH ₂	2.9 nM (1.7-4.2 ^c) (<i>K_d</i> , ITC)	PDB: 5SXM HMT-activity ^a : IC ₅₀ = 2.3 μM (1.6-3.3)
Ac-Abu-3mer ^{139,138}	Ac-ARAbu-NH ₂	6.0 ± 2.0 nM (<i>K_i</i> , FP)	-

MM-102¹³⁹

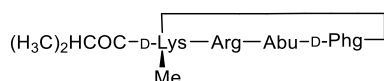


<1 nM (K_i , FP)

PDB: 4GM8

HMT-activity^a:
 $IC_{50} = 400 \pm 100$ nM
 cell viability MV4-11:
 $GI_{50} = 25$ μ M

MM-401¹⁴⁰

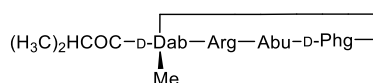


<1 nM (K_i , FP)

PDB: 4GM9

HMT-activity^a: $IC_{50} = 320$ nM;
 cell viability:
 MLL-AF9: $GI_{50} = 9.8 \pm 0.04$ μ M
 MV4-11: $GI_{50} = 12.4$ μ M
 MOLM13: $GI_{50} = 23.9$ μ M

MM-589¹⁴⁵



<1 nM (K_i , FP)

PDB: 5VFC

HMT-activity^b:
 $IC_{50} = 12.7 \pm 1.5$ nM
 cell viability MV4-11:
 $GI_{50} = 0.25 \pm 0.01$ μ M
 MOLM13:
 $GI_{50} = 0.21 \pm 0.02$ μ M

Besides peptidic WDR5-binding inhibitors, an approach utilizing WDR5-targeting monoclonal antibodies has been established in order to inhibit the WDR5-MLL1 interaction. Monoclonal antibodies are synthetic binding proteins utilized to accelerate target validation. The most promising variant displayed nanomolar WDR5 affinity ($K_d = 5 \pm 1$ nM), determined via titration using flow cytometry. This monoclonal antibody was additionally applied in an in vivo system, indicating that MLL1-WDR5 inhibition delays the onset of MLL-AF9 leukemia and increases the survival benefit in vivo.¹⁴⁶

In 2013, the first small molecule inhibitors for the WDR5-MLL1 interaction have been developed by Sinisterra et al.¹⁴⁷ They identified compounds through initial screening of a large library of small molecules (16000) upon their displacement ability of a FAM-labelled *Win* peptide from the WDR5 cavity. Following, a structure query in a database of over 6 million unique commercial molecules was performed, leading to two hit compounds that displayed direct binding to WDR5 (WDR5-103 **I-11**, WDR5-102 **I-12**). The candidate with the highest binding affinity (WDR-103 **I-11**) was further evaluated upon its ability to inhibit methylation activity of MLL1 (Table I-3). Structurally, the protonated N-methyl-piperazine moiety is inserting into the central WDR5 Arg pocket, forming a hydrogen bond with Cys261, while the benzamide ring is occupying the upper shallow region of the WDR5 cavity.¹⁴⁷ Further crystal structure guided optimizations of WDR-102 led to the identification of an improved WDR5 antagonist: WDR5-47 **I-13**.¹⁴⁸ There, key structural features as the central phenyl ring, the C-1 position amide and C-2 position 4-methylpiperazine were maintained, while additional introduction of a 4-fluoro and 3-methyl residue into the benzamide phenyl ring were critical features for increased binding affinity. Thus, the C-1 amide bond exerted hydrogen bonds to WDR5 side chains of Cys261 and Ser91 and the protonated 4-methylpiperazine displayed an essential water-mediated hydrogen bond to

Cys261. Since the C-1 amide as well as C-2 4-methylpiperazine substituent were crucial to maintain high WDR5 affinity, additional improvement and adjustment of C-1 and C-5 substituents on the benzamide core were performed. This led to the preparation of OICR-9429 **I-14** by Getlik et al.¹⁴⁹, the first highly potent small molecule antagonist for WDR5. In addition to its nanomolar binding affinity to WDR5, OICR-9429 was proven to suppress proliferation of N-terminal C/EBP α p30-expressing human AML cells.^{149,150} The increased WDR5 affinity mainly arises from novel direct hydrogen bonds of the pyridinone moiety to Asp107 and Ala65 as well as the trifluoromethyl group making additional van der Waals contacts to WDR5. Similar optimization studies, maintaining the same substituents and different alteration at C-1 and C-5 residues, performed by Li et al.^{151,152} resulted in the discovery of DDO-2117 **I-15**, a WDR5 antagonist with even higher affinity than OICR-9429. Additional interactions to WDR5 were claimed to arise from two hydrogen bonds of the added amine group to Asp107, as well as due to hydrophobic interactions of the methyl-, chloro- and fluoro substituents. DDO-2117 displayed abilities to inhibit HMT-activity of MLL1, as well as capabilities to repress MV4-11 leukemia cell viability (Table I-3).¹⁵¹

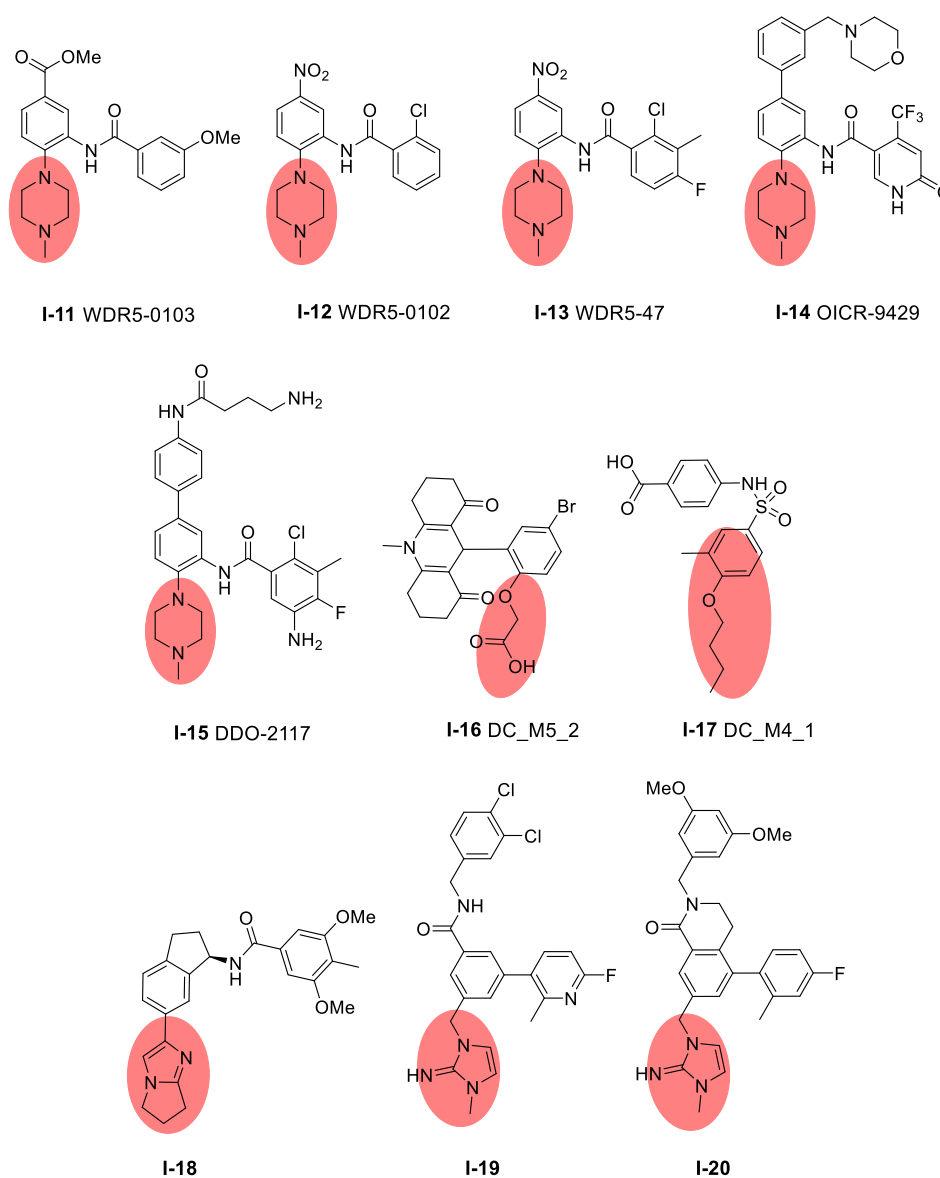


Figure I-13. Structures of discussed small molecule MLL1-WDR5 inhibitors **I-11** – **I-20**. The respective residue inserting into the WDR5 Arg binding cavity is highlighted in red.

Table I-3. Small molecule inhibitors for WDR5-MLL1 interaction, displaying their binding affinities to WDR5 and further investigations applied for the respective inhibitor (FP: Fluorescence Polarization, ITC: isothermal calorimetry, SPR: surface plasmon resonance, HMT-activity: histone methyltransferase activity of MLL1, ^a: radiometric assay).

Molecule	Binding Affinity	Further Investigations
WDR5-0103 ¹⁴⁷	$K_d = 450 \pm 20$ nM (ITC) $K_i = 3000 \pm 100$ nM (FP)	PDB: 3UR4; HMT-activity ^a : $IC_{50} = 280 \pm 12$ μ M
WDR5-0102 ¹⁴⁷	$K_d = 4000 \pm 1.1$ nM (ITC) $K_i = 11000 \pm 300$ nM (FP)	PDB: 3SMR
WDR5-47 ¹⁴⁸	$K_i = 270 \pm 10$ nM (FP)	PDB: 4IA9
OICR-9429 ^{149,150}	$K_d = 30$ nM (SPR) $K_i = 64 \pm 4.0$ nM (FP)	PDB: 4QLI; Cell viability AML: $GI_{50} = 5.0$ μ M
DDO-2117 ¹⁵¹	$IC_{50} = 7.6 \pm 0.1$ nM (FP) $K_d = 13.6 \pm 5.9$ nM (ITC)	HMT-activity ^a : $IC_{50} = 0.19 \pm 0.01$ μ M Cell viability MV4-11: $GI_{50} = 7.4 \pm 1.4$ μ M
DC_M5_2 ¹⁵³	$IC_{50} = 9630 \pm 1460$ nM (FP) $K_d = 13800$ nM (SPR)	-
DC_M4_1 ¹⁵³	$IC_{50} = 12970 \pm 200$ nM (FP) $K_d = 15800$ nM (ITC)	Cell viability MV4-11: $GI_{50} = 31.2$ μ M
I-18 ¹⁵⁴	$K_i < 1$ nM (FP) $K_i = 0.902 \pm 0.162$ nM (TR-FRET)	PDB: 6DAS HMT-activity ^a : $IC_{50} = 3.20 \pm 0.13$ μ M Cell viability MV4-11: $GI_{50} = 7.25 \pm 1.9$ μ M
I-19 ¹⁵⁵	$K_d = 0.1$ nM (FP) $K_d = 0.044 \pm 0.008$ nM (TR-FRET)	PDB: 6E23; HMT-activity ^a : $IC_{50} = 19.9 \pm 9.6$ nM Cell viability MV4-11: $GI_{50} = 3.20 \pm 0.21$ μ M MOLM13: $GI_{50} = 6.43 \pm 0.68$ μ M
I-20 ¹⁵⁶	$K_d < 0.02$ nM (TR-FRET)	PDB: 6UCS; HMT-activity ^a : $IC_{50} = 2.2 \pm 0.28$ nM Cell viability MV4-11: $GI_{50} = 38 \pm 9.0$ nM MOLM13: $GI_{50} = 78 \pm 13$ nM

Further investigations on the scaffolds of OICR-9429 and DDO-2117 led to a novel small molecule WDR5 binder, TRPH-395, developed by *Triphase Accelerator* together with its stake holder *FACIT (Fight Against Cancer Innovation Trust)*. Preclinical trials of this compound have been performed by *Propellon Therapeutics*^{157,158} in conjunction with the *Ontario Institute for Cancer Research (OICR)*. According to Triphase, these studies revealed that the compound displays broad anti-proliferative activity across a wide range of cancer cell lines. In 2019, *Celgene* spend almost 1 billion dollar on the rights on TRPH-395, who want use it to develop first-in-class WDR5 leukemia therapy,^{159–161} highlighting the great potential of WDR5-MLL1 inhibitors. The goal of *Triphase* is to advance the novel compound through phase 2 proof-of-concept clinical studies, though, at the moment TRPH-395 is still in preclinical trials.¹⁶²

An FP-based high throughput screening approach, in order to identify novel small-molecule inhibitors for the WDR5-MLL1 interactions was performed by Ye et al.¹⁵³ The structures of the two most potent high affinity WDR5 antagonists DC_M5_2 I-16 and DC_M4_1 I-17, are based on different scaffolds (Figure I-13). Docking studies revealed that the carboxy group of DC_M5_2 I-16 takes the place of the 4-methylpiperazine of previously published WDR5 binders, by interacting with WDR5 residues Cys261

and Tyr191 via two hydrogen bonds. For DC_M4_1 **I-17**, the hydrophobic residue on the aromatic ring inserts into the WDR5 Arg cavity, while the sulfamide group establishes two hydrogen bonds to residues Asp107 and Lys46. Besides their high WDR5 affinity, surface plasmon resonance (SPR) assays verified direct interaction to WDR5 and specificity to MLL1. Compound DC_M4_1 further displayed moderate inhibitory activities in cellular MV4-11 proliferation assays.¹⁵³

In 2018, a distinct lead series of small molecule WDR5 binders based on 6,7-dihydro-5-H-pyrrolo[1,2-a]imidazole fragments, was developed.¹⁵⁴ There, the cyclic imidazole group is responsible for binding the Arg cavity of WDR5, where the N-3 imidazole nitrogen similarly exerts hydrogen bonding to the carbonyl backbone of the Cys261. In addition, the methoxy and methyl groups of **I-18** exert favorable interactions to WDR5 Phe149, Pro173 and Tyr 191. Aho et al.¹⁵⁵ independently developed novel small molecule binders via fragment-based approaches and structure-based design. This led to the discovery of a highly potent WDR5-MLL1 inhibitor **I-19** with low nanomolar affinity to WDR5 ($K_d = 0.044$ nM (TR-FRET)). In this case, the cyclic guanidine binds into the Arg pocket of WDR5, exerting π - π stacking interactions with the Phe133 and 263, as well as the carbonyl oxygen forms a hydrogen bond with the backbone NH of the Cys261. Further stabilization is provided by hydrophobic interactions of the biaryl moiety. In addition to its high WDR5 binding affinity, it was demonstrated that this compound specifically inhibits methylation activity of MLL1, represses leukemia cell viability, displaces WDR5 from chromatin and decreases associated gene expression, which leads to nucleolar stress and p53 induction.¹⁵⁵ Additional optimizations of **I-19** resulted in compound **I-20**,¹⁵⁶ which displays even further improved binding affinity to WDR5 ($K_d < 0.02$ nM (TR-FRET)). The advanced features originate from the novel dihydroisoquinolinone bicyclic core, that was designed using X-ray structure guided modifications. Its increased binding affinity arises from preserving all favorable interactions of **I-19** to WDR5 plus additionally decreasing the conformational entropy due to fixation of the benzamide core. This novel compound displays improved specific MLL1 HMT-inhibition capacities, as well as repression of leukemia cell viability, making it the most promising small molecule candidate to date.¹⁵⁶

1.6 Light as a Trigger to Modulate Biological Functions

Light is essential to define life on earth through processes as photosynthesis and vision.¹⁶³ Due to its generally non-invasive and precise control (if applied correctly), its facile modulation, orthogonality and low costs, light further gained importance in scientific research. Thus, light is the ideal external trigger to gain spatiotemporal control over biological functions. In this context, light has been used as a tool in photopharmacology^{164–167} and to photoactivate photosensitive molecules,^{166,168–172} photocaged molecules^{173–177} or molecular photoswitches.^{164,178–187} Besides small photoactivatable molecules, photocontrollable peptides have emerged as versatile tools, especially to study protein-protein interactions (PPIs).^{180,188} The reader is here referred to Section 4.1 “Photoswitchable Peptides for Spatiotemporal Control of Biological Functions” in the cumulative part of this thesis. Our review¹⁸⁹ summarizes the key properties of molecular photoswitches, recapitulates structural studies of peptides linked to a photoswitchable moiety and provides a detailed overview of photoresponsive peptides used to study biological processes.

1.7 References

1. Waddington, C. H. The epigenotype. *Endeavour* **1**, 18–20 (1942).
2. Speybroeck, L. van. Linearization , Optimization , Protection Page inserted by evaluation version From Epigenesis to Epigenetics. *Ann. N.Y. Acad. Sci.* **981**, 61–81 (2002).
3. Waddington, C. H. A Discussion of Some Aspects of Theoretical Biology. in *The Strategy of the Gene* (1957).
4. Richards, E. J. Revisiting Soft Inheritance. *Nat. Rev. Genet.* **7**, 395–402 (2006).
5. Lacal, I. & Ventura, R. Epigenetic Inheritance: Concepts, Mechanisms and Perspectives. *Front. Mol. Neurosci.* **11**, 1–22 (2018).
6. Horsthemke, B. A critical view on transgenerational epigenetic inheritance in humans. *Nat. Commun.* **9**, 1–4 (2018).
7. Jablonka, E. & Lamb, M. J. PDFlib PLOP : PDF Linearization , Optimization , Protection Page inserted by evaluation version The Changing Concept of Epigenetics. *Ann. N.Y. Acad. Sci.* **981**, 82–96 (2002).
8. Non, A. L. & Thayer, Z. M. *Epigenetics and Human Variation. A Companion to Anthropological Genetics* (2019).
9. Fujita, R. *et al.* Nucleosome destabilization by nuclear non-coding RNAs. *Commun. Biol.* **3**, 60 (2020).
10. Pugh, J. E. & Holliday, R. Do chemical carcinogens act by altering epigenetic. *Heredity (Edinb).* **40**, 329 (1978).
11. Holliday, R. A new theory of carcinogenesis. *Br. J. Cancer* **40**, 513–522 (1979).
12. Holliday, R. The Inheritance of Epigenetic Defects. *Science* **238**, 163–170 (1987).
13. Holliday, R. Epigenetics: A historical overview. *Epigenetics* **1**, 76–80 (2006).
14. Feinberg, A. P. The key role of epigenetics in human disease prevention and mitigation. *N. Engl. J. Med.* **378**, 1323–1334 (2018).
15. Tzika, E., Dreker, T. & Imhof, A. Epigenetics and metabolism in health and disease. *Front. Genet.* **9**, (2018).
16. Grau-Perez, M., Agha, G., Pang, Y., Bermudez, J. D. & Tellez-Plaza, M. Mendelian Randomization and the Environmental Epigenetics of Health: a Systematic Review. *Curr. Environ. Heal. reports* **6**, 38–51 (2019).
17. Cavalli, G. & Heard, E. Advances in epigenetics link genetics to the environment and disease. *Nature* **571**, 489–499 (2019).
18. Berdasco, M. & Esteller, M. Clinical epigenetics: seizing opportunities for translation. *Nat. Rev. Genet.* **20**, 109–127 (2019).
19. Mastroeni, D. *et al.* Epigenetic changes in Alzheimer’s disease: Decrements in DNA methylation. *Neurobiol. Aging* **31**, 2025–2037 (2010).
20. Rodenhiser, D. & Mann, M. Epigenetics and human disease: translating basic biology into clinical applications. *C. Can. Med. Assoc. J.* **174**, 341–348 (2006).
21. Bennett, R. L. & Licht, J. D. Targeting Epigenetics in Cancer. *Annu. Rev. Pharmacol. Toxicol.* **58**, 187–207 (2018).
22. Bögershausen, N., Bruford, E. & Wollnik, B. Skirting the pitfalls: A clear-cut nomenclature for H3K4 methyltransferases. *Clin. Genet.* **83**, 212–214 (2013).

23. Armstrong, L. *Epigenetics*. (Garland Science, 2014).
24. Kornberg, R. D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285–294 (1999).
25. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* **49**, 1–22 (2017).
26. Avvakumov, N., Nourani, A. & Côté, J. Histone Chaperones: Modulators of Chromatin Marks. *Mol. Cell* **41**, 502–514 (2011).
27. Burgess, R. J. & Zhang, Z. Histone chaperones in nucleosome assembly and human disease. *Nat. Struct. Mol. Biol.* **20**, 14–22 (2013).
28. Narlikar, G. J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* **154**, 490–503 (2013).
29. Harshman, S. W., Young, N. L., Parthun, M. R. & Freitas, M. A. H1 histones: Current perspectives and challenges. *Nucleic Acids Res.* **41**, 9593–9609 (2013).
30. Klose, R. J. & Bird, A. P. Genomic DNA methylation: The mark and its mediators. *Trends Biochem. Sci.* **31**, 89–97 (2006).
31. Kohli, R. M. & Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**, 472–479 (2013).
32. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381–395 (2011).
33. Greer, E. L. & Shi, Y. Histone methylation: A dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13**, 343–357 (2012).
34. Hebbes, T. R., Thorne, A. W. & Crane-Robinson, C. A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* **7**, 1395–1402 (1988).
35. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349–352 (1997).
36. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
37. Barski, A. *et al.* High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129**, 823–837 (2007).
38. Wang, Z. *et al.* Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019–1031 (2009).
39. Lee, J. S., Smith, E. & Shilatifard, A. The Language of Histone Crosstalk. *Cell* **142**, 682–685 (2010).
40. Rothbart, S. B. & Strahl, B. D. Interpreting the language of histone and DNA modifications. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1839**, 627–643 (2014).
41. Dai, H. & Wang, Z. Histone Modification Patterns and Their Responses to Environment. *Curr. Environ. Heal. Reports* **1**, 11–21 (2014).
42. Dhalluin, C. *et al.* Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496 (1999).
43. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).
44. Nielsen, P. R. *et al.* Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**, 103–107 (2002).

45. Jacobs, S. A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080–2083 (2002).
46. Li, H. *et al.* Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91–95 (2006).
47. Peña, P. V. *et al.* Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* **442**, 100–103 (2006).
48. Shi, X. *et al.* ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**, 96–99 (2006).
49. Wysocka, J. *et al.* A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86–90 (2006).
50. Andrews, F. H., Strahl, B. D. & Kutateladze, T. G. Insights into newly discovered marks and readers of epigenetic information. *Nat. Chem. Biol.* **12**, 662–668 (2016).
51. Huang, H., Lin, S., Garcia, B. A. & Zhao, Y. Quantitative proteomic analysis of histone modifications. *Chem. Rev.* **115**, 2376–2418 (2015).
52. Stillman, B. Histone Modifications: Insights into Their Influence on Gene Expression. *Cell* **175**, 1–4 (2018).
53. Rice, J. C. *et al.* Histone Methyltransferases Direct Different Degrees of Methylation to Define Distinct Chromatin Domains. *Mol. Cell* **12**, 1591–1598 (2003).
54. Santos-Rosa, H. *et al.* Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407–411 (2002).
55. Ruthenburg, A. J., Allis, C. D. & Wysocka, J. Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. *Mol. Cell* **25**, 15–30 (2007).
56. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318 (2007).
57. Kim, T. S. & Buratowski, S. Dimethylation of H3K4 by Set1 Recruits the Set3 Histone Deacetylase Complex to 5' Transcribed Regions. *Cell* **137**, 259–272 (2009).
58. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
59. Kang, J. Y. *et al.* KDM2B is a histone H3K79 demethylase and induces transcriptional repression via sirtuin-1-mediated chromatin silencing. *FASEB J.* **32**, 5737–5750 (2018).
60. Trievel, R. C., Beach, B. M., Dirk, L. M. A., Houtz, R. L. & Hurley, J. H. Structure and catalytic mechanism of a SET domain protein methyltransferase. *Cell* **111**, 91–103 (2002).
61. Southall, S. M., Wong, P. S., Odho, Z., Roe, S. M. & Wilson, J. R. Structural Basis for the Requirement of Additional Factors for MLL1 SET Domain Activity and Recognition of Epigenetic Marks. *Mol. Cell* **33**, 181–191 (2009).
62. Rice, J. C. & Allis, C. D. Histone methylation versus histone acetylation: New insights into epigenetic regulation. *Curr. Opin. Cell Biol.* **13**, 263–273 (2001).
63. Dillon, S. C., Zhang, X., Trievel, R. C. & Cheng, X. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol.* **6**, 227 (2005).
64. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).
65. Ciccone, D. N. *et al.* KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* **461**, 415–418 (2009).

66. Lin, Y. *et al.* The SNAG domain of snail1 functions as a molecular hook for recruiting lysine-specific demethylase 1. *EMBO J.* **29**, 1803–1816 (2010).
67. Kong, X. *et al.* Catalytic mechanism investigation of lysine-specific demethylase 1 (LSD1): A computational study. *PLoS One* **6**, 1–11 (2011).
68. Sinha, K. M., Yasuda, H., Coombes, M. M., Yr Dent, S. & De Crombrughe, B. Regulation of the osteoblast-specific transcription factor Osterix by NO66, a Jumonji family histone demethylase. *EMBO J.* **29**, 68–79 (2010).
69. Christensen, J. *et al.* RBP2 Belongs to a Family of Demethylases, Specific for Tri- and Dimethylated Lysine 4 on Histone 3. *Cell* **128**, 1063–1076 (2007).
70. Yamane, K. *et al.* PLU-1 Is an H3K4 Demethylase Involved in Transcriptional Repression and Breast Cancer Cell Proliferation. *Mol. Cell* **25**, 801–812 (2007).
71. Tahiliani, M. *et al.* The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. *Nature* **447**, 601–605 (2007).
72. Min, G. L. *et al.* Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **318**, 447–450 (2007).
73. Yun, M., Wu, J., Workman, J. L. & Li, B. Readers of histone modifications. *Cell Res.* **21**, 564–578 (2011).
74. Sprangers, R., Groves, M. R., Sinning, I. & Sattler, M. High-resolution X-ray and NMR structures of the SMN Tudor domain: Conformational variation in the binding site for symmetrically dimethylated arginine residues. *J. Mol. Biol.* **327**, 507–520 (2003).
75. Nielsen, P. R. *et al.* Structure of the chromo barrel domain from the MOF acetyltransferase. *J. Biol. Chem.* **280**, 32326–32331 (2005).
76. Wang, W. K. *et al.* Malignant brain tumor repeats: A three-leaved propeller architecture with ligand/peptide binding pockets. *Structure* **11**, 775–789 (2003).
77. Hung, T. *et al.* ING4 Mediates Crosstalk between Histone H3 K4 Trimethylation and H3 Acetylation to Attenuate Cellular Transformation. *Mol. Cell* **33**, 248–256 (2009).
78. Su, X. *et al.* Molecular basis underlying histone H3 lysine – arginine methylation pattern readout by Spin / Ssty repeats of Spindlin1. *Genes Dev.* **28**, 622–636 (2014).
79. Musselman, C. A. & Kutateladze, T. G. Handpicking epigenetic marks with PHD fingers. *Nucleic Acids Res.* **39**, 9061–9071 (2011).
80. Jain, K. *et al.* Characterization of the plant homeodomain (PHD) reader family for their histone tail interactions. *Epigenetics Chromatin* **13**, 3 (2020).
81. Qian, S. *et al.* Dual recognition of H3K4me3 and H3K27me3 by a plant histone reader SHL. *Nat. Commun.* **9**, 1–11 (2018).
82. Sanchez, R. & Zhou, M. M. The PHD finger: A versatile epigenome reader. *Trends Biochem. Sci.* **36**, 364–372 (2011).
83. Schapira, M., Tyers, M., Torrent, M. & Arrowsmith, C. H. WD40 repeat domain proteins: A novel target class? *Nat. Rev. Drug Discov.* **16**, 773–786 (2017).
84. Wall, M. A. *et al.* The structure of the G protein heterotrimer G α 1 β 1 γ 2. *Cell* **83**, 1047–1058 (1995).
85. Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E. & Sigler, P. B. Crystal structure of a G α protein $\beta\gamma$ dimer at 2.1Å resolution. *Nature* **379**, 369–374 (1996).
86. Hamm, H. E. Structure and function of a heterotrimeric G protein. *FASEB J.* **10**, 311–319 (1996).

87. Sprague, E. R., Redd, M. J., Johnson, A. D. & Wolberger, C. Structure of the C-terminal domain of Tup1, a corepressor of transcription in yeast. *EMBO J.* **19**, 3016–3027 (2000).
88. Migliori, V., Mapelli, M. & Guccione, E. On WD40 proteins. *Epigenetics* **7**, 815–822 (2012).
89. Jain, B. P. & Pandey, S. WD40 Repeat Proteins: Signalling Scaffold with Diverse Functions. *Protein J.* **37**, 391–406 (2018).
90. Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859–872 (2005).
91. Couture, J. F., Collazo, E. & Trievel, R. C. Molecular recognition of histone H3 by the WD40 protein WDR5. *Nat. Struct. Mol. Biol.* **13**, 698–703 (2006).
92. Han, Z. *et al.* Structural Basis for the Specific Recognition of Methylated Histone H3 Lysine 4 by the WD-40 Protein WDR5. *Mol. Cell* **22**, 137–144 (2006).
93. Schuetz, A. *et al.* Structural basis for molecular recognition and presentation of histone H3 By WDR5. *EMBO J.* **25**, 4245–4252 (2006).
94. Ruthenburg, A. J. *et al.* Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. *Nat. Struct. Mol. Biol.* **13**, 704–712 (2006).
95. Monier, B., Tevy, M. F., Perrin, L., Capovilla, M. & Semeriva, M. Downstream of Homeotic Genes: In the Heart of Hox function. *Fly (Austin)*. **1**, 59–67 (2007).
96. Migliori, V. *et al.* Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat. Struct. Mol. Biol.* **19**, 136–145 (2012).
97. Hsieh, J. J. D., Cheng, E. H. Y. & Korsmeyer, S. J. Taspase1: A threonine aspartase required for cleavage of MLL and proper HOX gene expression. *Cell* **115**, 293–303 (2003).
98. Shilatifard, A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr. Opin. Cell Biol.* **20**, 341–348 (2008).
99. Wysocka, J. *et al.* No Title. *Nature* **442**, 86–90 (2006).
100. Parameswaran, S. *et al.* Molecular characterization of an MLL1 fusion and its role in chromosomal instability. *Mol. Oncol.* **13**, 422–440 (2019).
101. Winters, A. C. & Bernt, K. M. MLL-rearranged leukemias- An update on science and clinical approaches. *Front. Pediatr.* **5**, 11–13 (2017).
102. Meyer, C. *et al.* The MLL recombinome of acute leukemias in 2013. *Leukemia* **27**, 2165–2176 (2013).
103. Peterson, T. R. *et al.* DEPTOR Is an mTOR Inhibitor Frequently Overexpressed in Multiple Myeloma Cells and Required for Their Survival. *Cell* **137**, 873–886 (2009).
104. Hu, Y. *et al.* DEPTOR is a direct NOTCH1 target that promotes cell proliferation and survival in T-cell leukemia. *Oncogene* 1–10 (2016).
105. Milne, T. A. *et al.* Activity to Hox Gene Promoters University of Virginia Health System. *Mol. Cell* **10**, 1107–1117 (2002).
106. Kroon, E. *et al.* Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J.* **17**, 3714–3725 (1998).
107. Li, Z. *et al.* Consistent deregulation of gene expression between human and murine MLL rearrangement leukemias. *Cancer Res.* **69**, 1109–1116 (2009).
108. Lawrence, H. J. *et al.* Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid, and lymphoid hematopoiesis. *Blood* **89**, 1922–1930 (1997).
109. So, C. W., Karsunky, H., Wong, P., Weissman, I. L. & Cleary, M. L. Leukemic transformation of

- hematopoietic progenitors by MLL-GAS7 in the absence of Hoxa7 or Hoxa9. *Blood* **103**, 192–199 (2004).
110. Kawagoe, H., Humphries, K., Blair, A., Sutherland, H. & Hogge, D. Expression of hox genes and their cofactors in subpopulations of leukemic and normal human hematopoietic cells. *Leukemia* **13**, 687–698 (1999).
 111. Ferrando, A. A. *et al.* Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: Dominance of HOX dysregulation. *Blood* **102**, 262–268 (2003).
 112. Argiropoulos, B. & Humphries, R. K. Hox genes in hematopoiesis and leukemogenesis. *Oncogene* **26**, 6766–6776 (2007).
 113. Zhang, Y. *et al.* Evolving catalytic properties of the MLL family SET domain. *Structure* **23**, 1921–1933 (2015).
 114. Zhang, X. *et al.* Structure of the Neurospora SET domain protein DIM-5, a histone H3 lysine methyltransferase. *Cell* **111**, 117–127 (2002).
 115. Couture, J. F., Dirk, L. M. A., Brunzelle, J. S., Houtz, R. L. & Trievel, R. C. Structural origins for the product specificity of SET domain protein methyltransferases. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 20659–20664 (2008).
 116. Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 (2006).
 117. Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859–872 (2005).
 118. Xue, H. *et al.* Structural basis of nucleosome recognition and modification by MLL methyltransferases. *Nature* **573**, 445–449 (2019).
 119. Takahashi, Y. -h. *et al.* Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *Proc. Natl. Acad. Sci.* **108**, 20526–20531 (2011).
 120. Qu, Q. *et al.* Structure and Conformational Dynamics of a COMPASS Histone H3K4 Methyltransferase Complex. *Cell* **174**, 1117–1126 (2018).
 121. Hsu, P. L. *et al.* Crystal Structure of the COMPASS H3K4 Methyltransferase Catalytic Module. *Cell* **174**, 1106–1116 (2018).
 122. Park, S. H. *et al.* Cryo-EM structure of the human MLL1 core complex bound to the nucleosome. *Nat. Commun.* **10**, 3541–3554 (2019).
 123. Sun, Z. W. & Allis, C. D. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**, 104–108 (2002).
 124. Kim, J. *et al.* RAD6-Mediated Transcription-Coupled H2B Ubiquitylation Directly Stimulates H3K4 Methylation in Human Cells. *Cell* **137**, 459–471 (2009).
 125. Patel, A., Vought, V. E., Dharmarajan, V. & Cosgrove, M. S. A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J. Biol. Chem.* **283**, 32162–32175 (2008).
 126. Patel, A., Dharmarajan, V. & Cosgrove, M. S. Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *J. Biol. Chem.* **283**, 32158–32161 (2008).
 127. Song, J. J. & Kingston, R. E. WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket. *J. Biol. Chem.* **283**, 35258–35264 (2008).
 128. Avdic, V. *et al.* Structural and biochemical insights into MLL1 core complex assembly. *Structure* **19**, 101–108 (2011).

129. Odho, Z., Southall, S. M. & Wilson, J. R. Characterization of a novel WDR5-binding site that recruits RbBP5 through a conserved motif to enhance methylation of histone H3 lysine 4 by mixed lineage leukemia protein-1. *J. Biol. Chem.* **285**, 32967–32976 (2010).
130. Patel, A., Dharmarajan, V., Vought, V. E. & Cosgrove, M. S. On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *J. Biol. Chem.* **284**, 24242–24256 (2009).
131. Zhang, P. *et al.* A phosphorylation switch on RbBP5 regulates histone H3 Lys4 methylation. *Genes Dev.* **29**, 123–128 (2015).
132. Li, Y. *et al.* Structural basis for activity regulation of MLL family methyltransferases. *Nature* **530**, 447–452 (2016).
133. Kaustov, L. *et al.* The MLL1 trimeric catalytic complex is a dynamic conformational ensemble stabilized by multiple weak interactions. *Nucleic Acids Research* **47**, 9433–9447 (2019).
134. Shinsky, S. A., Monteith, K. E., Viggiano, S. & Cosgrove, M. S. Biochemical reconstitution and phylogenetic comparison of human SET1 family core complexes involved in histone methylation. *J. Biol. Chem.* **290**, 6361–6375 (2015).
135. Jiang, H. *et al.* Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* **144**, 513–525 (2011).
136. Yang, Z. *et al.* The DPY30 subunit in SET1/MLL complexes regulates the proliferation and differentiation of hematopoietic progenitor cells. *Blood* **124**, 2025–2033 (2014).
137. Yang, Z., Shah, K., Khodadadi-Jamayran, A. & Jiang, H. Dpy30 is critical for maintaining the identity and function of adult hematopoietic stem cells. *J. Exp. Med.* **213**, 2349–2364 (2016).
138. Karatas, H., Townsend, E. C., Bernard, D., Dou, Y. & Wang, S. Analysis of the binding of mixed lineage leukemia 1 (MLL1) and histone 3 peptides to WD repeat domain 5 (WDR5) for the design of inhibitors of the MLL1-WDR5 interaction. *J. Med. Chem.* **53**, 5179–5185 (2010).
139. Karatas, H. *et al.* High-Affinity, Small-Molecule Peptidomimetic Inhibitors of MLL1/WDR5 Protein–Protein Interaction. *J. Am. Chem. Soc.* **135**, 669–682 (2013).
140. Cao, F. *et al.* Targeting MLL1 H3K4 Methyltransferase Activity in Mixed-Lineage Leukemia. *Mol. Cell* **53**, 247–261 (2014).
141. Zhang, H. *et al.* MLL1 Inhibition Reprograms Epiblast Stem Cells to Naive Pluripotency. *Cell Stem Cell* **18**, 481–494 (2016).
142. Delgado, R. N. *et al.* Maintenance of neural stem cell positional identity by mixed-lineage leukemia 1. *Science* **368**, (2020).
143. Avdic, V. *et al.* Fine-tuning the stimulation of MLL1 methyltransferase activity by a histone H3-based peptide mimetic. *FASEB J.* **25**, 960–967 (2011).
144. Alicea-Velázquez, N. L. *et al.* Targeted disruption of the interaction between WD-40 repeat protein 5 (WDR5) and Mixed Lineage Leukemia (MLL)/SET1 Family Proteins Specifically Inhibits MLL1 and SETd1A Methyltransferase Complexes. *J. Biol. Chem.* **291**, 22357–22372 (2016).
145. Karatas, H. *et al.* Discovery of a Highly Potent, Cell-Permeable Macrocyclic Peptidomimetic (MM-589) Targeting the WD Repeat Domain 5 Protein (WDR5)–Mixed Lineage Leukemia (MLL) Protein–Protein Interaction. *J. Med. Chem.* **60**, 4818–4839 (2017).
146. Gupta, A. *et al.* Facile target validation in an animal model with intracellularly expressed monobodies. *Nat. Chem. Biol.* **14**, 895–900 (2018).
147. Senisterra, G. *et al.* Small-molecule inhibition of MLL activity by disruption of its interaction with WDR5. *Biochem. J.* **449**, 151–159 (2013).

148. Bolshan, Y. *et al.* Synthesis, Optimization, and Evaluation of Novel Small Molecules as Antagonists of WDR5 - MLL Interaction. *ACS Med. Chem. Lett.* **4**, 353–357 (2013).
149. Getlik, M. *et al.* Structure-Based Optimization of a Small Molecule Antagonist of the Interaction between WD Repeat-Containing Protein 5 (WDR5) and Mixed-Lineage Leukemia 1 (MLL1). *J. Med. Chem.* **59**, 2478–2496 (2016).
150. Grebien, F. *et al.* Pharmacological targeting of the Wdr5-MLL interaction in C/EBP α N-terminal leukemia. *Nat. Chem. Biol.* **11**, (2015).
151. Li, D. D. *et al.* High-affinity small molecular blockers of mixed lineage leukemia 1 (MLL1)-WDR5 interaction inhibit MLL1 complex H3K4 methyltransferase activity. *Eur. J. Med. Chem.* **124**, 480–489 (2016).
152. Li, D. D. *et al.* Structure-based design and synthesis of small molecular inhibitors disturbing the interaction of MLL1-WDR5. *Eur. J. Med. Chem.* **118**, 1–8 (2016).
153. Ye, X. *et al.* The identification of novel small-molecule inhibitors targeting WDR5-MLL1 interaction through fluorescence polarization based high-throughput screening. *Bioorganic Med. Chem. Lett.* **29**, 638–645 (2019).
154. Wang, F. *et al.* Discovery of Potent 2-Aryl-6,7-dihydro-5 H-pyrrolo[1,2- a]imidazoles as WDR5-WIN-Site Inhibitors Using Fragment-Based Methods and Structure-Based Design. *J. Med. Chem.* **61**, 5623–5642 (2018).
155. Aho, E. R. *et al.* Displacement of WDR5 from Chromatin by a WIN Site Inhibitor with Picomolar Affinity. *Cell Rep.* **26**, 2916–2928 (2019).
156. Tian, J. *et al.* Discovery and Structure-Based Optimization of Potent and Selective WD Repeat Domain 5 (WDR5) Inhibitors Containing a Dihydroisoquinolinone Bicyclic Core. *J. Med. Chem.* **5**, (2020).
157. Keown, A. Celgene Strikes Deal Worth Nearly 1 Billion for Blood Cancer Treatment. *Bio Space web page* (2019). Available at: <https://www.biospace.com/article/celgene-strikes-deal-worth-nearly-1-billion-with-triphase-accelerator-for-blood-cancer-treatment/>. (Accessed: 23rd February 2020)
158. TRPH-395: A NOVEL SMALL MOLECULE INHIBITOR OF WDR5. *Triphase web page* (2020). Available at: <http://triphaseco.com/trph-395/>. (Accessed: 23rd February 2020)
159. Hale, C. Celgene inks 1B deal for a preclinical epigenetic blood cancer drug from Canada. *fiercebiotech web page* (2019). Available at: <https://www.fiercebiotech.com/biotech/celgene-inks-1b-deal-for-a-preclinical-epigenetic-blood-cancer-drug-from-canada>. (Accessed: 23rd February 2020)
160. Celgene partners Triphase for preclinical blood cancer candidate. *Pharmaceutical Technology web page* (2019). Available at: <https://www.pharmaceutical-technology.com/news/celgene-partners-triphase/>. (Accessed: 23rd February 2020)
161. FACIT and Triphase Accelerator Announce New Partnership with Celgene for First-in-class WDR5 Leukemia Therapy. *globenewswire web page* (2019). Available at: <https://www.globenewswire.com/news-release/2019/01/29/1706697/0/en/FACIT-and-Triphase-Accelerator-Announce-New-Partnership-with-Celgene-for-First-in-class-WDR5-Leukemia-Therapy.html>. (Accessed: 23rd February 2020)
162. Our Pipeline. *Triphase web page* (2020). Available at: <http://triphaseco.com/our-pipeline/>. (Accessed: 23rd February 2020)
163. Berg, J. M., Tymoczko, J. L. & Stryer, L. *Biochemie*. (Elsevir, 2007).
164. Broichhagen, J., Frank, J. A. & Trauner, D. A Roadmap to Success in Photopharmacology. *Acc.*

- Chem. Res.* **48**, 1947–1960 (2015).
165. Hüll, K., Morstein, J. & Trauner, D. In Vivo Photopharmacology. *Chem. Rev.* **118**, 10710–10747 (2018).
166. Lerch, M. M., Hansen, M. J., van Dam, G. M., Szymanski, W. & Feringa, B. L. Emerging Targets in Photopharmacology. *Angew. Chemie - Int. Ed.* **55**, 10978–10999 (2016).
167. Velema, W. A., Szymanski, W. & Feringa, B. L. Photopharmacology: Beyond proof of principle. *J. Am. Chem. Soc.* **136**, 2178–2191 (2014).
168. Ann L CokerAbrahamse, H. & Hamblin, M. R. New Photosensitizers for Photodynamic Therapy. *Biochem. J.* **473**, 347–364 (2016).
169. Dougherty, T. J. *et al.* Photoradiation Therapy for the Treatment of Malignant Tumors. *Cancer Res.* **38**, 2628–2635 (1978).
170. Agostinis, P. *et al.* Photodynamic therapy of cancer: An update. *CA: A Cancer Journal for Clinicians* **61**, 250–281 (2011).
171. Linden, G. *et al.* Conditional Singlet Oxygen Generation through a Bioorthogonal DNA-targeted Tetrazine Reaction. *Angew. Chemie - Int. Ed.* **58**, 12868–12873 (2019).
172. Lovell, J. F., Liu, T. W. B., Chen, J. & Zheng, G. Activatable photosensitizers for imaging and therapy. *Chem. Rev.* **110**, 2839–2857 (2010).
173. Barton, D. H. R., Chow, Y. L., Cox, A. & Kirby, G. W. Photosensitive protection of functional groups. *Tetrahedron Lett.* **3**, 1055–1057 (1962).
174. Barltrop, J. A. & Schofield, P. photosensitive protecting groups. *Tetrahedron Lett.* **16**, 697–699 (1962).
175. Sheehan, J. C. & Wilson, R. M. Photolysis of Desyl Compounds. A New Photolytic Cyclization. *J. Am. Chem. Soc.* **86**, 5277–5281 (1964).
176. Engels, J. & Schlaeger, E. J. Synthesis, Structure, and Reactivity of Adenosine Cyclic 3',5'-Phosphate-Benzyl Triesters. *J. Med. Chem.* **20**, 907–911 (1977).
177. Silva, J. M., Silva, E. & Reis, R. L. Light-triggered release of photocaged therapeutics - Where are we now? *J. Control. Release* **298**, 154–176 (2019).
178. Mart, R. J. & Allemann, R. K. Azobenzene photocontrol of peptides and proteins. *Chem. Commun.* **52**, 12262–12277 (2016).
179. Lubbe, A. S., Szymanski, W. & Feringa, B. L. Recent developments in reversible photoregulation of oligonucleotide structure and function. *Chem. Soc. Rev.* **46**, 1052–1079 (2017).
180. Szymański, W., Beierle, J. M., Kistemaker, H. A. V., Velema, W. A. & Feringa, B. L. Reversible photocontrol of biological systems by the incorporation of molecular photoswitches. *Chem. Rev.* **113**, 6114–6178 (2013).
181. Beharry, A. A. & Woolley, G. A. Azobenzene photoswitches for biomolecules. *Chem. Soc. Rev.* **40**, 4422–4437 (2011).
182. Dong, M., Babalhavaeji, A., Samanta, S., Beharry, A. A. & Woolley, G. A. Red-Shifting Azobenzene Photoswitches for in Vivo Use. *Acc. Chem. Res.* **48**, 2662–2670 (2015).
183. Crespi, S., Simeth, N. A. & König, B. Heteroaryl azo dyes as molecular photoswitches. *Nat. Rev. Chem.* **3**, 133–146 (2019).
184. Komarov, I. V., Afonin, S., Babii, O., Schober, T. & Ulrich, A. S. Efficiently Photocontrollable or Not? Biological Activity of Photoisomerizable Diarylethenes. *Chem. Eur. J.* **24**, 11245–11254 (2018).

185. Klajn, R. Spiropyran-based dynamic materials. *Chem. Soc. Rev.* **43**, 148–184 (2014).
186. Kitzig, S., Thilemann, M., Cordes, T. & Rück-Braun, K. Light-Switchable Peptides with a Hemithioindigo Unit: Peptide Design, Photochromism, and Optical Spectroscopy. *ChemPhysChem* **17**, 1252–1263 (2016).
187. Wiedbrauk, S. & Dube, H. Hemithioindigo — an emerging photoswitch. *Tetrahedron Lett.* **56**, 4266–4274 (2015).
188. Pianowski, Z. L. Recent Implementations of Molecular Photoswitches into Smart Materials and Biological Systems. *Chem. - A Eur. J.* **25**, 5128–5144 (2019).
189. Albert, L. & Vazquez, O. Photoswitchable Peptides for Spatiotemporal Control of Biological Functions. *Chem. Commun.* **55**, 10192–10213 (2019).

2 Aim

Epigenetics is the collection of highly dynamic processes involving multiple chromatin-modifying and remodeling enzymes that control the accessibility to genes and their expression through a complex signaling network of protein-protein interactions (PPI). Importantly, epigenetic regulation is at the core of both natural and pathological states. Therefore, modulating the epigenome to tune transcription profiles and cellular phenotypes in a programmable manner is of wide interest, and may ultimately lead to novel epigenetic-based therapeutics. Despite the current advances of using small molecules as inhibitors of epigenetic enzymes, and their contribution to improve our insights into epigenetic mechanisms, the current methods are not sufficient to deal with the challenges of targeting the dynamic epigenome. Taking the importance of PPIs in epigenomic regulation into account as well as the precedents of the Mixed Lineage Leukemia 1 (MLL1) complex (section 1.4), in this thesis we pursued photopharmacological approaches to establish a proof-of-concept over light-controllable inhibition of MLL1. Thus, we planned to efficiently modulate the MLL1-WDR5 interaction by using photoswitchable peptidomimetics based on the MLL1 *Win*-derivative peptide NH₂-11mer **MB-3** as well as on the macrocyclic peptide MM-401 **A-2** (section 1.5). Such peptidomimetics would contain unique chemical motifs that enable a reversible switch between two distinct geometries under irradiation. This conformational trap should affect their capacity to bind to WDR5. Consequently, this would provide a light-controllable on/ off-activation of MLL1 and, in turn, an influence on the epigenetic states (Figure A-1).

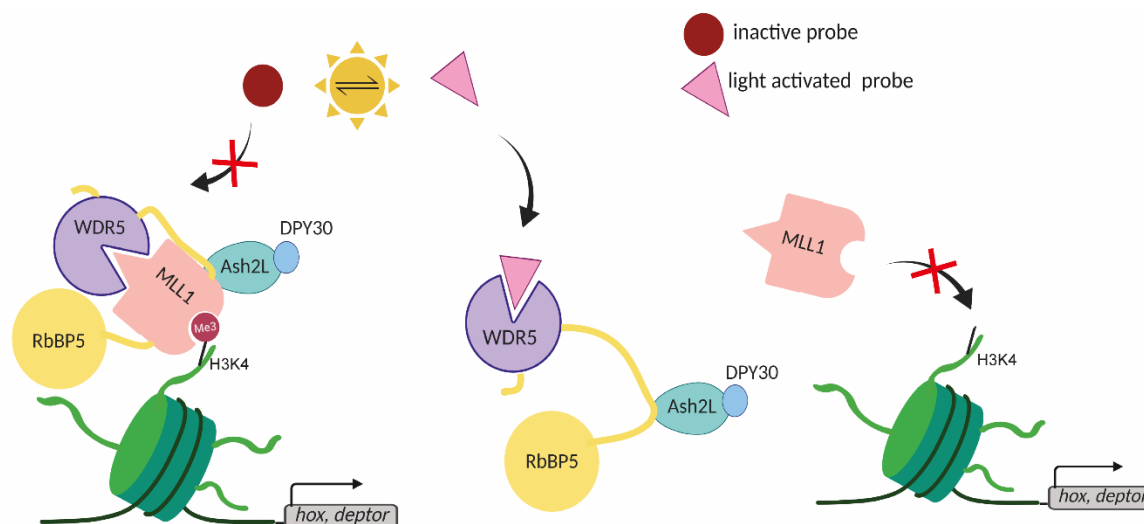


Figure A-1. Schematic representation of the indirect strategy to modulate methylation activity of MLL1 by controlling its interaction to WDR5 utilizing photoresponsive peptides.

The planned biological experiments will depend on the potential of our probes. At first, fluorescence polarization (FP) will be used as an *in vitro* assay to determine differences in WDR5 binding affinities between isomers. Promising candidates, which display both, low inhibition constants (K_i) and a significant affinity window between isomeric states will be evaluated on their ability to actually control MLL1 methylation via *in vitro* histone methyltransferase (HMT) assays and MLL1 complex disruption by pull-down assays and/or hydrogen-deuterium exchange (HDX) measurements. If promising *in vitro* results are observed, cell-penetrating candidates will be pursued by using transfection methods, or adding uptake vectors to our compounds. The effect of these compounds could be evaluated on

leukemia cell viability, MLL1-target gene expression and even in in vivo studies, as MLL1 is highly conserved in the optically transparent zebrafish model. Therefore, photocontrol of MLL1-dependent hematopoiesis inhibition could be achieved for the first time, to our knowledge. All together, these studies will demonstrate the potential of our probes as optoepigenetic tools and will shed light into hematopoiesis.

Initially, as proof of concept, we planned to introduce the classical azobenzene 4-[(4'-amino-methyl)phenylazo]benzoic acid (AMPB) into the peptide backbone of the *Win*-peptide derivative NH₂-11mer **MB-3**, whose WDR5 affinity is in the nanomolar range ($K_i = 20$ nM). I partially synthesized the peptide library during my master thesis (NH₂-11mer based peptides **MB-3** – **MB-12**) but did not study their biological activity. During this thesis, the battery of peptides will be completely characterized, i.e., photochemically and biologically evaluated. Several peptides will require resynthesis. Besides, the library will be extended to explore the effect when two residues are exchanged by AMPB (**MB-13**) and when AMPB is incorporated in improved parental peptides (K_i (Ac-10mer **MB-2**) = 3 nM; K_d (Win-6mer **A-1**) = 2.9 nM) than the NH₂-11mer **MB-3** (**MB-14** and **MB-15**).

parental NH₂-11mer:

MB-3: H₂N - S **ARA**EVHLRKS - CONH₂

based on NH₂-11mer:

MB-4: H₂N - S **X**ARA EVHLRKS - CONH₂

MB-5: H₂N - S **X**RA EVHLRKS - CONH₂

MB-6: H₂N - S AR **X**EVHLRKS - CONH₂

MB-7: H₂N - SARA **X**VHLRKS - CONH₂

MB-8: H₂N - SARA E **X**H L R K S - CONH₂

MB-9: H₂N - SARA EV **X**L R K S - CONH₂

MB-10: H₂N - SARA EVH **X**R K S - CONH₂

MB-11: H₂N - SARA EVHL **X**K S - CONH₂

MB-12: H₂N - SARA EVHLR **X**S - CONH₂

substitution of 2 aa:

MB-13: H₂N - SARA **X**H L R K S - CONH₂

based on parental Ac-10mer, **MB-2:** Ac-ARAEVHLRKS-CONH₂

MB-14: Ac - ARA **X**VHLRKS - CONH₂

based on parental Win-6mer, **A-1:** Ac-ARTEVY-CONH₂

MB-15: Ac - ART **X**V Y - CONH₂

X = AMPB moiety

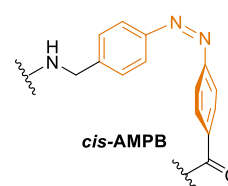
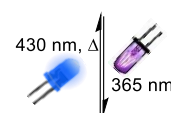
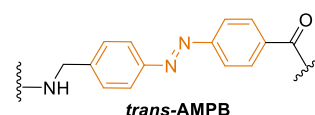


Figure A-2. Amino acid scan on NH₂-11mer peptide **MB-3** to yield our initial azo-photoactivatable peptide library. Highlighted in green, the indispensable motif for WDR5 recognition, the –ARA– sequence; X represents the 4-[(4'-amino-methyl)phenylazo]benzoic acid (AMPB).

Once we demonstrate the success of this main concept, we will strive to maximize the difference in biological output via:

- a) improving the photochemical properties of the molecular transducer: Thus, we will synthesize two novel visible-light-responsive peptide backbone photoswitches based on tetra-*ortho*-fluoro-azobenzene (*o*F₄Azo) and cyclic bridged azobenzene (cAzo) (Figure A-3). The cAzo photoswitch exerts inverted thermal stability, the *cis* isomer being the most stable one, while the *o*F₄Azo photoswitch displays almost quantitative isomer conversions in organic solvents.

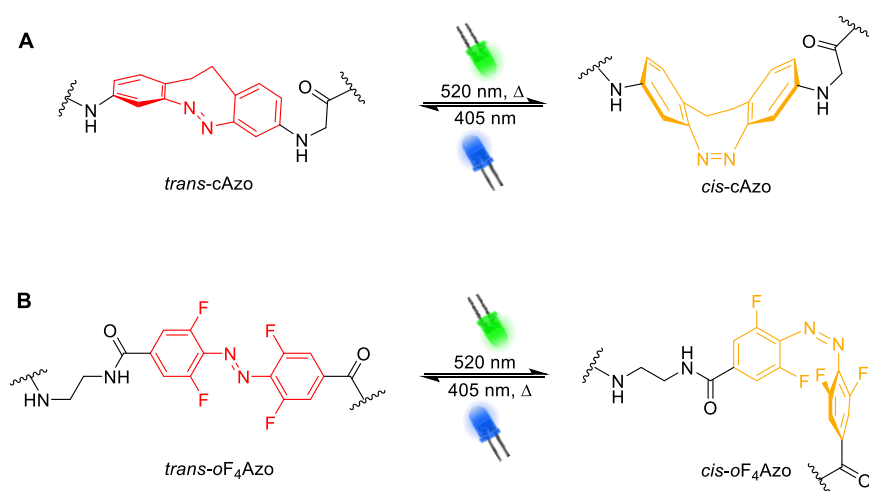


Figure A-3. Visible-light-responsive peptide backbone photoswitches based on **A**) cyclic bridged azobenzene (cAzo) and **B**) tetra-*ortho*-fluoro-azobenzene (oF₄Azo).

- b) using the conformational strained cyclic peptide MM-401 as scaffold. The battery of compounds will differ in the position where the best molecular photoswitch will be grafted, the presence of D-phenylglycine and the ring size. The latter will be achieved by varying the length of the connecting amino acid side chain (lysine, diaminobutyric acid, diaminopropionic acid, and none) between the interaction motif (—Arg-Abu—) and the molecular photoswitch.

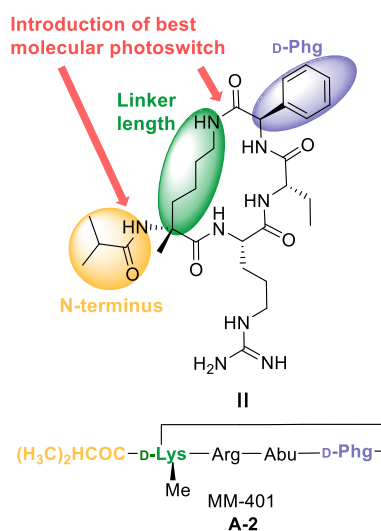


Figure A-4. Design of aimed cyclic peptide library based on the reported MM-401 A-2 macrocycle, and incorporation of the best molecular transducer.

3 Summaries

3.1 Summary in English

This cumulative thesis consists of two published articles in *Chemical Science* and *ChemBioChem*, a submitted manuscript and a published review in *Chemical Communications*, which provides a detailed overview of photoswitchable peptides capable of modulating biological functions. General aspects as used photoswitches, incorporation techniques and structural precedents are discussed too. Besides the articles, the thesis includes a general overview of epigenetics, which is focused on the methyltransferase MLL1 complex and its inhibitors that target the protein-protein interaction (PPI) MLL1-WDR5. The cumulative part of the thesis is followed by an outlook and conclusions, drawn from the published/submitted manuscripts, and an additional section, which deals with further developments, which were not discussed in the publications.

During the last years, we pursued the control of epigenetic states using photoswitchable molecules instead of genetically encoded photoreceptors. For this purpose, we focused on the histone methyltransferase Mixed-Lineage-Leukemia 1 (MLL1) also known as lysine [K]-specific Methyl Transferase 2A (KMT2A), which is a promising target in cancer therapy. Its activity depends on the formation of a protein-protein core complex (MLL1, WDR5, RbBP5, Ash2L, DPY30), where especially the WDR5-MLL1 interaction is of crucial importance. Such PPI is essential to facilitate the MLL1 binding to the RbBP5-Ash2L dimer, needed to maintain the catalytically active conformation of the SET1-domain of MLL1. Based on reported peptides, capable of binding to WDR5, we conceived photoswitchable PPI modulators for altering MLL1 activity. Thus, initially we hypothesized that the incorporation of the classical azobenzene 4-[(4'-aminomethyl)phenyl-azo]benzoic acid (AMPB) into the MLL1-derivative peptide: H₂N-SARAEVHLRKS-CONH₂ would cause a light-triggered conformational change, which may affect its WDR5 affinity and, thereby, MLL1 activity. Fluorescence polarization (FP)-based experiments revealed that the photoswitchable peptidomimetic H₂N-SARA-AMPB-VHLRKS-CONH₂ was the most promising candidate of our positional scanning library for the enzymatic activity studies. Hence, this compound displayed the highest affinity window between isomers: K_i (*trans*) = 1.25 ± 0.36 nM versus K_i (*cis*) = 6.50 ± 1.4 nM, as well as one order of magnitude tighter binding than the parental MLL1-derivative peptide. Our co-crystal structure (PDB: 5M23) could explain the latter result as a consequence of further stabilization due to an additional hydrogen bond between the first nitrogen of the diazenyl group to the Lys259 as well as van der Waals interactions between the second benzene ring of the AMPB and the Tyr260. Furthermore, our peptidomimetic did not only strongly bind to WDR5 but also inhibited the methylation activity of the MLL1 complex, according to radioactive histone methyltransferase (HMT) assays. Importantly, the difference between isomers was amplified (HMT activity IC₅₀ (*trans*) = 0.531 ± 0.023 μM; IC₅₀ (*cis*) = 8.23 ± 0.38 μM), demonstrating for the first time that photoswitchable peptidomimetics can control the activity of histone methyltransferases. The precise control of MLL1 engages a pivotal role in the transcriptional regulation of genes involved in hematopoiesis as H3K4 methylation is associated with transcriptionally active genes. Since MLL1-target genes, such as *hox* and *deptor*, are highly overexpressed in MLL-rearranged leukemias, we checked the effect of our peptide on leukemia cell viability. For this purpose, we prepared cell-penetrating versions of the best peptide candidate to explore the modulation of cell growth in leukemia mouse bone marrow cells (MLL-AF9 cells). Cell viability assays verified effective inhibition with a modest, yet clear, difference up to 1.52-fold in the cytotoxic behaviour of isomers (GI₅₀ (*trans*) = 2.14 ± 0.10 μM, GI₅₀ (*cis*) = 3.25 ± 0.23 μM). In addition, our most potent inhibitor was tested on its ability to downregulate the expression of *deptor* via reverse transcriptase real time quantitative PCR

(RT-qPCR). These experiments verified that our probe is, indeed, an optoepigenetic inhibitor because it differently downregulated *deptor* expression levels depending on the photoisomer (relative expression levels of *deptor*: *trans* = 0.23 %; *cis* = 0.36 %, $p = 0.008$). This proof-of-concept study resulted in an article published in *Chemical Science*.

Once that we demonstrated the potential of photopharmacological inhibition of MLL1 for external epigenetic control, we used this optimized biological system as a platform to improve the photoswitchable peptidomimetics. In particular, we strove for maximizing the difference in biological output between isomers and improving the photochemical properties. Thus, we established the synthesis of two visible-light responsive peptide backbone photoswitches: a Fmoc-protected cyclic azobenzene analogue of 5,6-dihydrodibenzol[c,g][1,2]diazocine (Fmoc-cAzoAA) and a Mtt-ethylenediamine protected tetra-*ortho*-fluoroazobenzene (Mtt-*o*F₄AzoAA). Apart from avoiding UV irradiation, the *o*F₄Azo switch displayed higher photoconversions than the AMPB derivative (95% *cis*; 86% *trans*). However, when they were introduced into the same MLL1-derivative peptide, they did not provide an increased difference in WDR5 binding compared to our former results (H₂N-SARA-cAzo-VHLRKS-CONH₂ K_i (*trans*) = 140 ± 35 nM, K_i (*cis*) = 207 ± 52 nM; H₂N-SARA-*o*F₄Azo-VHLRKS-CONH₂ K_i (*trans*) = 11.8 ± 1.4 nM, K_i (*cis*) = 30.8 ± 3.3 nM). Structure-activity relationships deduced from crystallization and theoretical calculations revealed substantial differences in the positioning of the azobenzene moieties. This project was published in *ChemBioChem*, and highlighted as a front cover.

Finally, we combined the visible-light *o*F₄Azo with a structurally-defined cyclic scaffold in an attempt to maximize the biological output. Thus, we designed and synthesized 14 photoswitchable cyclic peptide analogues based on the scaffold of the MLL1 inhibitor: MM-401 ((H₃C)₂CHCO-NH-c[D-Lys-Arg-Abu-D-Phg][#]). Following the same workflow, FP-based assays revealed that the candidate (H₃C)₂CHCO-NH-c[D-Dab-Arg-Abu-D-Phe-*o*F₄Azo][#] displayed an increase of 50% in the WDR5 affinity window between isomers (K_i (*trans*) = 78.6 ± 3.4 nM, K_i (*cis*) = 7.99 ± 0.79 nM) compared with our linear AMPB peptide. This confirms that a cyclic constrain leads to greater modulation potential than the lineal arrangement. Unlike the linear AMPB peptide, the *cis* isomer exhibited both higher WDR5 affinity and MLL1 inhibition activity compared to the *trans* ones: IC₅₀ (*trans*) = 6.68 ± 3.1 μM; IC₅₀ (*cis*) = 0.649 ± 0.3 μM. Crystallization studies, molecular docking (MD) in combination with virtual docking (VD) studies, pull-down assays and hydrogen deuterium exchange (HDX) mass spectrometry (MS) corroborated the increased difference in WDR5 affinity, and its effect on the whole MLL1 complex topology. Via HDX-MS, we discovered that the observed activity loss upon *cis*-cycle addition was yielded by i.a. allosteric rearrangements of the WDR5-RbBP5 binding interface, leading to the inability of the RbBP5-Ash2L dimer to stabilize the open SET1 conformation, which is required for its activity. Importantly, our cyclic PPI modulators presented improved photochromic properties, behaving as a quasi bistable system. Indeed, the photostationary states were stable for at least five months in darkness. Furthermore, we examined whether the loss of MLL1-enzymatic activity could modulate viability of human leukemia cells. Cell permeabilization via a peptide carrier varied for different cell types, and preliminary tests using the same MLL-AF9 cells as in the *Chemical Science*, proved that such cells were not suitable. The best results were obtained in MOLM-13 cells, where viability differences of *trans*: 37 %, *cis*: 0 % at 200 μM and *trans*: 57 %, *cis*: 26 % at 100 μM between isomers were detected. To exclude that the observed toxicity arises from the *o*F₄Azo-moiety itself, the enantiomer of our best *o*F₄Azo-containing cyclic peptide, where the Arg was exchanged by a D-Lys (D-Lys variant), was tested upon its effects on cell viability. Besides, a washing step after cargo transfection was included into the experiment to ensure that the observed effects resulted solely from intracellular mechanisms. Such experiments, with the D-Lys variant as a control proved that the detected toxicity is not originated from

the *o*F₄Azo-moeity since low cytotoxicity was observed at 100 μM, being even lower with the active *cis* isomer (viability: *trans*: 70 %, *cis*: 91 %). Also, no interaction of the D-Lys variant with WDR5 was detected by FP-based assays. In contrast, the active *cis* isomer of our best Arg-containing photoswitchable cyclic candidate caused high toxicity under the same conditions, while the *trans* isomer retained the same effect (viability: *cis*: 24 %, *trans*: 66 %). Ultimately, we tested the in vivo effect of our compound on the modulation of hematopoiesis. We chose the zebrafish developmental model, because MLL1 is conserved in zebrafish, and its knock out represses hematopoiesis. This lack of blood flow is observable by *o*-dianisidine staining. Gratifyingly, we demonstrated that the *cis* isomer is capable to affect hematopoiesis in 3 days post fertilization (dpf) zebrafish larva (*o*-dianisidine staining: 60% highly disrupted, 40% no staining), which displayed curved body axes, heart edema and a lack of response upon touch. In contrast, 100 % of *trans* incubated larva, as well as water and DMSO controls, did not exhibit any of these defects at all. The possibility of visible-light isomerization of our novel cyclic peptide allowed us to study whether *trans* incubated larvae could be in situ activated. Such experiment proved that the high significant differences between isomers could, indeed, be achieved via in situ irradiation to the *cis* state (*o*-dianisidine staining: 3% intact, 30% disrupted, 67%: no staining). All these results together demonstrate the possibility of modulating MLL1-dependent hematopoiesis in vivo using cyclic constraint photoresponsive peptides. This project has recently been submitted.

3.2 Zusammenfassung in Deutscher Sprache

Die vorliegende kumulative Dissertationsschrift besteht aus zwei publizierten Manuskripten in *Chemical Science* und *ChemBioChem*, einem eingereichten Manuskript und einem veröffentlichten Review in *Chemical Communications*, welches die optochemische Kontrolle über biologische Prozesse unter Anwendung fotoschaltbarer Peptide behandelt. Außerdem werden in dem Review grundlegende Prinzipien der eingesetzten Fotoschalter, die Techniken ihrer Inkorporierung in Peptide, sowie deren struktureller Eigenschaften diskutiert. Darüber hinaus beinhaltet diese Thesis eine allgemeine Übersicht der Epigenetik. Hierbei liegt der Schwerpunkt auf dem Methyltransferase MLL1-Komplex und dessen Inhibitoren, welche die Protein-Protein Interaktion (PPI) zwischen MLL1 und WDR5 als Target adressieren. Dem kumulativen Teil folgt ein Fazit und ein Ausblick sowie ein weiterer Abschnitt über zusätzlich durchgeführte Experimente, welche nicht in den Manuskripten diskutiert wurden.

Während der letzten Jahre verfolgten wir das Ziel, die Kontrolle über epigenetische Zustände mit Hilfe fotoschaltbarer Peptide zu erlangen. Hierbei lag unser Augenmerk auf der Histon Methyltransferase Mixed-Lineage-Leukemia 1 (MLL1), auch bekannt unter dem Namen Lysin [K]-spezifische Methyltransferase 2A (KMT2A), einem vielversprechenden Target in der Krebstherapie. Die enzymatische Aktivität von MLL1 ist abhängig von der Formation eines Protein-Protein Komplexes (MLL1, WDR5, RbBP5, Ash2L, DPY30), wobei insbesondere die Interaktion zwischen MLL1 und WDR5 von essentieller Bedeutung ist. Diese PPI ist notwendig um die Bindung des RbBP5-Ash2L Dimers an MLL1 zu gewährleisten und somit die katalytisch aktive offene Konformation der SET1-Domäne von MLL1 zu stabilisieren. Basierend auf bekannten WDR5-bindenden Peptiden haben wir fotoschaltbare Analoga entwickelt, um somit die MLL1-WDR5 Interaktion reversibel zu inhibieren. Zunächst haben wir das klassische Azobenzol 4-[(4'-aminomethyl)phenyl-azo]benzoic acid (AMPB) in ein MLL1 Peptidderivat ($\text{H}_2\text{N-SARA-EVHLRKS-CONH}_2$) eingebaut um somit eine lichtgesteuerte Konformationsänderung zwischen den beiden Isomeren zu erzeugen, was letztendlich zu einer kontrollierbaren Inhibierung der MLL1 Aktivität führen sollte. Fluoreszenz-Polarisations (FP) Assays zeigten, dass das Peptid: $\text{H}_2\text{N-SARA-AMPB-VHLRKS-CONH}_2$ den vielversprechendsten Kandidaten unserer Peptid-Bücherei darstellte, da es sowohl die höchste Differenz in Bindungsaffinität zwischen den Isomeren aufwies ($K_i(\text{trans}) = 1.25 \pm 0.36 \text{ nM}$ versus $K_i(\text{cis}) = 6.50 \pm 1.4 \text{ nM}$), als auch um eine Größenordnung gesteigerte Affinität zum originalen MLL1 Peptid besaß. Mit Hilfe der Co-Kristallstruktur (PDB: 5M23) konnten wir diese gesteigerte Affinität erklären. So sorgte eine weitere Wasserstoffbrückenbindung des ersten Stickstoffs der N=N-Azo Doppelbindung zum Lys259, sowie van der Waals Interaktionen zwischen dem zweiten Benzolring des AMPBs zum Tyr260 für zusätzliche Stabilität. Weiterhin wies unser Peptid nicht nur eine hohe Bindungsaffinität zu WDR5 auf, sondern war außerdem in der Lage die MLL1 Aktivität zu inhibieren. Hervorzuheben ist, dass die Differenz zwischen den Isomeren auf das 15-fache erhöht werden konnte (HMT Aktivität $\text{IC}_{50}(\text{trans}) = 0.531 \pm 0.023 \text{ }\mu\text{M}$; $\text{IC}_{50}(\text{cis}) = 8.23 \pm 0.38 \text{ }\mu\text{M}$). Dies zeigt erstmalig die lichtgesteuerte Modulierung von Histonmethyltransferasen. Die präzise Kontrolle von MLL1 spielt eine entscheidende Rolle in der Regulierung der Transkription von Genen, welche in der Hämatopoese und Entwicklung involviert sind. So ist die Trimethylierung von Histon 3 Lysin 4 (H3K4) mit transkriptional aktiven Genen assoziiert. Da MLL1 adressierte Gene, wie *hox* und *depor*, in diversen Leukämiearten stark überexprimiert vorliegen, haben wir den Effekt unseres Peptids auf die Viabilität von Leukämie Zellen (MLL-AF9 Maus Knochenmark Zellen) getestet. Hierfür verknüpften wir unsere Verbindung mit zellpenetrierenden Peptiden (*Cell Penetrating Peptides*, CPPs), um die Aufnahme in die Zelle zu gewährleisten. Die Zell-Viabilitäts Assays verifizierten die effektive Inhibierung des Zell-Wachstums mit einem deutlichen Unterschied zwischen den beiden Isomeren bis zu einem Faktor von 1.52 ($\text{GI}_{50}(\text{trans}) = 2.14 \pm 0.10 \text{ }\mu\text{M}$, $\text{GI}_{50}(\text{cis}) = 3.25 \pm 0.23 \text{ }\mu\text{M}$). Schließlich testeten

wir, anhand von Reversibler Transkriptase Echtzeit quantitativer PCR (RT-qPCR), ob unser vielversprechendster Kandidat in der Lage ist, die Transkription des MLL1 assoziierten Gens *depor* zu regulieren. Erfreulicherweise bewiesen diese Experimente, dass unser Peptid tatsächlich als optoepigenetischer Inhibitor dient, da er die Expression von *depor*, abhängig vom Fotoisomer, herabreguliert (relative Expressionslevel von *depor*: *trans* = 0.23%; *cis* = 0.36%, $p = 0.008$). Diese proof-of-concept Ergebnisse wurden in *Chemical Science* veröffentlicht.

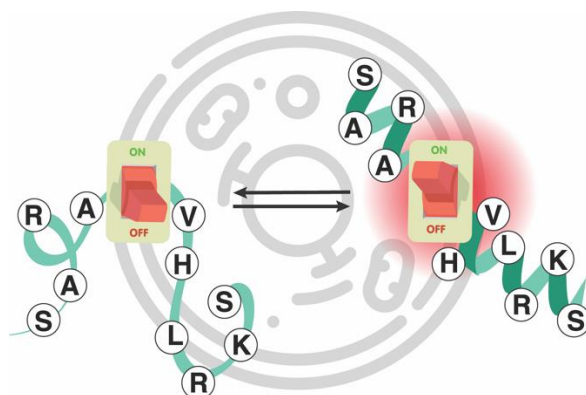
Nachdem das Potential der fotopharmakologischen Inhibierung von MLL1 etabliert war, wollten wir diese optimierte biologische Plattform nutzen, um unsere fotoschaltbaren Peptide weiter zu entwickeln. Spezifisch formuliert, sollte die Differenz des biologischen Outputs zwischen den Isomeren vervielfältigt, sowie die fotochromen Eigenschaften verbessert werden. Folglich haben wir die Synthese zweier neuartiger fotoschaltbarer Aminosäurebausteine entwickelt: ein Fmoc-geschütztes cyclisches Azobenzolanalogon von 5,6-dihydrodibenzol[c,g][1,2]diazocine (Fmoc-cAzoAA) sowie ein Mtt-ethylenediamin geschütztes tetra-*ortho*-fluoroazobenzol (Mtt-oF₄AzoAA). Abgesehen von der Vermeidung von UV-Licht für die Isomerisierung, weist der oF₄Azo Fotoschalter höhere Fotokonversionen auf als der vorige AMPB Schalter (95% *cis*; 86% *trans*). Inkorporierung in dasselbe MLL1 Peptid führte jedoch nicht zu einer erhöhten Differenz in WDR5-Bindungsdaffinität zwischen den Isomeren (H₂N-SARA-cAzo-VHLRKS-CONH₂ K_i (*trans*) = 140 ± 35 nM, K_i (*cis*) = 207 ± 52 nM; H₂N-SARA-oF₄Azo-VHLRKS-CONH₂ K_i (*trans*) = 11.8 ± 1.4 nM, K_i (*cis*) = 30.8 ± 3.3 nM). Struktur-Aktivitäts Relationen (SARs), abgeleitet von Kristallisationsstudien, sowie theoretischen Berechnungen wiesen erhebliche Unterschiede in der Positionierung der Azobenzoleinheiten auf. Dieses Projekt wurde in *ChemBioChem* veröffentlicht und als Front Cover ausgezeichnet.

Schließlich kombinierten wir den oF₄Azo Fotoschalter mit einem strukturell definierten Grundgerüst, um somit die Differenz zwischen den Isomeren zu maximieren. Basierend auf dem Grundgerüst des MLL1 Inhibitors MM-401: ((H₃C)₂CHCO-NH-c[D-Lys-Arg-Abu-D-Phg][#]) wurden insgesamt 14 cyclische Peptidderivate entworfen und synthetisiert. FP-basierte Assays zeigten, dass der Kandidat: (H₃C)₂CHCO-NH-c[D-Dab-Arg-Abu-D-Phe-oF₄Azo][#] eine Zunahme der WDR5-Affinitätsdifferenz zwischen Isomeren (K_i (*trans*) = 78.6 ± 3.4 nM, K_i (*cis*) = 7.99 ± 0.79 nM) um 50% im Vergleich zu unserem vorigen AMPB-Peptid aufwies. Dies bestätigte, dass cyclische Grundgerüste zu erhöhten Modulationspotentialen führen als lineare Analoga. Im Gegensatz zum linearen AMPB-Peptid zeigte das *cis* Isomer sowohl eine höhere WDR5-Affinität, als auch bessere MLL1-Inhibierungskapazitäten im Vergleich zum *trans* Isomer: IC₅₀ (*trans*) = 6.68 ± 3.1 μM; IC₅₀ (*cis*) = 0.649 ± 0.3 μM. Kristallisationsstudien sowie molekulare Dynamik-Simulationen (MD) kombiniert mit virtuellen Dockingstudien (VD), Pull-Down Assays und Wasserstoff-Deuteriumaustausch (Hydrogen Deuterium Exchange, HDX) verifizierten die Verbesserung der WDR5-Affinitäts Differenz sowie dessen Auswirkung auf die gesamte MLL1-Komplex Topologie. Mit Hilfe von HDX-MS Experimenten konnten wir zeigen, dass der Aktivitätsverlust durch Zugabe unseres *cis* Inhibitors unter anderem durch allosterische Umlagerungen der WDR5-RbBP5 Bindungsstelle hervorgerufen wird. Dies hat möglicherweise zur Folge, dass das RbBP5-Ash2L-Dimer nicht weiter in der Lage ist, die katalytisch aktive Konformation der SET1-Domäne von MLL1 zu stabilisieren, was zu deren Aktivitätsverlust führt. Des Weiteren besitzen unsere cyclischen Peptidkandidaten verbesserte photochrome Eigenschaften und verhalten sich wie ein quasi-bistabiles System. In der Tat sind beide Isomere bis zu fünf Monate beständig. Darüber hinaus haben wir untersucht, ob der Verlust der enzymatischen Aktivität von MLL1, die Viabilität menschlicher Leukämiezellen modulieren kann. Die Aufnahme in die Zelle, mit Hilfe eines Carrier-Peptids, variierte für verschiedene Zelltypen. So erwies sich die MLL-AF9 Zelllinie, welche für die Experimente im *Chemical Science* verwendet wurde, als untauglich. Die besten Ergebnisse wurden hingegen mit

MOLM-13 Zellen erzielt, bei welchen Viabilitätsunterschiede von: *trans*: 37%, *cis*: 0% bei 200 μ M and *trans*: 57%, *cis*: 26% bei 100 μ M festgestellt wurden. Um auszuschließen, dass die beobachtete Toxizität von der *o*F₄Azo Einheit selbst herrührt, wurde das Enantiomer unseres besten cyclischen Peptids, bei dem zusätzlich das Arg mit einem D-Lys substituiert wurde (D-Lys Variante), für die Zell Experimente eingesetzt. Zusätzlich wurde ein extra Waschschrift der Zellen nach der Cargo-Transfektion eingebaut, um sicherzustellen, dass die erfassten Effekte ausschließlich von intrazellulären Mechanismen resultieren. Diese Experimente mit der D-Lys Variante, welche außerdem keine WDR5 Affinität aufweist, bewiesen, dass die Toxizität nicht von der *o*F₄Azo Einheit herrührt, da bei 100 μ M eine geringe Toxizität beobachtet wurde, welche noch geringer für das aktive *cis* Isomer ausfiel (Viabilität: *trans*: 70%, *cis*: 91%). Im Gegensatz dazu verursachte das aktive *cis* Isomer unseres besten Arg-beinhaltenem Kandidaten unter gleichen Bedingungen eine hohe Toxizität, während das *trans* Isomer gleichbleibende Effekte aufwies (Viabilität: *trans*: 66%, *cis*: 24%). Schließlich wollten wir den in vivo Effekt unseres fotoschaltbaren Inhibitors auf die Modulierung von Hämatopoese untersuchen. Aufgrund der Konservierung von MLL1 haben wir uns hierbei für das Zebrafisch Modell entschieden. Dort ist bekannt, dass MLL1 Knockout zur Unterdrückung der Hämatopoese führt, was anhand eines verminderten Blutflusses mit Hilfe von *o*-Dianisidine Färbung beobachtet werden kann. Unsere Experimente konnten bestätigen, dass das *cis* Isomer die Hämatopoese in 3 dpf (days post fertilization) Zebrafischen beeinflusst (*o*-Dianisidine Färbung: 60% stark gestört, 40% keine Färbung). Außerdem wiesen *cis* inkubierte Zebrafische gekrümmte Körperachsen, Herzödeme und mangelndes Reaktionsvermögen auf. Im Gegensatz dazu wiesen 100% der *trans* inkubierten, sowie DMSO- und Wasser Kontrollfische keine dieser Defekte auf. Die Isomerisierung unseres Peptids mit sichtbarem Licht ermöglichte uns, zu untersuchen ob *trans* inkubierte Zebrafische in situ aktiviert werden können. Mit Hilfe dieses Experiments konnten wir beweisen, dass die signifikanten Unterschiede zwischen den Isomeren tatsächlich durch in situ Bestrahlung zum aktiven *cis* Isomer erreicht werden können (*o*-Dianisidine Färbung: 3% intakt, 30% gestört, 67%: keine Färbung). Zusammengefasst präsentieren unsere Ergebnisse die Möglichkeit der in vivo Modulierung von MLL1-abhängiger Hämatopoese unter Verwendung cyclischer fotoschaltbarer Peptide. Dieses Projekt wurde kürzlich eingereicht.

4 Cumulative Part

4.1 Photoswitchable peptides for spatiotemporal control of biological functions



Bioactive photoswitchable peptides as excellent photochemical tools for cellular processes.

This chapter was published as:

L. Albert, O. Vázquez*

Photoswitchable peptides for spatiotemporal control of biological functions

Chem. Commun. **2019**, 55, 10192-10213.

Copyright Royal Society of Chemistry 2019. Material in this chapter was reproduced from the above mentioned reference with permission from The Royal Society of Chemistry. This article was a contribution to the special themed collection: 2019 Emerging Investigators, published as a *Feature Article*.

In the following sections, the molecules included in this manuscript, will be named as **MA-X**, in which MA means manuscript A, and **X** means number of compound as given in this manuscript.



Cite this: *Chem. Commun.*, 2019, 55, 10192

Photoswitchable peptides for spatiotemporal control of biological functions

Lea Albert  and Olalla Vázquez *

Light is unsurpassed in its ability to modulate biological interactions. Since their discovery, chemists have been fascinated by photosensitive molecules capable of switching between isomeric forms, known as photoswitches. Photoswitchable peptides have been recognized for many years; however, their functional implementation in biological systems has only recently been achieved. Peptides are now acknowledged as excellent protein–protein interaction modulators and have been important in the emergence of photopharmacology. In this review, we briefly explain the different classes of photoswitches and summarize structural studies when they are incorporated into peptides. Importantly, we provide a detailed overview of the rapidly increasing number of examples, where biological modulation is driven by the structural changes. Furthermore, we discuss some of the remaining challenges faced in this field. These exciting proof-of-principle studies highlight the tremendous potential of photocontrollable peptides as optochemical tools for chemical biology and biomedicine.

Received 30th April 2019,
Accepted 1st August 2019

DOI: 10.1039/c9cc03346g

rsc.li/chemcomm

Introduction

Cellular processes are extraordinarily complex, requiring a strict spatiotemporal organization by well-tuned signalling networks. Light-triggered strategies are generally non-invasive and

offer precise control, facile modulation and orthogonality. Consequently, the use of light for the manipulation of biological phenomena had a significant impact on advancing 21st-century biology. Such implementation can be purely genetic,^{1–3} purely chemical^{4–9} or a combination of both by merging photo-sensitive small molecules with native or genetically engineered proteins.¹⁰ Despite the remarkable success of optogenetics, as chemists, we cannot resist the allure of reversibly controlling biological

Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Straße 4, 35043, Marburg, Germany. E-mail: olalla.vazquez@staff.uni-marburg.de



Lea Albert

Lea Albert received her Master degree in Chemistry from the University of Marburg, in 2015. She is currently pursuing her PhD under the supervision of Asst. Prof. Vázquez at the same university. In 2018 she was awarded with a Fulbright scholarship to join the lab of Prof. Yali Dou at the University of Michigan (USA). Abroad she broadened her knowledge in sophisticated biological and biochemical assays. Her research interest is focused on the

development, synthesis and evaluation of photoswitchable peptides to control epigenetic protein–protein interactions.



Olalla Vázquez

Olalla Vázquez obtained her PhD in Chemistry at the University of Santiago de Compostela (Prof. Mascareñas and Prof. Vázquez) in 2010. As FPU fellow, she was a visiting PhD student at both Harvard University (Prof. Verdine; 2006) and Humboldt Universität zu Berlin (Prof. Seitz; 2008). In 2011 she received a Marie Curie fellowship to conduct her postdoctoral studies (Prof. Seitz). She is currently Assistant Professor at the University

of Marburg working on optochemical tools for biological processes. During her career path, Olalla has received several national and international awards, including: the Research Lilly Award (2009), European Young Chemist Award (EYCA 2012), Fulbright-Cottrell Award (2016), Next Generation Optogenetics (2017) and MarBiNa Förderpreis (2018).

functions with photoswitchable small molecules. This strategy is even more appealing today, with the emergence of novel photopharmacological targets.^{11,12}

Previous reviews cover developments in the reversible photochemical regulation of biological systems, focusing on either a specific chromophore (azobenzenes,^{13–15} hetero-arylazodyes,¹⁶ diarylethenes¹⁷ spiropyran,¹⁸ and hemithioindigos^{19,20}) or oligonucleotide binders.²¹ More recent reviews have only briefly addressed proteins and peptides,^{7,9} or have limited the scope to the canonical azobenzene photoswitches.^{13,22} This contrasts with the current ‘peptide tidal wave’²³ according to which, peptides are efficient biological modulators^{24,25} and promising drugs.²⁶ In particular, synthetically-modified peptides capable of circumventing the limitations of natural peptides, while producing the same (or improved) biological effect—peptidomimetics—are the perfect tool for chemical biology and drug discovery.²⁷ Along these lines, grafting light-driven molecular transducers—converters of signals into different outputs—onto peptides, does not only provide a strategy to reduce off-target effects, but also spatiotemporal resolution to modulate and understand biological functions.

Here, we provide a complete and up-to-date discussion of photoswitchable peptides from structural to activity studies. Existing functional applications of these chemically modified peptides are used to highlight their potential, identify challenges, and highlight future directions.

Molecular photoswitches (PSs)

Molecular photoswitches (PSs) are chromophores capable of isomerizing by illumination at a specific wavelength between two (or more) species with different absorption spectra at their photostationary states (PSS).²⁵ We limit this review to non-natural compounds that have been incorporated (Fig. 1) into a peptide (Fig. 2): azobenzenes, stilbenes, hemithioindigos (HTI), overcrowded alkenes, fulgimides, spiropyran and diarylethenes (DAE). The group of Goodman first introduced PS 7 in peptides as amino acids with azobenzene-containing side chains²⁸ (Fig. 1A). This method was also used for HTI 22. Another strategy is based on Fmoc-protected backbone PS (Fig. 1B); there are Fmoc-azobenzenes 1–3, 15; -stilbenes 17–19, -DAEs 27–30, and -overcrowded alkene 23, to modify linear or cyclic peptides.

For helical peptides, the predominant approach is cross-linking halogenacetamide-modified PSs to cysteine side chains (Fig. 1C) at the desired spacing *i.e.*, cysteine mutations between the *i* and *i + n* positions. Such peptides are known as stapled

peptides, where one isomer ideally enhances the helical content of the parental peptide, while the other leads to unfolding. Mainly azobenzenes were used (8–14), although the HTIs (20–21) were also connected *via* amide crosslinkers. Isomer preference to stabilize the helical conformation depends on both, the cross-linking site and the length of the PS backbone (Table 2).

An ideal PS should meet the following requirements:²⁹ a fast switching process with sufficient thermal stability; high extinction coefficients and quantum yields; activity at penetrating and innocuous wavelength; stable and non-toxic (along with metabolites); different biological effects between isomers. As in other areas, azobenzenes are the prominent PS used in peptides due to their suitable photochemical properties, straightforward synthesis and the large conformational change upon irradiation. The classical azobenzenes, such as the backbone switches: 4-[(4-aminomethyl)phenylazo]benzoic (AMPB 1),³⁰ [3-(3-amino-methyl)phenyl-azo]phenyl-acetic acid (AMPP 2),³¹ 4-[(4-amino)phenylazo]benzoic acid (APB 3) and the crosslinking ones: sulfonated 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido)azo-benzene (BSBCA 11) and its derivatives 6-[2-sulfonato-4-(chloroacetyl amino)-phenylethynyl]benzene sulfonic acid (DDPBA 12) and 4,4'-bis(4-(2-chloroacetamido)phenyl)di-azanyl-biphenyl-2,2'-disulfonate (BPDBS 13) are readily synthesized. These classical azobenzenes typically produce up to 80% *cis* (thermal half-life ~2–5 days, Table 1) upon irradiation with the respective wavelength (Fig. 2) and up to 99% *trans* upon thermal relaxation (Table 2). The half-life of the thermally less stable isomer, however, is also dependent on the peptide structure into which the PS is introduced. For example, incorporation into cyclic peptides might lead to longer half-lives due to conformational constraints (Table 2). In recent years there have been great advances in developing and synthesizing³² tuned derivatives^{14,33–35} to gain optimized photochromic properties. Thus, introducing electron donating, or push/pull substituents in the *para*-position allows longer wavelength absorption, that is, accompanied by fast thermal relaxation rates (seconds). To overcome this obstacle, additional *ortho*-substituents, like -OMe, -Cl, or -F (as in the tetra-*ortho*-fluoroazobenzene *o*F₄Azo 15), were introduced, yielding more stable *cis* isomers (Tables 1–3). The -F derivatives display the slowest relaxation rates (years). Due to the separation of the *n*- π^* bands, almost quantitative photoconversion can be achieved (Table 3, entry 20). Stability against intracellular reducing agents such as glutathione (GSH) was demonstrated for -Cl, -F and -SEt derivatives.^{14,36} A subclass of azobenzenes are azopyrazoles (16), offering straightforward synthesis¹⁶ and broad structural diversity leading to wide wavelength coverage.¹⁶ As for azobenzenes, their photochromic properties largely depend on the substitution pattern of the benzene ring.³⁷ Thus, it was demonstrated that *m*-Me, *p*-OMe and *p*-OCF₃ substituents lead to quantitative photoconversion into both directions, while *o*-Me substitution leads to an increased thermal half-life (21 days, compared to 4 days of *p*-Me, Table 1). (1-Methyl-4-pyrazolyl)NNPh^{16,38} even displays a tremendously increased half-life of 1000 days. Contrarily to *o*-Me azobenzenes, a wide variety of azopyrazoles was proven to be stable against GSH reduction, which might be due to their higher electron density of the N=N double bond

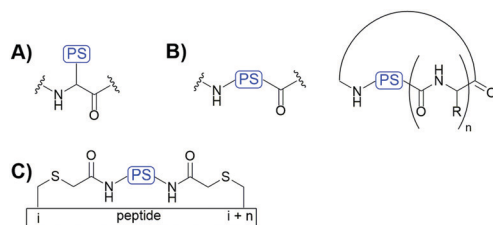


Fig. 1 PS incorporation methodologies into peptides: (A) side chain approach, (B) backbone strategy and (C) crosslinking method; PS = photoswitch.

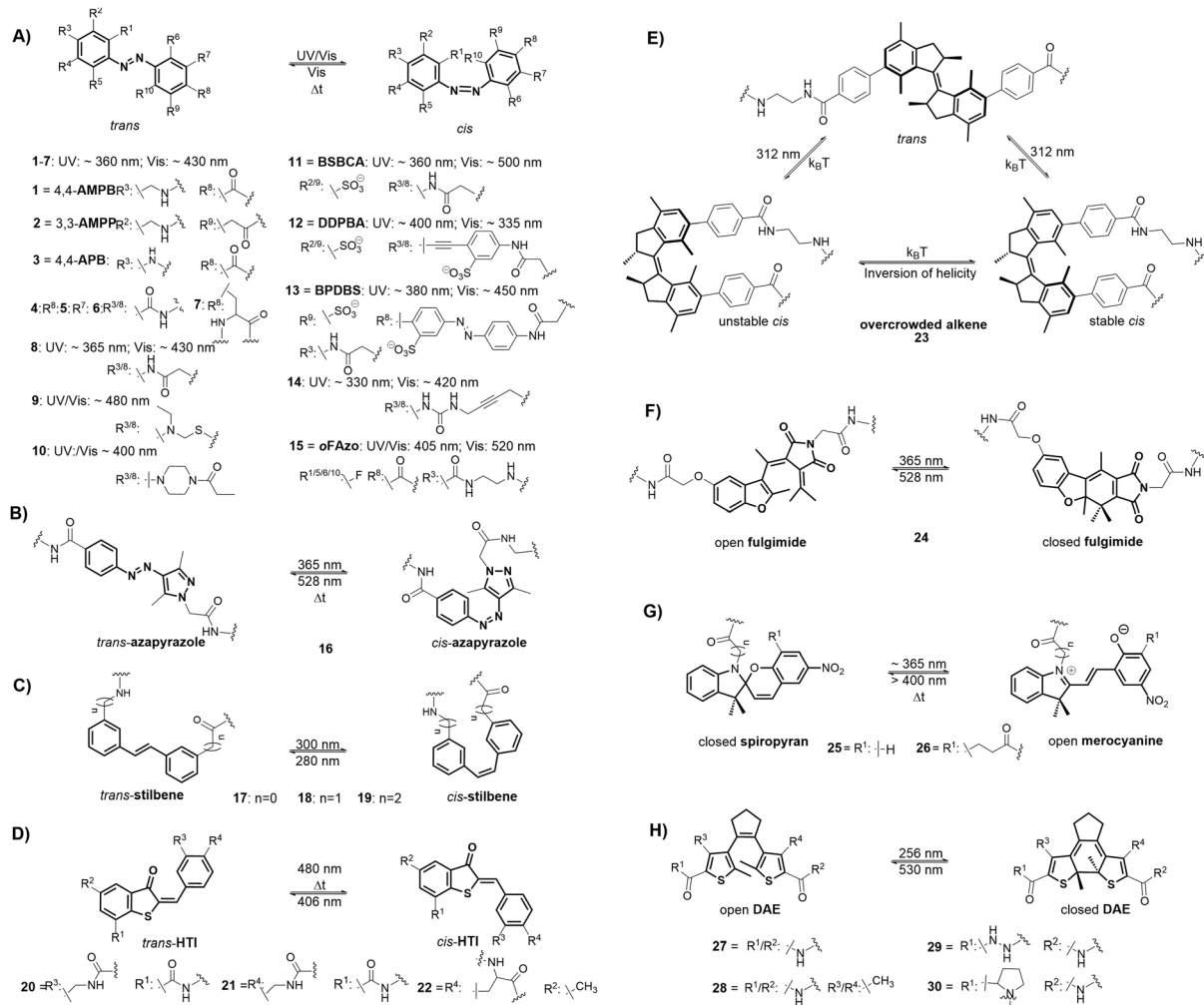


Fig. 2 Structure and photochromism of the photoswitches (PSs) grafted onto peptide scaffolds to control biological functions, and acronyms of classical used azobenzenes: (A) azobenzene; (B) azapyrazole; (C) stilbene; (D) hemithioindigo (HTI); (E) overcrowded alkene; (F) fulgimide; (G) spiropyran; (H) diarylethene (DAE).

because of the electron-rich pyrazole.³⁷ So far, the only azopyrazole introduced into a peptide, needs UV irradiation (Fig. 1) to reach its *cis* PSS exhibiting almost quantitative conversion into both directions (88% *cis*, 95% *trans*; Table 5, entry 7). As observed for azobenzenes and azopyrazoles, stilbenes (17–19) are planar in their *trans* PSS and non-planar in their *cis* state. In contrast to azobenzenes, stilbenes are thermally stable in both isomeric forms, requiring UV irradiation to trigger isomerization into both directions (Fig. 2), leading to acceptable ratios in their PSSs (Table 2). On the upside, they are always stable against GSH reduction.³⁹ However, the *cis* form is prone to undergo cyclization followed by irreversible oxidation to phenanthrenes. This hinders their application in biological contexts.⁵³ Contrarily to azobenzenes and stilbenes, both HTI isomers (20–22) are planar and can intercalate DNA. Furthermore, thermal relaxation of the *trans* to *cis* isomer is relatively slow (half-life of *trans*: several hours, some derivatives even several days; Table 1).^{46,54} In addition, HTIs hold the benefit of ultrafast visible light-driven photoisomerization in both directions (Fig. 1),⁵⁵ resulting in up to 78% *trans* and 99% *cis*

in their PSSs (Table 2). Furthermore, GSH reduction has not been reported to be an issue for HTI PSSs. All those PSs exert facile synthesis routes and are readily accessible. On the other hand, the synthesis of overcrowded alkenes as **23** is more complex; the crucial step is the formation of the central, sterically demanding double-bond. Once obtained, they can interconvert between several conformations under irradiation of the same single wavelength⁵⁶ (Fig. 2) yielding up to 93% *cis* (with a thermal half-life of 4 days, Table 1)⁴¹ and 99% *trans* in the respective PSS (Table 6, entry 2). This change of conformation is accompanied by a change of helicity around the double bond. Neither for overcrowded alkenes, nor for ring-chain isomerizable PSs like fulgimides **24**, spiropyranes **25–26** and diarylethanes (DAEs) **27–30**, reduction by GSH was reported to be a concern. The relevance of spiropyranes arises from the large change of polarity between isomers resulting in drastic effects on hydrophilicity/hydrophobicity. Their reversible isomerization is not only driven by light (UV-light required for open merocyanine; Fig. 1), but is also influenced through different solvents, metal ions, pH, temperature and mechanical forces.¹⁸

Table 1 Thermal half-lives of the non-thermodynamically stable isomers of selected PS; n.p. = not provided; aq. sol. = aqueous solution; $\tau_{1/2}$ = half-life

	$\tau_{1/2}$	Solvent/ T
PS		
Spiropyrane 25/26 ⁴⁰	n.p.	n.p.
Overcrowded alkene 23 ⁴¹	4.1 d	n.p./25 °C
HTI 20/21 ⁴²	hours	n.p.
Azopyrazole 16 ⁴³	8.1 d	aq. sol./n.p.
PS not incorporated into peptide		
Classical spiropyrane ⁴⁴	0.25 min	THF/n.p.
5'-Phenylpyrrole HTI ⁴⁵	9.3 h	Toluene/80 °C
HTI-based <i>N</i> -formyl amino acid methyl ester ⁴⁶	5 h	MeCN/25 °C
<i>p</i> -OMe-phenyl-4'-NMe2-benzothiophen-HTI ⁴⁷	30 d	n.p./25 °C
1-Methyl-5-phenylazoimidazole ⁴⁸	528 h	CDCl ₃ /25 °C
(1-Methyl-4-pyrazolyl)NNPh ^{16,38}	1000 d	DMSO/25 °C
(1,3,5-Trimethyl-4-pyrazolyl)NNPh ^{16,38}	10 d	DMSO/25 °C
(1,3,5-Trimethyl-4-pyrazolyl)NNPh- <i>o</i> -Me ³⁷	21 d	H ₂ O/25 °C
(1,3,5-Trimethyl-4-pyrazolyl)NNPh- <i>m</i> -Me ³⁷	13 d	H ₂ O/25 °C
(1,3,5-Trimethyl-4-pyrazolyl)NNPh- <i>p</i> -Me ³⁷	4 d	H ₂ O/25 °C
Plain azobenzene ⁴⁹	5 d	MeCN/25 °C
BSBCA 11 ⁵⁰	20 min	aq. sol./25 °C
2,2'-Dimethoxy-4,4'-diamidoazobenzene ⁵¹	1.9 d	DMSO/25 °C
2,2',6,6'-Tetramethoxy-4,4'-diamidoazobenzene ⁵¹	14 d	DMSO/25 °C
2,2',6,6'-Tetrafluoroazobenzene ⁵²	700 d	DMSO/25 °C

Generally, both isomers display short thermal half-lives (range of minutes/seconds, Table 1) leading to isomeric mixtures, where the open form is hydrolysed in aqueous solutions.⁵⁷ In the context of peptides, they have been used to modulate the α -helical conformation. Contrarily, the main advantage of fulgimides and DAEs is, that they exert two thermally stable isomers where isomerization needs to be driven by light (UV-light required to reach PSS of closed form; Fig. 1) resulting in up to 85% open/49% closed form in the PSSs for fulgimides and up to 98% open/95% closed form for DAEs (Table 5, entry 7). Neither of the DAE isomers undergoes *cis/trans* isomerization in the open form, nor are oxidized in the closed form.⁵⁸ The DAE moiety perfectly mimics β -turns, thus, has together with azobenzenes, HTIs, stilbenes and overcrowded alkenes, been used to modulate β -hairpin structures.⁵⁹ Alongside azobenzenes, DAEs are the most promising PSs for biological modulation today.⁵⁶ Other PSs such as thiophen-fulgides have not yet been introduced into a peptide scaffold.

Structural studies of photoswitchable peptides

Initial studies of photoresponsive peptides were focused on conformational changes caused by PS inclusion in a defined structure. The characterization of those changes can be studied by circular dichroism (CD), NMR, infrared spectroscopy (IR) and molecular dynamics (MD). The first PS introduced into a peptide backbone to study structural changes (Table 2, entry 1),⁶⁰ dated from 1995 by Chmielewski, who demonstrated increased β -turn content of a cyclic model peptide with the *cis* isomer of AMPB **1**. Following this initial breakthrough, stilbenes **17–19** and HTIs **20–22** have been incorporated into *de novo* designed model peptides as turn mimetics to photoregulate the β -turn content (Table 2, entry 2). Introduction of rigid HTIs **20** and **21**

showed stronger conformational directing effects compared to the more flexible stilbenes **18** and **19**.^{42,61} Additionally, 3,4-AMPB, AMPP **2**, stilbene **19**, HTIs **20** and **21** were used to demonstrate possible control of β -hairpin conformation of peptides derived from the tryptophan zipper (TrpZip) motif (Table 2, entries 3 and 4; β -amyloid section),^{61–64} and amino acid strands extracted from the TATA-box (Table 2, entries 5 and 6)^{61,65} and the β -finger of the nNOS peptide (Table 2, entries 7 and 8; enzyme section).^{66,67} Due to the high flexibility of **19**, neither of its isomers induces β -turn formation themselves. However, this property allows the adoption of the overall peptide conformation, maintaining the TrpZip hairpin structure in the *cis* state, while the *trans* form disrupts such a motif by 1.67-fold (Table 2, entry 3). Contrarily, improved β -turn induction is demonstrated using 3,4-AMPB; *i.e.*, increasing β -hairpin-content of the *cis* state up to 2.75-fold compared to *trans* TrpZip (Table 2, entry 4). Furthermore, the 3,4-AMPB incorporated into the cyclic nNOS sequence lead to the largest photoconversion reaching 90% *cis*. Stilbene **19** reached higher conversion in the cyclic TATA-box sequence than in the linear one (80% *versus* 63%). Additionally, the linear sequence did not show β -hairpin content in either isomeric form, while the cyclic version showed increased β -hairpin formation of the *cis* state stilbene **19** (while HTI **21** did not in either state) compared to the *trans* one. This highlights the significance of cyclic peptides as a conformational trap. To further elucidate the design principles to modulate conformational states of cyclic peptides, Moroder used cyclic sequences derived from the thioredoxin reductase (TRR, Table 2, entries 9–12) and the protein disulphide isomerase (PDI, Table 2, entries 13 and 14). These synthesized natural motives both contained a β -turn region and 3₁₀-helical turn motif and were discussed in detail in a review by this group.⁶⁹ Briefly, APB **3** and AMPB **1** were backbone implemented (Fig. 1B) into cycles and additionally constrained by disulphide bridging. Initial studies revealed that these bicyclic systems, incorporating APB moiety **2**, generated

Table 2 Photoswitchable peptides used in conformational studies. c[XXX]: cyclic peptides; bc[XXX]: bicyclic peptides; a disulphide bridge. For α -helical peptides, the crosslinking sites are depicted with a line from residue i to $i + n$; **X**: incorporated PS; n.p. = not provided in publication; aq. sol. = aqueous solution

Entry	Peptide sequence	PSS ratio	Structural change between isomers	Characterization technique
1 ⁶⁰	c[AAGGNAAX] X = AMPB 1 ⁶⁰	cis: 93% ($\tau_{1/2}$ = 7 d in DMSO, 25 °C) trans: 99%	cis: increased β -turn content	NMR (H, COSY, NOESY), MD
2 ⁴²	Ac-V-X-CONH-R ¹ R ¹ : CH ₃ ; X = stilbene 17, 18, 19; R ¹ : CH ₃ /H; X = HTI 20, 21 ⁴²	n.p. 20/21: trans: $\tau_{1/2}$ = several hours	cis: increased β -turn content (only in MD)	NMR (H, COSY, TOCSY, NOESY, ROESY), MD
3 ⁶¹	H ₂ N-KWTWK-X-EWTWS-COOH X = stilbene 19 ⁶¹	cis: 76% trans: 99%	trans: disrupts β -hairpin content (1.67-fold)	CD, NMR (H, NOESY, ROESY)
4 ⁶²⁻⁶⁴	H ₂ N-SWTWE-X-KWTWK-CNOH ₂ X = AMPB 2 ⁶²⁻⁶⁴	cis: 80% trans: 80%	cis: increased β -hairpin content (2.78-fold)	CD, IR, NMR (H, COSY, NOESY, ROESY), MD, T-jump
5 ^{61,65}	(H ₃ C-CH ₂) ₂ -N-(CH ₂) ₂ -CO-LAII-T-X-IVLL-CONH ₂ X = stilbene 19 ^{61,65}	cis: 63% trans: 99%	No change of β -hairpin content	CD, NMR (H, NOESY, ROESY), MD
6 ^{61,65}	c[IVLL ^D PGLAII-T-X] X = stilbene 19; ^{61,65} X = HTI 21 ⁶¹	19: cis: 80% 21: trans: 78%	19: cis: increased β -hairpin content; 21: less	CD, NMR (H, NOESY, ROESY), MD
7 ⁶⁷	c[TVRIKPTGDGTFTELHT-X] X = HTI 20 ⁶⁷	n.p. trans: ($\tau_{1/2}$ = hours)	cis: increased β -hairpin content	IR
8 ⁶⁶	c[LETFITGDGTPKTI-R-X] X = 3,4-AMPB ⁶⁶	cis: 90% ($\tau_{1/2}$ = 25 d in aq. sol.) trans: 99%	cis: increased β -hairpin content	IR
9 ⁶⁸⁻⁷¹	c[Ac(tBu)ATC(tBu)DGF-X] X = APB 3; X = AMPB 1 ⁶⁸⁻⁷¹ bc[ACATCDGF-X]	cis: 78% ($\tau_{1/2}$ = 3: 39 h; 1: 99 h in aq. sol., 25 °C) trans: 99%	trans: 2 turns, 1 helical turn; cis: S-shaped backbone	NMR (H, NOESY)
10 ⁶⁹⁻⁷²	X = APB 3; X = AMPB 1 ⁶⁹⁻⁷² bc[KCATCDKK-X] X = AMPB 1 ^{69,73-75}	cis: 78% ($\tau_{1/2}$ = 3: 45 h; 1: 87 h in aq. sol., 25 °C) trans: 99%	trans: γ -turns, α -helical turn; cis: S-shaped backbone	NMR (H, NOESY), MD
11 ^{69,73-75}	X = APB 3; X = AMPB 1 ⁶⁹⁻⁷² bc[KCATCDKK-X] X = AMPB 1 ^{69,73-75}	cis: 80% ($\tau_{1/2}$ = 102 h in aq. sol., 25 °C) trans: 99%	trans: β/γ -turns; both: α -helical turn	CD, NMR (H, NOESY), MD
12 ^{69,73-75}	c[KSATSDKK-X] X = AMPB 1 ^{69,73-75}	cis: 76% ($\tau_{1/2}$ = 160 h in aq. sol., 25 °C) trans: 99%	Both: β/γ -turns; cis: distorted helical loop	CD, NMR (H, NOESY), MD
13 ⁷⁶	bc[KCGHCCKK-X] X = AMPB 1 ⁷⁶	cis: 80% ($\tau_{1/2}$ = 182 h in aq. sol., 25 °C) trans: 99%	Both: β -turn; cis: 2nd β -turn & γ -turn; trans: extended backbone	CD, NMR (H, COSY, ROESY), MD
14 ⁷⁶	c[KSQHSKCK-X] X = AMPB 1 ⁷⁶	cis: 80% ($\tau_{1/2}$ = 253 h in aq. sol., 25 °C) trans: 99%	Both: disordered, trans: extended backbone	CD, NMR (H, COSY, ROESY), MD
Following: side chain cyclized peptides via crosslinking with PS X				
15 ⁷⁷	Ac-AACARAAARAAACRA-CONH ₂ X = AzO 8 ⁷⁷	cis: 65% trans: 80%	trans: increased α -helix content (4.62-fold)	CD, IR
16 ⁷⁸	Ac-AACAKAAAKAAACK-CONH ₂ X = AzO 8 ⁷⁸	cis: 75% trans: 99%	trans: increased α -helix content	IR, MD
17 ⁷⁹	Ac-EAARACARECAARQ-CONH ₂ X = AzO 8 ⁷⁹	cis: 90% ($\tau_{1/2}$ = 36 min in aq. sol., 25 °C) trans: 99%	cis: increases α -helix content (3.53-fold)	CD, MD
18 ⁷⁹⁻⁸¹	Ac-EACAREAAAREAACRQ-CONH ₂ X = AzO 8 ^{79,80} ; X = AzO 9 ⁸¹	cis: 80% ($\tau_{1/2}$ = 12 min in aq. sol., 25 °C) trans: 99%	trans: increases α -helix content (2.53-fold)	CD, MD, IR
19 ^{81,82}	Ac-EACARY-Aib-AACEAAARQ-CONH ₂ X = AzO 9 ⁸¹ ; X = AzO 10 ⁸²	cis: 50% ($\tau_{1/2}$ = 6.5 ms in aq. sol., 25 °C) trans: 99%	cis: increase of α -helix content	CD, MD
20 ⁸²	Ac-WGCA(EAAR) ₂ EAAARQ-CONH ₂ X = AzO 10 ⁸²	cis: 50% ($\tau_{1/2}$ = 1.8 s in aq. sol., 25 °C) trans: 99%	trans: increase of α -helix content	CD
21 ⁸²	Ac-WGAC(EAAR) ₂ EAAARQ-CONH ₂ X = AzO 10 ⁸²	cis: 50% ($\tau_{1/2}$ = 3.8 s in aq. sol., 25 °C) trans: 99%	trans: increase of α -helix content	CD, MD
22 ⁸²	Ac-WGEACAREAAAREAAARQ-CONH ₂ X = AzO 10 ⁸²	cis: 50% ($\tau_{1/2}$ = 1.7 s in aq. sol., 25 °C) trans: 99%	trans: increase of α -helix content	CD, MD
23 ⁸³	Ac-EACAREAAAREAAARQ-CONH ₂ X = BSECA 11 ⁸³	cis: 85% ($\tau_{1/2}$ = 35 min in aq. sol., 25 °C) trans: 99%	trans: increased α -helix content (1.86-fold)	CD, NMR (H, NOESY, ROESY)

Table 2 (continued)

Entry	Peptide sequence	PSS ratio	Structural change between isomers	Characterization technique
24 ⁸⁴	H ₂ N-WECAAREAAACR(EARR) ₄ Q-CONH ₂ X = BSBCA 11 ⁸⁴	<i>cis</i> : 80% ($\tau_{1/2}$ = 100 min in aq. sol., 25 °C) <i>trans</i> : 99%	<i>cis</i> : increased α -helix content (1.37-fold)	CD
25 ⁸⁴	H ₂ N-W(EAAR) ₂ ECAAAREAAACR(EAAAAR) ₂ Q-CONH ₂ X = BSBCA 11 ⁸⁴	<i>cis</i> : 80% ($\tau_{1/2}$ = 100 min in aq. sol., 25 °C) <i>trans</i> : 99%	<i>cis</i> : increased α -helix content (1.97-fold)	CD
26 ⁸⁴	H ₂ N-W(EAAR) ₃ ECAAAREAAACRQ-CONH ₂ X = BSBCA 11 ⁸⁴	<i>cis</i> : 80% ($\tau_{1/2}$ = 100 min in aq. sol., 25 °C) <i>trans</i> : 99%	<i>cis</i> : increased α -helix content (1.20-fold)	CD
27 ⁸⁴	H ₂ N-WEAAGREAAAAREAAAREAAACR(EAAAAR) ₂ Q-CONH ₂ X = BSBCA 11 ⁸⁴	<i>cis</i> : 80% ($\tau_{1/2}$ = 100 min in aq. sol., 25 °C) <i>trans</i> : 99%	<i>cis</i> : increased α -helix content (2.17-fold)	CD
28 ⁸⁴	H ₂ N-W(EAAR) ₂ ECAGREAAACR(EAAAAR) ₂ Q-CONH ₂ X = BSBCA 11 ⁸⁴	<i>cis</i> : 80% ($\tau_{1/2}$ = 100 min in aq. sol., 25 °C) <i>trans</i> : 99%	<i>cis</i> : increased α -helix content (1.60-fold)	CD
29 ⁸⁴	H ₂ N-W(EAAR) ₂ ECAAAREAAAREAAACRQ-CONH ₂ X = BSBCA 11 ⁸⁴	<i>cis</i> : 80% ($\tau_{1/2}$ = 100 min in aq. sol., 25 °C) <i>trans</i> : 99%	<i>cis</i> : increased α -helix content (2.50-fold)	CD
30 ⁸⁵	H ₂ N-WGAC(EAAAAR) ₂ EAAACRQ-CONH ₂ X = DDPBA 12 ⁸⁵	<i>cis</i> : 15% ($\tau_{1/2}$ = 159 min in aq. sol., 20 °C) <i>trans</i> : 99%	<i>cis</i> : increased α -helix content	CD
31 ^{85,86}	H ₂ N-WGAC(EAAAAR) ₄ CAQ-CONH ₂ X = BSBCA 11 ⁸⁴	<i>cis</i> : 15% ($\tau_{1/2}$ = 12: 102 min; 13: 16 min in aq. sol., 20 °C) <i>trans</i> : 99%	<i>trans</i> : increased α -helix content (1.12-fold)	UV/vis, CD, MD
32 ⁸⁶	X = DDPBA 12, ⁸⁵ X = BPDBS 13 ⁸⁶ H ₂ N-WAGC(EAAAAR) ₃ EAAACREAAQ-CONH ₂ X = BPDBS 13 ⁸⁶	n. p. <i>cis</i> : $\tau_{1/2}$ = 29 min in aq. sol., 20 °C <i>trans</i> : 99%	<i>trans</i> : increased α -helix content (1.10-fold)	UV/vis, CD, MD
33 ⁸⁶	H ₂ N-WGAC(EAAAAR) ₂ EAAACREAAAREAAQ-CONH ₂ X = BPDBS 13 ⁸⁶	n. p. <i>cis</i> : $\tau_{1/2}$ = 10 min in aq. sol., 20 °C	<i>cis</i> : increased α -helix content (1.06-fold)	CD, MD
34 ⁸⁶	Ac-WGAC(EAAAAR) ₂ C(AAARE) ₂ AQ-CONH ₂ X = BPDBS 13 ⁸⁶	n. p. <i>cis</i> : $\tau_{1/2}$ = 12 min in aq. sol., 20 °C	<i>cis</i> : increased α -helix content (2.54-fold)	CD, MD
35 ⁸⁷	Ac-WGAC(EAAAAR) ₂ C(AAARE) ₂ AQ-NOH ₂ X = benzylidene-pyrroline (BP) ⁸⁷	<i>cis</i> : 93% ($\tau_{1/2}$ = 10 h, 50 °C) <i>trans</i> : 73%	<i>trans</i> : increased α -helix content	CD, MD
36 ⁴⁰	Ac-KADAAHAKAAEA-CONH ₂ X = Spiropyrene 26 ⁴⁰	n. p.	Closed form: increased α -helix content (1.29-fold)	CD
37 ^{88,89}	H ₂ N-[EEEE]-CONH ₂ X = Spiropyrene 25	n. p.	Closed: increased α -helix content (by 85%)	CD
38 ⁹⁰	Ac-(GP-Hyp) ₃ GC-Hyp-PC(GP(Hyp)) ₅ GG-CONH ₂ X = Azo 14 ⁹⁰	<i>cis</i> : 35% <i>trans</i> : 99%	<i>trans</i> : stable triple helix; <i>cis</i> : unfolding of helix	CD, NMR (H, NOESY, FHSQC), MD

so called “frustrated” systems with little conformational preferences for either isomer. Replacement with the more flexible AMPB **1**, led to specific folding patterns of the *trans* isomers (Table 2, entries 9–12). In 2006 Moroder reported, for the first time, a peptide model where both isomers exist in a specific conformational state (Table 2, entry 13).⁷⁶ NMR experiments verified that both isomers exhibit different but defined structures of the bicycle, demonstrating the feasibility of such a model for allosteric rearrangements.

Contrary to these examples, α -helical peptides were exclusively studied by crosslinking (Fig. 1C) mutated cysteines mainly derived from the *de novo* designed Ac-(EAAAR)_nA-NH₂ sequence^{91,92} (Table 2, entries 15–35). This exists in a monomeric helical form in aqueous solution. Most of these examples rely on the use of azobenzenes **8–14**, that have been extensively studied and reviewed by Woolley.⁹³ In those cases the length of the crosslinker as well as its crosslinking sites/spacings in the peptide sequence are essential to predict which isomer stabilizes the α -helix. Thus, by utilization of shorter azobenzene moieties, such as **8–11**, shorter spacings can be chosen, where the *cis* isomer favours the α -helical conformation using the *i* to *i* + 4 and *i* to *i* + 7 spacings, while the *trans* isomer enhances helicity in the *i* to *i* + 11, *i* to *i* + 14 and *i* to *i* + 15 spacings (Table 2, entries 15–29). Using the azobenzene **8** in the *i* to *i* + 11 spacing led to the highest increase of helicity upon switching from the *cis* to the *trans* state (Table 2, entry 15). Furthermore, the crosslinking site was demonstrated to be essential by attaching BSBCA **11** in different positions of the same sequence (Table 2, entries 24–29), proving linking at the middle of the helix to be the most efficient. Additionally, Gly-Ala exchange demonstrated to affect conformational differences. Replacing those shorter PSs by longer derivatives **12** and **13**, in turn, allowing the use of longer spacings. α -helical content using *i* to *i* + 11 and *i* to *i* + 14 substitution is favoured in the *cis* isomer, while the *trans* isomer increases helicity using higher distances as *i* to *i* + 19 and *i* to *i* + 21 (Table 2, entries 30–34), at the expense of drastically reducing isomer differences. Here, the PSS *cis* ratio is decreased to 15% due to conformational constraints. Exchanging the BPDBS **13** with a benzylidene-pyrroline (BP) moiety in the same sequence (*i* to *i* + 11), lead to inversion of α -helical content between isomers and to worse isomer differences (Table 2, entry 35). Therefore, to address conformational modulation of α -helical peptides azobenzenes are at the state of the art. Of note, benzylidene-pyrroline (BP) and spiropyrane **26** were used to control α -helices of peptides too (Table 2, entries 35–37).

As it is clear from Table 2, peptide structures have a great impact on the yield of photoconversion and conformational changes. Thus, the same PS incorporated into different sequences leads to different PSS ratios. The biggest fold-change of conformation between isomers, determined by CD, was achieved by crosslinking azobenzene **8** to the α -helical Ac-(EAAAR)_nA-NH₂ motif either in an *i* to *i* + 4 (Table 2, entry 17) or *i* to *i* + 11 (Table 2, entry 15) spacing. According to Newton's third law, the more strain that is applied to a peptide, the more strain will be applied back to the PS, changing its properties. This might explain why cyclic constrained peptides are preferentially used to modulate conformational states. Here, the number of amino

acids used in the linear precursor peptides needs to be carefully chosen, to enable both cyclization and constraint at the same time. Seven to eight amino acids were found to be ideal for enabling sufficient differences between isomers.⁷³ However, regarding biological applications, even small conformational changes can trigger great biological responses due to their inherently nonlinear nature.

Approaches to modulate biological functions

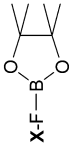
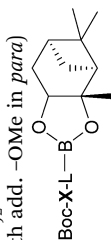
Cellular activity is governed by events at the molecular level, including spatial arrangement of the involved biomolecules.⁹⁴ Peptides are merely shorter fragments of proteins, and often lack the same stable secondary and tertiary structure as their parents. Some peptides possess native function-defining structure such as insulin, oxytocin, or growth factors. Others require interaction modulators to adopt a functional conformation. Whether a peptide has a stable fold or is intrinsically disordered, its conformational changes may trigger biological responses.

Protein–protein interactions (PPIs)

Protein–protein interactions (PPIs) mediate almost every critical molecular process in a cell, and consequently, targeting them, has enormous potential. Besides the fact that some PPIs exhibit hot-spots, the main question has always been how to target their large surfaces lacking binding pockets. Short peptides can modulate 15–40% of all PPIs⁹⁵ being promising PPI modulators.^{96,97} Therefore, it is not surprising that most examples belong to this section (Table 3).

Programmed cell death. The protein complexes Bcl-2/Bak and Bcl-x_L/Bid play a significant role in the regulation and initiation of apoptosis.⁹⁸ In particular, Bcl-x_L and Bcl-2 are inhibitors,⁹⁹ while Bak and Bid act as promoters.^{100,101} This interplay is mediated by dimerization, *via* the conserved Bcl-2 homology (BH) domains, composed of a bundle of amphipathic α -helices. A built-in photoswitch in a BH-peptide would control its helicity by light and, in turn, the interactions between subunits (Fig. 3). An initial study that employed the HTI **20**, crosslinked to a BH3-derived α -helical peptide in *i* to *i* + 4 position (Table 3, entry 1), suggested that such modification conferred helix stabilization and increased stability against proteases.¹⁰² Apoptosis assays in T-cell lymphoma Jurkat A3 cells demonstrated that the HTI-modified peptide increases cell apoptosis to 45% in comparison to the wild-type peptide (18%).¹⁰² The first modulation of Bcl-x_L-Bak/ Bcl-x_L-Bid binding was achieved by Allemann, using a similar approach, crosslinking two cysteine residues with the BSBCA **11** in BH3 peptides on the solvent-exposed face of the Bak and Bid recognising helices (Table 3, entries 2–4).¹⁰³ Fluorescence polarization (FP) assays demonstrated that helix-stabilized peptides (*cis* for *i* to *i* + 4 and *i* to *i* + 7; *trans* for *i* to *i* + 11) are selective for Bcl-x_L with affinities in the nM range and up to 20-fold difference in relation to the helix-destabilized ones. Noteworthy, the *cis* isomer stabilized helical peptides showed greater differences than the ones, where the *trans* was inducing

Table 3 Photoswitchable peptides used to target PPIs. X = PS; n.p. = not provided in publication

Entry	Sequence/PS	PSS ratio	Target	Δ interaction	Technique
1 ¹⁰²	BH3-derived; exact sequence n.p. X = HTI 20 (<i>i</i> to <i>i</i> + 4)	n.p.	Bcl-X _L	n.p.	Viability of Jurkat A3 cells
2 ¹⁰³	H ₂ N-GCYGRALAAEFGDCINR-CONH ₂ X = BSBCA 11	<i>cis</i> : 69% ($\tau_{1/2}$ = 22 min in aq. sol., 15 °C) <i>trans</i> : 99%	Bcl-X _L	<i>cis/trans</i> : 2.29	Affinity (FP)
3 ¹⁰³	H ₂ N-GAVGRCLAIFGDCINR-CONH ₂ X = BSBCA 11	<i>cis</i> : 76% ($\tau_{1/2}$ = 174 min in aq. sol., 15 °C) <i>trans</i> : 99%	Bcl-X _L	<i>trans/cis</i> : 19.6	Affinity (FP)
4 ¹⁰³	H ₂ N-DIIRNIARHLACDYGCDIRSI-CONH ₂ X = BSBCA 11	<i>cis</i> : 80% ($\tau_{1/2}$ = 190 min in aq. sol., 15 °C) <i>trans</i> : 99%	Bcl-X _L	<i>trans/cis</i> : 23.2	Affinity (FP)
5 ¹⁰⁴	Ac-GAVGRCLAIFGDCINR-CONH ₂ X = BSBCA 11	<i>cis</i> : 70% ($\tau_{1/2}$ = 22 min in aq. sol., 15 °C) <i>trans</i> : 99%	Bcl-X _L	<i>cis/trans</i> : 2.29	Cytochrome C release of SU-DHL-4 cells (flow-cytometry)
6 ¹⁰⁶	H ₂ N-DDDDVAALACFAVZIIACAAKKK-CONH ₂ X = BSBCA 11	<i>cis</i> : 75% <i>trans</i> : 99%	Bcl-X _L	<i>trans/cis</i> : 9.00	Inhibition (FP)
7 ¹¹²	X-F-COCOOMe X = 4; 5	4: <i>cis</i> : 54% <i>trans</i> : 72% 5: <i>cis</i> : 75% <i>trans</i> : 68% <i>cis</i> : 74% <i>trans</i> : 95%	α -Chymotrypsin	4: <i>trans/cis</i> : 2.00 5: <i>trans/cis</i> : 3.00	Inhibition (Dixon blot)
8 ¹¹³	 X-F-B X = 4		α -Chymotrypsin	<i>trans/cis</i> : 1.57	Inhibition (enzyme assay)
9 ¹¹⁵	X-F-L-B(OH) ₂ X = 4 (with add. -OMe in <i>para</i>)	<i>cis</i> : 97% ($\tau_{1/2}$ = 3.8 h in aq. sol., 37 °C) <i>trans</i> : 97% <i>cis</i> : 95% <i>trans</i> : 84%	Proteasome	<i>cis/trans</i> : 2.75	Inhibition (fluorescence competition)
10 ¹¹⁴	 Boc-X-L-B X = 7 (with add. -Ph in <i>para</i>)		Proteasome	<i>cis/trans</i> : 5.00	Inhibition (Dixon blot)
11 ¹¹⁷	Fynomer-protein		Chymase	<i>cis/trans</i> : 1.30	Inhibition (enzyme assay)
12 ^{120,121}	Crosslinked with X = BSBCA 11 H ₂ N-E-X-GG-COOH	<i>cis</i> : 80% ($\tau_{1/2}$ = 100 min in aq. sol., 22 °C) <i>trans</i> : 80% <i>cis</i> : 82% <i>trans</i> : 99%	Z1HO1 antibody	<i>trans/cis</i> : 4.10	Affinity (fluorescence titration)
13 ¹²²	H ₂ N-DYK-X-DK-COOH	n.p.	Anti-FLAG-tag antibody M1	<i>trans/cis</i> : 287	Affinity (SPR)
14 ¹²³	X = 3,4-AMPB	<i>cis</i> : 60% <i>trans</i> : 99%	Syk	<i>cis/trans</i> : 2.25	Affinity (SPR)
15 ¹²⁴	H ₂ N-P ³ YTGLG-X-G ¹ YEYL-XONH ₂ X = AMPB 1	<i>cis</i> : 66% <i>trans</i> : 99%	Syk	<i>cis/trans</i> : 3.65	Affinity (SPR)
16 ^{66,125}	H ₂ N-P ³ YTGL-X ¹ YEYL-CONH ₂ X = AMPB 1	<i>cis</i> : 90% ($\tau_{1/2}$ = 25 d in aq. sol.) <i>trans</i> : 99%	α -Synthrophin	<i>cis</i> : K _d : 10 μ M <i>trans</i> : —	Affinity (SPR, ITC)
17 ¹²⁶	c[LETFITGDTPTKIR-X] X = AMPB 1	<i>cis</i> : 88% <i>trans</i> : 80%	Kinesin motility	Complete ON/OFF	Inhibition of gliding motility (%)
18 ¹²⁷	X-IPKAIQASHGR-CONH ₂ X = 4	<i>cis</i> : 88% <i>trans</i> : 80%	Kinesin motility	<i>cis/trans</i> : 8.00	Affinity (motility assay)
19 ¹²⁸	X-AIPKAIR-CONH ₂ X = 4 (with -OMe in <i>para</i>)	<i>cis</i> : 88% <i>trans</i> : 80%	WDR5	<i>cis/trans</i> : 5.20	Inhibition (FP)
20 ³⁶	H ₂ N-SARA-X-VHLRKS-CONH ₂ X = AMPB 1	<i>cis</i> : 75% ($\tau_{1/2}$ > 11 d in aq. sol., 25 °C) <i>trans</i> : 87%	WDR5	<i>cis/trans</i> : 2.60	Inhibition (FP)
21 ¹²⁹	H ₂ N-SARA-X-VHLRKS-CONH ₂ X = OF ₄ Azo 15 Ac-NAAD-Nle-EAAIKHLARR-Nle-AAK-X	<i>cis</i> : 86% ($\tau_{1/2}$ > 7 d in aq. sol., 25 °C) <i>trans</i> : 93% <i>cis</i> : 80% <i>trans</i> : 90%	HIN1 hydrolase	<i>cis/trans</i> : 1.24	Rate constants enzymatic performance
22 ¹³⁰	H ₂ NCO-GARAFAEFRRLHEA-Nle-QAA-X Ac-X-EDDDWDF-CONH ₂ X = stilbene 19 X = stilbene 19	n.p.	R1 reductase	<i>cis/trans</i> : 2.64	Inhibition (FP)

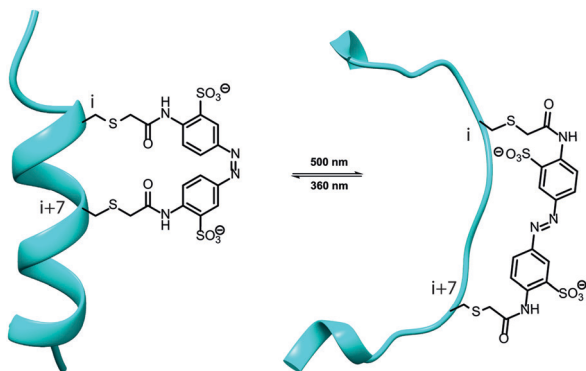


Fig. 3 Scheme of the conformational change illustrated for an α -helical peptide crosslinking the residues i , $i + 7$ with BSCBA **11**.

helicity.¹⁰³ Additionally, the i to $i + 7$ Bak-peptide-based nano-switch photoinduced apoptosis and cytochrome C release (Table 3, entry 5).¹⁰⁴ Later on, the NMR structure of the i to $i + 11$ photoswitchable peptide in complex with Bcl- x_L , was solved, explaining the improved affinities compared to the wild type.¹⁰⁵ Just recently, the group of Giralto and Gorostiza¹⁰⁶ identified generalized template peptides based on α -helices to develop photoswitchable PPI modulators using the same Bcl-2 PPI target as proof of concept together with MDM2-p53 and β -arrestin/adaptin interaction. Inhibition constants (K_i) determined by FP, displayed 10-fold increased binding affinity of the *cis* isomer for Bcl-2 binding (*trans* = 1.8 μ M; *cis* = 200 nM; Table 3, entry 6).¹⁰⁶

Enzymatic activity. A fundamental role of proteins is to act as catalysts—enzymes. Enzymatic activity is either governed directly by changes in the enzyme conformation, or by interfering with essential PPIs. Traditionally, enzyme activity has been controlled using small-molecule ligands, which bind to a well-defined pocket. The development of such ligands with photoswitchable properties was an extensive effort.^{12,107–109} However, the paradigm has now changed with the feasibility of turning PPIs into druggable targets.

The fact that peptides can be substrates of proteases may be the reason why those have been the preferred model to explore the light-driven enzymatic control by photoswitchable peptides; precisely, the abundant serine proteases. Interestingly, although these enzymes universally recognize β -strands,¹¹⁰ the reported examples do not take advantage of this feature for their design, neither incorporate photoswitchable β -turn mimetics. In 2000, Harvey and Abell explored structure–activity relationships of azobenzene substitutions at the N-termini of peptidyl α -ketoester inhibitors¹¹¹ to target α -chymotrypsin.¹¹² Evaluated *via* modified Dixon plots, most of the azobenzene 4/5-containing peptides maintained K_i values in the low μ M range. In all cases, the *cis* isomer was the most potent inhibitor (Table 3, entry 7). Particularly, the *meta* substituted dipeptidomimetic was the most active ($K_i = 40$ nM), while the *ortho* one provided the best switch (up to 3-fold difference in K_i).¹¹² This approach was also successful for peptidyl boronate esters.¹¹³ However, difference between isomers was not improved (Table 3, entry 8), and α -chymotrypsin inhibition was in the μ M range. In addition,

observed decomposition of the boronate group upon irradiation may harm further biological applications. Recently, this group and that of Feringa used analogue photoswitchable peptidomimetics to modulate proteasome activity, employing several PSs with different substitution patterns.^{114,115} The group of Feringa¹¹⁵ first demonstrated differences (up to 2.75-fold) in the inhibitory activity (IC_{50} *para*-OMe azobenzene 9: *trans* = 16 nM; *cis* = 46 nM) by fluorescence competition experiments (Table 3, entry 9), and slight specificity between catalytic units.¹¹⁵ The best of the subsequent analogues of Abell displayed activity comparable to bortezomib ($IC_{50} \sim$ nM) and besides, the *cis* isomer was five times less potent than the *trans* one (Table 3, entry 10).¹¹⁴ Interestingly, cell-line selectivity was observed in cytotoxicity assays, which culminated with the almost binary effect in colorectal colon cells (HCT-116) at 310 nM concentration (*cis* cell viability 27%; *trans* cell viability 91%).¹¹⁴ Last year, based on the strong interaction between the serine protease chymase and the 55 mer FynSH3-derived miniprotein—Fynomer—developed by Schlatter *et al.*,¹¹⁶ Woolley proposed a generalizable approach to obtain bioactive light-responsive peptide-based probes.¹¹⁷ Previously, the same group had studied, how irradiation affected the expressed double-cysteine Fynomer mutant crosslinked with BSBCA **11**,¹¹⁸ as well as the *in vivo* stability of BSBCA-modified peptides without biological function.¹¹⁹ They reported that the BSBCA **11**-modified Fynomer has a 10-fold decrease in its affinity to chymase compared to the wild-type protein. Although differences in binding capacities between isomers were observed (IC_{50} (*trans*) = 126 ± 4 nM; IC_{50} (*cis*) = 97 ± 2 nM; Table 3, entry 11), they might not be sufficient for cellular or *in vivo* applications.¹¹⁷

Related to photoresponsive miniprotein-based affinity reagents, since the 1990s there have been attempts to control antibody interactions by light (Fig. 4). In the pioneering work of Harada *et al.*,^{120,121} a monoclonal antibody (ZIH01) reversibly interacted with a hapten tetrapeptide, which carries the photo-responsive *L*-*p*-phenylazophenylalanine amino acid 7, upon *in situ* irradiation. Fluorescence titrations demonstrated binding of the *trans* isomer ($K_d = 5 \times 10^{-7} \text{ M}^{-1}$) while UV irradiation released it (Table 3, entry 12). Similar observations occurred in the recognition of the monoclonal anti-FLAG-tag antibody and the AMPB **1**-modified FLAG epitope;¹²² surface plasmon resonance (SPR) corroborated *in situ* reversible modulation with high isomer ratios (up to 280-fold) for two of the peptides (Table 3, entry 13).

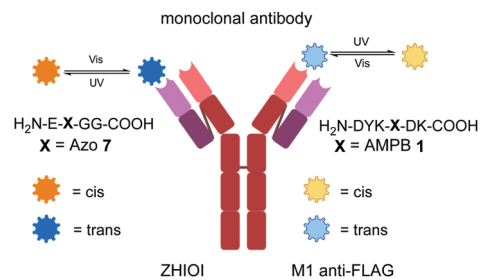


Fig. 4 Photocontrolled binding of antigens to monoclonal antibodies (*trans*: higher affinity). Left: ZIH01 antibody.¹²⁰ Right: M1 anti-FLAG antibody.¹²²

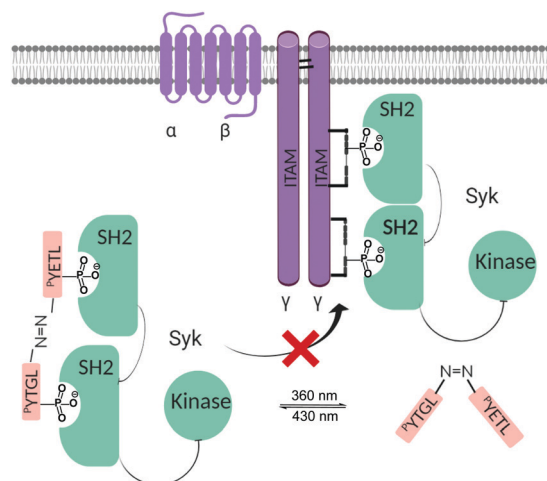


Fig. 5 Schematic representation of photocontrolled inhibition of Syk kinase, using AMPB 1-containing ITAM peptides.¹²⁴

In the context of immunology, the group of Liskamp synthesized bivalent photoswitchable mimetics (Fig. 5) of the immune-receptor tyrosine-based activation motif (ITAM), which bear the AMPB 1 derivative and a phosphotyrosine (^PY) within the backbone.^{123,124} ITAM recruits the spleen tyrosine kinase (Syk) and, in turn, enables its activity.¹³¹ A vital characteristic of the ITAM–Syk interaction is its bivalency; *i.e.*, two SH2 domains in Syk recognize two epitopes in the ITAM peptide. The incorporation of the AMPB 1 derivative as a connector between the two functional ITAMs, determined their distance and alignment to interact with Syk, resulting in activity changes. In the initial project, both isomers of the bivalent ITAM mimetics bound to the tandem SH2 receptor with higher affinities ($K_d \sim$ nM) than the monovalent analogue, according to SPR studies.¹²³ The best compound had a 2-fold difference between isomers, *trans* being the best (Table 3, entry 14). *In situ* irradiation during SPR enabled real-time measurements of the affinities. Reducing the flexibility of the system by using shorter and more constrained linkers, increased isomer difference to 3.6-fold while affinity of the *trans* isomer reached the nM range (Table 3, entry 15).¹²⁴ Conversely, more extended connectors inverted the tendency but with just slight differences (up to 1.3-fold).¹²⁴

In an unique related example, Hoppmann *et al.* achieved light-control of muscle contraction by the synthesis of a cyclic AMPB-containing peptide derived from the α -1-syntrophin binding site of the neuronal nitric oxide synthase (nNOS) (Table 2, entries 7 and 8).⁶⁶ The interaction between α -1-syntrophin and nNOS determines the catalytic activity and induces the synthesis of NO for muscle contraction.^{132,133} In this proof-of-concept project, the *cis* isomer adopted a β -hairpin structure analogous to the structure of the native protein (Fig. 6). Affinity to α -1-syntrophin was comparable to the unmodified control peptide ($K_d = 10.6 \mu\text{M}$ versus $K_d = 5.5 \mu\text{M}$, respectively), according to SPR measurements. Importantly, the *trans* isomer did not bind to α -syntrophin (Table 3, entry 16).⁶⁶ In subsequent work they demonstrated the stability of this photoswitchable peptide against GSH and the possibility of photoregulating muscle fibre

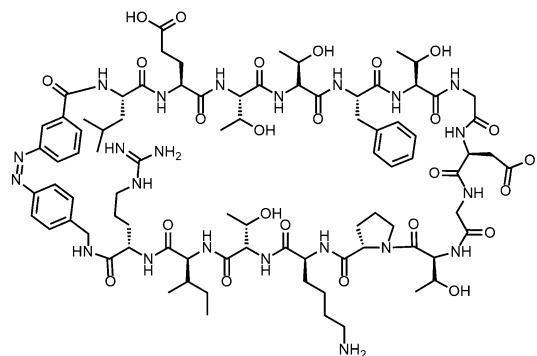


Fig. 6 Structure of the *cis* form of the AMPB-1-modified peptide based on the β -finger of nNOS.^{66,125}

contraction in living cells.¹²⁵ In concordance with the previous work, the *cis* isomer induced the NO released from C2C12 myotubes and, consequently, shortened muscle fibres. Conversely, the *trans* isomer almost did not affect them, as given by the basal contraction percentages (*cis*: 96%; *trans*: 24%). Despite the apparent breakthrough that represented this example, *in vivo* reversibility was not possible.¹²⁵ Last year, the same cyclic peptide was used as an efficient biosensor to detect and control α -syntrophin binding.¹³⁴

PPIs are responsible for the ATPase activity of kinesin, which is a protein biomolecular motor involved in cargo transport by gliding along microtubule tracks upon ATP hydrolysis and many other biological events.¹³⁵ In 2014, the first photoresponsive peptide for kinesin activity modulation was developed¹²⁶ inspired by its self-inhibition mechanism.¹³⁶ If there is no cargo, the C-terminal domain of kinesin bends to block unnecessary ATP consumption. Kumar *et al.* synthesized unsubstituted azobenzene 4-containing peptides derived from this C-terminal tail of kinesin. Importantly, they demonstrated binary reversible photoregulation of the “ON/OFF” motion states upon UV/vis illumination, respectively (Table 3, entry 17).¹²⁶ This outstanding example implies the enormous potential of photoswitchable peptides to both externally and precisely direct cargo transport by means of molecular machines. Furthermore, inhibitory capacities of derivatives were assessed using different peptide sequences, enabling a better insight into the controllable inhibition mechanism.¹²⁷ Several sequences were evaluated to identify the critical sequence required for binding. As a result, an optimized photoresponsive inhibitor of kinesin (X-Arg-Ile-Pro-Lys-Ala-Ile-Arg-OH; X represents the *para*-methoxy-azobenzene motif) was obtained. Upon UV irradiation, the *trans* isomer caused complete inhibition at a lower concentration than in the previous attempts with a recovery of 86% of gliding motility (Table 3, entry 18).¹²⁷

Recently, there has been an increasing interest in modulating epigenetic states by light—optoepigenetics.^{36,128,137–142} Epigenetic regulation involves multiple chromatin-modifying enzymes that control gene expression *via* different mechanisms such as PPI networks. As in proteases, peptides are the preferred substrates of the histone-remodelling enzymes; however, the initial examples of optical control of epigenetic enzymes were based on the incorporation of unsubstituted azobenzenes and azapyrazoles into

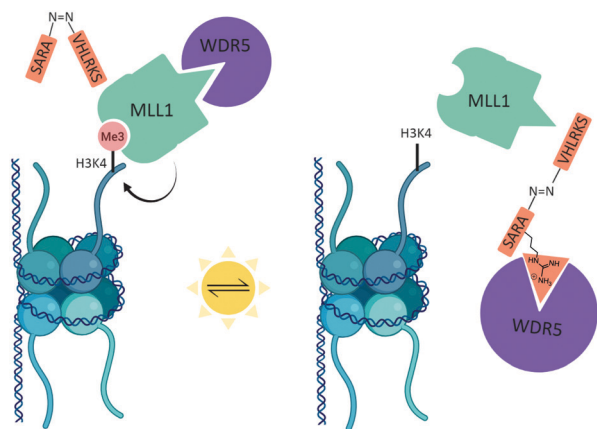


Fig. 7 Schematic representation of the reversible strategy to target MLL1 activity by blocking MLL1–WDR5 interaction.^{128,145}

non-peptidic inhibitors of histone deacetylases (HDAC).^{137,138,143} Our group contributed to this field by developing photoresponsive peptides as optoepigenetic tools to modulate the specific PPI within the methyltransferase mixed lineage leukaemia 1 (MLL1) core complex: MLL1–WDR5, by binding to WDR5 with high affinity ($K_i \sim \text{nM}$) (Fig. 7). A library of AMPB 1-modified peptides based on the previous reported WIN peptide was synthesized.¹⁴⁴ The best candidate displayed one order of magnitude higher affinity than the parental WIN, as well as a 5.2-fold difference between isomers ($K_i(\text{trans}) = 1.25 \text{ nM}$; $K_i(\text{cis}) = 6.50 \text{ nM}$), determined by FP (Table 3, entry 19).¹²⁸ This difference was amplified up to 15-fold when MLL1 inhibition was studied by radioactive histone methyltransferase (HMT) assays ($\text{IC}_{50}(\text{trans}) = 0.531 \mu\text{M}$; $\text{IC}_{50}(\text{cis}) = 8.23 \mu\text{M}$). Cell permeable analogues offered *in situ* photocontrol of leukaemia murine mouse bone marrow (MAF-9) viability with modest but clear differences between isomers ($\text{GI}_{50}(\text{trans}) = 4.98 \mu\text{M}$ versus $\text{GI}_{50}(\text{cis}) = 8.74 \mu\text{M}$). More importantly, the photoswitchable peptide influenced the expression of the MLL1 target gene *Deptor*, depending on the irradiation (*trans*: 0.5%; *cis*: 0.16% relative expression levels at $10 \mu\text{M}$).¹²⁸ Overall, this strategy, demonstrated the potential of photoswitchable AMPB 3-containing peptides as optoepigenetic elements to control gene expression in cellular setups. Recently we turned this approach into a visible-light one, by synthesizing the peptide backbone switches *cAzo* and *oF₄Azo* **15**, based on 5,6-dihydro-dibenzof[*c,g*][1,2]diazocine (*cAzo*)¹⁴⁶ and tetra-*ortho*-fluoroazobenzene (*oF₄Azo*),⁵² respectively. They were grafted onto the WIN peptide and further studied.³⁶ Unfortunately, the improved photochemical properties of our novel building blocks were not translated into an advance of controlling MLL1 activity (Table 3, entry 20). This highlights the challenge and the need for molecular information to address these projects.

Azobenzene moieties have been the preferred PSs to modulate enzymatic activities, but there are also examples using stilbene derivatives. The group of Gogoll incorporated Fmoc-stilbene into peptides to target hydrolases¹²⁹ and reductases.¹³⁰ In the case of the hydrolase, the PS **19** acted as a linker between the two parallel helical segments of the HN1 enzyme, achieving different rate constants between isomers. Thus, the *cis* isomer displayed a rate constant similar to the native HN-1. Upon *cis* → *trans* isomerization

the distance between the two helical motifs increased, leading to an activity loss of 42% in comparison to the *cis* state (Table 3, entry 21).¹²⁹ Recently, to control the activity of *M. tuberculosis* ribonucleotide reductase (RNR), they used the peptidic RNR inhibitor (Ac-Glu-Asp-Asp-Asp-Trp-Asp-Phe-OH),¹⁴⁷ which disrupts the interaction within the two subunits R1 and R2 of RNR by mimicking R2 and, thus, binding to R1. The stilbene **19** was incorporated at different positions into this previously reported R2-based inhibitor to modulate its binding to R1. FP experiments revealed that all stilbene-containing peptides have higher affinity than the parental peptide. The best compound in its *trans* state had a 2.6-fold affinity increase over the *cis* one (Table 3, entry 22).¹³⁰

Protein–nucleic acid interactions

The interaction of proteins with DNA and RNA is intimately connected to all involved aspects in gene expression, including chromatin compaction, transcription, and spatiotemporal localization of mRNA within a cell.¹⁴⁸ Apart from some exceptions,¹⁴⁹ proteins that bind to DNA or RNA are studied independently. However, they share the fundamental physical principles responsible for an effective recognition process.^{150–152} Here, we focus on protein interactions to DNA, since there is a lag in the development of peptide-based tools to photoregulate the RNA binding process. DNA recognition is commonly mediated by the insertion of an α -helix motif of proteins—transcription factors (TF)—into the major groove of the DNA. The existence of several structural DNA-binding domains and, above all, the simplicity of the recognition α -helix fostered the synthesis of artificial TF mimetics.^{152,153}

Photoswitchable bZIP-based peptides. Basic leucine zipper (bZIP) TF are 60-mer helicoidal homo-(hetero)dimers, whose DNA binding domain—basic region (BR)—consists of a positively charged recognition α -helix. However, dimerization by the leucine zipper (LZ) domain is essential for the recognition since both RB peptides and monomeric bZIP display low DNA-binding affinities ($\sim \mu\text{M}$).¹⁵⁴ bZIP dimers interact with DNA in the major groove, folding into chopstick-like structures. This structural simplicity and the crucial role of its helical conformation entails that bZIP has been the primary model for the design of light-responsive peptides to bind dsDNA. In 2000, the first example of photoregulated DNA recognition by a small version of the bZIP member GCN4 was reported (Fig. 8).¹⁵⁵ This pioneering work is based on the large conformational change undergone by an optimized RB homodimer, whose dimerization element is an unsubstituted azobenzene. The *cis* isomer had the appropriate geometry to specifically bind to the target DNA site with ~ 70 -fold higher affinity than the *trans* isomer, by EMSA (Table 4, entry 1). It was demonstrated that the *trans* → *cis* isomerization at 360 nm can be performed in the presence of DNA, although slower and with a slightly lower *cis* conversion than without DNA. Unfortunately, back isomerization in the presence of DNA was hampered due to the high affinity of the *cis* isomer.

In 2002, Woolleys group contributed to the photocontrol of GCN4 by crosslinking the azobenzene **8** to cysteine residues of the GCN4 LZ at positions *i* to *i* + 7.¹⁵⁶ Although the selected positions destabilized the coiled-coil conformation necessary

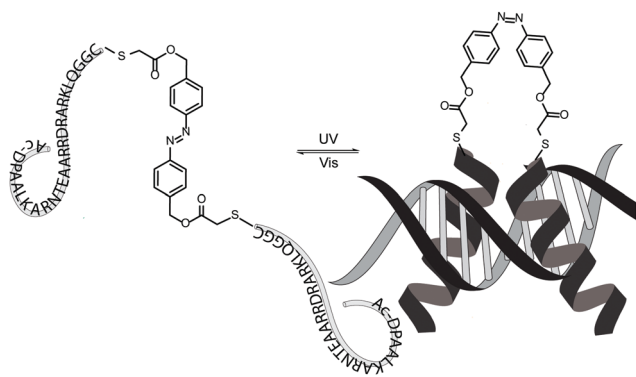


Fig. 8 Scheme of reversible DNA binding of the photoswitchable GCN4-based peptide; azobenzene **4** is the dimerizing unit between two BR.¹⁵⁵

for dimerization, the crosslink improved helicity in the *cis* isomer. This work offered an excellent starting point for further optimizations. The same group worked extensively in this direction, ultimately attaining fully reversible photocontrol of azobenzene-modified GCN4 (Table 4, entry 2).¹⁵⁷ Cysteine residues were crosslinked with the BSBCA **11** at different positions while maintaining their proximity. The *trans* isomer prevented LZ helical conformation and, therefore, GCN4 dimerization. This caused a 20-fold isomer difference in DNA-binding affinity in the kinetic capillary electrophoresis (CE) assays.¹⁵⁷ The group of Allemann used the same approach to control the MyoD TF, whose DNA-binding domain is the basic helix-loop-helix (bHLH).¹⁵⁸ The HLH motif is considered as a modification of the bZIP, in which RB are separated by a turn.¹⁵⁹ According to CD and FP assays, the incorporation of BSBCA **11** did not affect the conformation nor affinity, in the *trans* form. However, upon UV and FP assays, such modification did not affect the conformation nor affinity, in the *trans* form. However, upon UV irradiation the *cis* isomer adopted an α -helical structure with a 60-fold increase in DNA affinity compared to the *trans* one (Table 4, entry 3), which is the highest achieved by α -helix stabilization so far.¹⁵⁸

In 2010, the group of Woolley photocontrolled the activity of the bZIP heterodimer Fos/Jun of the AP-1 TF in living cells.¹⁶⁰ They hijacked the Jun monomer with a photoswitchable BSBCA **11**-modified Fos peptide, since the modified Fos/Jun interaction is stronger than the native one and the homodimeric interaction (Fig. 9). The *cis* peptide displayed 10 times higher potency to inhibit the Fos/Jun dimer than the *trans* one by EMSA (Table 4, entry 4). Importantly, in presence of the *in situ* yielded *cis* isomer, AP-1 activity decreased 40% in 293T cells.¹⁶⁰ Recently, Ali *et al.* reported a systematic study to test the effect of the crosslinking position using the azobenzene **8** over the bZIP CREB system.¹⁶¹ Some general rules could be deduced, *e.g.*, the photocontrol effect is directly correlated with the intrinsic helical propensity of the crosslinked regions. In addition, the affinity of CREB can be tuned by dominant negative versions, in which the N-terminal basic region is replaced by an acidic extension (A-CREB), adjusting the number of acidic heptads of such dominant negative CREB versions. Thus, they demonstrated that the extension of the acidic linker led to an

Table 4 Photoswitchable peptides for protein–DNA–interactions. X = PS; n.p. = not provided in publication

Entry	Sequence/PS	PSS ratio	Target	Δ affinity	Technique
1 ¹⁵⁵	Ac-DPAALKARNTEAARRDRARKLQGGC Ac-DPAALKARNTEAARRDRARKLQGGC X	<i>cis</i> : 95% <i>trans</i> : 99%	GCN4 target DNA (CREB20)	<i>trans/cis</i> : 70.0	Affinity (EMSA)
2 ¹⁵⁷	pGCN4-56 (S262C, N269C) protein ^a X = BSBCA 11	<i>cis</i> : 75% ($\tau_{1/2}$ = 50 min in aq. sol., 25 °C) <i>trans</i> : 99%	AP1 target DNA	<i>trans/cis</i> : 20.0	Affinity (capillary electroph.)
3 ¹⁵⁸	Ac-ADRRKAATCRRRRRLCKV-CONH ₂ X = BSBCA 11	<i>cis</i> : 61% ($\tau_{1/2}$ = 193 min in aq. sol., 15 °C) <i>trans</i> : 99%	MyoD target DNA (E-box)	<i>trans/cis</i> : 60.0	Affinity (FP)
4 ¹⁶⁰	cFos sequence (E38C, K45C) ^b X = BSBCA 11	<i>cis</i> : 87% ($\tau_{1/2}$ = 2 h in aq. sol., 20 °C) <i>trans</i> : 99%	AP-1 TRE DNA	<i>trans/cis</i> : 10.0	Inhibition (EMSA)
5 ¹⁶¹	A4-CREB-x4 sequence ^c X = Azo 8	<i>cis</i> : 82% ($\tau_{1/2}$ = 59 min in aq. sol.) <i>trans</i> : 99%	CREB target DNA	<i>trans/cis</i> : 5.40	Inhibition (EMSA)
6 ¹⁶⁵	H ₂ N-EAQCQKIAAKNARAKCKKA-CONH ₂ X = BSBCA 11	<i>cis</i> : 75% ($\tau_{1/2}$ = 150 min in aq. sol., 25 °C) <i>trans</i> : 99%	QRE target DNA	<i>cis/trans</i> : 22.7	Affinity (FP)
7 ¹⁶⁶	Ac-OTrn-EAAEA=O-m-QAEAKIAAKNARAKKA-CONH ₂ X = DAE 28	Closed: 88% Open: 100% n.p.	HF target DNA	Closed/open: 4.00	Affinity (QCM)
8 ¹⁶⁷	59 mer Sp1 TF ^d X = 4	n.p.	Sp1 target DNA	<i>trans</i> : K _d : 0.73 μ M <i>cis</i> : too low	Affinity (EMSA)
9 ¹⁶⁸	Ac-KKFAPECPKRFM-X-RSDHLKSHIKTHQNKK-CONH ₂ X = AMPB 1	<i>cis</i> : 56% ($\tau_{1/2}$ = 14 h in aq. sol., 25 °C) <i>trans</i> : 99%	Sp1 target DNA	<i>cis/trans</i> : 2.00	Affinity (fluorescence displacement)

^a GCN4-56-sequence (bold: crosslinking site); H₂N-MKDPAAALKARNTEAARRDRARKLQGGC-CONH₂. ^b Fos sequence: H₂N-NNELEKEAELEQLAEIQAEIQLECRNYALRCEEDLQKQLEKLGAP-CONH₂. ^c A4-CREB-x4 sequence: H₂N-MLEQQLELAQENELEKEAELEQLAEIQAEIQLECRNYALRCEEDLQKQLEKLGAP-CONH₂. ^d Sp1 TF: X-RPFMCTWSYCGKRFTRSDELQHRKTRHTGKGFACPEKRFMRSDHLSKSHIKTHQNKK-CONH₂.

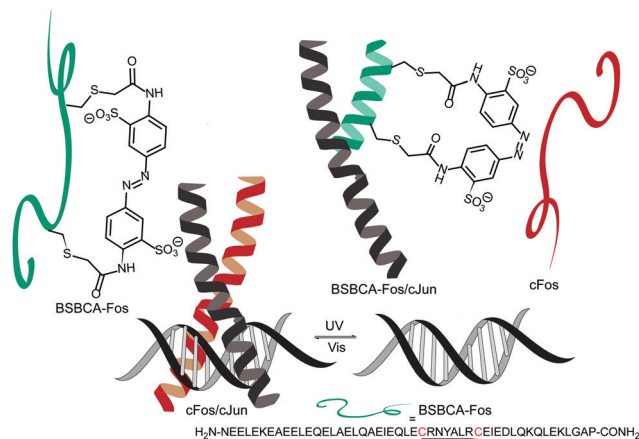


Fig. 9 Schematic representation of the photocontrol of AP-1 activity by hijacking Jun with a BSBCA **11**-modified Fos-peptide.¹⁶⁰

increase in the difference of DNA-binding affinity between isomers up to 5-fold (Table 4, entry 5).¹⁶¹ Finally, it is worth briefly mentioning that the group of Woolley achieved photocontrol of both CREB and GCN4 *via* an optogenetic approach based on the photoactive yellow proteins (PYP).^{162,163}

Photoswitchable HTH-based peptides. The helix-turn-helix (HTH) motif was the first DNA-binding domain discovered.¹⁶⁴ HTH is present in a bigger version in the eukaryotic TF called homeodomain (HD). This consists of 60 residues of three helices, where the N-terminus contacts with the minor groove of DNA and two helices lie parallel to each other, across the almost perpendicular ($\sim 120^\circ$) recognition α -helix—helix 3. In 2005, the group of Allemann successfully extrapolated Woolley's BSBCA **11**-crosslinking methodology for the HD Q50K mutant.¹⁶⁵ Thus, the whole Q50K TF was reduced into a photoswitchable 18-mer BSBCA **11**-crosslinked peptide from helix 3. The *trans* isomer maintained the DNA recognition properties of the wild-type TF with a 23-fold difference in isomer affinity (K_d (*trans*) = 7.5 nM; K_d (*cis*) = 140 nM) by FP-assays (Table 4, entry 6). Furthermore, the *trans* isomer specifically bound the QRE target DNA, while both the *cis* isomer and the non-crosslinked control, could not discriminate between DNA sequences.¹⁶⁵ In contrast to previous studies, the crosslinking was at the *i* to *i* + 11 positions, justifying the reverse tendency compared with the bZIP projects.

Azobenzenes are the most common PSs used to regulate the helicity of DNA-binding peptides. However, in 2012, Fujimoto *et al.* realised the potential of DAE-derivatives in biological systems due to isomer stability.¹⁶⁶ They systematically studied a DAE **28**-bridged peptide library obtained *via* a crosslinking reaction between ornithine residues and a succinimidyl DAE ester. Crosslinking sites in positions *i* to *i* + 6 led to maximized structural control compared to other substitution patterns, wherefore two different C-terminal DNA-binding sequences, from binding regions of homeodomains (QK) and HFN-3 γ (HF), were coupled to this peptide. UV irradiation triggered the photocyclization of DAE **28**, which resulted in lower helicity than in the open form, according to CD spectroscopy.

This change in the secondary structure of the DAE-based peptides was consistent with the 4-fold affinity difference between isomers for the QK (K_d (open) = 1.0 μ M; K_d (closed) = 4.0 μ M) by quartz crystal microbalance (QCM) (Table 4, entry 7). Additionally, irradiation during QCM allowed *in situ* DNA-binding regulation.¹⁶⁶ These examples demonstrate higher differences between isomers of azobenzene crosslinks compared to DAE.

Photoswitchable ZF-based peptides. Zinc-finger proteins (ZFs) are the most abundant TFs.¹⁶⁹ Their DNA-binding domain consists of two antiparallel β sheets and an α -helix, which is inserted in the major groove of the DNA. This folding is stabilized by tetrahedral coordination between a zinc ion and different combinations of cysteine and histidine residues. Most ZFs interact with the DNA as highly adaptable oligomers, which has been used to develop custom-designed ZFs¹⁷⁰ and more recently, ZF nucleases.¹⁷¹

Regarding photoswitchable ZF-based peptides, two different approaches were used, both relying on classical azobenzenes: **4** or AMPB **1**.^{167,168} On one hand, Nomura and Okamoto synthesized a 59 mer bivalent ZF peptide derived from the Sp1 TF with the azobenzene **4** modification in its N-terminus (Fig. 10).¹⁶⁷ The hypothesis was that the planar *trans* isomer should be more prone to stack between the DNA bases than the non-planar *cis* analogue, and consequently, affect DNA binding. According to CD experiments, the efficient *trans* \rightarrow *cis* isomerization did not affect the α -helix conformation while upon 420 nm irradiation the *trans* back-isomerization proceeded at the expense of protein denaturation. They explained this as being due to a possible interaction between the *cis*-azobenzene motif and the folded bivalent ZF. The obtained dissociation constants (K_d) from EMSA were comparable for both the *trans* isomer and the wild type (K_d (*trans*) = 0.73 μ M versus K_d (wild type) = 0.32 μ M). Unfortunately, the K_d value of the *cis* isomer could not be determined due to low-intensity bands (Table 4, entry 8). Furthermore, it was demonstrated that the designed photoswitchable peptide could control the binding of restriction enzymes to GC box sites, leading to photocontrollable DNA cleavage. While the *trans* isomer inhibited cleavage, UV irradiation enabled it, suggesting dissociation of the *cis* isomer from the GC box region.¹⁶⁷ On the

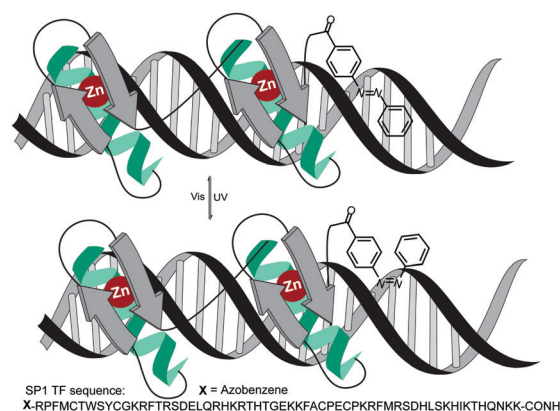


Fig. 10 Schematic representation of the photoswitchable Sp1 ZF TF; *trans* isomer binds to target DNA with higher affinity than the *cis* one.¹⁶⁷

other hand, Szymanski and Feringa developed a ZF-based system, where AMPB **1** was introduced into the turn region of one ZF of the Sp1 TF. Such modification aimed at affecting Zn²⁺ coordination and consequently, the ZF folding and recognition.¹⁶⁸ CD titrations confirmed that there was stronger Zn²⁺ coordination in the *cis* state than in the *trans* one (K_d (*trans*) = 81 μM; K_d (*cis*) = 53 μM). In addition, this ZF mimetic was compared with a photo-switchable version of the minor-groove binding domain AT-hook, containing a conserved Arg-Gly-Arg (RGR) sequence flanked by basic residues.¹⁷² In this design, the AMPB **1** unit bridged two RGR motifs. Surprisingly, fluorescence displacement assays demonstrated that for the ZF peptides the *trans* isomer had a 2-fold increased binding affinity compared to its corresponding *cis* isomer (IC_{50} (*trans*) = 0.72; IC_{50} (*cis*) = 0.36),[†] although Zn²⁺ coordination was favoured in the *cis* state (Table 4, entry 9). This might be due to optimal positioning of the α -helical binding motif of the ZF in its *trans* state. For the AT-hook, the expected inversed affinities were displayed (IC_{50} (*trans*) = 0.034; IC_{50} (*cis*) = 0.058),[†] where the *cis* isomer is the stronger DNA binder, due to its probable ability to interact through two hooks with the same DNA strand. This interaction recalls the first light-controllable bZIP example by Mascareñas.¹⁵⁵ Unfortunately, isomer differences were lower than in the previous examples of bZIP and HTH peptides.

Transmembrane proteins

In response to external stimuli, transmembrane proteins can (de)activate signalling cascades regulating intracellular communication and functions. Although the past two decades have witnessed spectacular progress in both atomic and functional characterization of these proteins, much has still to be done to completely understand them and their connections with diseases. Transmembrane proteins are the most popular family of targets for drug discovery.¹⁷³ Targeting these proteins with optical tools is a success story because light provides the fast time resolution needed to control these events dynamically. Furthermore, in humans, none of the transmembrane receptors naturally respond to light, except for rhodopsin. Therefore, both optogenetics and photoswitchable ligands enabled a fascinating collection of light-responsive proteins capable of assisting in structural elucidation and functional regulation.^{1,112,174,175} To date, photoswitchable small-molecule agonists tethered to endogenous receptors dominate this field.^{109,176,177} However, integrins and G-protein coupled receptors were successfully addressed by photoswitchable peptides (Table 5). Although beyond the scope of this review, azobenzene^{178–180} and HTI⁴⁶ moieties have been utilized to photocontrol the activity of gramicidin A ion channels. Additionally, spiropyran-based peptides caused light-stimulated lysis of artificial membranes based on bee venom miltettin¹⁸¹ and served as sensors.¹⁸²

Integrins. Integrins are the most prominent cell adhesion receptors. This heterodimeric complex consists of two

glycoproteins with both a transmembrane and extracellular domain. The sequence Arg-Gly-Asp (RGD) has been identified as the main integrin-binding motif. Short RGD-containing peptides reproduced integrin binding activity of adhesion proteins.¹⁸³ Regarding photocontrol, two different approaches, both using the azobenzenes **1** and **4**, were implemented: peptide conformation-dependency^{184,185} and photoswitchable peptide surface-coating systems.^{186,187} In 2003, the group of Moroder, incorporated the AMPB **1** into a cyclic heptapeptide adapted from the well-known $\alpha_v\beta_3$ integrin ligand. Surface plasmon-enhanced fluorescence spectroscopy (SPFS) demonstrated 40% higher binding affinity of the *trans* isomer than the *cis* one (Table 5, entry 1).¹⁸⁴ This was rationalized by NMR studies. Seeking maximal conformational restriction, a cyclic analogue with the AMPB **1** in different positions was prepared.¹⁸⁵ Although, shifting the AMPB **1** position led to a drastic decrease in affinity (up to 50%), isomer preference was conserved and isomer difference improved (2-fold *versus* 1.6-fold; Table 5, entry 2).¹⁸⁵ Regarding the second approach, poly-methylmethacrylate surfaces were linked with AMPB **1**-RGD-peptides, which displayed 2-fold lower cell adhesion after irradiation at 366 nm (*cis* isomer) in comparison to the *trans* (Table 5, entry 3).¹⁸⁷ Embedding RGD into a polyethylene glycol matrix onto a gold surface in a self-assembled monolayer (SAM), using the azobenzene **4** clearly improved this photoswitchable system. Complete reversibility of cell adhesion was achieved under irradiation (*trans*: 200 cells/field; *cis*: 10 cells per field) when murine NIH 3t3 fibroblast were cultured onto these surfaces (Table 5, entry 4).¹⁸⁶

G-protein coupled receptors (GPCRs). GPCRs are the most prevailing signalling proteins (>30% of encoded genome), representing most of the targets for approved drugs.¹⁹¹ General GPCR architecture consists of a polypeptide with an intracellular C-terminus followed by seven transmembrane α -helical domains and an extracellular N-terminus.^{192,193}

Among GPCRs, the endogenous glucagon-like peptide 1 receptor (GLP-1R) was the first to which, photosensitive peptides were shown to exert direct optical control.¹⁸⁸ GLP-1R is a well-established therapeutic target for the treatment of type-2 diabetes mellitus.¹⁹⁴ Upon incretin binding, GLP-1R is activated and, thereby, insulin is secreted. The need to precisely regulate insulin levels has fuelled impressive efforts towards its light-controlled release.^{188,195,196} Thus, taking as a starting point the peptidic GLP-1R agonist: liraglutide, Broichhagen *et al.* synthesized a photoswitchable version, incorporating the AMPB **2** switch as a bridge between the two α -helices.¹⁸⁸ The *trans* isomer retained the parental α -helix while the *cis* one unfolded the structure. Incubation of this compound in pancreatic β cells offered the possibility to switch between two distinctive pathways. Hence, the *trans* form engaged mainly with the calcium-ion flux, whereas the *cis* one favoured cAMP generation. A difference in insulin secretion was observed in pancreatic β -cells of rodent mouse islets. The *cis* isomer had similar secretion to the wildtype, while the *trans* one showed a 2.5-fold decrease (secretion (wild type) = 0.3 ng mL⁻¹ islet⁻¹ h⁻¹; (*cis*) = 0.4 ng mL⁻¹ islet⁻¹ h⁻¹; (*trans*) = 0.15 ng mL⁻¹ islet⁻¹ h⁻¹; Table 5, entry 5).¹⁸⁸

[†] IC_{50} : determined from ratio of the concentration of ethidium bromide/DAPI to concentration of peptide, needed to displace half of the intercalators from the DNA.

Table 5 Photoswitchable peptides for transmembrane proteins. X = PS; n.p. = not provided in publication

Entry	Sequence/PS	PSS ratio	Target	Δ affinity	Technique
1 ¹⁸⁴	c[KARGD ^D FV-X] X = AMPB 1	cis: 80% trans: 99%	$\alpha_v\beta_3$ -integrin	cis/trans: 1.60	Affinity (SPFS)
2 ¹⁸⁵	c[ARGD ^D FVK-X] X = AMPB 1	cis: 80% trans: 99%	$\alpha_v\beta_3$ -integrin	cis/trans: 2.00	Affinity (SPFS)
3 ¹⁸⁷	X-c[RGDFK] X = AMPB 1	n.p.	Integrin cell adhesion	trans/cis: 2.00	Cell adhesion assay
4 ¹⁸⁶	X-GRGDS-CONH ₂ X = Azo 4	n.p.	Integrin cell adhesion	trans/cis: 2.00	Cell adhesion assay
5 ¹⁸⁸	Liraglutide SQ ^a X = AMPP 2	n.p. cis: ($\tau_{1/2} \sim 5$ min)	GLP-1R	cis/trans: 2.50	Insulin secretion assay
6 ¹⁸⁹	H ₂ N-YRFSNC-X-QAGIRDMRGGFCSSRRLLA-COOH X = AMPP 2	n.p. cis: 79% ($\tau_{1/2} \sim 5$ min, 20 °C) trans: 96% ($\tau_{1/2} = 6.8$ d in aq. sol.)	NPR-A	cis/trans: 4.00	cGMP assay in HEK293t cells
7 ⁴³	H ₂ N-(YRLRY) ₂ -X X = Azo 6; X = azopyrazole 16; X = DAE 27; X = fulgimide 24	16: cis: 88% ($\tau_{1/2} = 8.1$ d in aq. sol.) trans: 94% 27: closed: 95% open: 98% 24: closed: 49% open: 85%	Y ₄ R	6: trans/cis: 2.20 16: trans/cis: 1.50 27: closed/open: 1.88 24: closed/open: 1.88	Affinity (radioligand binding assay)
8 ¹⁹⁰	H ₂ N-DDDDIVFECEARQLCGMDD-COOH X = BSBCA 11	n.p.	β -Adaptin	trans/cis: 12.6	Affinity (FP)

^a Liraglutide: H₂N-N-HAEGTFTSDVSSYLE-X-AAKEFIWLIV(RG)₂-COOH.

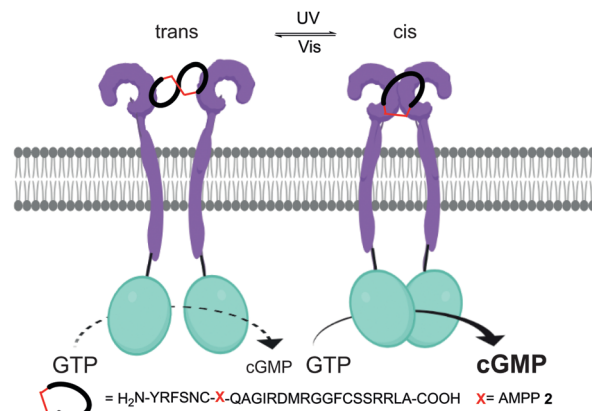


Fig. 11 Scheme of the modulation of the binding of NPR-A to AMPP 2-(displayed in red) peptide, which regulates guanyl cyclase activity.¹⁸⁹

These design principles were applicable to another peptide hormone: atrial natriuretic peptide (ANP), and its associated receptor: natriuretic peptide receptor A (NPR-A), which is not strictly a GPCR but linked to intracellular guanyl cyclase. As an agonist of NPR-A, ANP raises intracellular cGMP concentration and its misregulation is correlated to cardiovascular diseases, hypertension and heart failure.^{197,198} Recently, Trauner and Hoffmann-Röder reported a cyclic AMPP 2 peptide, which enabled the optical control of ANP-based cGMP production (Fig. 11).¹⁸⁹ Based on structure-activity studies, a photoswitchable peptide library was developed. The AMPP 2 moiety was systematically incorporated substituting two, three or four amino acids of the ANP backbone. The best compound, in which the AMPP 2 moiety replaced four amino acids of ANP C-terminus, exhibited 4-fold higher potency in cGMP production in its *cis* state than in *trans* one in HEK293T (Table 5, entry 6). Moreover, such differences enabled light-dependent muscle contraction and fluctuations in Ca²⁺ fluxes.¹⁸⁹

This year, the group of König turned the potent dimeric neuropeptide Y (NYP) into photoresponsive analogues by exchanging the inert aliphatic linker by four different backbone PSSs: azobenzene 6, azopyrazole 16, DAE 27 and fulgimide 24.⁴³ Decreasing the linker flexibility had a positive impact on the affinity to the receptor Y₄R, although with modest isomer differences. Competitive radioligand-binding assays displayed *K_i* values in the low nM range for all analogues. Azobenzene 6, which displayed slightly higher isomer differences (up to 2.2-fold) than the others, was the best candidate (Table 5, entry 7). Importantly, functional assays demonstrated that all peptides act as Y₄R partial agonists by measuring G-protein mediated signalling as well as by coupling of the receptor to arrestins.⁴³

Activated GPCRs can regulate the dynamics of their own endocytosis *via* different pathways.¹⁹⁹ One of the more extensively characterized routes is clathrin-mediated endocytosis (CME).²⁰⁰ The binding of clathrin-associated proteins depends on the direct interaction with the AP2 complex. Once bound, endocytosis is initiated by the formation of specialized membrane regions—clathrin-coated pits—into which, cell-surface receptors concentrate forming clusters. Afterwards, through highly regulated

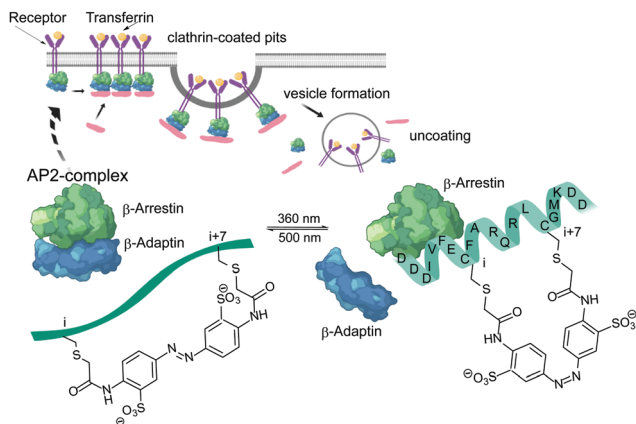


Fig. 12 Scheme of photocontrolled CME, using BSBCA **11** crosslinked β -arrestin C-terminal peptide.²⁰¹

steps, where the protein β -arrestin plays a major role, internalization takes place. In 2013, Giralt and Gorostiza reported an elegant example, demonstrating reversible modulation of CME PPI *via* BSBCA **11** stapled α -helical peptides (Fig. 12).¹⁹⁰ They studied two cross-linked variants (*i, i + 7* versus *i, i + 11*) of the helicoidal β -arrestin C-terminus. The wild-type peptide interacted with the β -appendage of AP2 (β -adaptin) with a $K_d = 2.1 \mu\text{M}$. The difference in the biological response was dramatically amplified. An *i* to *i + 7* spacing was demonstrated to be best with affinities in the μM range according to FP, and a 13-fold difference between isomers (K_d (*trans*) = 240 μM ; K_d (*cis*) = 19 μM ; Table 5, entry 8). Transferrin uptake and clathrin dynamic studies, validated that both peptides inhibited CME activity in a light-dependent fashion in living HeLa cells. Importantly, transferrin uptake of the *cis* BSBCA **11**-containing peptide was quantitative while it was reduced to 40% in its *trans* state.¹⁹⁰ Due to the interest of elucidating general features for the rationalized design of switchable PPI modulators, in a related project, they synthesized different 20-mer peptides crosslinked with BSBCA **11**, and studied their ability to bind β -adaptin.²⁰¹ This systematic analysis suggested that the stabilization of the α -helix is not crucial, demonstrated by the negative impact of peptide-rigidifying modifications at the positions *i, i + 4*, over the ones at *i, i + 7* and *i, i + 11*. Indeed, in this model the most flexible peptides were the better inhibitors. This gives importance to peptides lacking a stabilized secondary structure as potential modulators.²⁰¹

Cell uptake

To trigger biological functions, one main obstacle to overcome is the precise delivery of the active compound (cargo) to its target. Cell-penetrating peptides (CPPs), which are mostly polycationic and/or amphiphilic peptides,^{202–205} assist in cargo uptake. Depending on the CPP type, different cellular uptake mechanisms have been proposed although, they are still under scientific discussion.²⁰² In general, CPPs are hardly cell-type specific. Therefore, activatable CPPs would enable both deciphering of the uptake mechanisms and spatiotemporal control over the permeabilization, which would increase therapeutic CPP applicability.

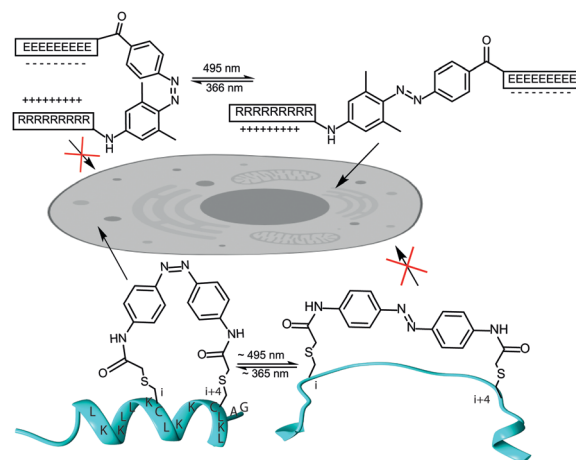


Fig. 13 Strategies to photocontrol CPP cellular uptake. (A) Deactivation of oligoarginine by neutralizing its positive charges with oligoglutamate. (B) Control of cellular uptake using stapled α -helices (*cis* isomer).^{206,207}

Two studies, using different approaches (Fig. 13),^{206,207} reported photocontrollable CPPs based on the classical azobenzene. In the first example, oligoarginine-based peptides with a photoswitchable uptake profile *via* electrostatic impairment were developed.²⁰⁶ The 9-mer oligo-arginine sequence was linked by an *ortho*-2,6-methylated azobenzene to an oligoglutamate analogue. In its *cis* form, the azobenzene adopted a turn-like structure enabling efficient neutralization, while the unmasked *trans* form, was capable of penetrating HeLa cells. Cell-uptake quantification by flow cytometry (FC) demonstrated binary difference between isomers and quantitative uptake for the *trans* isomer. Results were corroborated by confocal microscopy. The second example is based on the improvement of cellular uptake using stapled α -helical peptides.²⁰⁸ This methodology was implemented into the amphipathic KL peptide, using the crosslinked azobenzene **8** in *i* to *i + 4* positions.²⁰⁷ As expected, CD spectroscopy confirmed a stabilized α -helix of the *cis* isomer. Using a TAMRA-labelled analogue, fluorescence-activated cell sorting (FACS) proved just few uptake in HeLa cells for the unstapled peptide (15%) and the *trans* isomer (27%). The *cis* isomer led to a remarkable uptake increased to 76%.²⁰⁷

β -Amyloid aggregation

β -Amyloids (βA) are approximately 40-mer peptides derived from the membrane protein APP (amyloid precursor protein) by enzymatic proteolysis.²⁰⁹ βA s form fibrils, which are responsible for the pathogenesis of several diseases such as Alzheimer,^{210,211} Parkinson,²¹⁰ type 2 diabetes²¹² and amyloidosis.²¹¹ Although the precise mechanism of amyloid fibril formation remains unknown, the self-assembly process is considered to be dependent on the β -sheet conformation, stabilized by extensive intermolecular hydrogen-bond contacts— β -interactions. The β -cross morphology of the fibrils features a sheet-turn-sheet motif. Furthermore, the cytotoxicity was attributed to the early events of the amyloidogenesis.²⁰⁹ Investigation of these processes has profited from structural peptidic models, in which the conformational transition dynamics are controllable. Thus, photoresponsive

Table 6 Photoswitchable peptides for β -amyloid aggregation. PSS ratios, techniques. "n.p." = not provided; "X" = PS

Entry	Sequence/PS	PSS ratio	Target	Δ -isomers	Technique
1 ^{2,14}	Ac-SWTW-X-KWTWK-CONH ₂ X = AMPP 2	n.p.	β -Amyloid aggregation	<i>trans</i> : amyloid fiber; <i>cis</i> : disassembly	IR, TEM
2 ⁴¹	Ac-SWTW-X-KWTWK-CONH ₂	<i>cis</i> : 93% ($\tau_{1/2}$ = 4.1 d, 25 °C) <i>trans</i> : 99% n.p.	β -Amyloid aggregation	<i>trans</i> : sheet structures; <i>cis</i> : disassembly	cryoTEM
3 ²¹⁷	X = overcrowded alkene 23 Ac-CHGNCK-CONH ₂	n.p.	β -Amyloid aggregation	<i>trans/cis</i> : 2:00 (fluorescent signal)	TRLS, TEM, LCP-fluorescence
4 ²¹⁶	X = AzO 8 β A42-peptide ^a	<i>cis</i> : 65% <i>trans</i> : 99% n.p.	β -Amyloid aggregation; C17.2 neural cells β -Amyloid aggregation; SH-Sy5Y cells	<i>trans</i> : amyloid fiber (toxic); <i>cis</i> : non-amyloid aggregates cytotoxicity; <i>trans/cis</i> : 10:0 <i>trans</i> : amyloid fiber (non-toxic); <i>cis</i> : amyloid-oligomers (toxic) cytotoxicity: <i>cis/trans</i> : 4:11	TEM, LCP-fluorescence, cytotoxicity CD, TEM, cytotoxicity assay
5 ²¹⁸	X = AMPP 2 β A42-peptide ^b X = 3,4-AMPP				

^a β A42-peptide: H₂N-DAEFRHDSGYEVHHQKLVFFAEDV-X-KGAIIGLMVGGVVIA-CONH₂, ^b β A42-peptide: H₂N-DAEFRHDSGYEVHHQKLV-EDVGSNKGAIIGLMVGGVVIA-CONH₂.

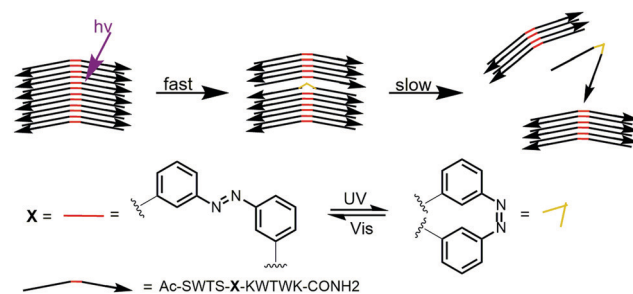


Fig. 14 Model of β -amyloid aggregation of AMPP 2 containing TrpZip. UV irradiation generates the *cis* PSS in picoseconds, while the disassembly of amyloid-like fibrils takes up to several minutes.

β -sheet mimetics have the potential to shed light into the process, responsible for cytotoxicity (Table 6).

The bases of the pioneering work on photocontrol of β -sheet aggregation comes from Moroder's discovery of stable TrpZip, in which the AMPP 2 moiety was incorporated in the β -turn region (Table 2, entry 4).^{62,213} Later on, IR spectroscopy and transmission electron microscopy (TEM) of the same, but acetylated peptide, confirmed amyloid-like fibre aggregation of the *trans* peptide, which was disassembled upon irradiation at 360 nm, leading to the reestablishment of β -hairpin structures (Table 6, entry 1).²¹⁴ Time-resolved IR spectroscopy demonstrated that the *cis*-induced disassembly occurs *via* thermally driven steps on much longer time scales than the actual isomerization of *trans* \rightarrow *cis*, which happens on a timescale of ns (Fig. 14).²¹⁵ In addition, this TrpZip motif has been the first system, in which an overcrowded alkene 23 was incorporated at the same position and used for chemical biology applications.⁴¹ They have the advantage of exerting two thermally stable isomers and higher geometrical change than azobenzenes. Analogue to the AMPP 2, the *cis* isomer engaged β -hairpin structures in contrast to the *trans*. Thus, cryogenic transmission electron microscopy (cryoTEM) confirmed aggregation of the *trans* isomer, and disassembly upon irradiation to the *cis* isomer (Table 6, entry 2). However, different structures were detected depending on the PSSs. Thus, while the *trans* isomer of the AMPP 2-modified peptide formed amyloid-like fibres, the over-crowded alkene analogue formed sheet-like structures, which were proposed to be a result of the high hydrophobic core.⁴¹

In 2012, the group of Pellarin performed MD simulations to identify, amongst hundreds of sequences, those peptides cross-linked to azobenzene 8, with the highest propensity to folding pathways (α helix \rightarrow β -sheet oligomer \rightarrow β -sheet fibrils) leading to cross β -fibrils and cytotoxic prefibrillar intermediates.²¹⁶ To overcome the restriction of β -turn peptides being able to enter such pathways, and the fact that both isomers form fibrils, slight changes in the photoswitch and in the incorporation position (3,4-AMPP instead of 3,3-AMPP 2 form β -sheet oligomers in their *trans* PSS, and the highest solubility in the corresponding *cis* PSS.²¹⁷ The most promising was the crosslinked hexapeptide Ac-Cys-His-Gly-Asn-Cys-Lys-CONH₂. Reversible aggregation was confirmed using time resolved light scattering (TRLS), TEM and luminescent conjugated polyelectrolyte (LCP)-based fluorescence

for amyloid fibrils.²¹⁷ In particular, upon UV irradiation fluorescence intensity decreased 2-fold due to disassembly of amyloid fibers *via cis*-triggered solubilisation (Table 6, entry 3). This approach was extended to the native β A42, where an AMPP 2-modified peptide derived from the 42 mer putative turn-region of β A was synthesized as a tool to elicit the role of turn nucleation in β A and evaluate its cytotoxicity in neural progenitor cells.²¹⁶ Surprisingly, TEM and LCP-based fluorescence revealed that the *trans* isomer behaved as in the former reports, *i.e.*, it readily assembled into fibrils, which were identical to the wild-type one. On the contrary, the *cis* isomer formed only amorphous, non-amyloid aggregates (Table 6, entry 4). This was unexpected, since it was hypothesized that the stabilized β -hairpin structure of the *cis* isomer would favour self-assembly, while the *trans* conformation would lie in an extended conformation, precluding self-assembly. Just the *trans*-induced oligomers displayed cytotoxicity ($\sim 50\%$), which was comparable to the wild type and consistent with the structural characterization. Of note, *cis*-induced ones were hardly toxic ($\sim 5\%$ cytotoxicity). This demonstrates that stable β -turn peptides are unable to enter folding pathways (α helix \rightarrow β -sheet oligomer \rightarrow β -sheet fibrils) leading to cross β -fibrils and cytotoxic prefibrillar intermediates.²¹⁶ To overcome the restriction of β -turn peptides being able to enter such pathway, and the fact that both isomers form fibrils, slight changes in the photoswitch and in the incorporation position (3,4-AMPP instead of 3,3-AMPP 2 substituted for residues 18–21 instead of 25–27) led to distinct observations.²¹⁸ Although, once again, it proved that the *trans* isomer is forming amyloid-like fibrils, those fibrils were, as expected, non-toxic in human neuroblastoma SH-Sy5Y cells (9% cytotoxicity). Isomerization to the *cis* isomer, led to disassembly of fibrils into toxic amyloid-like oligomers clustering into typical β A immature fibrils with increased (37%) cytotoxicity, similar to the wild-type oligomers (51% cytotoxicity; Table 6, entry 5).²¹⁸

Antibiotic peptides

Misuse and overuse of antibiotics have turned bacterial resistance into one of the major current threats to public health.²¹⁹ On-demand activatable versions of the current antimicrobials may result in more effective and precise treatments. In particular, photoswitchable antibiotics hold the promise to offer unprecedented spatiotemporal resolution, which may also contribute to decreasing environmental accumulation. An implication of this is the emergence of antibiotics as novel targets in photopharmacology.^{12,29,109} Regarding active peptides, one strategy to overcome bacterial resistance is to use naturally occurring antibiotic peptides. Optimizing their structures to gain

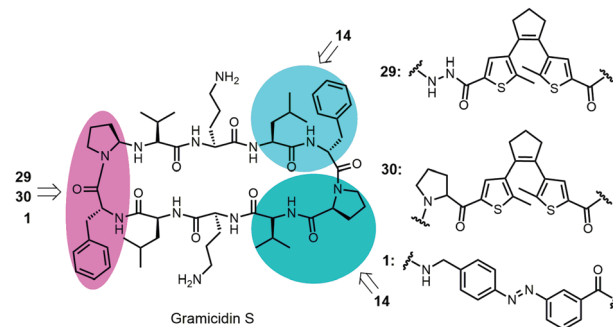


Fig. 15 Native cyclic gramicidin S antibiotic peptide. Substitution sites for the different PSs are displayed in different colours.^{223–225}

control over activity by light, represents a relatively unexplored approach, mainly relying on gramicidin S derivatives (Table 7).

Gramicidin S is an amphipathic cyclic decapeptide, whose sequence [(Val-Orn-Leu-DPhe-Pro)₂] forms an antiparallel β -sheet flanked by type II' β -turns stabilized by intramolecular hydrogen-bonds, enhancing its amphipathic properties and providing membranolytic activity against both, broad bacteria classes as well as drug-resistant strains and biofilms.^{220,221} Importantly, perturbations of this 3D scaffold causes a decrease, or complete, loss of function.²²² The critical role of the intramolecular hydrogen bonds for this stabilization led Komarov and Ulrich to envision DAE-modified gramicidin S analogues, in which the PS could act as a light-controllable β -turn mimetic (Fig. 15).²²³ In their proof-of-concept work, the replacement of non-polar residues by DAE 29 yielded three photoswitchable peptides, in which their open form was always as active as the parental antibiotic against *S. aureus*, *S. epidermidis* and *S. xyloso* cells. The photoswitchable peptide with the highest difference between isomers (16-fold for *S. aureus* and *S. xyloso* bacteria) was that, where DAE replaced the Leu and D-Phe residues (Table 7, entry 1).²²³ As demonstrated in photodynamic therapy, optical modulation of cytotoxicity has an important impact in selective cancer treatments. Therefore, the same group optimized the former peptide, in which the DAE replaced DPhe and Pro, with an additional D-proline residue coupled to the DAE 30.²²⁴ They reported a significant difference in the cytotoxicity of isomers (5.5 to 8-fold) in HeLa and COLO-205 cancer cells. Comparable to the native gramicidin S, the open analogues displayed cytotoxicity in the μ M range (Table 7, entry 2). Moreover, *in vivo* mice studies demonstrated the potential of this approach. In comparison to the non-irradiated strategy, *in situ* activation of the closed form entailed 60% increased animal survival. Last year, they performed a detailed structure

Table 7 Photoswitchable antibiotic peptides. PSS ratios, techniques. "n.p." = not provided; "X" = PS

Entry	Sequence/PS	PSS ratio	Target	Δ -isomers	Technique
1 ²²³	c[PV-Orn-L ^D FPV-Orn-X] X = DAE 29	Closed: 80% Open: 100%	<i>S. aureus</i> , <i>S. xyloso</i>	Open/closed: 16.0	Antimicrobial assay
2 ²²⁴	c[V-Orn-L ^D FPV-Orn-L-X] X = DAE 30	Closed: 80% Open: 100%	HeLa/COLO-205 cells	Open/closed: 8.00	Cytotoxicity assay
3 ²²⁵	c[V-Orn-L ^D FPV-Orn-X] X = AMPB 1	<i>cis</i> : 90% ($\tau_{1/2}$ = 84 h in DMSO) <i>trans</i> : 88%	<i>S. aureus</i>	<i>cis/trans</i> : 4.00	Antimicrobial assay

activity analysis with five different DAE-modified gramicidin S compounds.²²⁶ This study highlighted general conclusions for future optimized designs. Thus, it seems to be beneficial to limit the ring size to 10 or 12 residues, and a subtle increase in the hydrophilic face polarity. Other crucial aspects are the alteration of the hydrophobicity/hydrophilicity balance and the dimerization of peptides, since it would lower the homeolytic activity. Importantly, no correlation between antibacterial and cytotoxic activity in HeLa cells was observed.²²⁷ Last year, the group of Abell explored photophysical behaviour, which demonstrated reduced efficiency to form the closed DAE in reduced ring sizes of gramicidin peptides.²²⁸ They also explored the incorporation of the AMPB 1 into this system.²²⁵ The *cis* isomer exhibits higher antimicrobial activity compared to its *trans* form. However, in comparison to DAE, the AMPB 1-containing peptides displayed decrease in both, antimicrobial activity (8-fold) and isomer differences (down to 4-fold), which confirms the importance of photoswitchable β -turn analogues for this system (Table 7, entry 3).²²⁵ Of note, lysine rich antimicrobial peptides connected to a spiropyran moiety *via* different linker sizes were explored.²²⁹ This is the only example of a bioactive spiropyran-modified peptides. However, the spiropyran moiety merely acted as an aggregation sensor based on its naturally occurring isomerization depending on the environment. Additional conformational studies of peptides based on this PS do not display biological activity.¹⁸

Conclusion

A natural consequence of the widespread use of both peptides and light activation in chemical biology and biomedicine has been the convergence of these tools to enable the design of bioactive photoswitchable peptides. Synthetic and structural studies of azobenzene-containing peptides first emerged over 50 years ago, pioneered by Goodman.²⁸ Following this, great advances in understanding conformational changes of photoresponsive peptides have been driven mainly by Moroder and Woolley, demonstrating higher structural changes upon irradiation, for more constrained parental peptides. Early functional examples from Nakanishi and Woolley, which targeted antigen-antibody interactions and gramicidin channels, respectively date back 30 years.^{120,178} However, only in the last 5 years has the field of photoswitchable functional peptides started to bloom. Similar to examples of light-controlled systems from materials science and supramolecular chemistry, the unsubstituted azobenzenes 1–3 and sulfonated BSBCA 15 moieties, both of which require UV irradiation, dominate the reported examples of functional photoswitchable peptides to date (78%). Of note, DAEs as β -turn mimetics have already displaced azobenzenes for applications involving control over antimicrobial peptides. Furthermore, most examples are proof-of-principle studies at the cellular level, and rely on conformational alterations driven by UV irradiation. Undoubtedly, the use of UV light limits the possibilities of these to be translated into real clinic applications, although some remarkable examples of photoswitchable peptides have been used in live animals. For example, Babii *et al.* demonstrated light-controllable

survival rates of mice injected with a DAE-incorporated antibiotic peptide.²²⁴

Visible-light photoswitches have been implemented in peptides to control biological functions,^{145,230} but represent just 3% of those reported. Indeed, major challenges relating to spectral properties and physiological stability still need to be overcome for real clinical applications. Consequently, chemists must strive to expand the functional toolbox of molecular transducers to achieve more biologically compatible structures, capable of triggering specific biological responses on demand. Here, azopyrazoles represent rising stars due to their diversity and biological compatibility.^{16,37} In any case, despite being in their infancy, photoresponsive peptide tools display tremendous potential, especially considering the diverse biological targets and compatible photoswitches described herein. We hope that this review encourages other chemical biologists to join this exciting emerging field, since the best is yet to come!

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Dr David González-Rodríguez (Universidad Autónoma de Madrid), Prof. Dr Amanda E. Hargrove (Duke University) and Prof. Dr Jennifer M. Heemstra (Emory University) for support in the early stages of this review and Prof. Dr Kathryn Haas (Saint Mary's college Notre Dame) for her critical feedback, as well as Richard Lonsdale for proof-reading the manuscript. O. V. thanks German-American Fulbright Commission for the Fulbright-Cottrell Award 2016. This work was financially supported by the German-American Fulbright Commission and the DFG program TRR81: Chromatin Changes in Differentiation and Malignancies (TRR81/3, Z04).

Notes and references

- G. Miesenböck, *Annu. Rev. Cell Dev. Biol.*, 2011, **27**, 731–758.
- D. Tischer and O. D. Weiner, *Nat. Publ. Gr.*, 2014, **15**, 551–558.
- N. A. Repina, A. Rosenbloom, A. Mukherjee and D. V. Schaffer, *Annu. Rev. Chem. Biomol. Eng.*, 2017, **8**, 13–39.
- G. Mayer and A. Heckel, *Angew. Chem., Int. Ed.*, 2006, **45**, 4900–4921.
- A. A. Beharry and G. A. Woolley, *Chem. Soc. Rev.*, 2011, **40**, 4422–4437.
- C. Brieke, F. Rohrbach, A. Gottschalk, G. Mayer and A. Heckel, *Angew. Chem., Int. Ed.*, 2012, **51**, 8446–8476.
- W. Szymański, J. M. Beierle, H. A. V. Kistemaker, W. A. Velema and B. L. Feringa, *Chem. Rev.*, 2013, **113**, 6114–6178.
- N. Ankenbruck, T. Courtney, Y. Naro and A. Deiters, *Angew. Chem., Int. Ed.*, 2018, **57**, 2768–2798.
- Z. Pianowski, *Chem. – Eur. J.*, 2019, **25**, 1–18.
- T. Fehrentz, M. Schönberger and D. Trauner, *Angew. Chem., Int. Ed.*, 2011, **50**, 12156–12182.
- M. M. Lerch, M. J. Hansen, G. M. van Dam, W. Szymanski and B. L. Feringa, *Angew. Chem., Int. Ed.*, 2016, **55**, 10978–10999.
- K. Hüll, J. Morstein and D. Trauner, *Chem. Rev.*, 2018, **118**, 10710–10747.
- A. a Beharry and G. A. Woolley, *Chem. Soc. Rev.*, 2011, **40**, 4422–4437.
- M. Dong, A. Babalhavaeji, S. Samanta, A. A. Beharry and G. A. Woolley, *Acc. Chem. Res.*, 2015, **48**, 2662–2670.
- R. J. Mart and R. K. Allemann, *Chem. Commun.*, 2016, **52**, 12262–12277.

- 16 S. Crespi, N. A. Simeth and B. König, *Nat. Rev. Chem.*, 2019, **3**, 133–146.
- 17 I. V. Komarov, S. Afonin, O. Babii, T. Schober and A. S. Ulrich, *Chem. – Eur. J.*, 2018, **24**, 11245–11254.
- 18 R. Klajn, *Chem. Soc. Rev.*, 2014, **43**, 148–184.
- 19 S. Kitzig, M. Thilemann, T. Cordes and K. Rück-Braun, *ChemPhysChem*, 2016, **17**, 1252–1263.
- 20 S. Wiedbrauk and H. Dube, *Tetrahedron Lett.*, 2015, **56**, 4266–4274.
- 21 A. S. Lubbe, W. Szymanski and B. L. Feringa, *Chem. Soc. Rev.*, 2017, **46**, 1052–1079.
- 22 R. J. Mart and R. K. Allemann, *Chem. Commun.*, 2016, **110**, 257–269.
- 23 R. P. Kruger, *Cell*, 2017, **171**, 497.
- 24 A. J. Kastin and W. Pan, *Curr. Pharm. Des.*, 2010, **16**, 3390–3400.
- 25 B. L. Feringa and W. R. Browne, *Molecular Switches*, Wiley-VCH Verlag, Weinheim, 2011.
- 26 J. L. Lau and M. K. Dunn, *Bioorg. Med. Chem.*, 2018, **26**, 2700–2707.
- 27 E. Valeur, S. M. Guéret, H. Adihou, R. Gopalakrishnan, M. Lemurell, H. Waldmann, T. N. Grossmann and A. T. Plowright, *Angew. Chem., Int. Ed.*, 2017, **56**, 10294–10323.
- 28 M. Goodman and A. Kossoy, *J. Am. Chem. Soc.*, 1966, **88**, 5010–5015.
- 29 W. A. Velema, W. Szymanski and B. L. Feringa, *J. Am. Chem. Soc.*, 2014, **136**, 2178–2191.
- 30 L. Ulysse and J. Chmielewski, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 2145–2146.
- 31 J. Juodaityte and N. Sewald, *J. Biotechnol.*, 2004, **112**, 127–138.
- 32 M. J. Hansen, M. M. Lerch, W. Szymanski and B. L. Feringa, *Angew. Chem., Int. Ed.*, 2016, **55**, 13514–13518.
- 33 M. Cigl, A. Bubnov, M. Kaspar, F. Hampf, V. Hamplová, O. Pacherová and J. Svoboda, *J. Mater. Chem. C*, 2016, **4**, 5326–5333.
- 34 M. J. Hansen, M. M. Lerch, W. Szymanski and B. L. Feringa, *Angew. Chem., Int. Ed.*, 2016, **55**, 13514–13518.
- 35 M. Dong, A. Babalhavaeji, C. V. Collins, K. Jarrah, O. Sadovski, Q. Dai and G. A. Woolley, *J. Am. Chem. Soc.*, 2017, **139**, 13483–13486.
- 36 L. Albert, A. Peñalver, N. Djokovic, L. Werel, M. Hoffarth, D. Ruzic, J. Xu, L. O. Essen, K. Nikolich, Y. Dou and O. Vázquez, *ChemBioChem*, 2019, **20**, 1417–1429.
- 37 L. Stricker, M. Böckmann, T. M. Kirse, N. L. Doltsinis and B. J. Ravoo, *Chem. – Eur. J.*, 2018, **24**, 8639–8647.
- 38 C. E. Weston, R. D. Richardson, P. R. Haycock, A. J. P. White and M. J. Fuchter, *J. Am. Chem. Soc.*, 2014, **136**, 11878–11881.
- 39 D. H. Waldeck, *Chem. Rev.*, 1991, **91**, 415–436.
- 40 K. Fujimoto, M. Amano, Y. Horibe and M. Inouye, *Org. Lett.*, 2006, **8**, 285–287.
- 41 C. Poloni, M. C. A. Stuart, P. Van Der Meulen, W. Szymanski and B. L. Feringa, *Chem. Sci.*, 2015, **6**, 7311–7318.
- 42 M. Erdélyi, M. Varedian, C. Sköld, I. B. Niklasson, J. Nurbo, Å. Persson, J. Bergquist and A. Gogoll, *Org. Biomol. Chem.*, 2008, **6**, 4356–4373.
- 43 D. Lachmann, A. Konieczny, M. Keller and B. König, *Org. Biomol. Chem.*, 2019, **17**, 2467–2478.
- 44 H. R. Allcock and C. Kim, *Macromolecules*, 1991, **24**, 2846–2851.
- 45 J. E. Zweig and T. R. Newhouse, *J. Am. Chem. Soc.*, 2017, **139**, 10956–10959.
- 46 T. Loughheed, V. Borisenko, T. Hennig, K. Rück-Braun and G. A. Woolley, *Org. Biomol. Chem.*, 2004, **2**, 2798–2801.
- 47 F. Kink, M. P. Collado, S. Wiedbrauk, P. Mayer and H. Dube, *Chem. – Eur. J.*, 2017, **23**, 6237–6243.
- 48 T. Wendler, C. Schütt, C. Näther and R. Herges, *J. Org. Chem.*, 2012, **77**, 3284–3287.
- 49 X. M. Liu, X. Y. Jin, Z. X. Zhang, J. Wang and F. Q. Bai, *RSC Adv.*, 2018, **8**, 11580–11588.
- 50 D. C. Burns, F. Zhang and G. A. Woolley, *Nat. Protoc.*, 2007, **2**, 251–258.
- 51 A. a. Beharry, O. Sadovski and G. A. Woolley, *J. Am. Chem. Soc.*, 2011, **133**, 19684–19687.
- 52 D. Bleger, J. Schwarz, A. M. Brouwer and S. Hecht, *J. Am. Chem. Soc.*, 2012, **134**, 20597–20600.
- 53 M. Irie, *Chem. Rev.*, 2000, **100**, 1685–1716.
- 54 T. Yamaguchi, T. Seki, T. Tamaki and K. Ichimura, *Bull. Chem. Soc. Jpn.*, 1992, **65**, 649–656.
- 55 B. Maerz, S. Wiedbrauk, S. Oesterling, E. Samoylova, A. Nenov, P. Mayer, R. De Vivie-Riedle, W. Zinth and H. Dube, *Chem. – Eur. J.*, 2014, **20**, 13984–13992.
- 56 B. L. Feringa, R. A. van Delden, N. Koumura and E. M. Geertsema, *Chem. Rev.*, 2000, **100**, 1789–1816.
- 57 T. Stafforst and D. Hilvert, *Chem. Commun.*, 2009, 287–288.
- 58 S. Nakamura and M. Irie, *J. Org. Chem.*, 1988, **53**, 6136–6138.
- 59 I. V. Komarov, S. Afonin, O. Babii, T. Schober and A. S. Ulrich, *Chem. – Eur. J.*, 2018, **24**, 11245–11254.
- 60 L. Ulysse, J. Cubillos and J. Chmielewski, *J. Am. Chem. Soc.*, 1995, **117**, 8466–8467.
- 61 M. Varedian, M. Erdélyi, A. Presson and A. Gogoll, *J. Pept. Res.*, 2009, **15**, 107–113.
- 62 S.-L. Dong, M. Löweneck, T. E. Schrader, W. J. Schreier, W. Zinth, L. Moroder and C. Renner, *Chem. – Eur. J.*, 2006, **12**, 1114–1120.
- 63 T. E. Schrader, T. Cordes, W. J. Schreier, F. O. Koller, S. L. Dong, L. Moroder and W. Zinth, *J. Phys. Chem.*, 2011, **115**, 5219–5226.
- 64 A. A. Deeg, M. S. Rampp, A. Popp, B. M. Pilles, T. E. Schrader, L. Moroder, K. Hauser and W. Zinth, *Chem. – Eur. J.*, 2014, **20**, 694–703.
- 65 M. Erdélyi, A. Karlén and A. Gogoll, *Chem. – Eur. J.*, 2006, **12**, 403–412.
- 66 C. Hoppmann, A. Richter, P. Schmieder, K. Rück-Braun, S. Seedorff, H. Fabian and M. Beyermann, *Angew. Chem., Int. Ed.*, 2009, **48**, 6636–6639.
- 67 N. Regner, T. T. Herzog, K. Haiser, C. Hoppmann, M. Beyermann, J. Sauermaun, M. Engelhard, T. Cordes, K. Rück-Braun and W. Zinth, *J. Phys. Chem. B*, 2012, **116**, 4181–4191.
- 68 R. Behrendt, C. Renner, M. Schenk, F. Wang, J. Wachtveitl, D. Oesterheld and L. Moroder, *Angew. Chem., Int. Ed.*, 1999, **38**, 2771–2774.
- 69 C. Renner and L. Moroder, *ChemBioChem*, 2006, **7**, 868–878.
- 70 R. Behrendt, M. Schenk, H. J. Musiol and L. Moroder, *J. Pept. Sci.*, 1999, **5**, 519–529.
- 71 C. Renner, J. Cramer, R. Behrendt and L. Moroder, *Biopolymers*, 2000, **54**, 489–500.
- 72 P. H. Nguyen, R. D. Gorbunov and G. Stock, *Biophys. J.*, 2006, **91**, 1224–1234.
- 73 C. Renner, U. Kusebauch, M. Löweneck, A. G. Milbradt and L. Moroder, *J. Pept. Res.*, 2005, **65**, 4–14.
- 74 C. Renner, R. Behrendt, N. Heim and L. Moroder, *Biopolymers*, 2002, **63**, 382–393.
- 75 H. Satzger, C. Root, C. Renner, R. Behrendt, L. Moroder, J. Wachtveitl and W. Zinth, *Chem. Phys. Lett.*, 2004, **396**, 191–197.
- 76 M. Löweneck, A. G. Milbradt, C. Root, H. Satzger, W. Zinth, L. Moroder and C. Renner, *Biophys. J.*, 2006, **90**, 2099–2108.
- 77 R. Pfister, J. Helbing, P. Hamm, G. A. Woolley, J. A. Ihalainen, L. Chi, J. Bredenbeck and I. H. M. van Stokkum, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 5383–5388.
- 78 J. A. Ihalainen, B. Paoli, S. Muff, E. H. G. Backus, J. Bredenbeck, G. A. Woolley, A. Caflisch and P. Hamm, *PNAS*, 2008, **105**, 9588–9593.
- 79 D. G. Flint, J. R. Kumita, O. S. Smart and G. A. Woolley, *Chem. Biol.*, 2002, **9**, 391–397.
- 80 P. Hamm, J. Helbing, J. Bredenbeck, J. R. Kumita and G. A. Woolley, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 2379–2384.
- 81 L. Chi, O. Sadovski and G. A. Woolley, *Bioconjugate Chem.*, 2006, **17**, 670–676.
- 82 A. A. Beharry, O. Sadovski and G. A. Woolley, *Org. Biomol. Chem.*, 2008, **6**, 4323–4332.
- 83 Z. Zhang, D. C. Burns, J. R. Kumita, O. S. Smart and G. A. Woolley, *Bioconjugate Chem.*, 2003, **14**, 824–829.
- 84 A. M. Ali and G. A. Woolley, *Org. Biomol. Chem.*, 2013, **11**, 5325–5331.
- 85 F. Zhang, O. Sadovski and G. A. Woolley, *ChemBioChem*, 2008, **9**, 2147–2154.
- 86 S. Samanta and G. A. Woolley, *ChemBioChem*, 2011, **12**, 1712–1723.
- 87 M. Blanco-Lomas, S. Samanta, P. J. Campos, G. A. Woolley and D. Sampedro, *J. Am. Chem. Soc.*, 2012, **134**, 6960–6963.
- 88 A. Fissi, F. Ciardelli, Z. D. Fabbri, G. Ruggeri and K. Umezawa, *Biopolymers*, 1993, **33**, 1505–1517.
- 89 N. Angelini, B. Corrias, A. Fissi, O. Pieroni and F. Lenci, *Biophys. J.*, 1998, **74**, 2601–2610.
- 90 U. Kusebauch, S. A. Cadamuro, H. J. Musiol, L. Moroder and C. Renner, *Chem. – Eur. J.*, 2007, **13**, 2966–2973.
- 91 S. Marqusee, V. H. Robbins and R. L. Baldwin, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 5286–5290.
- 92 G. Merutka, W. Shalongo and E. Stellwagen, *Biochemistry*, 1991, **30**, 4245–4248.
- 93 G. A. Woolley, *Acc. Chem. Res.*, 2005, **38**, 486–493.
- 94 J. M. Thornton, A. E. Todd, D. Milburn, N. Borkakoti and C. A. Orengo, *Nat. Struct. Biol.*, 2000, **7**, 991–994.

- 95 N. London, B. Raveh and O. Schueler-Furman, *Curr. Opin. Chem. Biol.*, 2013, **17**, 952–959.
- 96 L. Nevola and E. Giralt, *Chem. Commun.*, 2015, **51**, 3302–3315.
- 97 L.-G. Milroy, T. N. Grossmann, S. Hennig, L. Brunsveld and C. Ottmann, *Chem. Rev.*, 2014, **114**, 4695–4748.
- 98 M. O. Hengartner, *Nature*, 2000, **407**, 770–776.
- 99 S. Cory, D. C. Huang and J. M. Adams, *Oncogene*, 2003, **22**, 8590–8607.
- 100 T. Chittenden, E. A. Harrington, R. O'Connor, C. Flemington, R. J. Lutz, G. I. Evan and B. C. Guild, *Nature*, 1995, **372**, 733–736.
- 101 S. J. Korsmeyer, K. Wang, D. T. Chao, C. L. Milliman and X. M. Yin, *Genes Dev.*, 1996, **10**, 2859–2869.
- 102 M. Fiillbeck, E. Michalsky, I. Stephanie, P. Henklein, H. Kuhn, K. Rück-Braun and R. Preissnerl, *Genome Inf.*, 2006, **17**, 141–151.
- 103 S. Kneissl, E. J. Loveridge, C. Williams, M. P. Crump and R. K. Allemann, *ChemBioChem*, 2008, **9**, 3046–3054.
- 104 R. J. Mart, R. J. Errington, C. L. Watkins, S. C. Chappell, M. Wiltshire, A. T. Jones, P. J. Smith and R. K. Allemann, *Mol. Biosyst.*, 2013, **9**, 2597–2605.
- 105 P. Wysoczanski, R. J. Mart, E. J. Loveridge, C. Williams, S. B. Whittaker, M. P. Crump and R. K. Allemann, *J. Am. Chem. Soc.*, 2012, **134**, 7644–7647.
- 106 L. Nevola, M. Varese, A. Martín-Quirós, G. Mari, K. Eckelt, P. Gorostiza and E. Giralt, *ChemMedChem*, 2019, **14**, 100–106.
- 107 U. Krauss, T. Drepper and K. E. Jaeger, *Chem. – Eur. J.*, 2011, **17**, 2552–2560.
- 108 W. Szymański, J. M. Beierle, H. A. V. Kistemaker, W. A. Velema and B. L. Feringa, *Chem. Rev.*, 2013, **113**, 6114–6178.
- 109 M. M. Lerch, M. J. Hansen, G. M. van Dam, W. Szymanski and B. L. Feringa, *Angew. Chem., Int. Ed.*, 2016, **55**, 10978–10999.
- 110 J. D. A. Tyndall, T. Nall and D. P. Fairlie, *Chem. Rev.*, 2005, **105**, 973–999.
- 111 M. R. Angelastro, S. Mehdi, J. P. Burkhart, N. P. Peet and P. Bey, *J. Med. Chem.*, 1990, **33**, 13–16.
- 112 A. J. Harvey and A. D. Abell, *Tetrahedron*, 2000, **56**, 9763–9771.
- 113 D. Pearson and A. D. Abell, *Org. Biomol. Chem.*, 2006, **4**, 3618–3625.
- 114 B. Blanco, K. A. Palasis, A. Adwal, D. F. Callen and A. D. Abell, *Bioorg. Med. Chem.*, 2017, **25**, 5050–5054.
- 115 M. J. Hansen, W. A. Velema, G. De Bruin, H. S. Overkleeft, W. Szymanski and B. L. Feringa, *ChemBioChem*, 2014, **15**, 2053–2057.
- 116 D. Schlatter, S. Brack, D. W. Banner, S. Batey, J. Benz, J. Bertschinger, W. Huber, C. Joseph, A. C. Rufer, A. Van Der Klooster, M. Weber, D. Grabulovski and M. Hennig, *MAbs*, 2012, **4**, 497–508.
- 117 A. Babalhavaeji and G. A. Woolley, *Chem. Commun.*, 2018, **54**, 1591–1594.
- 118 A. R. Davidson, F. Zhang, A. Zarrine-Afsar, M. S. Al-Abdul-Wahid, R. S. Prosser and G. A. Woolley, *J. Am. Chem. Soc.*, 2009, **131**, 2283–2289.
- 119 S. Samanta, A. A. Beharry, O. Sadovski, T. M. McCormick, A. Babalhavaeji, V. Tropepe and G. A. Woolley, *J. Am. Chem. Soc.*, 2013, **135**, 9777–9784.
- 120 M. Harada, M. Sisido, J. Hirose and M. Nakanishi, *FEBS Lett.*, 1991, **286**, 6–8.
- 121 M. Harada, M. Sisido, J. Hirose and M. Nakanishi, *Bull. Chem. Soc. Jpn.*, 1994, **67**, 1380–1385.
- 122 J. Parisot, K. Kurz, F. Hilbrig and R. Freitag, *J. Sep. Sci.*, 2009, **32**, 1613–1624.
- 123 J. Kuil, L. T. M. Van Wandelen, N. J. De Mol and R. M. J. Liskamp, *Bioorg. Med. Chem.*, 2008, **16**, 1393–1399.
- 124 J. Kuil, L. T. M. Van Wandelen, N. J. De Mol and R. M. J. Liskamp, *J. Pept. Sci.*, 2009, **15**, 685–691.
- 125 C. Hoppmann, P. Schmieder, P. Domaing, G. Vogelreiter, J. Eichhorst, B. Wiesner, I. Morano, K. Rück-braun and M. Beyermann, *Angew. Chem., Int. Ed.*, 2011, **50**, 7699–7702.
- 126 K. R. S. Kumar, T. Kamei, T. Fukaminato and N. Tamaoki, *ACS Nano*, 2014, **8**, 4157–4165.
- 127 A. S. Amrutha, K. R. S. Kumar, K. Matsuo and N. Tamaoki, *Org. Biomol. Chem.*, 2016, **14**, 7202–7210.
- 128 L. Albert, J. Xu, R. Wan, V. Srinivasan, Y. Dou and O. Vázquez, *Chem. Sci.*, 2017, **8**, 4612–4618.
- 129 N. J. V. Lindgren, M. Varedian and A. Gogoll, *Chem. – Eur. J.*, 2009, **15**, 501–505.
- 130 C. Karlsson, M. Blom, M. Johansson, A. M. Jansson, E. Scifo, A. Karlén, T. Govender and A. Gogoll, *Org. Biomol. Chem.*, 2015, **13**, 2612–2621.
- 131 R. P. Siraganian, J. Zhang, K. Suzuki and K. Sada, *Mol. Immunol.*, 2001, **38**, 1229–1233.
- 132 J. E. Brennan, D. S. Chao, S. H. Gee, A. W. McGee, S. E. Craven, D. R. Santillano, Z. Wu, F. Huang, H. Xia, M. F. Peters, S. C. Froehner and D. S. Bredt, *Cell*, 1996, **84**, 757–767.
- 133 S. C. Froehner, M. E. Adams, G. D. Thomas, I. S. Yuhanna and P. W. Shaul, *Circ. Res.*, 2003, **92**, 554–560.
- 134 K. Rück-Braun, C. Hoppmann, A. D. Abell, J. Yu, J. R. Horsley and K. L. Wegener, *Biosens. Bioelectron.*, 2018, **118**, 188–194.
- 135 R. D. Vale, T. S. Reese and M. P. Sheetz, *Cell*, 1985, **42**, 39–50.
- 136 D. L. Coy, W. O. Hancock, M. Wagenbach and J. Howard, *Nat. Cell Biol.*, 1999, **1**, 288–292.
- 137 S. A. Reis, B. Ghosh, J. A. Hendricks, D. M. Szantai-Kis, L. Törk, K. N. Ross, J. Lamb, W. Read-Button, B. Zheng, H. Wang, C. Salthouse, S. J. Haggarty and R. Mazitschek, *Nat. Chem. Biol.*, 2016, **12**, 317–323.
- 138 W. Szymanski, M. E. Ourailidou, W. A. Velema, F. J. Dekker and B. L. Feringa, *Chem. – Eur. J.*, 2015, **21**, 16517–16524.
- 139 B. Heinrich, K. Bouazoune, M. Wojcik, U. Bakowsky and O. Vázquez, *Org. Biomol. Chem.*, 2019, **17**, 1827–1833.
- 140 A. S. Madsen and C. a Olsen, *Nat. Chem. Biol.*, 2016, **12**, 1–2.
- 141 H. P. Nguyen, S. Stewart, M. N. Kukwikila, S. F. Jones, D. Offenbartl-Stiegert, S. Mao, S. Balasubramanian, S. Beck and S. Howorka, *Angew. Chem., Int. Ed.*, 2019, **131**, 1–6.
- 142 J. J. Day, *Dialogues Clin. Neurosci.*, 2014, **16**, 345–357.
- 143 C. E. Weston, A. Krämer, F. Colin, Ö. Yildiz, M. G. J. Baud, F. J. Meyer-Almes and M. J. Fuchter, *ACS Infect. Dis.*, 2017, **3**, 152–161.
- 144 H. Karatas, E. C. Townsend, D. Bernard, Y. Dou and S. Wang, *J. Med. Chem.*, 2010, **53**, 5179–5185.
- 145 L. Albert, A. Peñalver, N. Djokovic, L. Werel, M. Hoffarth, D. Ruzic, J. Xu, L.-O. Essen, K. Nikolic, Y. Dou and O. Vázquez, *ChemBioChem*, 2019, **20**, 1417–1429.
- 146 R. Siewertsen, H. Neumann, B. Buchheim-Stehn, R. Herges, C. Näther, F. Renth and F. Temps, *J. Am. Chem. Soc.*, 2009, **131**, 15594–15595.
- 147 J. Nurbo, A. K. Roos, D. Muthas, E. Wahlström, D. J. Ericsson, T. Lundstedt, T. Unge and A. Karlén, *J. Pept. Sci.*, 2007, **13**, 822–832.
- 148 T. H. Kim and B. Ren, *Annu. Rev. Genomics Hum. Genet.*, 2006, **7**, 81–102.
- 149 W. H. Hudson and E. A. Ortlund, *Nat. Rev. Mol. Cell Biol.*, 2014, **15**, 749–760.
- 150 M. E. Vázquez, A. M. Camano and J. L. Mascareñas, *Chem. Soc. Rev.*, 2003, **32**, 338–349.
- 151 E. Pazos, J. Mosquera, M. E. Vázquez and J. L. Mascareñas, *ChemBioChem*, 2011, **12**, 1958–1973.
- 152 S. Boga, D. Bouzada, D. García Peña, M. Vázquez López and M. E. Vázquez, *Eur. J. Org. Chem.*, 2018, 249–261.
- 153 D. A. Guarracino, B. N. Bullock and P. S. Arora, *Biopolymers*, 2011, **95**, 1–7.
- 154 X. Wang, W. Cao, A. Cao and L. Lai, *Biophys. J.*, 2003, **84**, 1867–1875.
- 155 A. Caamaño, M. Vázquez, J. Martínez-Costas, L. Castedo and J. Mascareñas, *Angew. Chem., Int. Ed.*, 2000, **39**, 3104–3107.
- 156 J. R. Kumita, D. G. Flint, G. A. Woolley and O. S. Smart, *Faraday Discuss.*, 2002, **122**, 89–103.
- 157 G. A. Woolley, A. S. I. Jaikaran, M. Berezovski, J. P. Calarco, S. N. Krylov, O. S. Smart and J. R. Kumita, *Biochemistry*, 2006, **45**, 6075–6084.
- 158 L. Guerrero, O. S. Smart, C. J. Weston, D. C. Burns, G. A. Woolley and R. K. Allemann, *Angew. Chem., Int. Ed.*, 2005, **44**, 7778–7782.
- 159 D. Yesudhas, M. Batool, M. A. Anwar, S. Panneerselvam and S. Choi, *Genes*, 2017, **8**, 1–22.
- 160 F. Zhang, K. A. Timm, K. M. Arndt and G. A. Woolley, *Angew. Chem., Int. Ed.*, 2010, **122**, 4035–4038.
- 161 A. M. Ali, M. W. Forbes and G. A. Woolley, *ChemBioChem*, 2015, **16**, 1757–1763.
- 162 S. A. Morgan and G. A. Woolley, *Photochem. Photobiol. Sci.*, 2010, **9**, 1320–1326.
- 163 A. M. Ali, J. M. Reis, Y. Xia, A. J. Rashid, V. Mercaldo, B. J. Walters, K. E. Brechun, V. Borisenko, S. A. Josselyn, J. Karanicolas and G. A. Woolley, *Chem. Biol.*, 2015, **22**, 1531–1539.
- 164 A. Laughon and M. P. Scott, *Nature*, 1984, **310**, 25–31.
- 165 L. Guerrero, O. S. Smart, G. A. Woolley and R. K. Allemann, *J. Am. Chem. Soc.*, 2005, **127**, 15624–15629.
- 166 K. Fujimoto, M. Kajino, I. Sakaguchi and M. Inouye, *Chem. – Eur. J.*, 2012, **18**, 9834–9840.

- 167 A. Nomura and A. Okamoto, *Chem. Commun.*, 2009, 1906–1908.
- 168 B. L. Feringa, W. Szymanski, C. Poloni, G. M. Murawska and N. A. Simeth, *Chem. – Eur. J.*, 2019, **25**, 4965.
- 169 M. Cassandri, A. Smirnov, F. Novelli, C. Pitolli, M. Agostini, M. Malewicz, G. Melino and G. Raschella, *Cell Death Discovery*, 2017, **3**, 17071.
- 170 S. Durai, M. Mani, K. Kandavelou, J. Wu, M. H. Porteus and S. Chandrasegaran, *Nucleic Acids Res.*, 2005, **33**, 5978–5990.
- 171 F. D. Urnov, E. J. Rebar, M. C. Holmes, H. S. Zhang and P. D. Gregory, *Nat. Rev. Genet.*, 2010, **11**, 636–646.
- 172 R. Hock, T. Furusawa, T. Ueda and M. Bustin, *Trends Cell Biol.*, 2007, **17**, 72–79.
- 173 T. Kenakin and L. J. Miller, *Pharmacol. Rev.*, 2010, **62**, 265–304.
- 174 T. Fehrentz, M. Schönberger and D. Trauner, *Angew. Chem., Int. Ed.*, 2011, **50**, 12156–12182.
- 175 W. Szymański, D. Yilmaz, A. Koçer and B. L. Feringa, *Acc. Chem. Res.*, 2013, **46**, 2910–2923.
- 176 P. Gorostiza and E. Isacoff, *Mol. Biosyst.*, 2007, **3**, 686–704.
- 177 T. Fehrentz, M. Schönberger and D. Trauner, *Angew. Chem., Int. Ed.*, 2011, **50**, 12222–12223.
- 178 C. J. Stankovic, S. H. Heinemann and S. L. Schreiber, *Biochim. Biophys. Acta*, 1991, **1061**, 163–170.
- 179 L. Lien, D. C. J. Jaikaran, Z. Zhang and G. A. Woolley, *J. Am. Chem. Soc.*, 1996, **118**, 12222–12223.
- 180 V. Borisenko, D. C. Burns, Z. Zhang and G. A. Woolley, *J. Am. Chem. Soc.*, 2000, **122**, 6364–6370.
- 181 T. Ueda, K. Nagamine, S. Kimura and Y. Imanishi, *J. Chem. Soc., Perkin Trans. 2*, 1995, 365–368.
- 182 E. Kato, T. Ueda, S. Kimura and Y. Imanishi, *Biophys. Chem.*, 1994, **49**, 215–222.
- 183 E. Ruoslahti, *Annu. Rev. Cell Dev. Biol.*, 1996, **12**, 697–715.
- 184 M. Schütt, S. S. Krupka, A. G. Milbradt, S. Deindl, E.-K. Sinner, D. Oesterheld, C. Renner and L. Moroder, *Chem. Biol.*, 2003, **10**, 487–490.
- 185 A. G. Milbradt, M. Löweneck, S. S. Krupka, M. Reif, E. K. Sinner, L. Moroder and C. Renner, *Biopolymers*, 2005, **77**, 304–313.
- 186 D. Liu, Y. Xie, H. Shao and X. Jiang, *Angew. Chem., Int. Ed.*, 2009, **48**, 4406–4408.
- 187 J. Auernheimer, C. Dahmen, U. Hersel, A. Bausch and H. Kessler, *J. Am. Chem. Soc.*, 2005, **127**, 16107–16110.
- 188 J. Broichhagen, T. Podewin, H. Meyer-Berg, Y. Von Ohlen, N. R. Johnston, B. J. Jones, S. R. Bloom, G. A. Rutter, A. Hoffmann-Röder, D. J. Hodson and D. Trauner, *Angew. Chem., Int. Ed.*, 2015, **54**, 15565–15569.
- 189 T. Podewin, J. Broichhagen, C. Frost, D. Groneberg, J. Ast, H. Meyer-Berg, N. H. F. Fine, A. Friebe, M. Zacharias, D. J. Hodson, D. Trauner and A. Hoffmann-Röder, *Chem. Sci.*, 2017, **8**, 4644–4653.
- 190 L. Nevola, A. Martin-Quiros, K. Eckelt, N. Camarero, S. Tosi, A. Llobet, E. Giralt and P. Gorostiza, *Angew. Chem., Int. Ed.*, 2013, **52**, 7704–7708.
- 191 J. M. Zhu, Y. Zhu and R. Liu, *Int. J. Equity Health*, 2008, **7**, 1–9.
- 192 P. Ghanouni, J. J. Steenhuis, D. L. Farrens and B. K. Kobilka, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 5997–6002.
- 193 G. Swaminath, Y. Xiang, T. W. Lee, J. Steenhuis, C. Parnot and B. K. Kobilka, *J. Biol. Chem.*, 2004, **279**, 686–691.
- 194 A. A. Kaspar and J. M. Reichert, *Drug Discovery Today*, 2013, **18**, 807–817.
- 195 J. Broichhagen, N. R. Johnston, Y. Von Ohlen, H. Meyer-berg, B. J. Jones, S. R. Bloom, G. A. Rutter, D. Trauner and D. J. Hodson, *Angew. Chem., Int. Ed.*, 2016, **55**, 5865–5868.
- 196 J. Broichhagen, M. Schönberger, S. C. Cork, J. A. Frank, P. Marchetti, M. Bugliani, A. M. J. Shapiro, S. Trapp, G. A. Rutter, D. J. Hodson and D. Trauner, *Nat. Commun.*, 2014, **5**, 1–11.
- 197 S. W. M. John, J. H. Krege, P. M. Oliver, J. R. Hagaman, J. B. Hodgkin, S. C. Pang, T. G. Flynn and O. Smithies, *Science*, 1995, **267**, 679–681.
- 198 M. R. Cowie, A. D. Struthers, D. A. Wood, A. J. S. Coats, S. G. Thompson, P. A. Poole-Wilson and G. C. Sutton, *Lancet*, 1997, **350**, 1349–1353.
- 199 Z. Y. Weinberg and M. A. Puthenveedu, *Traffic*, 2019, **20**, 121–129.
- 200 H. T. McMahon and E. Boucrot, *Nat. Rev. Mol. Cell Biol.*, 2011, **12**, 517–533.
- 201 A. Martin-Quiros, L. Nevola, K. Eckelt, P. Gorostiza and E. Giralt, *Chem. Biol.*, 2015, **22**, 31–37.
- 202 J. P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M. J. Gait, L. V. Chernomordik and B. Lebleu, *J. Biol. Chem.*, 2003, **278**, 585–590.
- 203 M. Lindgren, M. Halbrink, A. Prochiantz and U. Langel, *Trends Pharmacol. Sci.*, 2000, **21**, 99–103.
- 204 F. Milletti, *Drug Discovery Today*, 2012, **17**, 850–860.
- 205 S. Stalmans, N. Bracke, M. D'Hondt, B. Gevaert, K. Peremans, C. Burvenich, E. Wynendaele and B. De Spiegeleer, *PLoS One*, 2013, **8**, e71752.
- 206 A. Prestel and H. M. Möller, *Chem. Commun.*, 2016, **52**, 701–704.
- 207 G. C. Kim, J. H. Ahn, J. H. Oh, S. Nam, S. Hyun, J. Yu and Y. Lee, *Biomacromolecules*, 2018, **19**, 2863–2869.
- 208 L. Dietrich, B. Rathmer, K. Ewan, T. Bange, S. Heinrichs, T. C. Dale, D. Schade and T. N. Grossmann, *Cell Chem. Biol.*, 2017, **24**, 958–968.
- 209 F. Chiti and C. M. Dobson, *Annu. Rev. Biochem.*, 2006, **75**, 333–366.
- 210 R. L. Nussbaum and C. E. Ellis, *New Engl. J. Med.*, 2003, **348**, 1356–1364.
- 211 G. Merlini and V. Bellotti, *N. Engl. J. Med.*, 2003, **349**, 583–596.
- 212 M. Sunde, L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys and C. C. F. Blake, *J. Mol. Biol.*, 1997, **273**, 729–739.
- 213 T. E. Schrader, W. J. Schreier, T. Cordes, F. O. Koller, G. Babitzki, R. Denschlag, C. Renner, M. Loweneck, S.-L. Dong, L. Moroder, P. Tavan and W. Zinth, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 15729–15734.
- 214 A. A. Deeg, T. E. Schrader, S. Kempter, J. Pfizer, L. Moroder and W. Zinth, *ChemPhysChem*, 2011, **12**, 559–562.
- 215 A. A. Deeg, T. E. Schrader, J. Pfizer, L. Moroder and W. Zinth, *Adv. Biomed. Spectrosc.*, 2012, **27**, 387–391.
- 216 T. M. Doran, E. A. Anderson, S. E. Latchney, L. A. Opanashuk and B. L. Nilsson, *ACS Chem. Neurosci.*, 2012, **3**, 211–220.
- 217 S. A. Waldauer, S. Hassan, B. Paoli, P. M. Donaldson, R. Pfister, P. Hamm, A. Calfish and R. Pellarin, *J. Phys. Chem.*, 2012, **116**, 8961–8973.
- 218 C. Hoppmann, C. Barucker, D. Lorenz, G. Multhaup and M. Beyermann, *ChemBioChem*, 2012, **13**, 2657–2660.
- 219 F. Prestinaci, P. Pezzotti and A. Pantosti, *Pathog. Global Health*, 2015, **109**, 309–318.
- 220 L. H. Kondejewski, S. W. Farmer, D. S. Wishart, R. E. Hancock and R. S. Hodges, *Int. J. Pept. Protein Res.*, 1996, **47**, 460–466.
- 221 T. Kato and N. Izumiya, *Biochim. Biophys. Acta, Biomembr.*, 1977, **493**, 235–238.
- 222 G. Nagamurthi and S. Rambhav, *J. Biosci.*, 1985, **7**, 323–329.
- 223 O. Babii, S. Afonin, M. Berditsch, S. Reißer, P. K. Mykhailiuk, V. S. Kubyshevskii, T. Steinbrecher, A. S. Ulrich and I. V. Komarov, *Angew. Chem., Int. Ed.*, 2014, **126**, 3460–3463.
- 224 O. Babii, S. Afonin, L. V. Garmanchuk, V. V. Nikulina, T. V. Nikolaienko, O. V. Storozhuk, D. V. Shelest, O. I. Dsayukevich, L. I. Ostapchenko, V. Iurchenko, S. Zozulya, A. S. Ulrich and I. V. Komarov, *Angew. Chem.*, 2016, **128**, 5583–5586.
- 225 Y. Q. Yeoh, J. Yu, S. W. Polyak, J. R. Horsley and A. D. Abell, *ChemBioChem*, 2018, **19**, 2591–2597.
- 226 O. Babii, S. Afonin, A. Y. Ischenko, T. Schober, A. O. Negelia, G. M. Tolstanova, L. V. Garmanchuk, L. I. Ostapchenko, I. V. Komarov and A. S. Ulrich, *J. Med. Chem.*, 2018, **61**, 10793–10813.
- 227 O. Babii, S. Afonin, A. Y. Ischenko, T. Schober, A. O. Negelia, G. M. Tolstanova, L. V. Garmanchuk, L. I. Ostapchenko, I. V. Komarov and A. S. Ulrich, *J. Med. Chem.*, 2018, **61**, 10793–10813.
- 228 C. Schweigert, O. Babii, S. Afonin, T. Schober, J. Leier, N. Michenfelder, I. V. Komarov, A. Ulrich and A. N. Unterreiner, *ChemPhotoChem*, 2019, **3**, 1–9.
- 229 L. Chen, Y. Zhu, D. Yang, R. Zou, J. Wu and H. Tian, *Sci. Rep.*, 2014, **4**, 1–7.
- 230 S. Samanta, M. Convertino, N. V. Dokholyan, A. Deiters and J. Luo, *ChemBioChem*, 2018, **19**, 2178–2185.

4.1.1 Author Contribution

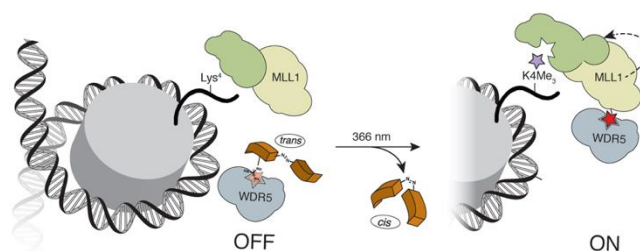
The idea of a need in the current literature for a recapitulation of the applications of photoresponsive peptides linked to the control of biological functions was envisioned by Jun.-Prof. Dr. Olalla Vázquez. I, Lea Albert, performed the detailed literature search as well as the collection, evaluation and organization of information. I designed and created most figures, and contributed to write the whole manuscript. Jun.-Prof. Dr. Olalla Vázquez had the role of corresponding author, i.e., she was leading and/or involved in all the 14-role taxonomy for author-contribution assignment.¹

Marburg,

Jun.-Prof. Dr. Olalla Vázquez

Lea Albert

4.2 Controlled inhibition of methyltransferases using photoswitchable peptidomimetics: towards an epigenetic regulation of leukemia



Shine light on epigenetics: we describe how photoswitchable peptidomimetics modulate the activity of the MLL1 enzyme affecting epigenetic states.

This chapter was published as:




L. Albert, J. Xu, R. Wan, V. Srinivasan, Y. Dou, O. Vázquez*,
Controlled inhibition of methyltransferases using photoswitchable peptidomimetics: towards an epigenetic regulation of leukemia
Chem. Sci. **2017**, *8*, 4612-4618.

Copyright Royal Society of Chemistry 2019. Material in this chapter was reproduced from the above mentioned reference with permission from The Royal Society of Chemistry.

In the following sections, the molecules included in this manuscript, will be named as **MB-X**, in which MB means manuscript B, and X means number of compound as given in this manuscript.

Cite this: *Chem. Sci.*, 2017, 8, 4612

Controlled inhibition of methyltransferases using photoswitchable peptidomimetics: towards an epigenetic regulation of leukemia†

Lea Albert, ^a Jing Xu, ^b Ruiwei Wan, ^b Vasundara Srinivasan, ^c Yali Dou ^b and Olalla Vázquez ^{*a}

We describe a cell-permeable photoswitchable probe capable of modulating epigenetic cellular states by disruption of an essential protein–protein interaction within the MLL1 methyltransferase core complex. Our azobenzene-containing peptides selectively block the WDR5–MLL1 interaction by binding to WDR5 with high affinity ($K_i = 1.25$ nM). We determined the co-crystal structure of this photoswitchable peptidomimetic with WDR5 to understand the interaction at the atomic level. Importantly, the photoswitchable *trans* and *cis* conformers of the probe display a clear difference in their inhibition of MLL1. We further demonstrate that the designed photo-controllable azo-peptidomimetics affect the transcription of the MLL1-target gene *Deptor*, which regulates hematopoiesis and leukemogenesis, and inhibit the growth of leukemia cells. This strategy demonstrates the potential of photopharmacological inhibition of methyltransferase protein–protein interactions as a novel method for external epigenetic control, providing a new toolbox for controlling epigenetic states.

Received 11th January 2017

Accepted 9th April 2017

DOI: 10.1039/c7sc00137a

rsc.li/chemical-science

Epigenetics, the study of variations in gene expression unrelated to changes in the DNA sequence, is one of the most promising fields in biomedical research, since genetics alone cannot explain human variation and disease.¹ Combinatorial post-translational modifications (PTMs) on histones, often referred to as the histone language,^{2–6} directly regulate the structure of chromatin and affect transcriptional activity by recruiting a large variety of proteins through protein–protein interactions (PPIs). One of the best-characterized histone PTMs is the specific methylation at lysine 4 of histone 3 (H3K4) by the mixed-lineage leukemia (MLL) enzymes. H3K4 trimethylation (H3K4me3) is commonly found at the promoter and enhancer regions of actively transcribed genes.⁷ Among different MLLs, MLL1 is essential for hematopoiesis⁸ and neurogenesis during embryonic development.^{9,10} It is also a promising therapeutic target; MLL1 deregulation has been linked to a subset of acute leukemia and solid tumors.^{11–14} In MLL1-rearranged leukemia cells, balanced chromosomal translocations lead to generation of MLL1 fusion proteins that include MLL1 N-terminal and C-terminal domains from several transcriptional elongation factors (AF4, AF9, ENL, and ELL).^{15,16} The C-terminus of MLL1 contains the catalytic SET domain, which is

regulated by PPIs within a conserved multi-component complex.¹⁷ It has been shown that binding of WD40-repeat protein 5 (WDR5) to arginine 3765 (R3765) of MLL1 is crucial to enzymatic activity.¹⁷ Both wild type and fusion MLL1 coexist in leukemia cells and contribute to the leukemic transcription program. Despite important recent advances, the exact function of MLL1 in leukemogenesis remains unclear, highlighting the importance of developing probes for MLL1.

In recent years, there have been many advances in the development of photoresponsive probes for biological intervention that open new avenues for biological and medicinal discoveries. These technologies range from classical caged compounds to the more recent optogenetic approaches. Optogenetics,^{18–21} the use of genetically encoded photoreceptors, has shown an unprecedented potential for controlling cellular behaviour in living tissues, although in some cases its application requires complex and time-consuming genetic modifications of the proteins under study, and in some instances a simpler approach with off-the-shelf reagents might be desirable. Photopharmacology^{22–25} overcomes these limitations by taking full advantage of the small-molecule photoswitches, providing excellent delivery properties and spatio-temporal resolution with affordable probes. Epigenetic regulation could also be manipulated through the development of specific epigenetic photoswitches, as exemplified by direct targeting of the histone-deacetylase enzyme (HDAC).^{26,27} However, to our knowledge, photo-controllable probes for histone methyltransferases have not yet been reported.

^aFachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Strasse 4, 35043 Marburg, Germany. E-mail: olalla.vazquez@staff.uni-marburg.de

^bDepartment of Pathology, University of Michigan, Ann Arbor, Michigan 48109, USA

^cLOEWE Zentrum für Synthetische Mikrobiologie SynMikro, Philipps-Universität Marburg, Hans-Meerwein-Strasse 4, 35043 Marburg, Germany

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7sc00137a



Herein, we report the design and synthesis of photo-responsive probes based on azobenzene-containing peptides capable of controlling the activity of MLL1 in a reversible manner. These photo-controllable peptidomimetics target the key PPI of the MLL1 core complex: WDR5-MLL1 (Fig. 1). Furthermore, we demonstrate the potential of this reversible approach without permanent knockout of the protein and increase in scope the available chemical optoepigenetic toolbox for analysing chromatin regulation.

The design of our approach is based on the recent identification of two truncations of the WDR5-Interacting peptide (WIN sequence: Ac-GSARAEVHLRKS-NH₂ (1)): Ac-ARAEVHLRKS-NH₂ (2) and NH₂-SARAEVHLRKS-NH₂ (3), which interact with higher affinity than the parent WIN ($K_i = 160$ nM, $K_i = 3$ nM and $K_i = 20$ nM for 1, 2 and 3 respectively).²⁸ The -ARA- sequence was identified as the indispensable motif for WDR5 recognition. Taking these promising peptide inhibitors of MLL1 as a starting point, we hypothesized that a conformational switch could reversibly modulate their affinity. Considering that the high affinity of peptide 2 could hamper the photoswitch by effectively trapping the active conformation in a stable complex with MLL1, we focused our synthetic efforts in using peptide 3 as a reference for the synthesis of our novel photoswitchable inhibitors.

As a molecular transducer we selected azobenzenes, owing to the large amount of information available on their photo-responsive properties.^{29–35} Specifically, as initial proof-of-concept, we chose the 4-[(4'-aminomethyl)phenylazo]benzoic acid (AMPB)³⁶ due to its synthetic simplicity, direct

incorporation into the peptide backbone^{37–39} and adequate spectroscopic properties (*cis* → *trans* irradiation at 430 nm; *trans* → *cis* irradiation at 366 nm). The fact that the WIN peptide is intrinsically disordered discouraged the rational design of a photo-responsive peptide and instead suggested a systematic amino acid scan approach for the incorporation of the photoswitch. The set of peptides was prepared following the standard Fmoc-solid phase methodology and the synthesis of AMPB was carried out following literature procedures.⁴⁰ The incorporation of this unnatural amino acid required optimization of the final TFA cleavage conditions to avoid side reactions with the azobenzene (see ESI†). All of the peptides displayed fast and reversible photoisomerization (Fig. S23†). Conversely, thermal *cis* → *trans* relaxation was a slow process (*cis/trans* ratio of 65 : 35 after four days in the dark; Fig. S24†), which allowed the performing of lengthy biological assays.

Once the battery of peptides was synthesized and the reversibility of their photoisomerization demonstrated, we explored whether the two photoisomers displayed suitable differences in binding affinity for WDR5 (Fig. 2). The peptide solutions were irradiated at 366 nm to generate the *cis* isomers, while the *trans* isomers were obtained through thermal relaxation since it is preferable to produce larger fold-changes between isomers.^{41,42} We determined the binding affinities using the fluorescence polarization (FP)-based competitive binding assay.²⁸ Dose-response curves provide IC₅₀ values, but since these data depend upon the experimental conditions, it is advisable to convert them to inhibition constants (K_i) for objective comparison (Fig. 2).⁴³ We verified that the excitation/emission wavelengths of the 5-carboxyfluorescein (5-FAM) tag (485/535 nm) did not interfere with the AMPB isomerization rates by recording HPLC chromatograms of the peptides before and after the assays (Fig. S29–S58†). By comparing these chromatograms with the non-irradiated ones (Fig. S1–S20†) photo-degradative processes could also be ruled out. In the same lines, the binding affinity of parent WIN 3 to WDR5 was calculated by non-irradiation and irradiation at 366 nm and 430 nm (Fig. 2 & Table S3†). Additional control experiments with AMPB itself further corroborated the absence of quenching artefacts (Fig. S27 & S28†).

As shown in Fig. 2, whenever there was a clear difference between both photoisomers, the *trans* isomer interacted more strongly with WDR5 than the analogous *cis* isomer. On the other hand, the effect of the AMPB switch on the WDR5-MLL1 interaction was more pronounced when it was located closer to the key arginine residue in the ARA minimal motif. Consequently, the N-terminal introduction of AMPB, or the replacement of both flanking alanine residues, had drastic consequences for the binding affinities, to the point that it was not possible to obtain measurable binding constants for peptidomimetics 4 and 6. Likewise, the influence of AMPB decreased when located away from the ARA motif. Intriguingly, the K_i of peptidomimetics 7 and 8 decreased by an estimated factor of 10 in comparison to peptide 3. These results agree with the best peptide inhibitors described (K_i for 2 is 3 nM). The decrease in K_i was stronger when the polar glutamic acid residue was replaced with the hydrophobic AMPB molecule. Curiously, the

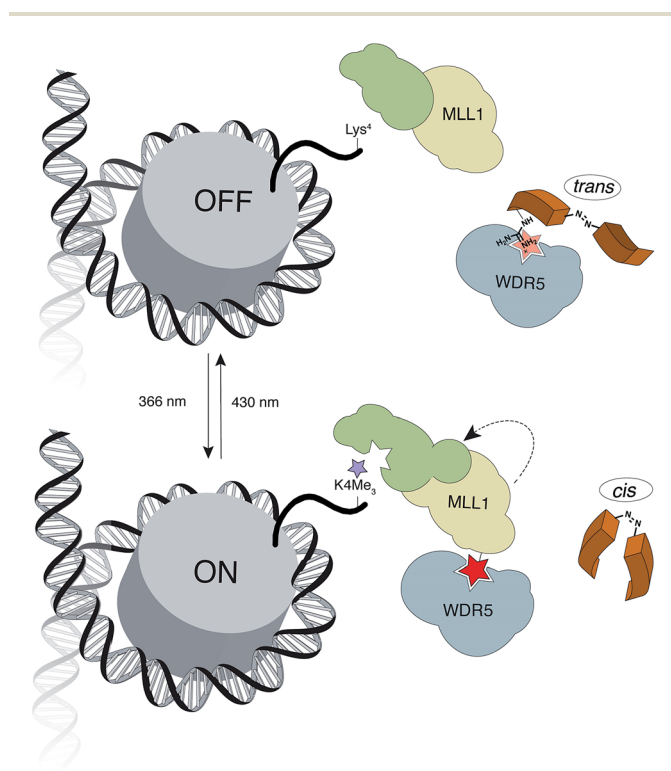


Fig. 1 Outline of the indirect strategy for MLL1 activity control through photoswitchable inhibitors of the MLL1-WDR5 interaction.



Table 1. Binding affinities of azo-photoactivatable peptides to WDR5

Peptides	Non-irradiation K_i /[nM]	Irradiation at 366 nm K_i /[nM]	<i>cis/trans</i>
H ₂ N-SARAEVHLRKS-CONH ₂ (3)	20.0 ± 0.98	24.5 ± 0.49	n.c
H ₂ N-SXARAEVHLRKS-CONH ₂ (4)	n.c	n.c	n.c
H ₂ N-SXRAEVHLRKS-CONH ₂ (5)	66.8 ± 10	432 ± 73	6.48
H ₂ N-SARXEVHLRKS-CONH ₂ (6)	n.c	n.c	n.c
H ₂ N-SARAXVHLRKS-CONH ₂ (7)	1.25 ± 0.36	6.50 ± 1.4	5.00
H ₂ N-SARAEVHLRKS-CONH ₂ (8)	2.40 ± 0.60	9.10 ± 2.2	3.79
H ₂ N-SARAEVXLRKS-CONH ₂ (9)	40.4 ± 3.9	39.7 ± 0.35	0.990
H ₂ N-SARAEVHXRKS-CONH ₂ (10)	17.0 ± 1.9	33.3 ± 4.2	1.96
H ₂ N-SARAEVHLXKS-CONH ₂ (11)	32.8 ± 4.5	70.4 ± 6.6	2.15
H ₂ N-SARAEVHLRXS-CONH ₂ (12)	19.8 ± 1.7	40.6 ± 2.5	2.05
H ₂ N-SARAXHLRKS-CONH ₂ (13)	0.293 ± 0.05	0.664 ± 0.08	3.63
Ac-ARAXVHLRKS-CONH ₂ (14)	1.45 ± 0.52	4.25 ± 0.83	2.93
Ac-ARTXVY-CONH ₂ (15)	0.109 ± 0.01	0.185 ± 0.02	0.767

X represents the incorporation of 4-[(4'-aminomethyl)phenylazo]benzoic acid in the peptide scaffold. Mean values from three independent experiments.

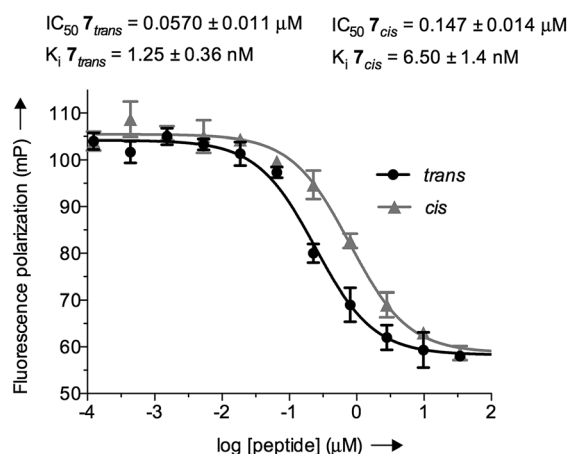
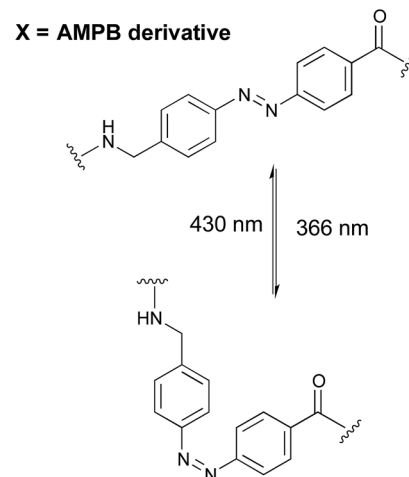


Fig. 2 Left: table of the competitive inhibition constants (K_i) to WDR5 for the azo-photoactivatable variants by FP-based competitive assays. Right: reversible photochromism of AMPB; example of the competitive binding curve for the most promising peptidomimetic 7. IC_{50} for 5 nM of tracer.

same modification in peptide 2, as well as in the recently published short version Win6mer,⁴⁴ had little influence on K_i , and smaller difference between isomers (peptidomimetics 14 and 15, respectively, Fig. 2), than the peptidomimetic 7. Furthermore, considering the *trans* isomer of AMPB is larger than glutamic acid residue, we explored the possibility of having an increase in the difference between isomers when two residues are exchanged by AMPB (peptide 13). However, this was not observed. Therefore, taken together these results led us to select peptidomimetic 7 for further structural studies.

To have a molecular interpretation of the increased affinity of 7 in comparison to 3, as well as of the affinity difference between isomers, we determined the co-crystal structure of WDR5 in complex with peptidomimetic 7, in two different states: I and II at 1.97 Å (PDB code: 5M23) and 2.43 Å (PDB code: 5M25) resolution, respectively (Fig. 3 & Table S6[†]). State I refers to the conditions where the *cis/trans* ratio of 7 was ~5 : 95 prior

to protein complexation, while in state II the *cis/trans* ratio of 7 was ~80 : 20 prior to protein complexation. Furthermore, in the latter state the crystallization experiments were performed in the absence of light, and the crystal plates were always kept in the dark. In both cases, the N-terminal part of the peptide –SARA– superimposes well, and the main difference between the two structures comes from the orientation of the azobenzene rings. Remarkably, in both states the *trans* isomer was the one trapped in the crystal. An interesting feature of the electron density maps of the WDR5-peptide complexes is that the C-terminal part of the peptide reveals static or molecular dynamic disorder, as reflected in the high B-values that indicate thermal displacement of atoms in crystal structures (Table S6[†]). The high binding affinity constants measured in solution for peptide 7 do not correlate with those seen in the crystal structures. This could be explained as an effect of the inherent dynamic nature of the peptide when bound to the protein in the



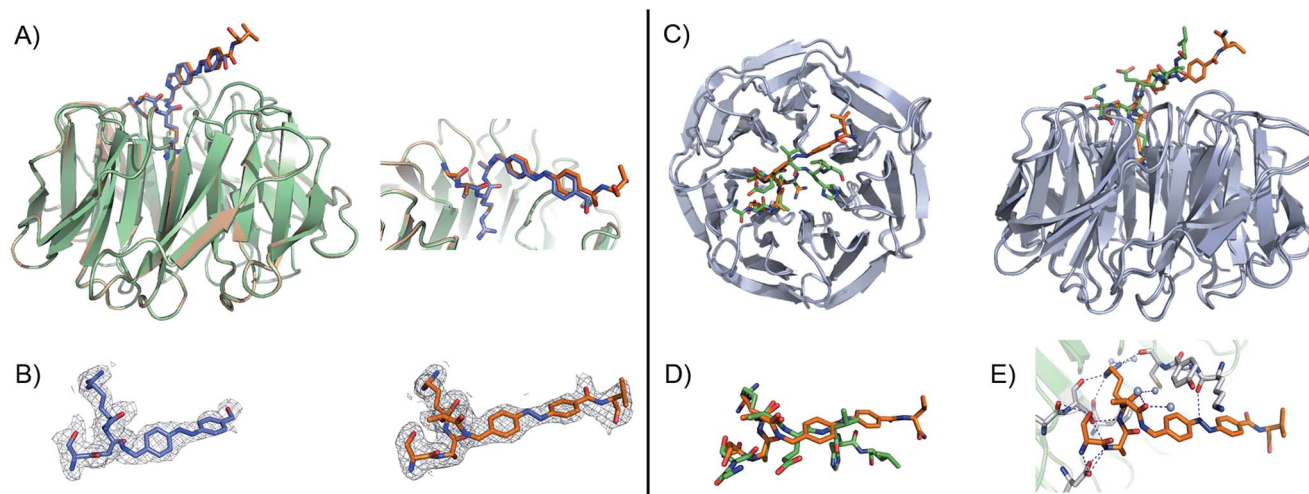


Fig. 3 (A) Close-up view and superimposition of the state I (orange) and state II (blue) of peptidomimetic 7; PDB code: 5M23 and 5M25, respectively. (B) A $(2F_o - F_c)$ experimental electron density map contoured at 1.0 sigma (grey) shows clear density for the peptides in both states: I (orange) and II (blue). (C) Superimposition and 90° rotation view of the WDR5 complex crystal structures with the peptidomimetic 7 (orange) and WIN peptide (green); PDB code 5M23 and 3EG6, respectively. (D) Close-up view of the two superimposed peptides: 7 and WIN. (E) Extensive hydrogen bond network of 7 with WDR5; 7 is depicted as orange sticks, interacting residues from WDR5 in white and water molecules as grey spheres.

crystallization buffer conditions. A similar observation has been previously described in the crystal structures of azobenzene ligands bound to streptavidin.⁴⁵

Comparing our peptidomimetic 7 with the parent WIN (PDB code: 3EG6), the key interactions of 7 with the WDR5 protein in both states agree well with those in the crystal structure of the WIN peptide–WDR5 complex.^{46,47} The overlay of the two structures clearly shows that the N-terminal part of the peptide (–SARA–) maintains similar interactions to the protein, as seen with the WIN peptide. However, from the glutamate onwards the orientation is different. The replacement of the solvent-exposed glutamic acid residue in the parent WIN peptide by the azobenzene ring provides additional stabilization through interaction with Lys259 and Tyr260 of MLL1 (Fig. 3E & S70†). The aromatic side chain of Tyr260 participates in van der Waals interactions with the second benzene ring of the azobenzene (CE2 and CD2 atoms of Tyr260 and CDH and CDI atoms of azobenzene, Fig. S70 and Table S7†) and Lys259 forms a hydrogen bond with the first nitrogen atom of the AMPB of peptide 7 (Fig. 3E & S70†).

The inherently nonlinear nature of biological interactions suggests that the relatively modest binding difference between the two peptidomimetic 7 isomers may be amplified and cause pronounced effects in a more realistic biological context. To test this, we studied the inhibition of MLL1 H3K4 methyltransferase activity through the previously reported *in vitro* HMT assay using the MLL1 core complex (*i.e.* MLL1, WDR5, RbBP5, and ASH2L) and the H3 20mer peptide as substrate.¹⁷ Gratifyingly, our assay showed that perturbation of the WDR5–MLL1 protein–protein interaction by the photoswitchable probe was sufficient to inhibit the methylation activity of MLL1. Furthermore, the difference in IC_{50} between both isomers was more than 15-fold (Fig. 4). This experiment confirmed that the photoswitchable

peptidomimetic could be used to modulate MLL1 methyltransferase activity. In addition, the IC_{50} obtained for the *trans* isomer was of the same order of magnitude as for MM401 (IC_{50} is 0.32 μ M),⁴⁸ which is a potent and highly specific cyclic inhibitor for MLL1, and the current published short version of Win6mer.⁴⁴

Next, we decided to investigate whether these conformationally photoswitchable peptides could be used to regulate expression of MLL1 target genes and ultimately, inhibit leukemia cell proliferation. For efficient cell internalization, two new probes with the oligo-arginine appendages: R8 (16) and (R-Ahx-R)₄ (17),^{49,50} as well as their respective cell-penetrating peptide controls (18 and 19) were synthesized.

We initially studied the cellular uptake and cell viability of our compounds in murine MLL-AF9-transduced mouse bone marrow cells (Fig. 5). In line with our expectations, peptidomimetics 16

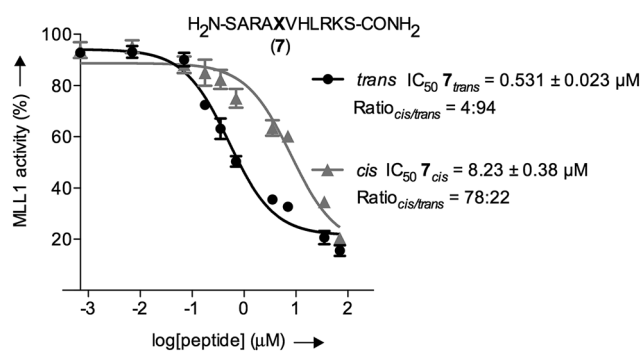


Fig. 4 *In vitro* functional HMT assay to evaluate actual potency of our conformational trap with both peptidomimetics 7 isomers; black straight line (●) represents the *trans* isomer, and grey straight line (▲) represents the *cis* isomer; X represents AMPB. The *cis/trans* ratios for τ_{trans} and τ_{cis} were 4 : 94 and 78 : 22, respectively.



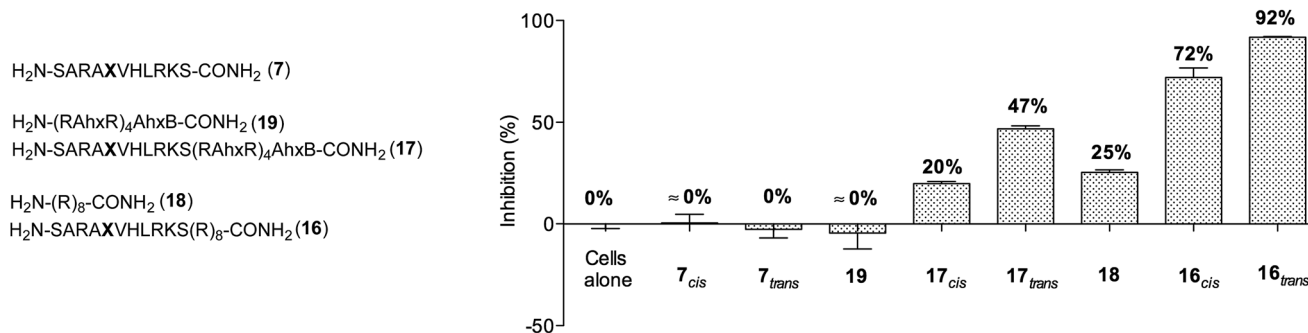


Fig. 5 Measurements of growth inhibition of murine MLL-AF9-transduced mouse bone marrow cells after 96 h of incubation with 5 μM concentration of different peptidomimetics. Ahx, B and X represent aminohexanoic acid, β -alanine and AMPB, respectively.

and 17 that contained peptide 7 within their sequence are the ones that inhibited the cell proliferation of leukemia. The *trans* isomers showed the highest inhibitory activity (92% for 16 and 47% for 17 at 5 μM concentration). The cell-penetrating peptide controls, as well as the unconjugated 7, had negligible effect in comparison. Furthermore, we also prepared two analogues to 16 and 17, which bear the parental WIN sequence 3 instead of the azo-containing peptide 7 (20 and 21, respectively) and compared their viability properties with our most active inhibitor (16_{trans}). As shown in Fig. 6, the difference between the WIN- and the azo-containing conjugates is obvious. Thus, while in the case of the *trans* isomer of peptidomimetic 16 we already determined less than 50% of viable cells at 2.5 μM concentration, with the WIN-containing conjugates 20 and 21 we measured 87% and 96% of viable cells at 50 μM concentration, respectively.

We then examined the potential of 16 and 17 to control externally the cellular response in proliferation assays. Thus, both isomers of both peptidomimetics were incubated at different concentrations for four days. After two days of incubation, the cells containing the *cis* isomers were irradiated at 430 nm for 90 seconds in a single isomerization cycle with a custom-made 6-well plate LED array. The possible cytotoxicity of different irradiation times, the thermal relaxation of the *cis*

isomer, and its stability against glutathione for four days were also evaluated, which showed a suitable *cis/trans* ratio whenever the *cis* isomer was kept in the dark during the time of the experiments (see ESI[†]).

As a control, we also included the *cis* isomers kept in the dark for four days without any irradiation. The dose-dependent growth inhibition assays indicated that in the case of both isomers the GI_{50} was better than the compound MM401, a previously reported MLL1 inhibitor (GI_{50} for 16_{trans} = 2.14 μM and GI_{50} for 17_{trans} = 4.98 μM versus GI_{50} for MM401 = 9.76 μM ; Fig. 7). More interestingly, we could observe modest, yet clear differences, up to 2-fold, in the cytotoxic behaviour of isomers in both compounds.

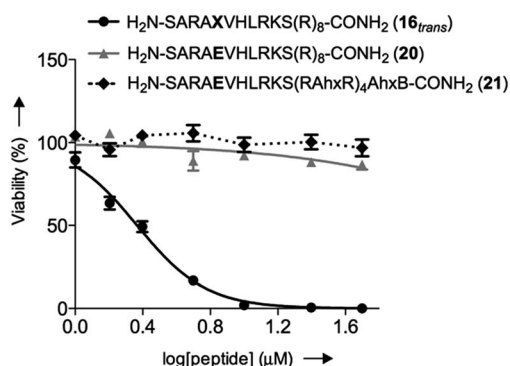


Fig. 6 Concentration-dependent viability curves of murine MLL-AF9-transduced mouse bone marrow cells after 96 h of incubation with different compounds: black straight line (●) represents the *trans* isomer 16, grey straight line (▲) represents 20 and dotted black line (◆) represents 21; Ahx, B and X represent aminohexanoic acid, β -alanine and AMPB, respectively.

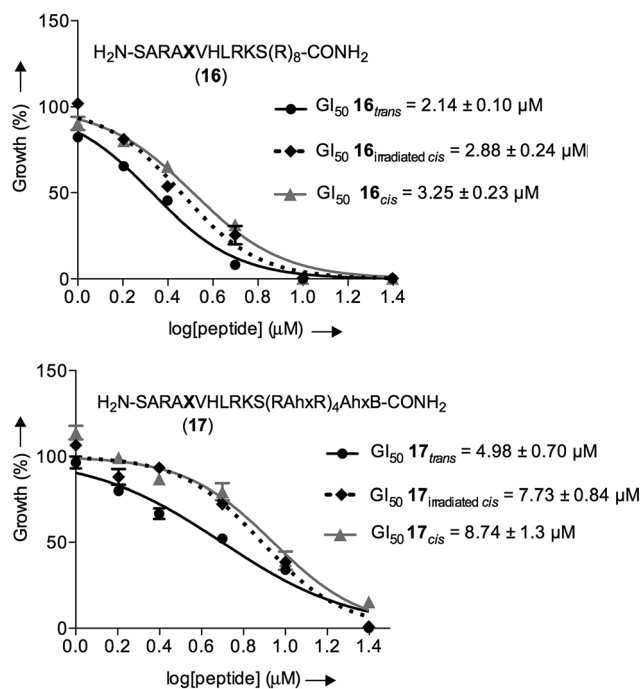


Fig. 7 Dose-dependent growth inhibition curves for the different isomers of 16 and 17 with murine MLL-AF9-transduced mouse bone marrow cells after 96 h of incubation; black straight line (●) represents the *trans* isomer, dotted black line (◆) represents the irradiated *cis* isomer and grey straight line represents the *cis* isomer (▲).



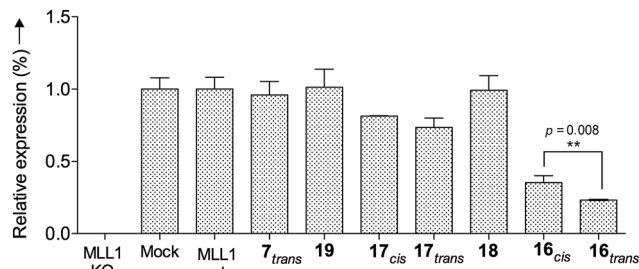


Fig. 8 Relative Deptor mRNA expression levels in murine MLL-AF9-transduced mouse bone marrow cells after 96 h of incubation at 10 μ M; they were analysed by RT-PCR and standardized to the endogenous GAPDH control. The significance is corroborated by the p -value.

These experiments confirm that our approach for targeting a key PPI with photoswitchable peptidomimetics has the potential to control enzymatic activity, leading to a distinct biological output. Our results show that it is possible to control cell proliferation through PPI photoswitches. Furthermore, our compound maintains good stability even after long periods of incubation, making potential therapeutic applications feasible.

Finally, we tested whether our probe acts as an optoepigenetic inhibitor and controls expression of MLL1 target genes. We used the Deptor gene as a reporter, given its importance in hematopoiesis and leukemogenesis,^{51,52} as well as significant regulation by MLL1 inhibition.⁵³ To this end, we examined Deptor expression by real-time PCR (RT-PCR). As expected, expression of Deptor was not affected by 7, 18 and 19, which also had no inhibitory effects on cell growth (Fig. 8). In contrast, treatment with peptidomimetic 16 led to significant downregulation of Deptor, comparable to MM401. Of note, a small but significant ($p = 0.008$) difference between *cis* and *trans* isomers was observed. Conjugate 17 also had a modest effect on Deptor expression, as compared to Mock (Fig. 8, $p < 0.001$). However, no statistically significant difference was found between the *trans* and *cis* isomers of 17, probably due to lower potency of this peptidomimetic, compared to 16. There is no statistically significant difference between *cis* isomer of 17 and the control.

Therefore, all these data support that our peptidomimetic 16 blocks WDR5 binding and inhibits the *in vitro* activity of MLL1 and cell proliferation of leukemia cells *via* reduction of MLL1 target gene expression.

Conclusions

In summary, we present results on a new strategy for external epigenetic control through targeting PPIs within a multi-protein complex with peptidomimetic photoswitches, which reversibly modulate the enzymatic activity of the histone methyltransferase MLL1, and consequently affect the transcription of the Deptor gene, and the growth of leukemia cells. Light is used to trigger the conformational switch of our peptidomimetics, what is ultimately responsible of the observed difference in activity. These discoveries highlight the importance of PPIs as druggable targets^{54–56} and demonstrate the possibility of using peptide

scaffolds as an efficient alternative to small-molecule inhibitors. We firmly expect that our results will encourage further developments in the use of peptidomimetics and small-molecule photoswitches in the field of epigenetics, increasing our understanding of chromatin phenomena at the molecular level.

Acknowledgements

We thank: Prof. Dr E. Meggers (Philipps-Universität Marburg) for access to his S1 laboratory, Ph.D. candidate Nathalie Nett (Philipps-Universität Marburg) for her assistance in the initial protein expression experiments, Prof. Dr M. E. Vázquez (CIQUS, University of Santiago de Compostela) for fruitful discussions during the elaboration of the manuscript, the staff at the X06SA beamline, Swiss Light Source, Paul Scherrer Institute for assistance, Tecan Group Ltd personnel aiding us with the plate reader. OxymaPure from Luxembourg Bio Technologies is gratefully acknowledged. Finally, we thank COST action CM1406 (Epigenetic Chemical Biology EPICHEMBO) for support.

References

- 1 R. Holliday, *Epigenetics*, 2006, **1**, 76–80.
- 2 Z. Su and J. M. Denu, *ACS Chem. Biol.*, 2016, **11**, 564–574.
- 3 B. D. Strahl and C. D. Allis, *Nature*, 2000, **403**, 41–45.
- 4 A. J. Ruthenburg, H. Li, D. J. Patel and C. D. Allis, *Nat. Rev. Mol. Cell Biol.*, 2007, **8**, 983–994.
- 5 J. S. Lee, E. Smith and A. Shilatifard, *Cell*, 2010, **142**, 682–685.
- 6 S. B. Hake, A. Xiao and C. D. Allis, *Br. J. Cancer*, 2004, **90**, 761–769.
- 7 M. G. Guenther, S. S. Levine, L. A. Boyer, R. Jaenisch and R. A. Young, *Cell*, 2007, **130**, 77–88.
- 8 B. Argiropoulos and R. K. Humphries, *Oncogene*, 2007, **26**, 6766–6776.
- 9 B. D. Yu, J. L. Hess, S. E. Horning, G. A. Brown and S. J. Korsmeyer, *Nature*, 1995, **378**, 505–508.
- 10 D. A. Lim, Y. C. Huang, T. Swigut, A. L. Mirick, J. M. Garcia-Verdugo, J. Wysocka, P. Ernst and A. Alvarez-Buylla, *Nature*, 2009, **458**, 529–533.
- 11 G. De Vita, P. Barba, N. Odartchenko, J. C. Givel, G. Freschi, G. Bucciarelli, M. C. Magli, E. Boncinelli and C. Cillo, *Eur. J. Cancer*, 1993, **29A**, 887–893.
- 12 D. Waltregny, Y. Alami, N. Clause, J. de Leval and V. Castronovo, *Prostate*, 2002, **50**, 162–169.
- 13 B. E. Li and P. Ernst, *Exp. Hematol.*, 2014, **42**, 995–1012.
- 14 M. S. Cosgrove and A. Patel, *FEBS J.*, 2010, **277**, 1832–1842.
- 15 C. W. Chen and S. A. Armstrong, *Exp. Hematol.*, 2015, **43**, 673–684.
- 16 R. L. Wright and A. T. Vaughan, *Crit. Rev. Oncol. Hematol.*, 2014, **91**, 283–291.
- 17 Y. Dou, T. A. Milne, A. J. Ruthenburg, S. Lee, J. W. Lee, G. L. Verdine, C. D. Allis and R. G. Roeder, *Nat. Struct. Mol. Biol.*, 2006, **13**, 713–719.
- 18 L. Fenno, O. Yizhar and K. Deisseroth, *Annu. Rev. Neurosci.*, 2011, **34**, 389–412.
- 19 K. Deisseroth, *Nat. Methods*, 2011, **8**, 26–29.
- 20 K. Deisseroth, *Sci. Am.*, 2010, **303**, 48–55.



- 21 D. Tischer and O. D. Weiner, *Nat. Rev. Mol. Cell Biol.*, 2014, **15**, 551–558.
- 22 M. Banghart, K. Borges, E. Isacoff, D. Trauner and R. H. Kramer, *Nat. Neurosci.*, 2004, **7**, 1381–1386.
- 23 M. M. Lerch, M. J. Hansen, G. M. van Dam, W. Szymanski and B. L. Feringa, *Angew. Chem., Int. Ed.*, 2016, **55**, 10978–10999.
- 24 J. Broichhagen, J. A. Frank and D. Trauner, *Acc. Chem. Res.*, 2015, **48**, 1947–1960.
- 25 W. A. Velema, W. Szymanski and B. L. Feringa, *J. Am. Chem. Soc.*, 2014, **136**, 2178–2191.
- 26 W. Szymanski, M. E. Ourailidou, W. A. Velema, F. J. Dekker and B. L. Feringa, *Chem.–Eur. J.*, 2015, **21**, 16517–16524.
- 27 S. A. Reis, B. Ghosh, J. A. Hendricks, D. M. Szantai-Kis, L. Tork, K. N. Ross, J. Lamb, W. Read-Button, B. X. Zheng, H. T. Wang, C. Salthouse, S. J. Haggarty and R. Mazitschek, *Nat. Chem. Biol.*, 2016, **12**, 317–323.
- 28 H. Karatas, E. C. Townsend, D. Bernard, Y. L. Dou and S. M. Wang, *J. Med. Chem.*, 2010, **53**, 5179–5185.
- 29 H. M. D. Bandara and S. C. Burdette, *Chem. Soc. Rev.*, 2012, **41**, 1809–1825.
- 30 M. X. Dong, A. Babalhavaeji, S. Samanta, A. A. Beharry and G. A. Woolley, *Acc. Chem. Res.*, 2015, **48**, 2662–2670.
- 31 R. J. Mart and R. K. Allemann, *Chem. Commun.*, 2016, **52**, 12262–12277.
- 32 J. Wachtveitl and A. Zumbusch, *ChemBioChem*, 2011, **12**, 1169–1170.
- 33 H. A. Wegner, *Angew. Chem., Int. Ed.*, 2012, **51**, 4787–4788.
- 34 G. A. Woolley, *Nat. Chem.*, 2012, **4**, 75–77.
- 35 F. Hamon, F. Djedaini-Pilard, F. Barbot and C. Len, *Tetrahedron*, 2010, **66**, 2538.
- 36 L. Ulysse and J. Chmielewska, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 2142–2146.
- 37 A. M. Caamaño, M. E. Vázquez, J. Martínez-Costas, L. Castedo and J. L. Mascareñas, *Angew. Chem., Int. Ed. Engl.*, 2000, **39**, 3104–3107.
- 38 A. A. Beharry, L. Wong, V. Tropepe and A. Woolley, *Angew. Chem., Int. Ed. Engl.*, 2011, **50**, 1325–1327.
- 39 J. Broichhagen, T. Podewin, H. Meyer-Berg, Y. von Ohlen, N. R. Johnston, B. J. Jones, S. R. Bloom, G. A. Rutter, A. Hoffmann-Röder, D. Hodson and D. Trauner, *Angew. Chem., Int. Ed. Engl.*, 2015, **54**, 15565–15569.
- 40 B. Priewisch and K. Ruck-Braun, *J. Org. Chem.*, 2005, **70**, 2350–2352.
- 41 N. Nishimura, T. Sueyoshi, H. Yamanaka, E. Imai, S. Yamamoto and S. Hasegawa, *Bull. Chem. Soc. Jpn.*, 1976, **49**, 1381–1387.
- 42 E. R. Talaty and J. C. Fargo, *J. Chem. Soc., Chem. Commun.*, 1967, 65–66.
- 43 Z. Nikolovska-Coleska, R. X. Wang, X. L. Fang, H. G. Pan, Y. Tomita, P. Li, P. P. Roller, K. Krajewski, N. G. Saito, J. A. Stuckey and S. M. Wang, *Anal. Biochem.*, 2004, **332**, 261–273.
- 44 N. L. Alicea-Velazquez, S. A. Shinsky, D. M. Loh, J. H. Lee, D. G. Skalnik and M. S. Cosgrove, *J. Biol. Chem.*, 2016, **291**, 22357–22372.
- 45 P. C. Weber, M. W. Pantoliano, D. M. Simons and F. R. Salemme, *J. Am. Chem. Soc.*, 1994, **116**, 2717–2724.
- 46 J. J. Song and R. E. Kingston, *J. Biol. Chem.*, 2008, **283**, 35258–35264.
- 47 A. Patel, V. Dharmarajan and M. S. Cosgrove, *J. Biol. Chem.*, 2008, **283**, 32158–32161.
- 48 F. Cao, E. C. Townsend, H. Karatas, J. Xu, L. Li, S. Lee, L. Liu, Y. Chen, P. Ouillette, J. Zhu, J. L. Hess, P. Atadja, M. Lei, Z. S. Qin, S. Malek, S. Wang and Y. Dou, *Mol. Cell*, 2014, **53**, 247–261.
- 49 S. Abes, H. M. Moulton, P. Clair, P. Prevot, D. S. Youngblood, R. P. Wu, P. L. Iversen and B. Lebleu, *J. Controlled Release*, 2006, **116**, 304–313.
- 50 E. A. Goun, T. H. Pillow, L. R. Jones, J. B. Rothbard and P. A. Wender, *ChemBioChem*, 2006, **7**, 1497–1515.
- 51 Y. Hu, H. Su, C. Liu, Z. Wang, L. Huang, Q. Wang, S. Liu, S. Chen, J. Zhou, P. Li, Z. Chen, H. Liu and G. Qing, *Oncogene*, 2017, **36**, 1038–1047.
- 52 T. R. Peterson, M. Laplante, C. C. Thoreen, Y. Sancak, S. A. Kang, W. M. Kuehl, N. S. Gray and D. M. Sabatini, *Cell*, 2009, **137**, 873–886.
- 53 J. Xu, L. Li, J. Xiong, A. denDekker, A. Ye, H. Karatas, L. Liu, H. Wang, Z. S. Qin, S. Wang and Y. Dou, *Cell Discovery*, 2016, **2**, 16008.
- 54 L. Nevola and E. Giralt, *Chem. Commun.*, 2015, **51**, 3302–3315.
- 55 M. Pelay-Gimeno, A. Glas, O. Koch and T. N. Grossmann, *Angew. Chem., Int. Ed. Engl.*, 2015, **54**, 8896–8927.
- 56 D. E. Scott, A. R. Bayly, C. Abell and J. Skidmore, *Nat. Rev. Drug Discovery*, 2016, **15**, 533–550.



Supporting Information

Controlled inhibition of methyltransferases using photoswitchable peptidomimetics: towards an epigenetic regulation of leukemia

*Lea Albert^a, Jing Xu^b, Ruiwei Wan^b, Vasundara Srinivasan^c, Yali Dou^b, Olalla Vázquez^{*a}*

Contribution from: ^aFachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Strasse 4, 35043 Marburg, Germany, ^bDepartment of Pathology, University of Michigan, Ann Arbor, Michigan 48109, United States and ^cLOEWE Zentrum für Synthetische Mikrobiologie SynMikro Philipps-Universität Marburg, Hans-Meerwein-Strasse 4, 35043 Marburg, Germany.

Table of Contents

Abbreviations	1
Materials and General Procedures.....	1
Solid Phase Peptide Synthesis	1
Synthesis and Characterization Data of the Peptides	4
Photoisomerization of the AMPB-containing Peptides	14
UV/Vis Characterization.....	15
RP-HPLC Characterization	16
Expression of WDR5Δ23	18
Synthesis of the Molecular Photoswitch	19
Fluorescence Polarization-based Assays.....	21
Saturation Binding Experiments	22
Corroboration of Absence of Quenching Artefacts	22
Competitive Fluorescence Polarization-based Binding Assays	24
Crystal Structure of WDR5Δ23-Peptide 7 Complex.....	40
Evaluation of Stability against GSH-reduction of Peptides	49
In Vitro Methyl Transferase Assay with MLL1 Core Complex	50
Protein Expression	50
In Vitro Histone Methyltransferase (HMT) Assay	50
Proliferation Assay	51
Real Time Quantitative PCR Analysis of Deptor Gene	52
References	53

Abbreviations

\mathcal{E}	extinction coefficient
λ	wavelength
ν	frequency
A	anisotropy
A_b	anisotropy of bound tracer
A_f	anisotropy of free tracer
Ahx	aminohexanoic acid
Alloc	alkyloxycarbonyl
AMPB	(4-aminomethyl)phenyl-azobenzoic acid
Azo	azobenzene
Boc	<i>tert</i> -butylcarbonyl
c	concentration
calcd.	calculated
DIC	diisopropylcarbodiimide
DIPEA	diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
EC ₅₀	half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
ESI	electrospray ionization
eq.	equivalents
FAM	fluorescein
FBS	fetal bovine serum
Fmoc	fluorenylmethyloxycarbonyl chloride
FP	fluorescence polarization
HPLC	high-performance liquid chromatography
HRMS	high resolution mass spectrometry
[I] ₅₀	concentration of free inhibitor at 50% inhibition
IC ₅₀	half inhibitory concentration
IMDM	Iscoe's Modified Dulbecco's Media
IPTG	isopropyl β -D-1-thiogalactopyranoside
irr.	irradiation
K_d	dissociation constant
K_i	inhibition constant
[L] ₅₀	concentration of free labelled ligand at 50% inhibition
L_F	concentration of free ligand
L_T	total ligand concentration
mAU	milli absorbance units
MALDI	matrix-assisted laser desorption ionization
mP	milli polarization value

MPD	2-Methyl-2,4-pentanediol
NMP	<i>N</i> -methyl-2-pyrrolidon
Ni-NTA	nickel nitrilotriacetid acid
$[P]_0$	concentration of the free protein at 0% inhibition
PEG	polyethylene glycol
R_f	retention factor
R_F	concentration of free receptor
R_T	total receptor concentration
RT-PCR	real time PCR
RP	reversed phase
rpm	revolutions per minute
rt	room temperature
SDS	sodium dodecyl sulphate polyacrylamide
t_0	time zero
t_R	retention time
TFA	trifluoroacetic acid
TIS	triisopropylsilane
Tris	Tris(hydroxymethyl)aminomethane
WDR5	WD-repeat protein-5

Materials and General Procedures

All commercially purchased reagents were used without further purification as delivered from the corresponding company. The respective reagents were purchased from the following companies: DMSO biograde, NaCl biograde, Tris, 2-mercaptoethanol and DIPEA from Carl Roth (Germany), 6-aminohexanoic acid and 4-aminobenzoic acid from Alfa Aesar (Germany), Oxone from Merck (Germany), picosulfonic acid from TCI (Germany), Fmoc-OSu from Fluorochem (UK), OxymaPure from Luxembourg Bio Technologies (Israel), DIC, glycerol and benzamidine from Acros (USA), KCl biograde, Triton and 4-aminobenzylamine from Sigma Aldrich (Germany). Fmoc-protected amino acids as well as the resin TentaGel S RAM were purchased from Iris Biotech (Germany). The rink amide MBHA resin was purchased from Fluorochem (UK). DMF was employed as peptide grade (Iris Biotech, Germany) and MeCN as HPLC grade (VWR, France). Water was purified with a Milli-Q Ultra Pure Water Purification System (TKA, Germany). All cell-based experiments were performed in compliance with University of Michigan guidelines and approved by the University Committee on the Use and Care of Animals (UCUCA).

Solid Phase Peptide Synthesis

Peptides were either synthesized manually in 2 mL polypropylen reactors with plunger and frit, (pore size 25 μm , Multi Syn Tech GmbH, Germany) or by an automated peptide synthesizer (Advanced ChemTech Apex 396). TentaGel S RAM resin (0.25 mmol/g) or a rink amide resin, where the loading of the first amino acid was determined, were used for synthesis.

Determination of the resin loading: The Fmoc group of a weighted amount of resin (~5 mg), previously loaded with the first amino acid [for 80-160 mg of rink amide resin (0.3-0.8 g/mmol): 2.00 eq. Fmoc-amino acid, 2.00 eq. OxymaPure (0.5 M in DMF), 2.00 eq. DIC were added to the resin for 45 min with a final concentration of 0.4 M of the amino acid], was deprotected by adding twice a solution of 20% piperidine in DMF for 5 min (500 μL). The filtrates were collected and the absorbance was measured at 300 nm (ϵ_{300} (Fmoc) = 7800 L mol⁻¹ cm⁻¹)^[1] on a Beckman-Coulter DU800 spectrophotometer and the loading was calculated. The remaining resin was swollen in DMF for 30 min and the non-reacted amino groups were capped following the general protocol of manual synthesis explained below.

Manual Solid Phase Protocol for Peptide Synthesis (protocol A)

The amounts of reagents of the following synthesis protocol correspond to 20-40 μmol scale. For shaking an Edmund Bühler Swip shaker was used.

Swelling: 160 mg of resin (40 μmol , loading = 0.25 mmol/g) were swelling in 1.0 mL DMF for 30 min.

Deprotection of the temporal Fmoc group: Piperidine (500 μL , 20% in DMF) was added to the resin and was shaken for 5 min. This step was repeated and the resin was filtered off and washed with DMF (5×1.5 mL), CH_2Cl_2 (5×1.5 mL), DMF (5×1.5 mL).

Coupling of amino acids: The Fmoc-amino acid (4.00 eq.) was dissolved in OxymaPure (0.5 M in DMF, except for the *N*-Fmoc-(4-aminomethyl)phenylazobenzoic acid (AMPB) coupling where OxymaPure was dissolved in NMP, 0.5 M, 4.00 eq.) and DIC (4.00 eq.) was added. The resulting solution was activated for 3 min and subsequently added to the resin. This suspension was shaken for 45 min. The resin was filtered off and washed with DMF (5×1.5 mL) and CH_2Cl_2 (5×1.5 mL) and DMF (5×1.5 mL). Monitoring of the coupling completion was confirmed by the TNBS-test^[2] on few resin-beads.

Capping: A solution of lutidine/ Ac_2O /DMF 6:5:89 (1.0 mL) was added to the resin and was shaken for 5 min. The resin was filtered off and washed with DMF (5×1.5 mL), CH_2Cl_2 (5×1.5 mL) and DMF (5×1.5 mL).

Automated Solid Phase Protocol for Peptide Synthesis (protocol B)

The amounts of reagents of the following synthesis protocol correspond to 20 μmol scale. The mixing was performed with a frequency of 300 rpm.

Swelling: 80 mg of resin (20 μmol , loading = 0.25 mmol/g) were swelling in 600 μL DMF for 30 min.

Deprotection of the temporal Fmoc group: Piperidine (550 μL , 20% in DMF) was added to the resin and was shaken for 5 min. This step was repeated and the resin was filtered off and washed with DMF (2×1.75 mL, and 2×1.00 mL with mixing for 1 min).

Coupling of amino acids: 215 μL of the respective Fmoc-amino acid (0.5 M in DMF), 215 μL of OxymaPure (0.5 M in DMF) and 121 μL of DIC (0.87 M in DMF) were added to the resin.

This suspension was mixed for 25 min. The resin was filtered and the coupling was repeated. Afterwards, the resin was washed with DMF (2×1.75 mL, and 2×1.00 mL with mixing for 1 min).

Capping: A solution of lutidine/ Ac_2O /DMF 6:5:89 (550 μL) was added to the resin and was mixed for 5 min. The resin was filtered off and washed with DMF (2×1.75 mL, and 2×1.00 mL with mixing for 1 min).

After the last coupling, the resin was washed with DMF (2×2.2 mL, and 2×1.00 mL with mixing for 1 min) and CH_2Cl_2 (2×2.2 mL, and 3×1.75 mL with mixing for 1 min).

Cleavage deprotection step: This step is the same for both synthesis protocols. All the used cleavage cocktails were TFA-based, however, depending on the specific peptide they differed in the concentration of TFA, type of scavengers and reaction times. The used cleavage cocktails and reaction times are listed below.

- Cleavage cocktail A: 90% TFA, 5% DCM, 2.5% H_2O , 2.5% TIS (3 h)
- Cleavage cocktail B: 94% TFA, 5% H_2O , 1% TIS (2 h)
- Cleavage cocktail C: 95% TFA, 5% H_2O (2 h)
- Cleavage cocktail D: 94% TFA, 5% H_2O , 0.5% *m*-Cresol, 0.5% H-Cys(OMe) x HCl (2 h)

The dried resin was treated with the corresponding mixture and was shaken for the specific time. After the defined time, the resin was filtered off and the filtrate was added to dry ice-cold Et_2O (1 mL of Et_2O for 100 μL cleavage cocktail). After 10 min, the precipitated peptide was centrifuged, the supernatant was discarded and the peptide pellet was dissolved in MilliQ in order to be purified.

Purification: The probes were purified at 25 °C by preparative or semipreparative HPLC, performed on a PLC 2020 Personal Purification System (GILSON) with a preparative Nucleodur C18 HTec-column (5 μm , 250×16 mm; Macherey Nagel) and a flow rate of 10 mL/min or with a semipreparative C18 3001 column (Phenomenex) with a flow rate of 6 mL/min. Detection of the signals was achieved with a UV detector at wavelength of 220 nm. The eluents were MilliQ H_2O (A) and MeCN (B) with addition of 0.1% TFA. The collected fractions, containing the desired peptides, were lyophilized and stored at -20 °C.

Characterization: The freeze-dried products were identified via analytical HPLC-MS on an Agilent 1200 Series HPLC-system (Agilent Technologies). If not stated differently, for all analytical HPLC runs an EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) using the linear gradient 5% to 40% B in 30 min with a flow rate of 0.20 mL/min was used and detection was monitored at 220 nm. Electrospray Ionization Mass Spectrometry (ESI-MS) was performed by direct injection on a *LTQ FT Finnigan* (Thermo Fischer Scientific) spectrometer. MALDI-TOF/MS spectra were recorded on an *UltraFLEX* (Bruker).

Synthesis and Characterization Data of the Peptides

Unless otherwise noted, the synthesis of the peptides was performed manually following the protocol A described above. If not noted differently, the linear gradient 5% to 40% B was used for analytical HPLC. The purity of the peptides was calculated from the integrated peak areas of the HPLC chromatograms, and is given as % for each peptide.

3: H₂N-Ser-Ala-Arg-Ala-Glu-Val-His-Leu-Arg-Lys-Ser-CONH₂; this peptide was synthesized according to protocol B (6 x 20 μmol, where the resins were combined after synthesis). For the final cleavage step, 551 mg of resin (loading = 0.25 mmol/g; peptide: 131.4 mg, 104.9 μmol) were treated with 14.0 mL cleavage cocktail A. After purification the product (62.7 mg, 34.4 μmol, 33%) was obtained as a white solid. $t_R = 15.32$ min. Purity $\geq 99\%$. Formula: C₅₂H₉₃N₂₁O₁₅. MALDI/TOF: [M+H]⁺ calcd.: 1253.4; found: 1253.4. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 1252.7233; found: 1252.7245.

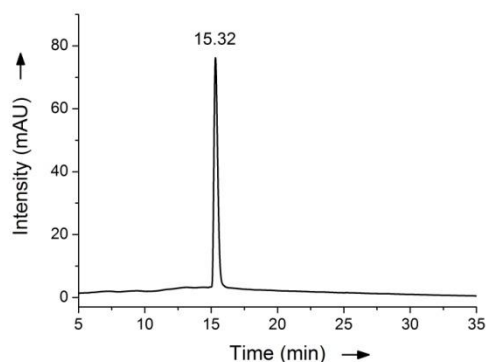


Figure S1. HPLC chromatogram of purified peptide 3.

4: H₂N-Ser-AMPB-Ala-Arg-Ala-Glu-Val-His-Leu-Arg-Lys-Ser-CONH₂; for the final cleavage step, 207 mg of resin (peptide: 56.2 mg, 37.7 μmol) were treated with 10.0 mL cleavage cocktail C. After purification, the product (14.8 mg, 7.19 μmol, 19%) was obtained as an orange solid. $t_R = 19.98$ min (*cis*), 21.69 min (*trans*). Purity ≥ 99%. Formula: C₆₆H₁₀₄N₂₄O₁₆. MALDI/TOF: [M+H]⁺ calcd.: 1490.6; found: 1490.6. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 745.4104; found: 745.4104.

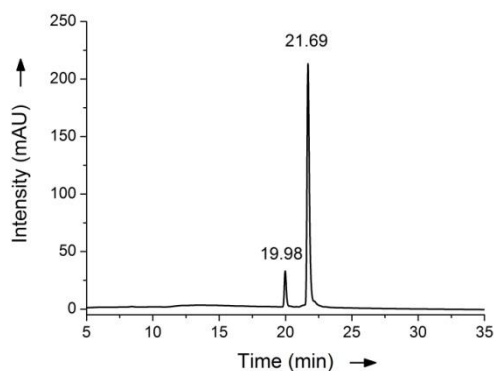


Figure S2. HPLC chromatogram of purified peptide 4.

5: H₂N-Ser-AMPB-Arg-Ala-Glu-Val-His-Leu-Arg-Lys-Ser-CONH₂; for the final cleavage step, 191 mg of resin (peptide: 50.1 mg, 35.3 μmol) were treated with 10.0 mL cleavage cocktail C. After purification, the product (14.1 mg, 7.09 μmol, 20%) was obtained as an orange solid. $t_R = 19.89$ min (*cis*), 21.87 min (*trans*). Purity ≥ 99%. Formula: C₆₃H₉₉N₂₃O₁₅. MALDI/TOF: [M+H]⁺ calcd.:1419.6; found:1419.6. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 709.8918; found: 709.8921.

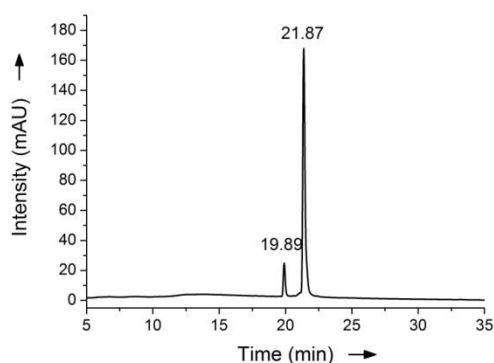


Figure S3. HPLC chromatogram of purified peptide 5.

6: H₂N-Ser-Ala-Arg-AMPB-Glu-Val-His-Leu-Arg-Lys-Ser-CONH₂; for the final cleavage step, 190 mg of resin (peptide: 49.7 mg, 35.0 μmol) were treated with 10.0 mL cleavage cocktail C. After purification, the product (14.4 mg, 7.24 μmol, 21%) was obtained as an orange solid. $t_R = 20.03$ min (*cis*), 22.26 min (*trans*). Purity ≥ 99%. Formula: C₆₃H₉₉N₂₃O₁₅. MALDI/TOF: [M+H]⁺ calcd.: 1419.6 found: 1419.6. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 709.8918; found: 709.8921.

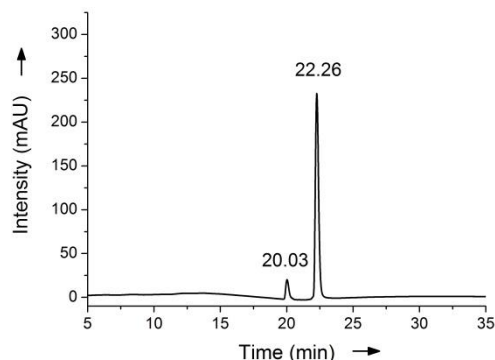


Figure S4. HPLC chromatogram of purified peptide 6.

7: H₂N-Ser-Ala-Arg-Ala-AMPB-Val-His-Leu-Arg-Lys-Ser-CONH₂; for the final cleavage step, the resin (scale = 40 μmol, loading = 0.25 mmol/g) was treated with 6.00 mL cleavage cocktail C. After purification, the product (43.0 mg, 22.3 μmol, 56%) was obtained as an orange solid. $t_R = 20.49$ min (*cis*), 22.96 min (*trans*). Purity ≥ 99%. Formula: C₆₁H₉₇N₂₃O₁₃. MALDI/TOF: [M+H]⁺ calcd.: 1361.7; found: 1361.7. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 1361.7736; found: 1360.7734.

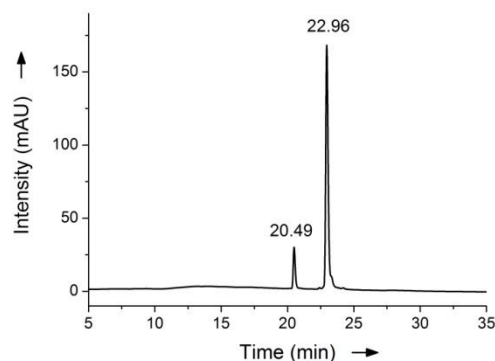


Figure S5. HPLC chromatogram of purified peptide 7.

8: H₂N-Ser-Ala-Arg-Ala-Glu-AMPB-His-Leu-Arg-Lys-Ser-CONH₂; for the final cleavage step, 62.0 mg of resin (loading = 0.25 mmol/g; peptide: 15.9 mg, 11.49 μmol) were treated with 7.00 mL cleavage cocktail B. After purification, the product (1.4 mg, 1.01 μmol, 9%) was obtained as an orange solid. t_R = 19.21 min (*cis*), 21.41 min (*trans*). Purity ≥ 99%. Formula: C₆₁H₉₅N₂₃O₁₅. MALDI/TOF: [M+H]⁺ calcd.: 1391.7; found: 1391.7. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 1390.7451; found: 1390.7491.

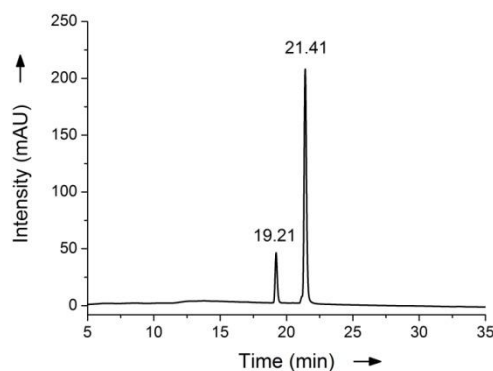


Figure S6. HPLC chromatogram of purified peptide **8**.

9: H₂N-Ser-Ala-Arg-Ala-Glu-Val-AMPB-Leu-Arg-Lys-Ser-CONH₂; for the final cleavage step, 109.0 mg of resin (loading = 0.25 mmol/g; peptide: 27.54 mg, 20.36 μmol) were treated with 8.00 mL cleavage cocktail D. After purification, the product (7.0 mg, 5.18 μmol, 25%) was obtained as an orange solid. t_R = 22.38 min (*cis*), 25.18 min (*trans*). Purity ≥ 99%. Formula: C₆₀H₉₇N₂₁O₁₅. MALDI/TOF: [M+H]⁺ calcd.: 1352.7; found: 1352.7. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 1353.7573; found: 1353.7567.

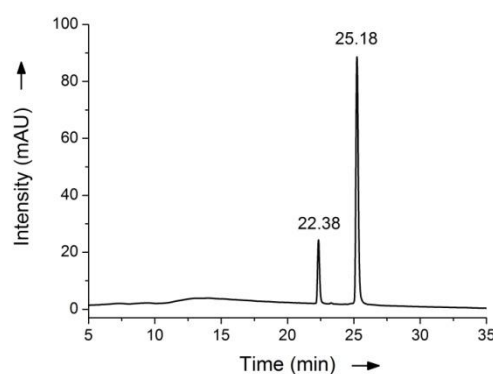


Figure S7. HPLC chromatogram of purified peptide **9**.

10: H₂N-Ser-Ala-Arg-Ala-Glu-Val-His-AMPB-Arg-Lys-Ser-CONH₂; for the final cleavage step, 32.0 mg of resin (loading = 0.25 mmol/g; peptide: 8.19 mg, 5.95 μmol) were treated with 3.80 mL cleavage cocktail B. After purification, the product (1.8 mg, 1.31 μmol, 22%) was obtained as an orange solid. t_R = 17.44 min (*cis*), 19.67 min (*trans*). Purity = 98%. Formula: C₆₀H₉₃N₂₃O₁₅. MALDI/TOF: [M+H]⁺ calcd.: 1377.7; found: 1377.7. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 1377.7321; found: 1377.7337.

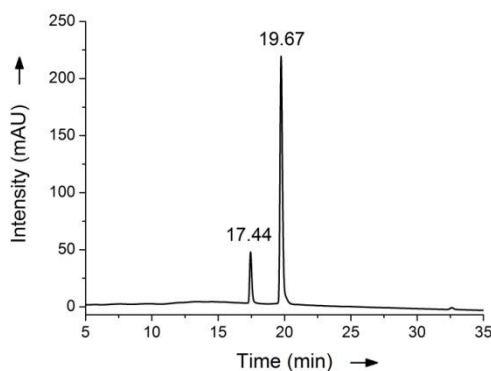


Figure S8. HPLC chromatogram of purified peptide **10**.

11: H₂N-Ser-Ala-Arg-Ala-Glu-Val-His-Leu-AMPB-Lys-Ser-CONH₂; for the final cleavage step, 71.5 mg of resin (rink amide, loading = 0.614 mmol/g, peptide: 32.2 mg, 24.10 μmol), were treated with 8.30 mL cleavage cocktail C. After purification, the product (9.4 mg, 5.05 μmol, 29%) was obtained as an orange solid. t_R = 22.13 min (*cis*), 23.74 min (*trans*). Purity = 99%. Formula: C₆₀H₉₂N₂₀O₁₅. MALDI/TOF: [M+H]⁺ calcd.: 1334.5; found: 1334.5. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 1334.7152; found: 1334.7157.

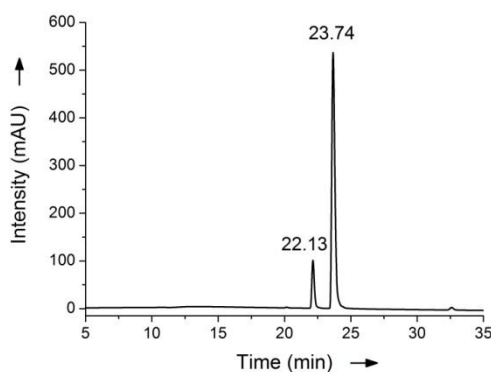


Figure S9. HPLC chromatogram of purified peptide **11**.

12: H₂N-Ser-Ala-Arg-Ala-Glu-Val-His-Leu-Arg-AMPB-Ser-CONH₂; for the final cleavage step, 26.0 mg of resin (link amide, loading = 0.614 mmol/g, peptide: 11.8 mg, 8.69 μmol), were treated with 3.1 mL cleavage cocktail C. After purification, the product (7.7 mg, 5.66 μmol, 65%) was obtained as an orange solid. t_R = 21.58 min (*cis*), 23.38 min (*trans*). Purity ≥ 99%. Formula: C₆₀H₉₂N₂₂O₁₅. MALDI/TOF: [M+H]⁺ calcd.: 1361.5; found: 1361.5. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 1362.7212; found: 1362.7289.

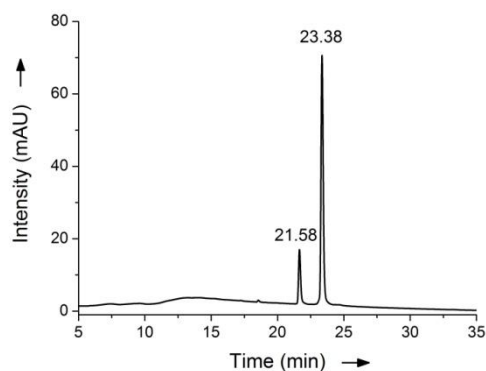


Figure S10. HPLC chromatogram of purified peptide **12**.

13: H₂N-Ser-Ala-Arg-Ala-AMPB-His-Leu-Arg-Lys-Ser-CONH₂; for the final cleavage step, the resin (scale = 20 μmol, loading = 0.25 mmol/g) was treated with 4.00 mL cleavage cocktail C. After purification, the product (15.7 mg, 8.57 μmol, 43%) was obtained as an orange solid. t_R = 18.57 min (*cis*), 20.64 min (*trans*). Purity = 98%. Formula: C₅₆H₈₈N₂₂O₁₂. MALDI/TOF: [M+H]⁺ calcd.: 1262.4; found: 1262.7. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 631.3549; found: 631.3546.

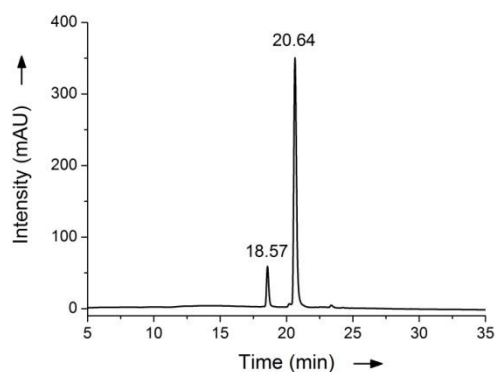


Figure S11. HPLC chromatogram of purified peptide **13**.

14: Ac-Ala-Arg-Ala-AMPB-Val-His-Leu-Arg-Lys-Ser-CONH₂; for the final cleavage step, 64.6 mg of resin (loading = 0.25 mmol/g; peptide: 16.0 mg, 12.0 μmol) were treated with 3.50 mL cleavage cocktail D. After purification, the product (2.70 mg, 1.52 μmol, 13%) was obtained as an orange solid. $t_R = 22.79$ min (*cis*), 25.21 min (*trans*) Purity ≥ 99%. Formula: C₆₀H₉₅N₂₂O₁₂. MALDI/TOF: [M+H]⁺ calcd.: 1316.5; found: 1316.6. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 658.3784; found: 658.3793.

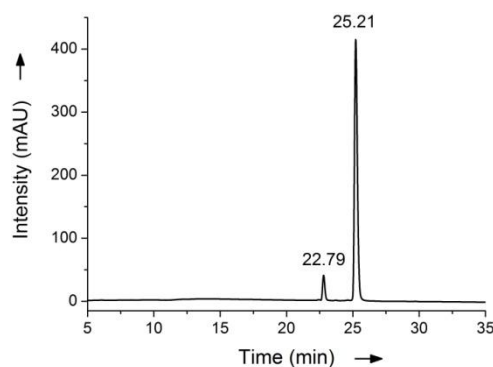


Figure S12. HPLC chromatogram of purified peptide **14**.

15: Ac-Ala-Arg-Thr-AMPB-Val-Tyr-CONH₂; for the final cleavage step, the resin (scale = 20 μmol, loading = 0.25 mmol/g) was treated with 4.00 mL cleavage cocktail C. After purification, the product (11.0 mg, 10.99 μmol, 55%) was obtained as an orange solid. $t_R = 26.14$ min (*cis*), 29.16 min (*trans*). Purity = 99%. Formula: C₄₃H₅₈N₁₂O₉. MALDI/TOF: [M+H]⁺ calcd.: 887.9; found: 887.9. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 887.4522; found: 887.4519.

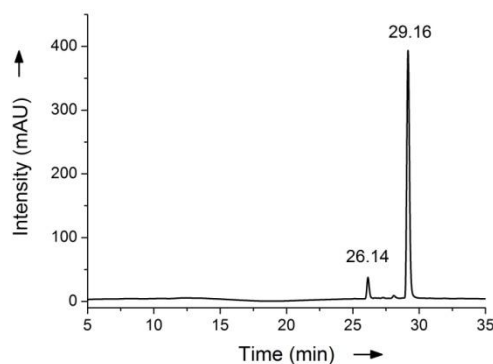


Figure S13. HPLC chromatogram of purified peptide **15**.

16: H₂N-Ser-Ala-Arg-Ala-AMPB-Val-His-Leu-Arg-Lys-Ser-(Arg)₈-CONH₂; this peptide was synthesized manually according to protocol A, but in a higher scale (0.1 mmol). For the final cleavage step, 821 mg of resin were treated with 20 mL cleavage cocktail B. After purification, the product (86.5 mg, 0.021 mmol, 21%) was obtained as an orange solid. $t_R = 20.25$ min (*cis*), 22.18 min (*trans*). Purity $\geq 99\%$. Formula: C₁₀₉H₁₉₃N₅₅O₂₁. MALDI/TOF: [M+H]⁺ calcd.: 2611.3; found: 2611.3. HRMS-ESI⁺ (m/z): [M+5H]⁵⁺ calcd.: 522.9223; found: 522.9231.

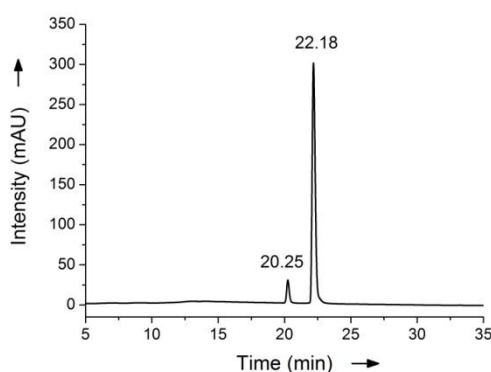


Figure S14. HPLC chromatogram of purified peptide **16**.

17: H₂N-Ser-Ala-Arg-Ala-AMPB-Val-His-Leu-Arg-Lys-Ser-(Arg-Ahx-Arg)₄-Ahx-βAla-CONH₂; this peptide was synthesized according to protocol A, but in a higher scale (0.1 mmol). For the final cleavage step, 748 mg of resin were treated with 20 mL cleavage cocktail B. After purification, the product (29.1 mg, 6.15 μmol, 6%) was obtained as an orange solid. $t_R = 20.85$ min (*cis*), 22.34 min (*trans*). Purity $\geq 99\%$. Formula: C₁₄₂H₂₅₃N₆₁O₂₇. MALDI/TOF: [M+H]⁺ calcd.: 3247.7; found: 3247.7. HRMS-ESI⁺ (m/z): [M+5H]⁵⁺ calcd.: 650.2138; found: 650.2136.

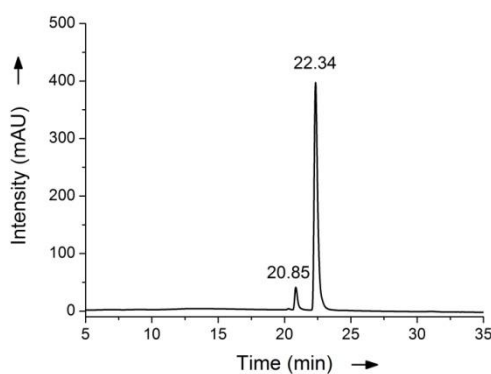


Figure S15. HPLC chromatogram of purified peptide **17**.

18: H₂N-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-CONH₂; for the final cleavage step, 216 mg of resin (peptide: 52 mg, 40 μmol) were treated with 11 mL cleavage cocktail A. After purification, the product (24.2 mg, 11.0 μmol, 27%) was obtained as a white solid. $t_R = 16.33$ min. Purity $\geq 99\%$. Formula: C₄₈H₉₉N₃₃O₈. MALDI/TOF: [M+H]⁺ calcd.: 1267.5; found: 1267.5. HRMS-ESI⁺ (m/z): [M+3H]³⁺ calcd.: 422.9524; found: 422.9509.

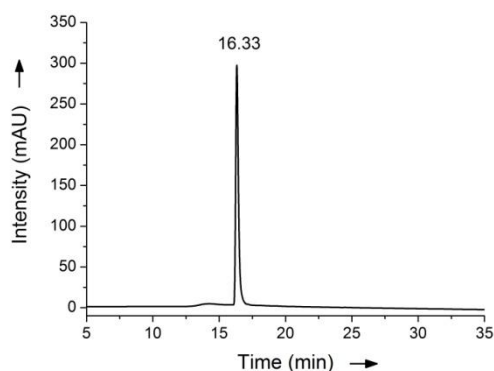


Figure S16. HPLC chromatogram of purified peptide **18** (gradient: 0-30% MeCN).

19: H₂N-(Arg-Ahx-Arg)₄-Ahx-βAla-CONH₂; for the final cleavage step, 105 mg of resin (peptide: 33.8 mg, 17.8 μmol) were treated with 5.3 mL cleavage cocktail A. After purification, the product (18.4 mg, 8.03 μmol, 45%) was obtained as a white solid. $t_R = 21.98$ min. Purity $\geq 99\%$. Formula: C₈₁H₁₅₉N₃₉O₁₄. MALDI/TOF: [M+H]⁺ calcd.: 1903.3; found: 1903.3. HRMS-ESI⁺ (m/z): [M+4H]⁴⁺ calcd.: 476.8311; found: 476.8318.

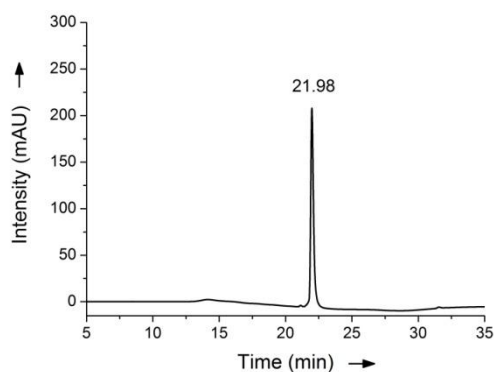


Figure S17. HPLC chromatogram of purified peptide **19** (gradient: 0-30% MeCN).

20: $\text{H}_2\text{N-Ser-Ala-Arg-Ala-Glu-Val-His-Leu-Arg-Lys-Ser-R}_8\text{-CONH}_2$; this peptide was synthesized according to protocol B (4 x 20 μmol , where the resins were combined after synthesis). For the final cleavage step, the resin (loading = 0.25 mmol/g) was treated with 16.0 mL cleavage cocktail A. After purification, the product (42.7 mg, 10.7 μmol , 13%) was obtained as a white solid. $t_R = 16.88$. Purity $\geq 99\%$. Formula: $\text{C}_{100}\text{H}_{193}\text{N}_{53}\text{O}_{23}$. MALDI/TOF: $[\text{M}+\text{H}]^+$ calcd.: 2502.9; found: 2502.4. HRMS-ESI⁺ (m/z): $[\text{M}+4\text{H}]^{4+}$ calcd.: 626.3891; found: 626.3881.

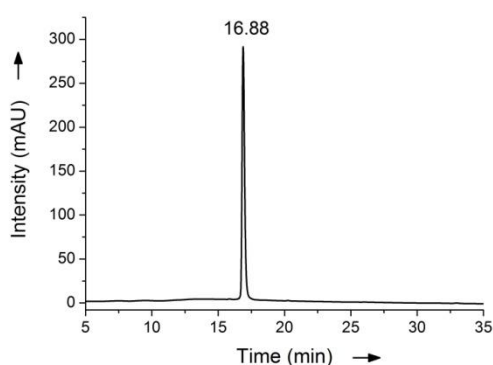


Figure S18. HPLC chromatogram of purified peptide **20**.

21: $\text{H}_2\text{N-Ser-Ala-Arg-Ala-Glu-Val-His-Leu-Arg-Lys-Ser-(Arg-Ahx-Arg)}_4\text{-Ahx-}\beta\text{Ala-CONH}_2$; this peptide was synthesized according to protocol B (3 x 20 μmol , where the resins were combined after synthesis). For the final cleavage step, the resin (loading = 0.25 mmol/g) was treated with 12.0 mL cleavage cocktail A. After purification, the product (6 mg, 1.3 μmol , 2%) was obtained as a white solid. $t_R = 18.39$. Purity $\geq 99\%$. Formula: $\text{C}_{133}\text{H}_{249}\text{N}_{59}\text{O}_{29}$. MALDI/TOF: $[\text{M}+\text{H}]^+$ calcd.: 3139.8; found: 3139.5. HRMS-ESI⁺ (m/z): $[\text{M}+5\text{H}]^{5+}$ calcd.: 628.6043; found: 628.6033.

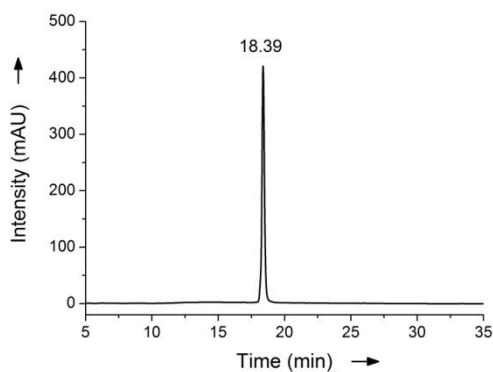


Figure S19. HPLC chromatogram of purified peptide **21**.

22: Ac-Ala-Arg-Thr-Glu-Val-His-Leu-Arg-Lys-Ser-Ahx-Ahx-Lys(FAM)-CONH₂; the fluorescently tagged peptide was synthesized in a 60 μmol scale using a rink amide resin (loading = 0.639 mmol/g). The *N*-terminus of the peptide was acetylated following the capping procedure explained before. The *C*-terminal lysine was introduced as Fmoc-Lys(Alloc)-OH, whose Alloc group was selectively removed in presence of standard conditions.^[3,4] Pd(PPh₃)₄ (1.00 eq.), morpholine (190 eq.) and a solution of DCM with 2% water (final: 1.5 mL) were added to the resin under nitrogen atmosphere and the mixture was shaken for 5 h. The resin was filtered off and washed with DMF (3 x 2.0 mL), a solution of sodiumdiethyldithiocarbamate (29.2 mM, 2 x 2.0 mL), DMF (3 x 2.0 mL) and CH₂Cl₂ (2 x 2.0 mL). Afterwards, the 5-carboxyfluorescein (FAM) was coupled to the selectively deprotected lysine side chain according to the coupling conditions of protocol A (20 μmol). For the final cleavage, 62.0 mg of resin (peptide: 34.3 mg, 17.64 μmol) were treated with 2.5 mL cleavage cocktail A. After purification, the product (19.1 mg, 9.79 μmol, 56%) was obtained as yellow solid. $t_R = 25.79$ min. Purity = 99%. Formula: C₉₁H₁₃₆N₂₄O₂₄. MALDI/TOF: [M+H]⁺ calcd.: 1951.2; found: 1951.2. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 975.5205; found: 975.5205.

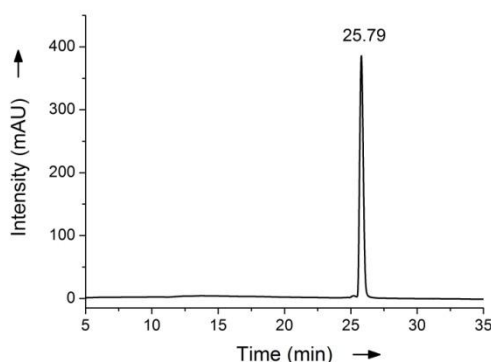


Figure S20. HPLC chromatogram of purified peptide **22**.

Photoisomerization of the AMPB-containing Peptides

A Benda (type: NU-4 KL) lamp and a Luminea spot LED E27 lamp were used for irradiation at 366 and 430 nm, respectively, in all experiments except for the proliferation assays and the RT-PCR experiments, where custom made 96- and 6-well plate LED arrays were used (LEDs: DLE-038-045 from Everlight Electronics Co. and H2A1-H435 from Roithner LaserTechnik). The photostationary states were studied by UV/Vis-spectroscopy and by analytical RP-HPLC.

UV/Vis Characterization

For the characterization via UV/Vis measurements of peptide **7**, a 500 μM solution of **7** in 0.1 M phosphate buffer (pH = 6.5) was irradiated at the respective wavelength. To obtain the *cis* isomer, the sample was irradiated at 366 nm for 1 h. For *cis* \rightarrow *trans* back isomerization, the same solution was irradiated at 430 nm for 30 min afterwards. The absorbance spectra are shown in Figure S21.

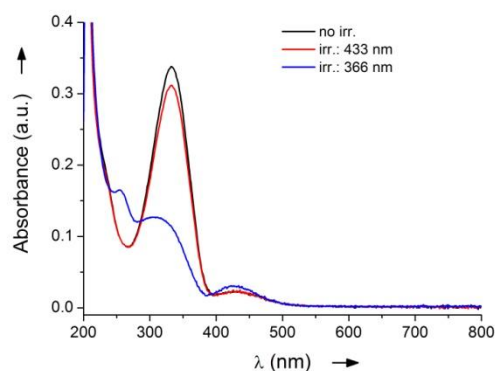


Figure S21. UV/Vis spectra of the: non-irradiated sample (black spectrum), 366 nm irradiated sample (blue spectrum) and the 430 nm irradiated sample (red spectrum).

As shown in Figure S21, the characteristic maximum of absorbance at $\lambda_{\text{max}} = 333$ nm for the π - π^* transition of the *trans* isomer of the azobenzene is clearly visible in the UV/Vis-spectrum of the non-irradiated sample (black spectrum). After irradiation at 366 nm (photostationary state of the *cis* isomer), this band decreases drastically due to the loss of the *trans* isomer and the generation of the *cis* isomer, which also becomes obvious in the rise of the absorbance bands at $\lambda_{\text{max}} = 255$ nm for the π - π^* and 425 nm for the n - π^* transition (blue spectrum). After irradiation at 430 nm, these two bands decrease and the band at 333 nm increases again. These spectral changes are ascribed to the *cis* \rightarrow *trans* back isomerization, what proves the fully reversibility of the photoisomerisation process with two isosbestic points at 287 and 394 nm. These results are consistent to the results published by MORODOER et al., who incorporated the AMPB into a different peptide.^[5,6]

RP-HPLC Characterization

In the case of AMPB it is known that the *trans* isomer is the thermodynamically more stable one.^[7] The maximum amount of *trans* isomer was directly obtained after purification for the whole battery of peptides. This was verified by irradiation at 430 nm (Figure S22). Both isomers could be detected by RP-HPLC, if not stated differently, using the same conditions detailed before (Section: General Materials and Methods).

Characterization of Isomerization in Phosphate Buffer (pH = 6.5)

For the isomerization, a 100 μM solution of the corresponding peptide in 0.1 M of phosphate buffer (pH = 6.5) was irradiated at the specific wavelength for different time periods. The photostationary states reached upon irradiation were determined by integrating the peak areas in the HPLC chromatograms at 287 nm (isosbestic point). All AMPB-containing peptides exhibited similar values, which are in turn also concordant with the published data.^[5,8] The photoisomerization of peptide **7** is illustrated in detail.

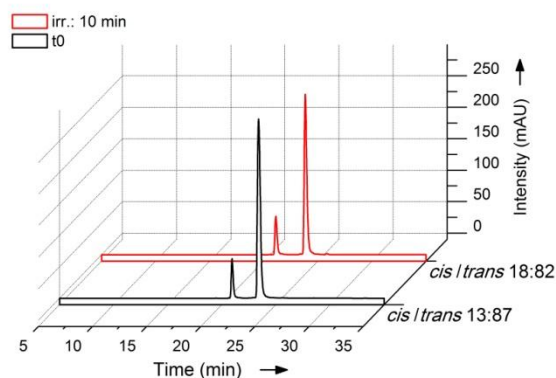


Figure S22. HPLC chromatograms of peptide **7**. After purification, the *trans* ratio is at 87% (black chromatogram). Via 10 min irradiation at 430 nm the *trans* ratio is not increasing (red chromatogram).

To assess the needed irradiation time to reach the maximum *cis* ratio, the same solution was irradiated at 366 nm for different time periods (Figure S23).

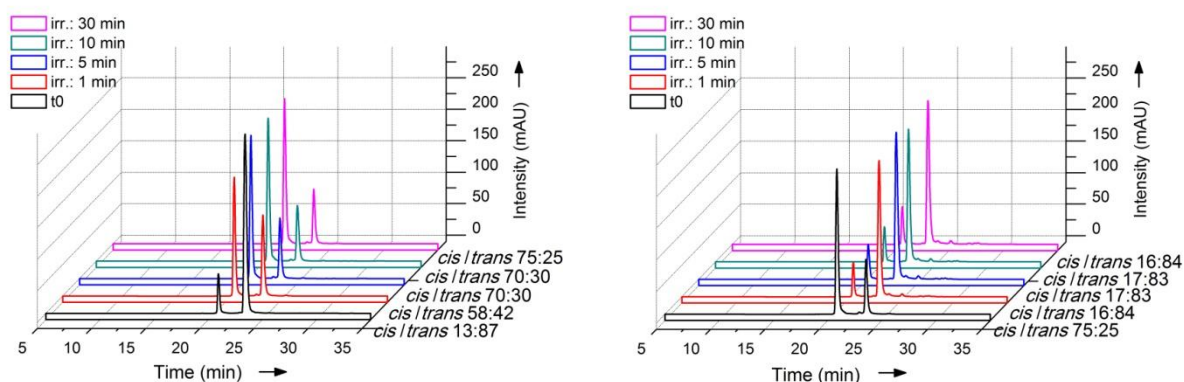


Figure S23. Left: HPLC chromatograms of peptide **7** after irradiation at 366 nm for different time periods. Right: HPLC chromatograms of peptide **7** with 75% *cis* after irradiation at 430 nm for different time periods.

Conversely, thermal *cis* \rightarrow *trans* relaxation is a rather slow process.^[5] A 100 μ M solution of peptide **7** in 0.1 M of phosphate buffer (pH = 6.5) was irradiated at 366 nm for 30 min to obtain 76% of the *cis* isomer. Afterwards, it was stored in total darkness at rt and HPLC chromatograms, under the same conditions as before, were recorded after different time periods (Figure S24). As shown in Figure S24, after 4 days the *cis/trans* ratio is still at 65:35, and even after 11 days, the half-life is not yet reached.

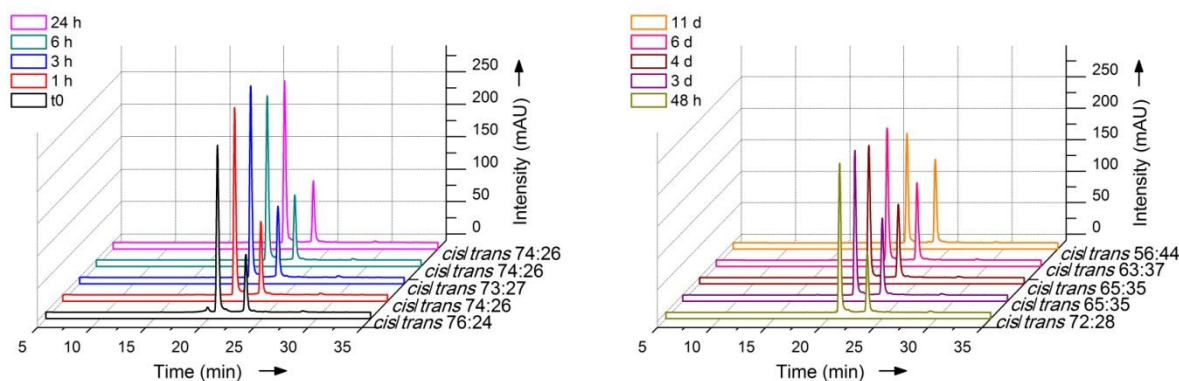


Figure S24. HPLC chromatograms of peptide **7** after the solution was irradiated at 366 nm for 30 min (black chromatogram, t0) and then stored in total darkness. HPLC chromatograms were recorded after the listed times.

Characterization of Isomerization in DMSO

The photoisomerization depends on several factors such as: solvent, concentration, intensity of the lamp etc. Therefore, the initial irradiation for the *trans* \rightarrow *cis* isomerization was optimized for each peptide in the Fluorescence-Polarization based assays. The photoisomerization of peptide **7** is exemplified. For the isomerization, a 3.0 mM solution of the corresponding peptide in DMSO was irradiated at 366 nm for different time periods

(Figure S25). The irradiation to obtain the *cis* isomer was done as described above. For HPLC measurements a gradient of 5% to 30% eluent B and an Eclipse XDB-C18 column (Agilent) were used.

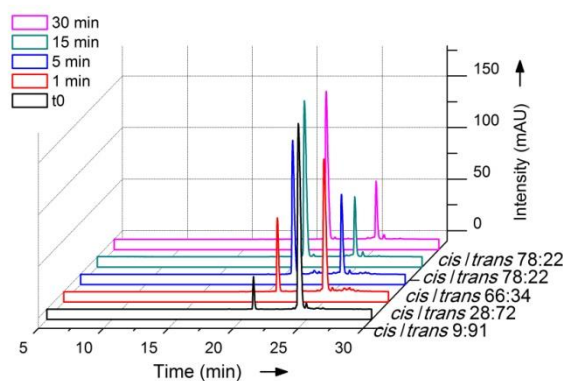


Figure S25. HPLC chromatograms after irradiation of peptide 7 in DMSO.

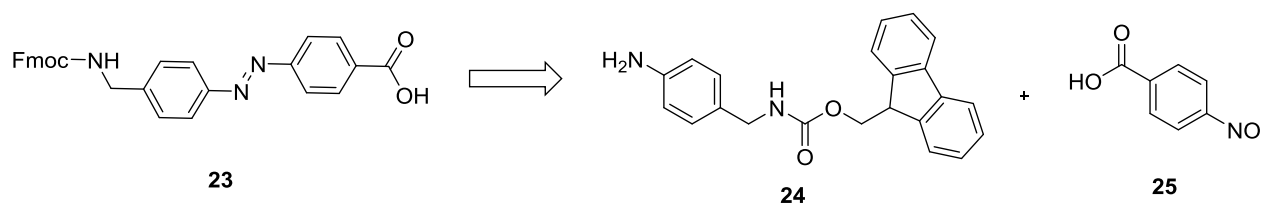
Expression of WDR5 Δ 23

The pET28a-based His-Sumo-WDR5 Δ 23 was expressed from the pET28-MHL vector in *E. coli* DH5 α cells. Expression was performed by inoculation of 640 mL LB medium (3 times) into 2 L baffled flasks with 50 μ g/mL kanamycin and 32 mL of preculture. The culture was incubated on a shaker (Multitron from Infors HT) at 37 $^{\circ}$ C and 200 rpm until an OD₆₀₀ of 0.5-0.6. The temperature was then reduced to 25 $^{\circ}$ C and the expression was induced by adding IPTG up to a final concentration of 0.3 mM. The cells were grown overnight and harvested through centrifugation (5000 rpm, 15 min, 4 $^{\circ}$ C, Sorvall RC6 Plus from Thermo Scientific). The cell pellet from the 2 L expression culture was resuspended in 40 mL lysis buffer (400 μ L PMSF (stock: 15 mg/mL), 400 μ L 10% NP-40, 4 μ L benzamidine (stock: 20mg/mL), 4 μ L 2-mercaptoethanole in 40 mL BC500 (25 mM Tris, pH = 8.0, 500 mM NaCl, 20% glycerol)). Lysis of the cells was carried out using a sonicator Sonoplus HD 220 from Bandelin (3 x 40%, output: 4, with 2 min pulses and 3 min rest). After sonification, the suspension was pelleted by centrifugation (12000 rpm, 20 min, 4 $^{\circ}$ C). The lysate was added to Ni-NTA agarose (Quiagen, Germany) and incubated for 1 h at 4 $^{\circ}$ C. Cleavage of the His₆-tag was achieved by incubation of the resin with 3.0 mL BC150 (25 mM Tris, pH = 8.0, 150 mM NaCl, 20% glycerol) + 18.1 μ L ULP-His₆. WDR5 Δ 23 was eluted from the resin with 3 mL of elution buffer (250 mM imidazole in BC150) and then loaded onto a Mono S 5/50 GL column (GE Healthcare) using an ÄKTA purification system (GE Healthcare). The flow rate was 1 mL/min with the Mono S buffers A (25 mM Tris, pH = 8.0, 20% glycerol) and B (25 mM

Tris, pH = 8.0, 1 M NaCl, 20% glycerol) with first an isocratic flow of 100% A for 10 min followed by a linear gradient of 0-100% B in 20 min. WDR5Δ23 eluted between 35-40% of buffer B. The fractions were checked via SDS-PAGE and pooled. The buffer was changed to the dialysis buffer (25 mM Tris, pH = 8.0, 150 mM NaCl) and the protein was concentrated up to a concentration of 275.57 μM (10.06 mg/mL). The concentrated protein was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

Synthesis of the Molecular Photoswitch

The synthesis of the molecular photoswitch **23**, was carried out via a MILLS-condensation (Scheme S1) according to the literature quoted below. All the analytical data were consistent with the literature.



Scheme S1. Retrosynthetic approach for the *N*-Fmoc-(4-aminomethyl)phenylazobenzoic acid (**23**).

(4-Aminobenzyl)carbamic 9H-fluoren-9-yl methyl ester (**24**)

Following the procedure of MORODER et al.,^[9] 9-fluorenylmethyl *N*-succinimidyl carbonate (101 mg, 0.30 mmol, 1.00 eq) dissolved in MeCN (0.70 mL) was added dropwise to a solution of 4-aminobenzylamine (0.03 mL, 0.30 mmol, 1.00 eq) and Et₃N (0.04 mL, 0.30 mmol, 1.00 eq) in MeCN/DMF 10:1 (0.40 mL). The solution was stirred for 1 h at rt. After filtration the precipitate was washed with methyl-*tert*-butyl-ether/trifluoroethanol 1:1 (2.0 mL). The crude product was dried and purified using flash column chromatography (*n*-Pentane/EtOAc 1:1) to yield the Fmoc-protected aminobenzylamine (48.0 mg, 0.14 mmol, 47%) as a white solid. **TLC:** R_f = 0.44 (*n*-Pentane/EtOAc 1:1) **¹H-NMR (300 MHz, DMSO_{d6}; δ):** 7.89 (2H, d, ³J = 7.4, 2 x CH_{ar,Fmoc}), 7.69 (2H, d, ³J = 7.4, 2 x CH_{ar,Fmoc}), 7.41 (2H, t, ³J = 7.3, 2 x CH_{ar,Fmoc}), 7.32 (2H, t, ³J = 7.7, 2 x CH_{ar,Fmoc}), 6.89 (2H, d, ³J = 8.2, 2 x CH_{ar}), 6.49 (2H, d, ³J = 8.2, 2 x CH_{ar}), 4.93 (s, 2H, CH₂NH), 4.31 (s, 1H, NH), 4.29 (s, 2H, NH₂), 4.21 (t, 1H, ³J = 6.6, CH₁₄), 3.99 (d, 2H, ³J = 5.9, CH₂O). **¹³C-NMR (75 MHz, DMSO_{d6}; δ):** 158.6 (CO₂NH), 141.9 (C_{ar,t}), 138.6 (C_{ar,NH2}), 135.6 (2 x C_{ar,t}), 131.9

(2 x $C_{ar,t}$), 127.9 (3 x $C_{ar,s,H}$), 127.5 (2 x $C_{ar,s,H}$), 126.9 (3 x $C_{ar,s,H}$), 119.9 (2 x $C_{ar,s,H}$), 113.6 (2 x $C_{ar,s,H}$), 74.1 (CH_2), 60.1 (CH), 43.2 (CH_2). **HRMS-ESI⁺ (m/z):** $[M+Na]^+$ calcl. for: $C_{22}H_{20}N_2O_2Na$, 367.1417; found: 367.1416.

4-Nitrosobenzoic acid (25)

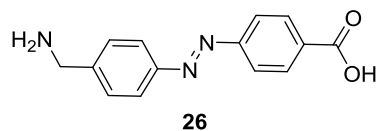
Following the procedure of PRIEWISCH et al.,^[10] aminobenzoic acid (3.43 g, 25.0 mmol, 1.00 eq) was suspended in DCM (40.0 mL). A solution of Oxone (30.74 g, 50.0 mmol, 2.00 eq) in H_2O (155 mL) was added and the suspension stirred for 1 h at rt. The precipitate was filtered and washed with H_2O . The crude product was freeze dried to yield 4-nitrosobenzoic acid (3.78 g, 25.0 mmol, >99.9%) as a yellow solid, which was directly used without any further purification. **TLC:** $R_f = 0.39$ (DCM/MeOH 9:1) **¹H-NMR (300 MHz, DMSO_{d6}; δ):** 13.41 (1H, s, OH), 8.27 (2H, d, $^3J = 8.6$, CH_{ar}), 8.03 (2H, d, $^3J = 8.6$, CH_{ar}). **¹³C-NMR (75 MHz, DMSO_{d6}; δ):** 168.9 (COOH), 166.9 ($C_{ar}NO$), 135.6 ($C_{ar,1}$), 130.4 (C_{ar}), 117.5 (C_{ar}).

N-Fmoc-(4-aminomethyl)phenylazobenzoic acid (23)

Following the procedure of PRIEWISCH et al.,^[10] Fmoc-protected aminobenzylamine (1.288 g, 3.74 mmol, 1.00 eq) was added to a suspension of 4-nitrosobenzoic acid (3.78 g, 25.0 mmol, 6.68 eq) in acetic acid/DMSO 1:1 (315 mL). The suspension was stirred for 36 h at rt. After careful addition of H_2O , the precipitate was filtered off and washed with H_2O . The crude product was purified by recrystallization in acetone to yield N-Fmoc-(4-aminomethyl)phenylazobenzoic acid (1.39 g, 2.92 mmol, 78%) as an orange solid. **TLC:** $R_f = 0.38$ (DCM/MeOH 9:1) **¹H-NMR (300 MHz, DMSO_{d6}; δ):** 13.28 (s, 1H, OH), 8.15 (d, $^3J = 8.5$, 2H, 2 x CH_{ar}), 7.93 (m, 6H, 3 x CH_{ar}), 7.71 (d, $^3J = 7.4$, 2H, 2 x CH_{ar}), 7.43 (t, $^3J = 8.8$, 2H, 2 x CH_{ar}), 7.34 (t, $^3J = 7.1$, 2H, 2 x CH_{ar}), 4.40 (d, $^3J = 6.7$, 2H, CH_2 -33), 4.30 (d, $^3J = 6.0$, 2H, CH_2 -17), 4.24 (t, $^3J = 6.8$, 1H, CH-23). **¹³C-NMR (75 MHz, DMSO_{d6}; δ):** 166.6 (COOH), 156.3 (CONH), 154.3 ($C_{ar}N$), 150.9 ($C_{ar}N$), 144.2 (C_{ar}), 143.8 (2 x C_{ar}), 140.7 (2 x C_{ar}), 132.7 (C_{ar}), 130.55 (2 x $C_{ar}H$), 127.9 (2 x $C_{ar}H$), 127.5 (2 x $C_{ar}H$), 126.9 (2 x $C_{ar}H$), 125.1 (2 x $C_{ar}H$), 122.8 (2 x $C_{ar}H$), 122.4 (2 x $C_{ar}H$), 120.0 (2 x $C_{ar}H$), 65.3 (CH_2O), 46.8 (CH), 43.5 (CH_2NH). **HRMS-ESI⁻ (m/z):** $[M-H]^-$ calcl. for: $C_{29}H_{22}N_3O_4$, 476.1616; found: 476.1603. **IR (thin film) $\tilde{\nu}$ (cm⁻¹):** 3305 (vw), 2826 (vw), 1686 (vs), 1603 (w), 1525 (m), 1462 (w), 1448 (w), 1423 (m), 1290 (s), 1257 (s), 1141 (w), 1103 (w), 1082

(vw), 1045 (w), 1010 (w), 986 (w), 936 (w), 865 (m), 838 (w), 775 (m), 757 (m), 734 (s), 690 (m), 666 (w), 641 (w), 620 (w), 591 (w), 556 (w), 541(m), 425 (m), 400 (w), 387 (w).

(4-Aminomethyl)phenylazobenzoic acid (**26**)



N-Fmoc-(4-aminomethyl)phenylazobenzoic acid (50.0 mg, 0.105 mmol, 1.00 eq) was dissolved in THF (4.12 mL) and piperidine (1.03 mL, 10.5 mmol, 100 eq) was added. The solution was stirred at rt for 2.5 h. After the reaction was completed, the solvent was evaporated under reduced pressure. The residue was washed with DCM, and the pure product (25.5 mg, 0.1 mmol, 95%) was obtained as an orange solid. **TLC:** $R_f = 0.03$ (DCM/MeOH 9:1) **$^1\text{H-NMR}$ (300 MHz, DMSO-d_6 ; δ):** 8.31 (s, 2H, NH_2), 8.17 (d, $^3J = 8.6$, 2H, 2 x CH_{ar}), 8.01 (d, $^3J = 3.3$, 2H, 2 x CH_{ar}), 7.98 (d, $^3J = 3.5$, 2H, 2 x CH_{ar}), 7.71 (d, $^3J = 8.5$, 2H, 2 x CH_{ar}), 4.18 (q, $^3J = 5.5$, 2H, $\text{CH}_2\text{-NH}_2$). **$^{13}\text{C-NMR}$ (75 MHz, DMSO-d_6 ; δ):** 166.6 (COOH), 154.1 (C_{ar}N), 151.7 (C_{ar}N), 137.9 (C_{ar}), 133.0 (C_{ar}), 130.6 (2 x C_{ar}), 129.9 (2 x C_{ar}), 122.9 (2 x C_{ar}), 122.6 (2 x C_{ar}), 41.9 (CH_2NH). **HRMS-ESI (m/z):** $[\text{M}+\text{H}]^+$ calcd. for: $\text{C}_{14}\text{H}_{13}\text{N}_3\text{O}_2\text{H}$, 256.1081; found: 256.1077.

Fluorescence Polarization-based Assays

All FP-based assays were performed in black 96-well microtiter plates (Greiner) and FP was measured as milipolarization (mP) units on a plate reader (Tecan Spark 20M); settings: excitation: 485 nm; emission: 530 nm; gain: optimal; Z-position: calculated from control well (0% inhibition: tracer and protein in assay buffer); number of flashes: 30; G-factor: 1.000. The data evaluation was done, using GraphPad Prism 6 software.

The concentration of the peptides was calculated through UV/Vis measurements (Beckman-Coulter DU800, AMPB-containing peptides (for *trans* isomer) $\epsilon_{335} = 25000 \text{ L mol}^{-1} \text{ cm}^{-1}$,^[6] fluorescently tagged peptide **22** $\epsilon_{494} = 76900 \text{ L mol}^{-1} \text{ cm}^{-1}$).^[11]

Saturation Binding Experiments

Dilutions of the WDR5Δ23 protein (2.0-0 μM, 25 μL) were added to 100 μL of a fixed concentration (20 nM) of the fluorescently tagged tracer peptide **22**. Both, protein and tracer, were dissolved in the assay buffer (0.1 M phosphate buffer pH = 6.5, 25 mM KCl, 0.01% Triton) with 4% of DMSO. The total volume was 125 μL. Each assay had two controls: blank (without tracer and protein) and tracer only. The plates were incubated at rt on a shaker (Edmund Bühler TiMix Control TH15) to reach equilibrium and the mP values were measured after 3 h. The measurements were done in triplicate and the standard deviation was calculated. The K_d -value was calculated by converting the mP values into their corresponding anisotropy (A) values. These values were normalized and plotted versus the respective concentrations of the protein with a nonlinear regression according to the following

$$\text{equation:}^{[12]} A = A_f + (A_b - A_f) \frac{(L_T + K_d + R_T) - \sqrt{[(L_T - K_d - R_T)^2 - 4L_T R_T]}}{2L_T}$$

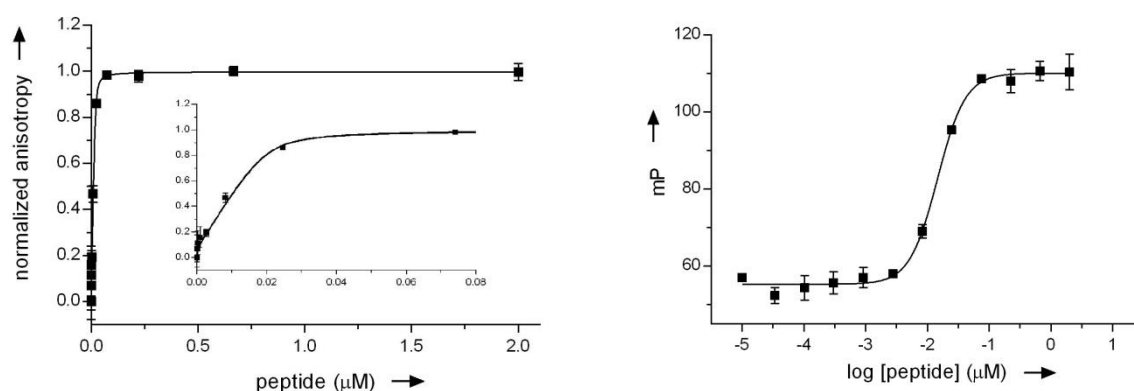


Figure S26. Left: saturation binding curve of peptide **18** obtained from the normalized anisotropy (20 nM tracer). The small picture is zoomed in to show the small x-values. Right: dose-response curve for the same experiment and conditions.

The dose-response curve provides the EC_{50} value of the tracer peptide **22**. To obtain the K_d value, the mP values had to be converted into their corresponding anisotropy (A) values, which were plotted versus the respective concentrations of the protein with a nonlinear regression according to the equation described above. The obtained values are the following: $EC_{50} = 0.0145 \pm 0.002 \mu\text{M}$; $K_d = 0.00104 \pm 0.0005 \mu\text{M}$.

Corroboration of Absence of Quenching Artefacts

To ultimately ensure, that the AMPB molecule incorporated into the peptide backbone does not have any quenching effects in the used FP-based assay, some controls were performed.

Thus, the anisotropy of a 20 nM solution of fluorescein dissolved in 0.1 M phosphate buffer pH = 6.5 with increasing concentrations of AMPB **26** dissolved in DMSO, was measured in triplicate. This measurement was repeated three times, independently. Figure S27 exemplifies one measurement. As shown in such figure, the increase of the amount of AMPB has no influence on the measured milli Polarization values.

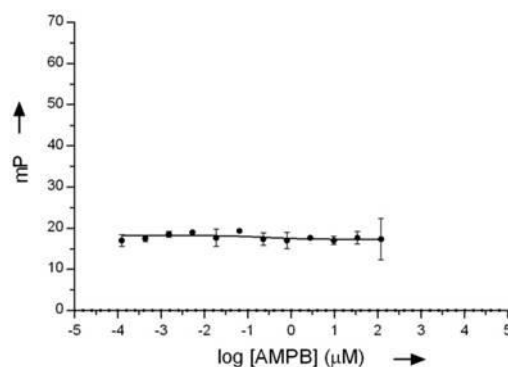


Figure S27. Measured mP values of a 20 nM fluorescein solution with increasing AMPB concentrations.

Furthermore, we also performed a control experiment where 120 μL of a pre-incubated complex solution of WDR5Δ23 and the tracer peptide **22** in assay buffer (0.1 M phosphate buffer pH = 6.5, 25 mM KCl, 0.01% Triton) plus peptide **3** were added to 5 μL of increasing dilutions of AMPB **26** in DMSO, reaching the final concentrations of 72 nM for WDR5Δ23, 20 nM for the tracer peptide and 0.5 μM for the peptide **3**. Again, the experiment was done in triplicate and performed three times independently. As illustrated in Figure S28, AMPB did not affect the milli Polarization values, which are constant during the whole experiment and the same as the previous value without extra AMPB addition as it is shown the overlay with the competitive Fluorescence Polarization-based Binding Assay (Figure S28 right).

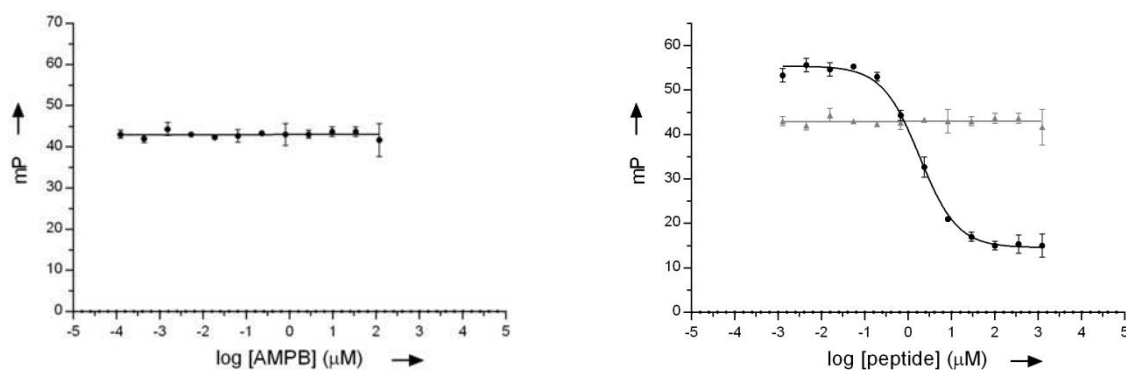


Figure S28. Left: measured mP values of FP-assay (20 nM tracer, 72 nM WDR5, 0.5 μM peptide **3**) with increasing concentrations of AMPB. Right: Overlay of dose response curve of peptide **3**, with left picture.

Competitive Fluorescence Polarization-based Binding Assays

The binding affinities of both isomers of each peptide were measured separately. The peptides were dissolved in DMSO. In order to calculate the binding affinity of the *trans* isomers, the peptide solutions were used directly, without previous irradiation since it was demonstrated that the maximum *trans* ratio is obtained directly from the purification. Contrarily, to determine the binding affinity of the *cis* isomers, the peptide solutions were irradiated at 366 nm for 2 h in a transparent glass vial, previously to the assay. Afterwards, the whole assay had to be done in darkness to prevent *cis* → *trans* relaxation. 120 μL of a preincubated complex solution of WDR5Δ23 and the tracer peptide **22** in assay buffer (0.1 M phosphate buffer pH = 6.5, 25 mM KCl, 0.01% Triton) were added to 5 μL of dilutions of the respective AMPB-containing peptide in DMSO, reaching the final concentrations of 72 nM for WDR5Δ23 and 20 nM for the tracer peptide **22**. Three control wells were included in each plate: blank (without tracer and protein), 100% inhibition (tracer only) and 0% inhibition (protein and tracer). The plates were incubated on a shaker (Edmund Bühler TiMix Control TH15) at rt for 5 h and the mP values were recorded. K_i -values were calculated using the following equation described previously by WANG et al. and the corresponding webpage

provided by them.^[13]

$$K_i = \frac{[I]_{50}}{\frac{[L]_{50}}{K_d} + \frac{[P]_0}{K_d} + 1}$$

To verify that the used assay conditions do not influence the *cis/trans* ratio, a control well (in triplicate) with 5 μL of the respective peptide in DMSO irradiated at 366 nm (*cis*), as well as the non-irradiated peptide (*trans*), each diluted with 120 μL assay buffer were included in each plate. Before and after the assay, HPLC chromatograms of such controls were recorded. We did not observe any influence on *cis/trans* ratio due to the assay conditions. Table S1 summarizes the *cis/trans* ratio of all AMPB-containing peptides after irradiation at 366 nm (*cis*) for 1 h of the respective peptide solution in DMSO and the non-irradiated peptide (*trans*) before the use in the FP-based assays, as well as the *cis/trans* ratios of the control wells after the FP-based assay.

Furthermore, to rule out any photo-degradative effects, we not just measured the inhibition constants of the non-irradiated control peptide **3**, but also of the irradiated peptide **3** at both, 366 nm and 430 nm. HPLC-chromatograms after the irradiation but before the FP-based assay were recorded as well as after the FP-based. These, and the chromatograms of the AMPB-containing peptides, are shown in Figure S29 to Figure S58.

Table S1. *cis/trans* ratio of the AMPB-containing peptides after purification, after irradiation at 366 nm and after the FP-assay.

peptide sequence	<i>cis/trans</i> ratio of non-irradiated peptides before FP-assay	<i>cis/trans</i> ratio of irradiated peptides (366 nm for 1 h) before FP-assay	<i>cis/trans</i> ratio of non-irradiated peptides after FP-assay	<i>cis/trans</i> ratio of irradiated peptides (366 nm) after FP-assay
H ₂ N-S X ARAEVHLRKS 4	12:88	75:25	20:80	71:29
H ₂ N-S X RAEVHLRKS 5	8:92	75:25	9:91	79:21
H ₂ N-SAR X EVHLRKS 6	8:92	79:21	11:89	81:19
H ₂ N-SARAX X VHLRKS 7	9:91	85:15	10:90	80:20
H ₂ N-SARAE X HRLRKS 8	10:90	80:20	11:89	79:21
H ₂ N-SARAEV X LRKS 9	9:91	79:21	9:91	79:21
H ₂ N-SARAEVH X RKS 10	10:90	81:19	11:89	81:19
H ₂ N-SARAEVHL X KS 11	7:93	83:17	7:93	83:17
H ₂ N-SARAEVHLR X S 12	15:85	79:21	15:85	80:20
H ₂ N-SARAX X HRLRKS 13	10:90	79:21	12:88	89:21
Ac-ARAX X VHLRKS 14	13:87	87:13	13:87	80:20
Ac-ARAX X VY 15	10:90	79:21	10:90	79:21
H ₂ N-SARAX X VHLRKS(R) ₈ 16	12:88	84:16	10:90	80:20
H ₂ N-SARAX X VHLRKS (RAhxR) ₄ AhxB 17	4:96	82:18	4:96	78:22

X = AMPB, B = β -alanine

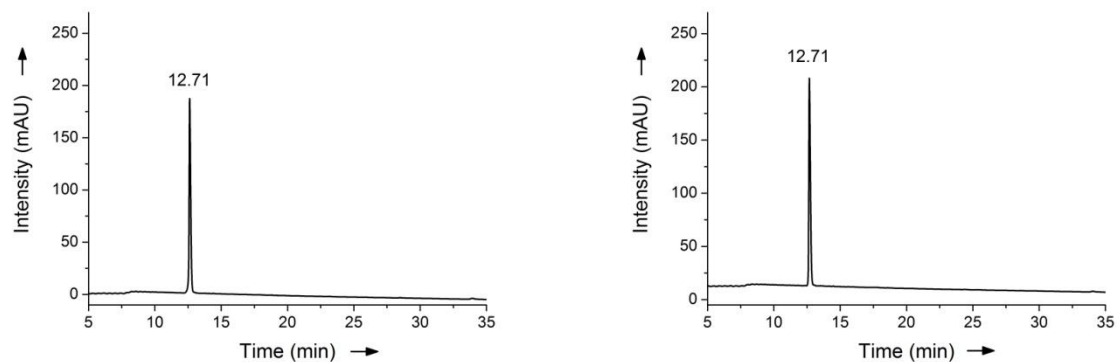


Figure S29. HPLC chromatograms of peptide **3** ($\text{H}_2\text{N-SARAEVHLRKS-NH}_2$) before the FP-based assay. Left: irradiated at 366 nm. Right: irradiated at 430 nm.

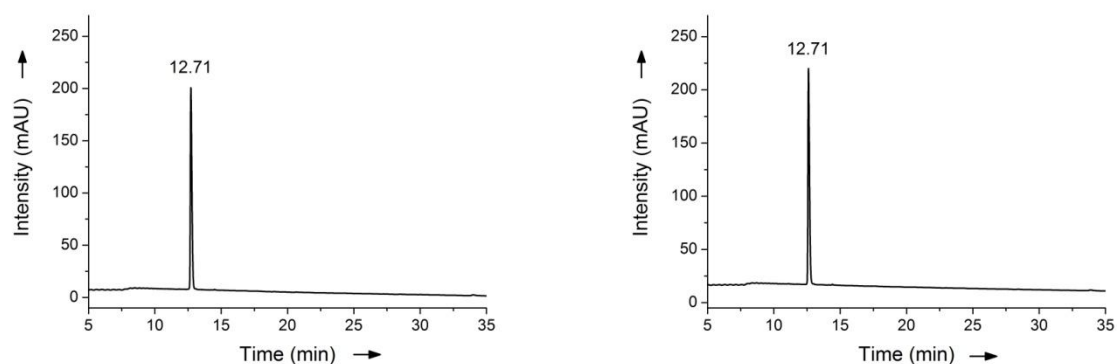


Figure S30. HPLC chromatograms of peptide **3** ($\text{H}_2\text{N-SARAEVHLRKS-NH}_2$) after the FP-based assay. Left: irradiated at 366 nm. Right: irradiated at 430 nm.

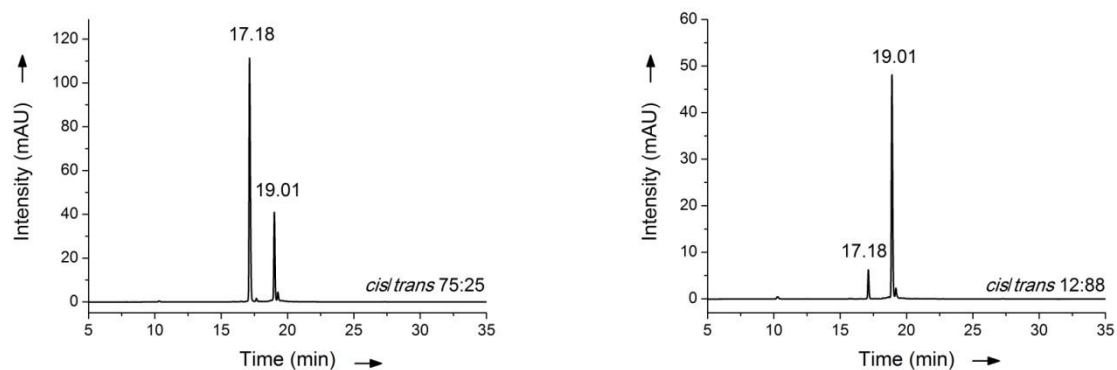


Figure S31. HPLC chromatograms of peptide **4** ($\text{H}_2\text{N-S-AMPB-ARAEVHLRKS-NH}_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

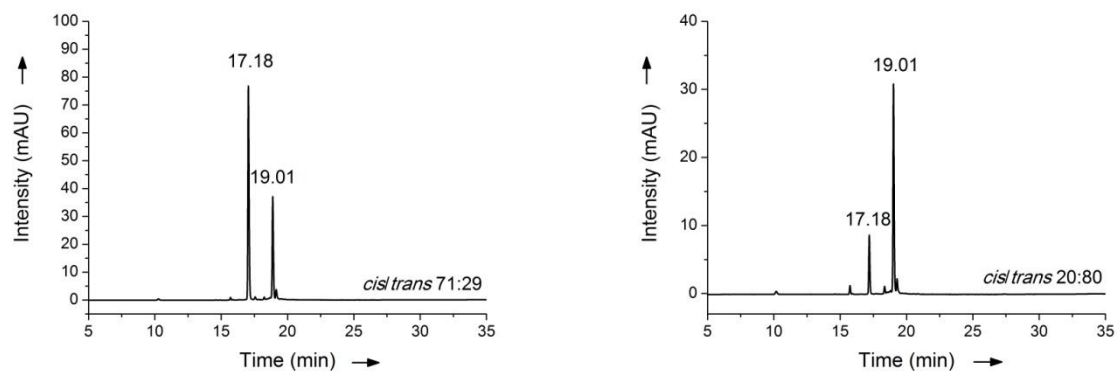


Figure S32. HPLC chromatograms of peptide 4 ($\text{H}_2\text{N-S-AMPB-ARAEVHLRKS-NH}_2$) after the FP-based assay. Left: irradiated peptide at 366 nm Right: non-irradiated peptide.

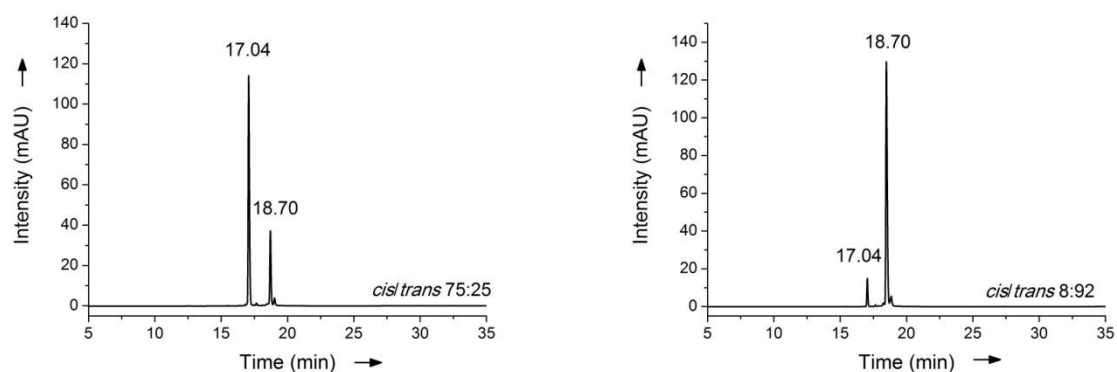


Figure S33. HPLC chromatograms of peptide 5 ($\text{H}_2\text{N-S-AMPB-RAEVHLRKS-NH}_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

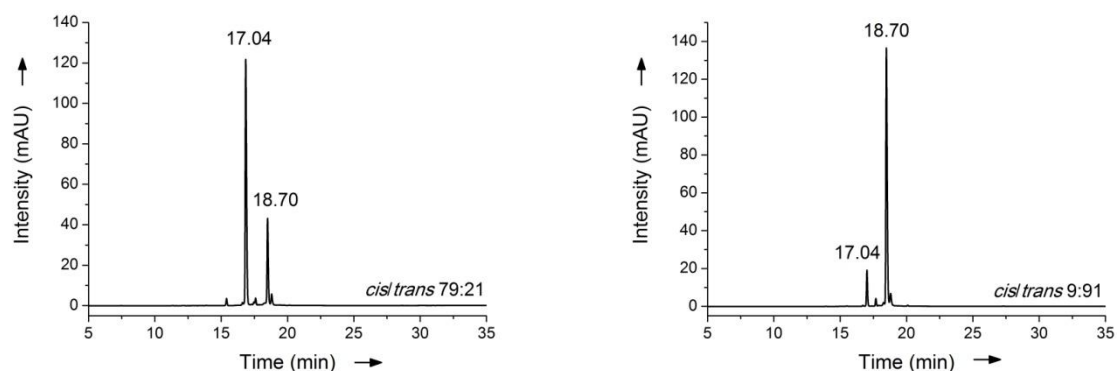


Figure S34. HPLC chromatograms of peptide 5 ($\text{H}_2\text{N-S-AMPB-RAEVHLRKS-NH}_2$) after the FP-based assay, left: irradiated peptide at 366 nm (*cis*). Right: non-irradiated peptide (*trans*).

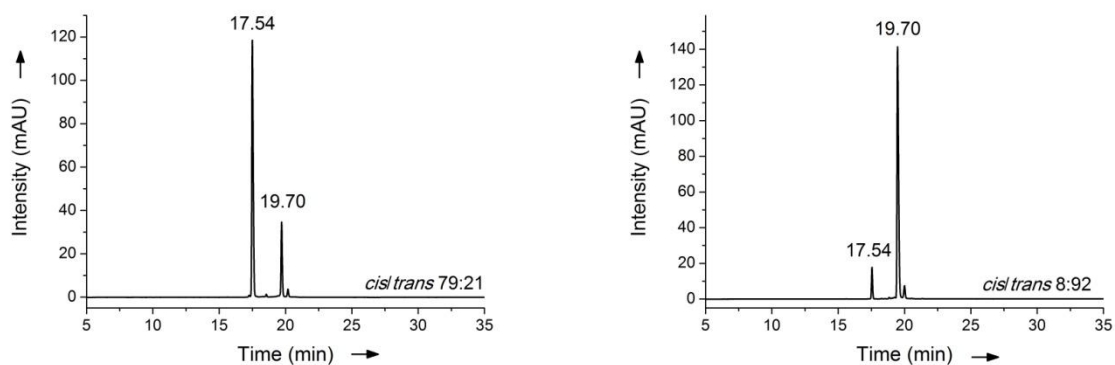


Figure S35. HPLC chromatograms of peptide 6 ($\text{H}_2\text{N-SAR-AMPB-EVHLRKS-NH}_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

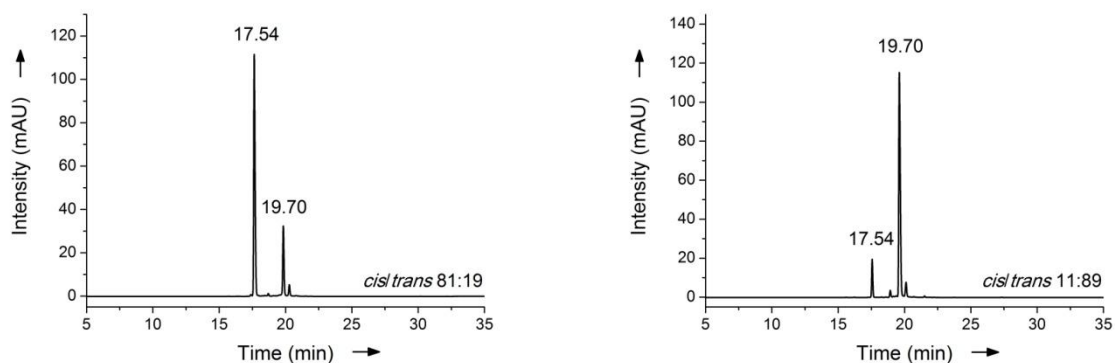


Figure S36. HPLC chromatograms of peptide 6 ($\text{H}_2\text{N-SAR-AMPB-EVHLRKS-NH}_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

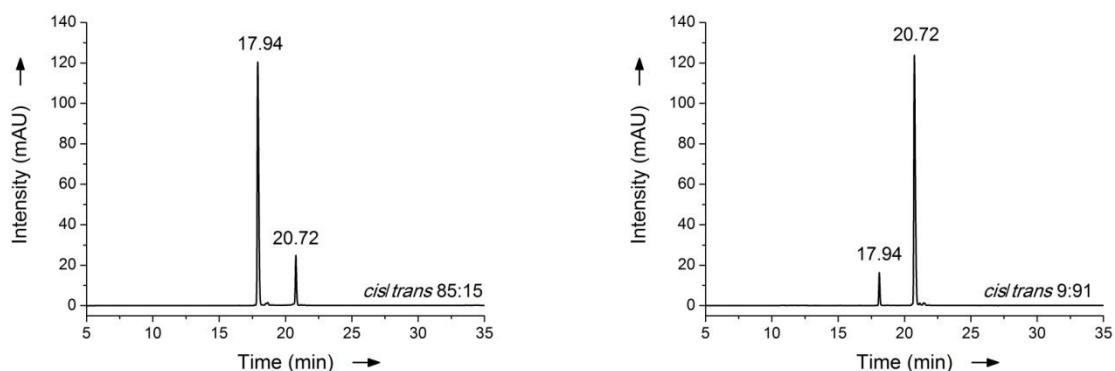


Figure S37. HPLC chromatograms of peptide 7 ($\text{H}_2\text{N-SARA-AMPB-VHLRKS-NH}_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

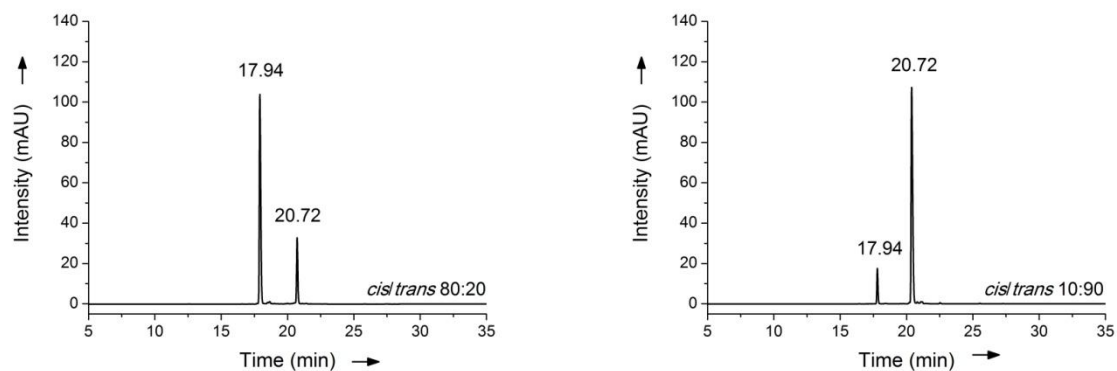


Figure S38. HPLC chromatograms of peptide **7** ($H_2N-SARA-AMPB-VHLRKS-NH_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

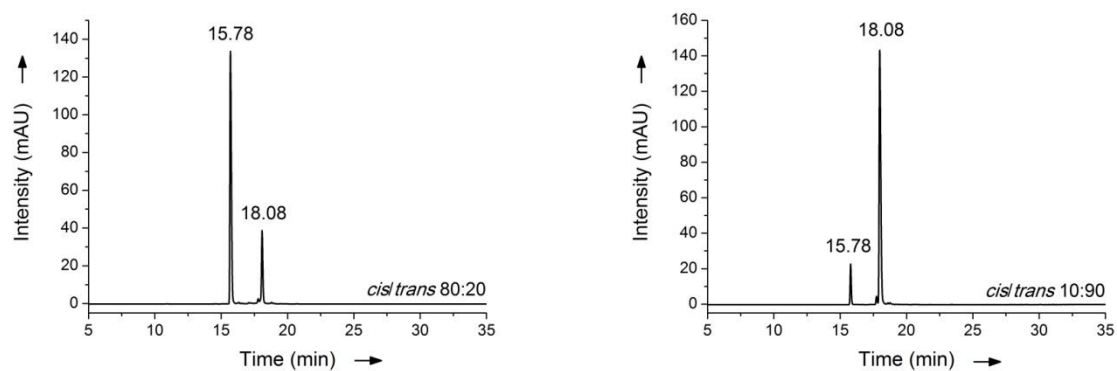


Figure S39. HPLC chromatograms of peptide **8** ($H_2N-SARAE-AMPB-HLRKS-NH_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

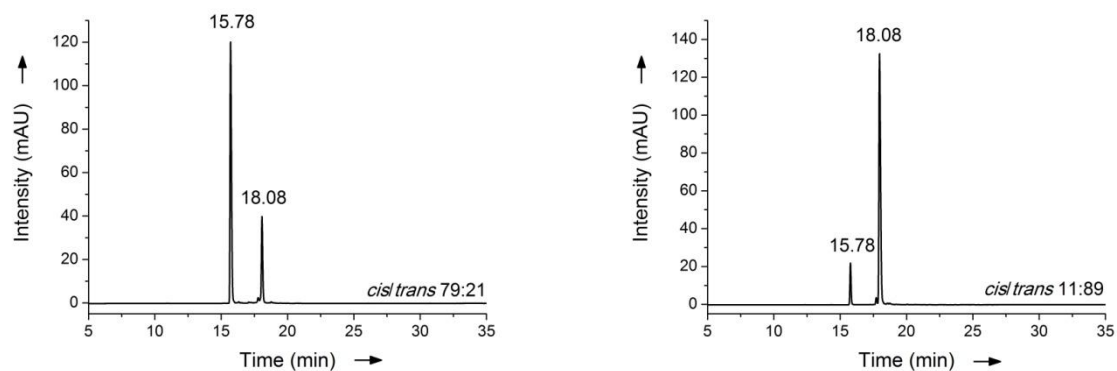


Figure S40. HPLC chromatograms of peptide **8** ($H_2N-SARAE-AMPB-HLRKS-NH_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

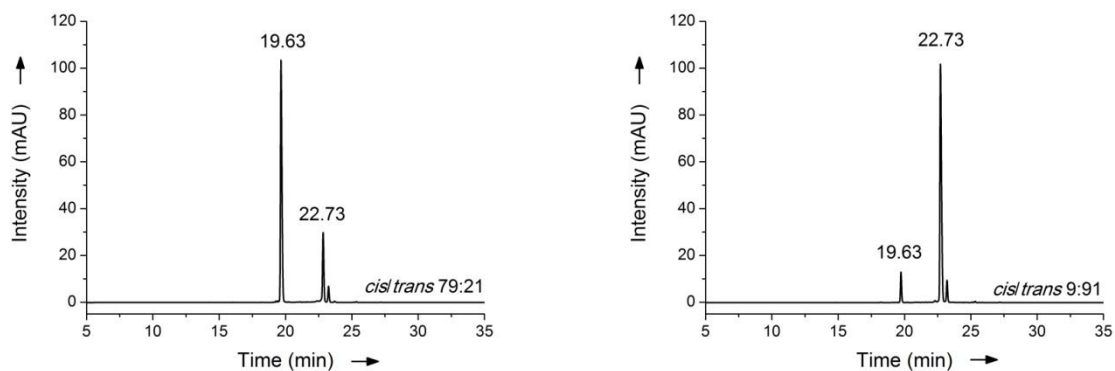


Figure S41. HPLC chromatograms of peptide **9** ($\text{H}_2\text{N-SARAEV-AMPB-LRKS-NH}_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

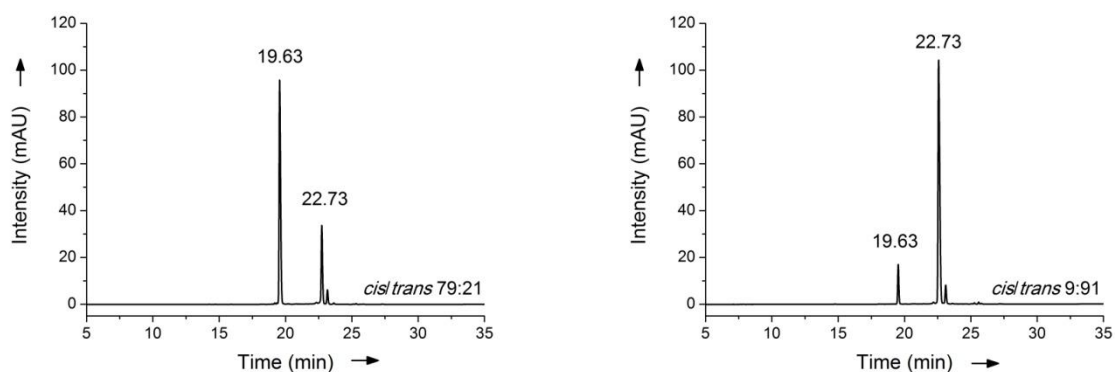


Figure S42. HPLC chromatograms of peptide **9** ($\text{H}_2\text{N-SARAEV-AMPB-LRKS-NH}_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

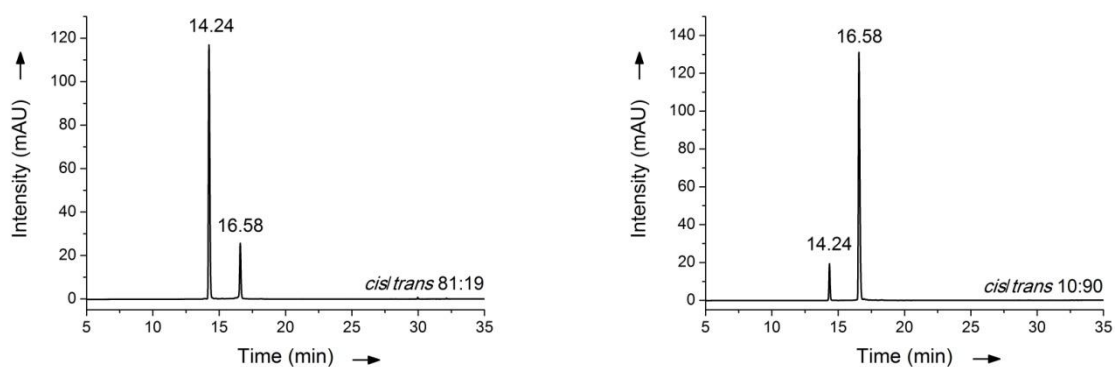


Figure S43. HPLC chromatograms of peptide **10** ($\text{H}_2\text{N-SARAEVH-AMPB-RKS-NH}_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

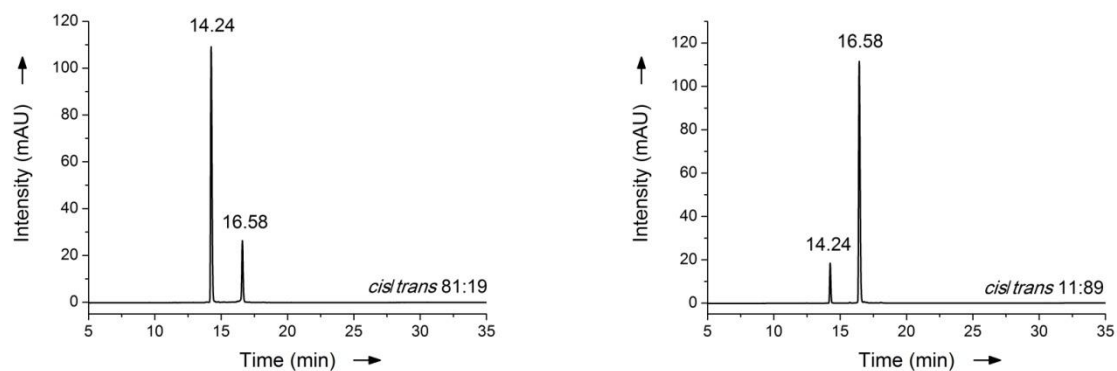


Figure S44. HPLC chromatograms of peptide **10** ($H_2N-SARAEVH-AMPB-RKS-NH_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

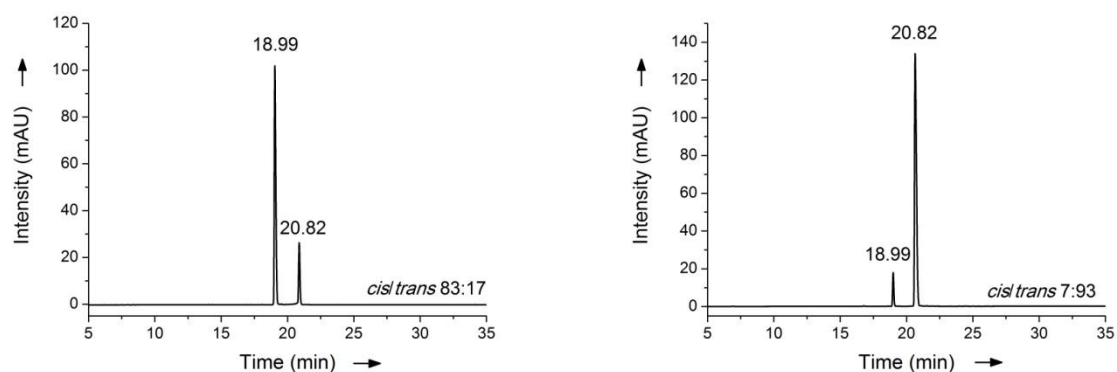


Figure S45. HPLC chromatograms of peptide **11** ($H_2N-SARAEVHL-AMPB-KS-NH_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

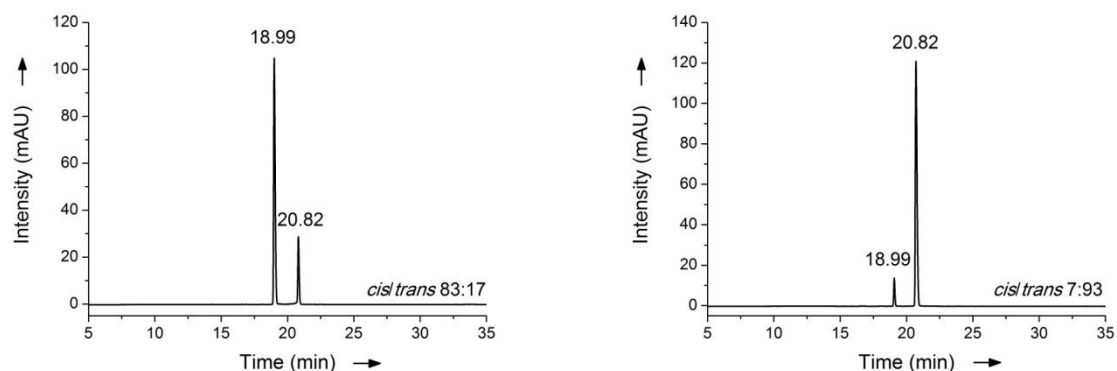


Figure S46. HPLC chromatograms of peptide **11** ($H_2N-SARAEVHL-AMPB-KS-NH_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

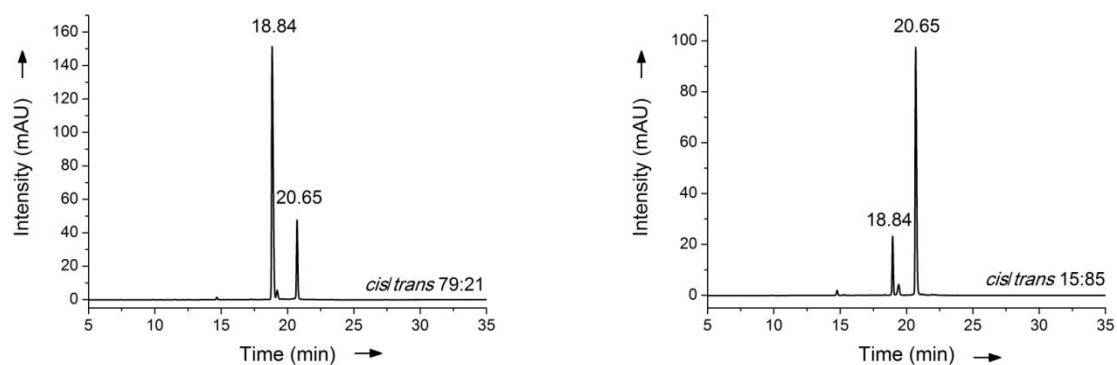


Figure S47. HPLC chromatograms of peptide **12** ($\text{H}_2\text{N-SARAEVHLR-AMPB-S-NH}_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

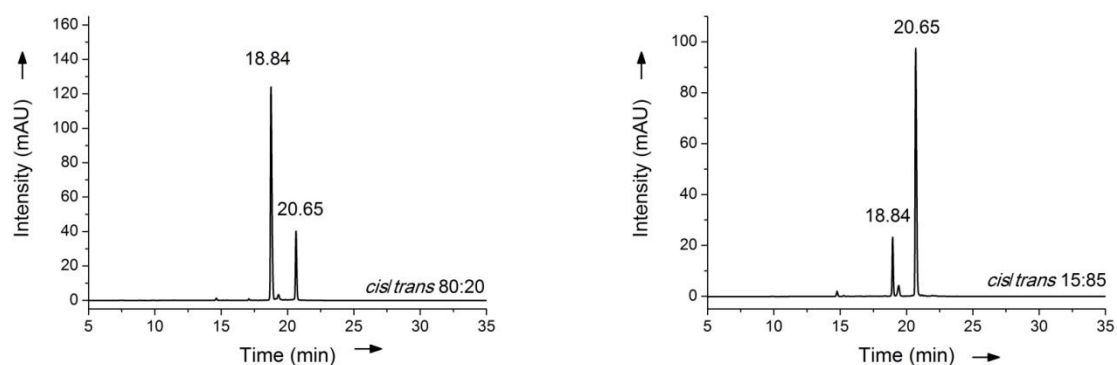


Figure S48. HPLC chromatograms of peptide **12** ($\text{H}_2\text{N-SARAEVHLR-AMPB-S-NH}_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

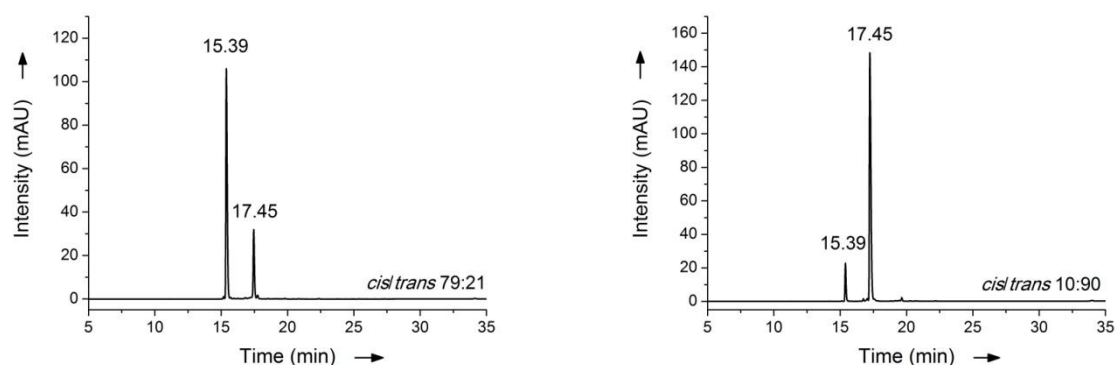


Figure S49. HPLC chromatograms of peptide **13** ($\text{H}_2\text{N-SARA-AMPB-HLRKS-NH}_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

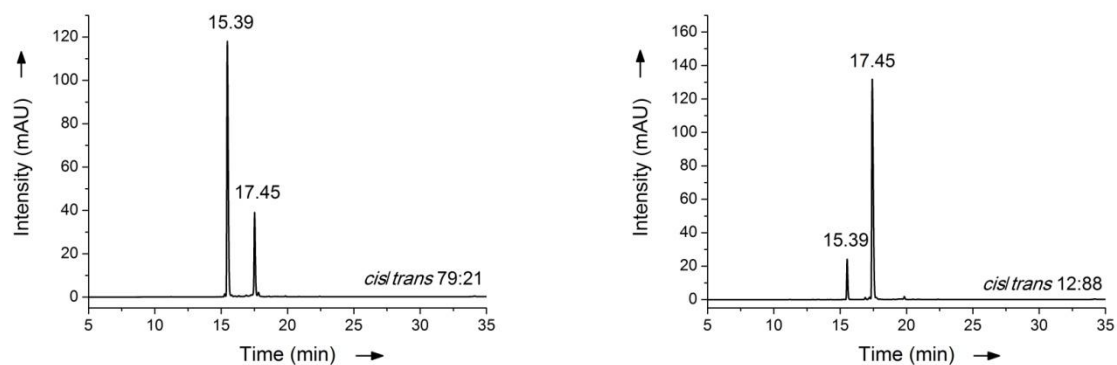


Figure S50. HPLC chromatograms of peptide **13** ($H_2N-SARA-AMPB-HLRKS-NH_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

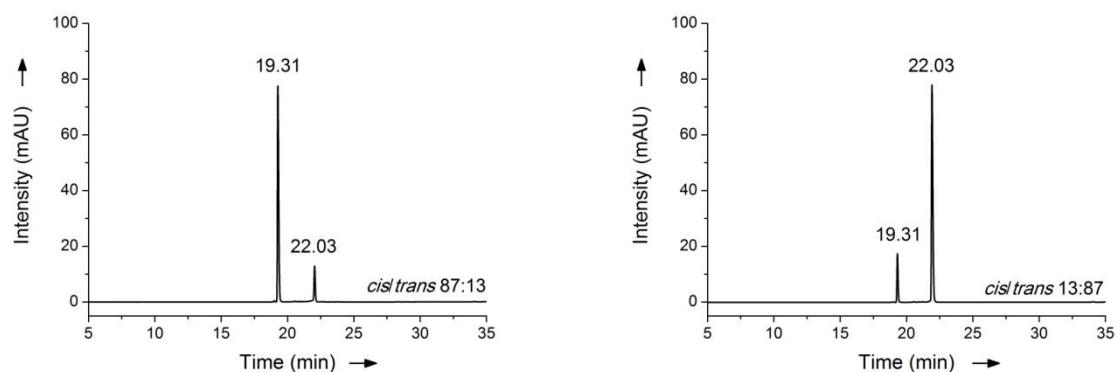


Figure S51. HPLC chromatograms of peptide **14** ($Ac-ARA-AMPB-VHLRKS-NH_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

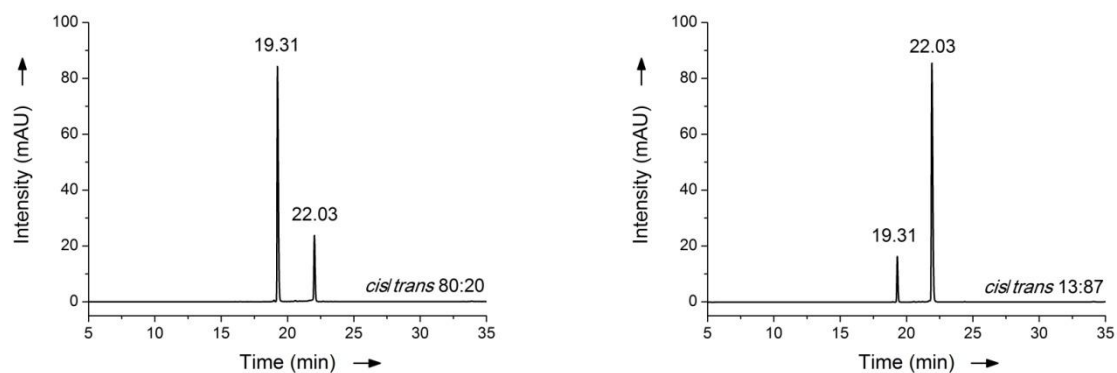


Figure S52. HPLC chromatograms of peptide **14** ($Ac-ARA-AMPB-VHLRKS-NH_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

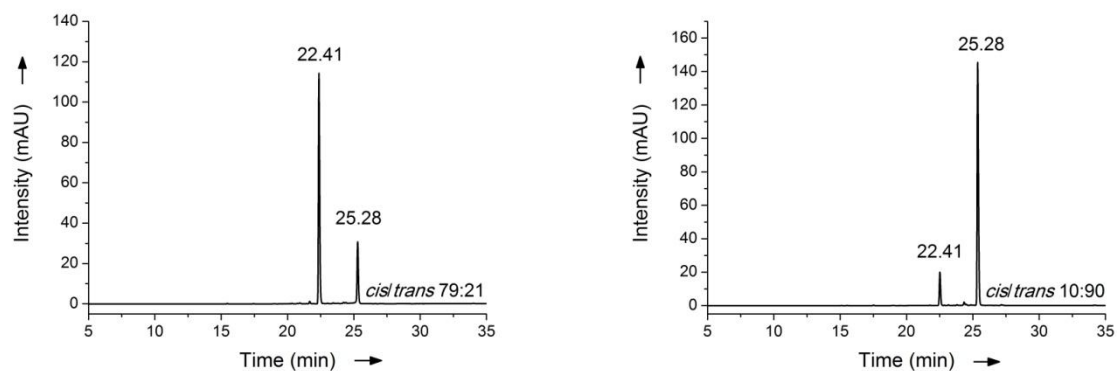


Figure S53. HPLC chromatograms of peptide **15** (Ac-ART-AMPB-VY-NH₂) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

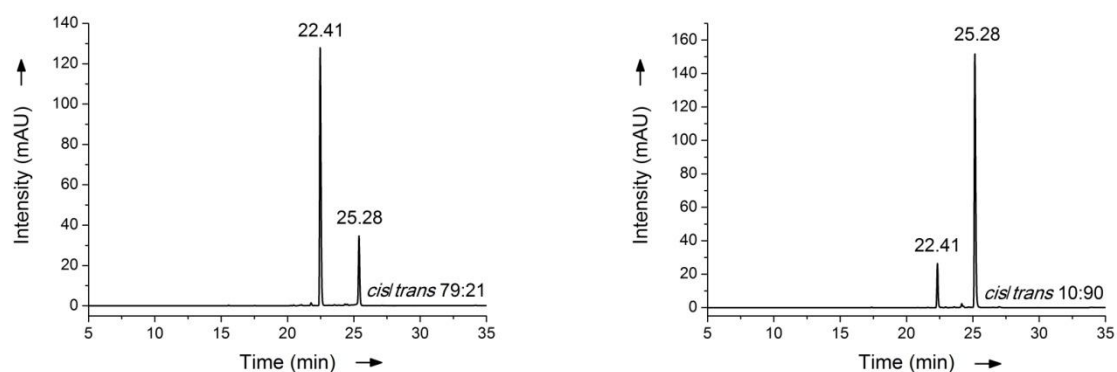


Figure S54. HPLC chromatograms of peptide **15** (Ac-ART-AMPB-VY-NH₂) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

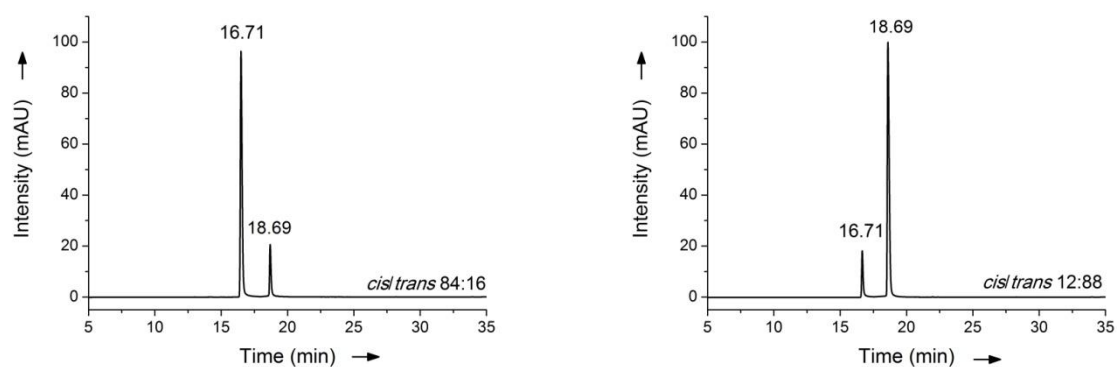


Figure S55. HPLC chromatograms of peptide **16** (H₂N-SARA-AMPB-VHLRKS-R₈-NH₂) before the FP-based assay, left: irradiated peptide at 366 nm (*cis*). Right: non-irradiated peptide (*trans*).

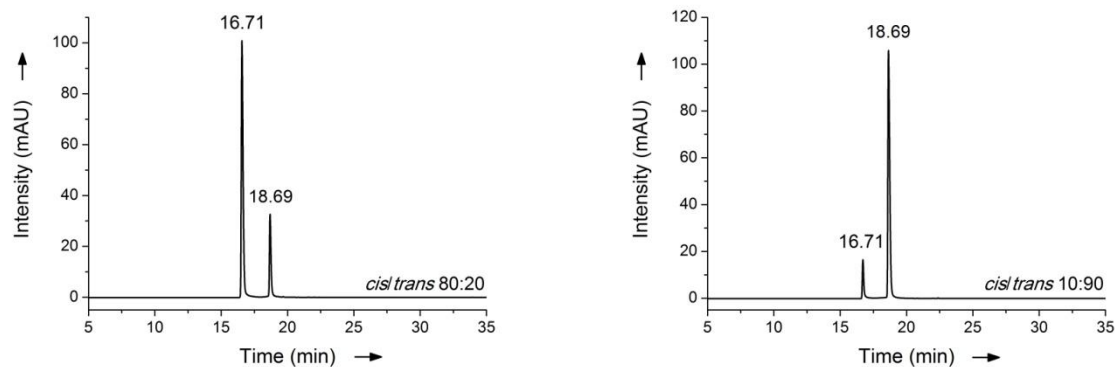


Figure S56. HPLC chromatograms of peptide **16** (H_2N -SARA-AMPB-VHLRKS- R_8 -NH $_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

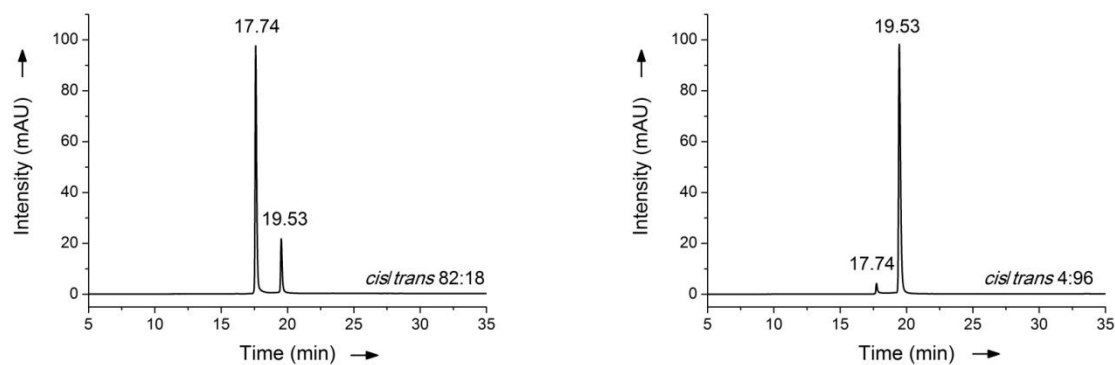


Figure S57. HPLC chromatograms of peptide **17** (H_2N -SARA-AMPB-VHLRKS-(RahxR) $_4$ RAhxB-NH $_2$) before the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

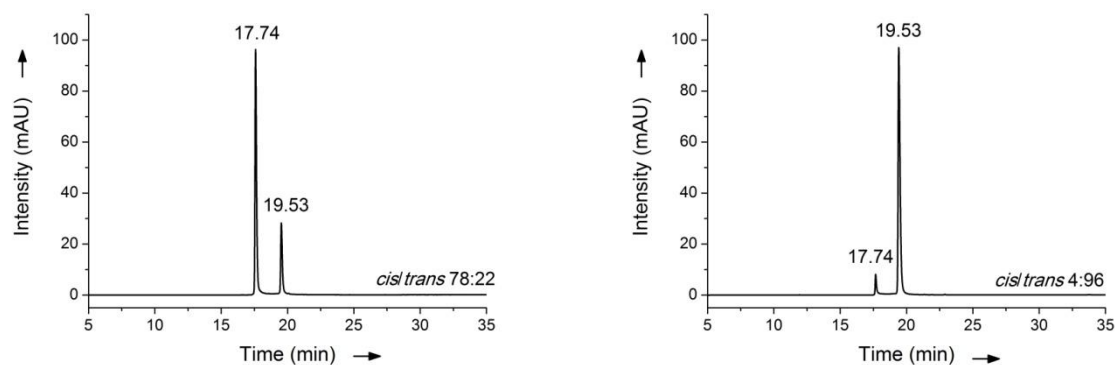


Figure S58. HPLC chromatograms of peptide **17** (H_2N -SARA-AMPB-VHLRKS-(RahxR) $_4$ RAhxB-NH $_2$) after the FP-based assay. Left: irradiated peptide at 366 nm. Right: non-irradiated peptide.

The FP measurements were done in triplicate and the competitive assay of both isomers of every peptide was measured three times independently. Here, the graphs of one measurement are exemplified for each peptide.

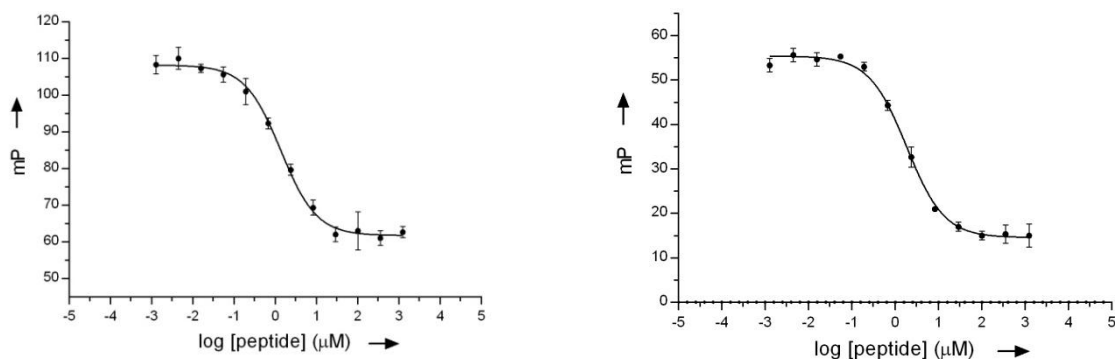


Figure S59. Dose-response curves of: left: peptide **3** (SARAEVHLRKS) non-irradiated, right: peptide **3** (SARAEVHLRKS) irr. at 430 nm.

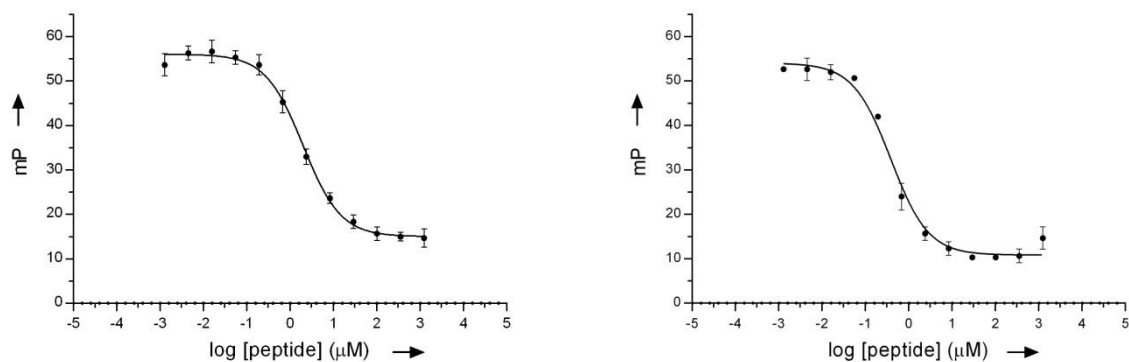


Figure S60. Dose-response curves of: left: peptide **3** (SARAEVHLRKS) irr. at 360 nm, right: peptide **20** (SARAEVHLRKS-R₈).

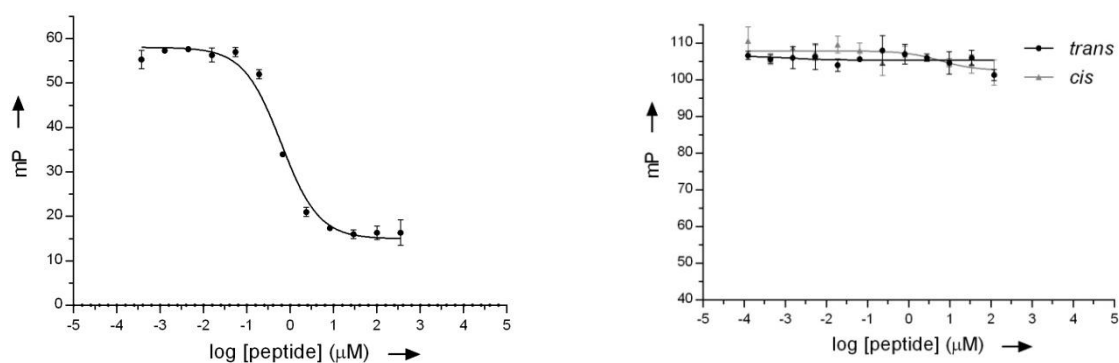


Figure S61. Dose-response curves of: left: **21** (SARAEVHLRKS-(RAhxR)₄RAhxB), right: peptide **4** (S-AMPB-ARAEVHLRKS).

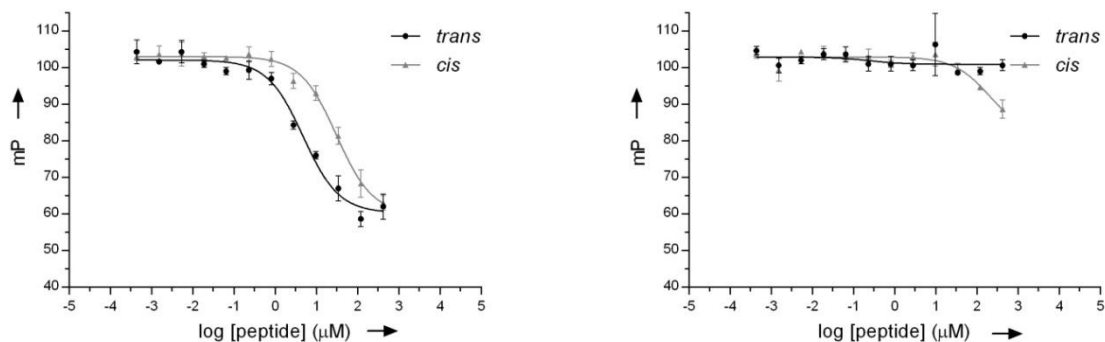


Figure S62. Dose-response curves of: left: peptide 5 (S-AMPB-RAEVHLRKS), right: peptide 6 (SAR-AMPB-EVHLRKS).

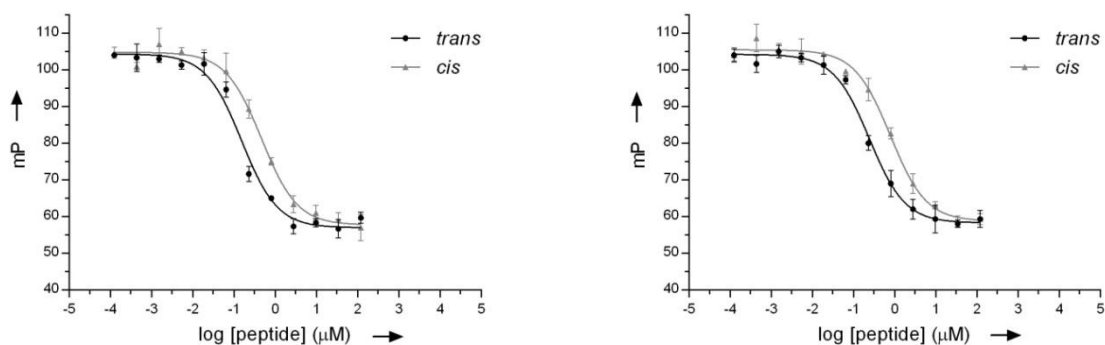


Figure S63. Dose-response curves of: left: peptide 7 (SARA-AMPB-VHLRKS), right: peptide 8 (SARAE-AMPB-HLRKS).

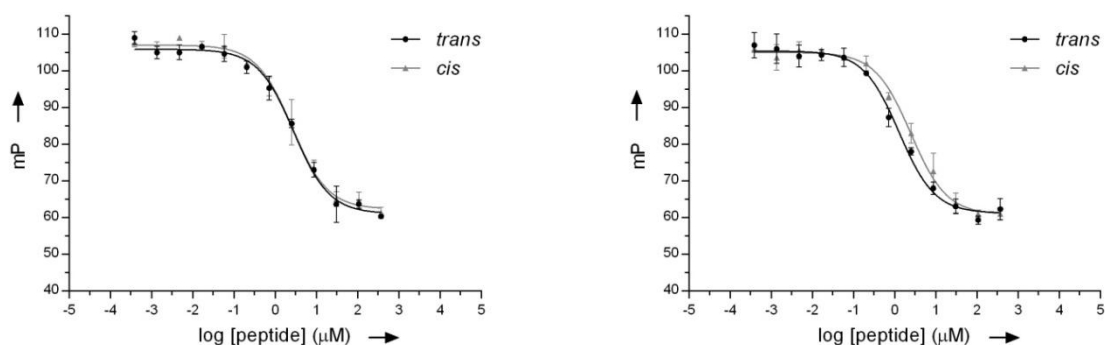


Figure S64. Dose-response curves of: left: peptide 9 (SARAEV-AMPB-LRKS), right: peptide 10 (SARAEVH-AMPB-RKS).

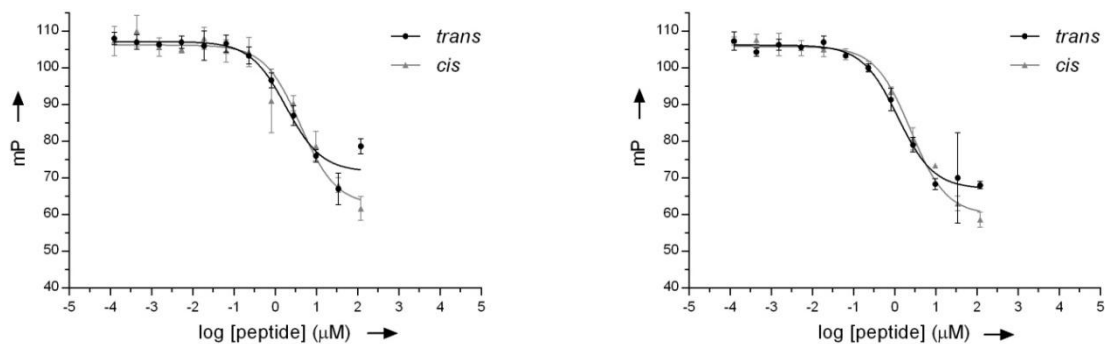


Figure S65. Dose-response curves of: left: peptide 11 (SARAEVHL-AMPB-KS), right: peptide 12 (SARAEVHLR-AMPB-S).

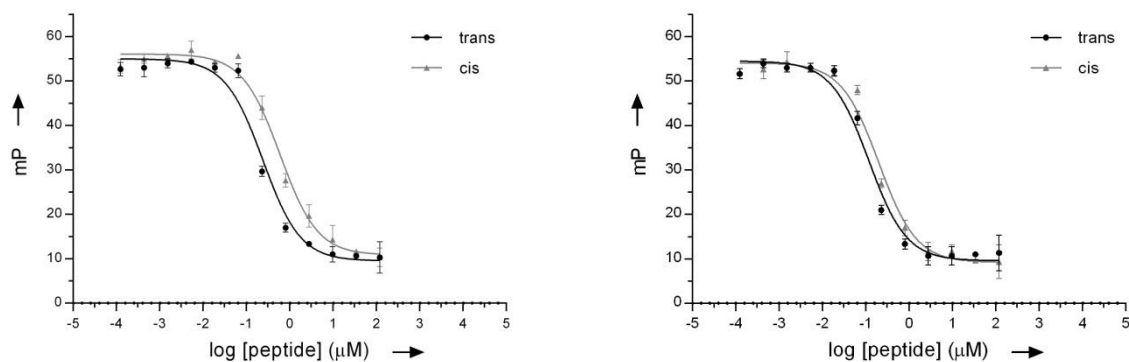


Figure S66. Dose-response curves of: left: peptide **13** (SARA-AMPB-HLRKS), right: peptide **15** (Ac-ART-AMPB-VY).

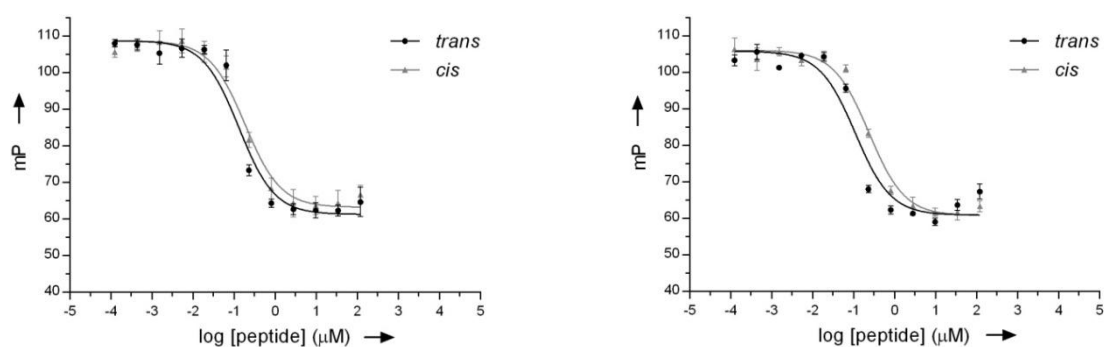


Figure S67. Dose-response curves of: left: peptide **16** (SARA-AMPB-VHLRKS(R)₈), right: peptide **17** (SARA-AMPB-VHLRKS(RAhxR)₄AhxB).

The mean and the standard deviation of the three independent measurements were calculated and are listed in the Table S2 to Table S4.

Table S2. Mean \pm SD of the IC₅₀ and K_i values of peptides **3**, **20** and **21**.

peptide	IC ₅₀ [μ M]	K _i [nM]
H ₂ N-SARAEVHLRKS 3	1.34 \pm 0.062	20.0 \pm 0.98
H ₂ N-SARAEVHLRKS-R ₈ 20	0.368 \pm 0.02	4.83 \pm 0.31
H ₂ N-SARAEVHLRKS- (RAhxB) ₄ RAhxB 21	0.612 \pm 0.02	8.67 \pm 0.32

Table S3. Mean \pm SD of the IC₅₀ and K_i values of the irradiated peptide **3**.

peptide	IC ₅₀ [μ M] (irr. at 430 nm)	IC ₅₀ [μ M] (irr. at 366 nm)	K _i [nM] (irr. at 430 nm)	K _i [nM] (irr. at 366 nm)
H ₂ N- SARAEVHLRKS 3	1.43 \pm 0.43	1.61 \pm 0.033	23.8 \pm 6.16	24.5 \pm 0.49

Table S4. Mean \pm SD of the IC₅₀ and K_i values of AMPB-containing peptides 4-17.

peptide	IC ₅₀ <i>trans</i> [μ M]	IC ₅₀ <i>cis</i> [μ M]	K _i <i>trans</i> [nM]	K _i <i>cis</i> [nM]	[<i>cis/trans</i>]
H ₂ N-S X ARA EVHLRKS 4	-	-	-	-	-
H ₂ N-S X RAEVHLRKS 5	3.91 \pm 0.83	24.9 \pm 5.7	66.8 \pm 10	433 \pm 73	6.48
H ₂ N-SAR X EVHLRKS 6	-	-	-	-	-
H ₂ N-SAR X VHLRKS 7	0.146 \pm 0.021	0.475 \pm 0.088	1.25 \pm 0.36	6.50 \pm 1.4	5.20
H ₂ N-SAR X AEVHLRKS 8	0.210 \pm 0.035	0.461 \pm 0.14	2.40 \pm 0.60	9.10 \pm 2.2	3.79
H ₂ N-SARAEV X LRKS 9	2.53 \pm 0.39	2.58 \pm 0.023	40.4 \pm 3.9	39.7 \pm 0.35	0.990
H ₂ N-SARAEVH X RKS 10	1.15 \pm 0.12	1.49 \pm 1.3	17.0 \pm 1.9	33.3 \pm 4.2	1.96
H ₂ N-SARAEVHL X KS 11	2.14 \pm 0.28	4.54 \pm 0.39	32.8 \pm 4.5	70.4 \pm 6.6	2.15
H ₂ N-SARAEVHLR X S 12	1.32 \pm 0.11	2.64 \pm 0.16	19.8 \pm 1.7	40.6 \pm 2.5	2.05
H ₂ N-SAR X AEVHLRKS 13	0.293 \pm 0.05	0.664 \pm 0.08	3.63 \pm 0.80	9.47 \pm 1.2	2.61
Ac-ART X VY 15	0.109 \pm 0.01	0.185 \pm 0.02	0.767 \pm 0.15	2.07 \pm 0.24	2.70
H ₂ N-SAR X AEVHLRKS-R ₈ 16	0.135 \pm 0.005	0.192 \pm 0.010	1.18 \pm 0.095	2.05 \pm 0.17	1.74
H ₂ N-SAR X AEVHLRKS- (RAhxR) ₄ AhxB 17	0.12 \pm 0.025	0.227 \pm 0.024	0.925 \pm 0.39	2.60 \pm 0.36	2.81

X = AMPB

For peptide **7** and additionally for peptide **14**, the exact same experiment was repeated just with lower final concentrations of WDR5 Δ 23 and the tracer peptide **22** (protein: 20 nM, tracer: 5 nM). This experiment was done, because the IC₅₀ value is depending on the assay conditions, while the K_i value should be in the same range, independently of the assay conditions.^[13] The measurement was also done in triplicate and the peptide was measured three times independently. Here, the graphs of one measurement for each peptide are exemplified (Figure S68).

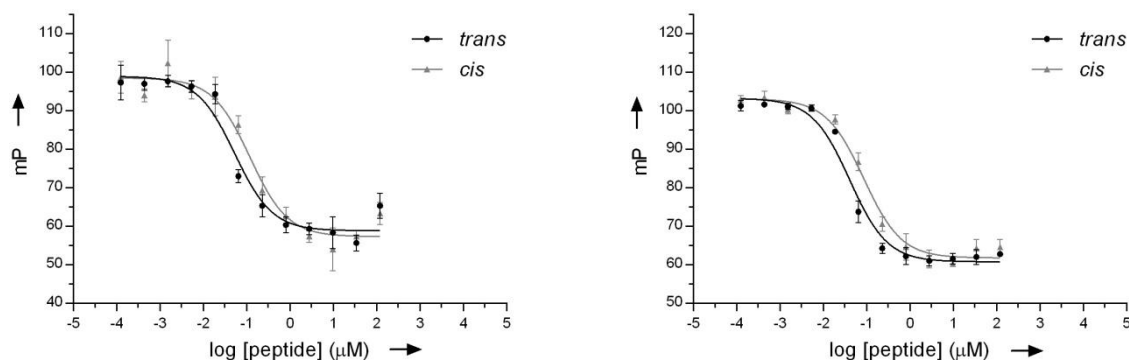


Figure S68. Dose-response curve of, left: peptide **7** (SARA-AMPB-VHLRKS), right: peptide **14** (Ac-ARA-AMPB-VHLRKS) with 5 nM of tracer.

The mean and the standard deviation of these three independent measurements were calculated and are listed in the Table S5.

Table S5. Mean \pm SD of the IC_{50} and K_i values of peptides **7** and **14** (with 5 nM tracer).

peptide	IC_{50} <i>trans</i> [μ M]	IC_{50} <i>cis</i> [μ M]	K_i <i>trans</i> [nM]	K_i <i>cis</i> [nM]	[<i>cis/trans</i>]
H ₂ N-SARA X VHLRKS 7	0.0570 ± 0.011	0.147 ± 0.014	2.03 ± 0.42	6.37 ± 1.0	3.14
Ac-ARA X VHLRKS 14	0.0450 ± 0.011	0.102 ± 0.015	1.50 ± 0.63	4.50 ± 0.82	3.00

X = AMPB, B = β -alanine

Crystal Structure of WDR5 Δ 23-Peptide **7** Complex

The expressed and purified WDR5 Δ 23 was mixed in a 4:1 ratio separately with two stock solutions (55.7 mM): *trans/cis* ratio (96:4) and *cis/trans* ratio (81:19) of the peptide **7** (H₂N-SARA-AMPB-VHLRKS-CONH₂), what we called ‘state I’ and ‘state II’, giving final concentrations of WDR5 = 229.62 μ M (8.5 mg/mL) and peptide **7** = 918.48 μ M. Previously, the two isomers of peptide **7** were dissolved in DMSO (55.7 mM) and 2.3 μ L of these stock solutions were added to 139.7 μ L of the protein in dialysis buffer (25 mM Tris, pH = 8.0, 150 mM NaCl). In the ‘state I’ the peptide solution was used without any irradiation while in the ‘state II’, the solution of peptide **7** was irradiated at 366 nm for 14 h before adding it to the protein and kept in the dark at all times during the crystallization experiments.

Crystallization trials were attempted with available commercial screening solutions with 300nl drop volume of protein-peptide complex at a concentration of 8.5 mg/mL protein and 918.48 μ M peptide **7** at 18°C with a Honeybee robot (Isogen Life Science). The hanging drop

vapor diffusion method was used for crystallization and crystals were obtained with 12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD, 0.02 M of each carboxylic acid, 0.1 M bicine/Trizma base pH 8.5 (Morpheus screen, Molecular Dimensions). Crystals were flash-frozen directly in the mother liquor solution and no cryoprotectant was necessary as the crystals grew in a cryoprotected condition. All manipulations for the dark state crystals were done under red light.

Diffraction data were collected at the Swiss Light Source (SLS) on the microfocus beamline X06SA. All data were processed using XDS^[14] in space group P21212. Data reduction and scaling was done using the CCP4 suite of programs.^[15] Data collection and processing statistics are given in Table S6. The structure of the complex of both states of the AMPB-containing peptide **7** in complex with WDR5 Δ 23 was determined by molecular replacement with PHASER^[16] using the coordinates of the previously determined structure of the WDR5 Δ 23-MLL1 WIN peptide complex (PDB code: 3EG6) as a search model. After an initial rigid body refinement, the structure was further refined with rounds of simulated annealing, energy minimization, and individual *B*-factor refinement with a maximum likelihood target using Phenix.^[17] Difference Fourier maps were calculated with Phenix and used to locate electron density corresponding to bound peptide, and the structure was built using COOT.^[18] PRODRG^[19] was used to generate geometry restraints for the AMPB-containing peptide refinement and Molprobity^[20] for structural validation. Final refinement statistics are given in Table S6. All structural figures were generated with Pymol.^[21] The coordinates and structure factors (PDB code 5M23 for state I WDR5-AMPB-containing peptide complex and PDB code 5M25 for state II complex) have been deposited to the Protein Data Bank.

Table S6. X-ray data collection and refinement statistics.

	WDR5 Δ 23-AMPB-containing peptide 7 ('light state')	WDR5 Δ 23-AMPB-containing peptide 7 ('dark state')
Wavelength (Å)	0.9788	0.9788
Resolution range (Å)	43.05 - 1.97 (2.11-1.97)	42.96 - 2.43 (2.63-2.43)
Space group	P 21 21 2	P 21 21 2
Unit cell (Å, Å, Å °, °, °)	80.553 86.098 40.407 90.00 90.00 90.00	80.553 86.098 40.407 90.00 90.00 90.00
Total reflections	258124	112937
Unique reflections	20077	17802
Multiplicity	12.7 (12.1)	6.3 (6.3)
Completeness (%)	99.78 (94.4)	98.8 (99.0)
Mean I/sigma(I)	14.4 (5.0)	8.5 (2.2)
Rmerge	0.070 (0.237)	0.090 (0.253)
R-factor	0.1664	0.1795
R-free	0.2143	0.2482
Number of atoms	2471	2469
Protein	2326	2325
ligands	51	44
RMS(bonds)	0.015	0.008
RMS(angles)	1.488	1.196
Ramachandran favored (%)	95.0	94.3
Ramachandran outliers (%)	0.0	0.0
Average B-factor	38.56	15.78

Figure S69 shows an overlay of the whole structures of the superimposed state I and state II of peptide 7.

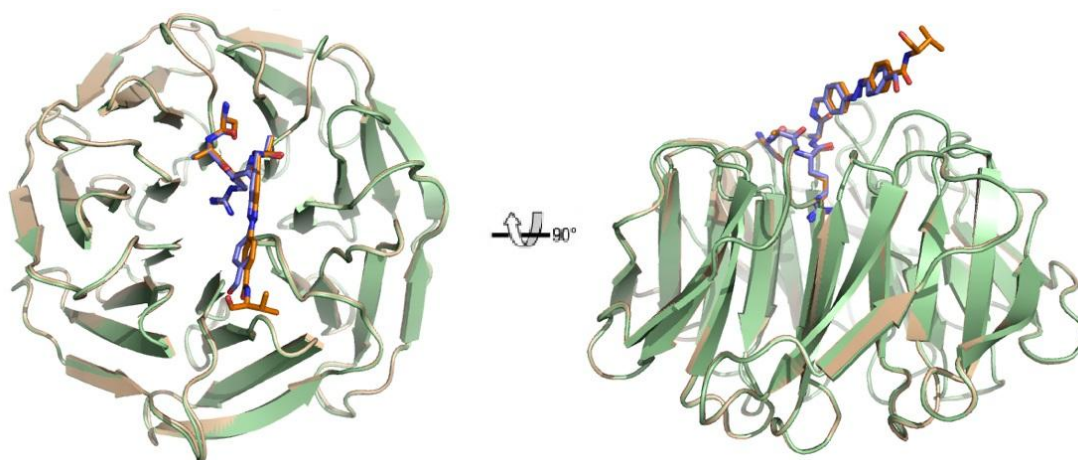


Figure S69. Superposition of the crystal structure of the WDR5 Δ 23-peptide 7 complex in the state I (peptide: orange, protein: wheat) overlaid with WDR5 Δ 23-peptide 7 complex in the state II (peptide: blue, protein: green).

The key interactions of peptide 7 with the WDR5 protein in both states agree to the ones seen in the crystal structure of the WIN peptide in complex with WDR5 (PDB code 3EG6).^[22,23] Thus, the hydrogen bonds between the Ser and Ala residue of the peptide 7 to Asp-107, as well as the one between the Ser and the Arg main chain residues of the peptide 7 with Ser-91 from WDR5, as well as two indirect hydrogen bonds, mediated by one water molecule, to Val-132 and Asp-107 could be found in our protein-peptide 7 complex. Also, the aromatic side chain of Tyr-260 participates in a water-mediated hydrogen bond to the second Ala residue. Furthermore, the crystal structure confirmed that the Arg side chain of peptide 7 inserts into a central tunnel of the WD40 β -propeller. The guanidinium moiety is sandwiched between the conserved aromatic side chains of Phe-133 and Phe-263, which is stabilized by a cation- π -interaction. Further stabilization is provided by a substantial network of direct and indirect (water mediated) hydrogen bonds between the Arg residue and the amino acids Ser-91, Phe-133, Ser-175, Ser-218 and Cys-261 of WDR5. In the C-terminus, the overlay of the superposition of the the state I AMPB-peptide and the WIN peptide structures shows quite a difference in the location of the two respective peptides. While the His-3769 of the WIN peptide points towards the Phe-149, Asp-172, Pro-173, and Tyr-191 residues of the protein,^[22] the azobenzene moiety of our peptide is mostly located outside of the protein structure (Figure 3 in the manuscript). All the interactions between both, the state of peptide 7 with WDR5 Δ 23 are shown in Figure S70 and Figure S71 and are additionally listed in Table S7 and Table S8.

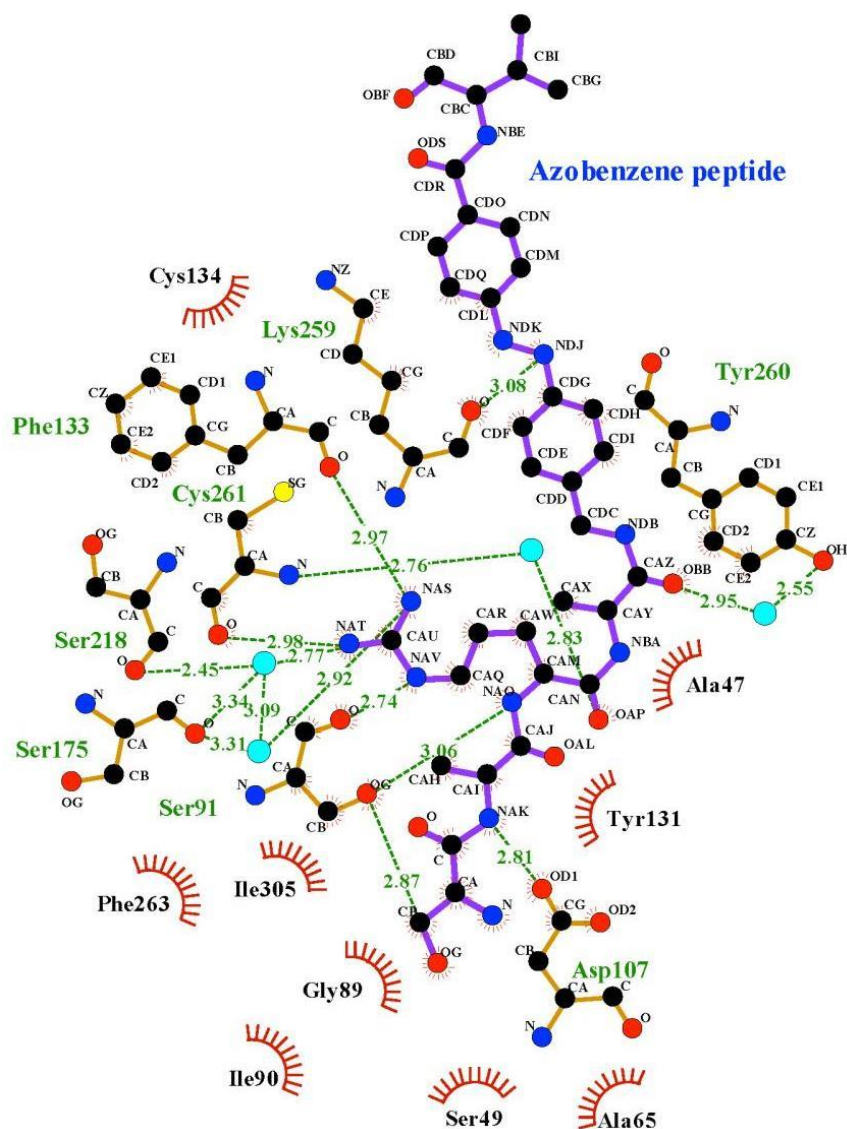


Figure S70. Schematic representation of the key interactions between the AMPB-peptide **7** in the state I and WDR5 Δ 23. Generated with LIGPLOT.^[24]

Table S7. Interactions between the AMPB-peptide **7** in the state I and WDR5 Δ 23. Interactions marked in red, are strong H-bonds.

peptide		Hydrogen-bond/ Van der Waals partner		
Residue	Atom	WDR5 residue	Atom	Distance (Å)
AMPB5	CDQ	Lys259	O	3.75
AMPB5	CDL	Lys259	CG	3.90
AMPB5	NDK	Lys259	CG	3.69
AMPB5	NDJ	Lys259	O	3.08
			CG	3.17
			CE	3.82
AMPB5	CDG	Lys259	O	3.88
			CG	3.93
			CE	3.83
AMPB5	CDH	Lys259	O	3.81
		Tyr260	CE2	3.82

		Tyr260	CD2	3.43
AMPB5	CDI	Tyr260	CE2	3.34
		Tyr260	CD2	3.59
AMPB5	CDF	Lys259	CE	3.74
AMPB5	CDD	Tyr260	CE2	3.93
Ala4	OBB	Tyr260	CE2	3.76
		Wat86 (bridge to Tyr260 OH)	O	2.95
Ala4	CAX	Leu321	CD2	3.93
		Ala47	CB	3.47
Arg3	OAP	Wat12 (bridge to Cys261 N)	O	2.83
		Tyr260	CD2	3.86
Arg3	CAM	Phe133	CE1	3.26
		Phe133	CZ	3.60
		Ser91	OG	3.92
Arg3	CAW	Phe133	CE1	3.61
		Phe133	CZ	3.70
		Ser91	CB	3.76
		Ser91	OG	3.54
Arg3	CAR	Cys261	O	3.81
		Phe133	CZ	3.96
		Ile305	CG1	3.82
		Ile305	CD1	3.90
Arg3	CAQ	Cys261	O	3.65
		Phe263	CE1	3.77
		Ile305	CG1	3.80
		Ser49	O	3.27
		Ser91	O	3.60
Arg3	NAV	Phe263	CE1	3.61
		Phe263	CZ	3.97
		Ser49	O	3.98
		Ser91	C	3.86
		Ser91	O	2.74
Arg3	CAU	Phe133	O	3.94
		Phe133	CE2	3.70
		Phe263	CE1	3.44
		Phe263	CZ	3.48
		Ser91	O	3.67
Arg3	NAT	Cys261	C	3.87
		Cys261	O	2.98
		Cys261	CB	3.70
		Wat33 (bridge to Ser218 O, Ser175 O)	O	2.77
		Phe133	CE2	3.85
		Phe263	CD1	3.94
		Phe263	CE1	3.40
		Phe263	CZ	3.68

Arg3	NAS	Wat45 (bridge to Ser91 O)	O	2.92
		Phe133	C	3.93
		Phe133	O	2.97
		Phe133	CE2	3.77
		Cys134	CB	3.83
		Ser175	O	3.93
		Phe263	CZ	3.62
		Ser91	O	3.77
Arg3	NAO	Phe133	CD1	3.98
		Phe133	CE1	3.24
		Ser91	OG	3.06
Ala2	CAJ	Phe133	CE1	3.79
Ala2	CAI	Asp107	OD1	3.74
Ala2	CAH	Phe133	CD1	3.92
		Phe133	CE1	3.98
		Tyr131	CD1	3.75
		Asp107	OD1	3.49
Ala2	NAK	Ser91	OG	3.21
		Asp107	CG	3.34
		Asp107	OD1	2.81
		Asp107	OD2	3.09
Ser1	C	Ser91	OG	3.79
		Asp107	CG	3.81
		Asp107	OD1	3.66
		Asp107	OD2	3.20
Ser1	CA	Ser91	OG	3.89
		Gly89	C	3.79
		Gly89	CA	3.93
		Asp107	CG	3.71
		Asp107	OD1	3.73
		Asp107	OD2	3.20
Ser1	N	Gly89	C	3.88
		Gly89	CA	3.46
		Asp107	CG	3.93
		Asp107	OD2	3.09
		Ala65	CA	3.89
		Ala65	CB	3.85
Ser1	CB	Ser91	OG	3.92
		Ile90	C	3.87
		Ile90	O	3.47
		Ala65	CA	3.69
		Ala65	CB	3.82
Ser1	OG	Ser49	CB	3.68
		Ser91	CA	3.49
		Ser91	CB	3.70
		Ser91	OG	2.87
		Ile90	C	3.78
		Ile90	O	3.40
		Ser91	N	3.86

H-bond cut off < 3.5 Å, Van der Waals: 3.6-4.0 Å.

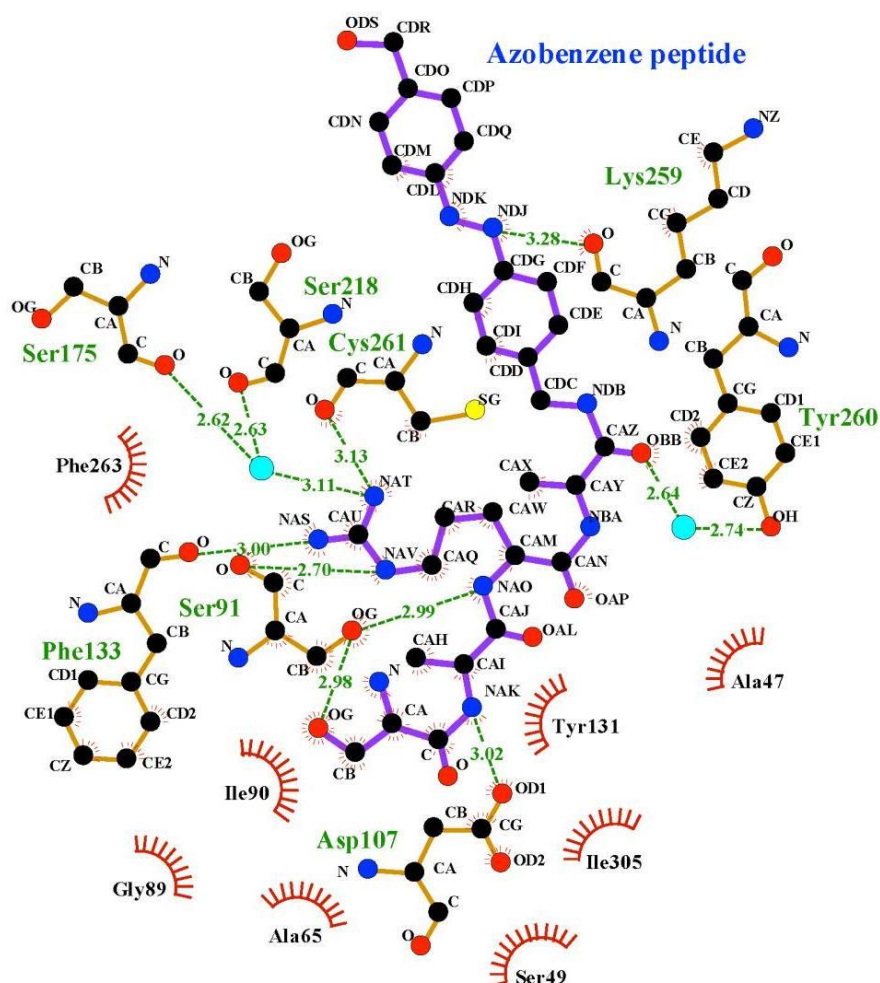


Figure S71. Schematic representation of the key interactions between the AMPB-peptide 7 in the state II and WDR5 Δ 23. Generated with LIGPLOT.^[24]

Table S8. Interactions between the AMPB-peptide 7 in the state II and WDR5 Δ 23. Interactions marked in red, are strong H-bonds.

peptide		Hydrogen-bond partner		
Residue	Atom	WDR5 residue	Atom	Distance (Å)
AMPB5	CDM	Lys259	O	3.63
AMPB5	CDL	Lys259	O	3.98
		Lys259	CG	3.67
AMPB5	NDK	Lys259	O	3.91
		Lys259	CG	3.35
		Lys259	CE	3.69
AMPB5	NDJ	Lys259	O	3.28
			CG	3.54
AMPB5	CDH	Tyr260	CE2	3.53
		Tyr260	CD2	3.72
AMPB5	CDI	Tyr260	CE2	3.49
		Tyr260	CD2	3.72
Ala4	OBB	Tyr260	CE2	3.65
		Wat71 (bridge to	O	2.64

Tyr260 OH)				
Ala4	CAX	Ala47	CB	3.82
Arg3	OAP	Tyr260	CD2	3.77
Arg3	CAM	Phe133	CE1	3.42
		Phe133	CZ	3.74
		Ser91	OG	3.87
Arg3	CAW	Phe133	CE1	3.57
		Phe133	CZ	3.60
		Ser91	CB	3.84
		Ser91	OG	3.50
Arg3	CAR	Cys261	O	3.99
		Ile305	CG1	3.78
		Ile305	CD1	3.71
Arg3	CAQ	Cys261	O	3.78
		Phe263	CE1	3.82
		Ile305	CG1	3.68
		Ser49	O	3.33
		Ser91	O	3.66
Arg3	NAV	Phe263	CE1	3.69
		Phe263	CZ	3.99
		Ser49	O	3.99
		Ser91	C	3.75
		Ser91	O	2.70
Arg3	CAU	Phe133	CE2	3.63
		Phe263	CE1	3.50
		Phe263	CZ	3.47
		Ser91	O	3.54
Arg3	NAT	Cys261	O	3.13
		Cys261	CB	3.73
		Wat27 (bridge to Ser218 O, Ser175 O)	O	3.11
		Phe133	CE2	3.65
		Phe263	CE1	3.48
		Phe263	CZ	3.63
Arg3	NAS	Asp92	OD2	3.92
		Phe133	C	3.91
		Phe133	O	3.00
		Phe133	CE2	3.60
		Phe133	CD2	3.62
		Phe263	CZ	3.58
		Ser91	O	3.56
Arg3	NAO	Phe133	CE1	3.43
		Ser91	OG	2.99
Ala2	CAJ	Ser91	OG	3.97
Ala2	CAI	Asp107	OD1	3.82
		Asp107	OD2	3.77
Ala2	CAH	Tyr131	CD1	3.78
		Asp107	OD1	3.45
Ala2	NAK	Ser91	OG	3.26
		Asp107	CG	3.32

		Asp107	OD1	3.02
		Asp107	OD2	2.83
Ser1	C	Ser91	OG	3.77
		Asp107	CG	3.95
		Asp107	OD1	3.95
		Asp107	OD2	3.15
Ser1	CA	Ser91	OG	3.87
		Asp107	CG	3.81
		Asp107	OD1	3.97
		Asp107	OD2	3.10
Ser1	N	Gly89	CA	3.69
		Asp107	OD2	3.13
		Ala65	CA	3.95
Ser1	CB	Ala65	CB	3.80
		Ser91	OG	3.87
		Ile90	N	3.94
		Ile90	C	3.63
		Ile90	O	3.07
		Ala65	CA	3.84
Ser1	OG	Ala65	CB	3.93
		Ser49	CB	3.45
		Ser91	CA	3.41
		Ser91	CB	3.73
		Ser91	OG	2.98
		Ile90	C	3.66
		Ile90	O	3.10
		Ser91	N	3.81

H-bond cut off < 3.5 Å, Van der Waals: 3.6-4.0 Å.

Evaluation of Stability against GSH-reduction of Peptides

To test the stability of the AMPB-containing peptides **16** and **17** against reduction by glutathione (GSH), the peptides were incubated with GSH for 4 days.

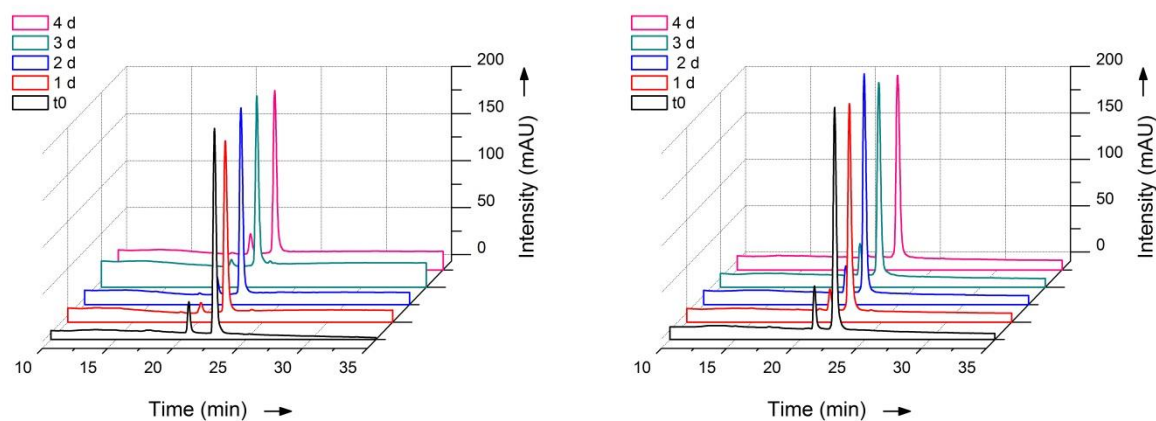


Figure S72. Left: HPLC chromatograms of peptides, incubated with 10 mM GSH, recorded after different time periods. Left: peptide **16**. Right: peptide **17**.

Therefore, 56 μL of a stock solution of GSH (0.25 M) and 1344 μL of a stock solution of peptides **16** and **17** (10 μM) in 0.1 M phosphate buffer pH = 7 were mixed giving final concentrations of 10 mM of GSH and 9.6 μM of the peptides, respectively. After different time periods, HPLC chromatograms of these mixtures were recorded (gradient: 5-40% B) (Figure S72).

In Vitro Methyl Transferase Assay with MLL1 Core Complex

Protein Expression

Full length constructs of both, RbBP5 (residues 1–538) and Ash2L (residues 1–635), were used for expression. Truncated WDR5 for robust HMT activity of the core complex, were used. MLL1, WDR5, RbBP5, and ASH2L were expressed as His-SUMO fusions from the pET28A-SUMO vector in BL21 DE3 pLyss codon (+) cells at 16 °C overnight. Expression was started with 0.1 mM IPTG in the mid-log phase of bacterial growth. For each protein, cells were harvested and the protein was purified by the His tag on Ni-NTA resin (Qiagen). The SUMO tag was removed from RbBP5, ASH2L, and MLL1 proteins by incubation with the ULP1 protease at 4 °C overnight. The protease and cleaved SUMO-His tag were collected by batch binding with the Ni-NTA resin for 1 h.

In Vitro Histone Methyltransferase (HMT) Assay

The HMT assay was performed, as previously described^[25] with slight modifications, in 50 mM Tris pH 8.0, 50 mM NaCl, 1.0 mM DTT, 5 mM MgCl_2 at room temperature. Each reaction contained 1.5 μM of the cofactor ^3H -S-adenosylmethionine (Perkin-Elmer). H3₁₋₂₁ peptide was used as the substrate at 250 μM . Both, the *trans* and the *cis* (*cis*: previously irradiated at 366 nm for 1 h) isomer, of peptide **7** were added at concentrations ranging from 0.705 nM to 70.5 μM and incubated with the preassembled MLL1 complex at a final concentration of 0.3 μM for 10 min on ice. Reactions were initiated by addition of the H3/SAM and allowed to proceed for 60 min before preparing scintillation counting. To count samples, reactions were spotted on separate squares of P81 filter paper (Millipore) and precipitated by submerging in freshly prepared 50 mM sodium bicarbonate buffer pH = 9.0. After washing and drying, the samples were vortexed in Ultima Gold scintillation fluid and counted.

Proliferation Assay

Previously to the cell viability assays, it was tested for how long the cells could be irradiated with the custom made 6-well LED array (430 nm), without damaging. Therefore, MLL-AF9-transduced mouse bone marrow cells were irradiated for different time periods (Figure S73).

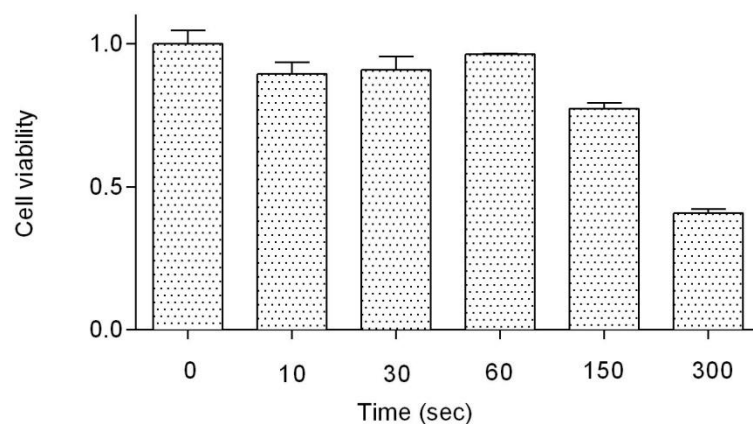


Figure S73. Cell viability after MLL-AF9-transduced mouse bone marrow cells were irradiated at 430 nm with the custom made 6-well LED array for different time periods.

MLL-AF9-transduced mouse bone marrow was cultured in IMDM medium + 15% FBS and supplemented with 10 ng/mL interleukin-3 (IL-3) at each passage, every other day.

Both, the *trans* and the *cis* (*cis*: previously irradiated at 366 nm for 4 h at 100 mM stock in DMSO) isomers, of peptides **7**, **16**, **17**, as well as the peptides **18** and **19** were diluted from stock to culture media containing a 0.1% DMSO final concentration. For viability assays, cells were cultured at 1×10^4 /mL in 6-well plates with 2.5% FBS and treated with each peptide at 1.0, 1.6, 2.5, 5.0, 10, 25, 50 and 100 μ M for 24 h and then added FBS up to 15% for 72 h and DMSO as control. For the *cis-trans* group, after 48 h treatment with *cis* peptides, the plate was irradiated at 430 nm for 90 sec and continued to culture for 2 days in the dark. Viability was determined using the CellTiter-Glo Kit (Promega) according to the manufacturer's directions. Luminescence was monitored on a Molecular Devices plate reader.

Real Time Quantitative PCR Analysis of Deptor Gene

Murine MLL1-AF9 transformed bone marrow cells were obtained by transducing normal murine bone marrow cells with MLL1-AF9 oncogene according to the procedures described by Tan et al.^[26] Both, the *trans* and the *cis* (*cis*: previously irradiated at 366 nm for 4 h for 100 mM stock) isomers, of peptides **7**, **16**, **17**, as well as the peptides **18** and **19** were dissolved in DMSO.

The transformed cells were treated with different concentrations of the respective peptide (1.0, 5.0, 10 μ M) and Mock (0.1% DMSO), giving a final concentration of 0.1% DMSO in all the samples. Total RNA was isolated from MLL1-AF9 transduced mouse bone marrow cells after 4 day treatment, using Trizol (Invitrogen) and the RNEASY kit (Qiagen) according to the protocol described earlier.^[27] The cDNA was generated using random priming with the SuperScript III kit (Invitrogen). Real-time PCR amplifications of Deptor and GAPDH genes were carried out with primers specific for each gene in the presence of SYBR dye. Relative quantification of each gene transcript was carried out as described previously.^[28] The results are presented in Fig8 in the manuscript as relative expression to Mock treatment after normalizing to an internal loading control (e.g. GAPDH).

References

- [1] W. J. Fang, T. Yakovleva, J. V. Aldrich, *Biopolymers* **2011**, *96*, 715–722.
- [2] W. S. Hancock, J. E. Battersby, *Anal. Biochem.* **1976**, *71*, 260–264.
- [3] N. Thieriet, J. Alsina, E. Giralt, F. Guib??, F. Albericio, *Tetrahedron Lett.* **1997**, *38*, 7275–7278.
- [4] P. Grieco, P. M. Gitu, V. J. Hruby, *J. Pept. Res.* **2001**, *57*, 250–256.
- [5] C. Renner, J. Cramer, R. Behrendt, L. Moroder, *Biopolymers* **2000**, *54*, 489–500.
- [6] M. Löweneck, A. G. Milbradt, C. Root, H. Satzger, W. Zinth, L. Moroder, C. Renner, *Biophys. J.* **2006**, *90*, 2099–2108.
- [7] H. Rau, *Stud. Org. Chem. Photochromism, Mol. Syst.* **1990**, *40*, 165–192.
- [8] D. A. James, D. C. Burns, G. A. Woolley, *Protein Eng.* **2001**, *14*, 983–991.
- [9] S.-L. Dong, M. Löweneck, T. E. Schrader, W. J. Schreier, W. Zinth, L. Moroder, C. Renner, *Chem. Eur. J.* **2006**, *12*, 1114–1120.
- [10] B. Priewisch, K. Rück-Braun, *J. Org. Chem.* **2005**, *70*, 2350–2352.
- [11] R. Sjöback, J. Nygren, M. Kubista, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **1995**, *51*, L7–L21.
- [12] Invitrogen Corporation, *Technical Resource Guide, Fluorescence Polarization, 4th Edition*, Madison, **2005**.
- [13] Z. Nikolovska-Coleska, R. Wang, X. Fang, H. Pan, Y. Tomita, P. Li, P. P. Roller, K. Krajewski, N. G. Saito, J. a. Stuckey, et al., *Anal. Biochem.* **2004**, *332*, 261–273.
- [14] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132.
- [15] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, et al., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 235–242.
- [16] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- [17] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221.
- [18] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- [19] A. W. Schüttelkopf, D. M. F. Van Aalten, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 1355–1363.
- [20] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 12–21.
- [21] W. L. DeLano, *The PyMOL User's Manual*, DeLano Scientific, **2002**.
- [22] A. Patel, V. Dharmarajan, M. S. Cosgrove, *J. Biol. Chem.* **2008**, *283*, 32158–32161.

- [23] J. J. Song, R. E. Kingston, *J. Biol. Chem.* **2008**, *283*, 35258–35264.
- [24] A. C. Wallace, R. A. Laskowski, J. M. Thornton, *Protein Eng.* **1995**, *8*, 127–134.
- [25] F. Cao, E. C. Townsend, H. Karatas, J. Xu, L. Li, S. Lee, L. Liu, Y. Chen, P. Ouillette, J. Zhu, et al., *Mol. Cell* **2014**, *53*, 247–261.
- [26] J. Tan, M. Jones, H. Koseki, M. Nakayama, A. G. Muntean, I. Maillard, J. L. Hess, *Cancer Cell* **2011**, *20*, 563–575.
- [27] X. Li, L. Li, R. Pandey, J. S. Byun, K. Gardner, Z. Qin, Y. Dou, *Cell Stem Cell* **2012**, *11*, 163–178.
- [28] Y. Dou, T. A. Milne, A. J. Ruthenburg, S. Lee, J. W. Lee, G. L. Verdine, C. D. Allis, R. G. Roeder, *Nat. Struct. Mol. Biol.* **2006**, *13*, 713–719.

4.2.1 Author Contribution

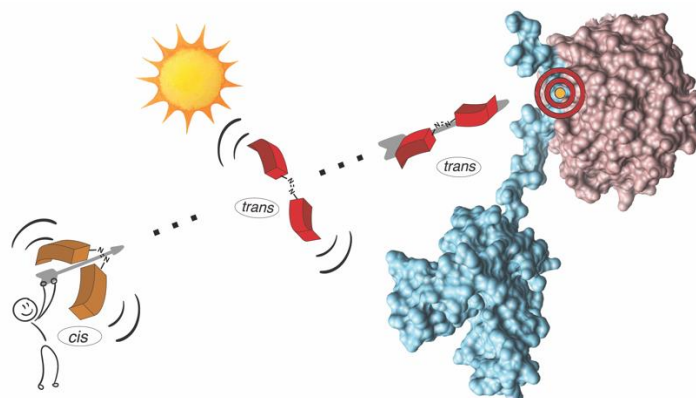
The idea of using photoresponsive peptides to regulate epigenetic states via the disruption of PPIs within a multi protein complex was envisioned by Jun.-Prof. Dr. Olalla Vázquez. I, Lea Albert, performed the synthesis, purification and characterization of the photoswitches as well as the photoactivatable peptide library. In addition, I performed the analysis of the photochromic properties of the photoswitchable peptides via UV-vis spectroscopy and HPLC-MS, and the establishment and interpretation of the fluorescence polarization (FP)-based assays. I expressed, purified and verified the correct folding of the WDR5 protein. In collaboration with Dr. Vasundara Srinivasan, I participated in setting up the crystallization trials as well as initial analysis. Crystal structure determinations, and corresponding figures and discussion of this part in the article was mainly performed by Dr. Vasundara Srinivasan. Biological evaluation of the photoresponsive peptides was performed in collaboration with the laboratory of Prof. Dr. Yali Dou. Thus, histone-methyl transferase assays, cell viability assays and RT-qPCR experiments were carried out by Dr. Jing Xu and Ruiwei Wan. I contributed to write the article. Jun.-Prof. Dr. Olalla Vázquez conceived of the idea for the study, designed, directed and interpreted experiments, wrote the manuscript and had the role of corresponding author, i.e., she was leading and/or involved in all the 14-role taxonomy for author-contribution assignment.¹

Marburg,

Jun.-Prof. Dr. Olalla Vázquez

Lea Albert

4.3 Modulating Protein-Protein Interactions with Visible-Light Responsive Peptide Backbone Photoswitches



Bright shot! Optical control of enzymatic activity by using visible-light photoswitchable protein–protein modulators. The synthesis and characterization of two visible-light-responsive peptide backbone photoswitches, based on azobenzene derivatives, are reported.

This chapter was published as:

L. Albert, A. Peñalver, N. Djokovic, L. Werel, M. Hoffarth, D. Ruzic, J. Xu, L.-O. Essen, K. Nikolic, Y. Dou, O. Vázquez*,
Modulating Protein-Protein Interactions with Visible-Light Responsive Peptide Backbone Photoswitches
ChemBioChem **2019**, *20*, 1417-1429.

Copyright John Wiley & Sons, Inc. The license for reprinting this article is attached at the end of the thesis.

Our manuscript has been highlighted as front cover of the special themed collection 2019 ChemBioTalents.



In the following sections, the molecules included in this manuscript, will be named as **MC-X**, in which MB means manuscript C, and **X** means number of compound as given in this manuscript.



Modulating Protein–Protein Interactions with Visible-Light-Responsive Peptide Backbone Photoswitches

Lea Albert,^[a] Alberto Peñalver,^[a] Nemanja Djokovic,^[b] Laura Werel,^[a] Malte Hoffarth,^[a] Dusan Ruzic,^[b] Jing Xu,^[c] Lars-Oliver Essen,^[a] Katarina Nikolic,^[b] Yali Dou,^[c] and Olalla Vázquez*^[a]

Life relies on a myriad of carefully orchestrated processes, in which proteins and their direct interplay ultimately determine cellular function and disease. Modulation of this complex crosstalk has recently attracted attention, even as a novel therapeutic strategy. Herein, we describe the synthesis and characterization of two visible-light-responsive peptide backbone photoswitches based on azobenzene derivatives, to exert optical control over protein–protein interactions (PPI). The novel peptidomimetics undergo fast and reversible isomerization with low photochemical fatigue under alternatively blue-/

green-light irradiation cycles. Both bind in the nanomolar range to the protein of interest. Importantly, the best peptidomimetic displays a clear difference between isomers in its protein-binding capacity and, in turn, in its potential to inhibit enzymatic activity through PPI disruption. In addition, crystal structure determination, docking and molecular dynamics calculations allow a molecular interpretation and open up new avenues in the design and synthesis of future photoswitchable PPI modulators.

Introduction

Proteins play a key role in the regulation of cellular behavior and are, therefore, also involved in the development of diseases. Despite the fact that some proteins perform their functions independently, the reality is that most of them are integrated in complex dynamic networks ruled by protein–protein interactions (PPIs).^[1] This crosstalk controls both protein production and activity, as well as signal transduction and metabolic pathways. Therefore, breaking the molecular code associated with PPIs will pave the way for on-demand modulation of functional outcomes capable of affecting pathogenic mechanisms. Historically, PPIs with relatively large buried interfaces have been considered “undruggable” targets.^[2] However, recent years have witnessed an outstanding success in targeting PPIs through many different approaches, such as small-molecule derivatives, recombinant proteins, antibodies, and peptides.^[3]

In particular, the use of peptidomimetics has acquired increasing relevance because they not only retain the advantages of peptides, but also overcome their intrinsic limitations to provide new features.^[4] For example, by grafting a light-driven molecular transducer onto a peptide scaffold, spatiotemporal resolution could be achievable to circumvent the common off-target effects in therapy.

Azobenzenes^[5] are by far the most extensively used photoswitches in the context of photopharmacology.^[6] However, their *trans*→*cis* isomerization wavelength at 366 nm is not ideal for in vivo applications and, furthermore, this photoisomerization is not complete. Therefore, in the last few years, there has been great interest in developing visible-light-shifting azobenzenes to overcome such limitations.^[7] Among these second-generation compounds, the cyclic azobenzene (cAzo) 5,6-dihydrodibenzo[*c,g*][1,2]diazocine is a very interesting one.^[8] The C2 bridge produces a highly twisted *trans* isomer that is less stable than the *cis* conformation; this contrasts with the situation in “normal” azobenzenes. *p*-Acetoamido substituents in the cAzo core enable slightly red-shifted isomerization wavelengths.^[9] More importantly, the separation of the $n\rightarrow\pi^*$ transition is large (85 nm), so that selective irradiation is possible. The same behavior was also observed with tetra-*ortho*-substituted azobenzenes.^[10] In particular, the group of Hecht has optimized the properties of classical azobenzenes by introducing σ -electron-withdrawing fluorine atoms in these positions, which leads to visible-light switches with high photoconversions and very long lived *cis* isomers.^[7c] Since this contribution, tetra-*ortho*-fluoroazobenzenes have been mainly used in materials science,^[11] and their applications in a biological context are scarce.^[12] Along these lines, there have been only a

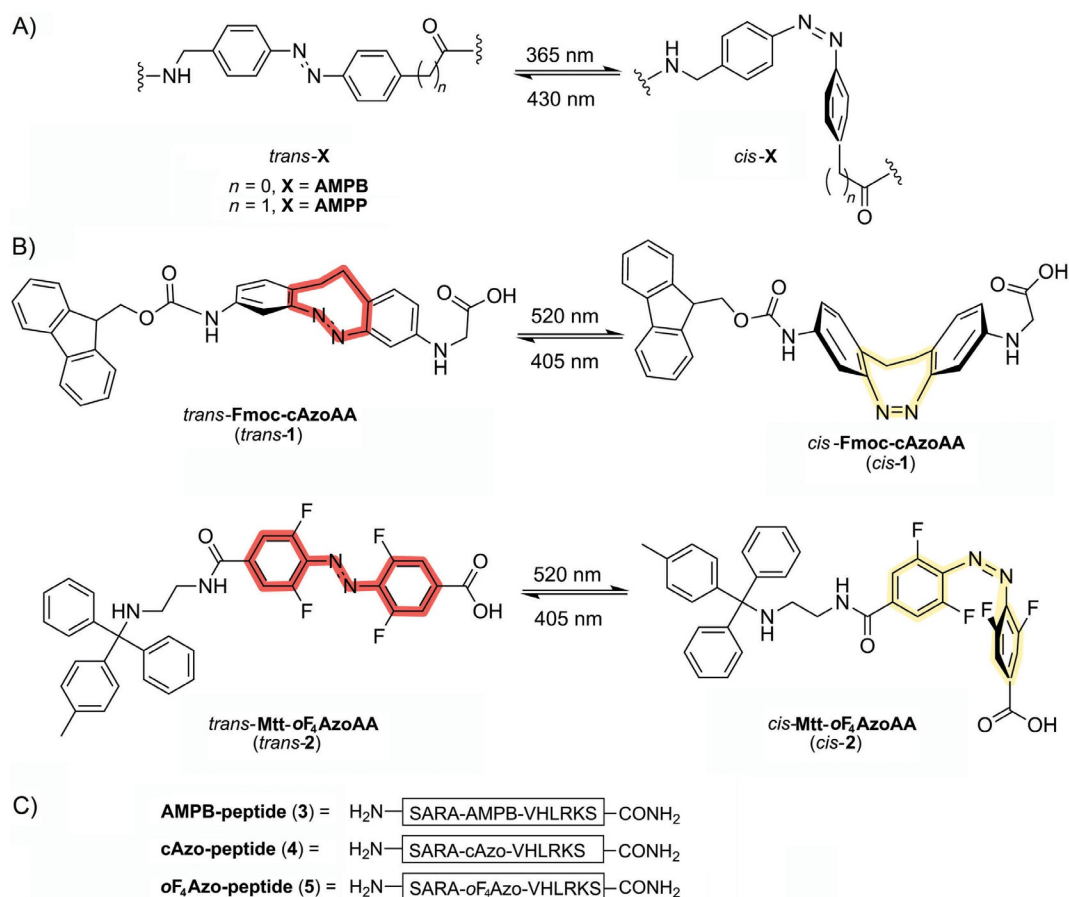
[a] L. Albert, A. Peñalver, L. Werel, M. Hoffarth, Prof. Dr. L.-O. Essen, Prof. Dr. O. Vázquez
Fachbereich Chemie, Philipps-Universität Marburg
Hans-Meerwein-Strasse 4, 35043, Marburg (Germany)
E-mail: olalla.vazquez@staff.uni-marburg.de

[b] N. Djokovic, D. Ruzic, K. Nikolic
Department of Pharmaceutical Chemistry
Faculty of Pharmacy, University of Belgrade
450 Vojvode Stepe, 11000 Belgrade (Serbia)

[c] Dr. J. Xu, Prof. Dr. Y. Dou
Department of Pathology, University of Michigan
Ann Arbor, MI 48109 (USA)

Supporting information and the ORCID identification numbers for the authors of this article can be found under <https://doi.org/10.1002/cbic.201800737>.

This article is part of the young researchers' issue. To view the complete issue, visit <http://chembiochem.org/chembiotalents>



Scheme 1. A) Previously reported light-responsive peptide backbone photoswitches. B) Structure and isomerization of the new visible-light photoswitchable amino acids reported herein: Fmoc-cAzoAA (1) and Mtt-oF₄AzoAA (2). C) Former photoswitchable peptidomimetic used for MLL1-WDR5 PPI disruption (3) and new ones reported herein, which contain visible-light-responsive peptide backbone photoswitches.

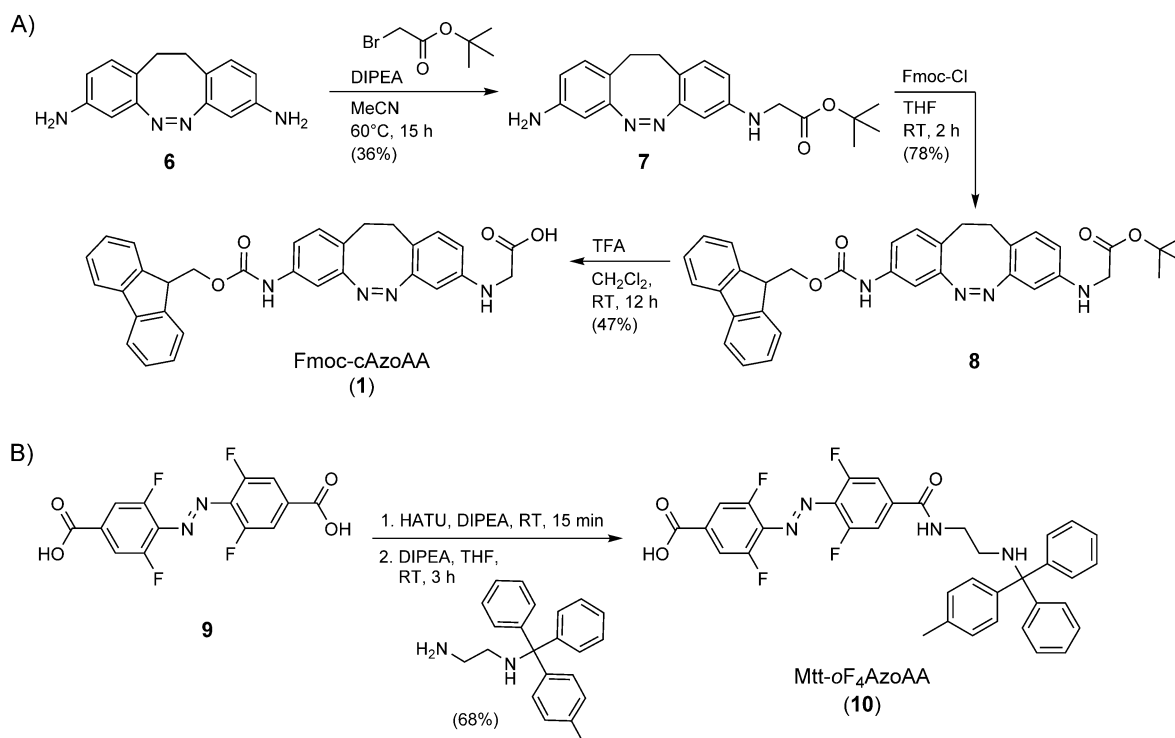
few cAzo derivatives used to modulate DNA hybridization^[13] and to control the helical conformation of peptides.^[9b]

Regarding azo-containing proteins and peptides, the strategies employed to introduce red-shifted azobenzenes are mainly restricted to cysteine-reactive tethered derivatives as side-chain crosslinkers^[14] and genetic encoding of the azoben-

zene amino acids.^[15] However, to the best of our knowledge, the direct inclusion of the azobenzene unit into peptide backbones is still limited to unsubstituted derivatives as classic fluorenylmethoxycarbonyl (Fmoc)-building blocks: [3-(3-aminomethyl)phenylazo]phenylacetic acid (AMPP),^[16] (4-aminomethyl)phenylazobenzoic acid (AMPB),^[17] and commercially available Fmoc-L-phenylalanine-4'-azobenzene.

Herein, we describe the synthesis and characterization of two novel visible-light photocontrollable amino acids based on the previous derivatives of Bléger et al.^[18] and Siewertsen et al.^[8] to be grafted onto a peptide scaffold. As our group has recently reported the possibility of externally controlling leukemia cell proliferation by disrupting the essential epigenetic PPI MLL1-WDR5, with photoswitchable azobenzene-containing peptides,^[19] we use this biological system to study its potential as novel PPI modulators. So far, to the best of our knowledge, no attempts to incorporate either tetra-*ortho*-fluoroazobenzenes or cyclic derivatives into peptide backbones have been reported. Exploring second-generation azobenzenes as visible-light-responsive peptide backbone photoswitches will complement information on systems in which the photoswitches are attached to the N terminus of the peptide sequence, or on the side chain of an amino acid as crosslinkers. Furthermore, it will contribute to gaining an insight into the design of peptidomimetics (Scheme 1).

Olalla Vázquez obtained her PhD in chemistry under the supervision of Profs. José Luis Mascareñas and Eugenio Vázquez, at Universidade de Santiago de Compostela (2010) working on synthetic transcription factors and fluorescent DNA binders. She was a visiting PhD student at Harvard University in Prof. Verdine' group (2006) and at Humboldt Universität zu Berlin with Prof. Seitz (2008). In 2011 she received a Marie Curie postdoctoral fellowship to explore RNA template-directed reactions in the context of cancer therapy. She is currently Assistant Professor at Philipps-Universität Marburg. Her research focuses on developing optochemical tools to control and understand biological processes at molecular level.



Scheme 2. Synthetic routes to the visible-light photoswitchable amino acids: A) Fmoc-cAzoAA (1) and B) Mtt-oF₄AzoAA (2). TFA: trifluoroacetic acid, HATU: 1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate. DIPEA: *N,N*-diisopropylethylamine

Results and Discussion

Synthesis of photoswitchable building blocks and inclusion into the peptide backbone

To simplify the synthetic work, even at the expense of moderate yields, we followed previous procedures of azobenzene cyclization^[9] and obtained compound **6** (Scheme 2). Afterwards, one of the primary amines was monoalkylated and the remaining one was protected with Fmoc-Cl. The final hydrolysis step yielded the corresponding Fmoc-cAzo AA (**1**).

The synthesis of peptide backbone switches would be more accessible if our molecular transducers were compatible with Fmoc solid-phase peptide synthesis (Fmoc-SPPS). However, the presence of halide substituents in the azobenzene derivatives makes them susceptible to aromatic nucleophilic substitution (S_NAr) under conventional Fmoc deprotection conditions (i.e., 20% piperidine/DMF).^[20] Consequently, we envisioned using the 4-methyltrityl (Mtt) group^[21] as an acid-labile amine protecting group because it is fully orthogonal to the side-chain ones. Thus, symmetrical dicarboxylic tetra-*ortho*-fluoroazobenzene (**9**)^[18] was monofunctionalized at the *para* position through condensation with previously synthesized Mtt-protected ethylenediamine^[22] under standard amidation conditions. This reaction provided **2** in moderate yields (Scheme 2).

Once the newly synthesized photoswitchable building blocks were characterized, with regard to purity and structural integrity (see the Supporting Information), they were incorporated into the peptide backbone of our previous PPI modulator, **3**, replacing the AMPB group. Synthesis of the cAzo pep-

tide (**4**) was straightforward. On the contrary, extension of the peptide chain after inclusion of **2** required optimization. Thus, we explored mild deprotection methods for achieving efficient removal of the Fmoc group without affecting the integrity of the molecular transducer. As shown in Table 1, the use of a

Table 1. Conditions tested for mild Fmoc deprotection of a tetra-*ortho*-fluoroazobenzene-containing peptide on a solid support at 25 °C.

Conditions	<i>t</i> [h]	Product	Starting material	Side products
A 50% Et ₃ N in CH ₂ Cl ₂	17	✓	✓	✓
B 50% Bu ₃ N in CH ₂ Cl ₂	17	✓	✓	✓
C 50% DIPEA in CH ₂ Cl ₂	17	✓	✓	✓
D 20 mM NaOH in 30% dioxane/methanol	1.5	✓	×	✓
E KF, [18]crown-6 in DMF	3	✓	✓	×
F KF, [18]crown-6 in DMF	16	✓	×	×

few weaker, but more sterically hindered, bases than piperidine (p*K*_a = 11.12),^[23] such as the tertiary amines *N,N*-diisopropylethylamine (DIPEA; p*K*_a 10.75), triethylamine (p*K*_a = 10.78), and tributylamine (p*K*_a = 10.89), caused prolonged reaction times, incomplete conversion, and detectable chemical degradation (Figures S11 and S12 in the Supporting Information). Instead, complete deprotection was accomplished with a 20 mM solution of sodium hydroxide in 30% dioxane/MeOH after only 30 min.^[24] Unfortunately, the reaction not only yielded the desired product, but also substitution side products (Figure S12, right). Likewise, other alternative conditions, such as piperazine

and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) treatment, were previously discharged due to the detection of substitution side products with **9** (Figure S9). Only a solution of potassium fluoride (6.5 equiv) in DMF with catalytic amounts of [18]crown-6 provided the deprotection product neatly (Figure S13).^[24] However, for time reasons, we resorted to peptide synthesis in solution through chemical ligation (see the Supporting Information) to obtain peptidomimetic **5** bearing the tetra-*ortho*-fluoroazobenzene derivative.

Photochemical behavior of photoswitchable peptidomimetics

Isomerization of the photoswitchable peptidomimetics **4** and **5** in aqueous solution was studied by means of UV-visible spectroscopy, HPLC, and NMR spectroscopy. A 0.128 mM solution of **4** prepared in the dark exhibited bands at $\lambda_{\max}=295$ and 402 nm for the $\pi\rightarrow\pi^*$ and $n\rightarrow\pi^*$ transitions, respectively. This is consistent with the photochemical behavior of the *cis* unsubstituted parent cAzo.^[8] Irradiation with a blue-light light-emitting diode (LED; $\lambda=405$ nm, see the Supporting Information for details) leads to spectroscopic changes in the solution: an increase of a band at $\lambda_{\max}=475$ nm. Thus, in the photostationary state (PSS) at $\lambda=405$ nm, the $n\rightarrow\pi^*$ transition is shifted to longer wavelengths. The PSS was reached after just 1 s of irradiation (Figure S16). Importantly, there is a spectral region ($\lambda=500\text{--}550$ nm) where essentially only this isomer absorbs, which consequently enables quantitative *cis* isomerization under green LED irradiation ($\lambda=520$ nm; see the Supporting Information). Additional spectral changes were detected in the region of the $\pi\rightarrow\pi^*$ ($\lambda_{\max}=295$ nm), where the absorption of the PSS at $\lambda=405$ nm was stronger than that of the PSS at $\lambda=520$ nm. Unlike the *trans* unsubstituted parental cAzo,^[8] as well as the *trans*-4,4'- and 3,3'-*p*-acetamido-substituted ones,^[9] no decrease of the band at $\lambda_{\max}=402$ nm was observed in **4**, which was also observed in the 3,3' polyurea-substituted ana-

logues.^[25] Upon irradiation with green light, the spectrum reverted to that of the thermodynamically favored *cis* isomer. In the case of **5**, the *trans* isomer was the most stable thermodynamically. Thus, a 0.105 mM solution of **5** prepared in the dark exhibited characteristic absorption bands of the *trans* tetra-*ortho*-fluoroazobenzene isomers:^[7c] an intense band at $\lambda_{\max}=319$ nm and a weaker one at $\lambda_{\max}=457$ nm assigned to the $\pi\rightarrow\pi^*$ and $n\rightarrow\pi^*$ transitions, respectively. After 30 s of continuous irradiation with the green LED, the *cis* PSS was reached (Figure S15), which led to a drastic decrease of the band at $\lambda_{\max}=319$ nm, a shift of the weaker band to shorter wavelengths ($\lambda_{\max}=412$ nm), and an increase of a band at $\lambda_{\max}=250$ nm. Irradiation at $\lambda=405$ nm reverted the spectrum to that of the *trans* isomer.

We next monitored the UV-visible absorptions of aqueous solutions of **4** and **5** after alternating irradiation cycles at $\lambda=405$ and 520 nm (Figure 1). We demonstrated the reversibility of photoisomerization (≥ 14 cycles) for our peptide backbone photoswitches without any signs of photodegradation and photochemical fatigue (Figure 1B).

To determine both the ratio and stability of the isomers in the PSS, we used HPLC. In case of **5**, we integrated the peak area of the different chromatograms at the isosbestic point ($\lambda=275$ nm) and corroborated the isomer conversion ratio by means of ¹H NMR spectroscopy (Figure S19); for **4**, the HPLC chromatograms were recorded at $\lambda=395$ nm because both isomers exhibited the same intensity in absorption and, therefore, the same extinction coefficient. There was no clear isosbestic point. In both cases, the photoconversions were efficient with ratios over 85% for the corresponding isomers (Figure 1C), except for the isomerization of peptidomimetic **4** upon irradiation with blue LEDs, for which only 16% of the *trans* isomer was detected at the PSS. This rather low conversion rate could be attributed to overlap of the $\pi\rightarrow\pi^*$ and $n\rightarrow\pi^*$ transitions in the *cis* isomer.^[9a,25]

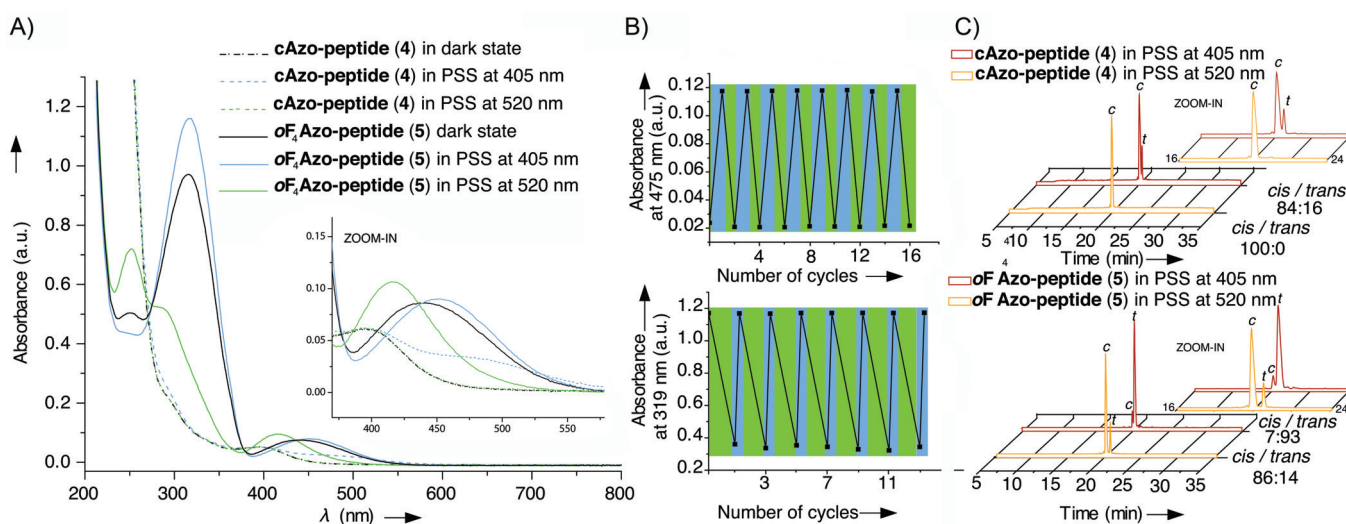


Figure 1. A) UV/Vis spectra of aqueous solutions of peptidomimetics **4** (0.128 mM; ----) or **5** (0.105 mM; —) in the PSSs at $\lambda=405$ (blue) and 520 nm (green) at 25 °C. Inset: a magnification of the $\lambda=350\text{--}600$ nm region of the spectra. B) Reversible photochromism of **4** (top) and **5** (bottom) at $\lambda=475$ and 319 nm, respectively, upon alternating intervals of irradiation at $\lambda=405$ (blue) or 520 nm (green) at 25 °C. C) HPLC chromatograms and isomer ratios in the PSS at $\lambda=405$ (red) and 520 nm (orange) of **4** (top) and **5** (bottom) at 25 °C. Insets: magnifications of the region at 16–24 min.

Regarding the thermal stability, at room temperature, the isomer ratios at the PSSs are constant for at least 24 h in all cases, as long as the solutions are stored in total darkness (Figures S17 and S18). After 4 days, there is no dramatic change in isomer distribution, but side products are detected (< 8% in the PSS at $\lambda=520$ nm of both compounds and in the PSS at $\lambda=405$ nm of **5**; Figures S17 and S18). Interestingly, this degradation is more pronounced if the cAzo-containing peptide (**4**) is in the PSS at $\lambda=405$ nm (*cis/trans* 84:16). The detected side products amounted to 29% and the substrate further degraded to 50% after one week (Figure S18). Nevertheless, stability for up to 3 days is compatible with the vast majority of biological assays.

Certain azobenzene derivatives are susceptible to reduction by intracellular thiols, whereas others are not. Glutathione (GSH) is the primary intracellular reducing agent, the expected highest intracellular concentration of which is 10 mM.^[26] To have the possibility to use our compounds in a biological context, we evaluated the stability of both **4** and **5** in the presence of 10 mM GSH. Interestingly, the behavior of the peptidomimetics is different: whereas **5** was completely resistant to GSH reduction in both PSSs for 4 days (Figure S20), peptide **4** was found to be sensitive after just 1 day (Figure S21). This degradation effect was more severe in the PSS at $\lambda=405$ nm, which was in agreement with our previous stability tests.

Binding affinity of the photoswitchable peptidomimetics to WDR5 and inhibition of MLL1 activity

In our previous study, we demonstrated that the AMPB-containing peptide **3** was able to efficiently inhibit MLL1 activity through strong binding to the protein WDR5 (K_i in low-nM range for both isomers).^[19] WDR5 belongs to the MLL1 protein core complex and is essential because WDR5 knockdown totally abolishes the histone methyltransferase (HMT) function of MLL1.^[27] The newly synthesized peptidomimetics **4** and **5** are analogues of **3** bearing either the cAzo derivative or *o*F₄Azo, respectively. Thus, once these visible-light photoswitchable peptidomimetics were synthesized and photochemically characterized, we investigated their WDR5-binding capacities and possible differences between isomers. To this end, we used our previously optimized fluorescence polarization (FP) competitive assay,^[19] with alteration to the irradiation step. We included the corresponding controls to rule out any interference with the isomerization (Figures S22–S25). The obtained results are summarized in Figure 2 and Table 2; the IC_{50} values (Table S1) were transformed into inhibition constants (K_i) to enable objective comparisons,^[28] in particular, with our former peptide **3**.

All our novel peptidomimetics interacted with WDR5 with high affinity, in the nanomolar range (Table 2). In comparison with our previous results, peptide **3** proved to be the photoswitchable peptidomimetic with the highest affinity to WDR5, as well as the biggest difference between isomers ($K_i=1.25$ nM for *trans*-**3**; $K_i=6.50$ nM for *cis*-**3**; ratio: 5.0). Thus, exchange of the azobenzene with the cyclic analogue caused a more than 100-fold decrease in affinity to WDR5, relative to that with **3**, whereas, as expected, the incorporation of the small fluorine

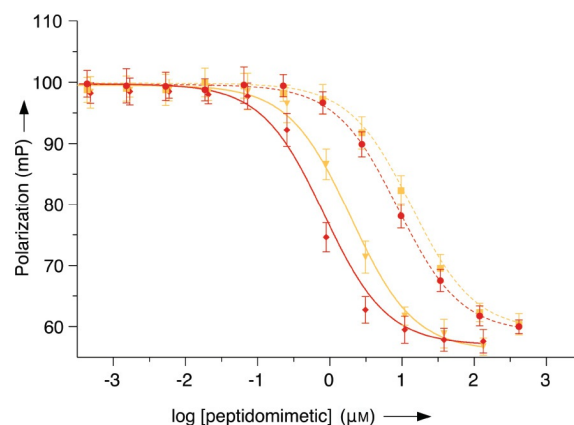


Figure 2. FP-based competitive assays of the peptidomimetics **4** (---) and **5** (—) in the PSSs at $\lambda=405$ nm (● and ◆, respectively) and 520 nm (■ and ▼, respectively) at 25 °C. Mean data points and standard deviations are derived from three independent experiments.

Table 2. Binding affinities of the photoswitchable peptidomimetics to WDR5 at 25 °C.			
Peptidomimetic	K_i [nM]		Ratio
	PSS at 405 nm	PSS at 520 nm	
4	140 ± 35	207 ± 52	1.5
5	11.8 ± 1.4	30.8 ± 3.3	2.6

atoms into the azobenzene unit (**5**) affected the binding properties only slightly. Furthermore, the improved photoconversion properties of **5** were not directly translated into higher differences between isomers in our FP-binding assays compared with **3**. Nevertheless, it should be remembered that even low photoconversions can trigger significant biological effects. From all of our results so far, we selected **5** for further biological assays.

Next, we explored if visible-light photoswitchable peptidomimetic **5** could modulate the essential PPI of the MLL1 core complex (MLL1-WDR5) and, in turn, MLL1 activity. To this end and to avoid radioactivity, we used an amplified luminescent proximity homogeneous (AlphaLISA) assay^[29] with the tetramer reconstituted MLL1 core complex (i.e., MLL1, WDR5, RbBP5, and Ash2L), the H3-21-mer peptide as a substrate, and the co-factor S-adenosyl methionine (SAM) as a universal methyl group donor. Gratifyingly, the functional data obtained from our AlphaLISA-based MLL HMT assay demonstrated that **5** effectively inhibited MLL1 activity (Figure 3; IC_{50} for *trans* **5** = (0.927 ± 0.034) μ M, IC_{50} for *cis* **5** = (1.73 ± 0.12) μ M) through disruption of the WDR5-MLL1 PPI due to its high affinity to WDR5. Notably, these determined potencies are based on the ability of **5** to inhibit the mono- and dimethylation function of the reconstituted MLL1 core complex. In addition, we could observe a modest, yet clear, difference between isomers of up to 1.8-fold. To further corroborate effective PPI disruption by **5** and the slight difference between isomers, we performed GST pull-down experiments. We incubated 0.262 μ M GST-tagged MLL1 protein with the remaining proteins of the core complex to a final concentration of 0.4 μ M. Afterwards, increasing

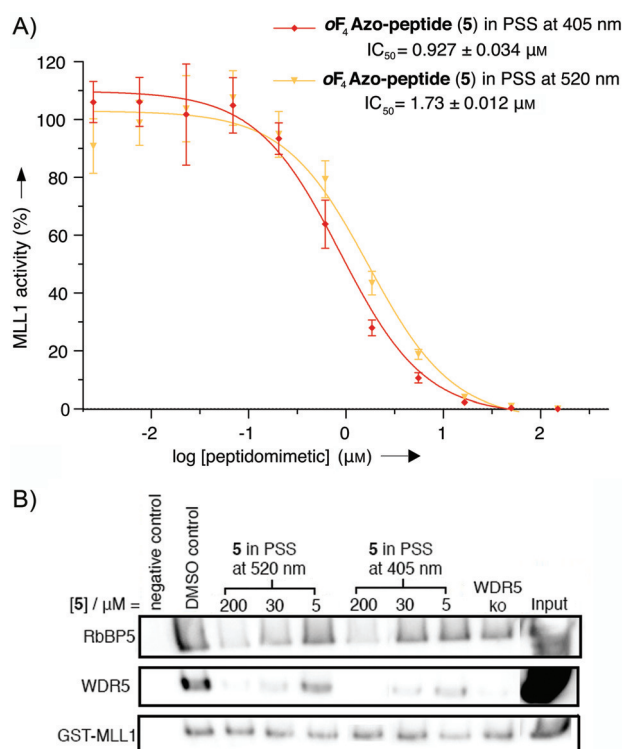


Figure 3. A) In vitro functional HMT AlphaLISA assay to evaluate the potency of peptidomimetic **5** to inhibit MLL1 enzymatic activity in the PSSs at $\lambda = 405$ nm (♦) and 520 nm (▼). Mean data points and standard deviations are derived from three independent experiments. B) In vitro glutathione–sepharose–transferase (GST) pull-down assay with the reconstituted 4-mer MLL1 core complex in the presence of **5** in the PSSs at $\lambda = 405$ and 520 nm; ko = knockout.

amounts of isomers of **5** were added. Both isomers were able to disrupt the MLL1 complex in a dose-dependent fashion. Importantly, we demonstrated different behavior between isomers: the *trans* isomer, again, showed a higher dissociation potential through its higher affinity to WDR5. These results are consistent with our FP-based experiments, as well as with the AlphaLISA-based MLL HMT assays. Therefore, collectively, our results confirm that visible-light irradiation triggers conformational changes in peptidomimetic **5**, which affect its capacity to modulate the WDR5–MLL1 PPI and, consequently, its potential to inhibit MLL1.

Crystal structure determination and structural basis of the peptidomimetic–WDR5 interaction

To understand the molecular principles behind the observed affinity deviations, we aimed to crystallize WDR5 in complex with peptidomimetics **4** and **5**. No crystals have been obtained for **5** in any PSS, to date. Cococrystallization with **4** in the PSS at $\lambda = 520$ nm (*cis* isomer) yielded crystals diffracting to a resolution of 1.51 Å (PDB ID: 6IAM).

The complex of WDR5 with **4** appears in a monoclinic crystal form (Table 3) that is distinct from those of previous crystal packings observed for WDR5–peptide complexes; this is most likely due to packing of the cAzo moiety against Pro168 of a

Table 3. Data collection and refinement of the cococrystal structure of WDR5 in complex with **4**.

Data collection	Values
λ [Å]	0.873 Å
resolution range [Å]	37.49–1.51 (1.564–1.51)
space group	$P12_11$
unit cell [Å, °]	$a = 46.53, b = 46.56, c = 66.17,$ $\alpha = 90, \beta = 107, \gamma = 90$
total reflections	167527 (15726)
unique reflections	42377 (4073)
multiplicity	4.0 (3.9)
completeness (%)	99.3 (96.2)
mean $I/\sigma(I)$	9.04 (1.50)
Wilson B factor (Å ²)	11.27
R_{merge} [%]	0.0899 (0.713)
Refinement statistics	
$R_{\text{work}}/R_{\text{free}}$ [%]	0.148/0.181
number of atoms	2895
protein	2377
ligands	100
solvent	418
protein residues	322
RMS (bonds/angles) [Å/°]	0.009/1.08
Ramachandran favored [%]	96.15
Ramachandran outliers [%]	0.00
rotamer outliers [%]	0.36
clashscore	2.20
average B factor [Å ²]	16.1
PDB ID	6IAM

symmetry-related complex. Nevertheless, the WDR5 structure is almost unaffected, with a low root-mean-square deviation (RMSD) of 0.200 Å for 262 C α atoms upon superimposition on the WDR5–histone 3 peptide complex (PDB ID: 2CO0).^[30] Peptide-bound WDR5 adopts the well-established β -propeller conformation, with a central cavity that accommodates the Arg3 side chain^[31] that is part of the N-terminal recognition motif of **4**. The superimposition of **3** and **4** (Figure 4B) reveals almost identical interactions of the N-terminal stretches (SARA) with WDR5, including bridging water molecules and an α -helical main-chain trace. However, the central cAzo moiety packs differently to WDR5 than that of the linear azobenzene group of **3**. For example, the cAzo group lacks any direct interactions with the side chains of Lys259 and Tyr260, which are otherwise found for the WDR5–**3** complex. The C-terminal stretching of **4** (Val6–Ser11) is fully ordered, which might be due to interactions of the C-terminal carboxamide group with a symmetry-related WDR5 molecule, and adopts an α -helical conformation not observed previously for any cococrystals between WDR5 and MLL-derived peptides (Figure S26, right). Apart from additional packing interactions, peptide **4** establishes three water-mediated hydrogen bonds: the side-chain carbonyl group of Asp172 interacts with the backbone carbonyl group of Arg9 from the ligand. Likewise, the side-chain hydroxy group of Tyr191 and the amino group of the Lys259 coordinate a water molecule that bridges to the C-terminal backbone carbonyl of the cAzo entity (Figures S27 and S28).

We reported **3** to have an increased affinity in comparison to the WIN peptide (PDB ID: 3EG6) in a previous publication.^[19]

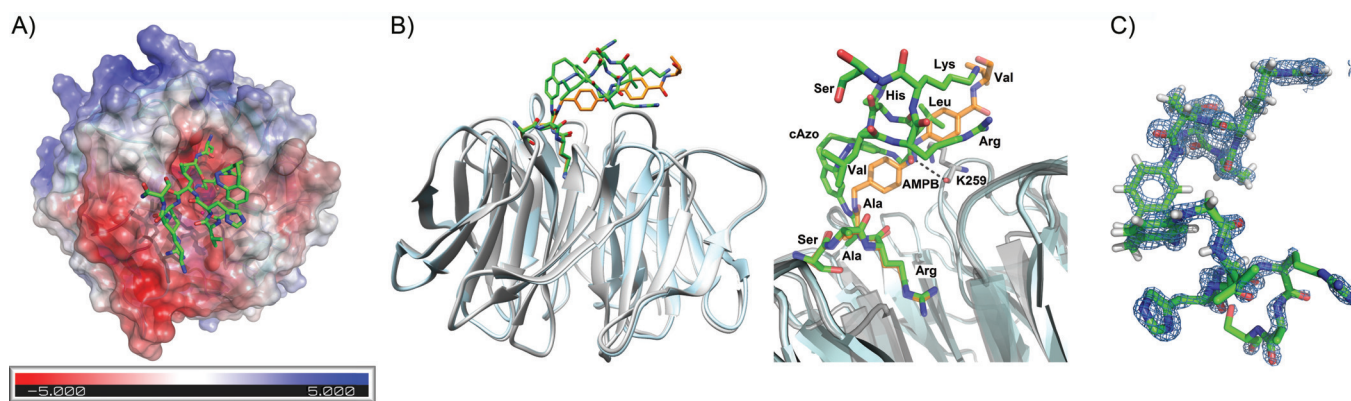


Figure 4. A) Crystal structure of WDR5 in complex with **4** (PDB ID: 6IAM); WDR5 is shown as a surface representation of its electrostatic potential with the color code from red (negative) to blue (positive) in dimensionless units of $k_B = T/e_c$ in which k_B is the Boltzmann constant, T is the temperature, and e_c is the charge of an electron (generated by using the APBS PyMOL plug-in). B) Left: side view of the superimposed cocrystal structures of **3** (orange) and **4** (green) to WDR5. Right: Close-up view of the left overlay B, in which the key hydrogen bond between **3** and WDR5 (K259) is highlighted, along with labeled residues and azobenzene motifs. C) The $2F_o - F_c$ experimental electron density map contoured at 1.0σ (blue) shows the density for **4**.

Because **4** lacks several of the interactions of **3**, it is not surprising that it has a lower binding affinity, which is, indeed, more comparable to the WIN peptide than that to **3** ($K_i = 120$ nM for WIN peptide; $K_i = 140$ nM for *cis*-**4**, and $K_i = 1.25$ nM for *trans*-**3**)

Molecular modeling of peptidomimetics–WDR5 interactions

To further rationalize the obtained experimental results and to set up a workflow for further design of novel photoswitchable peptidomimetics, a series of virtual docking (VD) and molecular dynamics (MD) calculations were performed.

For all six ligands, *cis/trans*-**3**, *cis/trans*-**4**, and *cis/trans*-**5**, initial poses for MD calculations were obtained through molecular docking. The docking protocol was validated with heavy atoms RMSD = 0.8915 calculated for the part of *trans*-**3** already resolved in the crystal structure (PDB ID: 5M23; Figure S29).^[19] Considering the significant conformational flexibility of peptidomimetics, several initial protein–ligand complexes were generated for each ligand. Only protein–ligand complexes in which the N-terminal part of the peptidomimetics (SARA-) remained stabilized during 20 ns MD production runs were used for further analysis. Results of MD simulations expressed as RMSD fluctuations of the N-terminal part of peptidomimetics (SARA-; Figure 5), and the whole protein (Figure S30) during 20 ns of production runs indicate reasonably stabilized complexes converged to the equilibration state.

The molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) calculations of the binding energy were performed on the last 7 ns of each simulation to further validate predicted binding modes of the peptidomimetics and to investigate the relationship between experimental K_i values (transformed in $pK_{i, \text{exptl}} = -\log K_i$) derived for the synthesized photoswitchable peptidomimetics and their calculated MM/PBSA scores. A linear regression model with $R^2 = 0.88$ between aforementioned parameters was established (Table 4); thus indicating that this model could be used as a predictive tool with discriminative properties between *cis* and *trans* isomers for structurally similar peptidomimetics.

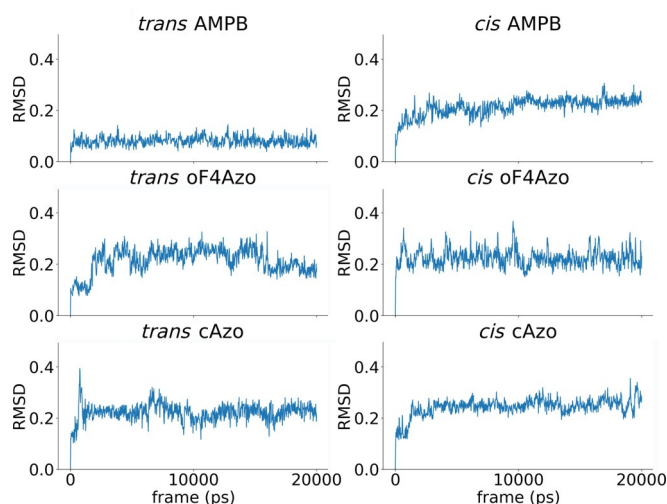


Figure 5. RMSD fluctuations during 20 ns of MD production runs calculated for the SARA- sequence of the ligands.

Table 4. Results obtained from MM/PBSA calculations and regression analyses.

Peptidomimetic	$pK_{i, \text{exptl}}^{[a]}$	MM/PBSA score [kJ mol]
<i>trans</i> - 3	8.903	−51.282
<i>cis</i> - 3	8.187	−21.136
<i>trans</i> - 4	6.854	−5.909
<i>cis</i> - 4	6.684	−5.680
<i>trans</i> - 5	7.923	−28.388
<i>cis</i> - 5	7.511	−15.371

[a] $pK_{i, \text{exptl}} = -19.2458(\text{MM/PBSA}) + 126.4877$, $R^2 = 0.88$.

To gain an insight into details of the molecular interactions, after a trajectory clustering procedure, obtained cluster representatives were further analyzed. For all of the examined ligands, the largest fluctuations were detected in the C-terminal part of the peptidomimetics (−VHLRKS; Figure S31). These re-

sults are in accordance with the previously published crystal structure of the 3–WDR5 complex (PDB ID: 5M23), for which the C-terminal part of the peptidomimetics (-VHLRKS) remain unresolved.

The binding mode of *trans*-3 remained similar to the binding mode observed in the crystallographic structure (PDB ID: 5M23; Figure S32), which further supported our computational workflow. Better stabilization of the SARA- sequence of *trans*-3 relative to the other ligands was confirmed through inspection of RMSD plots (Figure 5) and calculation standard deviations of RMSD over the 20 ns of MD simulations (Table S3). According to MD simulations of *trans*-3, valine, as the least fluctuating residue from the C terminus of the peptidomimetic (-VHLRKS; Figure S31), was recognized as an important residue for stabilization of the protein–ligand complex. Upon comparing poses of *cis*-3 and *trans*-3 (Figure 6), the N-terminal part of both peptidomimetics (SARA-) remained similar; the largest difference was in the orientation of the azobenzene part of the ligand, and consequently, in the C-terminal part (-VHLRKS).

The binding mode of *trans*-5 remained similar to that of *trans*-3 with regard to the N-terminal domain, as expected, whereas the tetra-*ortho*-fluoroazobenzene moiety was partially shifted relative to the azobenzene moiety of *trans*-3 (Figure S33). Leucine within the C-terminal part (-VHLRKS) showed transient intramolecular interactions with the tetra-*ortho*-fluoroazobenzene moiety during MD simulations (Figures 7A, S33, and S35). This particular intramolecular interaction was seen as the main reason for increased C-terminal valine fluctuations compared with the same valine from *trans*-3 during MD simulations (Figure S31). Valine retained a similar binding mode inside the WDR5 binding site to that of *trans*-3 (Figure S33). Nevertheless, the similar orientation and position of the C-terminal valine from *trans* peptides 3 and 5 indicate its impor-

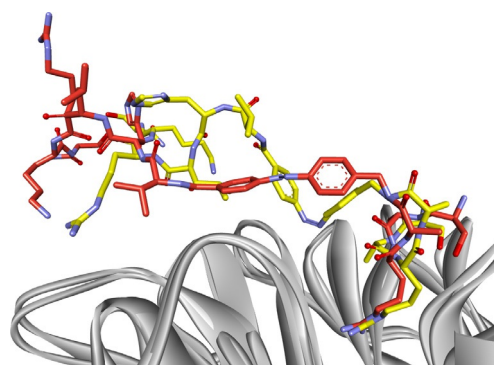


Figure 6. Difference in the binding modes obtained for *trans*-3 (red) and *cis*-3 (yellow) inside of the WDR5 binding site.

tance in the overall stabilization of the protein–ligand complex.

A significant shift in the binding conformation of the SARA-domain was observed for *cis*-3 (Figure S35). Overall, the difference in observed binding modes between *cis/trans*-3 and *cis/trans*-5 can be interpreted as being due to different connectors between the SARA- domain and the novel fluorinated azobenzene. Thus, novel peptidomimetic 5 has a 2-aminoethyl-carbamoyl linker connecting the fluorinated azobenzene and SARA-domain, instead of the methylene group of peptidomimetic 3, which connects the SARA- domain and unsubstituted azobenzene.

MD calculations performed for *trans*-4 indicated that this isomer partially mimicked the positioning of other studied *trans* ligands (Figures 7C and S36). The predicted MD binding mode of *cis*-4 is comparable to that of its crystal structure (PDB ID: 6IAM) in terms of the orientation of the SARA-

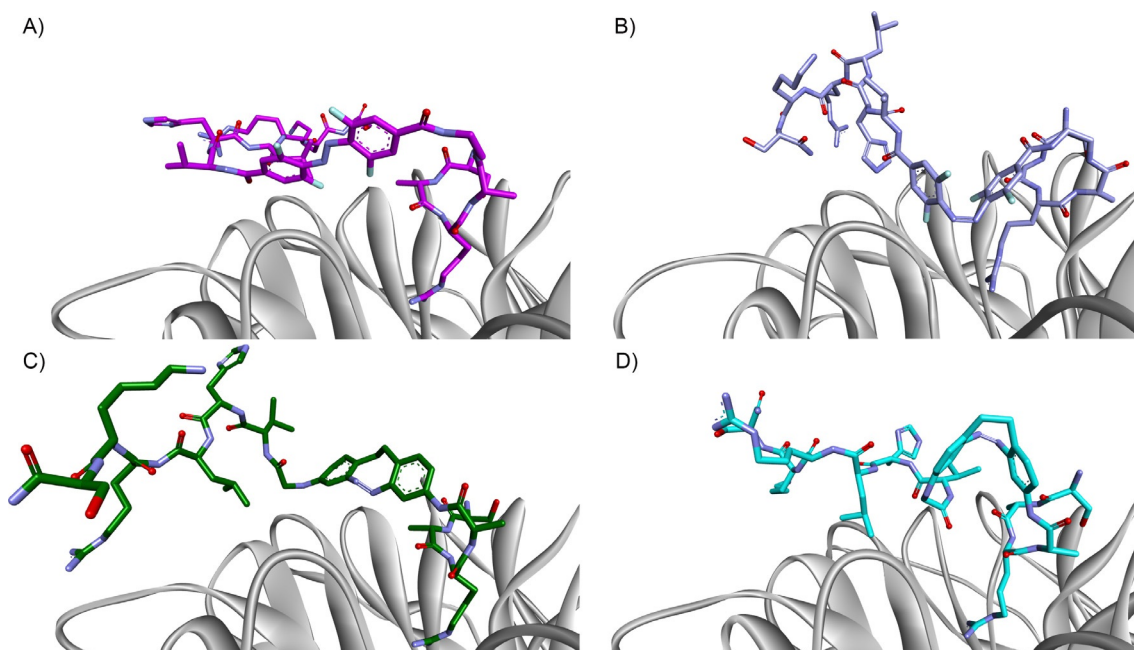


Figure 7. Binding modes of WDR5 and A) *trans*-5 (dark purple); B) *cis*-5 (light purple); C) *trans*-4 (green), and D) *cis*-4 (cyan).

sequence and cAzo moiety. This supports our computational workflow. However, some disagreement in the orientation of the -VHLRKS sequence was observed (Figure S37). Interestingly, as a part of the C-terminal -VHLRKS sequence, valine positioning in our crystal structure is comparable to the position of isoleucine of the predicted structure; thus signifying the importance of this region (residues Tyr131, Phe149) of the protein surface for the stabilization of a protein–ligand complex. Compared with *trans*-4, and to other ligands, the C-terminal part of *cis*-4 interacted with a smaller part of the protein binding surface, which could be a partial explanation for the inability to crystallize protein–ligand complexes of the other peptides (Table S4).

The novel crystal structure of *cis*-4 bound to WDR5, as well as linear regression derived for experimental results (pK_i values for the synthesized peptides) and MM/PBSA scores, makes this procedure promising for further in silico design of novel WDR5-MLL1 PPI photoswitchable disruptors. There is a need to design and synthesize novel derivatives of aforementioned peptidomimetics to increase the overall predictive power of such a computational study.

Conclusion

We have reported the synthesis of two visible-light photoswitchable amino acids for grafting onto a peptide scaffold. The goal was to use them as visible-light-responsive peptide backbone photoswitches and explore their potential as PPI modulators based on our former photoswitchable ligand 3.^[19] We found 5 particularly interesting because the only difference from the parental compound was the *ortho* incorporation of four fluorine atoms in the azobenzene motif and one additional methylene group in the linker of the azobenzene to alanine. Such modification improved the photochemical characteristics of this second-generation photoswitchable PPI modulator, which allowed better photoconversion under visible-light irradiation. This improvement, however, did not lead to a higher difference in the studied biological responses: neither in WDR5 affinity nor in MLL1 inhibition. This interesting observation highlights the difficulty in straightforwardly rationalizing the design of effective photoswitchable PPI inhibitors and plays down the quantitative photoconversion of the molecular transducer to achieve a functional binary system (on/off) in biological environments. For this purpose, having molecular information at our disposal is essential. Our structure–activity relationships determined from the novel crystal structure and computational methods shed light on binding to WDR5. The crystal structure and results from the molecular modeling study revealed substantial differences in the orientation of the azobenzene core for the characterized peptidomimetics. Agreement between experimental K_i values and MD results turns the developed in silico procedure into a promising tool for the further design of novel closely related peptidomimetics, which is our goal for further studies.

Although quantitative photoconversion of the photoswitches might not be essential for biological applications, visible-light irradiation is an indispensable requirement for in vivo use

and to avoid phototoxicity. In addition, our photocontrollable peptide backbone switches offer benefits, such as fast isomerization and reversibility without degradation and high stability. Consequently, we believe that they will find applications in a wide range of light-driven molecular processes beyond protein control and related to photopharmacology, optobiology, materials science, antisense chemistry, or antimicrobial activity.

Experimental Section

General: All commercially reagents and solvents were purchased and used without further purification (see the Supporting Information). NMR spectra were recorded at 300 K on Bruker AV III HD 300 and 600 MHz spectrometers, whereas HRMS (ESI) results were acquired with a LTQ-FT Ultra mass spectrometer (Thermo Fischer Scientific), as specified in the Supporting Information.

Peptide synthesis: Peptides were synthesized according to the standard Fmoc-SPPS methodology and purified by means of preparative or semipreparative HPLC. Characterization was performed through HPLC-MS and HRMS (ESI). Peptide 5 was synthesized through the ligation reaction of two precursor peptides: Boc-S(tBu)-A-R(Boc)₂-A-OH and H₂N-oF₄Azo-V-H(Trt)-L-R(Boc)₂-K(Boc)-S(tBu)-CONH₂ (Boc: *tert*-butyloxycarbonyl, Trt: trityl).

Organic synthesis: Precursors 6 and 9 for the preparation of photoswitchable amino acids 1 and 2 were synthesized according to literature procedures.^[8–9,32]

Compound 7: Compound 6 (200 mg, 0.840 mmol, 1.00 equiv) was dissolved in MeCN (4.26 mL), under a nitrogen atmosphere, and DIPEA (217 mg, 1.64 mmol, 2.00 equiv) was added to this solution. After the solution was heated to 60 °C, *tert*-butyl bromoacetate (163 mg, 0.84 mmol, 1.00 equiv) was added and stirred for 15 h. Afterwards, the solvent was removed under reduced pressure and the product was isolated by flash column chromatography (pentane/EtOAc 2:1). Product 7 was obtained as a yellow solid (107 mg, 0.304 mmol, 36%). TLC: R_f = 0.41 (pentane/EtOAc 1:1); ¹H NMR (300 MHz, CDCl₃): δ = 6.74 (t, ³ J = 8.2 Hz, 2H; 2 × CH_{ar}), 6.32 (dd, ³ J = 8.1 Hz, ⁴ J = 2.4 Hz, 1H; CH_{ar}), 6.26 (dd, ³ J = 8.2 Hz, ⁴ J = 2.5 Hz, 1H; CH_{ar}), 6.13 (d, ³ J = 2.4 Hz, 1H; CH_{ar}), 6.02 (d, ³ J = 2.5 Hz, 1H; CH_{ar}), 4.20 (s, 1H; NH), 3.68 (s, 2H; NHCH₂), 3.57 (s, 2H; NH₂), 2.84–2.78 (m, 2H; CH₂), 2.61–2.55 (m, 2H; CH₂), 1.46 ppm (s, 9H; 3 × CH₃); ¹³C NMR (75 MHz, CDCl₃): δ = 170.1 (CO), 156.4 (C_{ar}), 156.3 (C_{ar}), 145.8 (C_{ar}), 144.9 (C_{ar}), 131.7 (C_{ar}H), 130.6 (C_{ar}H), 118.6 (C_{ar}), 117.8 (C_{ar}), 114.2 (C_{ar}H), 112.3 (C_{ar}H), 105.3 (C_{ar}H), 102.9 (C_{ar}H), 82.2 (C(CH₃)₃), 46.6 (CH₂NH), 31.2 (2 × CH₂), 28.2 ppm (3 × CH₃); HRMS (ESI+): m/z calcd for C₂₀H₂₄N₄O₂H [M+H]⁺: 353.1972; found: 353.1974.

Compound 8: Compound 7 (124 mg, 0.352 mmol, 1.00 equiv) dissolved in THF (2.30 mL) was cooled to 0 °C and pyridine (31.2 μ L, 0.388 mmol, 1.10 equiv) was added. Then, Fmoc-Cl (100 mg, 0.288 mmol, 1.10 equiv) was added portionwise and the mixture was stirred at RT for 1.5 h. The mixture was diluted with 1 M HCl (2.0 mL), extracted with EtOAc (3 × 20.0 mL), washed with brine, and dried over anhydrous MgSO₄. The solvent was removed in vacuo. Product 8 was isolated by flash column chromatography (pentane/EtOAc 2:1) and obtained as a yellow solid (158 mg, 0.275 mmol, 78%). TLC: R_f = 0.47 (pentane/EtOAc 2:1); ¹H NMR (300 MHz, CDCl₃): δ = 7.77 (d, ³ J = 7.8 Hz, 2H; 2 × CH_{ar}), 7.58 (d, ³ J = 7.3 Hz, 2H; 2 × CH_{ar}), 7.41 (t, ³ J = 7.3 Hz, 2H; 2 × CH_{ar}), 7.31 (td, ³ J = 7.4 Hz, ⁴ J = 1.0 Hz, 2H; 2 × CH_{ar}), 6.92 (d, ³ J = 8.4 Hz, 1H; CH_{ar}), 6.78 (d, ³ J = 8.2 Hz, 1H; CH_{ar}), 6.63 (m, 1H; CH_{ar}), 6.39 (dd, ³ J = 8.3 Hz,

$^4J = 2.4$ Hz, 1 H; CH_{ar}), 6.32–6.22 (m, 1 H; CH_{ar}), 6.13 (d, $^3J = 2.4$ Hz, 1 H; CH_{ar}), 4.51 (d, $^3J = 6.4$ Hz, 2 H; OCH_2), 4.23 (t, $^3J = 6.4$ Hz, 1 H; OCH_2CH), 3.67 (s, 2 H; $NHCH_2$), 3.03–2.90 (m, 2 H; CH_2), 2.90–2.68 (m, 2 H; CH_2), 1.44 ppm (s, 9 H; $3 \times CH_3$); ^{13}C NMR (75 MHz, $CDCl_3$): $\delta = 168.5$ (CO), 155.8 (CO), 143.8 ($2 \times C_{ar}H$), 141.4 ($2 \times C_{ar}H$), 136.7 ($2 \times C_{ar}H$), 130.5 ($2 \times C_{ar}$), 128.0 ($2 \times C_{ar}$), 127.7 ($2 \times C_{ar}$), 127.2 ($2 \times C_{ar}$), 127.1 ($2 \times C_{ar}$), 125.0 ($2 \times C_{ar}$), 124.9 ($2 \times C_{ar}$), 120.2 ($2 \times C_{ar}$), 120.1 ($2 \times C_{ar}$), 82.2 (CCH₃), 68.5 (CH₂), 52.9 (CH), 47.2 (CH₂), 31.9 (CH₂), 31.0 (CH₂), 28.2 ppm ($3 \times CH_3$); HRMS (ESI+): m/z calcd for $C_{35}H_{34}N_4O_4Na$ [$M+Na$]⁺: 597.2472; found: 597.2473.

Compound 1: Compound **8** (202 mg, 0.352 mmol, 1.00 equiv) was dissolved in CH_2Cl_2 (2.50 mL) and cooled to 0 °C. Then TFA (5.00 mL) was added dropwise. After stirring for 10 h at RT, the solvent was removed under reduced pressure. The obtained residue was dissolved in EtOAc and water, extracted with EtOAc (3×10.0 mL), and dried over anhydrous Na_2SO_4 . The solvent was removed under reduced pressure and the product was isolated by flash column chromatography ($CH_2Cl_2/MeOH$ 15:1). Product **1** was obtained as a yellow solid (85.8 mg, 0.165 mmol, 47%). TLC: $R_f = 0.06$ ($CH_2Cl_2/MeOH$ 15:1); 1H NMR (300 MHz, $CDCl_3$): $\delta = 7.74$ (d, $^3J = 7.5$ Hz, 2 H; $2 \times CH_{ar}$), 7.54 (d, $^3J = 7.2$ Hz, 2 H; $2 \times CH_{ar}$), 7.38 (t, $^3J = 7.4$ Hz, 2 H; $2 \times CH_{ar}$), 7.31–7.21 (m, 3 H; $3 \times CH_{ar}$), 6.82 (s, 2 H; $2 \times CH_{ar}$), 6.71 (d, $^3J = 8.3$ Hz, 1 H; CH_{ar}), 6.21 (dd, $^3J = 8.2$ Hz, $^4J = 2.4$ Hz, 1 H; CH_{ar}), 5.96 (t, $^3J = 4.2$ Hz, 1 H; CH_{ar}), 4.48 (d, $^3J = 6.4$ Hz, 2 H; OCH_2), 4.20 (d, $^3J = 6.4$ Hz, 1 H; OCH_2CH), 3.74 (s, 2 H; $NHCH_2$), 2.91–2.79 (m, 2 H; CH_2), 2.67–2.54 ppm (m, 12 H; CH_2); ^{13}C NMR (75 MHz, $CDCl_3$): $\delta = 174.3$ (CO), 156.2 (CO), 155.8 ($2 \times C_{ar}$), 145.6 ($2 \times C_{ar}$), 143.7 ($2 \times C_{ar}$), 141.5 ($2 \times C_{ar}$), 136.1 (C_{ar}), 130.8 ($2 \times C_{ar}H$), 130.4 ($2 \times C_{ar}H$), 128.0 ($2 \times C_{ar}H$), 127.3 ($2 \times C_{ar}H$), 125.0 ($2 \times C_{ar}H$), 120.2 ($2 \times C_{ar}H$), 118.2 (C_{ar}), 112.7 ($C_{ar}H$), 102.8 ($C_{ar}H$), 67.1 (CHCH₂), 47.2 (CHCH₂), 45.8 (NHCH₂), 31.3 (CH₂), 30.9 ppm (CH₂); HRMS (ESI+): m/z calcd for $C_{31}H_{26}N_4O_4H$ [$M+H$]⁺: 519.2038; found: 519.2019.

Compound 2: Compound **9** (275. mg, 0.804 mmol, 1.50 equiv) was suspended in THF (20.0 mL). HATU (206 mg, 0.536 mmol, 1.00 equiv) and DIPEA (187 μ L, 1.61 mmol, 2.00 equiv) were added and the resulting solution was stirred at RT. After 15 min, compound **10** (210 mg, 0.665 mmol, 1.20 equiv) in THF (7.00 mL) was added and the solution was stirred for 3 h at RT. Afterwards, the solvent was removed under reduced pressure and the crude product was purified by flash column chromatography ($CH_2Cl_2/MeOH$ 45:1 + 3% NEt_3) to yield the desired product **2** as the triethylamine salt as a red solid (244 mg, 0.329 mmol, 61%). To obtain the NMR spectra without traces of triethylamine, a small amount was further purified by flash column chromatography ($MeCN/MeOH$ 45:1 + 3% NH_3). TLC: $R_f = 0.16$ ($CH_2Cl_2/MeOH$ 45:1 + 3% NEt_3); 1H NMR (300 MHz, DMSO): $\delta = 8.87$ (s, 1 H; CONHCH₂), 7.80 (d, $^3J = 10.6$ Hz, 4 H; $4 \times CH_{ar}$), 7.42 (d, $^3J = 7.4$ Hz, 4 H; $4 \times CH_{ar}$), 7.28 (t, $^3J = 6.7$ Hz, 6 H; $6 \times CH_{ar}$), 7.20 (d, $^3J = 7.1$ Hz, 2 H; $2 \times CH_{ar}$), 7.09 (d, $^3J = 8.0$ Hz, 2 H; $2 \times CH_{ar}$), 3.48–3.46 (m, 2 H; CH_2NH), 2.26 (s, 3 H; CH_3), 2.16 ppm (m, 2 H; CH_2NH); ^{13}C NMR (75 MHz, DMSO): $\delta = 164.5$ (CO), 162.8 (CO), 156.1 ($C_{ar}F$), 155.9 ($C_{ar}F$), 152.6 ($C_{ar}F$), 152.4 ($C_{ar}F$), 150.3 ($C_{ar}C$), 146.1 ($2 \times C_{ar}C$), 144.0 ($C_{ar}C$), 142.9 ($C_{ar}C$), 138.7 ($C_{ar}C$), 135.7 ($C_{ar}C$), 135.0 ($C_{ar}C$), 128.3 ($4 \times C_{ar}H$), 128.2 ($2 \times C_{ar}H$), 127.6 ($4 \times C_{ar}H$), 127.5 ($2 \times C_{ar}H$), 125.9 ($2 \times C_{ar}H$), 113.9 ($C_{ar}HCF$), 113.6 ($C_{ar}HCF$), 112.2 ($C_{ar}HCF$), 111.9 ($C_{ar}HCF$), 82.4 (NHCC₂), 70.1 (NHCH₂), 43.1 (NHCH₂), 20.4 ppm (CH₃); HRMS (ESI+): m/z calcd for $C_{36}H_{27}F_4N_4O_3$ [$M-H$]⁻: 639.2025; found: 639.2025.

UV/Vis spectroscopy and FP measurements: Concentration determinations, UV/Vis, and FP-based measurements were performed on a Tecan (Switzerland) Spark 20M multimode microplate reader at RT. All measurements for concentration determinations were performed in a 1400 μ L quartz cuvette (Hellma Analytics (104F-QS)

with a pathlength of 1 cm. FP assays were performed as described previously^[19] in black 96-well microtiter plates (Greiner, ref. no. 655900) with excitation at $\lambda = 485$ nm and emission at $\lambda = 530$ nm. A detailed procedure of the FP-based assays is provided in the Supporting Information. K_i values were calculated by using the equation described previously by Wang et al. and the corresponding webpage provided by them.^[28]

AlphaLISA HMT assay: The AlphaLISA HMT assay was performed in white 384-well plates (Corning, ref. no. 4512) and with the AlphaLISA buffer (50 mM Tris pH 8.0, 50 mM NaCl, 5 mM $MgSO_4$, 10% glycerol, 0.01% Tween-20, 1 mM dithiothreitol (DTT)). The MLL1 complex (MLL1, WDR5 Δ 23, Ash2L, RbBP5) was reconstituted in AlphaLISA buffer to obtain a 205 nM solution. Each reaction contained final concentrations of 100 nM MLL1 complex, 1.70 μ M substrate H3-21-mer peptide (Anaspec), and 2.00 μ M SAM (PerkinElmer). Both *trans* and *cis* isomers of peptidomimetic **5** (*trans*: previously irradiated at $\lambda = 405$ nm for 5 min; *cis*: previously irradiated at $\lambda = 520$ nm for 5 min, as discussed in the Supporting Information) were added at concentrations ranging from 1.70 nM to 150.0 μ M and incubated with the preassembled MLL1 complex for 20 min on ice. Each plate contained triplicates of a negative DMSO biograde control, a positive control with the already methylated H3 peptide, and a blank with only buffer and DMSO. Reactions were initiated by the addition of the H3/SAM mixture and incubated for 3 h at RT, before addition of the acceptor and donor Alpha-beads (equilibrated with the AlphaLISA buffers). After another 2 h of incubation at RT, luminescence was measured on a PerkinElmer EnVision plate reader (mirror: 444, emission: $\lambda = 570$ nm). The experiment was performed in triplicate three times, independently.

GST pull-down assay: For the GST pull-down assay GST-tagged MLL1 (500 μ L) in assay buffer (50 mM Tris, 150 mM NaCl, 1.00 mM phenylmethylsulfonyl fluoride (PMSF) protease inhibitor) with 0.05% NP-40 and 1 \times BSA were preincubated on a rotator with GSH-sepharose beads (20.0 μ L; equilibrated with assay buffer) with a final concentration of 0.262 μ M for 1.5 h at 4 °C. After three washing steps of the beads with lysis buffer (assay buffer + 0.5% NP-40), the remaining proteins (WDR5 Δ 23, Ash2L, RbBP5) in assay buffer with 0.05% NP-40 and 1 \times BSA were added to the beads with a final concentration of 0.4 μ M. After incubation for 5 min on ice, the respective concentrations (200, 30, 5 μ M) of peptidomimetic **5** in DMSO (*cis* and *trans*: photoisomerization was performed as outlined in the Supporting Information) were added and the mixture was incubated on the rotator at 4 °C for 3 h. A negative control, to which no GST-MLL1 was added, and a positive DMSO biograde control were included in each assay. After 3 h, the beads were washed with lysis buffer (5 \times), incubated with SDS-loading dye (80 μ L; 0.5 M Tris pH 6.8), 5% SDS, 25% glycerol, 5% bromophenol blue in MilliQ water), and the proteins were denatured by incubation at 95 °C for 5 min. After SDS-PAGE (12%, following a protocol of the *Lab FAQs* from Roche), the bound proteins were detected by means of immunoblots (Western blot buffers, following a protocol of the *Lab FAQs* from Roche and the *Western Blotting Protocol* (Tank Transfer) from Sigma Aldrich) by using appropriate antibodies (rabbit antibodies from Bethyl, USA: anti-MLL1 (A300-375A), anti-WDR5 (A302-429A), anti-RbBP5 (A300-109A)).

Crystallization, data collection, structure determination and analysis

Crystallization of WDR5 Δ 23–4 Complex: WDR5 was concentrated to 200 μ M, mixed with 25.9 mm **4** to yield final concentrations of 194 μ M protein and 792.5 μ M **4** (ratio: 1:4 protein/peptide), and crystallized in SWISSCI MRC two-well crystallization plates (Jena

Bioscience). The reservoir solution volume was 50 μL , the drops contained 1 μL of a 1:1 mixture of protein and crystallization solution. Crystal growth took place at 4 $^{\circ}\text{C}$ in a solution containing 10% (w/v) poly(ethylene glycol) (PEG) 20000, 20% (v/v) PEG 550 monomethyl ether (MME), 0.02 M sodium formate, 0.02 M ammonium acetate, 0.02 M trisodium citrate, 0.02 M sodium potassium L-tartrate, and 0.02 M sodium oxamate. Crystals were obtained after 2 weeks and flash-cooled in liquid nitrogen without additional cryoprotectant. Diffraction data were collected at the European Synchrotron Radiation Facility (ESRF) by using beamline ID23-2 equipped with a microdiffractometer and Pilatus3 \times 2M detector. The crystals diffracted to 1.51 \AA with X-ray radiation of 14.2 keV (0.873 \AA). The data were processed by using XDS^[33] in space group P12₁. Data reduction and scaling were performed by using CCP4i2 (Version 7.0.065),^[34] with AIMLESS (Version 0.7.3).^[35] The phases were solved by molecular replacement with the previously determined structure of the WDR5 in complex with another peptidomimetic inhibitor (PDB ID: 5M23)^[19] by using PHASER.^[36] The model was built by using COOT (Version 0.8.9)^[37] and refinement in PHENIX (Version 1.11.1)^[38] until $R_{\text{work}} = 16.7\%$ and $R_{\text{free}} = 17.2\%$ were achieved. The peptidomimetic restraints were generated through ReadySet.^[38] Final refinement statistics are given in Table 3. All structural representation were generated by using PyMOL (Version 2.2). The coordinates and structure factors have been deposited with the Protein Data Bank under PDB ID: 6IAM.

Molecular docking: For VD of ligands into WDR5, GOLD 5.6 software^[39] was used. Docking of peptides as PPI inhibitors still represents a challenging task due to significant flexibility of peptides and challenges in identifications of hot-spot regions on protein surfaces.^[40] Although there are spatialized algorithms for the docking of peptides, the GOLD algorithm, which is primarily intended for small-molecule docking, has proven to be accurate and a fast alternative,^[41] especially in this particular case in which it is reasonable to assume that the ARA sequence retains a similar binding mode. Application of more specialized methodologies for peptide-protein docking was hampered by the fact that the photoswitch in our ligands was introduced in the middle of peptide sequence; thus making them unsuitable for these procedures. Exploiting the known SARA- sequence conformation obtained in PDB ID: 5M23, a docking procedure in GOLD was set in two steps, in which a fragment-based approach for molecular docking was employed. The docking protocol was set up in two steps: 1) Fragment SARA-Azo (Azo indicates azobenzenes/diazocines) was docked in PDB ID: 5M23 by using substructure constraints on the SARA- sequence. The SARA- sequence conformation was retained from the PDB structure, whereas Azo was manually added to the SARA- sequence and the final conformation for first docking was generated in a 20 ns MD run of this ligand with position restraints added to the SARA- sequence with Gromacs 5.1.4 software, as explained in more detail in the next section (MD calculations). 2) Docking of the whole molecule was performed by constraining the SARA-Azopose obtained in the previous docking step. The peptide fragment (-VHLRKS sequence) was built in DS Visualizer in the β -sheet conformation. The final conformation of the ligand VHLRKS sequence was generated after a 20 ns MD run of the ligand with position restraints added to the SARA- sequence, with the Gromacs 5.1.4 software, as explained in the next section (MD calculations). For molecular docking, the PLP scoring function was used, with 50 poses generated and maximum flexibility accounted for ligands. The binding site was selected as the area of 8 \AA around the ligand. Constraint weights were 20 for the SARA- sequence and 5 for azobenzene. Only poses with the smallest RMSD of the SARA-

sequence, relative to PDB ID: 5M23, were retained for further MD studies.

MD calculations: The MD protocol included a definition of azobenzenes as novel residues. Parameters for azobenzenes were obtained from the ParamChem web server^[42] and manually included in the Charmm36 force-field.^[43] Gromacs 5.1.4 software was used for all MD simulations. Protonation for protein-ligand complexes were determined with the PROPKA server.^[44] The protein was parametrized by the CHARMM36 force field for the amino acids of the ligands. The water model employed was TIP3P and an octahedron simulation box was used. The solvated system was preliminarily minimized by 5000 steps of steepest descent. The system was then heated to 310 K during 250 ps in an NVT ensemble with 1 fs time steps. Subsequently, the pressure was equilibrated to 1 atm during the 500 ps NPT simulation with 2 fs time steps. In the equilibration steps, harmonic positional restraints were set on the backbone of the protein with a spring constant of 1000 $\text{kJ mol}^{-1} \text{\AA}^{-2}$. The position restraints used in equilibration were gradually removed during 1 ns simulation with 2 fs time steps. The production run was performed in the NPT ensemble at 310 K without any restraint, except for simulations with ligands only where position restraints were kept on the -SARA- sequence. The Verlet cutoff scheme, the Nose-Hoover thermostat, the Parrinello-Rahman barostat, LINCS for constraints (all bonds with H atoms), and the particle mesh Ewald for electrostatics were applied. Upon completion of the simulation, a Gromos clustering process^[45] on the MD trajectory and relevant cluster representatives were obtained. MD simulations were used for refinement of complexes generated by molecular docking. Only simulations in which the ARA sequence remained stable in the active site were retained for further MM/PBSA analyses.

MM/PBSA scoring: The MM/PBSA method was used for postprocessing of data generated from MD simulations. This method, as an end-point free energy calculation methodology, has been proven to be a useful tool for postprocessing results obtained by means of molecular docking followed by MD refinement.^[46] Although this method is not accurate enough for the prediction of absolute binding free energies, mostly because MM/PBSA counts on severe thermodynamic approximations, it has been proven useful for the rationalization of experimental results on a series of similar ligands by ranking relative binding affinities.^[46,47] Here, we used 20 ns production MD runs for the refinement of obtained docking poses, and the GMXPBSA 2.1 tool for the calculation of binding energies from every 18th snapshot extracted in the last 7 ns of simulations (total of 21 snapshots per trajectory) with the single-trajectory approach. The dielectric constant of the solute was switched to one, whereas all other parameters were retained on default values. The entropic term was neglected from calculations.

Acknowledgements

We gratefully acknowledge Prof. Dr. Eric Meggers for S1 laboratory accessibility; Norbert Frommknecht for design assistance and construction of the LED lamps. Numerical simulations were run on the PARADOX-IV supercomputing facility at the Scientific Computing Laboratory, National Center of Excellence for the Study of Complex Systems, Institute of Physics, Belgrade, supported, in part, by the Ministry of Education, Science, and Technological Development of the Republic of Serbia under project no. ON171017. N.Dj., D.R. and K.N. acknowledge project no. 172033 of the Minis-

try of Science and Technological Development of the Republic of Serbia. Finally, we thank COST action CM1406 (Epigenetic Chemical Biology EPICHEMbio) for support.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: azobenzenes · isomerization · peptidomimetics · photochemistry · protein–protein interactions

- [1] a) L. Bonetta, *Nature* **2010**, *468*, 851–854; b) H. C. Lu, A. Fornili, F. Fraternali, *Expert Rev. Proteomics* **2013**, *10*, 511–520.
- [2] a) D. E. Scott, A. R. Bayly, C. Abell, J. Skidmore, *Nat. Rev. Drug Discovery* **2016**, *15*, 533–550; b) S. Surade, T. L. Blundell, *Chem. Biol.* **2012**, *19*, 42–50.
- [3] a) N. Sawyer, A. M. Watkins, P. S. Arora, *Acc. Chem. Res.* **2017**, *50*, 1313–1322; b) I. S. Moreira, O. Sensoy, *Curr. Top. Med. Chem.* **2018**, *18*, 645–646; c) T. Berg, *Angew. Chem. Int. Ed.* **2003**, *42*, 2462–2481; *Angew. Chem.* **2003**, *115*, 2566–2586; d) S. Gul, K. Hadian, *Expert Opin. Drug Discovery* **2014**, *9*, 1393–1404; e) L. G. Milroy, T. N. Grossmann, S. Hennig, L. Brunsveld, C. Ottmann, *Chem. Rev.* **2014**, *114*, 4695–4748; f) J. A. Wells, C. L. McClendon, *Nature* **2007**, *450*, 1001–1009; g) D. C. Fry, *Methods Mol. Biol.* **2015**, *1278*, 93–106; h) L. C. Cesa, A. K. Mapp, J. E. Gestwicki, *Front. Bioeng. Biotechnol.* **2015**, *3*, 119.
- [4] a) L. Nevola, E. Giral, *Chem. Commun.* **2015**, *51*, 3302–3315; b) M. Pelay-Gimeno, A. Glas, O. Koch, T. N. Grossmann, *Angew. Chem. Int. Ed.* **2015**, *54*, 8896–8927; *Angew. Chem.* **2015**, *127*, 9022–9054; c) A. Barnard, K. Long, H. L. Martin, J. A. Miles, T. A. Edwards, D. C. Tomlinson, A. Macdonald, A. J. Wilson, *Angew. Chem. Int. Ed.* **2015**, *54*, 2960–2965; *Angew. Chem.* **2015**, *127*, 3003–3008; d) P. Wójcik, L. Berlicki, *Bioorg. Med. Chem. Lett.* **2016**, *26*, 707–713.
- [5] a) R. J. Mart, R. K. Allemann, *Chem. Commun.* **2016**, *52*, 12262–12277; b) H. M. D. Bandara, S. C. Burdette, *Chem. Soc. Rev.* **2012**, *41*, 1809–1825.
- [6] a) M. M. Lerch, M. J. Hansen, G. M. van Dam, W. Szymanski, B. L. Feringa, *Angew. Chem. Int. Ed.* **2016**, *55*, 10978–10999; *Angew. Chem.* **2016**, *128*, 11140–11163; b) J. Broichhagen, J. A. Frank, D. Trauner, *Acc. Chem. Res.* **2015**, *48*, 1947–1960; c) Z. B. Mehta, N. R. Johnston, M. S. Nguyen-Tu, J. Broichhagen, P. Schultz, D. P. Larner, I. Leclerc, D. Trauner, G. A. Rutter, D. J. Hodson, *Sci. Rep.* **2017**, *7*, 291; d) A. A. Beharry, G. A. Woolley, *Chem. Soc. Rev.* **2011**, *40*, 4422–4437.
- [7] a) M. X. Dong, A. Babalhavaei, S. Samanta, A. A. Beharry, G. A. Woolley, *Acc. Chem. Res.* **2015**, *48*, 2662–2670; b) H. A. Wegner, *Angew. Chem. Int. Ed.* **2012**, *51*, 4787–4788; *Angew. Chem.* **2012**, *124*, 4869–4871; c) C. Knie, M. Utecht, F. L. Zhao, H. Kulla, S. Kovalenko, A. M. Brouwer, P. Saalfrank, S. Hecht, D. Blegler, *Chem. Eur. J.* **2014**, *20*, 16492–16501; d) M. Wegener, M. J. Hansen, A. J. M. Driessen, W. Szymanski, B. Feringa, *J. Am. Chem. Soc.* **2017**, *139*, 17979–17986; e) Z. Pianowski, *Chem. Eur. J.* **2019**, <https://doi.org/10.1002/chem.201805814>.
- [8] R. Siewertsen, H. Neumann, B. Buchheim-Stehn, R. Herges, C. Nather, F. Renth, F. Temps, *J. Am. Chem. Soc.* **2009**, *131*, 15594–15595.
- [9] a) H. Sell, C. Nather, R. Herges, *Beilstein J. Org. Chem.* **2013**, *9*, 1–7; b) S. Samanta, C. Qin, A. J. Lough, G. A. Woolley, *Angew. Chem. Int. Ed.* **2012**, *51*, 6452–6455; *Angew. Chem.* **2012**, *124*, 6558–6561.
- [10] A. A. Beharry, O. Sadvovskii, G. A. Woolley, *J. Am. Chem. Soc.* **2011**, *133*, 19684–19687.
- [11] a) M. Cigl, A. Bubnov, M. Kaspar, F. Hampl, V. Hamplova, O. Pacherova, J. Svoboda, *J. Mater. Chem. C* **2016**, *4*, 5326–5333; b) J. Karcher, Z. L. Pianowski, *Chem. Eur. J.* **2018**, *24*, 11605–11610.
- [12] J. Luo, S. Samanta, M. Convertino, N. V. Dokholyan, A. Deiters, *ChemBioChem* **2018**, *19*, 2178–2185.
- [13] F. Eljabu, J. Dhruval, H. B. Yan, *Bioorg. Med. Chem. Lett.* **2015**, *25*, 5594–5596.
- [14] D. C. Burns, F. Zhang, G. A. Woolley, *Nat. Protoc.* **2007**, *2*, 251–258.
- [15] a) M. Bose, D. Groff, J. M. Xie, E. Brustad, P. G. Schultz, *J. Am. Chem. Soc.* **2006**, *128*, 388–389; b) A. A. John, C. P. Ramil, Y. L. Tian, G. Cheng, Q. Lin, *Org. Lett.* **2015**, *17*, 6258–6261.
- [16] R. Behrendt, M. Schenk, H. J. Musiol, L. Moroder, *J. Pept. Sci.* **1999**, *5*, 519–529.
- [17] L. Ulysse, J. Chmielewski, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2145–2146.
- [18] D. Bléger, J. Schwarz, A. M. Brouwer, S. Hecht, *J. Am. Chem. Soc.* **2012**, *134*, 20597–20600.
- [19] L. Albert, J. Xu, R. W. Wan, V. Srinivasan, Y. L. Dou, O. Vazquez, *Chem. Sci.* **2017**, *8*, 4612–4618.
- [20] a) Z. Ahmed, A. Siiskonen, M. Virkki, A. Priimagi, *Chem. Commun.* **2017**, *53*, 12520–12523; b) M. D. Wendt, A. R. Kunzer, *Tetrahedron Lett.* **2010**, *51*, 641–644.
- [21] B. Sax, F. Dick, R. Tanner, J. Gosteli, *Pept. Res.* **1992**, *5*, 245–246.
- [22] D. Choukhi, M. Ciobanu, C. Zambaldo, V. Duplan, S. Barluenga, N. Winsinger, *Chem. Eur. J.* **2012**, *18*, 12698–12704.
- [23] K. Ralhan, V. G. KrishnaKumar, S. Gupta, *RSC Adv.* **2015**, *5*, 104417–104425.
- [24] G. B. Fields, *Methods Mol. Biol.* **1994**, *35*, 17–27.
- [25] S. Z. Li, G. Han, W. Q. Zhang, *Macromolecules* **2018**, *51*, 4290–4297.
- [26] a) H. R. López-Mirabal, J. R. Winther, *Biochim. Biophys. Acta Mol. Cell Res.* **2008**, *1783*, 629–640; b) C. Boulégué, M. Lowenack, C. Renner, L. Moroder, *ChemBioChem* **2007**, *8*, 591–594.
- [27] Y. Dou, T. A. Milne, A. J. Ruthenburg, S. Lee, J. W. Lee, G. L. Verdine, C. D. Allis, R. G. Roeder, *Nat. Struct. Mol. Biol.* **2006**, *13*, 713–719.
- [28] Z. Nikolovska-Coleska, R. Wang, X. Fang, H. Pan, Y. Tomita, P. Li, P. P. Roller, K. Krajewski, N. G. Saito, J. A. Stuckey, S. Wang, *Anal. Biochem.* **2004**, *332*, 261–273.
- [29] H. Karatas, Y. Li, L. Liu, J. Ji, S. Lee, Y. Chen, J. Yang, L. Huang, D. Bernard, J. Xu, E. C. Townsend, F. Cao, X. Ran, X. Li, B. Wen, D. Sun, J. A. Stuckey, M. Lei, Y. Dou, S. Wang, *J. Med. Chem.* **2017**, *60*, 4818–4839.
- [30] A. J. Ruthenburg, W. Wang, D. M. Graybosch, H. Li, C. D. Allis, D. J. Patel, G. L. Verdine, *Nat. Struct. Mol. Biol.* **2006**, *13*, 704–712.
- [31] a) N. L. Alicea-Velázquez, S. A. Shinsky, D. M. Loh, J. H. Lee, D. G. Skalnik, M. S. Cosgrove, *J. Biol. Chem.* **2016**, *291*, 22357–22372; b) J. J. Song, R. E. Kingston, *J. Biol. Chem.* **2008**, *283*, 35258–35264; c) A. Patel, V. Dharmarajan, M. S. Cosgrove, *J. Biol. Chem.* **2008**, *283*, 32158–32161.
- [32] D. Bléger, S. Hecht, *Angew. Chem. Int. Ed.* **2015**, *54*, 11338–11349; *Angew. Chem.* **2015**, *127*, 11494–11506.
- [33] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132.
- [34] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 235–242.
- [35] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1204–1214.
- [36] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- [37] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 486–501.
- [38] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221.
- [39] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727–748.
- [40] M. Ciemny, M. Kurcinski, K. Kamel, A. Kolinski, N. Alam, O. Schueler-Furman, S. Kmiecik, *Drug Discovery Today* **2018**, *23*, 1530–1537.
- [41] a) A. S. Hauser, B. Windshugel, *J. Chem. Inf. Model.* **2016**, *56*, 188–200; b) C. Lamm, C. Zanoni, G. Aiello, A. Arnoldi, G. Grazioso, *Sci. Rep.* **2016**, *6*, 29931.
- [42] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell, Jr., *J. Comput. Chem.* **2010**, *31*, 671–690.
- [43] J. Huang, A. D. MacKerell, Jr., *J. Comput. Chem.* **2013**, *34*, 2135–2145.
- [44] M. H. Olsson, C. R. Sondergaard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 525–537.

- [45] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, A. E. Mark, *Angew. Chem. Int. Ed.* **1999**, *38*, 236–240; *Angew. Chem.* **1999**, *111*, 249–253.
- [46] T. Hou, J. Wang, Y. Li, W. Wang, *J. Chem. Inf. Model.* **2011**, *51*, 69–82.
- [47] a) H. Y. Sun, Y. Y. Li, M. Y. Shen, S. Tian, L. Xu, P. C. Pan, Y. Guan, T. J. Hou, *Phys. Chem. Chem. Phys.* **2014**, *16*, 22035–22045; b) S. Genheden, U. Ryde, *Expert Opin. Drug Discovery* **2015**, *10*, 449–461.

Manuscript received: November 29, 2018

Accepted manuscript online: January 24, 2019

Version of record online: April 25, 2019

CHEMBIOCHEM

Supporting Information

Modulating Protein–Protein Interactions with Visible-Light-Responsive Peptide Backbone Photoswitches

Lea Albert,^[a] Alberto Peñalver,^[a] Nemanja Djokovic,^[b] Laura Werel,^[a] Malte Hoffarth,^[a]
Dusan Ruzic,^[b] Jing Xu,^[c] Lars-Oliver Essen,^[a] Katarina Nikolic,^[b] Yali Dou,^[c] and
Olalla Vázquez*^[a]

cbic_201800737_sm_miscellaneous_information.pdf

Table of Contents

Abbreviations	II
Experimental Procedures	1
Materials.....	1
Nuclear Magnetic Resonance Spectroscopy (NMR)	1
Mass Spectrometry	1
High Performance Liquid Chromatography (HPLC)	2
UV-Vis Spectroscopy	2
Photoisomerization of <i>o</i> F ₄ Azo and <i>c</i> Azo Peptides	2
Fluorescence Polarization-based Assay	2
Transformation, Expression and Purification of Proteins	3
Synthesis	6
Synthesis of Fmoc-protected cyclic azobenzene amino acid (Fmoc- <i>c</i> AzoAA, 1)	6
Synthesis of Mtt-protected tetra- <i>ortho</i> -fluoroazobenze amino acid (Mtt- <i>o</i> F ₄ AzoAA, 2)	11
Solid Phase Peptide Synthesis.....	15
Synthesis and Characterization Data of the Peptides	17
Results and Discussion	19
Screening for Suitable Fmoc-deprotection Conditions	19
Extinction Coefficients	22
Stability against Glutathione.....	25
Fluorescence Polarization-based Assay	26
Crystallisation of WDR5Δ23-Peptide 4 Complex.....	28
Docking and Molecular Dynamics (MD) Studies	30
References	36

Abbreviations

ε	extinction coefficient
λ	wavelength
ν	frequency
A	anisotropy
a.u.	aberration units
Ahx	aminohexanoic acid
Alloc	alkyloxycarbonyl
Azo	azobenzene
Boc	<i>tert</i> -butylcarbonyl
c	concentration
cAzo	cyclic azobenzene
calcd.	calculated
conc.	concentration
CuCN	copper (I) cyanide
DBU	1,8-diazabicyclo[5,4,0]undeca-7-ene
DIC	<i>N, N'</i> -diisopropylcarbodiimide
DIPEA	<i>N, N'</i> -diisopropylethylamine
DMF	<i>N, N'</i> -dimethylformamide
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EtOAc	ethyl acetate
EDTA	ethylenediaminetetraacetic acid
ESI	electrospray ionization
eq.	equivalents

oF_4Azo	tetra- <i>ortho</i> -fluorazobenzene
FAM	fluorescein
Fmoc	fluorenylmethyloxycarbonyl chloride
FP	fluorescence polarization
GSH	glutathione
HFIP	hexafluoroisopropanol
HMT	histone methyltransferase
HPLC	high-performance liquid chromatography
HRMS	high resolution mass spectrometry
IC_{50}	half inhibitory concentration
IPTG	isopropyl β -D-1-thiogalactopyranoside
irr.	irradiation
K_d	dissociation constant
K_i	inhibition constant
M	molar
mAU	milli absorbance units
MD	Molecular Dynamics
MLL1	mixed lineage leukaemia 1
mP	milli polarization value
MPD	2-methyl-2,4-pentanediol
Mtt	4-methyltrityl
NMP	<i>N</i> -methyl-2-pyrrolidon
NMR	nuclear magnetic resonance
Ni-NTA	nickel nitrilotriacetic acid
PMSF	phenylmethylsulfonyl fluoride
PSS	photostationary state

PyBop	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
R_f	retention factor
RbBP5	retinoblasma-bindin-protein 5
RP	reversed phase
rpm	revolutions per minute
r.t.	room temperature
SAM	S-adenosylmethionine
sec.	seconds
SPPS	solid phase peptide synthesis
t_0	time zero
t_R	retention time
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TIS	triisopropyl silane
Tris	tris(hydroxymethyl)aminomethane
Trt	trityl
UV	ultraviolet
vis	visible
WDR5	WD-repeat protein-5

Experimental Procedures

Materials

All commercially reagents were purchased from the following companies and used without further purification: DMSO biograde, NaCl biograde, Tris, 2-mercaptoethanol, Fmoc-Cl, FeSO₄ × 7 H₂O and DIPEA from Carl Roth (Germany); 6-aminohexanoic acid and KMnO₄ from Alfa Aesar (Germany); OxymaPure from Luxembourg Bio Technologies (Israel); DIC, ammonium persulfate and glycerol from Acros (USA); Sieber amide resin, Mtt-Cl, 4-bromo-2,6-difluoraniline and MBHA resin from Fluorochem (UK); glucose monohydrate from Applichem (Germany); CuCN and HCl from TCI (Japan); PyBop from Novabiochem (Germany); KCl biograde, Triton, 2-Cl-Trt resin, NaNO₃, NaOH, KOH, 4,4'-ethylenedianiline, IPTG, ampicillin and dithiothreitol from Goldbio (USA); LB-broth, LB-agar, bromphenolblue and Tween-20 from Thermo Fisher (USA); ZnSO₄ from Fluka (USA); acrylamide from National diagnostics (USA); glutathione-sepharose beads from GE-Healthcare (USA); protease inhibitor from Roche diagnostics (Germany), S-adenosylmethionine from Biolabs (USA); *tert*-butyl-2-bromoacetate, ethylenediamine, PMSF, kanamycin, NP-40 and MgSO₄, from Sigma Aldrich/ Merck (Germany); Fmoc-protected amino acids, DMF peptide grade, HATU and TentaGel S RAM resin from Iris Biotech (Germany); MeCN HPLC grade, pyridine, H₂SO₄ and triethylamine from VWR (France). The H₃₁₋₂₁ peptide and dimethylated H₃₁₋₂₁ peptide were purchased from Anaspec (USA, ref. codes: AS-61702; AS-64356); The protein marker (ref. nr.: 1610373) and BSA stocks (ref. nr.: 500-0207) from Bio-Rad (USA). The plasmid preparation kit was purchased from Quiagen (ref. code: 27106, USA) and the corresponding mini-prep column from Denville Scientific (ref. nr.: CM-400-250, USA). Beads and buffers used for AlphaLISA Assay were purchased from Perkin Elmer (USA): 5 × buffer: AlphaLISA 5 × epigenetics buffer 1 (AL008F1); 30 × buffer: AlphaLISA 30 × epigenetics buffer supplement (AL008L2); acceptor bead: AlphaLISA anti-H3K4me1-2 acceptor beads (AL116M); donor beads: AlphaScreen streptavidin donor beads (6760002). Water was purified with a Milli-Q Ultra Pure Water Purification System (TKA, Germany).

Nuclear Magnetic Resonance Spectroscopy (NMR)

NMR spectra were recorded at 300 K either on a Bruker AV III HD 300 MHz at a frequency of 300 MHz (¹H), 75 MHz (¹³C, ¹³C-APT) or on a Bruker AV III HD 500 MHz at a frequency of 500 MHz (¹H), 125 MHz (¹³C) or on a Bruker AV III HD 600 MHz at a frequency of 600 MHz (¹H). The ¹H, ¹³C and ¹³C-APT NMR spectra were referenced to solvent residue peaks. As internal standards, deuterated chloroform (CDCl₃), deuterated dimethyl sulfoxide ([D₆]DMSO) with TMS or deuterated water (D₂O) were used. Solvent shifts (ppm): (CDCl₃, δ) = 7.26 ppm (¹H) and 77.16 ppm (¹³C), ([D₆]DMSO, δ) = 2.50 ppm (¹H) and 39.52 (¹³C), D₂O = 4.79 ppm (¹H).

Mass Spectrometry

High resolution electrospray ionization (ESI) mass spectra were acquired with a LTQ-FT Ultra mass spectrometer (Thermo Fischer Scientific). The resolution was set to 100,000. Field desorption (FD) time-of-flight (FD-TOF/MS) mass spectra were acquired with an AccuTOF GCv 4G (JEOL) Time of Flight (TOF) mass spectrometer. FD-emitters were purchased from Linden Chromaspec GmbH (Bremen, Germany).

High Performance Liquid Chromatography (HPLC)

For preparative purposes, a PLC 2020 Personal Purification System (Gilson) with a preparative Nucleodur C18 HTec-column (5 μm , 250 x 16 mm; Macherey & Nagel) with a flow rate of 10 mL/min or a semipreparative C18 3001 column (Phenomenex) with a flow rate of 8 mL/min using an isocratic regime during the first five minutes, for column equilibration, followed by the respectively stated linear gradient in 30 min at 25°C. The detection was carried out by measuring the absorption at wavelengths: 220 nm and 260 nm. Milli-Q water (A) and MeCN (B) were employed as eluents with an addition of 0.1% of TFA in both solvents.

Analytical HPLC-MS were recorded on an Agilent 1200 Series HPLC-system (Agilent Technologies), with a column oven set to 55 °C. Detection was monitored at 220 nm. Milli-Q water (A) and MeCN (B) were employed as eluents with an addition of 0.05% and 0.03% of TFA. Different analytical columns and methods were used:

- column 1: EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) using the stated linear gradient in 30 min with a flow rate of 0.20 mL/min
- column 2: Eclipse XBD-C18 column (Agilent Technologies) using the stated linear gradient in 30 min with a flow rate of 1.00 mL/min
- column 3: Kinetex 5 μm XB-C18 100Å LC column 150 x 4.6 mm (Phenomenex) using the stated linear gradient in 30 min with a flow rate of 1.00 mL/min

UV-Vis Spectroscopy

Concentration determinations and UV-vis measurements were performed on a Tecan (Switzerland) Spark 20M multimode microplate reader at r.t. All measurements were performed in a 1400 μL quartz cuvette (Hellma Analytics (104F-QS) with a pathlength of 1 cm. The, to be measured, compound was added to 700 μL of the respective blank measurement in the cuvette and absorbance spectra were recorded. To determine the concentration of the peptides the following extinction coefficients were used: *o*F₄Azo-containing peptides (for *trans* isomer) $\epsilon_{319} = 25000 \text{ L mol}^{-1} \text{ cm}^{-1}$ in MeCN,^[1] *c*Azo-containing peptides (for *cis* isomer) $\epsilon_{238} = 25990 \text{ L mol}^{-1} \text{ cm}^{-1}$ in MilliQ water, fluorescently tagged peptides $\epsilon_{494} = 76900 \text{ L mol}^{-1} \text{ cm}^{-1}$ in 0.1 M phosphate buffer, pH = 9^[2]).

Photoisomerization of *o*F₄Azo and *c*Azo Peptides

For isomerization an UV-Emitter 405 nm SMD Roschwege Star-UV405-03-00-00, 3 W (to reach the *trans* photostationary state) and a HighPower-LED Green 87 lm 130° 3.8 V, 1000 mA Roschwege (to reach the *cis* photostationary state) were used for irradiation at 405 and 525 nm, respectively. For the irradiation at 525 nm an additional cut-on filter (OG515 coloured glass filter, 515 nm longpass, Thorlabs) was applied. Experiments were performed at r.t.

Fluorescence Polarization-based Assay

All FP-based assays were performed at r.t. in black 96-well microtiter plates (Greiner) and the FP was measured as millipolarization (mP) units on a Tecan (Switzerland) Spark 20M multimode microplate reader; settings: used filters, excitation: 485 nm; emission: 530 nm; gain: optimal; Z-position: calculated from control well (0% inhibition: tracer and protein in assay

buffer); number of flashes: 30; G-factor: 1.067. The data evaluation was done, using GraphPad Prism 6 software.

Competitive Fluorescence Polarization-based Binding Assay

The binding affinities of both isomers of each peptide were measured separately as reported by us previously.^[3] The isomerization, however, was done differently as detailed above. The peptides were dissolved in DMSO biograde. In order to calculate the binding affinity of the *trans* isomers, the peptide solutions were irradiated at 405 nm for 10 sec. To determine the binding affinity of the *cis* isomers, the peptide solutions were irradiated at 525 nm for 2 min in a transparent glass vial, previously to the assay. Afterwards, the whole assay had to be done in darkness to prevent *cis* → *trans* and *trans* → *cis* relaxation. 120 µL of a preincubated complex solution of WDR5Δ23 and the tracer peptide in assay buffer (0.1 M phosphate buffer pH = 6.5, 25 mM KCl, 0.01% Triton) were added to 5 µL of dilutions of the respective peptide in DMSO in triplicate, reaching the final concentrations of 72 nM for WDR5Δ23 and 20 nM for the tracer peptide. Three control wells in triplicate were included in each plate: blank (without tracer and protein), 100% inhibition (tracer only) and 0% inhibition (protein and tracer). The plates were incubated on a shaker (Edmund Bühler TiMix Control TH15) at r.t. for 5 h and the mP values were recorded. GraphPad Prism 6 software was used for data evaluation. K_i values were calculated using the following equation described previously by Wang et al. and the corresponding webpage provided by them:^[4]

$$K_i = \frac{[I]_{50}}{\frac{[L]_{50}}{K_d} + \frac{[P]_0}{K_d} + 1}$$

where $[I]_{50}$ denotes the concentration of the free inhibitor at 50 % inhibition, $[L]_{50}$ is the concentration of free labelled ligand at 50 % inhibition, $[P]_0$ is the concentration of the free protein at 0 % inhibition, and K_d is the dissociation constant of the protein-ligand complex.

Transformation, Expression and Purification of Proteins

Cultivation of Pre-cultures

Pre-cultures of the pET28 plasmids of WDR5Δ23, MLL1(3762-end), Ash2L (1-534), RbBP5 (1-538) and the pGEX plasmid of GST-MLL1, used for inoculation of heterologous expression and starting material for plasmid preparation were prepared in test tubes with 5.00 mL LB medium with kanamycin for the pETH28 plasmids and with ampicillin for the pGEX plasmid. Then, either a colony from a plate or some cell material from a glycerol stock was scratched off with a toothpick and added to the test tube. The culture was incubated on a shaker at 37 °C with 200 rpm for 16 h. For better oxygen transfer, test tubes were attached in an inclined way. Glycerol stocks were prepared by mixing 700 µL preculture with 300 µL 80% glycerol. They were flash frozen in liquid nitrogen and stored at -80 °C.

Transformation

Transformation was performed thawing aliquots of competent *E. coli* BL21 codon + cells on ice. To 25 µL of these cells, 1 µL of the respective plasmids were added. The mixture was incubated for 30 min on ice, followed by a heat shock in a 42 °C water bath for 45 sec. After

2 min incubation on ice, 250 μ L of LB medium was added and the suspension was shaken at 37 °C, 200 rpm for 1 h. Then, 50 μ L were plated on LB-agar-kanamycin/ampicillin plates and incubated for 16 h at 37 °C.

Plasmid Preparation

The plasmid preparation was performed by the use of the Quiagen kit (see Materials section). Therefore, a 5.00 mL culture is prepared with the particular plasmid containing cells and incubated at 37 °C for 16 h. After harvesting the cells, the pellet is lysed with 250 μ L resuspension buffer P1, which consists of RNase that cleaves RNA and EDTA, which complexes the cations that normally stabilize the cell membrane. By adding 250 μ L lysis buffer P2, which contains SDS and NaOH to disrupt the cell membrane and denatures dsDNA. Due to their *supercoil* structure, plasmids are, in contrast to genomic DNA, able to recover the denaturation. After mixing and incubation for 5 min, 350 μ L neutralization buffer N3 was added. Such buffer containing KOAc, neutralizes the pH and leads to precipitation of chromosomal DNA and proteins. The precipitant is pelleted during a 10 min centrifugation (13000 rpm) step. The supernatant is then transferred onto a small column, washed once with 500 μ L PB buffer, once with 750 μ L PE buffer and finally eluted with 30 μ L EB buffer or MilliQ water. Concentrations were measured photometric at 260 nm, which is the absorption maxima of nucleic acids.

Expression and Purification of His-SUMO tagged Proteins

The pET28a-based His-Sumo-Proteins (WDR5 Δ 23, MLL1 (3762-end), Ash2L (1-534), RbBP5 (1-538)) were expressed from the pET28-MHL vector in *E. coli* BL21 codon+ cells. Expression was performed by inoculation of 800 mL LB medium into 2 L baffled flasks with 50 μ g/mL kanamycin and 50 mL of pre-culture. For the expression of MLL1 and Ash2L, ZnSO₄ in MilliQ water to a final concentration of 10 μ M was added to ensure proper folding of the enzyme. The culture was incubated on a shaker (Multitron from Infors HT) at 37 °C and 200 rpm until an OD₆₀₀ of 0.5-0.6 was reached. The temperature was then reduced to 20 °C and the expression was induced by adding IPTG up to a final concentration of 0.3 mM. The cells were grown for 16 h and harvested through centrifugation (5000 rpm, 15 min, 4 °C, Sorvall RC6 Plus from Thermo Scientific). The cell pellet from the 800 mL expression culture was resuspended in 40 mL lysis buffer (400 μ L PMSF (stock: 100 mM), 400 μ L 10% NP-40, 4 μ L 2-mercaptoethanole and one tablet of protease inhibitor (Roche diagnostics) cocktail in 40 mL BC500). Lysis of the cells was carried out using a sonicator Sonoplus HD 220 from Bandelin (3 x 40%, output: 4, with 2 min pulses and 3 min rest). After sonification, the suspension was pelleted by centrifugation (12000 rpm, 20 min, 4 °C). The lysate was added to Ni-NTA agarose (Quiagen, Germany) and incubated for 1 h at 4 °C. Cleavage of the His₆-tag was achieved by incubation of the resin with 3.0 mL BC500 + 18.1 μ L ULP-His₆. Proteins used for the HMT-Assay were used without further purification.

WDR5 Δ 23 used for the FP-binding assays and for crystallization was further purified by loading onto a Mono S 5/50 GL column (GE Healthcare) using an ÄKTA purification system (GE Healthcare). The flow rate was 1 mL/min with the Mono S buffers A (25 mM Tris, pH = 8.0, 20% glycerol) and B (25 mM Tris, pH = 8.0, 1 M NaCl, 20% glycerol) with first an isocratic flow of 100% A for 10 min followed by a linear gradient of 0-100% B in 20 min. WDR5 Δ 23 eluted between 35-40% of buffer B. The fractions were checked via SDS-PAGE and pooled. The

buffer was changed to the dialysis buffer (25 mM Tris, pH = 8.0, 150 mM NaCl) and the protein was concentrated to 200 μ M. The concentrated protein was aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

Expression and Purification of GST tagged Proteins

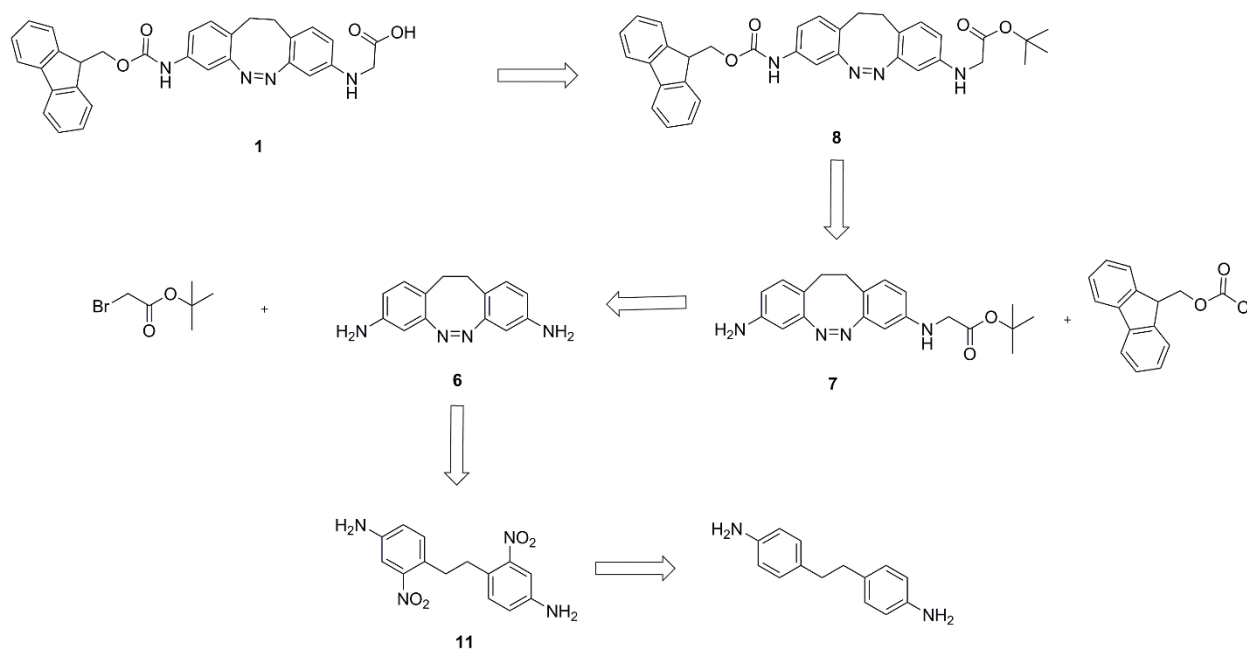
The pGEX-based GST-proteins were expressed in *E. coli* BL21 codon+ cells. Expression was performed by inoculation of 800 mL LB medium into 2 L baffled flasks with 50 μ g/mL ampicillin and 50 mL of pre-culture. For the expression of GST-MLL1, 10 μ M ZnSO₄ was added to ensure proper folding of the enzyme. The culture was incubated on a shaker (Multitron from Infors HT) at 37 °C and 200 rpm until an OD₆₀₀ of 0.5-0.6 was reached. The temperature was then reduced to 20 °C and the expression was induced by adding IPTG up to a final concentration of 0.2 mM. The cells were grown for 16 h and harvested through centrifugation (5000 rpm, 15 min, 4 °C, Sorvall RC6 Plus from Thermo Scientific). The cell pellet from the 800 mL expression culture was resuspended in 40 mL lysis buffer (400 μ L PMSF (stock: 100 mM), 400 μ L 10% NP-40, 4 μ L 2-mercaptoethanole and one tablet of protease inhibitor cocktail in 40 mL BC350 (25 mM Tris, pH = 8.0, 350 mM NaCl, 20% glycerol)). Lysis of the cells was carried out using a sonicator Sonoplus HD 220 from Bandelin (3 x 40%, output: 4, with 2 min pulses and 3 min rest). After sonification, the suspension was pelleted by centrifugation (12000 rpm, 20 min, 4 °C). The lysate was added to glutathione-sepharose beads (GE Healthcare, USA) and incubated for 1 h at 4 °C. Cleavage of the GST-tag was achieved by incubation of the resin with 1.0 mL BC350 + 10 μ M reduced glutathione for 1 h at 4°C. This step was repeated twice. Finally, the protein was dialysed with BC-350 without glycerol to wash out remaining glutathione.

Synthesis

The names of the chemical compounds were generated with ChemDraw Professional (2017, Perkin Elmer). The commercial compounds do not have numbers, only the synthesized ones.

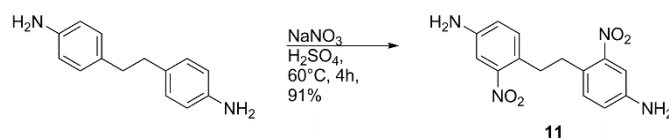
Synthesis of Fmoc-protected cyclic azobenzene amino acid (Fmoc-cAzoAA, 1)

The retrosynthesis route of the **Fmoc-cAzoAA 1** is illustrated in Scheme S1. In the Supporting Information only the procedures for the literature reported compounds are shown. The procedures for the novel compounds are detailed in the Manuscript, and the NMR spectra are shown here in the Supporting Information.



Scheme S1. Retrosynthesis route of the Fmoc-protected cyclic azobenzene amino acid **1** (**Fmoc-cAzoAA**).

4,4'-(ethane-1,2-diyl)bis(3-nitroaniline) (**11**)

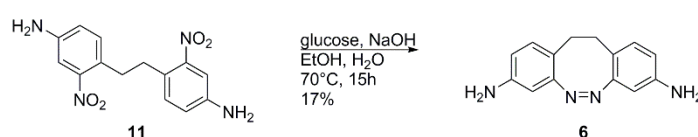


Scheme S2. Nitration of 1,2-bis(4-aminophenyl)ethane to the compound **11**.

Following the procedure of Sell et al.^[5], the 1,2-bis(4-aminophenyl)ethane (1.00 g, 4.71 mmol, 1.00 eq.) was dissolved in 8.50 mL H_2SO_4 (conc.) and was heated to 60°C . A solution of NaNO_3 (881 mg, 10.4 mmol, 2.20 eq.) in 9.10 mL of H_2SO_4 (conc.) was added. After stirring at

60 °C for 4 h the solution was poured into 100 mL of ice water. The resulting green suspension was neutralized by the addition of an aqueous ammonia solution (25%). The orange precipitate was filtered off, washed with distilled water and dried in vacuo. The product **11** was obtained as an orange solid (1.30 g, 4.30 mmol, 91%). The characterization is in agreement with the literature.^[5] **TLC:** $R_f = 0.05$ ($\text{CH}_2\text{Cl}_2/\text{MeOH}$ 1:1 + 1% NEt_3) **¹H NMR (300 MHz, DMSO, δ):** 7.06 (d, $^3J = 2.4$ Hz, 2H, 2 \times CH_{ar}), 6.99 (d, $^3J = 8.3$ Hz, 2H, 2 \times CH_{ar}), 6.80 (d, $^3J = 2.4$ Hz, 1H, CH_{ar}), 6.77 (d, $^3J = 2.4$ Hz, 1H, CH_{ar}), 5.58 (s, 4H, 2 \times NH_2), 2.86 (s, 4H, 2 \times CH_2). **¹³C NMR (75 MHz, DMSO, δ):** 149.4 (2 \times C_{ar}), 148.1 (2 \times C_{ar}), 132.3 (2 \times C_{ar}), 121.4 (2 \times C_{ar}), 118.8 (2 \times C_{ar}), 108.0 (2 \times C_{ar}), 32.7 (2 \times CH_2). **HRMS-ESI+ (m/z):** calcd. for $[\text{M}+\text{Na}]^+$ $\text{C}_{14}\text{H}_{14}\text{N}_4\text{O}_4\text{Na}$: 325.0907; found: 325.0901.

(Z)-11,12-dihydrodibenzo[c,g][1,2]diazocine-3,8-diamine (**6**)



Scheme S3. Cyclization to form the cAzo derivative **6**.

Following the procedure of Sell at al.^[5], compound **11** (800 mg, 2.65 mmol, 1.00 eq.) was dissolved in 106 mL EtOH. NaOH (6.65 g, 166 mmol, 62.8 eq.) in 26.5 mL distilled water was added and the mixture was warmed up to 70 °C. Then, glucose monohydrate (5.39 g, 27.2 mmol, 10.3 eq.) in 15 mL of distilled water was added. The solution was stirred at 70 °C for 15 h. The reaction was quenched by addition of 250 mL distilled water. The product was extracted with EtOAc (6 \times 100 mL), washed with brine and dried over anhydrous MgSO_4 . The crude was purified by flash column chromatography (pentane/EtOAc 1:1). The product **6** was obtained as a yellow solid (107.5 mg, 0.451 mmol, 17%). The characterization is in agreement with the literature.^[5] **TLC:** $R_f = 0.38$ (pentane/EtOAc 1:1) **¹H NMR (300 MHz, CDCl_3 , δ):** 6.73 (d, $^3J = 8.1$ Hz, 2H, 2 \times CH_{ar}), 6.35 (d, $^3J = 2.4$ Hz, 1H, CH_{ar}), 6.32 (d, $^3J = 2.4$ Hz, 1H, CH_{ar}), 6.14 (d, $^3J = 2.4$ Hz, 2H, 2 \times CH_{ar}), 3.55 (s, 4H, 2 \times NH_2), 2.85-2.78 (m, 2H CH_2), 2.62-2.55 (m, 2H CH_2). **¹³C NMR (75 MHz, CDCl_3 , δ):** 156.4 (2 \times C_{ar}), 144.9 (2 \times C_{ar}), 130.6 (2 \times CH_{ar}), 118.6 (2 \times C_{ar}), 114.2 (2 \times CH_{ar}), 105.3 (2 \times CH_{ar}), 31.2 (2 \times CH_2). **HRMS-ESI+ (m/z):** calcd. for $[\text{M}+\text{H}]^+$ $\text{C}_{14}\text{H}_{14}\text{N}_4\text{H}$: 239.1291; found: 239.1291.

NMRs of tert-butyl (Z)-(8-amino-11,12-dihydrodibenzo[c,g][1,2]diazocin-3-yl)glycinat (**7**)

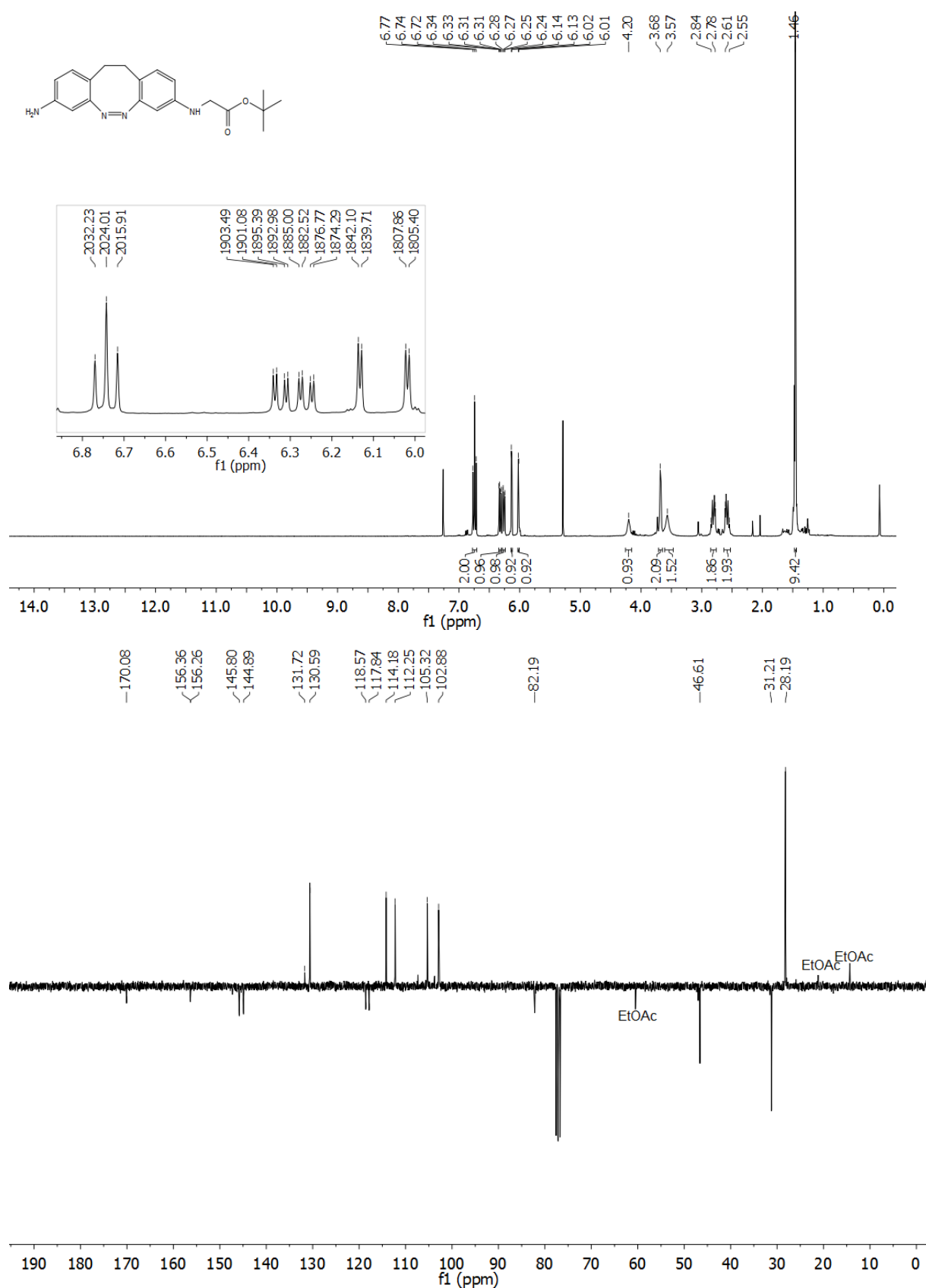


Figure S1. Top: ¹H-NMR spectrum of the compound **7**. Insert: zoom of signals in the aromatic region to highlight the shifts in Hz units for calculating *J*-constants. Bottom: ¹³C-APT-NMR spectrum of compound **7**.

NMRs of tert-butyl (Z)-(8-Fmoc-11,12-dihydrodibenzo[c,g][1,2]diazocin-3-yl)glycinate (**8**)

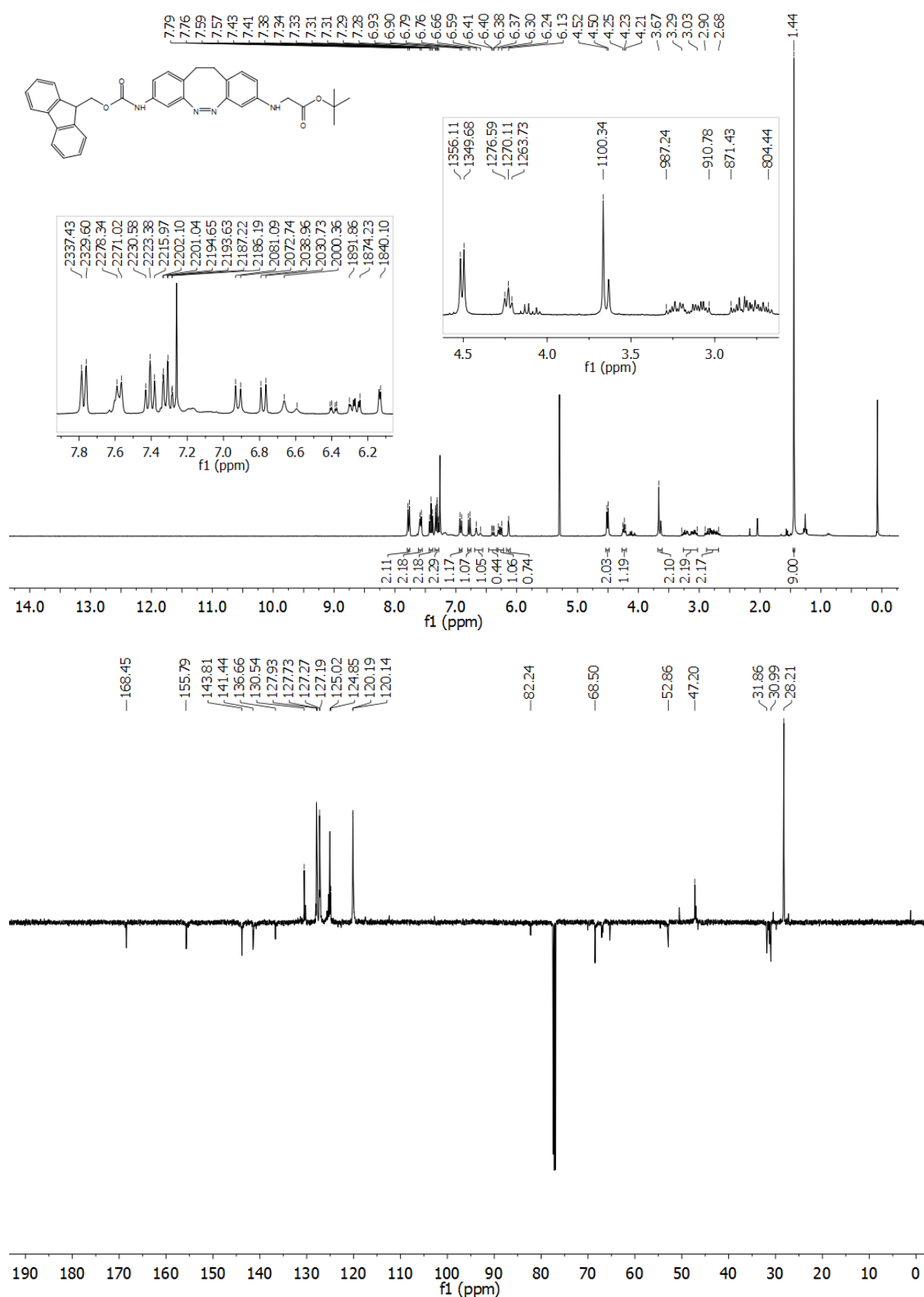


Figure S2. Top: ¹H-NMR spectrum of the compound **8**. Insert: zoom-in of signals in the aromatic and CH region to highlight the shifts in Hz units for calculating *J*-constants. Bottom: ¹³C-APT-NMR spectrum of compound **8**.

NMRs of (Z)-(8-Fmoc-11,12-dihydrodibenzo-[c,g][1,2]diazocin-3-yl)glycinate (**1**)

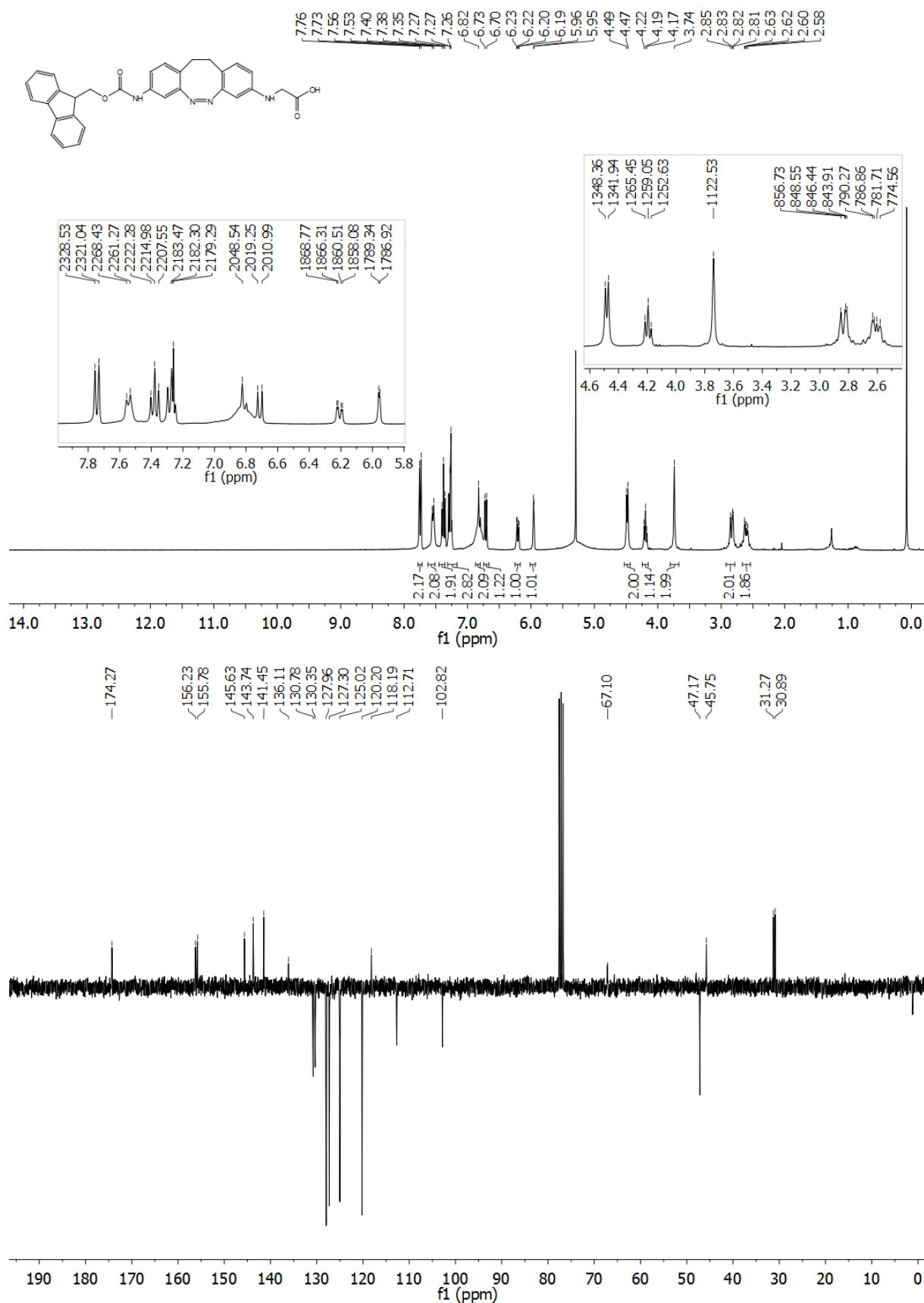
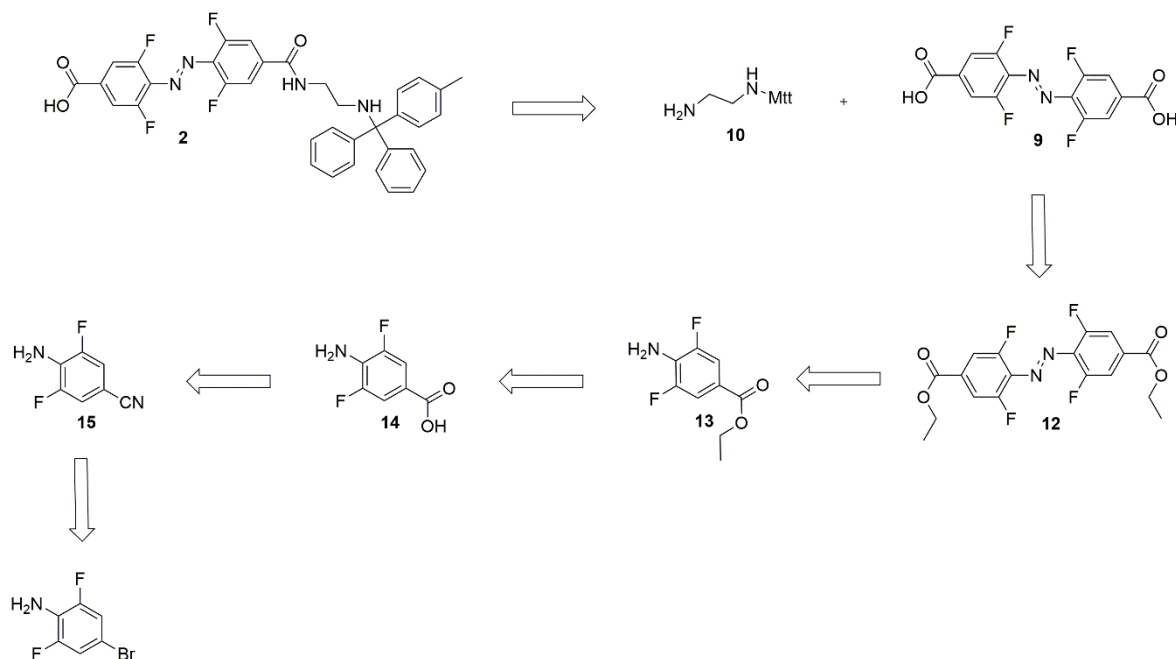


Figure S3. Top: ¹H-NMR spectrum of the compound **1**. Insert: zoom-in of signals in the aromatic and CH region to highlight the shifts in Hz units for calculating *J*-constants. Bottom: ¹³C-APT-NMR spectrum of compound **1**.

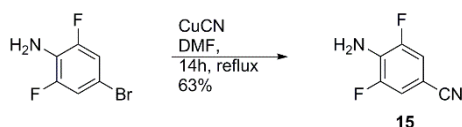
Synthesis of Mtt-protected tetra-*ortho*-fluoroazobenze amino acid (**Mtt-oF₄AzoAA**, **2**)

The retrosynthesis route of **Mtt-oF₄AzoAA** (**2**) is illustrated in Scheme S4. In the Supporting Information only the procedures for the literature reported compounds are shown. The procedures for the novel compounds are detailed in the Manuscript, and the NMR spectra are shown here in the Supporting Information.



Scheme S4. Retrosynthesis route for the Mtt-protected tetra-*ortho*-fluoroazobenze amino acid (**2**) (**Mtt-oF₄AzoAA**).

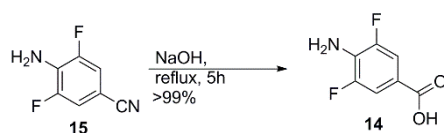
4-amino-3,5-difluorobenzonitrile (**15**)



Scheme S5. Synthesis of oF₂Azo derivative **15**.

Following the procedure of Bléger et al.^[6], the 4-bromo-2,6-difluoroaniline (5.00 g, 24.0 mmol, 1.00 eq.) and CuCN (6.45 g, 72.0 mmol, 3.00 eq.) were dissolved in 50.0 mL DMF and heated to 160 °C for 16 h. After cooling to r.t., the mixture was poured into a NH₃ aqueous solution (12 %, 250 mL), filtered and the precipitate was washed several times with EtOAc. The organic phases were washed with the same NH₃ solution, distilled water, brine and dried over anhydrous MgSO₄. The crude was purified by flash column chromatography (CH₂Cl₂/*n*-pentane 2:1) to yield the product **15** (2.33 g, 15.1 mmol, 63%) as a white solid. **TLC**: R_f = 0.39 (CH₂Cl₂/pentane 2:1). The characterization is in agreement with the literature.^[6] **¹H NMR (300 MHz, CDCl₃, δ)**: 7.13 (dd, ³J = 6.0 Hz, ⁴J = 2.2 Hz, 2H, 2 × CH_{ar}), 4.29 (s, 2H, NH₂). **¹³C NMR (75 MHz, CDCl₃, δ)**: 152.3 (F_{Car}), 149.1 (F_{Car}), 129.7 (C_{ar}), 118.0 (CN), 115.7 (C_{ar}H), 115.4 (C_{ar}H); 98.4 (C_{ar}). **HRMS-EI+ (m/z)**: calcd. for [M+H]⁺ C₇H₄F₂N₂, 154.03425; found, 154.03252.

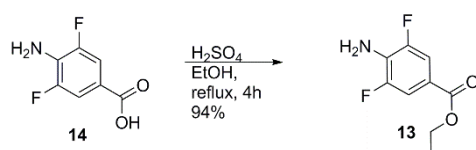
4-amino-3,5-difluorobenzoic acid (**14**)



Scheme S6. Synthesis of oF₂Azo derivative **14**.

Following the procedure of Bléger et al.^[6], the 4-amino-3,5-difluorobenzonitrile (**15**, 1.99 g, 12.9 mmol, 1.00 eq.) was suspended in 1 M NaOH (80.0 mL) and stirred at 110 °C for 4 h. The reaction mixture was cooled to r.t. and afterwards acidified with 1 M HCl (10.0 mL). The resulting precipitate was filtered, washed with distilled water and dried to yield the desired product **14** (2.23 g, 12.9 mmol, 99%) as a white solid. The characterization is in agreement with the literature.^[6] **¹H NMR (300 MHz, DMSO-d₆, δ)**: 12.68 (s, 1H, COOH), 7.40 (dd, ³J_{HH} 7.2, ⁴J_{HH} 2.5, 2H, 2 × CH_{ar}), 6.05 (s, 2H, NH₂). **¹³C NMR (75 MHz, DMSO-d₆, δ)**: 166.0 (COOH), 151.3 (FC_{ar}), 148.2 (FC_{ar}), 130.6 (C_{ar}), 115.6 (C_{ar}), 112.3 (C_{ar}H), 112.1 (C_{ar}H). **HRMS-ESI- (m/z)**: calcd. for [M-H]⁻ C₇H₄F₂N₁O₂, 172.0216; found, 172.0217.

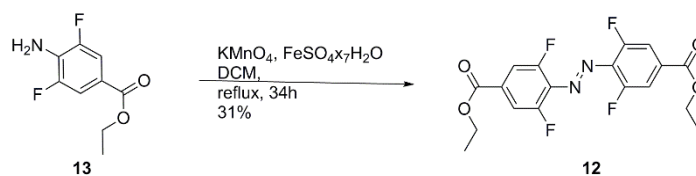
ethyl 4-amino-3,5-difluorobenzoate (**13**)



Scheme S7. Synthesis of oF₂Azo derivative **13**.

Following the procedure of Bléger et al.,^[6] the 4-amino-3,5-difluorobenzoic acid (**14**, 1.90 g, 11.0 mmol, 1.00 eq.) was dissolved in EtOH (40.0 mL) and conc. H₂SO₄ (6.00 mL) and refluxed for 6 h. The solution was neutralized with saturated NaHCO₃, extracted with CH₂Cl₂, dried over anhydrous MgSO₄, filtered and the solvent was removed under reduced pressure to yield the desired product (**13**, 2.07 g, 10.3 mmol, 94%) as a pale brown solid. The characterization is in agreement with the literature.^[6] **¹H NMR (300 MHz, CDCl₃, δ)**: 7.53 (dd, ³J = 7.2 Hz, ⁴J = 2.1 Hz, 2H, 2 × CH_{ar}), 4.33 (q, ³J = 7.1 Hz, 2H, OCH₂), 3.86 (s, 2H, NH₂), 1.37 (t, ³J = 7.1, 3H, CH₃). **¹³C NMR (75 MHz, CDCl₃, δ)**: 165.3 (CO), 152.4 (FC_{ar}), 149.3 (FC_{ar}), 128.9 (C_{ar}), 118.8 (C_{ar}), 112.8 (C_{ar}H), 112.6 (C_{ar}H), 61.2 (CH₂), 14.5 (CH₃). **HRMS-ESI+ (m/z)**: calcd. for [M+H]⁺ C₉H₉F₂N₁O₂H, 202.0674; found, 202.0675.

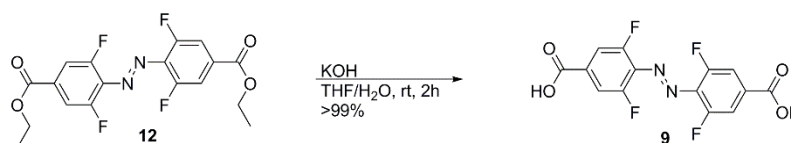
Diethyl 4,4'-(diazene-1,2-diyl)-bis(3,5-difluorobenzoate) (**12**)



Scheme S8. Synthesis of the oF₄Azo derivative **12**.

Following the procedure of Bléger et al.^[6], the ethyl-4-amino-3,5-difluorobenzoate (**13**, 2.43 g, 12.1 mmol, 1.00 eq.) was dissolved in CH_2Cl_2 (300 mL) and KMnO_4 (20.0 g, 127 mmol, 10.5 eq) and $\text{FeSO}_4 \times 7 \text{H}_2\text{O}$ (20.0 g, 71.9 mmol, 5.94 eq.) were added. The reaction mixture was refluxed for 28 h and filtered over celite. The solvent was removed under reduced pressure and the crude was purified by flash column chromatography (CH_2Cl_2 /pentane 1:1) to yield the desired product **12** (748 mg, 1.88 mmol, 31%) as a red solid. The characterization is in agreement with the literature.^[6] **TLC**: $R_f = 0.75$ (CH_2Cl_2); **¹H NMR (300 MHz, CDCl_3 , δ)**: 7.75 (d, $^3J = 8.9$ Hz, 4H, 4 × CH_{ar}), 4.43 (q, $^3J = 7.1$ Hz, 4H, 2 × CH_2), 1.43 (t, $^3J = 7.1$ Hz, 6H, 2 × CH_3). **¹³C NMR (75 MHz, CDCl_3 , δ)**: 163.8 (2 × CO), 156.9 (2 × C_{ar}), 153.4 (2 × C_{ar}), 134.0 (2 × C_{ar}), 133.9 (2 × C_{ar}), 114.2 (2 × C_{ar}), 113.9 (2 × C_{ar}), 62.3 (2 × CH_2), 14.3 (2 × CH_3). **HRMS-ESI+ (m/z)**: calcd. for $[\text{M}+\text{H}]^+$ $\text{C}_{18}\text{H}_{14}\text{F}_4\text{N}_2\text{O}_4\text{H}$, 399.0962; found: 399.0963.

Diethyl 4,4'-(diazene-1,2-diyl)-bis(3,5-difluorobenzoate) (**9**)



Scheme S9. Synthesis of oF₄Azo derivative **9**.

Following the procedure of Bléger et al.^[7], the diethyl 4,4'-(diazene-1,2-diyl)-bis(3,5-difluorobenzoate) (**12**, 457 mg, 1.22 mmol, 1.00 eq.) was dissolved in THF (30.0 mL) and a solution of KOH (242 mg, 4.31 mmol, 3.53 eq.) in distilled water (15.0 mL) was added. The reaction mixture was stirred for 4 h, and then acidified with 1 M HCl (5.00 mL). The formed precipitate was filtered, washed with distilled water and dried on the freeze drier for 10 h to yield the desired product **9** (417 mg, 1.22 mmol, <99%) as a red solid. The characterization is in agreement with the literature.^[7] **¹H NMR (300 MHz, DMSO-d_6 , δ)**: 13.91 (s, 2H, 2 × COOH), 7.82 (d, $^3J = 9.3$ Hz, 4H, 4 × CH_{ar}). **¹³C NMR (75 MHz, DMSO-d_6 , δ)**: 164.5 (2 × CO), 156.0 (2 × C_{ar}), 152.5 (2 × C_{ar}), 135.2 (2 × C_{ar}), 133.0 (2 × C_{ar}), 114.0 (2 × C_{ar}), 113.7 (2 × C_{ar}). **HRMS-ESI- (m/z)**: calcd. for $[\text{M}-\text{H}]^-$ $\text{C}_{14}\text{H}_5\text{F}_4\text{N}_2\text{O}_4$, 341.0191; found, 341.0203.

NMR of 4-((4-((2-((diphenyl(p-tolyl)methyl)amino)ethyl)carbamoyl)-2,6-difluorophenyl) diazenyl) -3,5-difluorobenzoic acid (2**)**

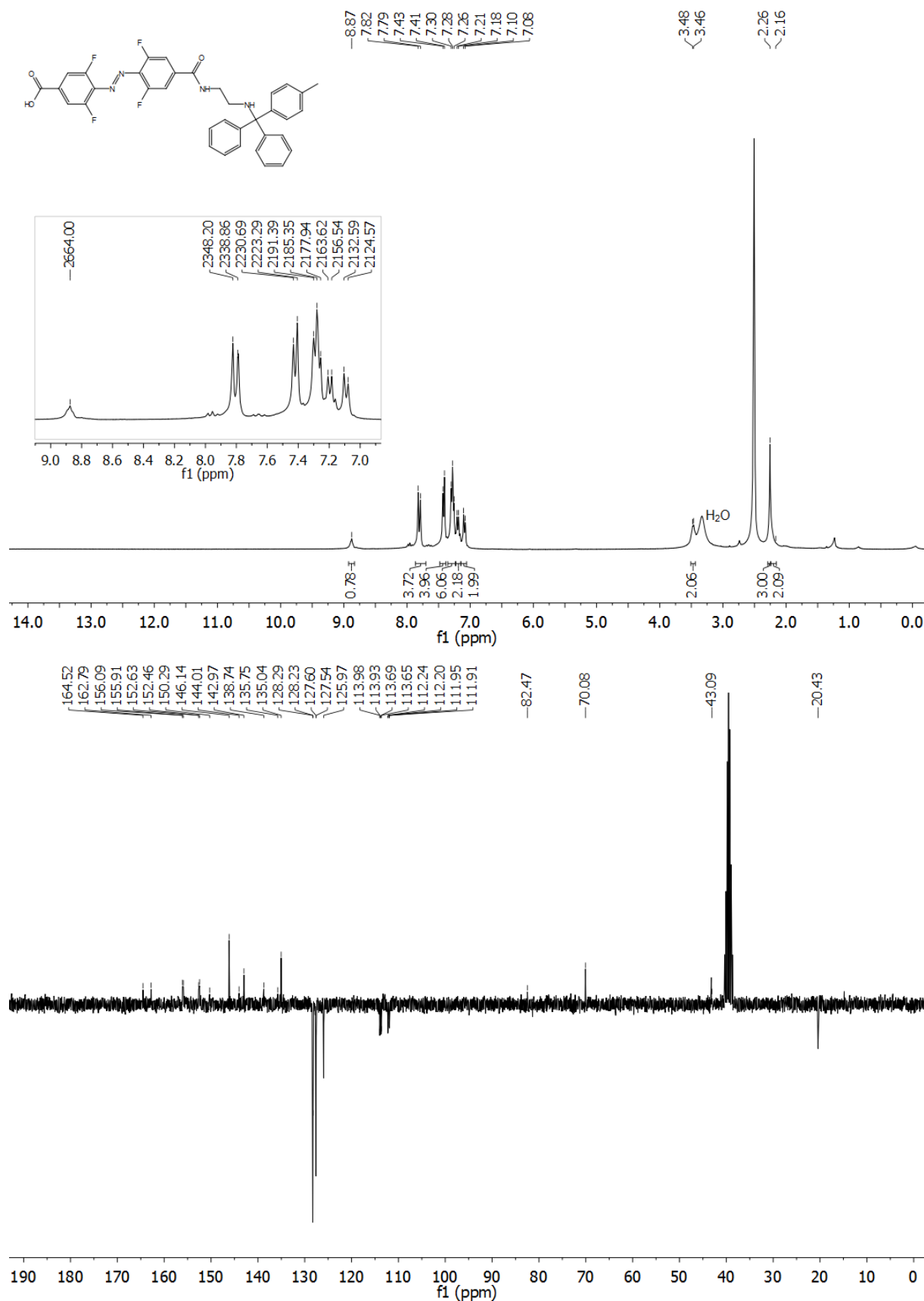
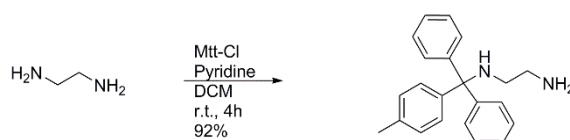


Figure S4. Top: ¹H-NMR spectrum of the compound **2**. Insert: zoom-in of signals in the aromatic region to highlight the shifts in Hz units for calculating *J*-constants. Bottom: ¹³C-APT-NMR spectrum of compound **2**.

N1-(diphenyl(p-tolyl)methyl)ethane-1,2-diamine (10)



Following the procedure of Chouiki et al.^[8] ethylene diamine (6.20 mL, 92.2 mmol, 9.00 eq) was dissolved in DCM (20.0 mL) and pyridine (14.4 mL) was added portionwise. Methyltrityl chloride (3.00 g, 10.25 mmol, 1.00 eq) was added at 0°C. The solution was stirred 4 h at r.t., before it was quenched by adding MeOH (5.50 mL) and the solvent was evaporated. The crude was dissolved in DCM (1% NEt₃)/ distilled water (1:1; 100 mL) and the product was extracted with DCM (2 x 50.0 mL). The combined organic layers were washed with brine, dried over anhydrous MgSO₄ and the solvent was removed under reduced pressure. The crude was purified by flash column chromatography (EtOAc + 3% NEt₃). The product **10** was obtained as a colourless oil (2.99 g, 9.45 mmol, 92%). The characterization is in agreement with the literature.^[8] **TLC:** R_f = 0.2 (EtOAc + 3% NEt₃) **¹H NMR (300 MHz, CDCl₃, δ):** 7.48 (dd, ³J = 5.3 Hz, 4H, 4 × CH_{ar}), 7.35 (d, ³J = 8.2 Hz, 2H, 2 × CH_{ar}), 7.30 – 7.22 (m, 4H, 4 × CH_{ar}), 7.21 – 7.12 (m, 2H, 2 × CH_{ar}), 7.07 (d, ³J = 8.1 Hz, 2H, 2 × CH_{ar}), 2.79 (t, ³J = 5.9 Hz, 2H, CH₂), 2.30 (s, 3H, CH₃), 2.20 (t, ³J = 6.0 Hz, 2H, CH₂), 1.42 (s, 2H, NH₂). **¹³C NMR (75 MHz, CDCl₃, δ):** 146.5 (2 × C_{ar}), 143.4 (C_{ar}), 135.87 (C_{ar}), 128.8 (4 × CH_{ar}), 128.7 (2 × CH_{ar}), 128.6 (4 × CH_{ar}), 127.9 (2 × CH_{ar}), 126.3 (2 × CH_{ar}), 70.6 (CNH), 46.7 (CH₂), 43.0 (CH₂), 21.1 (CH₃). **HRMS-ESI⁺ (m/z):** calcd. for [M + H]⁺ C₂₂H₂₄N₂H: 317.2021; found: 317.2011.

Solid Phase Peptide Synthesis

Peptides were synthesized manually in 2 mL polypropylen reactors with plunger and frit, (pore size 25 μm, Multi Syn Tech GmbH, Germany; for bigger scales, 15 mL reactors were used). TentaGel S RAM resin (0.25 mmol/g), Cl-Trt resin (1.0-1.8 mmol/g) or Sieber-amide resin (0.3-0.6 mmol/g), where the loading of the first amino acid for both Sieber-amide and Cl-Trt resin was determined, were used for synthesis.

Determination of the resin loading: The Fmoc group of a weighted amount of resin (~5.00 mg), previously loaded with the first amino acid was deprotected by adding twice a solution of 20% piperidine in DMF for 5 min (500 μL). The filtrates were collected and the absorbance was measured at 300 nm (ϵ_{300} (Fmoc) = 7800 M⁻¹ cm⁻¹)^[9] on a TECAN 20M Spark plate reader and the loading was calculated. Loading of the resins was done as follows:

Sieber-amide resin: 2.00 eq. Fmoc-amino acid, 2.00 eq. OxymaPure (0.5 M in DMF), 2.00 eq. DIC were added to the resin for 45 min with a final concentration of 0.4 M of the amino acid. After washing, the non-reacted amino groups were capped following the general protocol of manual synthesis explained below

Cl-Trt resin: 2.00 eq. Fmoc-amino acid, 8.00 eq. DIPEA in DMF were added to the resin for 45 min with a final concentration of 0.4 M of the amino acid. The Cl-Trt resin was capped by washing with 3 × DMF, 3 × DCM, 3 × DMF, 3 × MeOH, 3 × DCM.

Manual Solid Phase Protocol for Peptide Synthesis

The amounts of reagents of the following synthesis protocol correspond to 20-40 μmol scale. For higher scales, the reagents were scaled up correspondingly. For shaking an Edmund Bühler Swip shaker was used.

Swelling: The appropriate amount of resin was swollen in 1.0 mL DMF for 30 min.

Deprotection of the temporal Fmoc group: Piperidine (500 μL , 20% in DMF) was added to the resin and was shaken for 5 min. This step was repeated and the resin was filtered off and washed with DMF (5 \times 1.50 mL), CH_2Cl_2 (5 \times 1.50 mL), DMF (5 \times 1.50 mL).

Coupling of amino acids: The Fmoc-amino acid (4.00 eq.) was dissolved in OxymaPure (0.5 M in DMF, except for the coupling of the the azobenzene derivatives, where OxymaPure was dissolved in NMP, 0.5 M, 4.00 eq.), and DIC (4.00 eq.) was added. The resulting solution was activated for 3 min and subsequently added to the resin. This suspension was shaken for 45 min. The resin was filtered off and washed with DMF (5 \times 1.50 mL) and CH_2Cl_2 (5 \times 1.50 mL) and DMF (5 \times 1.50 mL). Monitoring of the coupling completion was confirmed by the TNBS-test^[10] on few resin-beads.

Capping: A solution of lutidine/ Ac_2O /DMF 6:5:89 (1.0 mL) was added to the resin and was shaken for 5 min. The resin was filtered off and washed with DMF (5 \times 1.50 mL), CH_2Cl_2 (5 \times 1.50 mL) and DMF (5 \times 1.50 mL).

Cleavage deprotection step: Depending on the specific peptide the cleavage cocktail differed in the concentration of TFA or HFIP, type of scavengers and reaction times. The used cleavage cocktails and reaction times are listed below.

- Cleavage cocktail A: 95% TFA, 5% H_2O (2 h)
- Cleavage cocktail B: HFIP/ CH_2Cl_2 1:4 (3 h)
- Cleavage cocktail C: 1% TFA in CH_2Cl_2 (2 \times 30 min)
- Cleavage cocktail D: 90% TFA, 10% H_2O (90 min)

The dried resin was treated with the corresponding mixture and was shaken for the specific time. After the defined time, the resin was filtered off and the filtrate was added to dry ice-cold Et_2O (1 mL of Et_2O for 100 μL cleavage cocktail). After 10 min, the precipitated peptide was centrifuged, the supernatant was discarded and the peptide pellet was dissolved in MilliQ in order to be purified.

Purification: The probes were purified as specified above. The collected fractions, containing the desired peptides, were lyophilized and stored at $-20\text{ }^\circ\text{C}$.

Characterization: The freeze-dried products were identified via analytical HPLC-MS on an Agilent 1200 Series HPLC-system (Agilent Technologies) as detailed before. Electrospray Ionization Mass Spectrometry (ESI-MS) was performed as detailed before.

Synthesis and Characterization Data of the Peptides

Unless otherwise noted, the synthesis of the peptides was performed manually following the protocol described above. The purity of the peptides was calculated from the integrated peak areas of the HPLC chromatograms, and is given as % for each peptide. In the chromatograms, the different isomers of azobenzene-containing peptides are labelled with their respective retention times above the respective peak and assigned in the text into brackets.

4: H₂N-Ser-Ala-Arg-Ala-**cAzo**-Val-His-Leu-Arg-Lys-Ser-CONH₂; For the final cleavage step, the resin (scale = 20 μmol, loading = 0.25 mmol/g) was treated with 2.00 mL cleavage cocktail A. After purification, the product 5 × TFA salt (3.8 mg, 1.83 μmol, 9%) was obtained as a yellow solid. t_R = 23.36 min (*cis*). Purity ≥ 99%. Formula: C₆₃H₁₀₀N₂₄O₁₃. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 701.4026; found: 701.4028.

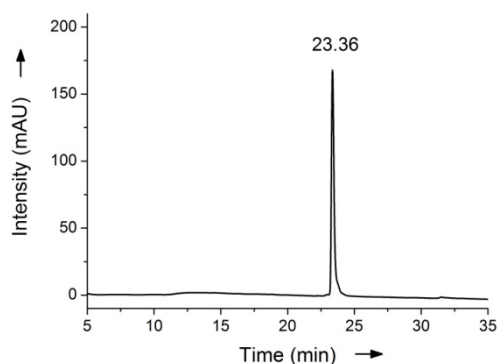


Figure S5. HPLC chromatogram of purified peptide **4**. Gradient: 5-40% B on column 1.

16: Boc-Ser(tBu)-Ala-Arg(Boc)₂-Ala-COOH; For the final cleavage step, the resin (scale = 0.35 mmol, loading = 0.832 mmol/g) was treated with 14.00 mL cleavage cocktail B. After purification, the product (146.5 mg, 0.193 mmol, 55%) was obtained as a white solid. t_R = 24.09 min. Purity = 99%. Formula: C₃₄H₆₁N₇O₁₂. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 760.4451; found: 760.4445.

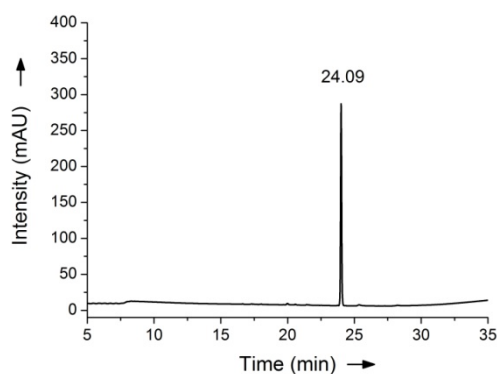


Figure S6. HPLC chromatogram of purified peptide **16**. Gradient: 5-95% B on column 1.

17: H₂N-**oF₄Azo**-Val-His(Trt)-Leu-Arg(Boc)₂-Lys(Boc)-Ser(tBu)-CONH₂; For the final cleavage step, the resin (scale = 80 μmol, loading = 0.513 mmol/g) was treated with 6.00 mL cleavage cocktail C, 2 × for 30 min. After purification, the product 1 × TFA-salt (33.9 mg, 18.7 μmol, 23%) was obtained as an orange solid. *t_R* = 14.95 min (*cis*, product - 1Boc), 15.70 min (*trans*, product - 1Boc), 19.55 min (*cis*), 20.39 min (*trans*). Purity = 91%. Formula: C₈₆H₁₁₆F₄N₁₇O₁₅. HRMS-ESI⁺ (m/z): [M+H]⁺ calcd.: 1703.8798; found: 1703.8778.

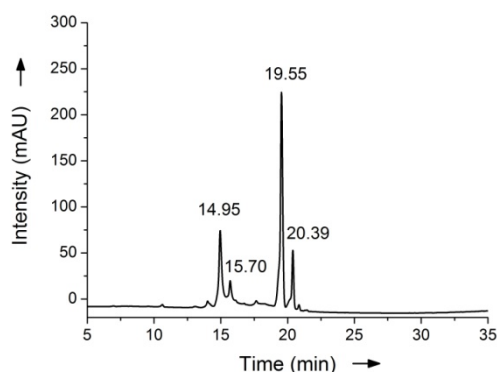
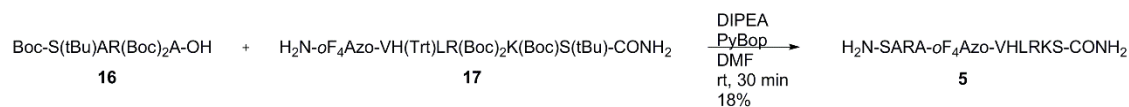


Figure S7. HPLC chromatogram of purified peptide **17**. Gradient: 35-95% B on column 2.

Ligation of fragments **16** and **17**



Scheme S10. Ligation of fragments **16** and **17** to form final **oF₄Azo-peptide 5**.

5: H₂N-Ser-Ala-Arg-Ala-**oF₄Azo**-Val-His-Leu-Arg-Lys-Ser-CONH₂; Ligation of the two fragments **16** and **17** was performed in solution. Therefore, **17** (33.9 mg, 0.0187 mmol, 1.00 eq.) was dissolved in 187 μL DMF and DIPEA (3.30 μL, 0.0187 mmol, 1.00 eq) was added. Peptide **17** (42.6 mg, 0.056 mmol, 3.00 eq.) was separately, dissolved in 187 μL DMF, where PyBop (29.1 mg, 0.056 mmol, 3.00 eq.) and DIPEA (26.1 μL, 0.15 mmol, 8.00 eq) were added and the mixture was activated for 3 min before it was added to the solution of **17**. The reaction mixture was shaken for 30 min. After purification, the side chain protecting groups were deprotected by treatment of the crude with 20 mL cleavage cocktail D. The product was precipitated in ice cold Et₂O, dissolved in water and irradiated at 520 nm as explained before to obtain the *cis* isomer, which was isolated via preparative HPLC. The product 5 × TFA-salt (6.9 mg, 3.35 μmol, 18%) was obtained as an orange solid. *t_R* = 21.74 min (*cis*), 22.25 min (*trans*). Purity = 99%. Formula: C₆₃H₉₈F₄N₂₄O₁₄. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 745.3810; found: 745.3804.

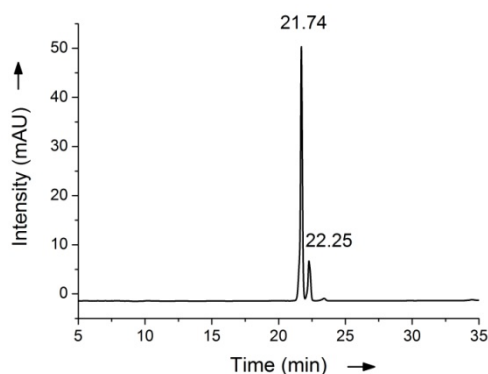


Figure S8. HPLC chromatogram of purified peptide **5**. Gradient: 5-40% B on column 2.

Results and Discussion

Screening for Suitable Fmoc-deprotection Conditions

To test whether the oF_4Azo motif could be introduced into the peptide backbone as an Fmoc-amino acid building block, first the stability of the compound **9** against 5% piperazine in DMF and 2% DBU in DMF was evaluated since these conditions could be used as mild protocol for Fmoc-deprotection in SPPS.

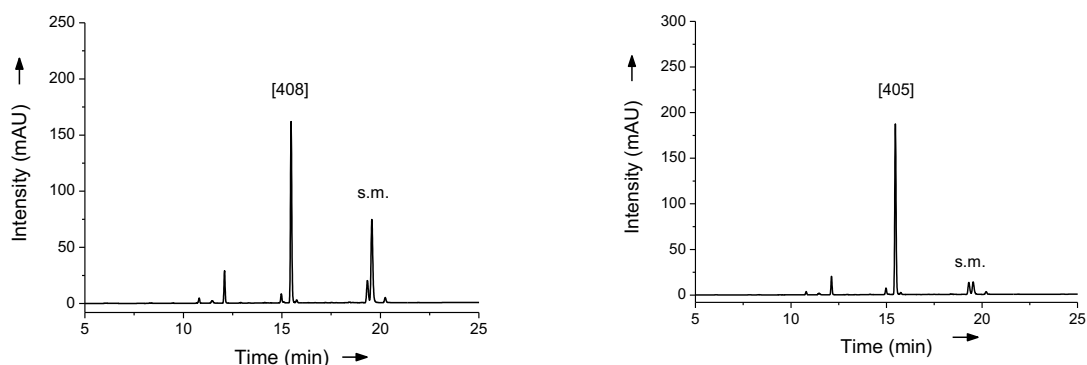
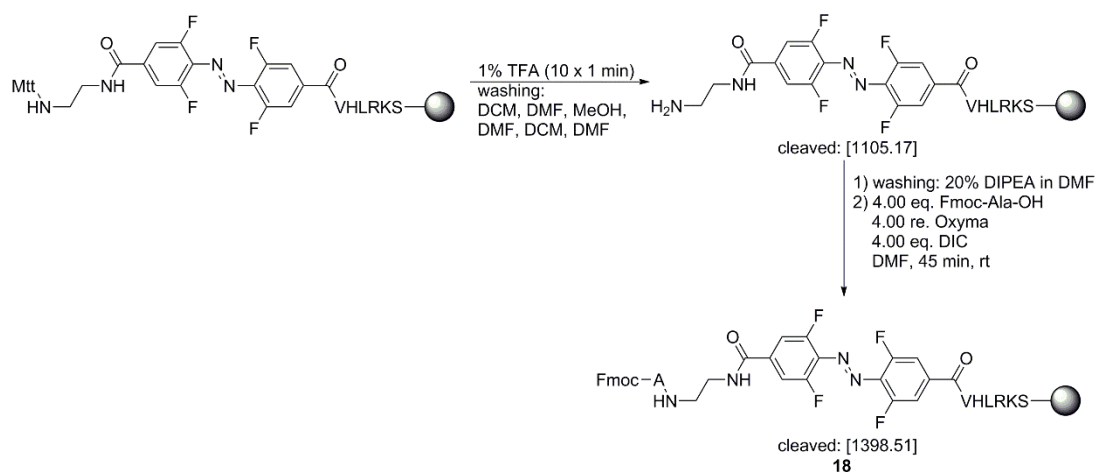


Figure S9. HPLC chromatograms of oF_4Azo compound **5**. Gradient 5-95% B on column 2. Left: 5% piperazine in DMF, 10 min. Right: 2% DBU in DMF, 10 min. Mass of degradation products indicated on top of the peak in square brackets and g/mol.

As obvious from the HPLC chromatograms these conditions are not compatible with the stability of **9**, since they lead to adduct formation as side products in the case of the piperazine and substitution side products in the case of DBU.

To further evaluate this, we incorporated the **Mtt-oF₄AzoAA 2** to the peptide backbone of the VHLRKS-peptide fragment via the general SPPS method detailed before. As shown in Scheme S11, the Mtt-protecting group was removed on the resin via washing the resin with 1% TFA in DCM (10 × 1 min) followed by the washing with DCM, DMF, MeOH, DMF, DCM and DMF (each 3 × 1 min). Then, the resin was again washed with 20% DIPEA in DMF (5 × 3 min), before Fmoc-Ala-OH was coupled according to the usual SPPS protocol.



Scheme S11. On resin Mtt-deprotection and coupling of Fmoc-Ala-OH to the oF₄Azo-peptide fragment.

The HPLC chromatogram and the ESI spectrum of the cleaved compound **18** are shown in Figure S10 (19.12 and 19.89 min: *cis* and *trans* isomer of **18**, respectively).

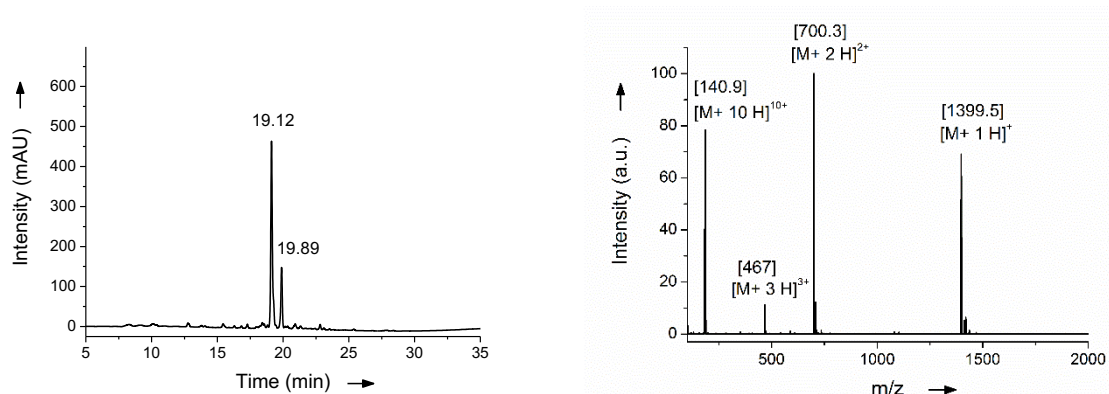
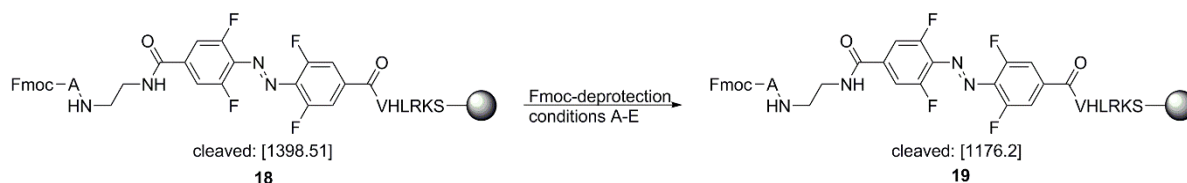


Figure S10. Left: HPLC chromatogram of Fmoc-protected product **18** with a mass of 1398.51 g/mol. Gradient 5-95% B, column 2. Right: ESI spectrum of compound **18**.

The resin was then incubated with the respective Fmoc-deprotection conditions (see Scheme S12 and Table 1, Manuscript) and then washed and dried as in the general SPPS protocol. Next, a test cleavage with the cocktail D (90% TFA, 10% H₂O, 90 min) was performed and HPLC chromatograms were recorded (5-95% B, column 2), which are shown in Figure S11 till Figure S13.



Scheme S12. Reaction for the evaluation of different Fmoc-deprotection conditions on the peptide fragment **18**.

The peaks at 19.12 and 19.89 min correspond to the Fmoc-protected peptide **18** with a mass of 1398.51 g/mol. The peak at 14.25 min corresponds to the desired Fmoc-deprotected product **19** with a mass of 1176.2 g/mol. The masses of the additional new peaks are indicated in square brackets in g/mol above the respective peak.

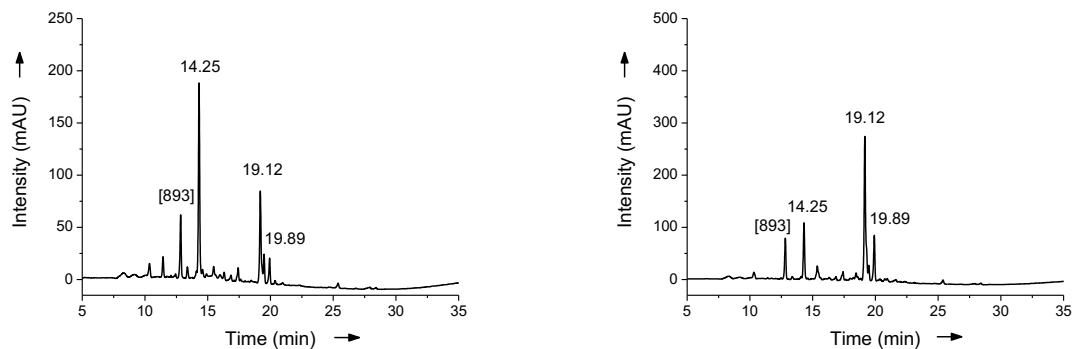


Figure S11. HPLC chromatograms after Fmoc-deprotection and resin cleavage with cocktail D. Gradients 5-95% B, column 2. Left: Conditions A incubated for 17 h. Right: Conditions B incubated for 17 h.

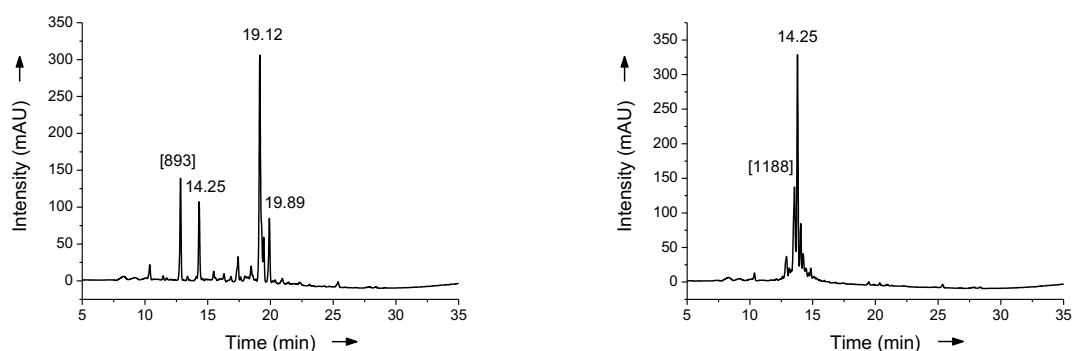


Figure S12. HPLC chromatograms after Fmoc-deprotection and resin cleavage with cocktail D. Gradient 5-95% B, column 2. Left: Conditions C incubated for 17 h. Right: Conditions D incubated for 90 min.

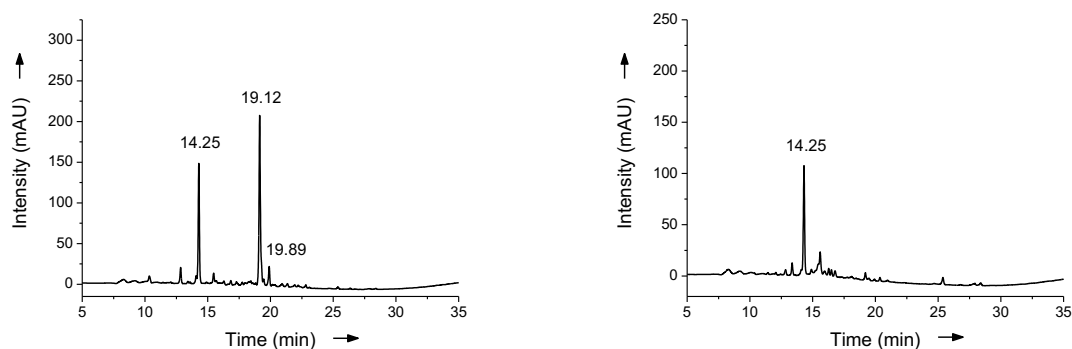


Figure S13. HPLC chromatograms after Fmoc-deprotection and resin cleavage with cocktail D. Gradient 5-95% B, column 2. Left: Conditions E incubated for 3 h. Right: Conditions F incubated for 16 h.

As can be seen, almost none of the above detailed conditions could prove to be efficient for Fmoc deprotection without the degradation of the peptide (A, B, C, D). The only exception is

condition E/F, with which it was possible to deprotect the Fmoc group without a notably high degradation of the peptide.

Extinction Coefficients

The extinction coefficient of cAzo-containing peptide **4**, was measured in MilliQ water at r.t. A specific amount of the compound **4** (> 5.00 mg) was accurately weighted on a Mettler Toledo XP6 micro balance to prepare a solution of known concentration. Increasing amounts of the corresponding stock solution were added without exceeding 10 % of the initial volume of the cuvette; absorbance was measured as explained before. In all the cases the recorded absorbance was between 0.1 and 1 a.u. to be in concordance with Lambert-Beer law:

$$A = \varepsilon \cdot c \cdot l$$

where A = absorbance; ε = molar extinction coefficient; c = concentration; l = path length.

The lineal regression of the obtained absorbance measurements versus the concentrations of the corresponding compound allowed us to obtain the molar extinction coefficient and regression value (R^2). The background signal was subtracted.

The obtained molar extinction coefficient in MilliQ water is: $\varepsilon_{494} = 25990 \text{ L mol}^{-1} \text{ cm}^{-1}$ with $R^2 = 0.995$

In the case of the other peptides the following extinction coefficients were used: oF₄Azo containing peptides (for *trans* isomer) $\varepsilon_{319} = 25000 \text{ L mol}^{-1} \text{ cm}^{-1}$ in MeCN,^[1] cAzo containing peptides (for *cis* isomer) $\varepsilon_{238} = 25990 \text{ L mol}^{-1} \text{ cm}^{-1}$ in MilliQ water, fluorescently tagged peptides $\varepsilon_{494} = 76900 \text{ L mol}^{-1} \text{ cm}^{-1}$ ^[2] in 0.1 M phosphate buffer, pH = 9).

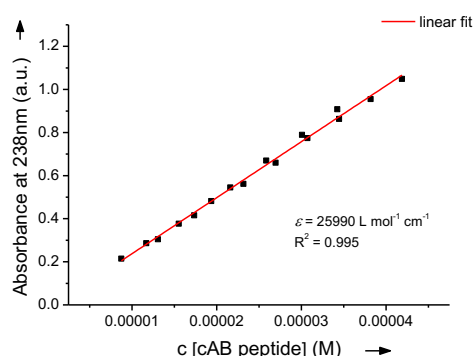


Figure S14. Linear fit of the triplicate measurements to obtain the extinction coefficient of cAzo-containing peptide **4**.

UV-vis Characterization

For the characterization via UV-vis measurements of the oF₄Azo-containing peptide **5** and cAzo containing peptide **4**, a Tecan (Switzerland) Spark 20M multimode microplate reader was used. 200 μL of a 0.262 mM solution of oF₄Azo-containing peptide **5** in MilliQ water was added to a blank of 300 μL MilliQ water in a cuvette (final conc.: 0.105 mM). The cuvette was irradiated at the respective wavelengths for the indicated times, followed by measuring the absorbance spectra at r.t. To evaluate the required time of irradiation to reach the photostationary state

(PSS) of one isomer, the two isomers were irradiated for different times. The absorbance spectra are shown in Figure S15.

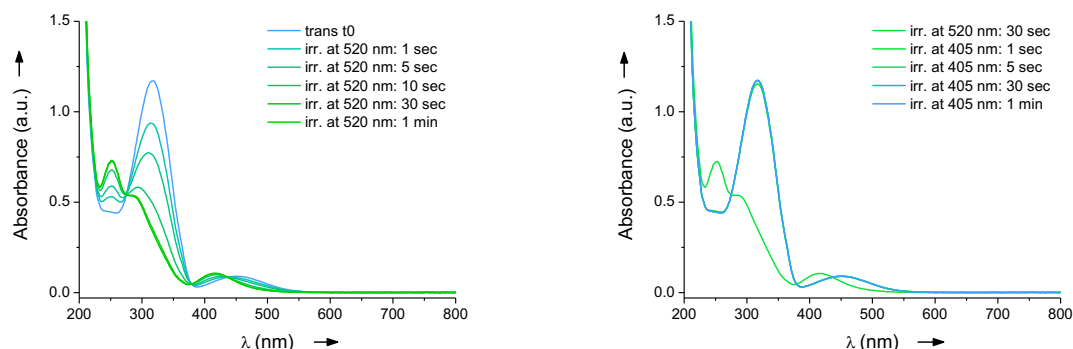


Figure S15. UV-vis spectra of *oF*₄Azo-containing peptide **5**. Left: irradiation of *trans* isomer at 520 nm to obtain the *cis* isomer. Right: irradiation of *cis* isomer at 405 nm to obtain the *trans* isomer.

For the characterization via UV-vis measurements of the *cAzo*-containing peptide **4**, 100 μ L of a 0.51 mM solution of *cAzo*-containing peptide **4** in MilliQ water was added to a blank of 300 μ L water in a cuvette (final conc.: 0.128 mM). The procedure was the same as detailed above. The absorbance spectra are shown in Figure S16.

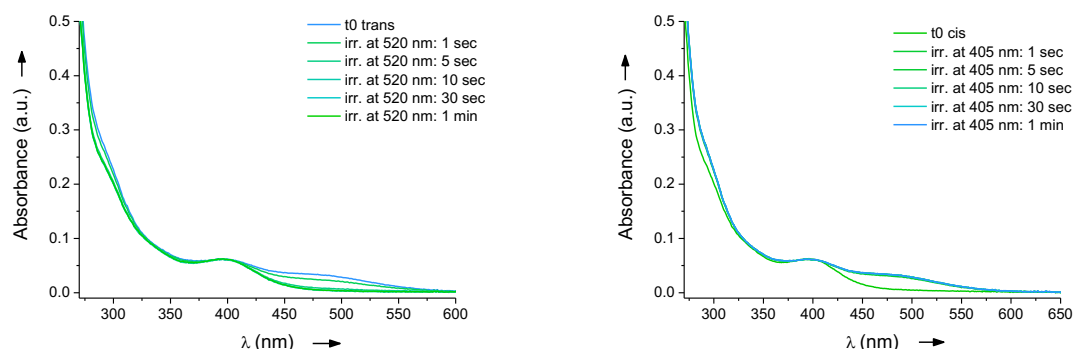


Figure S16. UV-vis spectra of *cAzo*-containing peptide **4**. Left: irradiation of *trans* isomer at 520 nm to obtain the *cis* isomer. Right: irradiation of *cis* isomer at 405 nm to obtain the *trans* isomer.

RP-HPLC Characterization

In the case of the *oF*₄Azo-containing peptide **5** the *trans* isomer is the thermodynamically more stable one. Contrarily for the *cAzo*-containing peptide **4**, the *cis* isomer is thermodynamically favoured due to the ring strain of the *trans* isomer. To evaluate the ratio of each isomer at their photostationary states, as well as examining the stability of each isomer in total darkness, we used RP-HPLC, if not stated differently, the column 3 and a gradient of 10-35% B for *oF*₄Azo-containing peptide **5** and 5-40% B for *cAzo*-containing peptide **4** were used.

For the isomerization, a 120 μ M solution of the corresponding peptide in MilliQ water was irradiated at 405 nm for 5 sec to obtain the *trans* isomer, and at 520 nm for 1 min to obtain the *cis* isomer for the *oF*₄Azo-containing peptide **5**. For the *cAzo*-containing peptide **4**, at 405 nm for 1 min to obtain the *trans* isomer, and at 520 nm for 10 sec. to obtain the *cis* isomer. The

photostationary states reached upon irradiation were determined by integrating the peak areas in the HPLC chromatograms at 275 nm (oF₄Azo- containing peptide **5**, isosbestic point) and at 395 nm (cAzo-containing peptide **4**); due to the low ratio of the *trans* isomer in its photostationary state, the isosbestic point is not clearly visible in these spectra. Because of this, the *cis/trans* ratio of peptide **4** was determined via HPLC at 395 nm, which is the wavelength where both isomers show the same height of absorbance and have, therefore, the same extinction coefficient). To evaluate the relaxation times, the peptide solutions were stored in total darkness at r.t. and HPLC chromatograms were recorded after the indicated time periods, shown in Figure S17 for the oF₄Azo-containing peptide **5** and in Figure S18 for the cAzo-containing peptide **4**.

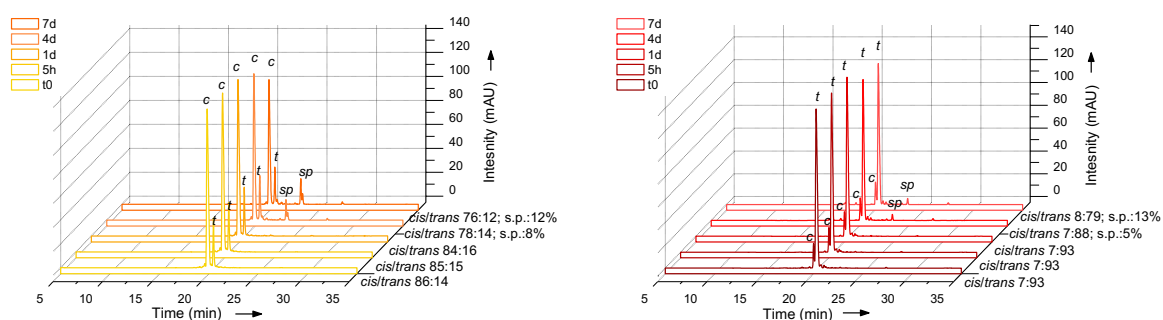


Figure S17. HPLC chromatograms of oF₄Azo-containing peptide **5** incubated in total darkness. HPLC chromatograms were recorded after the indicated time. c: *cis*, t: *trans*, sp: side product. Left: after irradiation at 520 nm to reach the photostationary state of the *cis* isomer. Right: after irradiation at 405 nm to reach the photostationary state of the *trans* isomer. s.p.: indication of ratio of side products.

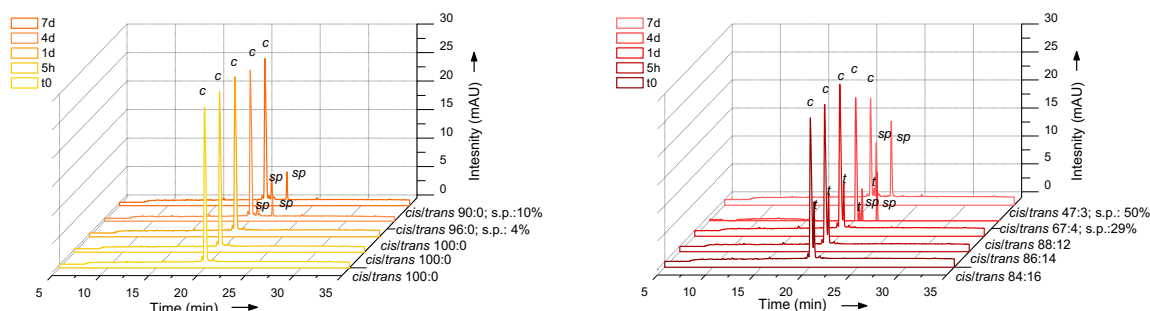


Figure S18. HPLC chromatograms of cAzo-containing peptide **4** incubated in total darkness. HPLC chromatograms were recorded after the indicated time. c: *cis*, t: *trans*, sp: side product. Left: no irradiation; photostationary state of the *cis* isomer. Right: after irradiation at 405 nm to reach the photostationary state of the *trans* isomer. s.p.: indication of ratio of side products.

NMR Characterization

To confirm the ratio between isomers of the oF₄Azo-containing peptide **5**, obtained by HPLC characterization, we used ¹H-NMR. The NMR spectra of peptide **5** was recorded in D₂O at 300 K a Bruker AV III HD 600 MHz at a frequency of 600 MHz (¹H), as detailed above. The peptide was dissolved in D₂O to give a solution of 1.34 mM concentration. Then, the solution was irradiated at the corresponding wavelength as detailed above and ¹H-NMRs of each isomer were recorded. The aromatic region of the spectra are shown in Figure S19. The *cis/trans* ratios were calculated by integrating the peak areas of these aromatic proton signals, (**5**: irr. at 405 nm: *cis/trans*: 8:92; at 520 nm: 84:16 by NMR) and they corroborate the HPLC

ratios. Of note, under our conditions neither the amide protons^[11–13] nor the aromatic histidine protons^[14,15] were visible.

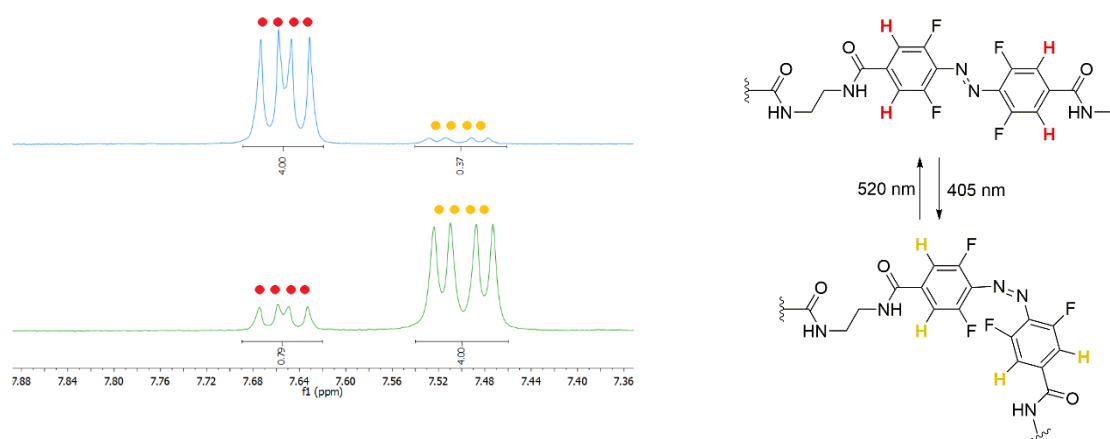


Figure S19. Selected aromatic region of the ^1H -NMR spectra (600 MHz) of a 1.34 mM solution of **oF₄Azo-peptide (5)** in D_2O after irradiation to the *trans*- (top, blue) and *cis* (bottom, green) isomer. Aromatic protons are marked in colours on the spectra and molecules.

Stability against Glutathione

The stability of the peptides **5** and **4** against glutathione (GSH) was evaluated by incubating 3.5 μL of a 250 mM stock of GSH with 80 μL of a 100 μM solution of the respective peptide in MilliQ water giving final concentrations of 10 mM of GSH and 96 μM of the peptides, respectively at r.t. After different time periods, HPLC chromatograms of these mixtures were recorded and are shown in Figure S20 to Figure S21 (column 3, gradient of 10-35% B for **oF₄Azo**-containing peptide **5** and 5-40% B for **cAzo**-containing peptide **4**).

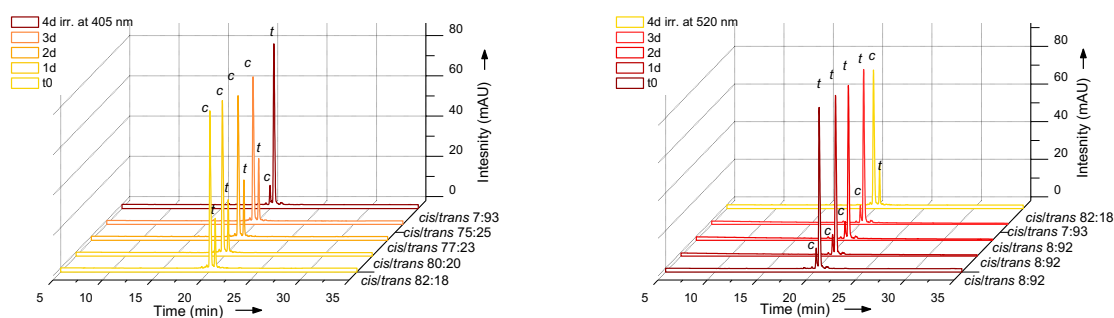


Figure S20. HPLC chromatograms of **oF₄Azo-peptide (5)** incubated with 10 mM reduced glutathione at 275 nm. The solutions were stored in darkness and HPLC chromatograms were recorded after the indicated time. c: *cis*, t: *trans*, sp: side product. Left: after irradiation at 520 nm to reach the photostationary state of the *cis* isomer. After 4 days incubation, the solution was irradiated at 405 nm to switch back to the photostationary state of the *trans* isomer (red). Right: after irradiation at 405 nm to reach the photostationary state of the *trans* isomer. After 4 days incubation, the solution was irradiated at 520 nm to switch back to the photostationary state of the *cis* isomer (yellow).

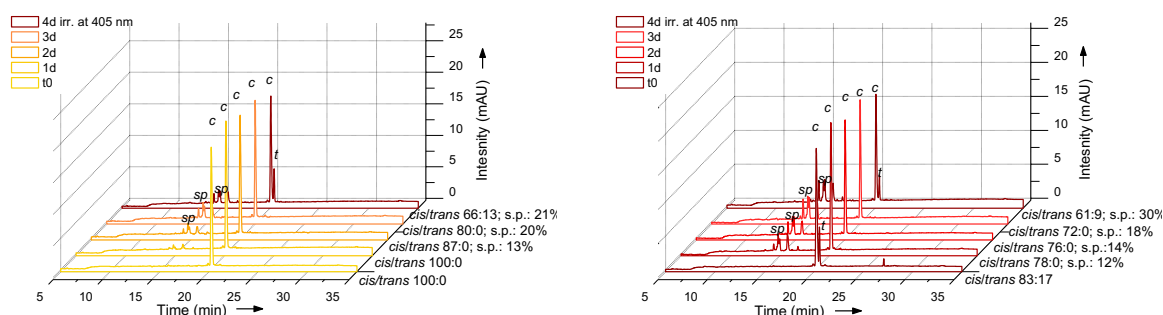


Figure S21. HPLC chromatograms of **cAzo-peptide (4)** incubated with 10 mM reduced glutathione. The solutions were stored in darkness and HPLC chromatograms were recorded after the indicated time. c: *cis*, t: *trans*, sp: side product. Left: after irradiation at 520 nm to reach the photostationary state of the *cis* isomer. After 4 days incubation, the solution was irradiated at 405 nm to switch back to the photostationary state of the *trans* isomer (red). Right: after irradiation at 405 nm to reach the photostationary state of the *trans* isomer. After 4 days incubation, the solution was again irradiated at 405 nm to switch back to the photostationary state of the *trans* isomer (red). s.p.: indication of ratio of side products.

Fluorescence Polarization-based Assay

Competitive Fluorescence Polarization-based Binding Assays

The FP-binding experiments were performed in triplicate and three times independently. The results are represented in Table S1 and the FP spectra are shown in Figure 2 (Manuscript).

Table S1. Mean \pm SD of the IC_{50} and K_i values of peptides **4** and **5** of three independent measurements.

peptide	Exp.	PSS at 405 nm IC_{50} [μ M]	PSS at 520 nm IC_{50} [μ M]	PSS at 405 nm K_i [μ M]	PSS at 520 nm K_i [μ M]	ratio
H ₂ N-SARA oF₄Azo VHLRKS (5)	1	0.757	1.88	10.9	28.7	2.63
	2	0.759	1.91	11.0	29.1	2.65
	3	0.913	2.27	13.4	34.7	2.59
	mean	0.810 \pm 0.089	2.02 \pm 0.21	0.0118 \pm 0.0014	0.0308 \pm 0.0033	2.61
H ₂ N-SARA cAzo VHLRKS (4)	1	7.07	9.58	0.110	0.150	1.36
	2	8.53	16.1	0.133	0.252	1.89
	3	11.4	14.0	0.178	0.220	1.24
	mean	8.98 \pm 2.18	13.2 \pm 3.3	0.140 \pm 0.035	0.207 \pm 0.052	1.48

To verify that the used assay conditions do not influence the *cis/trans* ratio, a control well with 5 μ L of the respective peptide in DMSO irradiated at 525 nm (*cis*), as well at 405 nm (*trans*), each diluted with 120 μ L assay buffer were included in each plate. Before and after the assay, HPLC chromatograms of such controls were recorded at the respective isosbestic points (gradient: 5-40% B, column 2). We did not observe any influence on *cis/trans* ratio due to the assay conditions. The chromatograms are shown in Figure S22 to Figure S25. *Cis* and *trans*

isomers are labelled with their respective retention times (**5**, *cis*: 20.07/19.91 min; **5**, *trans*: 20.40/20.49 min; **4**, *cis*: 20.02 min; **4**, *trans*: 20.42 min).

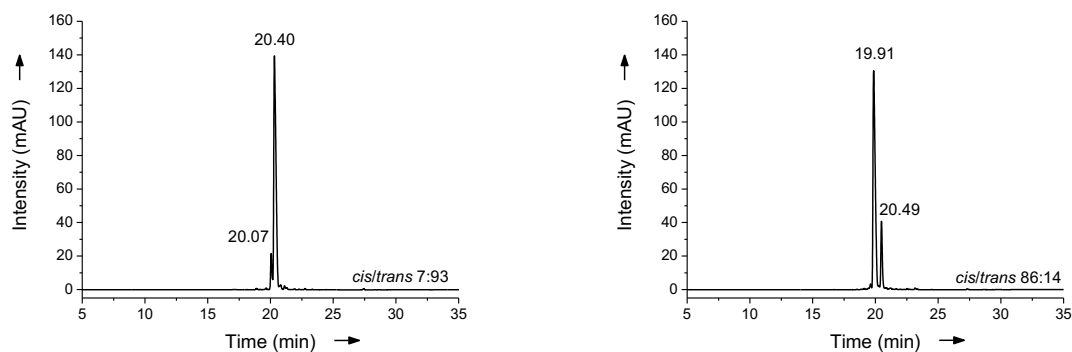


Figure S22. HPLC chromatograms of peptide **5** (SARA-oF₄Azo-VHLRKS) before the FP-based assay. Left: irradiated peptide at 405 nm. Right: irradiated peptide at 520 nm.

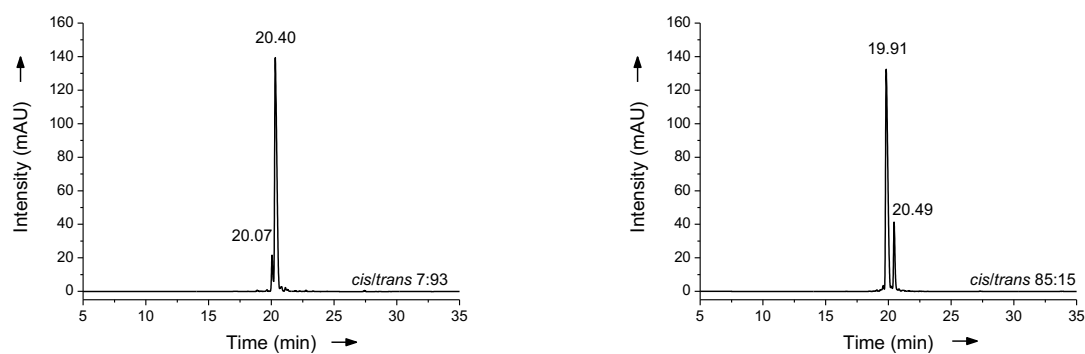


Figure S23. HPLC chromatograms of peptide **5** (SARA-oF₄Azo-VHLRKS) after the FP-based assay. Left: irradiated peptide at 405 nm. Right: irradiated peptide at 520 nm.

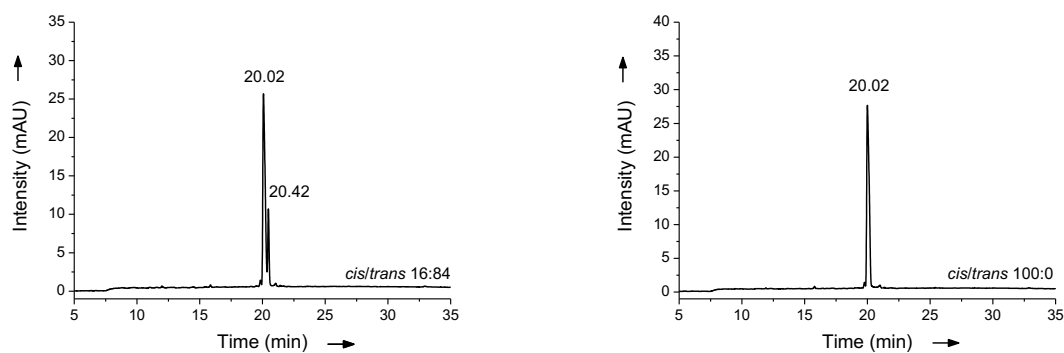


Figure S24. HPLC chromatograms of peptide **4** (SARA-cAzo-VHLRKS) before the FP-based assay. Left: irradiated peptide at 405 nm. Right: irradiated peptide at 520 nm.

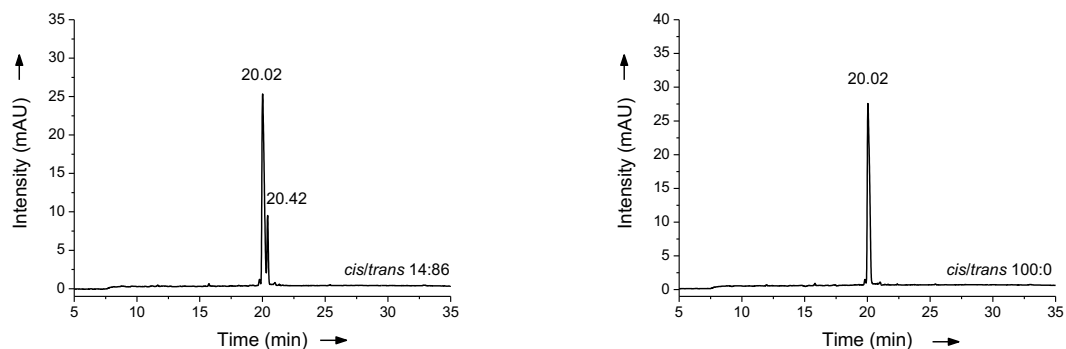


Figure S25. HPLC chromatograms of peptide **4** (SARA-**cAzo**-VHLRKS) after the FP-based assay. Left: irradiated peptide at 405 nm. Right: irradiated peptide at 520 nm.

Crystallisation of WDR5 Δ 23-Peptide **4** Complex

Figure S26 shows an overlay of the whole structures of the superimposed **cAzo-peptide (4)** with **AMPB-peptide (3)**, as well as the ribbon representation of **4** showing its helical conformation. Figure S27 represents the hydrogen bonds of peptide **4** to WDR5.

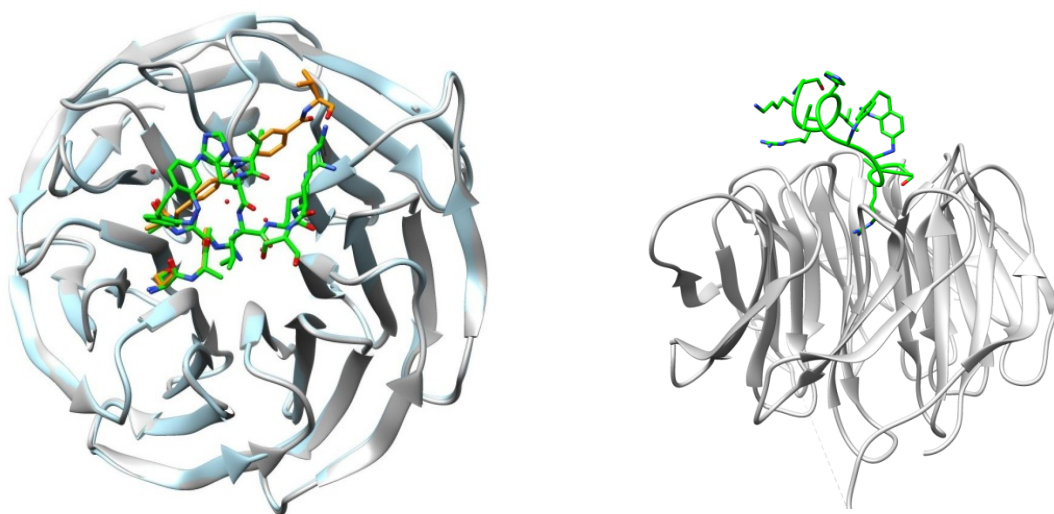


Figure S26. Right: Superposition of the crystal structure of the WDR5 Δ 23/cAzo peptide **4** and WDR5 Δ 23/AMPB-peptide (**3**) (peptide **3**: orange, peptide **4**: green). Left: Ribbon representation of cAzo-peptide (**4**) (green) showing its helical conformation.

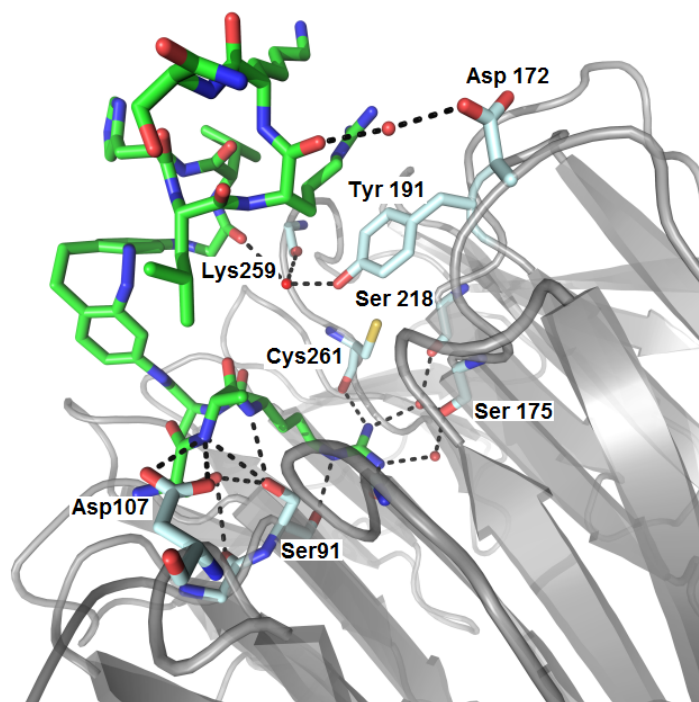


Figure S27. Hydrogen bond network of peptide 4 to WDR5Δ23 (peptide 4: green). Interacting protein residues are labelled.

All the interactions between peptide 4 with WDR5Δ23 are shown in Figure S28 and are additionally listed in Table S2.

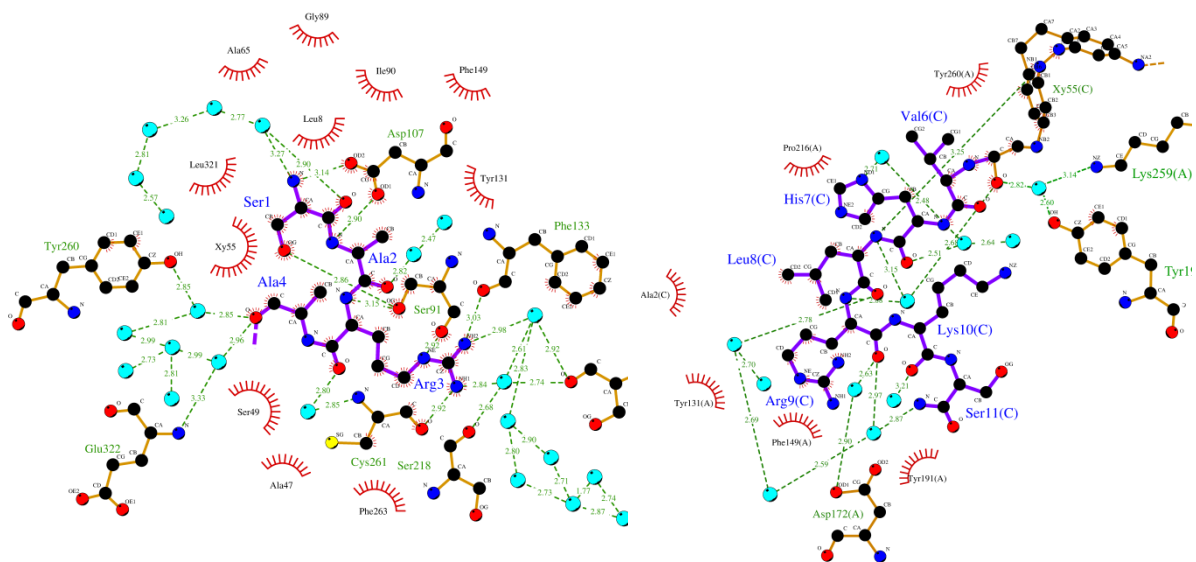


Figure S28. Schematic representation of the key interactions between the cAzo-peptide 4 and WDR5Δ23. Left: SARA part. Right: cAzo-VHLRKS part. Generated with LIGPLOT.^[16]

Table S2. Interactions between the **cAzo-peptide (4)** and WDR5Δ23.

peptide		Hydrogen-bond/ Van der Waals partner		
Residue	Atom	WDR5 residue	Atom	Distance (Å)
Ser1	OG	Ser91	OG	2.86
Ser1	N	Asp107	OD2	3.14
Ala2	N	Asp107	OD1	2.90
Arg3	O	water (bridge to Cys261, N)	O	2.80
Arg3	NH1	Cys261	O	2.92
		water (bridge to Ser175, O)	O	2.84
Arg3	NH2	Phe133	O	3.03
		water (bridge to Ser175, O and via another water to Ser218, O)	O	2.98
Ala4	O	water (bridge to Tye260, O)		2.85
		water (bridge to Glu322, N)		2.96
cAzo	O	water (bridge to Tyr191, OH)		2.82
		water (bridge to lys259, NH2)		
Arg9	O	Water (bridge to Asp172, OD1)	O	2.63

H-bond cut off < 3.5 Å, Van der Waals: 3.6-4.0 Å.

Docking and Molecular Dynamics (MD) Studies

The following Figures show the differences between poses obtained from molecular docking, molecular dynamics (MD) and the crystal structure of **AMPB-peptide (3)**, PDB code: 5M23), as well as RMSD fluctuations during 20 ns of molecular dynamics, and fluctuations of residues from peptidomimetics during 20 ns of molecular dynamics simulations. Furthermore, the measured distances between leucine and tetra-*ortho*-fluorazobenzene, the difference in binding modes obtained for *cis* **AMPB-peptide (3)** and *cis* **oF₄Azo-peptide (5)** and the surface area of protein-peptidomimetics binding. Additionally, Figures that show the different binding modes of the **AMPB-peptide (3)** and the *cis/trans* **oF₄Azo-peptide (5)** obtained by MD, are represented, as well as the difference of computational predicted binding modes between *cis* and *trans* **cAzo-peptide (4)**.

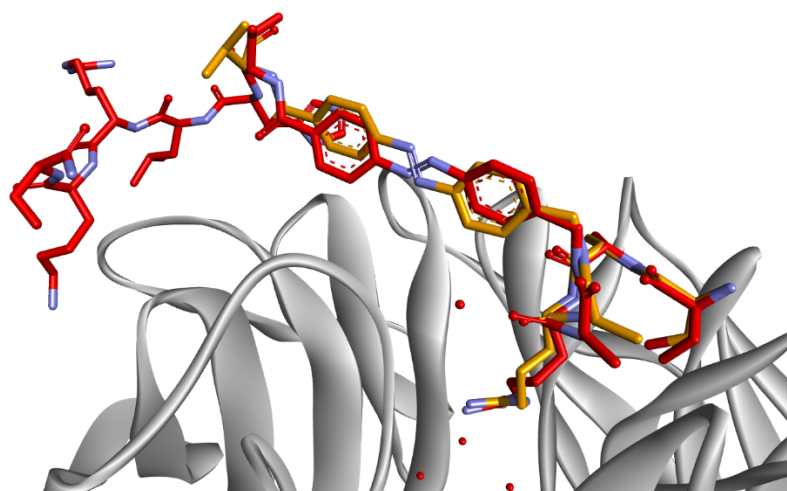


Figure S29. Difference between pose of *trans* AMPB-peptide (3) obtained by molecular docking (red sticks) and pose from PDB ID:5M23 (orange sticks).

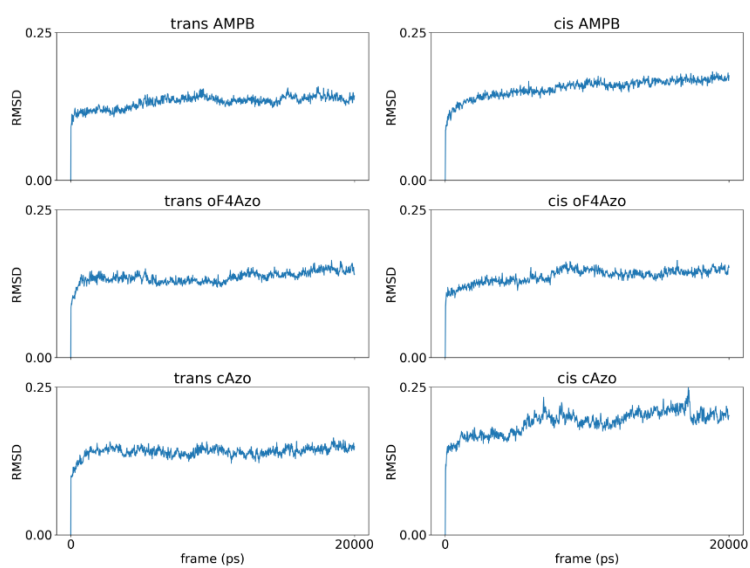


Figure S30. RMSD fluctuations during 20 ns of molecular dynamics production runs calculated for whole protein.

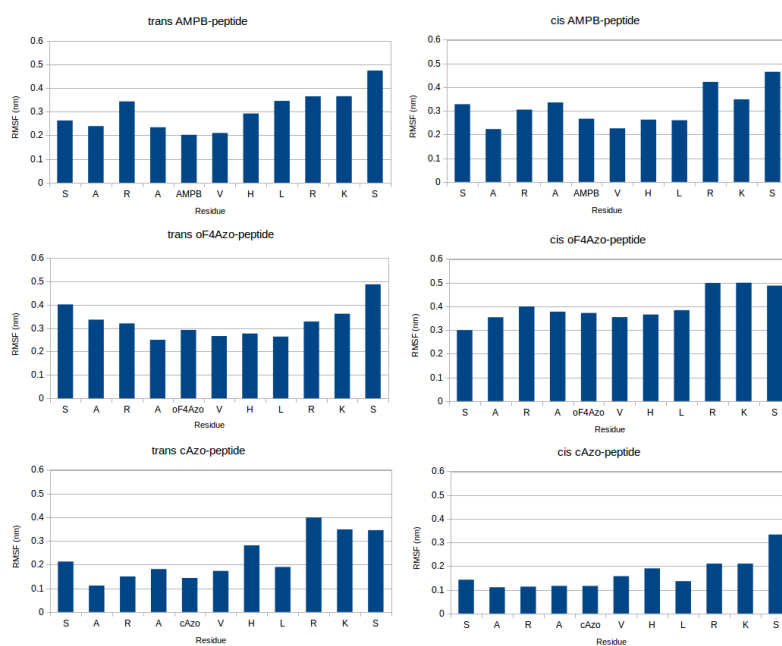


Figure S31. Fluctuations of residues from peptidomimetics expressed as root mean square fluctuations (RMSF) during 20 ns of molecular dynamics simulations.

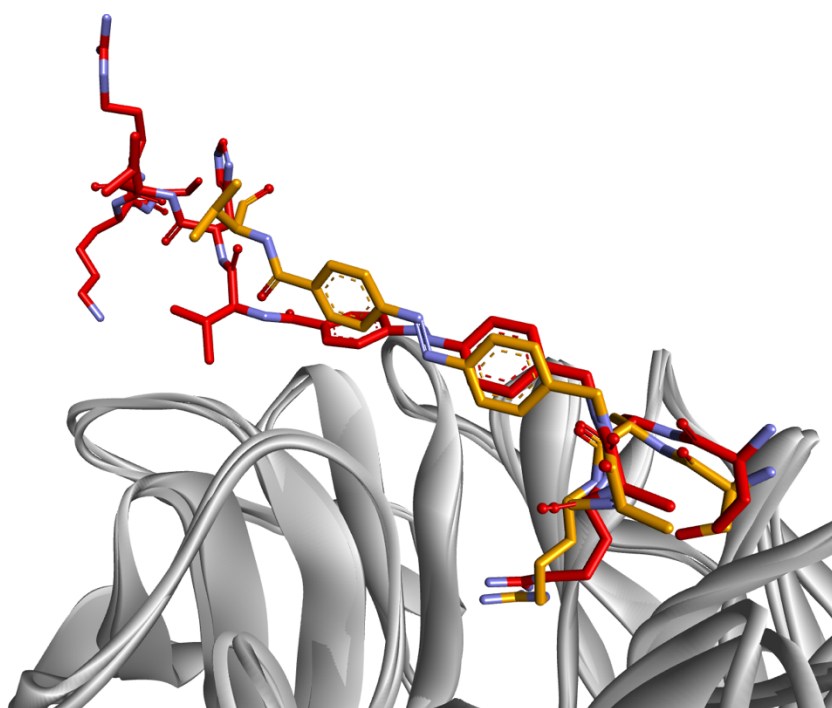


Figure S32. Comparison between binding modes of *trans* AMPB-peptide (3) predicted in MD simulation (red) and the described in the crystal structure PDB code: 5M23 (orange).

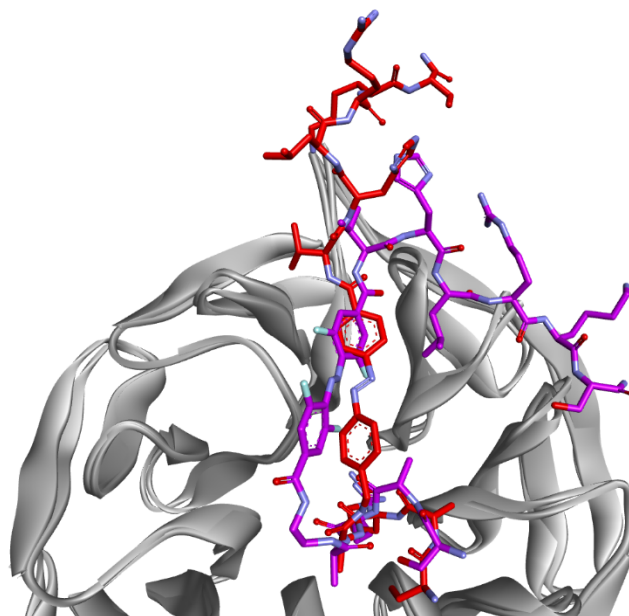


Figure S33. Comparison between binding modes predicted in MD simulations for *trans* AMPB-peptide (3) (red) and for *trans* oF₄Azo-peptide (5) (dark purple).

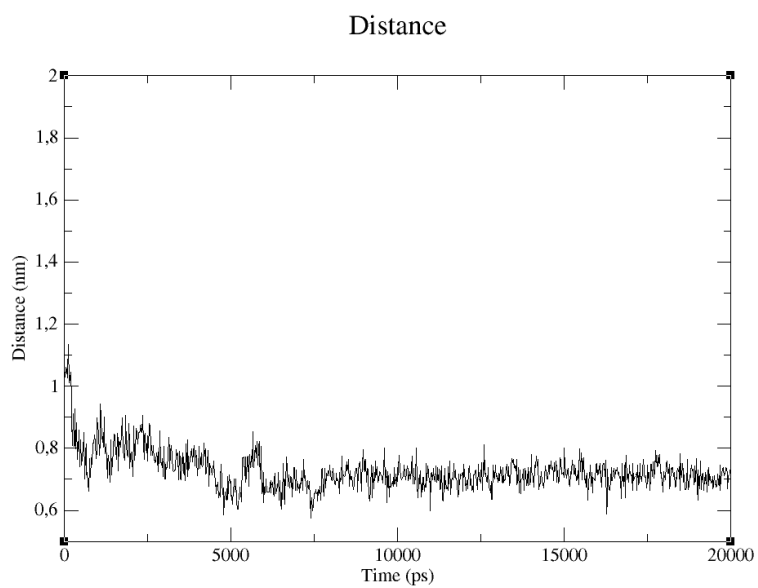


Figure S34. Measured distances between leucine of *trans* oF₄Azo-peptide (5) (-VHLRKS sequence of peptides) and tetra-ortho-fluoroazobenzene during 20 ns of simulation. Stabilization of intramolecular interaction was achieved after 7 ns.

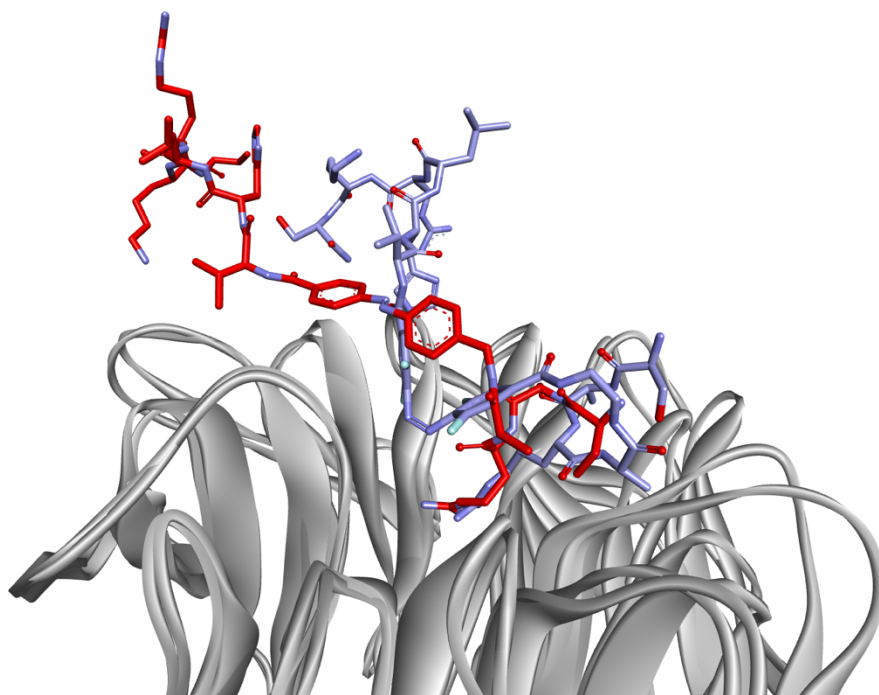


Figure S35. Comparison between binding modes predicted in MD simulations for *trans* AMPB-peptide (3) (red) and for *cis* oF4Azo-peptide (5) (light purple).

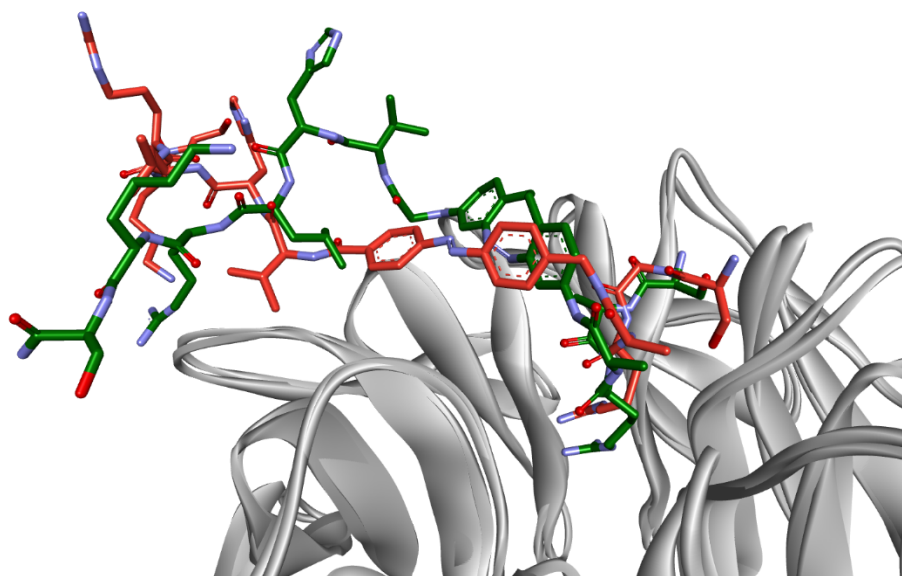


Figure S36. Comparison between binding modes predicted in MD simulations for *trans* AMPB-peptide (3) (red) and for *trans* cAzo-peptide (4) (dark green).

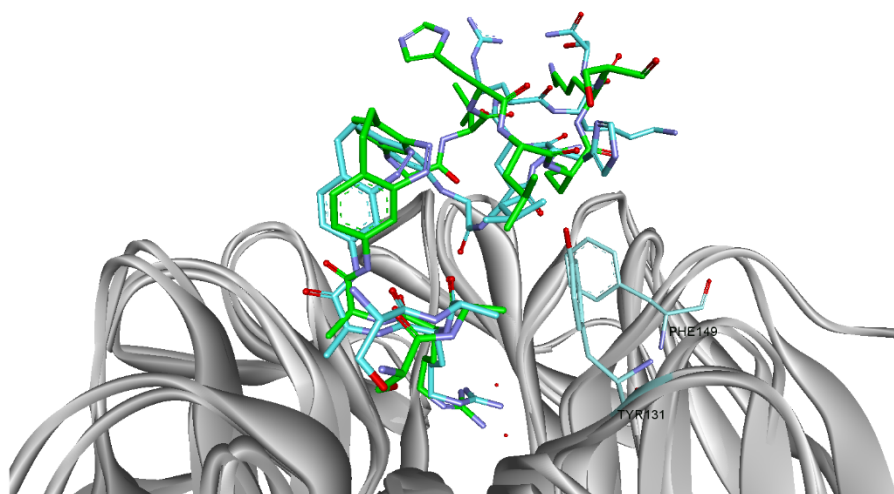


Figure S37. Comparison between binding modes of *cis* **cAzo-peptide (4)** computationally predicted (turquoise) and the described in the crystal structure PDB code: 6IAM (green).

Table S3. Average and standard deviation for RMSD values through 20 ns of MD runs.

ID	Average	Standard Deviation
<i>trans</i> AMPB-peptide (3)	0.080306	0.017255
<i>cis</i> AMPB-peptide (3)	0.213613	0.034478
<i>trans</i> oF₄Azo-peptide (5)	0.214039	0.04802
<i>cis</i> oF₄Azo-peptide (5)	0.224058	0.033466
<i>trans</i> cAzo-peptide (4)	0.221753	0.033309
<i>cis</i> cAzo-peptide (4)	0.243952	0.086662

Table S4. Surface area of protein-peptidomimetics binding surface calculated by Voliera.^[17]

ID	Binding surface area (Å ²)
<i>trans</i> AMPB-peptide (3)	2153.999
<i>cis</i> AMPB-peptide (3)	1643.134
<i>trans</i> oF₄Azo-peptide (5)	1617.568
<i>cis</i> oF₄Azo-peptide (5)	1533.619
<i>trans</i> cAzo-peptide (4)	1827.268
<i>cis</i> cAzo-peptide (4)	1287.747

References

- [1] C. Knie, M. Utecht, F. Zhao, H. Kulla, S. Kovalenko, A. M. Brouwer, P. Saalfrank, S. Hecht, D. Bléger, *Chem. - A Eur. J.* **2014**, *20*, 16492–16501.
- [2] R. Sjöback, J. Nygren, M. Kubista, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **1995**, *51*, L7–L21.
- [3] L. Albert, J. Xu, R. Wan, V. Srinivasan, Y. Dou, O. Vázquez, *Chem. Sci.* **2017**, *8*, 4612–4618.
- [4] Z. Nikolovska-Coleska, R. Wang, X. Fang, H. Pan, Y. Tomita, P. Li, P. P. Roller, K. Krajewski, N. G. Saito, J. a. Stuckey, et al., *Anal. Biochem.* **2004**, *332*, 261–273.
- [5] H. Sell, C. Näther, R. Herges, *Beilstein J. Org. Chem.* **2013**, *9*, 1–7.
- [6] D. Bleger, J. Schwarz, A. M. Brouwer, S. Hecht, *J. Am. Chem. Soc.* **2012**, *134*, 20597–20600.
- [7] D. Bléger, S. Hecht, *Angew. Chemie - Int. Ed.* **2015**, *54*, 11338–11349.
- [8] D. Chouikhi, M. Ciobanu, C. Zambaldo, V. Duplan, S. Barluenga, N. Winssinger, *Chem. - A Eur. J.* **2012**, *18*, 12698–12704.
- [9] W. J. Fang, T. Yakovleva, J. V. Aldrich, *Biopolymers* **2011**, *96*, 715–722.
- [10] W. S. Hancock, J. E. Battersby, *Anal. Biochem.* **1976**, *71*, 260–264.
- [11] R. S. Molday, S. W. Englander, R. G. Kallen, *Biochemistry* **1972**, *11*, 150.
- [12] Y. Bai, J. S. Milne, L. Mayne, S. W. Englander, *Proteins* **1993**, *86*, 75–86.
- [13] Y. Zhang, Y. Paterson, H. Roder, *Protein Sci.* **1995**, *9*, 804–814.
- [14] N. Hayashi, H. Kuyama, C. Nakajima, K. Kawahara, M. Miyagi, O. Nishimura, H. Matsuo, T. Nakazawa, **2014**, *Biochemistry*, *53*, 1818–1826
- [15] M. Kielmas, J. Adamczyk, *Anal. Bioanal. Chem.* **2014**, *406*, 8013–8020.
- [16] A. C. Wallace, R. A. Laskowski, J. M. Thornton, *Protein Eng.* **1995**, *8*, 127–134.
- [17] J. V. Ribeiro, J. A. C. Tamames, N. M. F. S. A. Cerqueira, P. A. Fernandes, M. J. Ramos, *Chem. Biol. Drug Des.* **2013**, *82*, 743–755.

4.3.1 Author Contribution

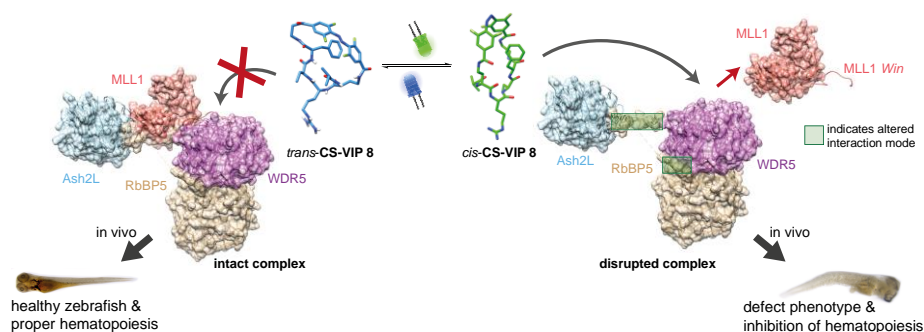
The idea of providing novel visible-light responsive peptide backbone photoswitches to control MLL1 functions was envisioned by Jun.-Prof. Dr. Olalla Vázquez. Synthesis, purification and characterization of these switches were performed by me, Lea Albert. I incorporated these new compounds into peptides, purified and characterized them. I conducted the complete photochemical characterization as well as FP-based assays, for which I expressed and purified the WDR5 protein. In addition, I expressed and purified the remaining MLL1 complex proteins (MLL1, Ash2L, RbBP5) and reconstituted the whole MLL1 complex. I performed interpretation of the obtained co-crystal structure and, after initial introduction to the principles of histone-methyltransferase assay and GST-pull down assays by Dr. Jing Xu, these experiments were also performed by myself in the laboratory of Prof. Yali Dou. Under my supervision, Alberto Peñalver (Bachelor student in our laboratory) and especially Malte Hoffarth (Master student in our laboratory) performed synthesis of high amounts of the molecular photoswitches, which were then used utilized for peptide incorporation and stability testing of the *o*F₄Azo. Malte Hoffarth further participated in synthesizing a linear precursor peptide. In collaboration with Laura Werel (group of Prof. Dr. Lars-Oliver Essen) we set up crystallization trials. The crystal structure was solved by Laura Werel and Prof. Dr. Lars-Oliver Essen, who also supported evaluation and analysis of the obtained structure. Molecular modelling studies were performed in the laboratory of Prof. Dr. Katarina Nikolić by Nemanja Djoković and Dušan Ružić. I contributed to write the article. Jun.-Prof. Dr. Olalla Vázquez conceived of the idea for the study, designed, directed and interpreted experiments, wrote the manuscript and had the role of corresponding author, i.e., she was leading and/or involved in all the 14-role taxonomy for author.¹

Marburg,

Jun. Prof. Dr. Olalla Vázquez

Lea Albert

4.4 *In Vivo* Optochemical Control of Hematopoiesis by MLL1 Complex Disruption



We report *in vivo* photocontrol of hematopoiesis via optochemical modulation of MLL1 complex architecture utilizing conformationally strained visible-light photoswitches (CS-VIPs).

This chapter was submitted as:

L. Albert, W. Steinchen, L. Werel, J. Nagpal, N. Djokovic, D. Ruzic, M. Hoffarth, J. Xu, A. Royant, G. Bange, K. Nikolic, S. Ryu, Y. Dou, L.-O. Essen, O. Vázquez*,
In Vivo Optochemical Control of Hematopoiesis by MLL1 Complex Disruption
Nature Chemical Biology 2020, manuscript submitted.

Upon acceptance of the article: Material in this chapter was reproduced from the above mentioned reference with permission from Nature Research.

In the following sections, the molecules included in this manuscript, will be named as **MD-X**, in which MB means manuscript D, and X means number of compound as given in this manuscript.

***In Vivo* Optochemical Control of Hematopoiesis by MLL1 Complex Disruption**

Lea Albert¹, Wieland Steinchen^{1,2}, Laura Werel¹, Jatin Nagpal³, Nemanja Djokovic⁴, Dusan Ruzic⁴, Malte Hoffarth¹, Jing Xu⁵, Antoine Royant^{6,7}, Gert Bange^{1,2}, Katarina Nikolic⁴, Soojin Ryu^{3,8}, Yali Dou⁹, Lars-Oliver Essen^{1,2}, Olalla Vázquez^{1,2*}

Conditional modulation of the epigenome to tune transcription profiles and, in turn, phenotypes is a novel approach to study disease-causing mechanisms, and may ultimately lead to epigenetics-based therapeutics, e.g. in hematopathology. In blood homeostasis the histone methyltransferase MLL1 (KMT2A) plays a crucial, yet controversial, role. Here, we present a photopharmacological approach based on new Conformationally Strained, Visible-light Photoswitches (CS-VIPs). Among our CS-VIPs, CS-VIP 8 optimally fulfils the requirements of an *on-off* switch for hematopoiesis by photoinhibition. Combining HDX-MS with MD-calculations, X-ray crystallization and *in vivo* analyses, we uncover the molecular mechanism of WDR5-binding inhibitors within the intact MLL1 complex. Upon WDR5 binding, CS-VIP 8 causes MLL1 release from the core complex with concomitant allosteric rearrangements in the WDR5/RbBP5 interface. Using zebrafish as a model, we can now control hematopoiesis on an organismic level without genetic intervention, thus promoting further studies on the complex MLL1 biology.

Throughout life, blood cellular components must be in continuous regeneration to sustain blood system homeostasis as well as to respond to increased demands due to injury or infection¹. Epigenetics coordinates the gene expression programs responsible for balancing hematopoiesis². Improper orchestration of these epigenetic mechanisms can cause aberrant hematopoietic stem cell (HSC) differentiations, malfunction, mutations and abnormal recruitments leading to severe hematological malignancies like leukemia³. The histone methyltransferase MLL1 (also known as KMT2A) is part of the so-called MLL1 complex (MLL1, WDR5, RbBP5, Ash2L, DPY30), where subunit interactions control MLL1's activity. As chromatin-modifier, the MLL1 complex is vital for sustaining both hematopoiesis⁴ and HSC self-renewal⁵. Furthermore, the MLL1 complex is a potent oncogenic driver in cancers of the hematopoietic systems by influencing the expression of *Hox* genes⁶. Interestingly, conditional knockout studies showed controversial results regarding the need of a functional MLL1 complex during hematopoiesis⁷. MLL1 may maintain its hematopoietic roles using mechanisms that do not always depend on its histone methyltransferase (HMT) activity⁸, but on protein-protein interactions (PPI) within the MLL1 core complex⁹. Nevertheless, in hematopathology the functional consequences of PPI modulations within the whole MLL1 complex remains elusive. To decode the complex biology underlying MLL1-dependent hematopoiesis, we require integrative approaches involving *in vivo* models, which can be interrogated in a highly controlled fashion.

The precise photocontrol of molecular activity to systematically trigger phenotypic traits opens unprecedented venues for the *in vivo* elucidation of complex biological processes¹⁰⁻¹³. Since the emergence of the first

genetically encoded photoreceptors in neuroscience^{14,15}, the field has evolved. A more recent approach, photopharmacology^{16,17}, provides synthetic photoswitches that can control functional outputs within unmodified biological targets, bypassing the delivery of genes into target cells^{18,19}. This strategy holds the promise to minimize off-target effects and resistance by on-demand therapeutic action. Recently, true *in vivo* applications where photoswitches reversibly modulate vision²⁰, membrane transport²¹ or cytoskeleton dynamics²² have been reported^{11,23}. However, there are no precedents of *in vivo* photocontrol of hematopoiesis. Such tools would shed light into the molecular mechanisms behind the signalling pathways involved in blood cell formation, and may ultimately lead to novel therapeutics for hematological malignancies.

Herein, we report Conformationally Strained Visible-light Photoswitches (CS-VIPs), which represent a genuine reversible binary system in their biological response. CS-VIPs undergo exceptional slow relaxation in aqueous solution. Employing X-ray crystallography, hydrogen-deuterium exchange mass spectrometry (HDX-MS) molecular dynamics (MD) calculations and biological assays, we define the action mechanism of these WDR5-binding inhibitors on the whole MLL1 multiprotein complex. Importantly, our most potent photopharmacological agent, **CS-VIP 8**, is a true *on-off* switch that enables *in vivo* photocontrol over the hematopoietic function for the first time, to our knowledge.

Results

Design and synthesis of the conformationally strained visible-light photoswitches (CS-VIPs). Recently, we reported *in vitro* photocontrol of MLL1 activity, which

¹Department of Chemistry, University of Marburg, Germany. ²Center for Synthetic Microbiology (SYNMIKRO), University of Marburg, Germany. ³University Medical Center, Johannes Gutenberg University Mainz, Germany. ⁴Department of Pharmaceutical Chemistry, University of Belgrade, Serbia. ⁵Department of Pathology, University of Michigan, USA. ⁶Institut de Biologie Structurale, University Grenoble Alpes, CEA, CNRS, France. ⁷European Synchrotron Radiation Facility, Grenoble, France. ⁸College of Medicine and Health, University of Exeter, UK. ⁹Norris Comprehensive Cancer Center, University of Southern California, USA e-mail: olalla.vazquez@staff.uni-marburg.de

ultimately affected the growth of leukemia cells²⁴. However, the modest difference in activity between isomers precluded the possibility of using our probes for *in vivo* experiments. To tackle this issue, we envisioned that the inclusion of a molecular transducer into a structurally-defined scaffold should impart higher restraints to favor or disfavor conformational preferences and, thereby, to have a strong influence on the different biological output between photoisomers. The potent macrocyclic MLL1 inhibitors²⁵ are hence ideal to test if strained photoswitchable cyclopeptides can exert an effective photomodulation of the molecular recognition process of MLL1, which affects *in vivo* hematopoiesis. In addition, cyclic peptidomimetics provide pharmacological advantages, namely higher cellular potency, metabolic stability and resistance against exopeptidases. Thus, we synthesized photoswitchable cyclic peptidomimetics based on MM-401²⁵ (Fig. 1a) bearing the *o*F₄Azo photoswitch developed by Hecht due to its excellent photochemical properties²⁶. Initially, we explored a set of 13 CS-VIPs, which differed in the position where the *o*F₄Azo moiety is grafted, the presence of D-phenylglycine and the

ring size. The latter was achieved by varying the length of the connecting amino acid side chain (lysine, diamino-butyric acid, diaminopropionic acid, or none) between the interaction motif (-Arg-Abu-) and the *o*F₄Azo unit. D-Phenylglycine is often considered as troublesome amino acid due to the risk of epimerization during solid-phase synthesis²⁷. Therefore, we considered its elimination as well as its exchange by D-phenylalanine. In brief, solid phase peptide synthesis (SPPS; Fig. 1b) enabled the rapid generation of our linear precursors, in which *o*F₄Azo was always on-resin incorporated at last, and the key arginine was Boc-protected to circumvent both lability under standard Fmoc deprotection conditions²⁸ and side reactions during Pbf-deprotection²⁹, respectively. Afterwards, they were either head-to-tail or side-chain-to-tail cyclized at high dilution (1 mM) and irradiated at 520 nm for 3 min to steer the yield towards the intramolecular product via *cis* isomerization³⁰.

Photochemical characterization. As an illustrative example, the spectroscopic properties of CS-VIP 8 (Fig. 2)

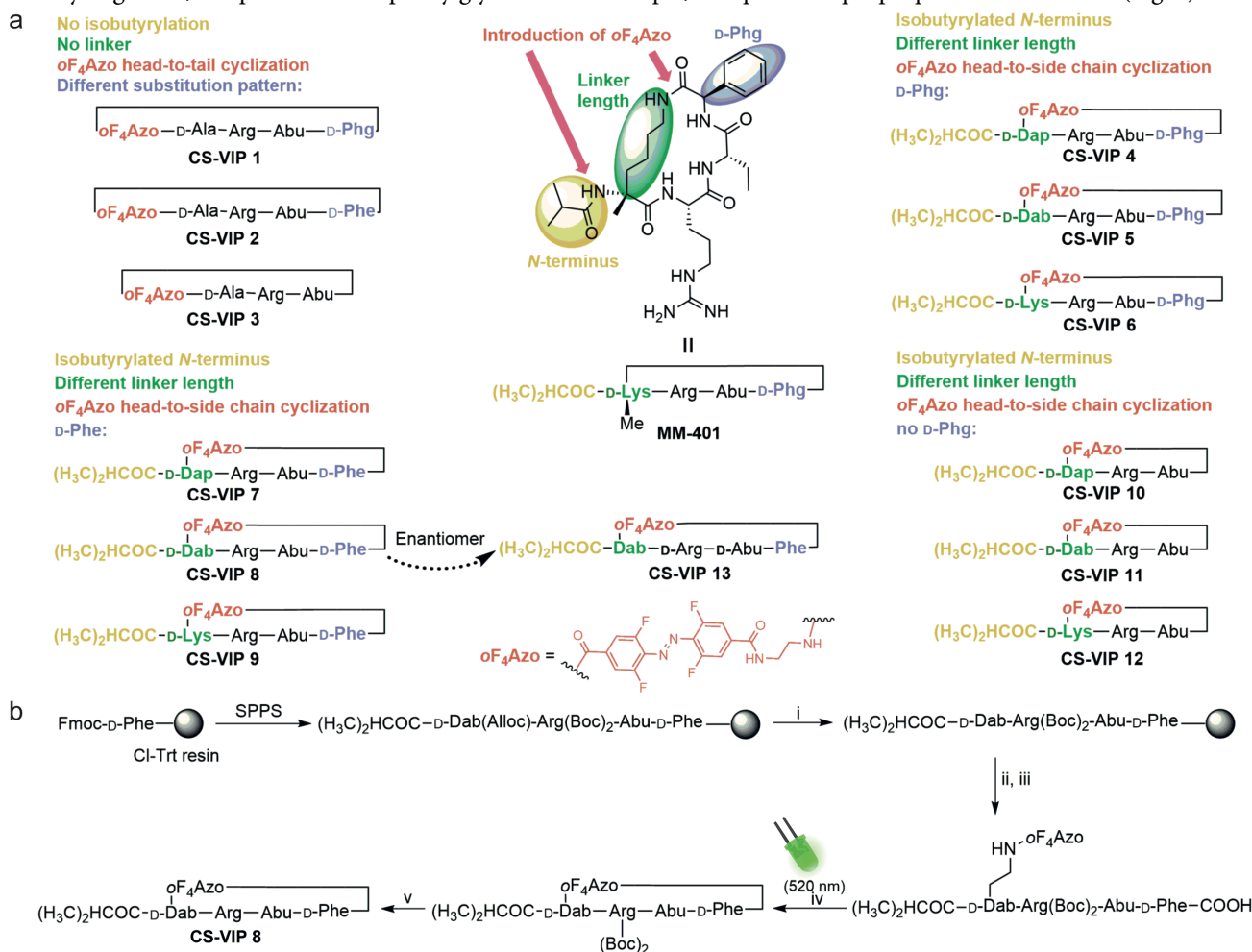


Figure 1 | Design and synthesis of CS-VIPs. a, Design of photoswitchable cyclopeptides based on MM-401. b, Synthesis route exemplified by CS-VIP 8. i: 1.00 eq. Pd(PPh₃)₄, 190 eq. morpholine, DCM, 2 h, rt; ii: 4.00 eq. oxyma, 4.00 eq. DIC, 4.00 eq. Mtt-*o*F₄Azo, DMF, 45 min, rt; iii: HFIP/DCM 1:4, 2 h, rt; iv: 3.00 eq. HATU, 4.00 eq. DIPEA, THF/DMF 98:2, 45 min, rt, at 520 nm 3 min; v: 4 M HCl in dioxane, 3 h, rt. rt: room temperature; CS-VIPs: conformationally strained visible-light photoswitches; DIC: diisopropylcarbodiimide, Mtt: mehyltrityl, *o*F₄Azo: tetra-*ortho*-fluoroozobenzene, HFIP: hexafluoroisopropanol, HATU: 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate, DIPEA: *N,N*-diisopropylethylamine, THF: tetrahydrofuran, DMF: dimethylformamide, D-Phe: D-phenylalanine, Cl-Trt: trityl chloride, SPPS: solid phase peptide synthesis, D-Dab: D-diaminobutyric acid, Abu: aminobutyric acid, Boc: butyloxycarbonyl, Alloc: allyloxycarbonyl.

were analyzed in detail. Its UV-vis spectra displayed the expected absorbance bands of *o*F₄Azo-containing compounds in water (Fig. 2b)^{28,31}. Photoisomerization was fully reversible for up to 12 cycles without any observable photodegradation (Fig. 2c). Interestingly, the thermodynamically unstable *cis* isomer of the *o*F₄Azo building block is kinetically highly stable in organic solvents (*o*F₄Azo half-life: 700 days at room temperature in DMSO, 92 h at 60 °C in acetonitrile)³². In contrast, thermal *cis* → *trans* conversion proceeds apparently faster in aqueous solution (*o*F₄Azo linear peptide *cis* content 86 % at 520 nm photostationary state, (PSS); 76 % after 7 days)²⁸. Importantly, when we incorporated the *o*F₄Azo moiety into a conformationally restricted system such as **CS-VIP 8**, its extraordinary stability was not only retained in water, but surpassed the reported non-restricted *o*F₄Azo molecules^{28,33} as well as unsubstituted azobenzene-containing cyclopeptides³⁰. Once irradiated and stored in darkness, we did not observe any change in the *cis:trans* isomer ratios for at least five months (Figs. 2d,e). Notably, the 520 nm photostationary state (PSS) achieved an accumulation of 95% *cis* isomer (*cis:trans* 19:1), whereas the 405 nm PSS formed a *cis:trans* mixture in a ratio of ~1:10. Consequently, a switch between these PSS is capable to cause a ~ 9.5/18-fold change of the *cis/trans*-isomer concentration, respectively. Finally, considering biologically relevant *in vivo* studies, we found that the stability against reducing agents such as glutathione (GSH) was also improved (Fig. S18). Preliminary photochemical studies of the other CS-VIPs showed a similar behavior.

Optimization of CS-VIP binding to WDR5. We employed fluorescence polarization (FP)-based assays to evaluate the

binding potential of our CS-VIPs to WDR5 (Table 1, Fig. S19). Importantly, the incorporation of *o*F₄Azo generally retained mostly the parental nanomolar binding of MM-401 to WDR5 ($K_i < 1$ nM) for most CS-VIPs (**4-9** and **11**). The variation of the apparent affinities depended on the substitution pattern, and the ring size of the CS-VIPs (Table 1) as well as on the effective concentration of the binding-active isomer as set by the chosen PSS. **CS-VIPs 1-3**, which unlike MM-401, lack the isobutyrylated *N*-terminus and were head-to-tail cyclized, displayed either no detectable (**CS-VIP 1** and **3**) or low micromolar (**CS-VIP 2**) binding. *In silico* binding poses of both isomers for the non-binding CS-VIPs (**1** and **3**) revealed that the key arginine is not inserted in the central cavity of WDR5 (Fig. S36). Besides, our experiments reassert the reported positive effect of the isobutyrylated *N*-terminus on WDR5 binding affinity³⁴.

Considering CS-VIPs with the same ring size, we observed that the phenyl group is crucial for high binding affinities, i.e., **CS-VIP 10-12** devoid of this moiety displayed only micromolar affinities to WDR5 and lacked light-dependent switching. In our virtual docking (VD) study, the binding poses of **CS-VIP 10-12** were also ranked lower (Table S6, Fig. S39) than the ones bearing D-Phg (**CS-VIP 4-6**, Fig. S37) or D-Phe (**CS-VIP 7-9**, Fig. S38). Again, the D-Phe-containing CS-VIPs displayed the highest affinity differences between isomers. Also, for the inactive **CS-VIP 1**, the substitution of D-Phg for D-Phe (**CS-VIP 2**) increased the affinity being consistent with docking simulations (Table S6, Fig. S36). In contrast to **CS-VIP 1** and **CS-VIP 3**, packing interactions between the arginine and the WDR5 residues F133, F263 and C261 could be

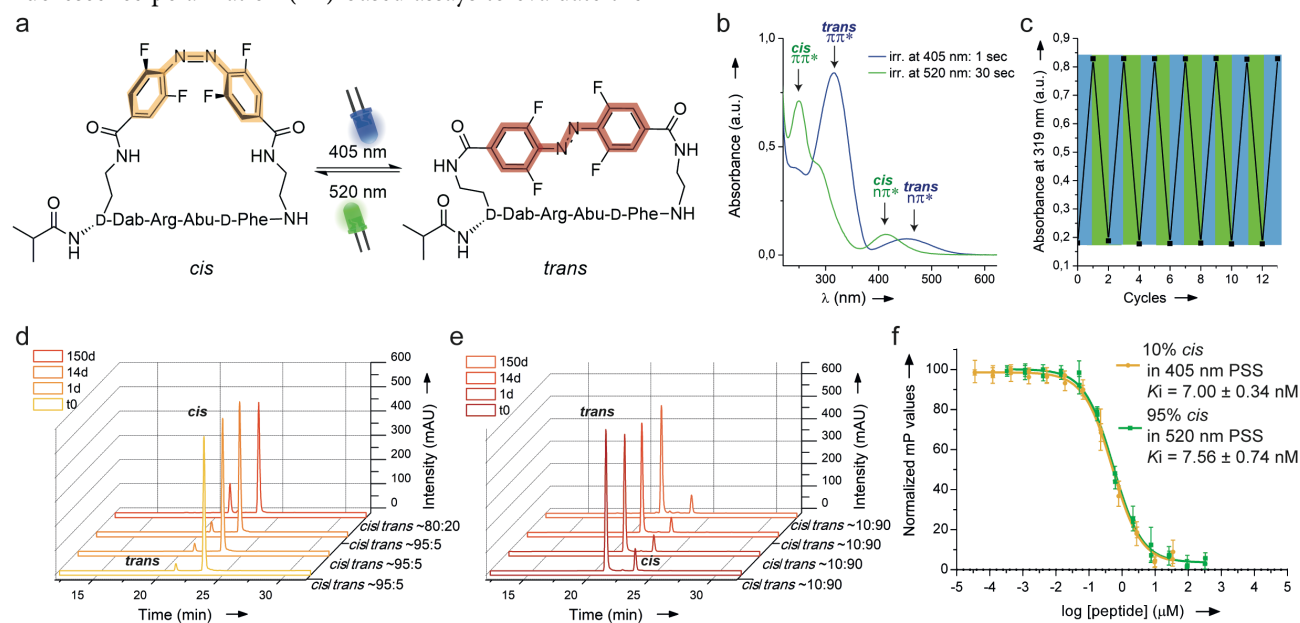


Figure 2 | Photoisomerization studies of CS-VIPs and interaction to WDR5. a, Isomerization of **CS-VIP 8**. b, UV-vis spectra in aqueous solution of **CS-VIP 8** after irradiation at 405 nm and 520 nm to reach the photostationary states (PSS) achieving predominantly the *trans* and *cis* isomer, respectively. c, Reversible *cis:trans* cycling between photoisomers by alternating illumination at 405 (blue)/520 (green) nm. d, Stability of the *cis* isomer of **CS-VIP 8** at the 520 nm PSS at r.t. in water via HPLC measurements. e, Stability of the *trans* isomer of **CS-VIP 8** at the 405 nm PSS at r.t. in water via HPLC measurements. f, Results of FP-based assays of **CS-VIP 8** obtained by generating 405 nm and 520 nm PSS considering only the residual 10% of the *cis* isomer present in solution (yellow) and compared with the *cis* state as obtained by illumination at 520 nm (green) including the calculated K_i values.

could be predicted for **CS-VIP 2** (Fig. S36). Consequently, our D-Phe modification not only improved the synthetic accessibility of the CS-VIPs but also its functional properties.

An optimal ring size is decisive for maximizing both the interaction and the conformational change between the photoisomers. In our case, the calculated CHEMPLP values (Table S6) suggested that the 2-carbon linker (D-Dab) had the ideal length to supply the best **CS-VIPs: 5 and 8**, which we verified experimentally (Table 1, Fig. S19). Furthermore, the latter excels with an almost 10-fold switch of binding affinity upon change of the PSS. This is in agreement with the theoretical limit for a compound being only active in the *cis*-state. To confirm this, we adjusted the curve fit parameters of our FP assay to the effective *cis* isomer concentration of **CS-VIP 8** (Fig. 2f). We obtained an excellent concordance between the K_i values of the 520 nm PSS (green) and the 405 nm PSS (yellow), when only considering the concentration of the residual *cis* isomer (10%). In addition, when we conducted FP-based competitive assays where samples were illuminated after complexation with WDR5 (Fig. S20), only the *trans* PSS mixture of **CS-VIP 8** isomerized to the *cis* isomer yielding comparable K_i values (Fig. S20). In contrast, *cis* \rightarrow *trans* photoisomerization is apparently slightly less effective when **CS-VIP 8** is bound to WDR5 (K_i 405 nm PSS = 64.7 nM; K_i *cis* \rightarrow *trans* = 47.1 nM; Fig. S20).

Taken all together, our **CS-VIP 8** variant is the most promising candidate for *in vivo* experiments as it maintains the nanomolar parental WDR5 affinity with a \sim 10-fold difference between its photostationary isomers. These features significantly improved our previous results with photoresponsive linear peptides^{24,28}. As for the MM-401 enantiomer²⁵, the **CS-VIP 8** enantiomer, **CS-VIP 13**, exerted a 244/803-fold diminished binding to WDR5 relative to **CS-VIP 8** judging from its K_i values recorded at 405 nm and 520 nm PSSs, respectively (Table 1).

Table 1 | Inhibition constants (K_i) of CS-VIPs to WDR5 by fluorescence polarization (FP)-based competitive assays.

CS-VIP	PSS at 405 nm K_i [μ M]	PSS at 520 nm K_i [μ M]	ratio
1	n.c.	n.c.	-
2	\sim 3.35*	\sim 1.31*	2.56
3	n.c.	n.c.	-
4	0.043 ± 0.006	0.079 ± 0.016	1.84
5	0.051 ± 0.010	0.008 ± 0.002	6.38
6	0.394 ± 0.023	0.239 ± 0.071	1.65
7	\sim 0.448*	0.104 ± 0.010	4.31
8	0.079 ± 0.004	0.008 ± 0.0008	9.88
9	0.831 ± 0.054	0.137 ± 0.012	6.07
10	\sim 1.10*	\sim 1.52*	1.38
11	0.688 ± 0.004	0.459 ± 0.060	1.50
12	\sim 3.69*	\sim 2.90*	1.27
13	\sim 19.3*	\sim 6.42*	3.00

Mean values from at least two independent measurements (in triplicates). *estimated values due to lack of the bottom plateau; n.c.: not calculable.

To confirm the interaction mode of our CS-VIPs, we determined the co-crystal structures of the WDR5•**CS-VIP 8** complex (Fig. 3). Two novel WDR5 crystal forms for complexes with *cis*-**CS-VIP 8** have been obtained using different *cis:trans* ratios during crystallization. One by co-crystallization with **CS-VIP 8** set to a 1:10 *cis:trans* mixture formed by 405 nm PSS (*cis1*, PDB entry 7AXS), the other with the predominant *cis* isomer formed at the 520 nm PSS (*cis2*, PDB entry 7AXP). Both states share the same interaction mode, which closely mimics the binding mode of MM-401 (PDB: 4GM9, Fig. S27) along the isobutyryl-D-Dab-Arg-Abu-DPhe motif. Interestingly, the superposition of these *cis*-**CS-VIP 8** complexes (Fig. 3a) revealed different orientations of both the *o*F₄Azo and D-Phe moieties, which are caused by different crystal packing contacts involving the *o*F₄Azo group. In the *cis1* crystal form, the D-Phe ring enables π -stacking contacts with Y260 like D-Phg in MM-401. The *cis2* crystal form has a higher solvent content and features less H-bonds and no π -stacking of *cis*-**CS-VIP 8** to Y260 (Figs. 3b, S25). UV-vis microspectroscopy displayed only the *cis* isomer of **CS-VIP 8** within *cis1* crystals (Fig. S28).

In order to reveal structural consequences of *cis:trans* isomerization within WDR5•**CS-VIP 8** complexes, we irradiated the *cis1* crystallization setups of WDR5•*cis*-**CS-VIP 8** after crystal formation at 405 nm to enforce *cis:trans* switching (PDB entry: 7AXU). UV-vis measurements verified *cis* \rightarrow *trans* photoisomerization within the crystals (Fig. S28) like in solution. However, this procedure also caused complete diffusion of **CS-VIP 8** out of the WDR5-binding site as indicated by a lack of difference electron density for the ligand (Fig. 3d). Finally, we tested whether *in crystallo* irradiation at 405 nm of isolated co-crystals enables us to obtain intermediary states for the *off*-diffusion of the *trans*-**CS-VIP 8** isomer out of the WDR5 binding site. *In crystallo* switching to *trans*-**CS-VIP 8** was observed at both 180 K (PDB entry 7AXQ, Fig. 3e) and 293 K (7AXX, Fig. S28). However, both of the resulting structures indicated X-ray mediated *trans* \rightarrow *cis* back isomerisation by displaying again the *cis* isomer of **CS-VIP 8** within the binding site that was confirmed by microspectroscopic analysis after X-ray exposure (Figs. 3f, S26, S28). We only observed increased thermal B factors for the *o*F₄Azo group, especially in the co-crystal cryo-trapped at 180 K. These observations further verified our previous FP experiments, which were additionally substantiated by our computational calculations (SN1). These corroborated that WDR5-affinity to *trans*-**CS-VIP 8** is about five magnitudes lower compared to *cis*-**CS-VIP 8**, and predicted that the presence of WDR5 largely increases the energy barrier for *cis* \rightarrow *trans* isomerization. Overall, our biophysical analyses showed that only the *cis* isomer of **CS-VIP 8** is active towards WDR5, which encouraged us to perform additional experiments in realistic biological contexts.

MLL1 inhibition and leukemogenesis suppression. Compounds with the ability to effectively bind to WDR5 have displayed inhibition of MLL1 activity by blocking

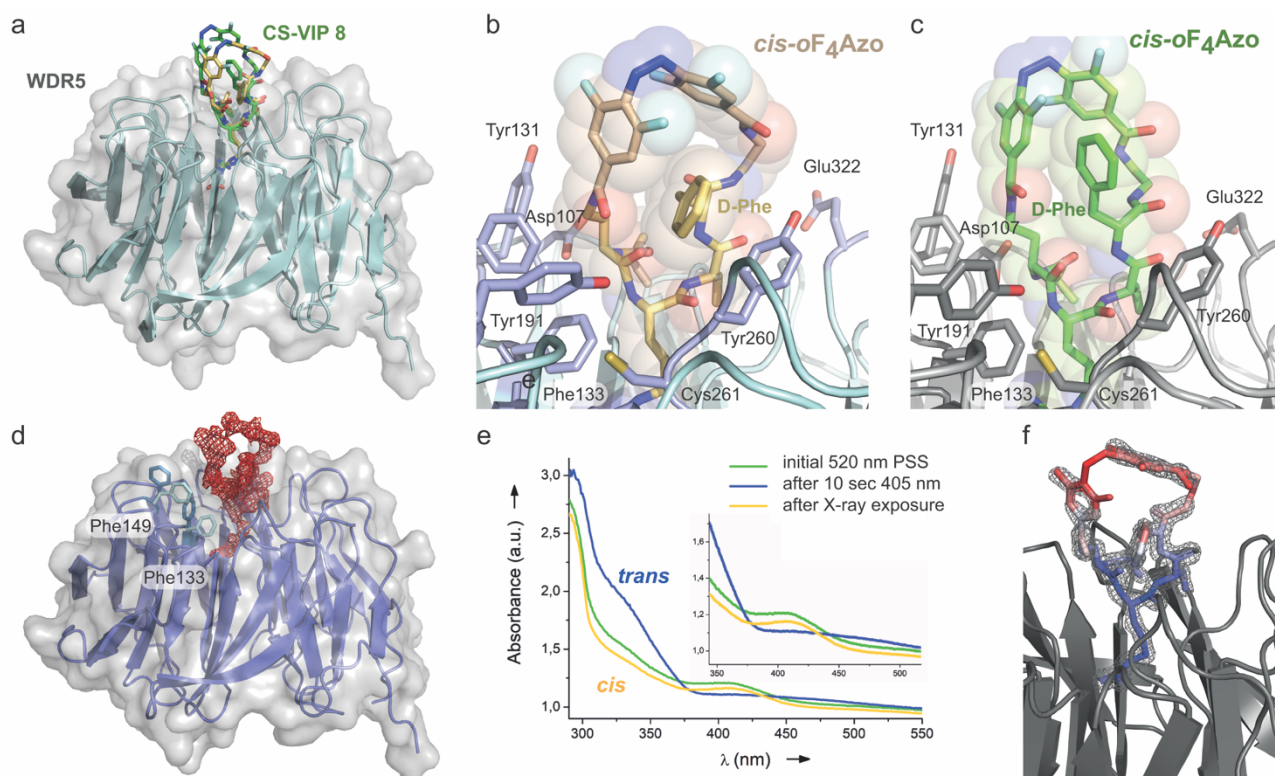


Figure 3 | Structural characterization of WDR5•CS-VIP 8 interaction. **a**, Co-crystal structures of WDR5•CS-VIP 8 complexes: overlay of *cis1* (WDR5: ribbon in light cyan, molecular surface in grey, CS-VIP 8: yellow) and *cis2* (CS-VIP 8: green) crystal forms demonstrating trapping of *cis*-CS-VIP 8 at both PSS. **b**, WDR5 binding site for CS-VIP 8 (yellow) in crystal form 1 (*cis1* state). **c**, WDR5 binding site for CS-VIP 8 (green) in crystal form 2 (*cis2*). **d**, Apo-structure of WDR5 as generated by 405 nm illumination *in situ*, i.e., after crystallization of WDR5•*cis*-CS-VIP 8 complex. The shown $F_{obs,1}-F_{obs,2}$ difference electron density (red, contouring level 2σ) was calculated for *cis1*-apo states using phases of the *cis1* WDR5•CS-VIP 8 complex. Notably, the side chains of Phe133 and Phe149 (marine blue: apo form; light cyan: *cis1* form) adopt a different conformation upon CS-VIP 8 release. **e**, Absorbance spectra of frozen WDR5 co-crystals recorded at the iCOS cryobench. Displayed are spectra for a WDR5•CS-VIP 8 crystal at 520 nm PSS (green), after irradiation at 405 nm for 10 sec at 180 K (blue) and after data collection of the same co-crystal (yellow). **f**, Binding site of the cryo-trapped WDR5•CS-VIP 8 structure from **e**; the CS-VIP 8 ligand is colored according to the B-factors ranging from 13 \AA^2 (blue) to 32 \AA^2 (red); SIGMAA-weighted $2mF_{obs}-DF_{calc}$ electron density (grey) is shown at a contouring level of 2σ .

MLL1/WDR5 interaction, which led to a decrease in leukemia cell proliferation³⁵. *In vitro* functional radiometric histone methyltransferase (HMT)-assays using the MLL1 complex (MLL1, WDR5, RbBP5, Ash2L)²⁴ demonstrated that *cis*-CS-VIP 8 is a potent inhibitor of MLL1 methylation activity with a 12-fold difference depending on illumination (IC_{50} 405 nm PSS = $9.20 \pm 3.04 \mu\text{M}$; IC_{50} 520 nm PSS = $0.792 \pm 0.259 \mu\text{M}$; Fig. S21). Along these lines, we tested the effect of CS-VIP 8 on different leukemia-related cells after 3-day incubation at different PSSs. The poor cellular uptake of CS-VIP 8 demanded the use of the pep-1 carrier as transfection method³⁶ as well as relatively high concentrations. Compared to the 405 nm PSS, our photoswitchable CS-VIP 8 clearly gained inhibitory potency upon irradiation at 520 nm in a dose-dependent manner in different leukemia cell lines (Figs. 4a, S22). In acute myeloid leukemia cells, MOLM-13, only CS-VIP 8 set to its 520 nm PSS entirely abolished cell viability at $200\mu\text{M}$ concentration. At $100 \mu\text{M}$ it repressed the leukemia cell growth in more than 70 % while CS-VIP 8 at the 405 nm PSS did not reach such inhibition even at twice concentration (Fig. 4a). Analogous behaviour was observed in MV4-11 cells (Fig. S22).

Isomers are distinct compounds with different pharmacokinetic properties. In our case, cell-uptake might depend on the photoisomer³⁷. Consequently, to exclude the possibility that the observed differences might have emerged only due to poor cell-permeability, we included a washing step after transfection in our protocol. To confirm specific interaction, we synthesized a close mimetic of CS-VIP 8 as negative control, an enantiomer, in which the key L-Arg residue was replaced by a D-Lys (CS-VIP 14). This compound lacked any detectable binding to WDR5 (Fig. S19). For that reason, CS-VIP 14 was the perfect control to evaluate both the effect of the *oF*₄Azo moiety itself, and the toxicity specificity via an MLL1-dependent inhibition mechanism. These new experiments are consistent with our previous *in cellulo* assays without additional washing (Fig. 4b), confirming that the detected cytotoxicity of *cis*-CS-VIP 8 is only attributed to an intracellular process. Importantly, the significant functional difference between isomers was maintained too. In contrast, the cytotoxicity of the *cis*-CS-VIP 14 was negligible, which ruled out unspecific effects from the *oF*₄Azo moiety, and highlighted the decisive role of the guanidinium group of the key L-Arg residue. Interestingly, this compound even exerted an inverse activity profile, i.e., higher biological output for the

trans than *cis* isomer, suggesting different mechanisms involved.

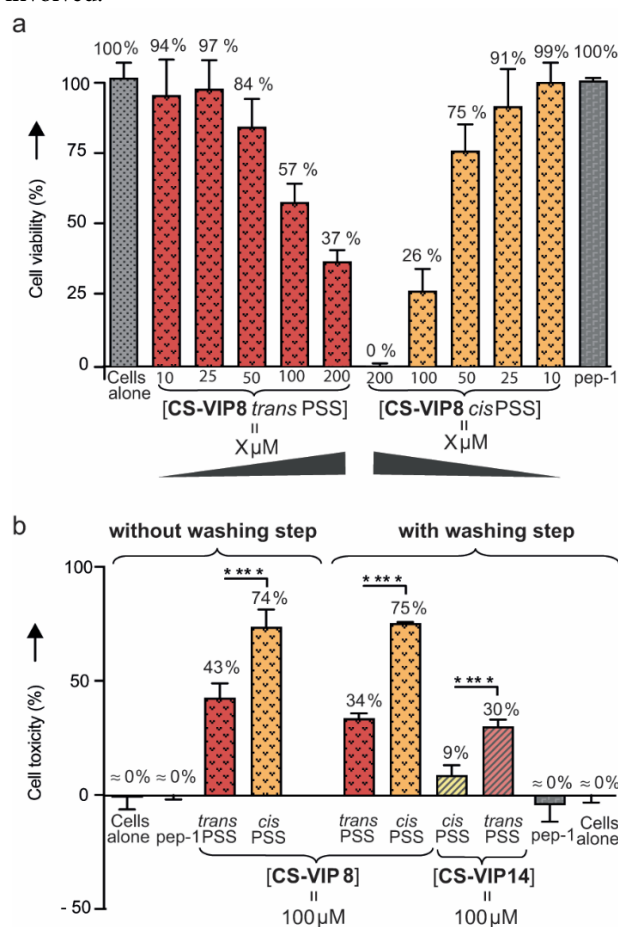


Figure 4 | Cell-based assays studying the effect of CS-VIPs on leukemia MOLM-13 cells. a, Cell viability after treatment with both isomeric states of CS-VIP 8 for 3 days; X: corresponding concentration. **b**, Comparison of cell toxicity of 100 μM CS-VIP-8 between unwashed and washed cells, as well as between CS-VIP-8 and the negative control CS-VIP-14 in washed-cell experiments. All mean data-points and standard deviations are derived from at least two independent experiments, each concentration in triplicate. **** means $p < 0.0001$

***cis*-CS-VIP 8 affects MLL1-complex assembly.** Despite the crucial contribution of MLL1 activity in regulating the leukemia transcription program²⁵, its HMT function seems to be dispensable for normal hematopoiesis⁸. However, depleting MLL1 complex components such as DPY30 and Ash2L displayed severe defects of lineage specification and differentiation^{9,38}. Having established the structural basis of efficient MLL1/WDR5 PPI distortion via the *cis*-CS-VIP 8 binding to WDR5, we employed hydrogen-deuterium exchange mass spectrometry (HDX-MS) to further corroborate our findings in solution by analyzing WDR5 in either absence or presence of *cis*-CS-VIP 8. Hereby, regions of reduced H/D-exchange of WDR5 clustered around the binding site as evidenced by our crystal structures and virtual docking (Figs. 5a, S29). We next aimed to study the impact of *cis*-CS-VIP 8 binding to WDR5 on the assembly of the whole MLL1 complex. Initially, we determined the H/D-exchange profiles of all constituents of the MLL1 complex, i.e. WDR5, MLL1, RbBP5, Ash2L and DPY30,

either within the complex or individually. Regions of reduced H/D-exchange correlated well with reported interaction interfaces³⁹⁻⁴³, thus, corroborating the integrity of our recombinant MLL1 complex (Figs. S30-S34). We then evaluated the H/D-exchange rates of the MLL1 complex in presence of *cis*-CS-VIP 8 and compared it with the untreated MLL1 complex. We obtained altered profiles for all complex components except for DPY30, and mapped those differences onto available crystal structures (Fig. 5b). The MLL1 polypeptide, *Win*, displayed increased H/D-exchange rates as well as the core of the MLL1 SET domain, suggesting loss of interaction with WDR5 and RbBP5/Ash2L (Figs. 5b, S30). Correspondingly, amino acids 336-354 of RbBP5 bridging between MLL1 and Ash2L exhibited likewise elevated H/D-exchange suggesting that MLL1 is displaced from the complex by *cis*-CS-VIP 8 (Figs. 5b, S32). Conformational changes observed for Ash2L are in proximity of its RbBP5/MLL1 interface (Fig. S33) which further strengthens the notion of MLL1 release. Pull-down assays with purified RbBP5, WDR5, Ash2L and MLL1 proteins clearly indicate MLL1 vacating the complex in presence of *cis*-CS-VIP 8 (Fig. S35). Although *cis*-CS-VIP 8-dependent H/D-exchange could not be detected at the *Win* binding site of WDR5, we postulate that this is due to opposing H/D-exchange differences induced by *cis*-CS-VIP 8 binding (Fig. 5a) and *Win* peptide release, roughly cancelling each other out. Notably, *cis*-CS-VIP 8 induced elevated HDX in WDR5 residues 220-230, which are remote from the ligand binding site (Figs. 5b, S31) and coincide with parts of the RbBP5-binding interface (PDB entry 3P4F)⁴². Our MD simulations, in which a comparison of changes in solvent accessible surface areas (SASA) was conducted, verified that the 220-230 region of WDR5 is indeed prone to deuterium exchange, as increased differences in SASA were calculated for hydrogens from peptide bonds of WDR5 alone vs. WDR5-*cis*-CS-VIP 8 (Figs. S42, SN2). Most importantly, conformational analysis of MD trajectories of the MLL1 complex either alone or complexed to *cis*-CS-VIP 8 substantiated the partial unbinding of RbBP5 from WDR5 residues 220-230 during the course of MD simulation. We found that allosteric effects of *cis*-CS-VIP 8 caused different dynamical behaviour, for the MLL1 polypeptide, which resulted in partial unbinding of RbBP5 (Figs. S41, 42, SN4). However, we conclude that complete dissociation of WDR5/RbBP5 is unlikely because no difference in deuterium uptake was observed at the second binding site of WDR5 to RbBP5 (residues 240-250, 289) upon *cis*-CS-VIP 8 addition. As WDR5, RbBP5 and Ash2L still interacted in pull-down assays (Fig. S32), this analysis suggests an altered quaternary structure of the remaining MLL1-devoid complex in presence of *cis*-CS-VIP 8.

Allosteric communication through WDR5. Further analysis on WDR5, using perturbation-response scanning (PRS) revealed that “sensor regions”, i.e., protein sites that are conformationally changed upon perturbation of effector residues, are positioned along the WDR5/RbBP5

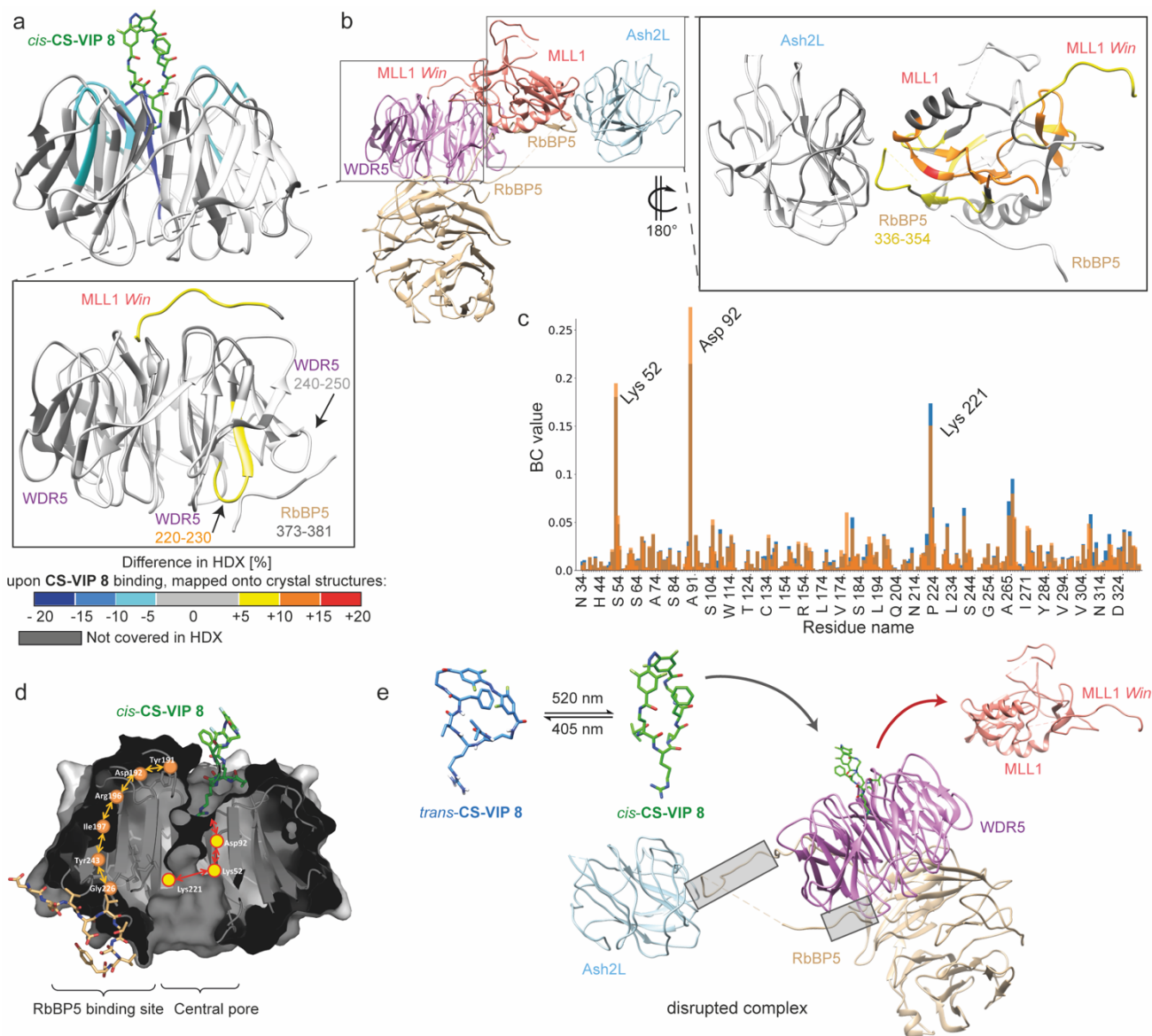


Figure 5 | MLL1 complex disruption triggered by CS-VIP 8 binding to WDR5 evaluated via HDX-MS and MD-studies. **a**, Difference in deuterium (D)-uptake in WDR5 alone vs. bound to *cis*-CS-VIP 8 (green) mapped onto our crystal structure (PDB: 7AXP). **b**, Structure of the 4-mer MLL1 complex (PDB: 6KIU)⁴⁴; WDR5: light purple, MLL1: salmon, RbBP5: tan, Ash2L: light blue. Zoomed-in inserts display difference in D-uptake between respective proteins in MLL1 complex without addition of CS-VIP 8 and after addition of CS-VIP 8, mapped onto crystal structures of: WDR5-MLL1-*Win*-RbBP5, (PDB: 3P4F)⁴² and of MLL1-RbBP5-Ash2L, (PDB: 5F6L)³⁹; dark grey coloured regions represent amino acids, that have not been covered in the HDX peptide coverage map. **c**, Betweenness-centrality (BC) values observed for: MLL1-*Win*-WDR5-RbBP5 (orange bars), and: *cis*-CS-VIP 8-RbBP5 (blue bars). **d**, Schematic representation of CS-VIP 8 mode of action via displaying shortest paths connecting MLL1 and RbBP5 binding sites of residue interaction network identified through network analysis (CS-VIP 8: green, RbBP5 peptide: tan, WDR5: grey, main communication path, initiated by arginine binding, crossing through central pore is depicted in yellow circles, while additional pathway across WDR5 surface is depicted in orange circles). **e** Model of the mode of action of the WDR5-binding inhibitors within the whole complex, where grey boxes highlight altered interaction interfaces between proteins upon CS-VIP 8 binding.

interface, and effectors around its MLL1 binding site as in the central pore of WDR5 (Fig. S43, SN3). Thereby, we imply a direct allosteric communication between MLL1 and RbBP5 binding sites via WDR5, which further indicates the importance of WDR5 as mediator of RbBP5/MLL1 assembly in functional MLL1 complexes³⁹. Using graph-theory based residue interaction network analysis on conformational ensembles obtained by MD simulations, communication pathways between MLL1 and RbBP5 were studied. Betweenness centrality (BC) analysis of obtained networks (Figs. 5c,d, S47, SN4) revealed that WDR5

residues D92 and K52 at the MLL1 binding site, as well as RbBP5 binding site residue K221 are major contributors to the overall information flow along all derived networks (Figs. 5c,d). Allosteric communication of these residues with the RbBP5 binding site through the central pore of WDR5 represents apparently as the main cause of partial RbBP5 unbinding. Additionally, residues around Y191 were also identified to contribute for RbBP5 stabilisation (Fig. 5d, SN4, Table S7), which is in accordance to the observed increased D-uptake of WDR5 residues 195-205 upon *cis*-CS-VIP 8 binding as well as a reported mutational analysis,

in which Y191F was identified as a cause of low MLL1 activity⁴⁵. All together, the observed activity loss upon *cis*-CS-VIP 8 addition as yielded by allosteric rearrangements of the WDR5-RbBP5 interface apparently destabilizes the closed SET1 conformation, highlighting its importance for MLL1 activity and uncovering the mode of action of the WDR5-binding inhibitors within the whole multiprotein MLL1 complex (Fig. 5e).

***In vivo* optochemical control of hematopoiesis.**

Encouraged by the on-target functional window we obtained for the *cis:trans* isomers of CS-VIP 8 and the severe disruption of MLL1 complex, we evaluated the potential of our conformationally strained visible-light photoswitch for *in vivo* photomodulation of hematopoiesis. Genetic programs of hematopoiesis are highly conserved across vertebrates⁴⁶. Therefore, we considered zebrafish as best model organism for our purpose because it is not only optically transparent, but also encodes most of the functional domains of human *ml1* gene products⁴⁷. Furthermore, it is known that MLL1 depletion in zebrafish causes severe defects related to hematopoiesis, such as lack of blood flow⁴⁸, or acute developmental abnormalities⁴⁹.

Once CS-VIP 8 isomers reached their corresponding PSS in pure DMSO, each was diluted with egg water to a final concentration of 500 μ M (1.5 % DMSO). These solutions were directly added to 1-phenyl-2-thiourea (PTU)-treated 3 day post-fertilization (dpf) larvae and incubated for 18 hours in darkness at 28.5 °C. Strikingly, only the exposure of the *cis*-CS-VIP 8 as obtained from the 520 nm PSS caused a complete lack of responsivity upon tail touch in the larvae, compared with untreated fish.

Furthermore, under these experimental conditions, abnormal developmental phenotypes such as curved body axis and heart edema were observed. Remarkably, larvae incubated with either 1.5 % DMSO or CS-VIP 8 from 405 nm PSS (mainly *trans* state) behaved as untreated ones, i.e., they were all alive with fast response to touch and without any apparent abnormalities. To determine whether the observed physiological changes due to *cis*-CS-VIP 8 exposure were related to hematopoiesis, we stained larvae with *o*-dianisidine for hemoglobin detection⁵⁰. The staining pattern remained unaffected for the vehicle, i.e., 1.5 % DMSO as well as for 97 % of larvae incubated with CS-VIP 8 at the 405 nm PSS. In contrast, staining of larvae treated with *cis*-CS-VIP 8 at 520 nm PSS, was highly diminished (60 %) or even absent (40 %, Fig. 6a,b). These reduced-blood flow and abnormal phenotypes are in agreement with reported antisense knockdowns of *ml1*⁴⁸. Together with our structural data, which display specific MLL1 complex disruption, these functional data strongly suggest that only *cis*-CS-VIP 8, can affect hematopoiesis *in vivo*.

Finally, to release the full potential of our photo-responsive probes, we explored the possibility to externally control the *in vivo* inhibition of hematopoiesis via on-demand photoactivation. Thus, larvae incubated with the innocuous *trans*-CS-VIP 8 for one hour were irradiated *in situ* at 520 nm for 30 sec. After 18 h incubation, we detected effects comparable to the positive control, where larvae were exposed to *cis*-CS-VIP 8. Precisely, only 3% of larvae displayed an intact *o*-dianisidine staining pattern while in the remainder, labelling was either partially disrupted (30 %) or completely affected (67 %). Negative controls

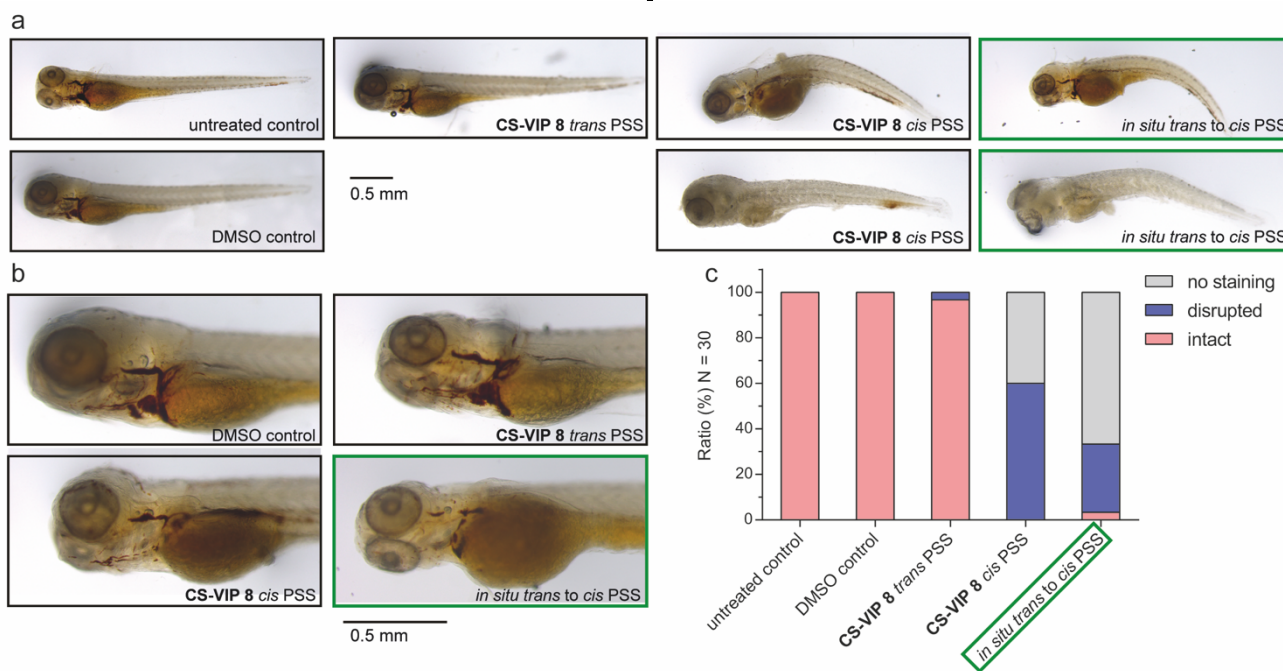


Figure 6 | *In vivo* optochemical inhibition of hematopoiesis in 3 days post fertilization (dpf) zebrafish larvae. a, Microscopy images of 3-dpf larvae incubated with 500 μ M of CS-VIP 8 at each PSS and negative controls. CS-VIP 8 *trans* PSS incubated larvae develop normally without green light exposure, while *in situ* irradiation of those zebrafish at 520 nm results in reduced blood flow analogue to CS-VIP 8 *cis* PSS incubated larvae. **b,** Magnified micrographs of the heart region. **c,** Bar chart comparing the blood flow of larvae incubated with CS-VIP 8 at the respective PSS and controls via graphical analysis of *o*-dianisidine staining patterns. Mean of three independent measurements.

of untreated fish or containing only the vehicle (1.5 % DMSO) verified the harmlessness of our visible-light irradiation (Fig. S24). Taken together, for the first time, we developed a light-controlled enzyme modulator capable to achieve *in vivo* optochemical control of hematopoiesis.

Discussion

Our understanding of gene regulatory mechanisms at the molecular level is often limited due to the lack of *in vivo* structure-function studies, capable of coping with the redundancy and ubiquitous expression of chromatin-modifying complexes. In this context, our work provides a successful proof-of-concept of *in vivo* photopharmacology to gain insight into the hematopoietic function of the MLL1 complex. For the first time, we introduced a family of Conformationally Strained Visible-light Photoswitches, named as CS-VIPs, that cannot only tightly interact with WDR5 as target, but also light-control *in vivo* hematopoiesis in zebrafish without additional genetic modifications. Our exhaustive molecular studies provide a detailed view on the action of MLL1-WDR5 inhibitors in the context of the intact MLL1 complex. Upon **CS-VIP 8** binding to WDR5, MLL1 dissociates from the core-complex, while the remaining complex subunits undergo conformational changes, e.g. by rewiring residues interaction networks in the WDR5-RbBP5 binding interface. Although the specific goal of this project was to unravel MLL1-related hematopoiesis, we expect our CS-VIPs to find further applications in diverse areas where *in situ* temporal control of MLL1 complex architecture is required or in which, β -propeller proteins, such as WDR5, act as interactions hubs⁵¹. Our methodology goes beyond the traditional reductionist mindset for dissecting biological systems, especially by providing new tools for analyzing the assembly of complex biological machineries. Our workflow can be readily expanded for other drugable targets and will further increase the scare toolbox of light-controlled enzyme modulators^{33,52-54}. Finally, both the implication of epigenetics in the development of hematopoietic malignancies and the appearance of epigenetic modifiers in clinical trials provides a perfect timeline for exploiting the modulation of hematopoiesis via epigenetic networks.

References

- 1 Laurenti, E. & Gottgens, B. From haematopoietic stem cells to complex differentiation landscapes. *Nature* **553**, 418-426, doi:10.1038/nature25022 (2018).
- 2 Antoniani, C., Romano, O. & Miccio, A. Concise Review: Epigenetic Regulation of Hematopoiesis: Biological Insights and Therapeutic Applications. *Stem Cells Transl Med* **6**, 2106-2114, doi:10.1002/sctm.17-0192 (2017).
- 3 Chung, Y. R., Schatoff, E. & Abdel-Wahab, O. Epigenetic alterations in hematopoietic malignancies. *Int J Hematol* **96**, 413-427, doi:10.1007/s12185-012-1181-z (2012).
- 4 Artinger, E. L. *et al.* An MLL-dependent network sustains hematopoiesis. *Proc Natl Acad Sci U S A* **110**, 12000-12005, doi:10.1073/pnas.1301278110 (2013).
- 5 McMahon, K. A. *et al.* Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell* **1**, 338-345, doi:10.1016/j.stem.2007.07.002 (2007).
- 6 Li, B. E. & Ernst, P. Two decades of leukemia oncoprotein epistasis: the MLL1 paradigm for epigenetic deregulation in leukemia. *Exp Hematol* **42**, 995-1012, doi:10.1016/j.exphem.2014.09.006 (2014).
- 7 Gan, T., Jude, C. D., Zaffuto, K. & Ernst, P. Developmentally induced Mll1 loss reveals defects in postnatal haematopoiesis. *Leukemia* **24**, 1732-1741, doi:10.1038/leu.2010.171 (2010).
- 8 Mishra, B. P. *et al.* The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis. *Cell Rep* **7**, 1239-1247, doi:10.1016/j.celrep.2014.04.015 (2014).
- 9 Luscher-Firzlaff, J. *et al.* Hematopoietic stem and progenitor cell proliferation and differentiation requires the trithorax protein Ash2l. *Sci Rep* **9**, 8262, doi:10.1038/s41598-019-44720-3 (2019).
- 10 Ankenbruck, N., Courtney, T., Naro, Y. & Deiters, A. Optochemical Control of Biological Processes in Cells and Animals. *Angew Chem Int Ed Engl* **57**, 2768-2798, doi:10.1002/anie.201700171 (2018).
- 11 Hull, K., Morstein, J. & Trauner, D. In Vivo Photopharmacology. *Chem Rev* **118**, 10710-10747, doi:10.1021/acs.chemrev.8b00037 (2018).
- 12 Liu, X. & Tonegawa, S. Optogenetics 3.0. *Cell* **141**, 22-24, doi:10.1016/j.cell.2010.03.019 (2010).
- 13 Gautier, A. *et al.* How to control proteins with light in living systems. *Nat Chem Biol* **10**, 533-541, doi:10.1038/nchembio.1534 (2014).
- 14 Fenno, L., Yizhar, O. & Deisseroth, K. The development and application of optogenetics. *Annu Rev Neurosci* **34**, 389-412, doi:10.1146/annurev-neuro-061010-113817 (2011).
- 15 Deisseroth, K. Controlling the brain with light. *Sci Am* **303**, 48-55, doi:10.1038/scientificamerican1110-48 (2010).
- 16 Velema, W. A., Szymanski, W. & Feringa, B. L. Photopharmacology: beyond proof of principle. *J Am Chem Soc* **136**, 2178-2191, doi:10.1021/ja413063e (2014).
- 17 Borowiak, M. *et al.* Optical Manipulation of F-Actin with Photoswitchable Small Molecules. *J Am Chem Soc* **142**, 9240-9249, doi:10.1021/jacs.9b12898 (2020).
- 18 Mayer, G. & Heckel, A. Biologically active molecules with a "light switch". *Angew Chem Int Ed Engl* **45**, 4900-4921, doi:10.1002/anie.200600387 (2006).
- 19 Szymanski, W., Beierle, J. M., Kistemaker, H. A., Velema, W. A. & Feringa, B. L. Reversible photocontrol of biological systems by the incorporation of molecular photoswitches. *Chem Rev* **113**, 6114-6178, doi:10.1021/cr300179f (2013).
- 20 Polosukhina, A. *et al.* Photochemical restoration of visual responses in blind mice. *Neuron* **75**, 271-282, doi:10.1016/j.neuron.2012.05.022 (2012).
- 21 Pittolo, S. *et al.* An allosteric modulator to control endogenous G protein-coupled receptors with light. *Nat Chem Biol* **10**, 813-815, doi:10.1038/nchembio.1612 (2014).
- 22 Borowiak, M. *et al.* Photoswitchable Inhibitors of Microtubule Dynamics Optically Control Mitosis and Cell Death. *Cell* **162**, 403-411, doi:10.1016/j.cell.2015.06.049 (2015).
- 23 Broichhagen, J. & Trauner, D. The *in vivo* chemistry of photoswitched tethered ligands. *Curr Opin Chem Biol* **21**, 121-127, doi:10.1016/j.cbpa.2014.07.008 (2014).
- 24 Albert, L. *et al.* Controlled inhibition of methyltransferases using photoswitchable peptidomimetics: towards an epigenetic regulation of leukemia. *Chem Sci* **8**, 4612-4618, doi:10.1039/c7sc00137a (2017).
- 25 Cao, F. *et al.* Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia. *Mol Cell* **53**, 247-261, doi:10.1016/j.molcel.2013.12.001 (2014).
- 26 Bleger, D., Schwarz, J., Brouwer, A. M. & Hecht, S. o-Fluoroazobenzenes as readily synthesized photoswitches offering nearly quantitative two-way isomerization with visible light. *J Am Chem Soc* **134**, 20597-20600, doi:10.1021/ja310323y (2012).
- 27 Liang, C., Behnam, M. A. M., Sundermann, T. R. & Klein, C. D. Phenylglycine racemization in Fmoc-based solid-phase peptide synthesis: Stereochemical stability is achieved by choice of reaction conditions. *Tetrahedron Lett* **58**, 2325-2329, doi:10.1016/j.tetlet.2017.04.047 (2017).

- 28 Albert, L. *et al.* Modulating Protein-Protein Interactions with Visible-Light-Responsive Peptide Backbone Photoswitches. *Chembiochem* **20**, 1417-1429, doi:10.1002/cbic.201800737 (2019).
- 29 Schutt, M. *et al.* Photocontrol of cell adhesion processes: model studies with cyclic azobenzene-RGD peptides. *Chem Biol* **10**, 487-490, doi:10.1016/s1074-5521(03)00128-5 (2003).
- 30 Renner, C., Kusebauch, U., Loweneck, M., Milbradt, A. G. & Moroder, L. Azobenzene as photoresponsive conformational switch in cyclic peptides. *J Pept Res* **65**, 4-14, doi:10.1111/j.1399-3011.2004.00203.x (2005).
- 31 Zhang, L., Linden, G. & Vazquez, O. In search of visible-light photoresponsive peptide nucleic acids (PNAs) for reversible control of DNA hybridization. *Beilstein J Org Chem* **15**, 2500-2508, doi:10.3762/bjoc.15.243 (2019).
- 32 Knie, C. *et al.* ortho-Fluoroazobenzenes: visible light switches with very long-Lived Z isomers. *Chemistry* **20**, 16492-16501, doi:10.1002/chem.201404649 (2014).
- 33 Aggarwal, K. *et al.* Visible Light Mediated Bidirectional Control over Carbonic Anhydrase Activity in Cells and in Vivo Using Azobenzenesulfonamides. *J Am Chem Soc* **142**, 14522-14531, doi:10.1021/jacs.0c05383 (2020).
- 34 Karatas, H. *et al.* High-affinity, small-molecule peptidomimetic inhibitors of MLL1/WDR5 protein-protein interaction. *J Am Chem Soc* **135**, 669-682, doi:10.1021/ja306028q (2013).
- 35 Wang, Z. H., Li, D. D., Chen, W. L., You, Q. D. & Guo, X. K. Targeting protein-protein interaction between MLL1 and reciprocal proteins for leukemia therapy. *Bioorg Med Chem* **26**, 356-365, doi:10.1016/j.bmc.2017.11.045 (2018).
- 36 Morris, M. C., Depollier, J., Mery, J., Heitz, F. & Divita, G. A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nat Biotechnol* **19**, 1173-1176, doi:10.1038/nbt1201-1173 (2001).
- 37 Schober, T. *et al.* Controlling the Uptake of Diarylethene-Based Cell-Penetrating Peptides into Cells Using Light. *Chemphotochem* **3**, 384-391, doi:10.1002/cptc.201900019 (2019).
- 38 Yang, Z. *et al.* The DPY30 subunit in SET1/MLL complexes regulates the proliferation and differentiation of hematopoietic progenitor cells. *Blood* **124**, 2025-2033, doi:10.1182/blood-2014-01-549220 (2014).
- 39 Li, Y. *et al.* Structural basis for activity regulation of MLL family methyltransferases. *Nature* **530**, 447-452, doi:10.1038/nature16952 (2016).
- 40 Patel, A., Dharmarajan, V. & Cosgrove, M. S. Structure of WDR5 bound to mixed lineage leukemia protein-1 peptide. *J Biol Chem* **283**, 32158-32161, doi:10.1074/jbc.C800164200 (2008).
- 41 Odho, Z., Southall, S. M. & Wilson, J. R. Characterization of a novel WDR5-binding site that recruits RbBP5 through a conserved motif to enhance methylation of histone H3 lysine 4 by mixed lineage leukemia protein-1. *J Biol Chem* **285**, 32967-32976, doi:10.1074/jbc.M110.159921 (2010).
- 42 Avdic, V. *et al.* Structural and biochemical insights into MLL1 core complex assembly. *Structure* **19**, 101-108, doi:10.1016/j.str.2010.09.022 (2011).
- 43 Kaustov, L. *et al.* The MLL1 trimeric catalytic complex is a dynamic conformational ensemble stabilized by multiple weak interactions. *Nucleic Acids Res* **47**, 9433-9447, doi:10.1093/nar/gkz697 (2019).
- 44 Xue, H. *et al.* Structural basis of nucleosome recognition and modification by MLL methyltransferases. *Nature* **573**, 445-449, doi:10.1038/s41586-019-1528-1 (2019).
- 45 Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* **13**, 713-719, doi:10.1038/nsmb1128 (2006).
- 46 Lu, J. W. *et al.* Zebrafish as a Model for the Study of Human Myeloid Malignancies. *Biomed Res Int* **2015**, 641475, doi:10.1155/2015/641475 (2015).
- 47 Robinson, B. W. *et al.* mll ortholog containing functional domains of human MLL is expressed throughout the zebrafish lifespan and in haematopoietic tissues. *Br J Haematol* **152**, 307-321, doi:10.1111/j.1365-2141.2010.08398.x (2011).
- 48 Wan, X., Hu, B., Liu, J. X., Feng, X. & Xiao, W. Zebrafish mll gene is essential for hematopoiesis. *J Biol Chem* **286**, 33345-33357, doi:10.1074/jbc.M111.253252 (2011).
- 49 Kim, J. D. *et al.* Proper Activity of Histone H3 Lysine 4 (H3K4) Methyltransferase Is Required for Morphogenesis during Zebrafish Cardiogenesis. *Mol Cells* **38**, 580-586, doi:10.14348/molcells.2015.0053 (2015).
- 50 Paffett-Lugassy, N. N. & Zon, L. I. Analysis of hematopoietic development in the zebrafish. *Methods Mol Med* **105**, 171-198, doi:10.1385/1-59259-826-9:171 (2005).
- 51 Schapira, M., Tyers, M., Torrent, M. & Arrowsmith, C. H. WD40 repeat domain proteins: a novel target class? *Nat Rev Drug Discov* **16**, 773-786, doi:10.1038/nrd.2017.179 (2017).
- 52 Hoppmann, C. *et al.* Photocontrol of contracting muscle fibers. *Angew Chem Int Ed Engl* **50**, 7699-7702, doi:10.1002/anie.201101398 (2011).
- 53 Shankar, S. *et al.* Light-controlled switching of the spin state of iron(III). *Nat Commun* **9**, 4750, doi:10.1038/s41467-018-07023-1 (2018).
- 54 Krauss, U., Lee, J., Benkovic, S. J. & Jaeger, K. E. LOVely enzymes - towards engineering light-controllable biocatalysts. *Microb Biotechnol* **3**, 15-23, doi:10.1111/j.1751-7915.2009.00140.x (2010).

Acknowledgements

We thank Prof. Dr. E. Meggers for access to cell laboratory, Dr. V. Srinivasan for initial crystallization attempts, N. Frommknecht for tailored LED lamps from Marburg University (UMR); P. Erramsetti and Dr. Y. Wang for assisting with MOLM-13 cell line from Zentrum für Tumor und Immun-biologie (UMR). We acknowledge the DFG for co-financing the AccuTOF GCv 4G (JEOL) Time of Flight (TOF) mass spectrometer (INST 160/622-1 FUGG) as well as the support of the Institute of Molecular Biology (IMB, Mainz, Germany) Microscopy Core Facility. Numerical simulations were run on the PARADOX-IV supercomputing facility at the Scientific Computing Laboratory, National Center of Excellence for the Study of Complex Systems, Institute of Physics Belgrade, supported in part by the Ministry of Education, Science, and Technological Development of the Republic of Serbia under project No. ON171017. ND, DR and KN acknowledge project of Ministry of Science and Technological Development of the Republic of Serbia for Faculty of Pharmacy, University of Belgrade, No. 451-03-68/2020-14/200161. A. R. acknowledges the /ic/OS platform of the Grenoble Instruct-ERIC center (ISBG; UMS 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-05-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). Finally, this work was financially supported by the DFG programs: SPP1926 'Next Generation Optogenetics': Young Investigator program (grant # GO1011/11-1) and TRR81: 'Chromatin Changes in Differentiation and Malignancies' (TRR81/3, Z04).

Author contributions

L. A. synthesized the CS-VIPs, characterized the photoisomerization, expressed all proteins, conducted the FP-, HMT-, pull-down *in vivo* and cell-based assays, HDX interpretation; L. W. and L.-O. E. solved crystal structures;

W. S. performed HDX analysis and interpretation; G. B. performed HDX interpretation; J. N. designed, performed and interpreted *in vivo* experiments; S. R. designed and interpreted *in vivo* experiments; N. D., D. R. and K. N. designed, conducted and interpreted MD simulations; L.-O. E. and A. R. performed the UV-vis measurements of co-crystals at the cryobench; M. H. expressed WDR5, synthesized the *o*F₄Azo used for initially evaluated CS-VIPs library, and conducted preliminary co-crystal experiments, O. V. conceived of the idea for the study, designed, directed and interpreted experiments, and has the role of corresponding author.⁵⁵ L. A., W. S., K. N., L.-O. E., and O. V. wrote the manuscript. All authors commented and approved the manuscript.

Competing interests

The authors declare no competing interests.

Methods

General procedure for cyclization of linear precursor peptides to CS-VIPs. Linear precursor peptides were synthesized according to the standard Fmoc-SPPS using Oxyma/DIC as coupling agents and Cl-Trt resin. Mtt-*o*F₄Azo was on-resin incorporated under standard coupling conditions but using NMP instead of DMF as solvent. For cyclic peptides CS-VIP 1 to CS-VIP 3, the Mtt-*o*F₄Azo was the last amino acid added, for cyclic peptides CS-VIP 4 to CS-VIP 14, the Mtt-*o*F₄Azo was incorporated on the side chain of the corresponding Alloc-protected amino acid, after orthogonal deprotection. In both cases capping step was omitted. The final linear precursor peptides were cleaved from the resin with HFIP/DCM (1:4) for 3 h, and purified using a preparative RP-HPLC (eluent A: H₂O + 0.1% TFA, eluent B: MeCN + 0.1% TFA, gradient of 5-95% B). Intramolecular cyclization was performed under irradiation. The corresponding peptide (1.00 eq.) was dissolved in DMF ([peptide] = 75 mM), and was irradiated at 520 nm for 3 min. Afterwards, it was diluted with THF to an approximated 1.5 mM solution in THF/DMF (98:2). HATU (3.00 eq.) and DIPEA (4.00 eq.), were dissolved in THF containing 2% DMF and added to the peptide solution under 520 nm irradiation reaching a final peptide concentration of 1 mM. The mixture was stirred for 45 min in total darkness, until complete conversion was confirmed via analytical HPLC (gradient: 5-95% B). The solvent was removed via evaporation at 28 °C using a rotary evaporator. The crude was purified via preparative RP-HPLC (gradient of 15-95% B). After lyophilisation, the protecting group of arginine was removed. In the cases where Pbf was the protecting group (all cyclic peptides except CS-VIP 8), the peptide was treated with TFA/H₂O (90:10) for 2 h (535 µL / 1 mg peptide). In the case of CS-VIP 8, Boc was the protecting group and, consequently, 4 M HCl in dioxane for 3 h was used for deprotection (535 µL / 1 mg peptide). Complete conversion was ensured via analytical HPLC. Final purification via preparative RP-HPLC (gradient of 5-95% MeCN) yielded the CS-VIPs as orange solids. Analytical characterization data are detailed in the supporting information (Supplementary Note: Peptide Characterization).

Photoisomerization of CS-VIPs. For the isomerization studies an Emitter 405 nm SMD Roschwege Star-UV405-03-00-00, 3 W (to reach the *trans* PSS) and a HighPower-LED Green 87 lm 130° 3.8 V, 1000 mA Roschwege (to reach the *cis* PSS) were used at r.t. For the latter, an additional cut-on filter (OG515 coloured glass filter, 515 nm longpass, Thorlabs) was applied. Custom made 24-well LED arrays, composed of 24 LED's of either 405 nm or 520 nm wavelengths (same LED's + cut-on filter for the green LEDs, as above), were utilized for the irradiation of 96-well clear bottom plates (used for fluorescence polarization assays with *in situ* irradiation), as well as for clear 24-well plates utilized for *in vivo* zebrafish experiments.

Co-crystallization of WDR5•CS-VIP 8. An aliquot of purified frozen WDR5 protein (193 µM in 25 mM Tris, pH = 8.0, 150 mM NaCl) was thawed on ice and mixed with a solution of *o*F₄Azo CS-VIP 8 at each PSS (37.8 mM in DMSO yielding final concentrations of 189 µM WDR5 and

756 µM CS-VIP 8 (ratio: 1:4 protein/peptide). Crystallization was performed in SWISSCI MRC two-well crystallization plates (Jena Bioscience). The reservoir solution volume was 50 µL and the drops contained 1 µL of a 1:1 mixture of protein-peptide complex and the crystallisation solution. Crystal growth occurred at 4 °C in the Morpheus[®] G2 condition from Molecular Dimensions crystallization screen, which contains 0.1 M carboxylic acids (0.1 M sodium formate; 0.1 M ammonium acetate; 0.1 M sodium citrate tribasic dihydrate; 0.1 M potassium sodium tartrate tetrahydrate; 0.1 M sodium oxamate), 0.1 M buffer system 1, pH 6.5 (imidazole; MES monohydrate (2-(N-morpholine)ethansulfoic acid)), 30% v/v precipitant mix 2 (20% v/v ethylene glycol; 10% w/v PEG 8000). Crystallization plates were stored in total darkness, crystals were obtained after two weeks and flash-frozen in liquid nitrogen without additional cryoprotectant. Diffraction data of the crystals obtained were collected without any further irradiation of the crystal, once it was formed and named: "*cis2*" (PDB: 7AXP) and "*cis1*" (PDB: 7AXS). In an attempt to obtain the structure of *trans*-CS-VIP 8 bound to WDR5, the crystallization plate, was irradiated at 405 nm for 5 sec after crystal formation, but before freezing the crystal and named "*cis1*, 405 nm, *in situ*" (PDB: 7AXU). Finally, the obtained WDR5•CS-VIP 8 crystals were irradiated in their isolated frozen states at different temperatures and for different time periods at the cryobench and named:

"*cis1*, 405 nm/180 K" was irradiated at a temperature gradient from 180 (PDB: 7AXQ) to 293 K for 10, 15 and 45 sec at 405 nm, followed by back irradiation at 520 nm for 30 and 90 sec, followed by, again, back irradiation at 405 nm. The last crystal was irradiated at 293 K for 5 and 10 sec at 405 nm and is referred to as "*cis1*, 405 nm/293 K" (PDB: 7AXX). UV-vis absorbance spectra have been, as well, recorded at the cryobench after each respective irradiation point. Final refinement statistics are given in Tables S3.

UV-vis measurements of WDR5•CS-VIP 8 co-crystals. UV-vis absorption spectra of protein co-crystals were collected offline on the goniometer of the /ic/OS Lab at the ESRF (von Stetten et al., 2015). Briefly, the experimental setup is composed of two reflective mirror objectives of 4x magnification placed 180° from each other compared to the sample position. A 200 µm optical fibre connecting a balanced deuterium-halogen lamp (Mikropack DH-2000-BAL, Ocean Optics) and the illuminating objective results in a 50 µm focal spot on the crystal. A 400 µm optical fibre connects the collecting objective and a fixed-grating spectrophotometer equipped with a CCD detector (QE65Pro, Ocean Optics). Crystals were flashcooled in the evaporating nitrogen flux of a cryostream (Cryostream 700, Oxford Cryosystems) at 100 K. Photoswitching was triggered using a 405 nm LED (Thorlabs) connected to the actinic objective (90° from the two other objectives) via a 1000 µm optical fibre. Photoswitching experiments were also performed at room temperature using a humidity controller.⁵⁶

Competitive fluorescence polarization (FP)-based assays. FP-based assays were conducted as we described previously.²⁴ The ones, where isomerization of CS-VIP 8 was performed in presence of WDR5, implemented some small adjustments. Thus, clear bottom 96-well microtiter plates were used, to enable *in situ* irradiation. The assay for each PSS of CS-VIP 8 was set up as the previous described procedure but a separate plate was utilized for each individual state. The plate containing CS-VIP 8 in its *trans*-PSS and in its *cis*-PSS are named as *plate A* and *plate B*, respectively. After 3 h incubation of CS-VIP 8 with tracer peptide bound to WDR5, initial mP values were recorded (isomer PSS t0 in Fig. S20) time points of the *trans* and *cis* PSSs). Then, both plates, were irradiated at the respective wavelength for *in situ* isomerization. Precisely, plate A containing CS-VIP 8 at *trans* PSS was irradiated at 520 nm for 30 sec, while the plate B, containing CS-VIP 8 at *cis* PSS was irradiated at 405 nm for 5 sec to trigger isomerization. After irradiation, the plates were incubated for another hour, before mP values were, again, recorded. After this second measurement, the plates were again irradiated, to induce the back isomerization to their initial PSSs. Specifically, plate A was irradiated at 405 nm, while plate B was irradiated at 520 nm as detailed above. After another hour of incubation, mP values were, for a third time, recorded. Evaluation of the data was performed as detailed previously.

Hydrogen-deuterium exchange mass spectrometry (HDX-MS). Purified individual proteins or the MLL1 complex (50 µM) were incubated with *cis*-

CS-VIP 8 (150 μM) in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.2% (v/v) DMSO for 30 min on ice. The same volume of DMSO was employed in samples not containing *cis*-CS-VIP 8. Sample generation was aided by an autosampler (LEAP Technologies). 7.5 μL protein solution was mixed with 67.5 μL of D_2O -containing buffer and incubated for 10/95/1,000/10,000 sec at 25 $^\circ\text{C}$. 55 μL of this reaction was added to 55 μL quench buffer (400 mM $\text{KH}_2\text{PO}_4/\text{H}_3\text{PO}_4$, 2 M guanidine-HCl, pH 2.2) kept at 1 $^\circ\text{C}$ and 95 μL of the resulting mixture injected into an ACQUITY UPLC M-class system with HDX technology (Waters).⁵⁷ Peptic peptides were generated by an online pepsin column and separated by reversed-phase HPLC followed by mass spectrometric analysis and quantification of deuterium incorporation, as described previously^{58,59} aided by the ProteinLynx Global Server (PLGS) and DynamX 3.0 softwares (Waters). Plots of hydrogen-deuterium exchange profiles were generated with MEMHDX.⁶⁰

Virtual docking (VD) / molecular dynamics (MD). Three-dimensional conformations of CS-VIPs were generated in ChemBio3D software by Hartree-Fock method with 6-31G basis set (CambridgeSoft Corporation, 2013. ChemBio3D Ultra, Version 13.0. Cambridge, MA, USA). (S)-arginine constrained docking simulations were performed in GOLD 5.8.1 software⁶¹, whereas the *cis*-CS-VIPs were docked in novel crystal structure of WDR5 (PDB: 7AXP), and *trans*-CS-VIPs were docked in MD simulated complex of *trans*-CS-VIP 8 into WDR5.

Plain MD simulations were performed on four different systems: System A: apo WDR5; System B: WDR5:*cis*-CS-VIP 8; System C: MLL1 peptide-WDR5-RbBP5 peptide; System D: *cis*-CS-VIP 8-WDR5-RbBP5 peptide. Photoisomerization of CS-VIP 8 was studied using a “diabatic switching potential” non-equilibrium MD simulations; b) well-tempered metadynamics (WT-metaD). MD simulations were performed using GROMACS 2016.6 software⁶². Initial coordinates were obtained from PDB: 7AXP, or from PDB: 3P4F for systems containing RbBP5. Studied systems were represented using AMBERff14SB⁶³ force field, GAFF2 force field⁶⁴ and TIP3P water model. Bonded parameters of azobenzene (including parameters for excited S1 state) were acquired from literature⁶⁵. Non-protein residue’s atomic charges were fitted using restrained electrostatic potentials (RESP/HF-6-31G*)⁶⁶. After neutralization and minimization (5000 steps of steepest descent), equilibration of systems was performed firstly in NVT ensemble (310 K, Nosé-Hoover thermostat) and then in NPT ensemble (1.0 bar, Parrinello-Rahman barostat). Position restraints were gradually removed (from 1000 kJ/mol nm⁻²) before production runs. Particle Mesh Ewald (PME) for long-range electrostatics and non-bonded cut-off of 10 Å were used together with SHAKE bonds constrain algorithm. Obtained trajectories were analysed using built-in Gromacs tools. Perturbation-response scanning (PRS) was performed using ProDy software⁶⁷. To identify allosteric pathways, gRINN software⁶⁸ was utilized and interaction networks were additionally analyzed using Cytoscape 3.7.2. DyNet plug-in and D_n -score^{69,70}. Diabatic switching potential MD simulations were performed according to adapted protocol described elsewhere⁷¹. WT-metaD simulations were performed using PLUMED 2.3.5 software ref patched with GROMACS 2016.6 and CNND dihedral angle as collective variable. The bias factor, the deposition rate, the initial hill height, and the hill width were set to 25, 0.5 ps⁻¹, 2.5 kJ·mol⁻¹, and 0.35 rad, respectively.

In vivo experiments using zebrafish model. For *in vivo* experiments, zebrafish embryos were incubated in egg water (3 g sea-salt/10 L ultrapure water) dosed with 0.0003% w/v of PTU. After hatching, the larvae were kept in this solution for three days in a 28.5 $^\circ\text{C}$ incubator until three days post fertilization (dpf) zebrafish larval stage. Afterwards, a 714.3 μM solution of CS-VIP 8 in egg water, containing 2.14% DMSO, was irradiated at 520 nm/405 nm wavelength to obtain either *cis* or *trans* PSSs, respectively. After that, the assay was performed in darkness using a red safety light. 700 μL of the solution of each PSS were added to a clear 24-well plate in duplicate. Then, 5 larvae in 300 μL egg water were added to each well leading to a final volume of 1 mL ([CS-VIP 8] = 500 μM with 1.5% DMSO). For each experiment, two controls without any addition of CS-VIP 8 were included in duplicates: a DMSO control and untreated control. The plate was then incubated at 28.5 $^\circ\text{C}$ for 18 h. For the experiments, where CS-VIP 8 was *in situ* isomerized, larvae were initially added to a solution of CS-VIP 8 in its *trans* PSS and incubated at 28.5 $^\circ\text{C}$ for 1 h. Afterwards, the whole plate was irradiated at 520 nm for 30 sec. Next, the plate was incubated at 28.5 $^\circ\text{C}$ for 17 h and subsequently

examined under a microscope (Zeiss Stemi DV4) to evaluate their response to perturbation. Afterwards, *o*-dianisidine staining was performed: a 5.85 mM *o*-dianisidine stock solution EtOH was prepared one day previously to its use. At the day of use, 2.0 mL of the stock solution were mixed with 500 μL of 0.1 M NaOAc pH = 4.5, 2.0 mL sterile water and 100 μL H_2O_2 to obtain the *o*-dianisidine staining solution. After 18 h incubation, the larvae solution was exchanged for 1.0 mL ultrapure water. Next, the water was removed from the wells. Immediately after, 1 mL of the *o*-dianisidine staining solution was added to each well and incubated in the dark at r.t. for 20 min. Then, the *o*-dianisidine staining solution was removed from the larvae, and they were, once again, washed with 1.0 mL ultrapure water. Afterwards, 1 mL of a 4% paraformaldehyde (PFA) solution in ultrapure water was added to the emptied wells to fix the larvae. For imaging, the PFA solution was removed, and larvae were washed with 1.0 mL PBS buffer followed by the addition of 1.0 mL 80% glycerol in ultrapure water. Microscopic images were acquired with a Leica M205 FA Fluorescence Stereo Microscope (settings: HDR average: automatic image formats: highest pixel: 3264 \times 2448; Input: DCF4955257; automatic exposure (80% brightness; complete larvae were 34.9 times zoomed in; heart region was 84.8 times zoomed in).

His-Sumo pull-down assay. 1 nmol (final concentration 0.262 μM) of His-SUMO-tagged RbBP5 in 200 μL assay buffer (25 mM Tris, 150 mM NaCl, 20 mM imidazole, pH = 8) were preincubated on a rotator with 20.0 μL of Ni-NTA beads (equilibrated with assay buffer) for 1 h at 4 $^\circ\text{C}$. After washing the beads three times with assay buffer, 2 nmol of the remaining proteins (WDR5, MLL1, Ash2L) in 200 μL assay buffer were added to the beads. After 5 min incubation on ice, *cis*-CS-VIP 8 at 200 μM in DMSO and the sample mixtures were incubated on the rotator at 4 $^\circ\text{C}$ for 1 h. As controls, one sample did not contain His-Sumo-RbBP5 whereas one contained DMSO instead of peptide. After 1 h, the beads were washed three times with the assay buffer containing 40 mM imidazole, followed by elution with 250 mM imidazole. Bound proteins were on 15% SDS-PAGE gels, following protocol of the *Lab FAQs* from Roche.

Cell viability assays. Suspension cells were cultured under standard conditions in RPMI Glutamax medium supplemented with FBS. For cell uptake, we used the transfection agent pep-1³⁶, which we synthesized (15, SI). Pep-1 was applied following the the Active Motif’s Chariot™ protocol with slight modifications. The whole assays were performed in darkness with a red safety light, which did not affect the isomerization ratio of the peptides.

Preparation of CS-VIP:pep-1 complex: The CS-VIP stock solution at the respective concentration in DMSO was irradiated at the specific wavelength to reach the desired PSS. Then, 1 μL of these stocks was further diluted by the addition of 49 μL of PBS buffer in an Eppendorf tube. In addition DMSO control tubes, which contained 1 μL of DMSO instead of a CS-VIP dilution (pep-1) were included. Preparation of the pep-1 stock solution was performed by dissolving 0.3 mg in 150 μL sterile ultrapure water. This stock solution was frozen, aliquoted, stored at -20 $^\circ\text{C}$ and thawed immediately before usage. For each planned pep-1 reaction, 2 μL of the 653 μM pep-1 stock solution were further diluted with 48 μL sterile ultrapure water. Then, 50 μL of such pep-1 solution were added to the 50 μL of each previously prepared CS-VIP concentration or DMSO, and mixed. In addition to the control pep-1 control, two more were included, without the addition of pep-1. One containing 1 μL DMSO in 49 μL PBS and 50 μL sterile ultrapure water (DMSO control w/o pep-1), and one containing neither DMSO nor pep-1, only 50 μL PBS and 50 μL ultrapure water (cells alone). The tubes were incubated on a shaker at r.t. for 30 min.

Preparation of the cells: The cultured cells were centrifuged (350 rcf, 25 $^\circ\text{C}$, 5 min) and the culturing media was exchanged for pure RPMI without any supplements. Cells were mixed with a 5 mL serological pipette 20 times, before they were counted. Then, 100 μL of 20×10^4 cells/mL in RPMI without supplements (final density after addition of media up to 1 mL volume: 2×10^4 cells/mL) were seeded into sterile clear non-binding 24-well plates.

Transfection of CS-VIPs into the cells: The 100 μL of incubated cargo: pep-1 complexes, including the controls, were added to the 100 μL of the seeded cells in the 24-well plates and mixed by pipetting 100 μL of the solutions up and down 10 times. Besides the cyclic peptide dilution wells, each plate also contained the control wells described above. After 5 min incubation on a shaker, the plate was incubated in the cell culture incubator at 37 $^\circ\text{C}$

for 1 h. After 1 h incubation, 800 μ L RPMI containing 3.125 % FBS (final FBS content : 2.5 %) were added to each well, and the plate was incubated in the cell culture incubator for 3 h, for experiments including the washing step. For experiments excluding the washing step, after 24 h incubation, FBS was added up to 10 % (for MOLM-13 cells) / 15 % (for MV4-11 cells), followed by another 48 h of incubation, before Cell Titer-Glo[®] read-out.

Washing of the cells after cargo transfection: After the 3 h incubation of the transfected cells, the wells of each respective condition were combined into one sterile falcon tube, respectively. Thus, the solution of each well was mixed by pipetting up and down 500 μ L of the respective solution 20 times in order to dispense the cells, before adding it to the falcon tube. Once all wells of each condition were combined, the empty wells were washed with 500 μ L of RPMI media (with 10 % FBS). The falcon tubes were centrifuged (300 rcf, 3 min, rt) and the supernatant was discarded. 500 μ L of fresh RPMI media (with 10 % FBS) were added to each falcon to re-dissolve the cell pellet by mixing 20 times. The solution was transferred into a 2 mL Eppendorf tube and the falcon was washed with another 500 μ L of fresh RPMI media (with 10 % FBS). After centrifugation (300 rcf, 3 min, rt) of the Eppendorf tubes, the supernatant was, again, discarded and the cell pellet re-dissolved in 800 μ L of RPMI (with 10 % FBS) by mixing 20 times. The cells of each condition were re-counted, followed by reseeding 1 mL of each condition into a new 24-well plate with a density of: 2×10^4 cells/mL. The plate was incubated at 37 °C in the cell incubator for 2 days (48 h), followed by Cell Titer-Glo[®] evaluation.

References

- 55 Credit where credit is due. *Nature* **462**, 825, doi:10.1038/462825a (2009).
- 56 von Stetten, D. *et al.* In crystallo optical spectroscopy (icOS) as a complementary tool on the macromolecular crystallography beamlines of the ESRF. *Acta Crystallogr D Biol Crystallogr* **71**, 15-26, doi:10.1107/S139900471401517X (2015).
- 57 Wales, T. E., Fadgen, K. E., Gerhardt, G. C. & Engen, J. R. High-speed and high-resolution UPLC separation at zero degrees Celsius. *Anal Chem* **80**, 6815-6820, doi:10.1021/ac8008862 (2008).
- 58 Skotnicka, D. *et al.* CdbA is a DNA-binding protein and c-di-GMP receptor important for nucleoid organization and segregation in *Myxococcus xanthus*. *Nat Commun* **11**, 1791, doi:10.1038/s41467-020-15628-8 (2020).
- 59 Osorio-Valeriano, M. *et al.* ParB-type DNA Segregation Proteins Are CTP-Dependent Molecular Switches. *Cell* **179**, 1512-1524 e1515, doi:10.1016/j.cell.2019.11.015 (2019).
- 60 Hourdel, V. *et al.* MEMHDX: an interactive tool to expedite the statistical validation and visualization of large HDX-MS datasets. *Bioinformatics* **32**, 3413-3419, doi:10.1093/bioinformatics/btw420 (2016).
- 61 Jones, G., Willett, P., Glen, R. C., Leach, A. R. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* **267**, 727-748, doi:10.1006/jmbi.1996.0897 (1997).
- 62 Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845-854, doi:10.1093/bioinformatics/btt055 (2013).
- 63 Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696-3713, doi:10.1021/acs.jctc.5b00255 (2015).
- 64 Vassetti, D., Pagliai, M. & Procacci, P. Assessment of GAFF2 and OPLS-AA General Force Fields in Combination with the Water Models TIP3P, SPCE, and OPC3 for the Solvation Free Energy of Druglike Organic Molecules. *J Chem Theory Comput* **15**, 1983-1995, doi:10.1021/acs.jctc.8b01039 (2019).
- 65 Duchstein, P., Neiss, C., Gorling, A. & Zahn, D. Molecular mechanics modeling of azobenzene-based photoswitches. *J Mol Model* **18**, 2479-2482, doi:10.1007/s00894-011-1270-6 (2012).
- 66 Woods, R. J. & Chappelle, R. Restrained electrostatic potential atomic partial charges for condensed-phase simulations of carbohydrates. *Theochem* **527**, 149-156, doi:10.1016/S0166-1280(00)00487-5 (2000).
- 67 Bakan, A., Meireles, L. M. & Bahar, I. ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **27**, 1575-1577, doi:10.1093/bioinformatics/btr168 (2011).
- 68 Sercinoglu, O. & Ozbek, P. gRINN: a tool for calculation of residue interaction energies and protein energy network analysis of molecular dynamics simulations. *Nucleic Acids Res* **46**, W554-W562, doi:10.1093/nar/gky381 (2018).
- 69 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).
- 70 Goenawan, I. H., Bryan, K. & Lynn, D. J. DyNet: visualization and analysis of dynamic molecular interaction networks. *Bioinformatics* **32**, 2713-2715, doi:10.1093/bioinformatics/btw187 (2016).
- 71 Biswas, M. & Burghardt, I. Azobenzene photoisomerization-induced destabilization of B-DNA. *Biophys J* **107**, 932-940, doi:10.1016/j.bpj.2014.06.044 (2014).

***In Vivo* Optochemical Control of Hematopoiesis by MLL1 Complex Disruption**

**Lea Albert,¹ Wieland Steinchen,^{1,2} Laura Werel,¹ Jatin Nagpal,³ Nemanja Djokovic,⁴
Dusan Ruzic,⁴ Malte Hoffarth,¹ Jing Xu,⁵ Antoine Royant,^{6,7} Gert Bange,^{1,2} Katarina
Nikolic,⁴ Soojin Ryu,^{3,8} Yali Dou,⁹ Lars-Oliver Essen,^{1,2} Olalla Vázquez^{1,2*}**

¹Department of Chemistry, University of Marburg, Germany.

²Center for Synthetic Microbiology (SYNMIKRO), University of Marburg, Germany,

³University Medical Center, Johannes Gutenberg University Mainz, Germany

⁴Department of Pharmaceutical Chemistry, University of Belgrade, Serbia.

⁵Department of Pathology, University of Michigan, USA.

⁶Institut de Biologie Structurale, University Grenoble Alpes, CEA, CNRS, France

⁷European Synchrotron Radiation Facility, Grenoble, France

⁸College of Medicine and Health, University of Exeter, UK.

⁹Norris Comprehensive Cancer Center, University of Southern California, USA.

e-mail: olalla.vazquez@staff.uni-marburg.de

Table of Contents

1 Abbreviations	3
1 Supplementary Note: Peptide Characterization	5
1.1 Cyclic peptides	6
1.2 Pep-1 carrier	12
2 Supplementary Figures and Tables	13
2.1 Characterization of isomerization of CS-VIP 8	13
2.2 Stability against glutathione (GSH)	14
2.3 Fluorescence polarization-based assays.....	14
2.4 Histone methyltransferase assay.....	16
2.5 Cell viability assays.....	17
2.6 <i>In vivo</i> zebrafish studies	17
2.7 Structures of WDR5: CS-VIP 8 complexes	18
2.8 Microspectroscopy of WDR5• CS-VIP 8 co-crystals	21
2.9 Hydrogen-deuterium exchange mass spectrometry (HDX-MS)	22
2.10 His-SUMO pull-down assay	29
2.12 Virtual docking and molecular dynamic studies	29
2.12.1 Supplementary note N1	33
2.12.2 Supplementary note N2	36
2.12.3 Supplementary note N3	37
2.12.4 Supplementary note N4	38
5 References	44

1 Abbreviations

λ	Wavelength
2-Cl-Trt	2-Chloro-trityl
Ash2L	Absent, small or homeotic 2-like protein
a.u.	Absorbance units
BC	Betweenness-centrality
Boc	<i>tert</i> -Butylcarbonyl
calcd.	Calculated
d	Days
DMSO	Dimethyl sulfoxide
ESI	Electrospray ionization
<i>o</i> F ₄ Azo	tetra- <i>ortho</i> -Fluorazobenzene
FP	Fluorescence polarization
GSH	Glutathione
H3	Histone 3
HDX	Hydrogen deuterium exchange
HFIP	Hexafluoroisopropanol
HMT	Histone methyltransferase
HPLC	High-performance liquid chromatography
HRMS	High resolution mass spectrometry
IC ₅₀	Half inhibitory concentration
irr.	Irradiation
K_i	Inhibition constant
MD	Molecular dynamics
MLL1	Mixed lineage leukaemia 1
mP	Milli polarization value
Pbf	Pentamethyldihydrobenzofuran-5-sulfonyl
PCA	Principal component analysis

Pep-1	Carrier peptide ¹
PRS	Perturbation-response scanning
PSS	Photostationary state
R_f	Retention factor
RbBP5	Retinoblasma-bindin-protein 5
SASA	Solvent accessible surface areas
SUMO	Small ubiquitin-related modifier
t_0	Initial time
t_R	Retention time
TFA	Trifluoroacetic acid
Tris	Tris(hydroxymethyl)aminomethane
Trt	Trityl
UV	Ultraviolet
vis	Visible
WDR5	WD-repeat protein-5
WT-metaD	Well-tempered metadynamics

1 Supplementary Note: Peptide Characterization

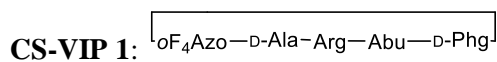
Table S1. Depending on the specific peptide, the cleavage / deprotection cocktail differed in the concentration of TFA or HFIP, type of scavengers and reaction times. The used cocktails and reaction times are listed in this table. For 40 mg of resin, 1 mL of cleavage cocktail was utilized.

Cocktail	Deprotection cocktail composition and application
Cleavage cocktail A	95% TFA, 5% CH ₂ Cl ₂ , 2.5% H ₂ O, 2.5% triisopropylsilane (3 h) It has been utilized to cleave the pep-1 peptide from the cysteamide resin, while at the same time deprotecting the side chain protecting groups.
Cleavage cocktail B	HFIP/CH ₂ Cl ₂ 1:4 (3 h) It has been utilized to cleave the linear precursor peptides from the Cl-Trt resin, while maintaining the protecting groups of the amino acid side chains.
Deprotection cocktail C	90% TFA, 10% H ₂ O (2 h) It has been utilized to deprotect the Pbf-group at the Arg-side chain of the final cyclized peptides in solution.
Deprotection cocktail D	4 M HCl in dioxane (3 h) It has been utilized to deprotect the Boc-group at the Arg-side chain of the final cyclized peptide CS-VIP 8 in solution.

Table S2. Columns for analytical HPLC-MS, which were recorded on an Agilent 1200 Series HPLC-system (Agilent Technologies), with a column oven set to 55 °C. Detection was monitored at 220 nm. Ultrapure water (A) and MeCN (B) were employed as eluents with an addition of 0.05% and 0.03% of TFA, respectively. Different analytical columns and gradients were used (the specific column and gradient used for each analytical HPLC run are noted at the respective chromatograms).

Column	Column specification
Column 1	Eclipse XBD-C18 column (Agilent Technologies) using the stated linear gradient in 30 min with a flow rate of 1.00 mL/min
Column 2	EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) using the stated linear gradient in 30 min with a flow rate of 0.20 mL/min
Column 3	Kinetex 5µm XB-C18 100 Å LC column 150 × 4.6 mm (Phenomenex) using the stated linear gradient in 30 min with a flow rate of 1.00 mL/min

1.1 Cyclic peptides



For the lineal precursor of **CS-VIP 1**, Arg(Pbf) was employed. Cyclization was performed in a 10.0 μmol scale as detailed above. After purification, the 1 \times TFA salt product (0.2 mg, 0.216 μmol , 2%) was obtained as an orange solid. **HPLC:** $t_R = 17.10$ min (*trans*); 18.32 min (*cis*). Purity = 97%. Formula: $\text{C}_{37}\text{H}_{41}\text{F}_4\text{N}_{11}\text{O}_6$. **HRMS-ESI⁺ (m/z):** $[\text{M}+1\text{H}]^+$ calcd.: 812.3256; found: 812.3247.

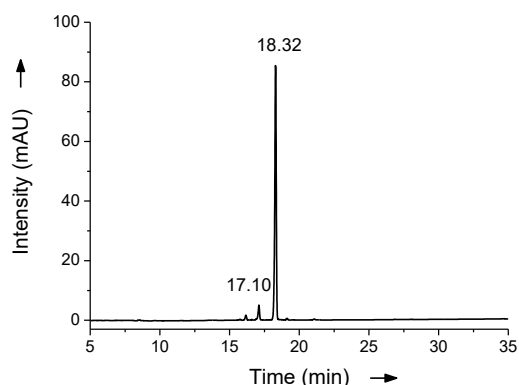
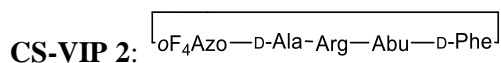


Figure S1. HPLC chromatogram of the purified peptide CS-VIP 1. Gradient: 5-95% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 2**, Arg(Pbf) was employed. Cyclization was performed in a 12.0 μmol scale as detailed above. After purification, the 1 \times TFA salt product (0.8 mg, 0.852 μmol , 7%) was obtained as an orange solid. **HPLC:** $t_R = 17.55$ min (*trans*); 19.11 min (*cis*). Purity = 95%. Formula: $\text{C}_{38}\text{H}_{43}\text{F}_4\text{N}_{11}\text{O}_6$. **HRMS-ESI⁺ (m/z):** $[\text{M}+1\text{H}]^+$ calcd.: 826.3407; found: 826.3410.

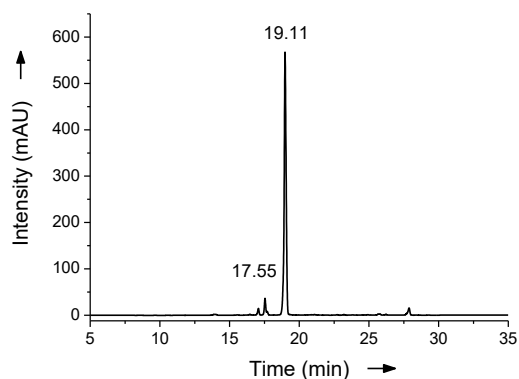
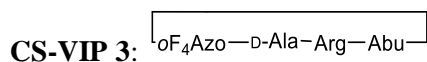


Figure S2. HPLC chromatogram of the purified peptide CS-VIP 2. Gradient: 5-95% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 3**, Arg(Pbf) was employed. Cyclization was performed in a 12.0 μmol scale as detailed above. After purification, the 1 \times TFA salt product (0.8 mg, 1.01 μmol , 8%) was obtained as orange solid. **HPLC:** $t_R = 14.46$ min (*trans*); 16.86 min (*cis*). Purity > 99%. Formula: $\text{C}_{29}\text{H}_{34}\text{F}_4\text{N}_{10}\text{O}_5$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 679.2728; found: 679.2719.

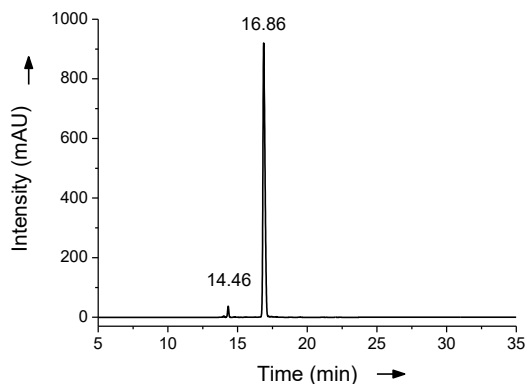
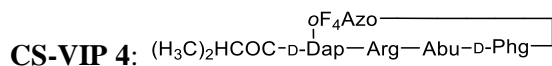


Figure S3. HPLC chromatogram of the purified peptide **CS-VIP 3**. Gradient: 5-95% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 4**, Arg(Pbf) was employed. Cyclization was performed in a 11.0 μmol scale as detailed above. After purification, the 1 \times TFA salt product (0.7 mg, 0.692 μmol , 6%) was obtained as an orange solid. **HPLC:** $t_R = 21.61$ min (*cis*). Purity = 98%. Formula: $\text{C}_{41}\text{H}_{48}\text{F}_4\text{N}_{12}\text{O}_7$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 897.3778; found: 897.3779.

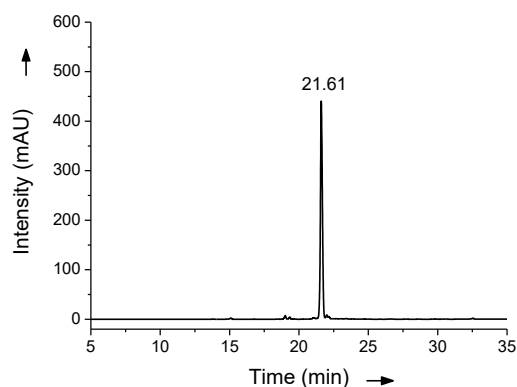
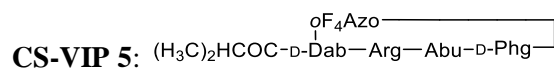


Figure S4. HPLC chromatogram of the purified peptide **CS-VIP 4**. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 5**, Arg(Pbf) was employed. Cyclization was performed in a 12.0 μmol scale as detailed above. After purification, the 1 \times TFA salt product (1.2 mg, 1.17 μmol , 10%) was obtained as an orange solid. **HPLC:** $t_R = 19.53$ min (*trans*); 21.65 min (*cis*). Purity = 92%. Formula: $\text{C}_{42}\text{H}_{50}\text{F}_4\text{N}_{12}\text{O}_7$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 911.3934; found: 911.3937.

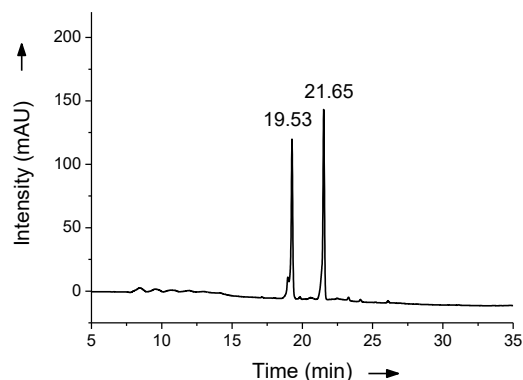
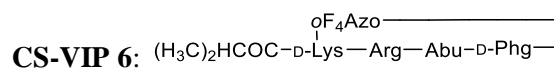


Figure S5. HPLC chromatogram of the purified peptide CS-VIP 5. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 6**, Arg(Pbf) was employed. Cyclization was performed in a 14.0 μmol scale as detailed above. After purification, the 1 \times TFA salt product (2.0 mg, 1.90 μmol , 14%) was obtained as an orange solid. **HPLC:** $t_R = 20.47$ min (*trans*); 22.14 min (*cis*). Purity = 95%. Formula: $\text{C}_{44}\text{H}_{54}\text{F}_4\text{N}_{12}\text{O}_7$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 939.4247; found: 939.4249.

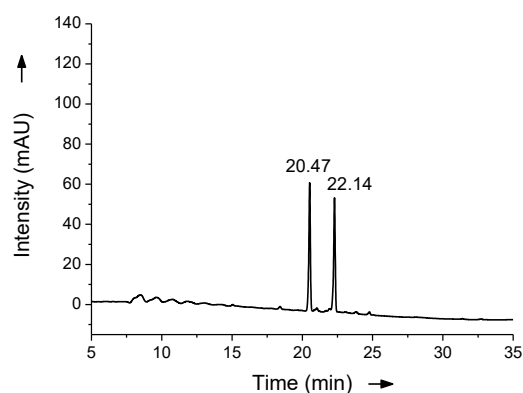
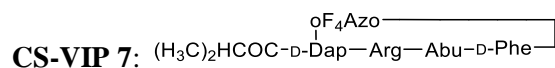


Figure S6. HPLC chromatogram of the purified peptide CS-VIP 6. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 7**, Arg(Pbf) was employed. Cyclization was performed in a 23.0 μmol scale as detailed above. After purification, the 1 \times TFA salt product (2.9 mg, 2.33 μmol , 12%) was obtained as an orange solid. **HPLC:** $t_R = 21.26$ min (*trans*); 22.25 min (*cis*). Purity = 94%. Formula: $\text{C}_{42}\text{H}_{50}\text{F}_4\text{N}_{12}\text{O}_7$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 911.3934; found: 911.3936.

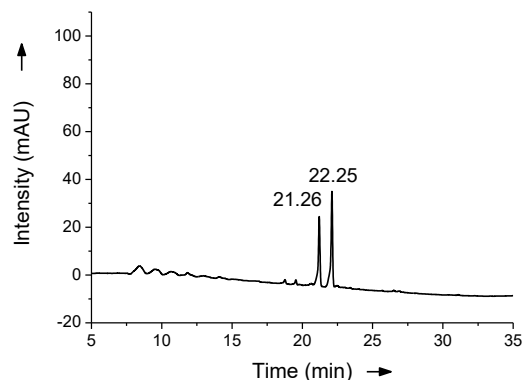
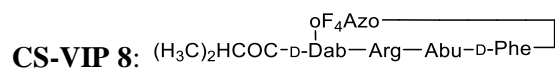


Figure S7. HPLC chromatogram of the purified peptide **CS-VIP 7**. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 8**, Arg(Boc)₂ was employed. Cyclization was performed in a 150 μmol scale as detailed above. After purification, the 1 \times TFA salt product (50.2 mg, 48.3 μmol , 32%) was obtained as an orange solid. **HPLC:** $t_R = 20.02$ min (*trans*); 23.18 min (*cis*). Purity > 99%. Formula: $\text{C}_{43}\text{H}_{52}\text{F}_4\text{N}_{12}\text{O}_7$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 925.4091; found: 925.4095.

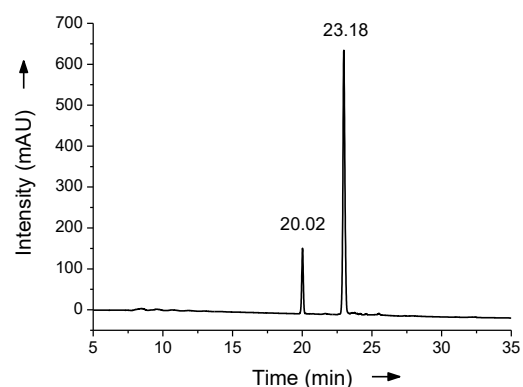
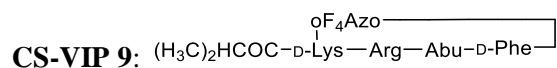


Figure S8. HPLC chromatogram of the purified peptide **CS-VIP 8**. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 9**, Arg(Boc)₂ was employed. Cyclization was performed in a 58.0 μmol scale as detailed above. After purification, the 1 × TFA salt product (16.6 mg, 15.6 μmol, 27%) was obtained as an orange solid. **HPLC:** $t_R = 20.84$ min (*trans*); 23.35 min (*cis*). Purity > 99%. Formula: C₄₅H₅₆F₄N₁₂O₇. **HRMS-ESI⁺ (m/z):** [M+H]⁺ calcd.: 953.4404; found: 953.4409.

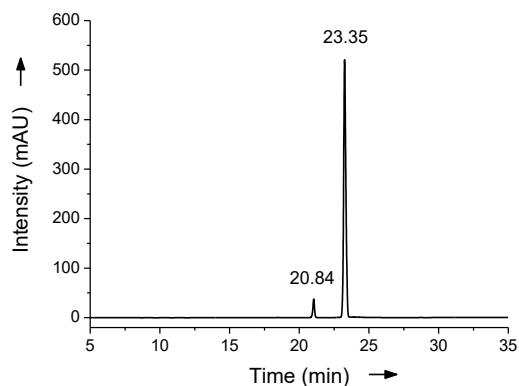
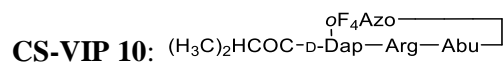


Figure S9. HPLC chromatogram of the purified peptide **CS-VIP 9**. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the lineal precursor of **CS-VIP 10**, Arg(Pbf) was employed. Cyclization was performed in a 27.0 μmol scale as detailed above. After purification, the 1 × TFA salt product (3.3 mg, 3.41 μmol, 13%) was obtained as an orange solid. **HPLC:** $t_R = 17.28$ min (*trans*); 19.72 min (*cis*). Purity = 99%. Formula: C₃₃H₄₁F₄N₁₁O₆. **HRMS-ESI⁺ (m/z):** [M+H]⁺ calcd.: 764.3250; found: 764.3253.

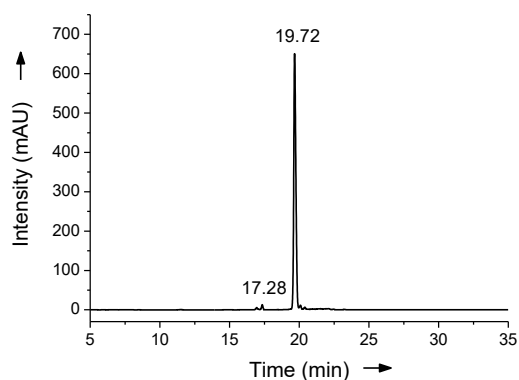
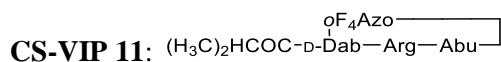


Figure S10. HPLC chromatogram of the purified peptide **CS-VIP 10**. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the linear precursor of **CS-VIP 11**, Arg(Pbf) was employed. Cyclization was performed in a 27.0 μmol scale as detailed above. After purification, the $1 \times$ TFA salt product (3.0 mg, 2.92 μmol , 11%) was obtained as an orange solid. **HPLC:** $t_R = 17.19$ min (*trans*); 19.49 min (*cis*). Purity = 99%. Formula: $\text{C}_{34}\text{H}_{43}\text{F}_4\text{N}_{11}\text{O}_6$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 778.3; found: 778.3.

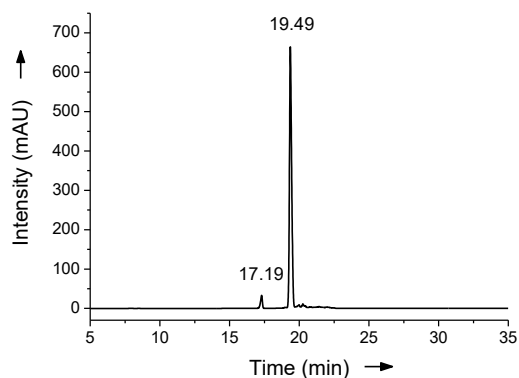
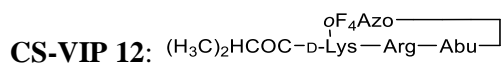


Figure S11. HPLC chromatogram of the purified peptide **CS-VIP 11**. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the linear precursor of **CS-VIP 12**, Arg(Pbf) was employed. Cyclization was performed in a 25.0 μmol scale as detailed above. After purification, the $1 \times$ TFA salt product (1.4 mg, 1.52 μmol , 6%) was obtained as an orange solid. **HPLC:** $t_R = 17.39$ min (*trans*); 19.37 min (*cis*). Purity = 99%. Formula: $\text{C}_{36}\text{H}_{47}\text{F}_4\text{N}_{11}\text{O}_6$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 806.3720; found: 806.3722.

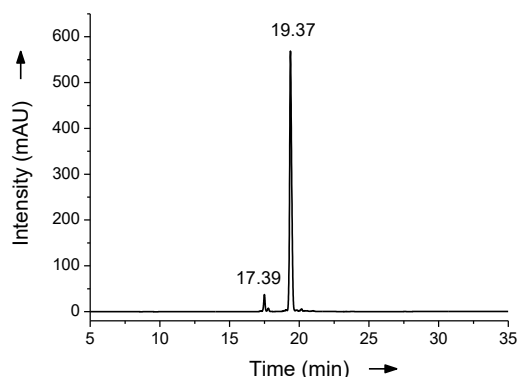
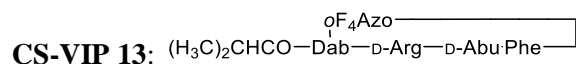


Figure S12. HPLC chromatogram of the purified peptide **CS-VIP 12**. Gradient: 5-75% B in column 1, monitored at 220 nm.



For the linear precursor of **CS-VIP 13**, D-Arg(Pbf) was employed. Cyclization was performed in a 58.6 μmol scale as detailed above. After purification, the 1 \times TFA salt product (2.30 mg, 2.21 μmol , 4%) was obtained as an orange solid. **HPLC:** $t_R = 18.10$ min (*trans*); 20.59 min (*cis*). Purity = 98%. Formula: $\text{C}_{43}\text{H}_{52}\text{F}_4\text{N}_{12}\text{O}_7$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 925.4109; found: 925.4108.

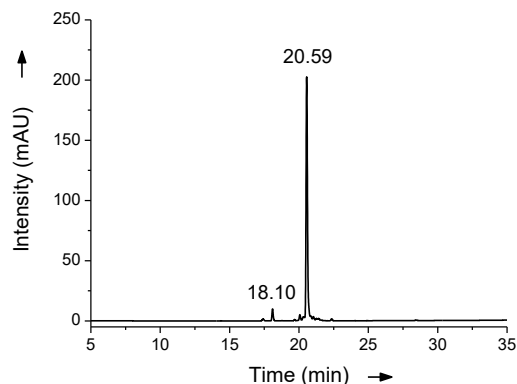
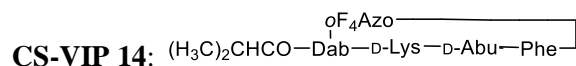


Figure S13. HPLC chromatogram of the purified peptide CS-VIP 13. Gradient: 5-95% B in column 1, monitored at 220 nm.



For the linear precursor of **CS-VIP 14**, D-Lys(Boc) was employed. Cyclization was performed in a 28.4 μmol scale as detailed above. After purification, the 1 \times TFA salt product (5.36 mg, 5.30 μmol , 19%) was obtained as an orange solid. **HPLC:** $t_R = 17.96$ min (*trans*); 20.33 min (*cis*). Purity = 98%. Formula: $\text{C}_{43}\text{H}_{52}\text{F}_4\text{N}_{10}\text{O}_7$. **HRMS-ESI⁺ (m/z):** $[\text{M}+\text{H}]^+$ calcd.: 897.4029; found: 897.4058.

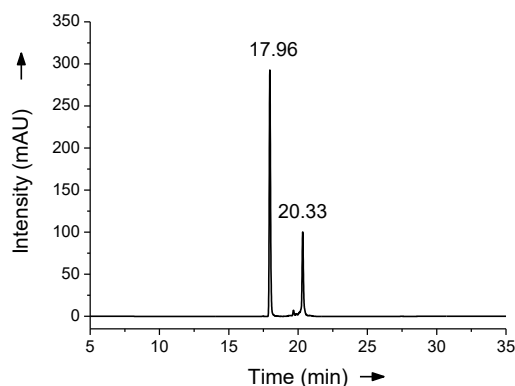


Figure S14. HPLC chromatogram of the purified peptide CS-VIP 14. Gradient: 5-95% B in column 1, monitored at 220 nm.

1.2 Pep-1 carrier

(15): Ac-Lys-Glu-Thr-Trp-Trp-Glu-Thr-Trp-Trp-Thr-Glu-Trp-Ser-Gln-Pro-Lys-Lys-Lys-Arg-Lys-Val-NHCH₂CH₂SH; for the final cleavage step, the resin (scale = 20 μmol , loading = 0.84 mmol/g) was treated with 2.00 mL cleavage cocktail A. After purification, the 6 \times TFA salt product (2.1 mg,

0.578 μmol , 3%) was obtained as a white solid. **HPLC**: $t_R = 17.50$ min. Purity = 98%. Formula: $\text{C}_{140}\text{H}_{202}\text{N}_{36}\text{O}_{33}\text{S}$. **HRMS-ESI⁺ (m/z)**: $[\text{M}+3\text{H}]^{3+}$ calcd.: 983.8401; found: 983.8416.

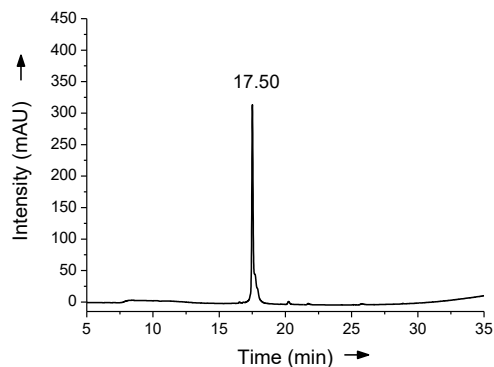


Figure S15. HPLC chromatogram of the purified pep-1 peptide 15. Gradient: 5-95% B in column 1, monitored at 220 nm.

2 Supplementary Figures and Tables

2.1 Characterization of isomerization of CS-VIP 8

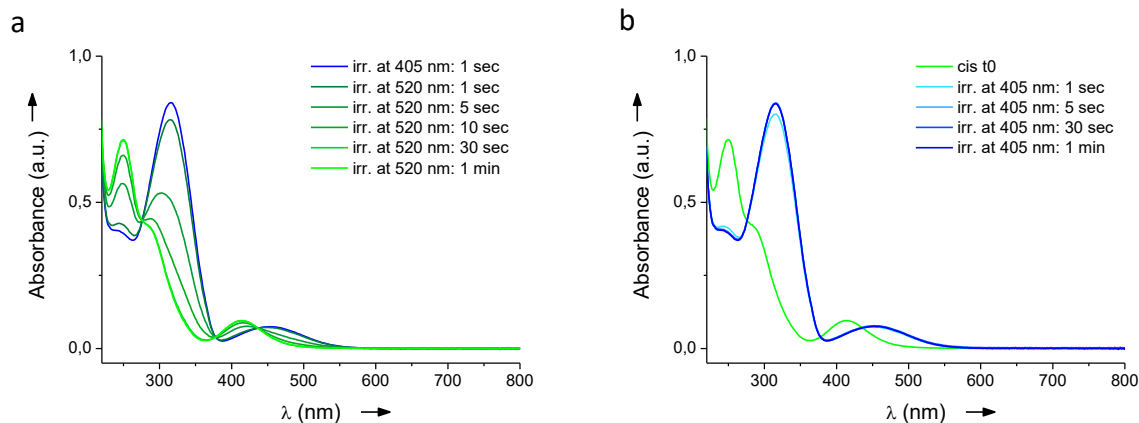


Figure S16. UV-vis absorbance spectra of *oF*₄Azo-containing CS-VIP 8. The solution (34.5 μM in ultrapure water) was irradiated in the cuvette at the respective wavelength; **a**, irradiation of the solution to achieve its PSS at 520 nm for populating the *cis* isomer; **b**, irradiation of the solution at 405 nm to produce the PSS predominantly occupied by the *trans* isomer.

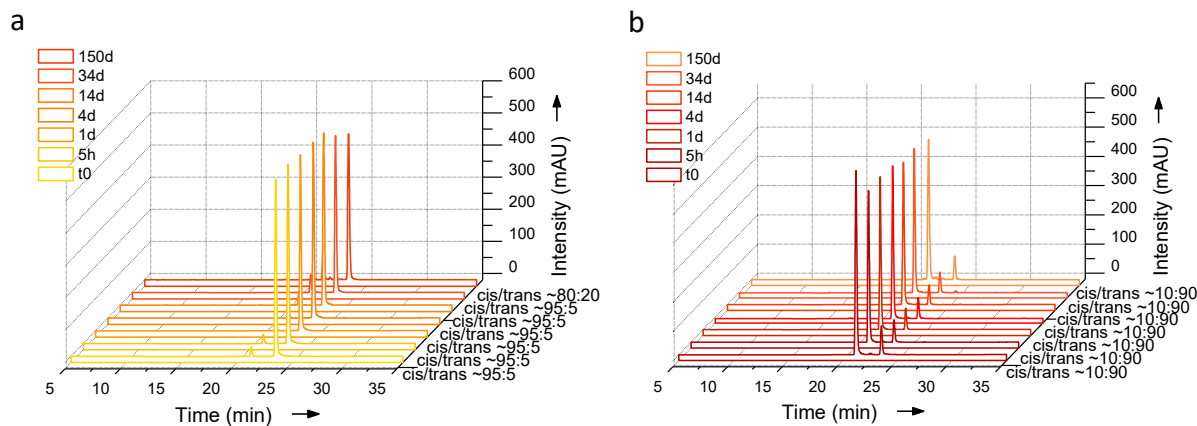


Figure S17. Stability of CS-VIP 8 isomeric states. HPLC chromatograms of the *oF*₄Azo-containing CS-VIP 8 at 100 μM in ultrapure water after irradiation at the respective wavelength followed by storage in total darkness. Column 2 and gradient 5-95% B, at 275 nm (isosbestic point); **a**, after irradiation at 520 nm; **b**, after irradiation at 405 nm.

2.2 Stability against glutathione (GSH)

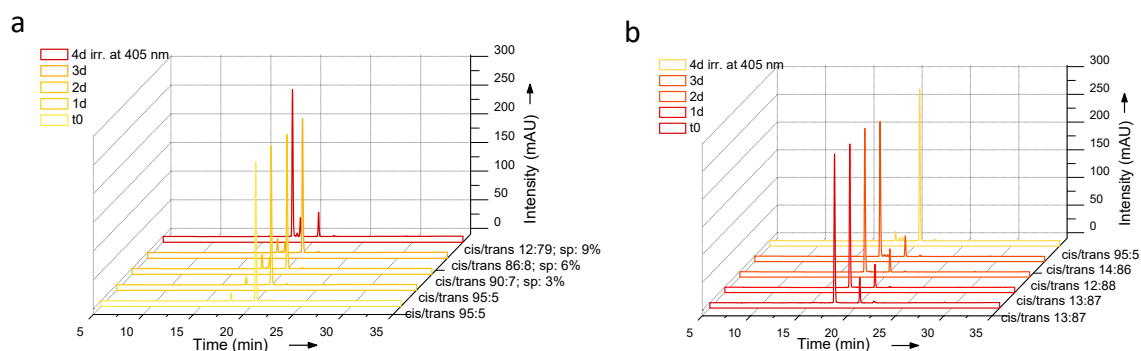
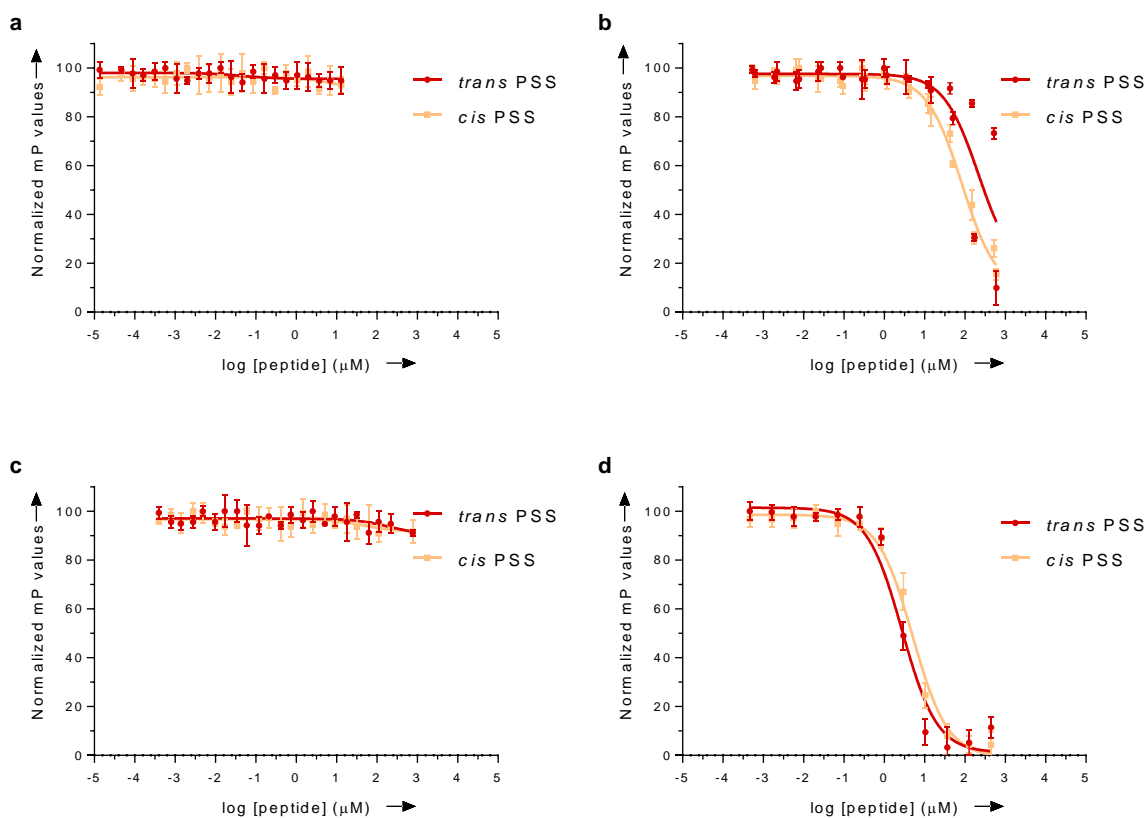
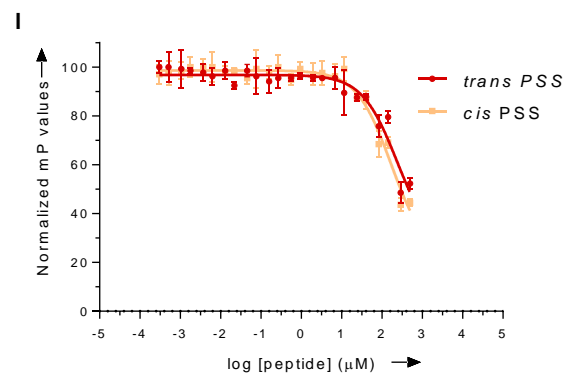
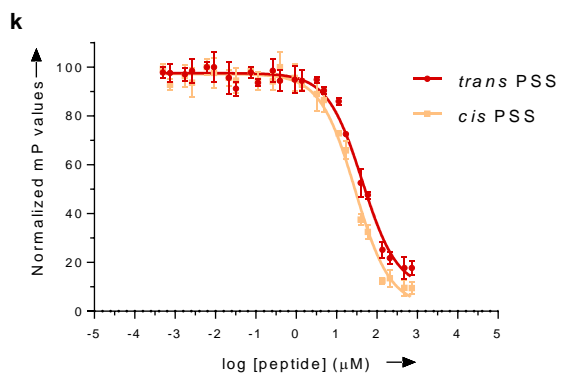
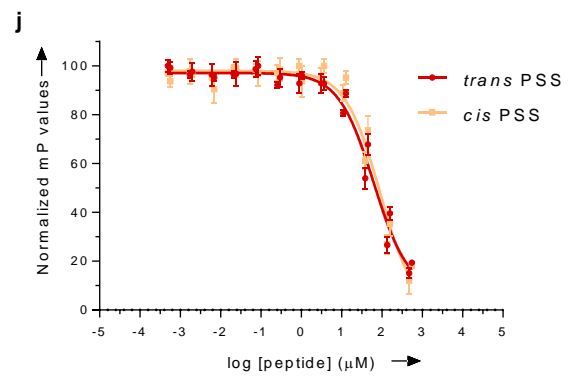
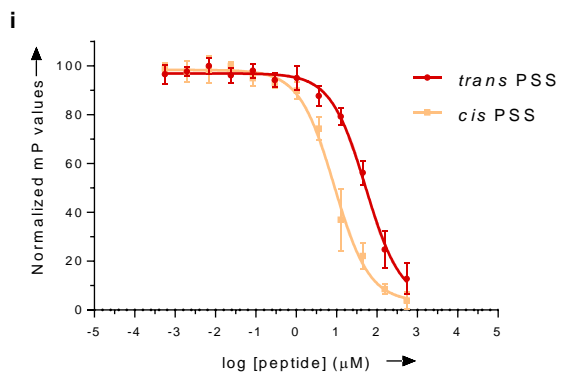
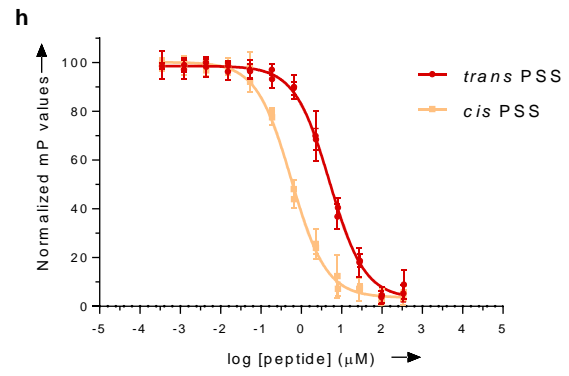
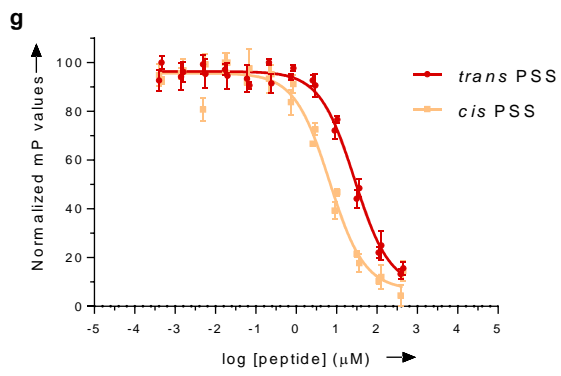
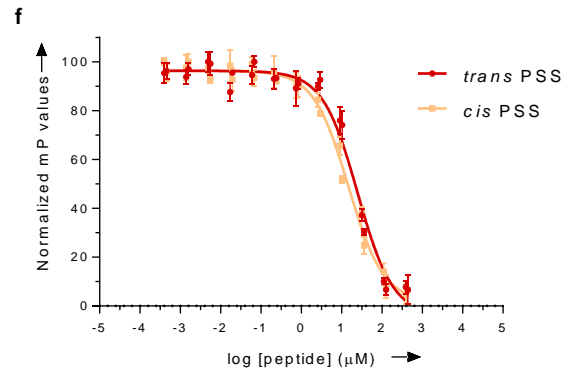
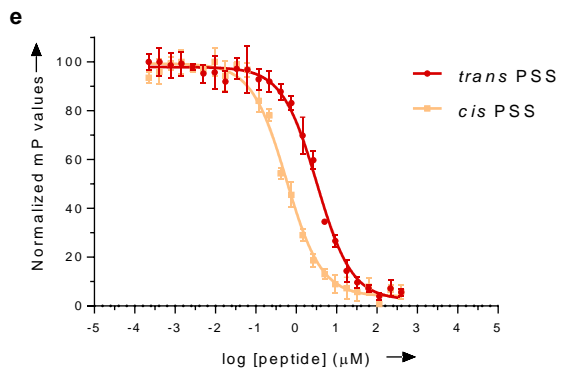


Figure S18. Stability of CS-VIP 8 against GSH. HPLC chromatograms of *oF*₄Azo-peptide **CS-VIP 8** at 96 μ M in ultrapure water incubated with 10 mM glutathione monitored at 275 nm (isosbestic point) using column 3 and a gradient 5-95% B. The solutions were stored in darkness, HPLC chromatograms were recorded after the indicated time (sp: reduced hydrazo species, d: days, irr.: irradiated). **a**, after irradiation at 520 nm to achieve PSS populated by *cis* isomer. After 4 days incubation, the solution was irradiated at 405 nm to switch back to the PSS mostly occupied by *trans* isomer (red); **b**, after irradiation at 405 nm to reach the PSS with *trans* isomer. After 4 days incubation, the solution was irradiated at 520 nm to switch back to PSS state with *cis* isomer (yellow).

2.3 Fluorescence polarization-based assays





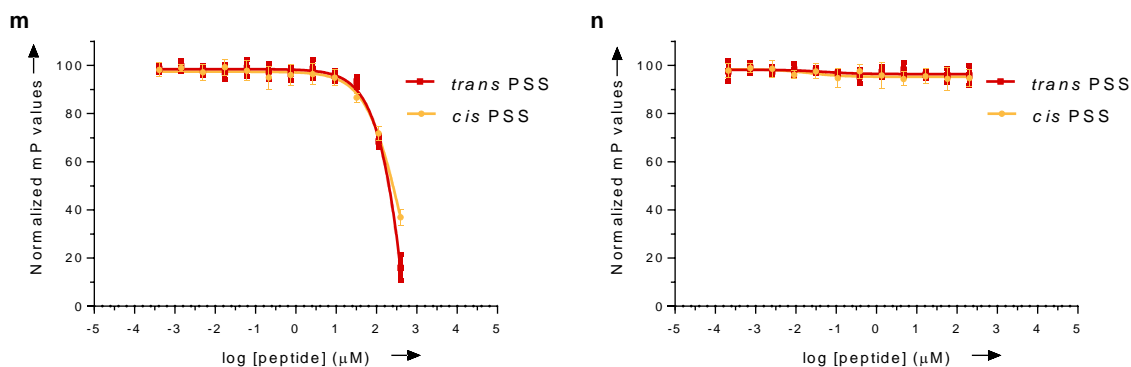


Figure S19. Dose-dependent inhibition curves obtained from FP-based competitive assays. Displaying the mean of the independent experiments (normalized mP values). Dose dependent curves of **a**, peptide CS-VIP 1. **b**, peptide CS-VIP 2. **c**, peptide CS-VIP 3. **d**, peptide CS-VIP 4. **e**, peptide CS-VIP 5. **f**, peptide CS-VIP 6. **g**, peptide CS-VIP 7. **h**, peptide CS-VIP 8. **i**, peptide CS-VIP 9. **j**, peptide CS-VIP 10. **k**, peptide CS-VIP 11. **l**, peptide CS-VIP 12. **m**, peptide CS-VIP 13. **n**, peptide CS-VIP 14. All mean data-points and standard deviations are derived from at least 2 independent experiments, each concentration in triplicate.

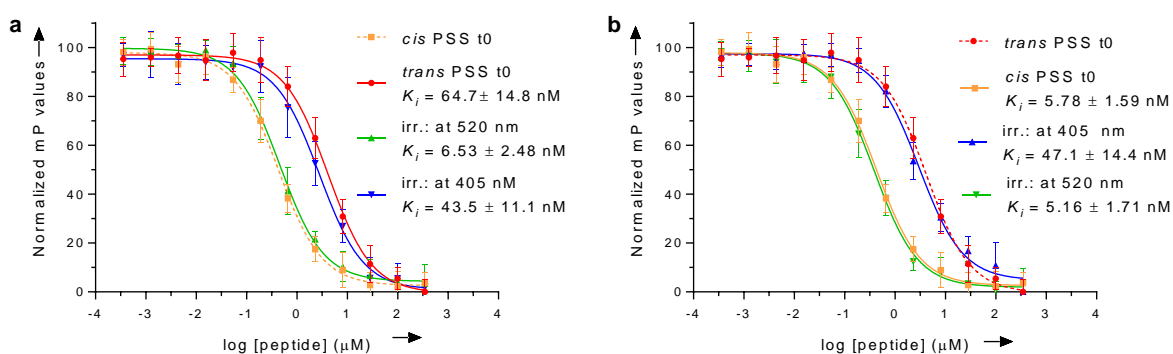


Figure S20. FP-based assays with in situ irradiation. Dose-dependent inhibition curves obtained from FP-based competitive assays, where cyclic peptide CS-VIP 8 was irradiated in situ in presence of WDR5 to switch between isomeric states. Graphs display mean of three independent experiments (normalized mP values). **a**, starting with CS-VIP 8 in its *trans* PSS, the plate was irradiated to switch CS-VIP 8 in situ between the isomeric states. **b**, starting with CS-VIP 8 in its *cis* PSS, the plate was irradiated to switch CS-VIP 8 in situ between the isomeric states. All mean data-points and standard deviations are derived from at least 2 independent experiments, each concentration in triplicate.

2.4 Histone methyltransferase assay

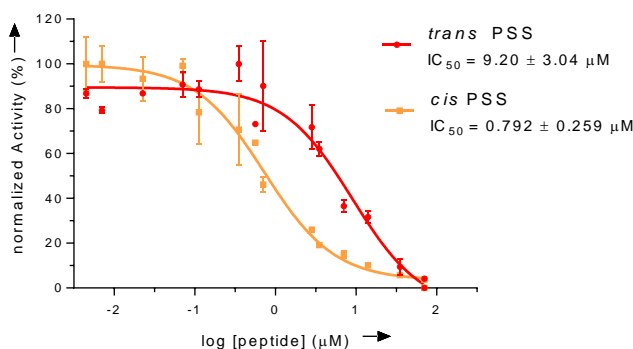


Figure S21. Radiometric histone-methyltransferase (HMT) assay of cyclic peptide CS-VIP 8. All mean data-points and standard deviations are derived from at least 2 independent experiments, each concentration in duplicate.

2.5 Cell viability assays

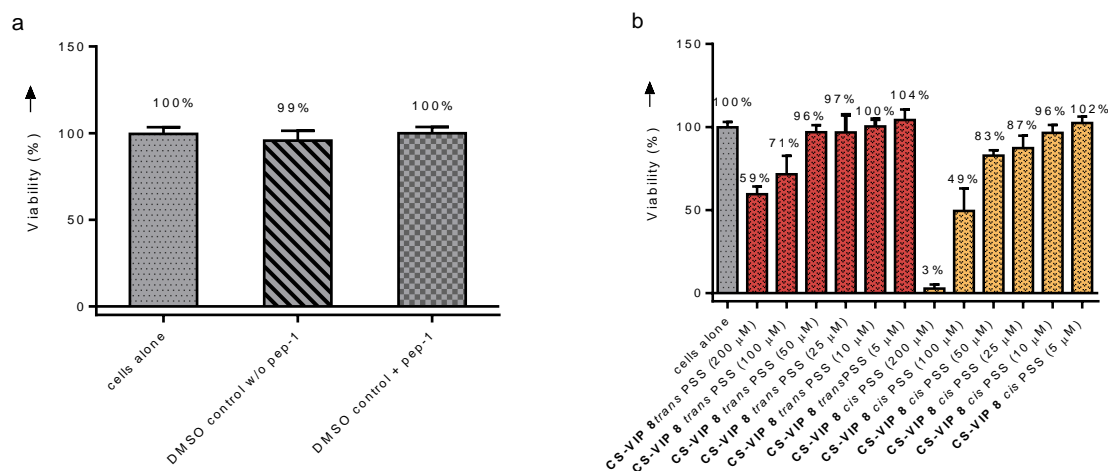


Figure S22. Cell viability of MV4-11 cells after 3 day incubation. **a**, Verifying non-toxicity of DMSO control and DMSO-pep-1 carrier. **b**, Dose dependent toxicity of CS-VIP 8 at its 520 nm and 405 nm PSSs with its predominant *cis* and *trans* isomers. All mean data-points and standard deviations are derived from at least 3 independent experiments, each concentration in triplicate.

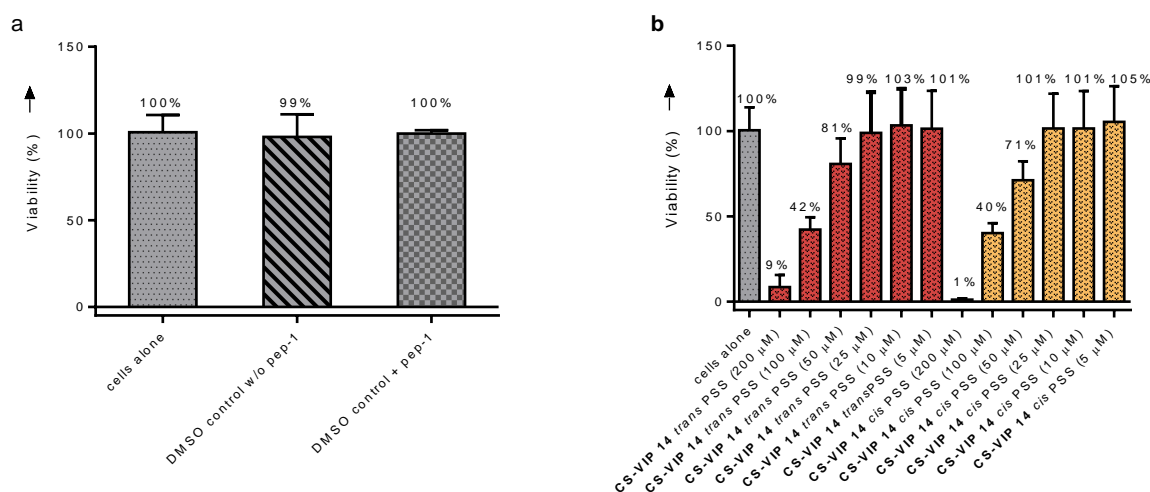


Figure S23. Cell viability of MOLM-13 cells after 3 day incubation. **a**, Cell viability of MOLM-13 cells, verifying non-toxicity of DMSO control and DMSO + pep-1 vehicle after 3 d incubation **b**, Toxicity after 3 day incubation with negative control CS-VIP 14. All mean data-points and standard deviations are derived from at least 2 independent experiments, each concentration in triplicate.

2.6 In vivo zebrafish studies

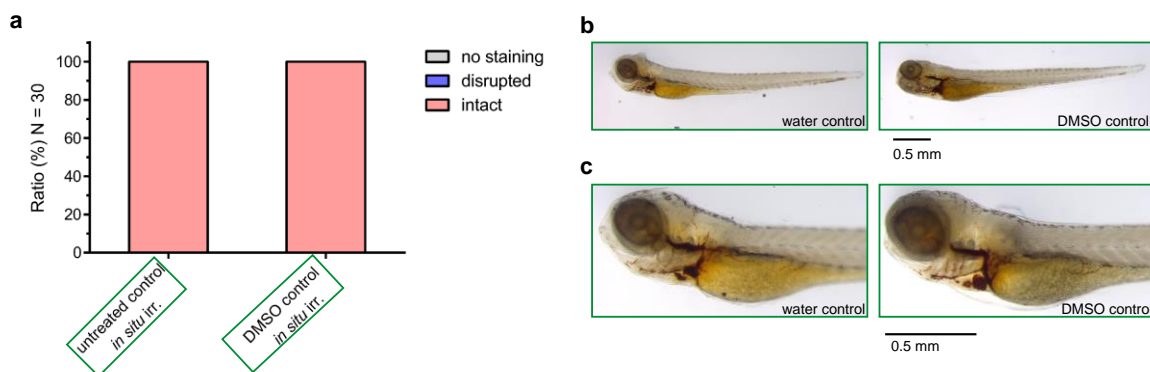


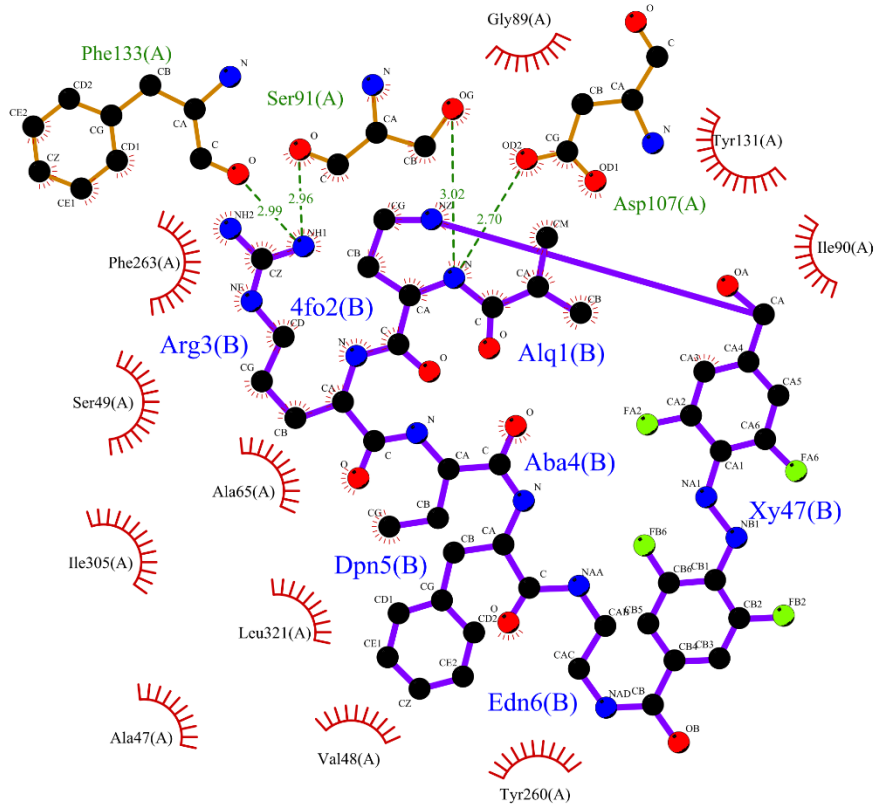
Figure S24. *In vivo* evaluation of CS-VIP 8 using the zebrafish developmental model. Controls of *in situ* irradiated larvae at 520 nm, that do not contain compound CS-VIP 8. **a**, Blots displaying staining patterns of control larvae. **b**, Images of control *in situ* irradiated larvae displaying intact phenotype and staining pattern. **c**, Zoom in of the heart region. Microscopic images of zebrafish larvae were taken at a Leica M205 FA fluorescence stereo microscope (settings: HDR average: automatic; image formats: highest pixel: 3264 × 2448; Input: DCF4955257; automatic exposure (80% brightness); zoom for full larvae: 34.9; zoom for heart region: 84.8.). Experiments have been performed three independent times, each time 2 × 5 larvae / well.

2.7 Structures of WDR5: CS-VIP 8 complexes

Table S3. Data collection and final refinement statistics of the co-crystal structure of WDR5 in complex with CS-VIP 8 in the different states. Diffraction data were processed using XDS.² Data reduction and scaling were done using CCP4i2 (Version 7.0.065),³ and AIMLESS (Version 0.7.3).⁴ The previously published structure of WDR5 in complex with another peptidomimetic inhibitor (PDB code 6IAM) was refined without ligand to the respective datasets via rigid body fit using phenix.refine.⁵ The model was built using COOT (Version 0.8.9)⁶ and further refined in PHENIX (Version 1.11.1).⁵ The peptidomimetic restraints were generated via ReadySet.⁵ *These data have been anisotropically scaled by the STARANISO server; statistics are taken from stats_obs.log. Values in parentheses correspond to the highest resolution shell(s).

Data collection	WDR5•CS-VIP8, <i>cis</i> 1	WDR5•CS-VIP8, <i>cis</i> 1, 405 nm <i>in situ</i>	WDR5•CS-VIP8, <i>cis</i> 1, 405 nm /293 K	WDR5•CS-VIP8, <i>cis</i> 1, 405 nm /180 K	WDR5•CS-VIP8, <i>cis</i> 2
PDB code	7AXS	7AXU	7AXX	7AXQ	7AXP
X-ray source	ESRF ID23-2	ESRF ID23-2	SLS PXI	SLS PXI	DESY P13
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P3 ₂ 2 ₁
Unit-cell parameters (Å)	<i>a</i> =47.0, <i>b</i> =47.2, <i>c</i> =133.5	<i>a</i> =47.4, <i>b</i> =47.2, <i>c</i> =133.5,	<i>a</i> =46.0, <i>b</i> =46.9, <i>c</i> =126.3	<i>a</i> =47.0, <i>b</i> =47.3, <i>c</i> =133.6	<i>a</i> = <i>b</i> =99.2, <i>c</i> =98.5
Wavelength (Å)	1.000	1.000	1.000	1.000	0.976
Resolution (Å)	44.4 – 1.88 (1.95 – 1.88)	44.7 – 1.68 (1.74 – 1.68)	43.3 – 1.80 (1.86 – 1.80)	44.7 – 1.56 (1.618 – 1.56)	64.9 – 3.26x3.26x2.44 (3.34-3.26, 2.72-2.44)*
Completeness (%)	0.895 (0.777)	0.998 (0.998)	0.962 (0.674)	0.976 (0.768)	0.928 (0.927, 0.790)*
Observed reflections	42789 (3521)	69755 (6864)	151557 (6981)	267568 (26016)	155702 (8515, 6310)*
Unique reflections	22301 (1893)	34930 (3439)	25815 (1740)	42057 (3264)	120956 (606, 603)*
Multiplicity	1.9 (1.9)	2.0 (2.0)	5.9 (3.1)	6.2 (6.1)	12.9 (14.0, 10.5)*
<i>R</i> _{merge} (%)	0.029 (0.224)	0.028 (0.198)	0.085 (0.618)	0.090 (1.096)	0.086 (0.555, 1.316)*
Mean <i>I</i> / σ (<i>I</i>)	13.1 (2.78)	15.4 (3.50)	11.3 (1.45)	10.8 (1.34)	16.5 (5.40, 1.75)*
CC _{1/2} (%)	0.999 (0.894)	0.998 (0.922)	0.998 (0.618)	0.998 (0.552)	0.996 (0.994, 0.702)*
Refinement					
<i>R</i> _{work} / <i>R</i> _{free} (%)	0.200/0.293	0.159/0.193	0.180/0.223	0.167/0.200	0.264/0.324
Average <i>B</i> factor (Å ²)	24.92	18.40	25.79	23.17	53.88
No. of atoms					
total	2812	2842	2693	2775	2427
protein	2461	2450	2403	2473	2341
ligand	66	0	66	66	66
solvent	285	392	224	236	20
r.m.s.d., bond lengths (Å)	0.007	0.016	0.007	0.012	0.012
r.m.s.d., bond angles (Å)	1.40	1.77	1.41	1.37	1.32
Ramachandran plot					
favoured (%)	92.86	95.82	94.43	96.14	78.81
allowed (%)	6.17	0.64	4.92	3.54	15.56
outliers (%)	0.97	0.36	0.66	0.32	5.63

a



b

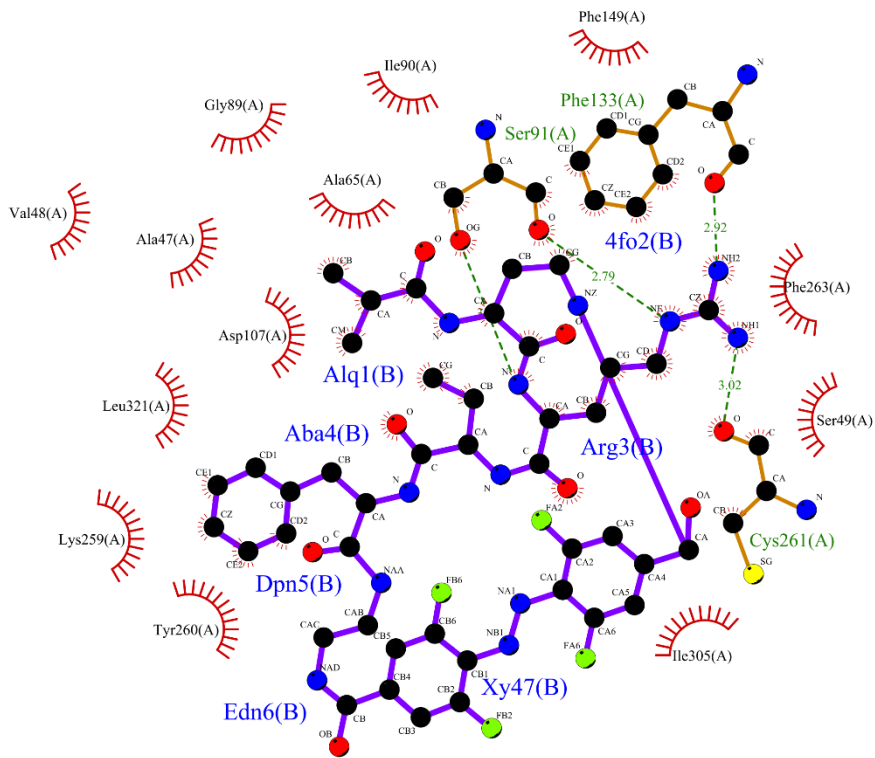


Figure S25. Schematic representation of the key interactions between *cis*-states of CS-VIP 8 and WDR5 generated with LIGPLOT.⁷ a, *cis2* crystal form. b, *cis1* crystal form.

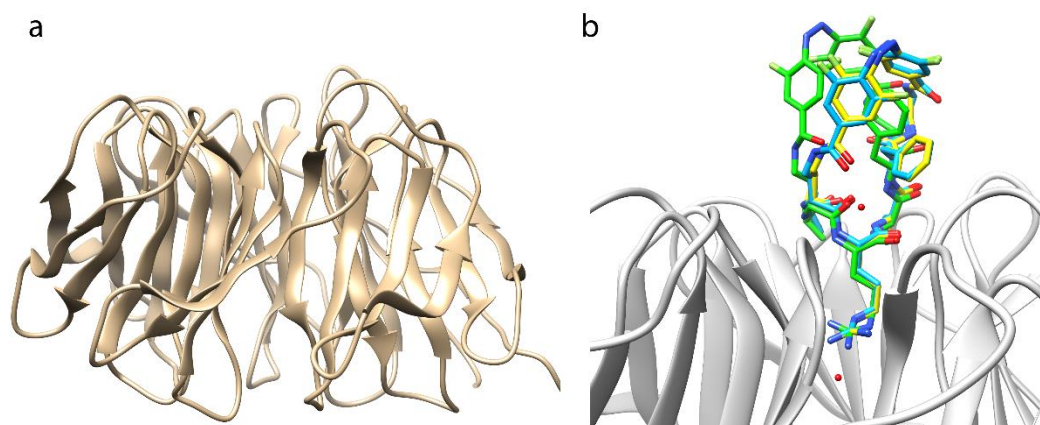


Figure S26. Representation of WDR5•CS-VIP 8 co-crystal structures, where crystals have been additionally irradiated at 405 nm after crystal formation. **a**, Structure of *in situ* illuminated WDR5•*cis*-CS-VIP 8 (*cis*1, 405 nm *in situ*). Only WDR5 is structurally defined, but no bound CS-VIP 8 ligand. **b**, Superimposition of *cis*1 crystal form (yellow), *cis*2 crystal form (green) and the *cis*1 crystal form illuminated with 405 nm at 293 K (*cis*1, 405 nm/298K, dodger blue).

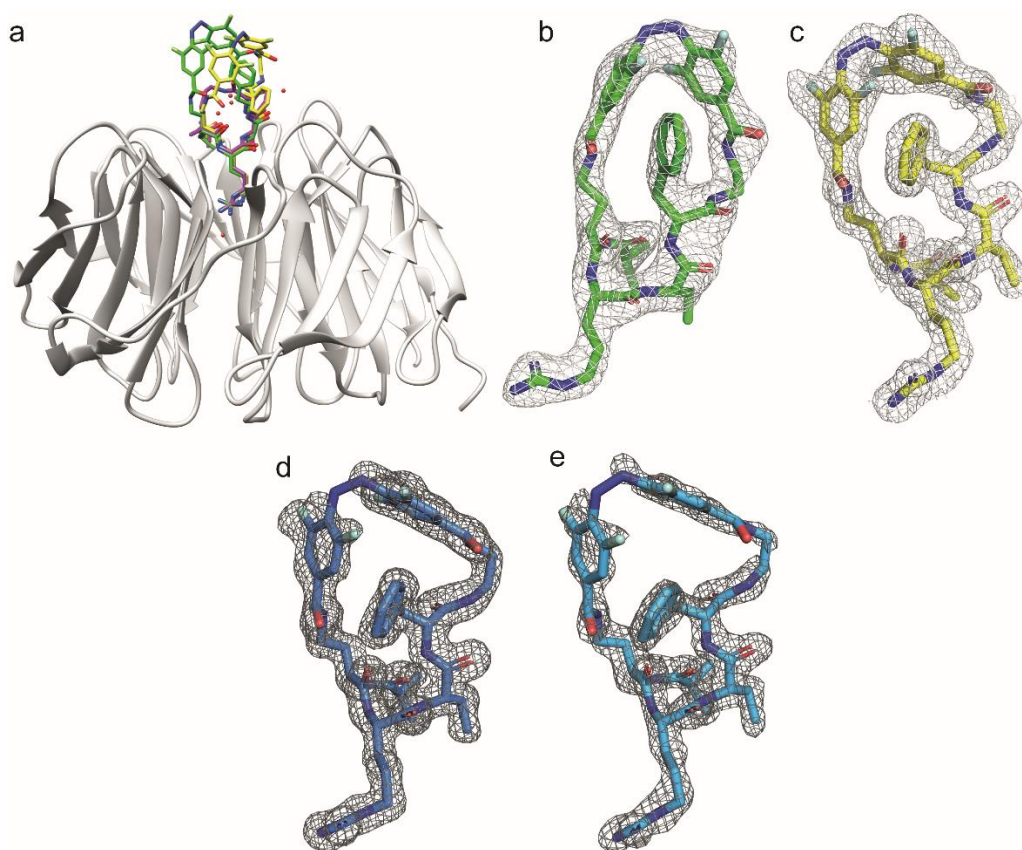


Figure S27. Structural investigation of CS-VIP 8 ligands bound to WDR5. **a**, Experimental $F_{obs,1} - F_{obs,2}$ difference electron density map of *cis*1-state crystals subtracted by 405 nm *in situ* illuminated *cis*1 crystals (red, contouring level 2.0 σ); the difference indicates *off*-diffusion of CS-VIP 8 ligand out of the binding site. **b-e**, $2mF_{obs} - DF_{calc}$ experimental density maps of *cis*-CS-VIP 8 ligand; **b**, *cis*2 crystal form (contouring level 2.0 σ); **c**, *cis*1 crystal form (contouring level 1.0 σ); **d**, *cis*1 crystal form illumination at 180 K (contouring level 1.0 σ); **e**, *cis*1 crystal form illumination at r.t. (contouring level 1.0 σ); **f**, Superposition of crystal structures of *cis*1 state (yellow), *cis*2 state (green) and MM-401 (pink, PDB: 4GM9).

2.8 Microspectroscopy of WDR5•CS-VIP 8 co-crystals

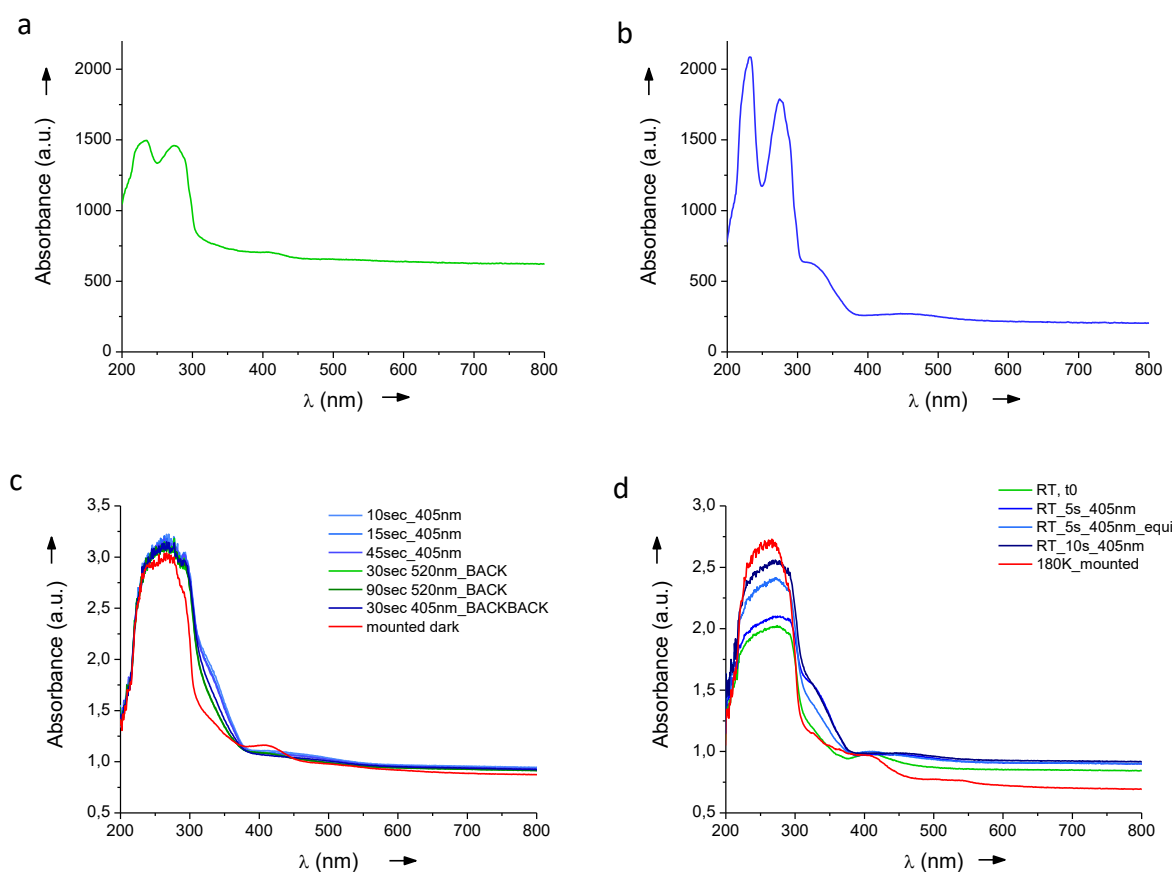


Figure S28. UV-vis absorbance spectra of WDR5•CS-VIP 8 complexes measured at icOS cryobench. **a**, *cis1* state WDR5•CS-VIP 8 co-crystal shows the absence of *trans*-characteristic absorption maximum at 305 nm. **b**, Additional *in situ* irradiation at 405 nm of *cis1* co-crystal within the crystallization plate shows formation of *trans* isomer. **c**, Irradiation at 405 nm of isolated *cis1* state co-crystal at 180 K for different time periods. As obvious by appearance of the absorption maximum at 305 nm, isomerization from *cis* to *trans* is observed. However, subsequently X-ray exposed co-crystals (*mounted dark*) display again the ligand in its *cis* state. **d**, Irradiation at 405 nm of isolated *cis1* state co-crystal at 293 K for different time periods. Again, isomerization from *cis* to *trans* is observed, but, interestingly, after equilibrating the crystal for 5 sec in darkness at r.t. (*RT_10s_405nm_equi*), the ligand reverts partly back into its *cis* state.

2.9 Hydrogen-deuterium exchange mass spectrometry (HDX-MS)

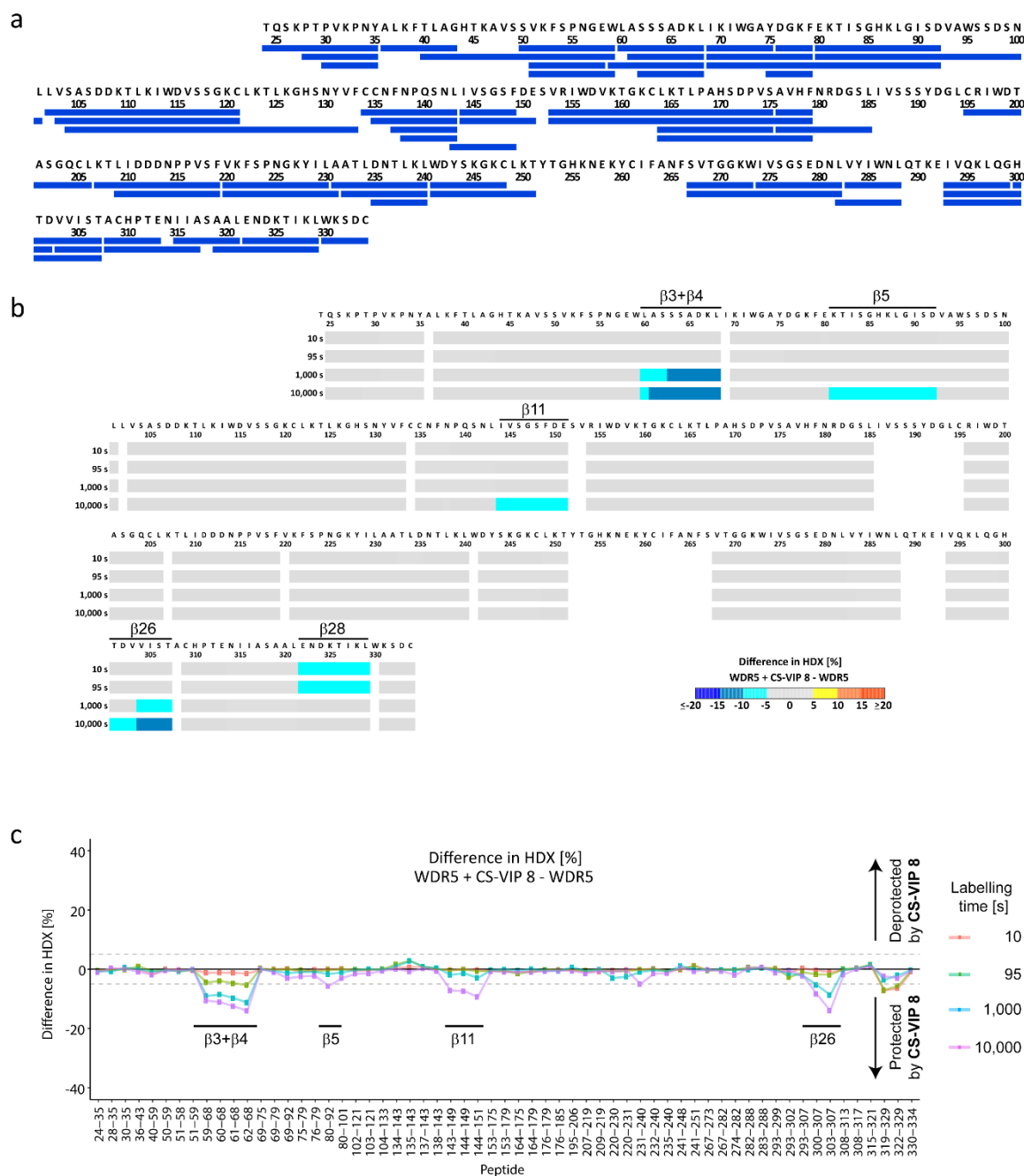


Figure S29. CS-VIP 8-binding to WDR5. S: β -strand. **a**, Blue bars represent peptides of WDR5 that were identified and analyzed for their HDX. **b**, Difference in HDX between CS-VIP 8-bound WDR and apo-WDR5. Blue color represents regions of decreased HDX in presence of CS-VIP 8. The figure was adapted from DynamX 3.0 (Waters). In case of overlapping peptides, the exposure data rendered is determined by the shortest peptide. Where multiple peptide are of shortest length, the peptide with the residue closest to the C-terminus is determining rendering. **c**, Difference in HDX between CS-VIP 8-bound WDR and apo-WDR5 per peptides indicated by their starting and ending amino acid residue on x-axis. The dashed line represents a 5 % threshold.

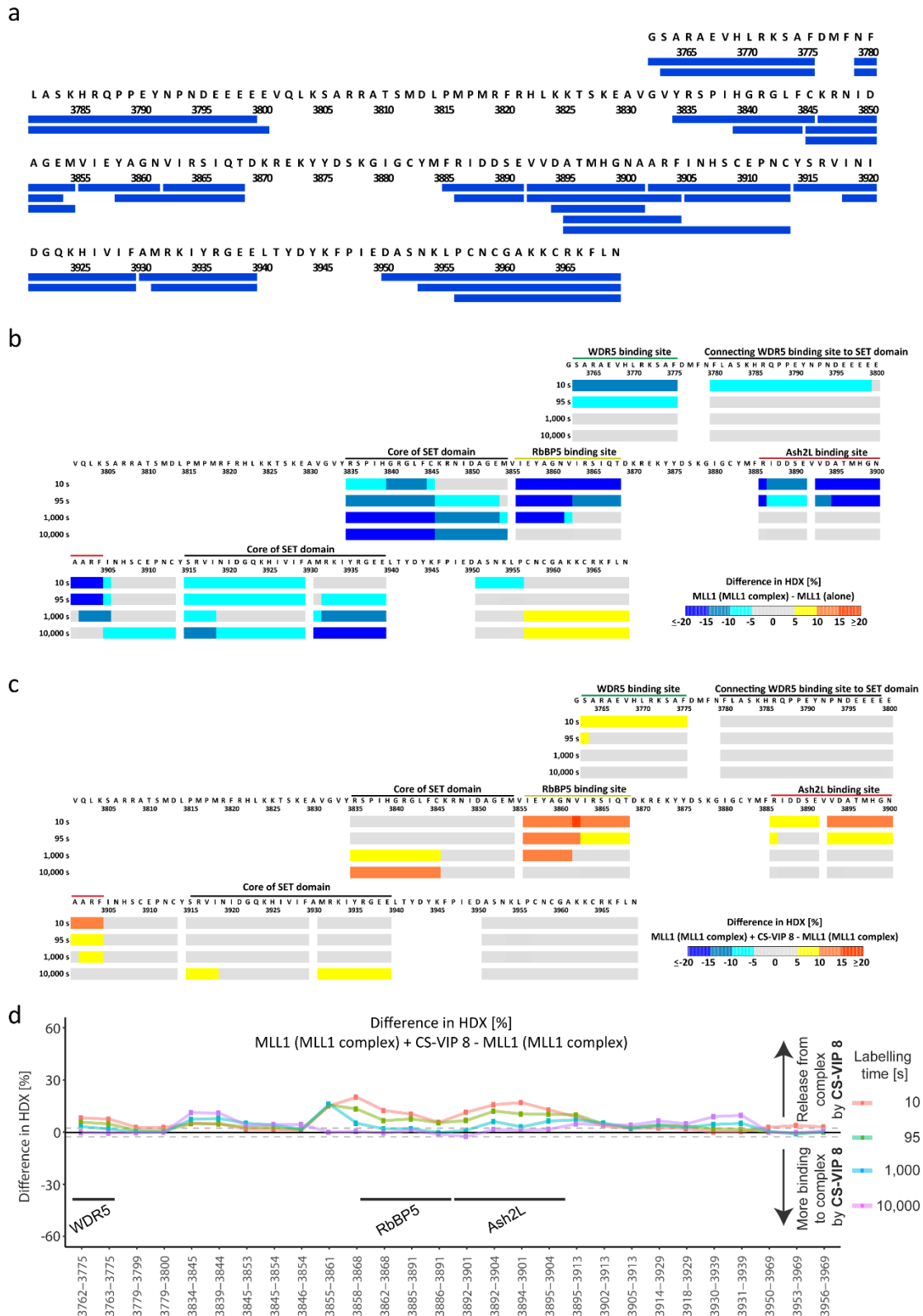


Figure S30. HDX of MLL1. **a**, Blue bars represent peptides of MLL1 that were identified and analyzed for their HDX. **b**, Difference in HDX between MLL1 bound in the MLL1 complex and individual MLL1. Blue and red color represent regions of decreased and increased HDX of MLL1 in the MLL1 complex, respectively. **c**, Difference in HDX of MLL1 in the MLL1 complex in presence and absence of CS-VIP 8. Blue and red color represents regions of decreased and increased HDX in presence of CS-VIP 8. **b-c**, The figures were adapted from DynamX 3.0 (Waters). In case of overlapping peptides, the exposure data rendered is determined by the shortest peptide. Where multiple peptide are of shortest length, the peptide with the residue closest to the C-terminus is determining rendering. **d**, Difference in HDX between MLL1 in complex with and without addition of CS-VIP 8, per peptides indicated by their starting and ending amino acid residue on x-axis. The dashed line represents a 5 % threshold.

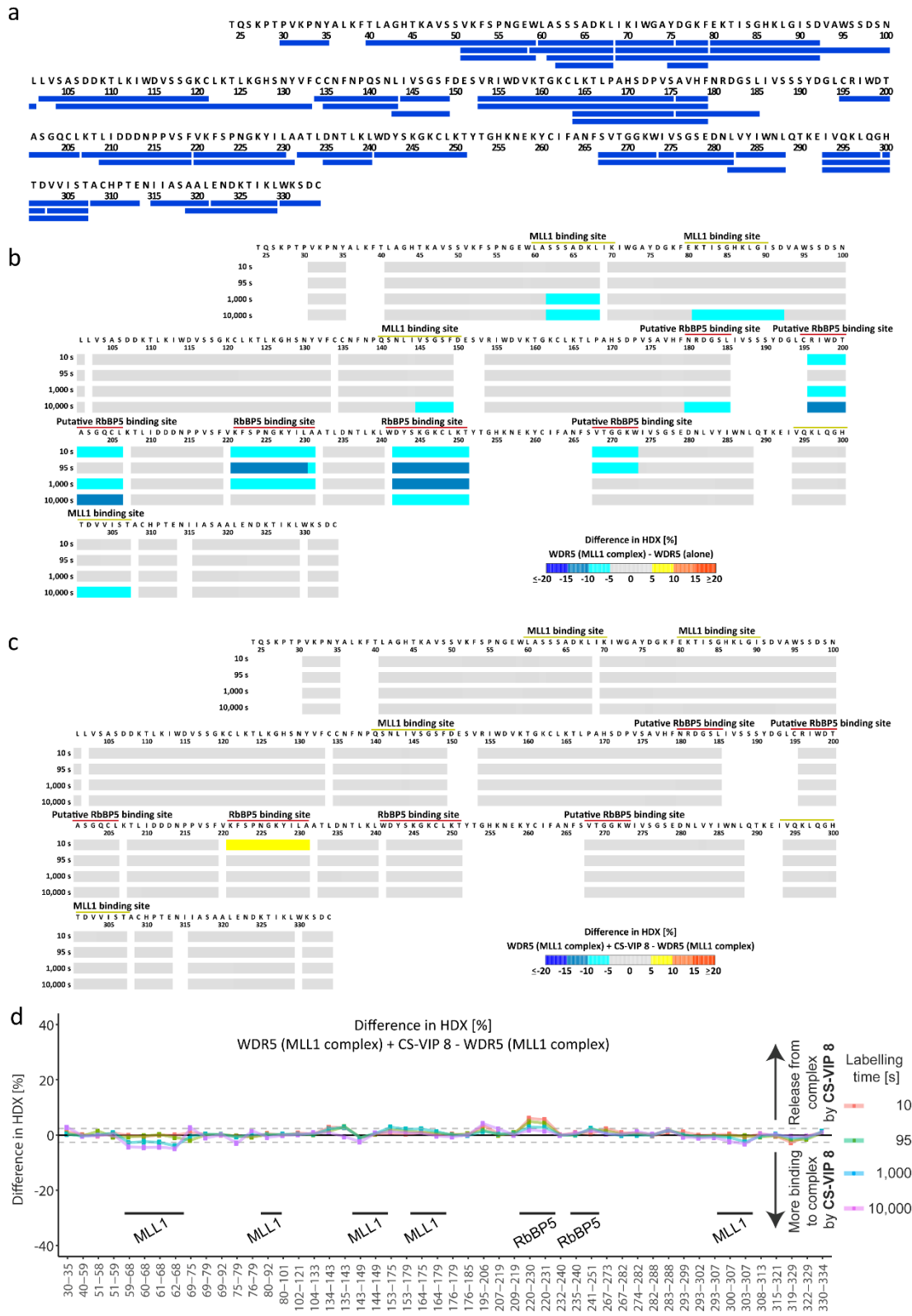


Figure S31. HDX of WDR5. **a**, Blue bars represent peptides of WDR5 that were identified and analyzed for their HDX. **b**, Difference in HDX between WDR5 bound in the MLL1 complex and individual WDR5. Blue color represents regions of decreased HDX of WDR5 in the MLL1 complex. **c**, Difference in HDX of WDR5 in the MLL1 complex in presence and absence of CS-VIP 8. Red color indicates regions of increased HDX in presence of CS-VIP 8. **b-c**, The figures were adapted from DynamX 3.0 (Waters). In case of overlapping peptides, the exposure data rendered is determined by the shortest peptide. Where multiple peptide are of shortest length, the peptide with the residue closest to the C-terminus is determining rendering. **d**, Difference in HDX between WDR5 in complex with and without addition of CS-VIP 8, per peptides indicated by their starting and ending amino acid residue on x-axis. The dashed line represents a 5% threshold.

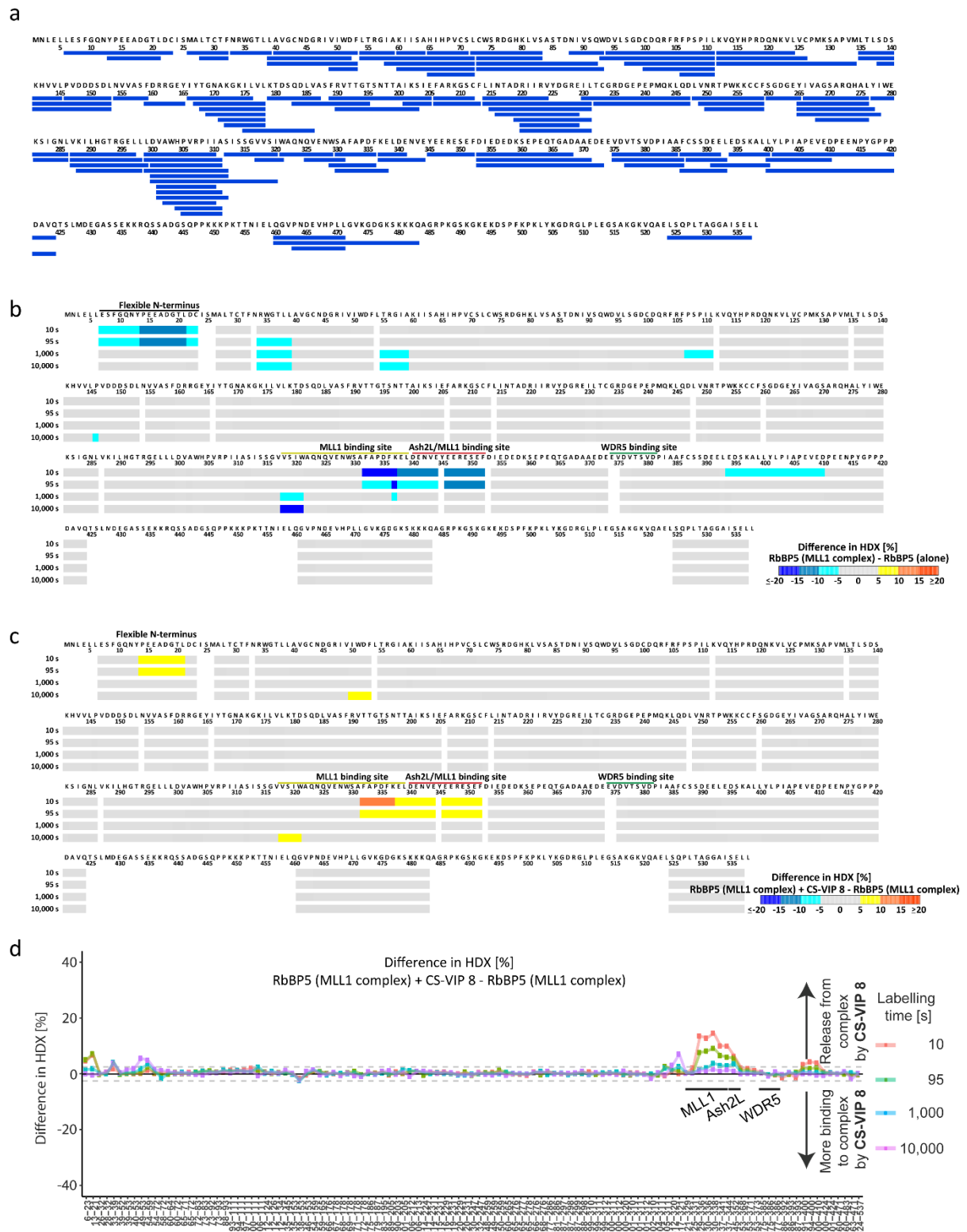


Figure S32. HDX of RbBP5. **a**, Blue bars represent peptides of RbBP5 that were identified and analyzed for their HDX. **b**, Difference in HDX between RbBP5 bound in the MLL1 complex and individual RbBP5. Blue color represents regions of decreased HDX of RbBP5 in the MLL1 complex. **c**, Difference in HDX of RbBP5 in the MLL1 complex in presence and absence of **CS-VIP 8**. Red color represents regions of increased HDX in presence of **CS-VIP 8**. **b-c**, The figures were adapted from DynamX 3.0 (Waters). In case of overlapping peptides, the exposure data rendered is determined by the shortest peptide. Where multiple peptide are of shortest length, the peptide with the residue closest to the C-terminus is determining rendering. **d**, Difference in HDX between RbBP5 in complex with and without addition of **CS-VIP 8**, per peptides indicated by their starting and ending amino acid residue on x-axis. The dashed line represents a 5 % threshold.

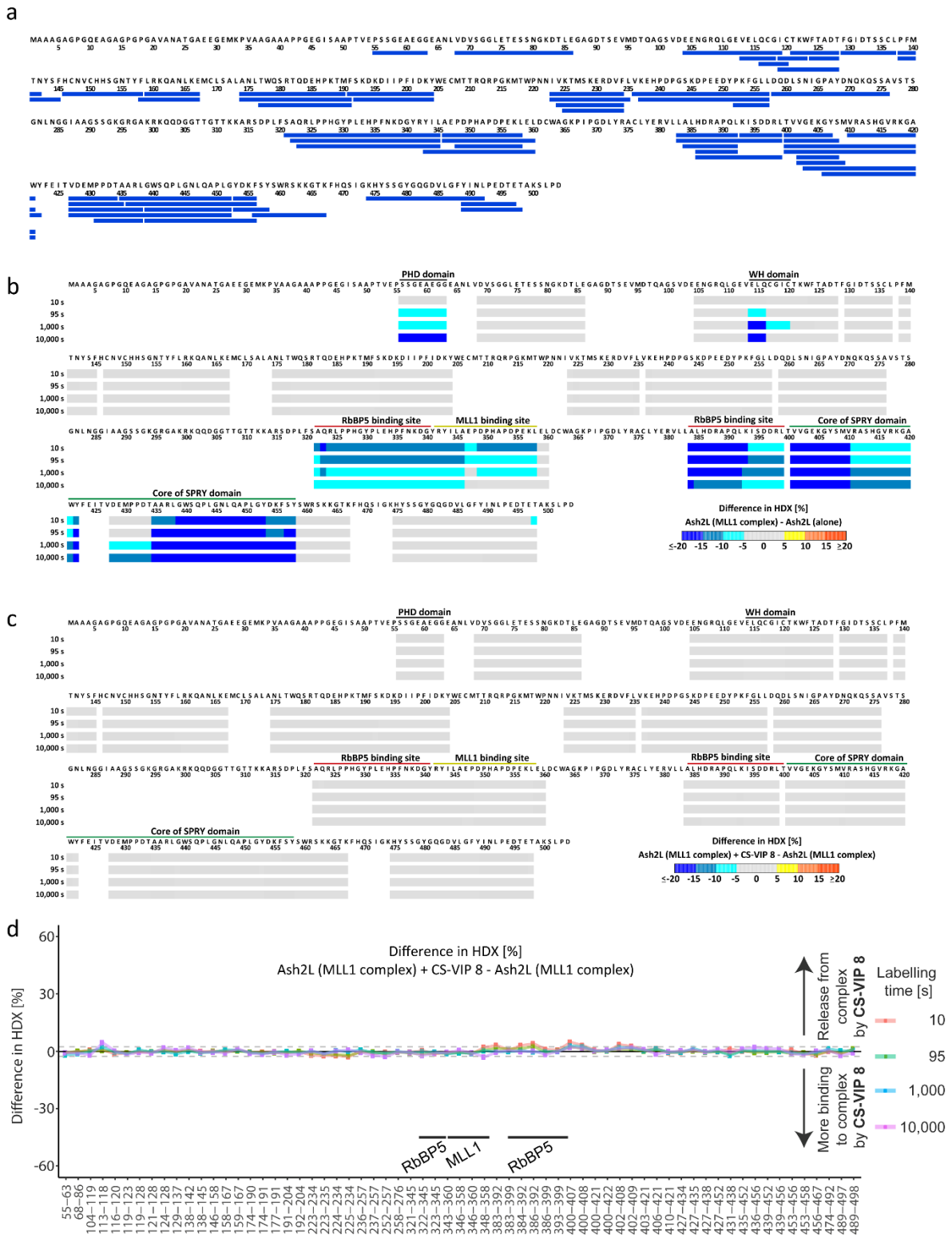


Figure S33. HDX of Ash2L. **a**, Blue bars represent peptides of Ash2L that were identified and analyzed for their HDX. **b**, Difference in HDX between Ash2L bound in the MLL1 complex and individual Ash2L. Blue color represents regions of decreased HDX of Ash2L in the MLL1 complex. **c**, Difference in HDX of Ash2L in the MLL1 complex in presence and absence of CS-VIP 8. No CS-VIP 8-dependent conformational changes of Ash2L in the MLL1 complex are observed. **b-c**, The figures were adapted from DynamX 3.0 (Waters). In case of overlapping peptides, the exposure data rendered is determined by the shortest peptide. Where multiple peptide are of shortest length, the peptide with the residue closest to the C-terminus is determining rendering. **d**, Difference in HDX between Ash2L in complex with and without addition of CS-VIP 8, per peptides indicated by their starting and ending amino acid residue on x-axis. The dashed line represents a 5 % threshold.

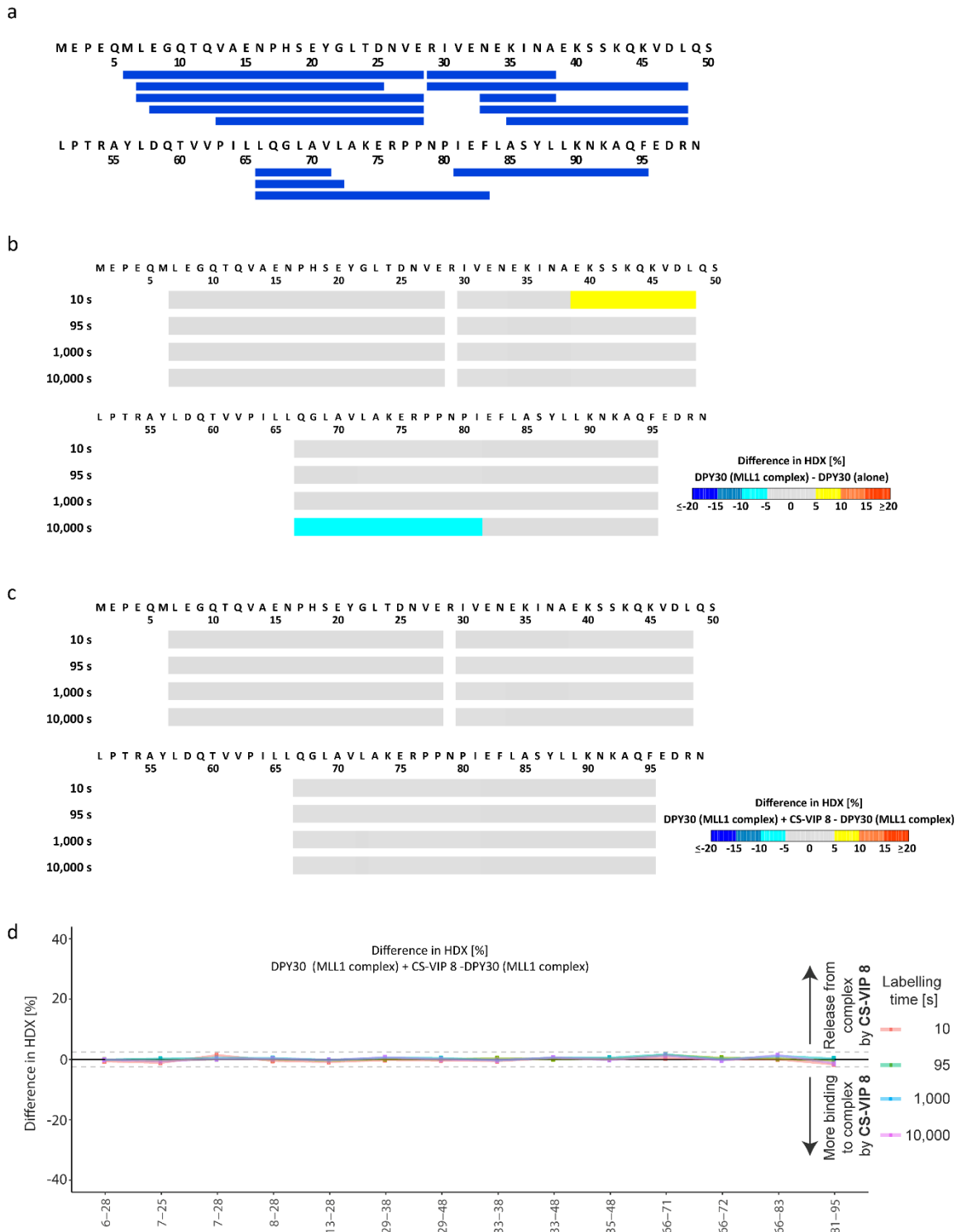


Figure S34. Blue bars HDX of DPY30. **a**, Blue bars represent peptides of DPY30 that were identified and analyzed for their HDX. **b**, Difference in HDX between DPY30 in context of the MLL1 complex and individual DPY30. Blue and red color represents regions of decreased and increased HDX of DPY30 in the MLL1 complex, respectively. **c**, Difference in HDX of DPY30 in the MLL1 complex in presence and absence of CS-VIP 8. No CS-VIP 8-dependent conformational changes of Ash2L in the MLL1 complex are observed. **b-c**, The figures were adapted from DynamX 3.0 (Waters). In case of overlapping peptides, the exposure data rendered is determined by the shortest peptide. Where multiple peptide are of shortest length, the peptide with the residue closest to the C-terminus is determining rendering. **d**, Difference in HDX between DPY30 in complex with and without addition of CS-VIP 8, per peptides indicated by their starting and ending amino acid residue on x-axis. The dashed line represents a 5 % threshold.

Table S4. Overview of data obtained by hydrogen-deuterium exchange mass spectrometry (HDX-MS). *50 peptides of WDR5 were identified in both datasets. 11 peptides were identified only for WDR5 alone but not when in the MLL1 complex owing to overlaps with peptides from the other components of the complex.

Dataset 1	WDR5 ± CS-VIP 8 520 nm PSS (<i>cis</i>)
Conditions of H/D exchange	25 °C in 20 mM Tris-Cl pH 8.0, 150 mM NaC, 0.2 % (v/v) DMSO; final D2O = 90 %
Timecourse of H/D exchange	10/95/1000/10000 s
Replicates	3, technical (separate H/D exchange reactions measured)
Number of Peptides	61*
Sequence Coverage [%]	90.7
Redundancy per amino acid	2.32

Table S5. Overview of data obtained by hydrogen-deuterium exchange mass spectrometry (HDX-MS).

Dataset 2	MLL1 complex ± CS-VIP 8 520 nm PSS (<i>cis</i>)
Conditions of H/D exchange	25 °C in 20 mM Tris-Cl pH 8.0, 150 mM NaC, 0.2 % (v/v) DMSO; final D2O = 90 %
Timecourse of H/D exchange	10/95/1000/10000 s
Replicates	3, technical (separate H/D exchange reactions measured)
Number of Peptides	50 (WDR5)*
	28 (MLL1)
	113 (RbBP5)
	65 (Ash2L)
	14 (DPY30)
Sequence coverage [%]	86.5 (WDR5)
	70.2 (MLL1)
	84.4 (RbBP5)
	64.1 (Ash2L)
	73.7 (DPY30)
Redundancy per amino acid	2.04 (WDR5)
	3.38 (MLL1)
	2.91 (RbBP5)
	2.60 (Ash2L)
	2.92 (DPY30)

2.10 His-SUMO pull-down assay

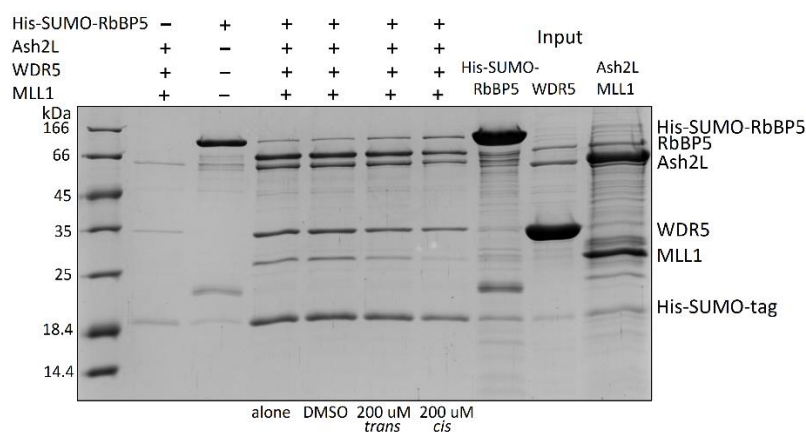


Figure S35. In vitro evaluation of CS-VIP 8. His-Sumo pull-down assay of 4-mer MLL1 complex (His-SUMO-RbBP5, Ash2L, MLL1, WDR5) using His-SUMO-RbBP5, incubated alone, with the DMSO control and 200 μ M of cyclic peptide CS-VIP 8 in its 405 nm and 520 nm PSSs.

2.12 Virtual docking and molecular dynamic studies

Table S6. Results of docking procedure for WDR5:*trans* CS-VIP 8 complex (left table) and for WDR5:*cis* CS-VIP 8 (right table). ChemPLP is an empirical scoring function in GOLD which is used to rank the binding poses of docked compounds. ChemPLP of MM-401 = 101.5614; pK_i of MM-401 = 9.00.

<i>trans</i>	pK_i	CHEMPLP
CS-VIP 2	5.4750	64.7386
CS-VIP 4	7.3665	70.4741
CS-VIP 5	7.2924	70.5104
CS-VIP 6	6.4045	60.9488
CS-VIP 7	6.3487	64.7224
CS-VIP 8	7.1024	69.8156
CS-VIP 9	6.0804	61.6970
CS-VIP 10	5.9586	60.4900
CS-VIP 11	6.1624	64.0815
CS-VIP 12	5.4330	56.9175
correlation coefficient $R^2 = 0.8365$		

<i>cis</i>	pK_i	CHEMPLP
CS-VIP 2	5.8827	70.7245
CS-VIP 4	7.1024	73.8027
CS-VIP 5	8.0969	75.3565
CS-VIP 6	6.6216	72.2703
CS-VIP 7	6.9830	75.7999
CS-VIP 8	8.0969	78.6736
CS-VIP 9	6.8633	76.2567
CS-VIP 10	5.8182	61.7735
CS-VIP 11	6.3382	61.0028
CS-VIP 12	5.5376	60.0407
correlation coefficient $R^2 = 0.765$		

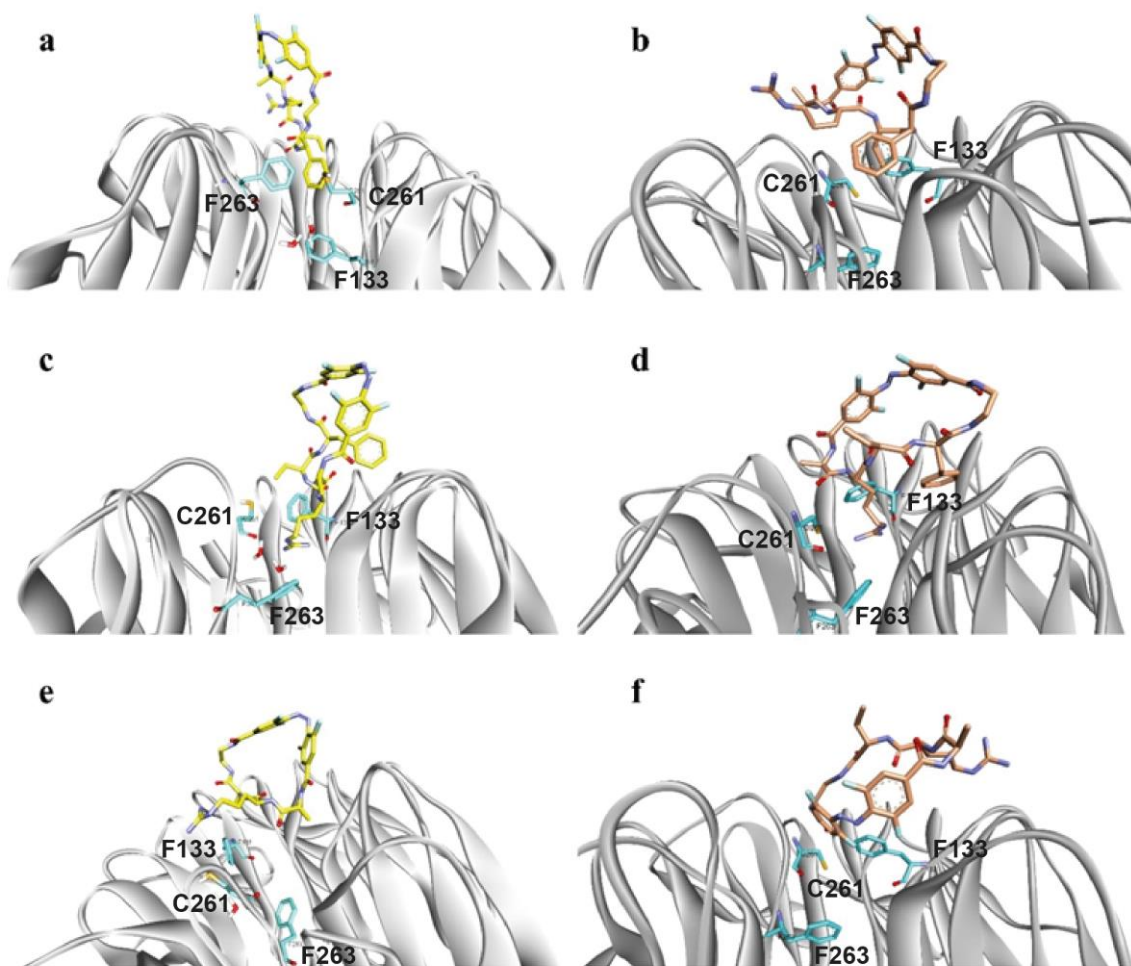


Figure S36. Presentation of binding poses of different ligands (class A) docked onto WDR5. **a**, *cis*-CS-VIP 1. **b**, *trans*-CS-VIP 1. **c**, *cis*-CS-VIP 2. **d**, *trans*-CS-VIP 2. **e**, *cis*-CS-VIP 3. **f**, *trans*-CS-VIP 3. *Cis* isomers are presented on left-side hand (carbon atoms are labelled in yellow), whereas *trans* isomers are presented on right-side hand (carbon atoms are labelled in crimson red).

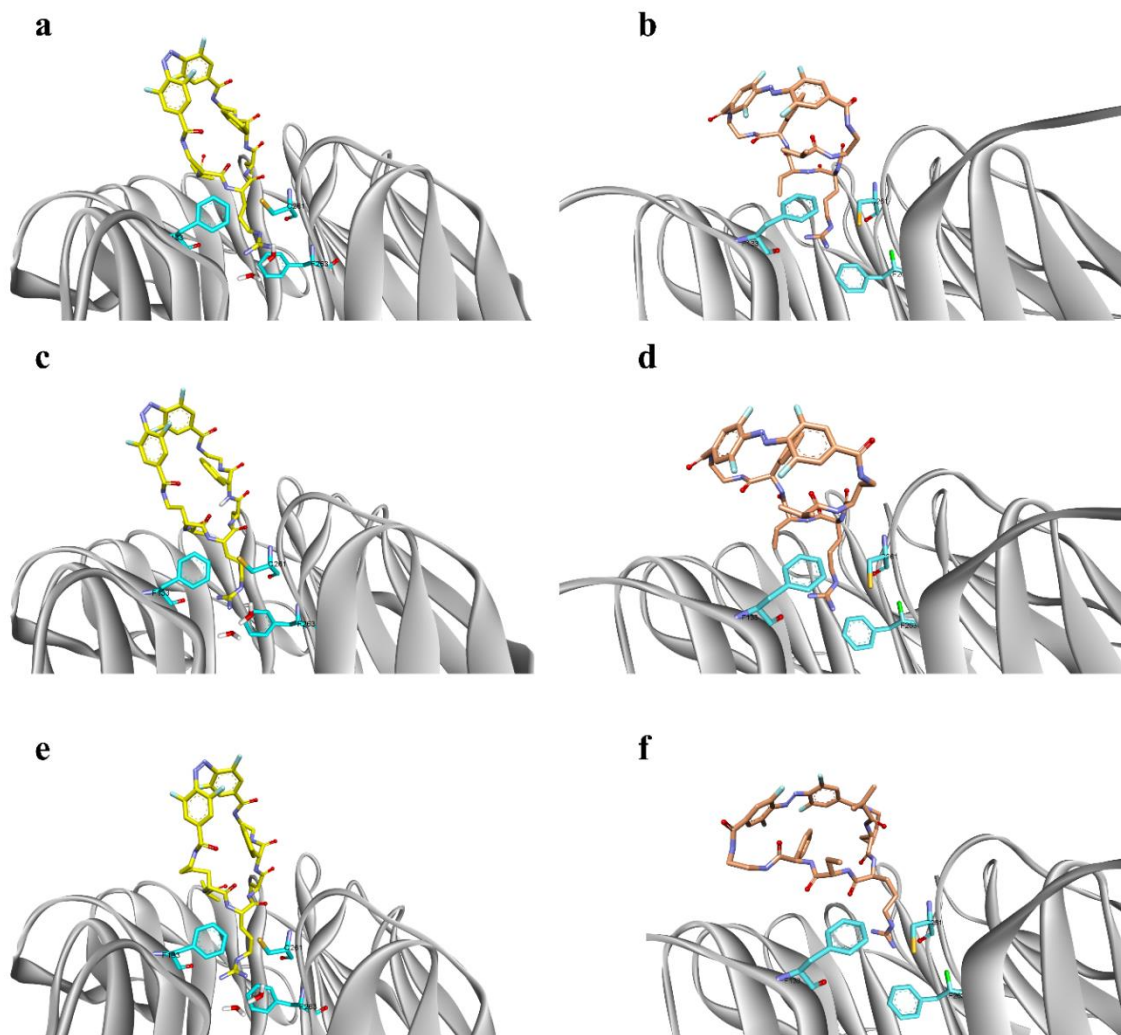


Figure S37. Presentation of binding poses of different ligands (class B) docked onto WDR5. **a**, *cis*-CS-VIP 4. **b**, *trans*-CS-VIP 4. **c**, *cis*-CS-VIP 5. **d**, *trans*-CS-VIP 5. **e**, *cis*-CS-VIP 6. **f**, *trans*-CS-VIP 6. *Cis* isomers are presented on left-side hand (carbon atoms are labelled in yellow), whereas *trans* isomers are presented on right-side hand (carbon atoms are labelled in crimson red).

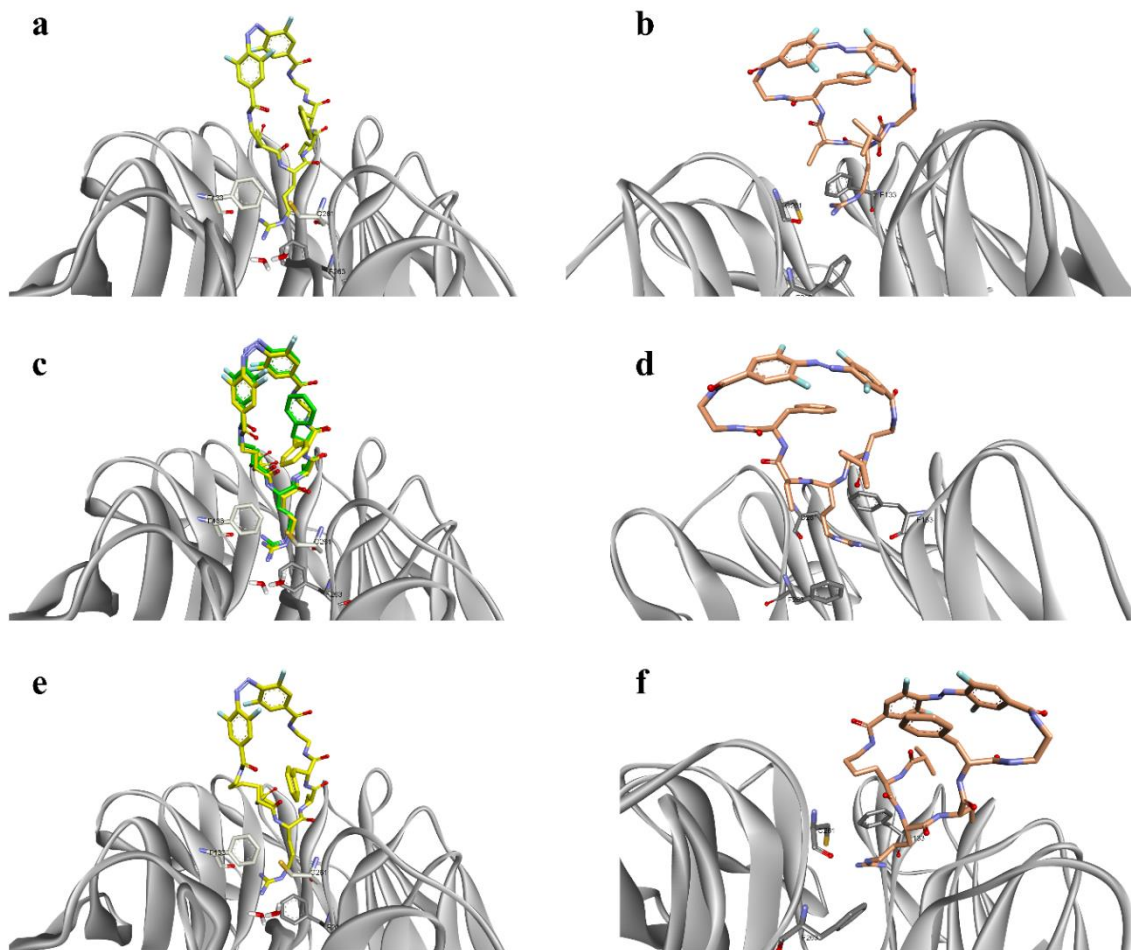


Figure S38. Presentation of binding poses of different ligands (class C) docked onto WDR5. **a**, *cis*-CS-VIP 7. **b**, *trans*-CS-VIP 7. **c**, *cis*-CS-VIP 8 (*cis*-CS-VIP 8 from crystal structure is presented in green, while docked *cis*-CS-VIP 8 is presented in yellow). **d**, *trans*-CS-VIP 8. **e**, *cis*-CS-VIP 9. **f**, *trans*-CS-VIP 9. *Cis* isomers are presented on left-side hand (carbon atoms are labelled in yellow), whereas *trans* isomers are presented on right-side hand (carbon atoms are labelled in crimson red).

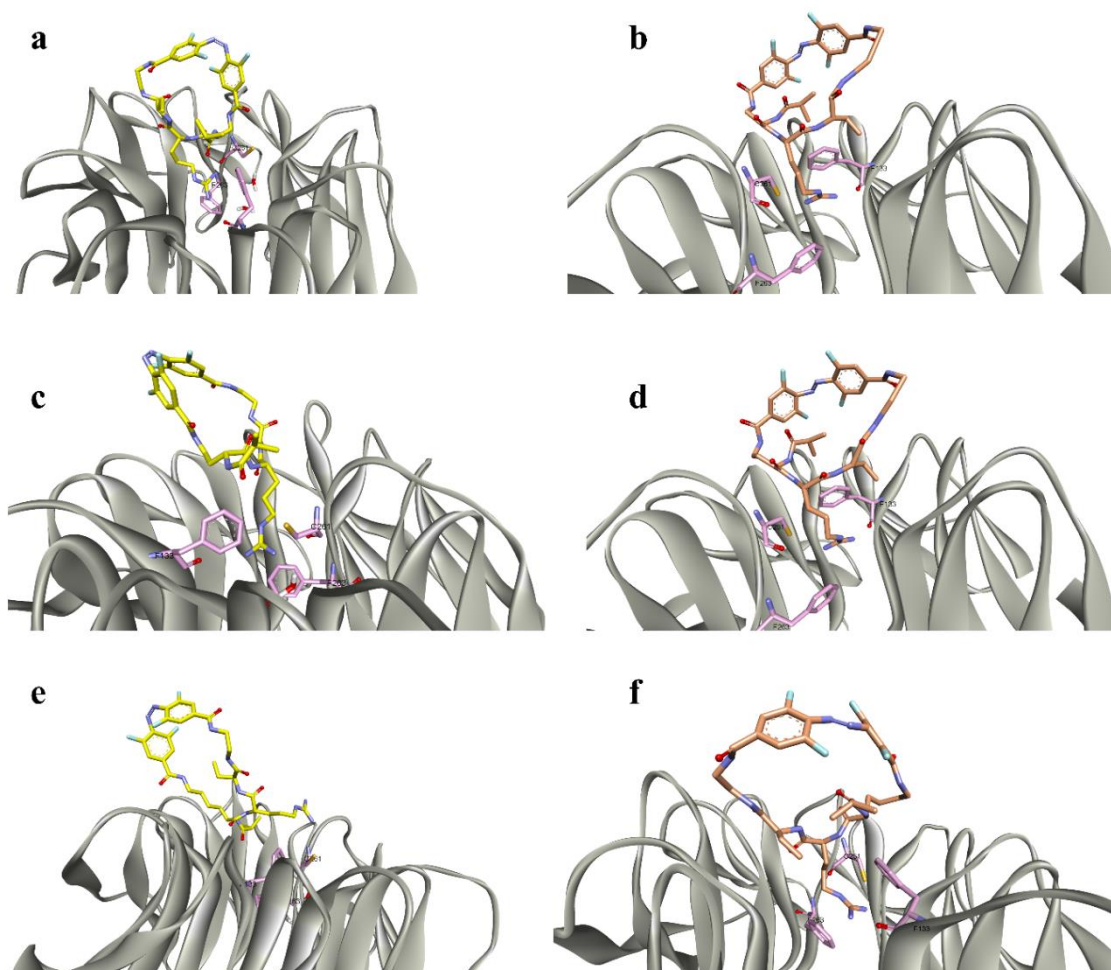


Figure S39. Presentation of binding poses of different ligands (class D) docked onto WDR5. **a**, *cis*-CS-VIP 10. **b**, *trans*-CS-VIP 10. **c**, *cis*-CS-VIP 11. **d**, *trans*-CS-VIP 11. **e**, *cis*-CS-VIP 12. **f**, *trans*-CS-VIP 12. *Cis* isomers are presented on left-side hand (carbon atoms are labelled in yellow), whereas *trans* isomers are presented on right-side hand (carbon atoms are labelled in crimson red).

2.12.1 Supplementary note N1

Molecular modelling of WDR5:*trans*-CS-VIP 8 complex and its photoinduced isomerization:

To ultimately get an idea of the conformational and binding free energy changes associated with photoisomerization of **CS-VIP 8**, two independent approaches were used: a) “diabatic switching potential” MD simulation, and b) well-tempered metadynamics (WT-metaD) simulation. Both approaches have been successfully used in simulations of azobenzene-containing systems.^{8,9}

a) In switching potential approach, starting from WDR5:*cis*-**CS-VIP 8** system, S1 state was initiated and simulated up to 25 ns in this state (Figure S40). However, diffusion of the ligand observed in our X-ray study was not observed through simulations, probably due to differences in experimental and simulated conditions as well as limited sampling of MD. Further inspection of observed behavior revealed that the -D-Dab-Arg-Abu- region of **CS-VIP 8** remained in the bioactive-like conformation during simulations either in S1 state or during application of switching potentials, with preservation of all crucial intra and intermolecular bonds (Figure S40). After application of switching potential, system

switched to the S0 state in *trans*-**CS-VIP 8** configuration. From this point, the system was simulated using standard, equilibrium simulations (100 ns of plain MD simulation, total of 3 replicas), in order to assess reliability of predicted structure of WDR5: *trans*-**CS-VIP 8** complex. RMSD analysis (Figure S40 b) indicated proper stabilization of the system. However, observed RMSD standard deviations were the largest among all studied systems (Figures S37 b, S41 and S43) which could indicate significant drop in affinity of *trans* state of **CS-VIP 8** compared to *cis* state. To examine this further, WT-metaD was used.

b) In order to confirm binding mode of *trans*-**CS-VIP 8** and to gain deeper insight into energetics behind *cis/trans* isomerization, WT-metaD was used. The ability of this approach to explore free-energy changes associated to the end states of thermodynamic cycle was utilized for the description of the influence that WDR5 has on process of *cis/trans* isomerization. Comparison of free-energy profiles obtained through WT-metaD for both, solo ligand **CS-VIP 8** and WDR5:**CS-VIP 8** complex, revealed significant differences in $\Delta G^{(trans-cis)}$ between bound and unbound ligand (Figure S41 a). Thermodynamic cycle used in these calculations is presented in Figure S41 B. Systems were simulated until convergence of all replicas (Figure S41 c) and obtained free-energies of *cis/trans* isomerization ($\Delta G^{(trans-cis)}$) revealed that the bound state was associated with much lower (more positive) change in free energy ($\Delta G^{(trans-cis)}$), in comparison to the free ligand. This indicates that WDR5 is able to stabilize the *cis* state to much larger extent, compared to corresponding *trans* state. As a consequence, the *cis* state of **CS-VIP 8** has a much larger binding affinity to WDR5. Namely, predicted $\Delta\Delta G^{(trans-cis)}_{binding} = -9.35$ kcal/mol indicates several hundred thousand times increase in affinity of *cis*-**CS-VIP 8** state under simulated conditions – (Figure S41 a). This result further confirmed experimental observation that binding affinities observed for *trans*-**CS-VIP 8** may result from the residual *cis*-isomer (Figure 3f). Additionally, according to the obtained free-energy profiles, presence of WDR5 largely increase energy barrier for transition from *cis* state into *trans* state (Figure S41 A). This result is aligned to the proposed hypothesis from in situ back isomerization experiments (Figure S20) which states that once *cis* isomer is tightly bound, isomerization to the *trans* state becomes more difficult. The *trans* stereoisomer was thermodynamically more stable in both states, which is aligned to the standard behavior of azobenzenes.

Considering predicted *trans*-**CS-VIP 8** poses, results from WT-metaD and switching potential simulation were in agreement. After clustering of all obtained trajectories, the most populated centroid representative corresponding to *trans*-**CS-VIP 8** was compared to co-crystal structure of *cis*-**CS-VIP 8** state (Figure S40). According to our results, ligand's conformational shift during photoisomerization appears to be large, but the -D-Dab-Arg-Abu- region of **CS-VIP 8** was able to remain in its bioactive-like conformation. The largest difference between *cis* and *trans* states was observed in the positioning of azobenzene part of the **CS-VIP 8** (Figure S37). However, this conformational shift resulted in a large drop in the affinity of *trans*-isomer towards WDR5. Taken together, results from “diabatic switching potential” MD simulation, and WT-metaD corroborate the inability of trapping the *trans*-isomer in crystals, proposed residual *cis*-isomer affinity as well as results from in situ irradiation FP assays.

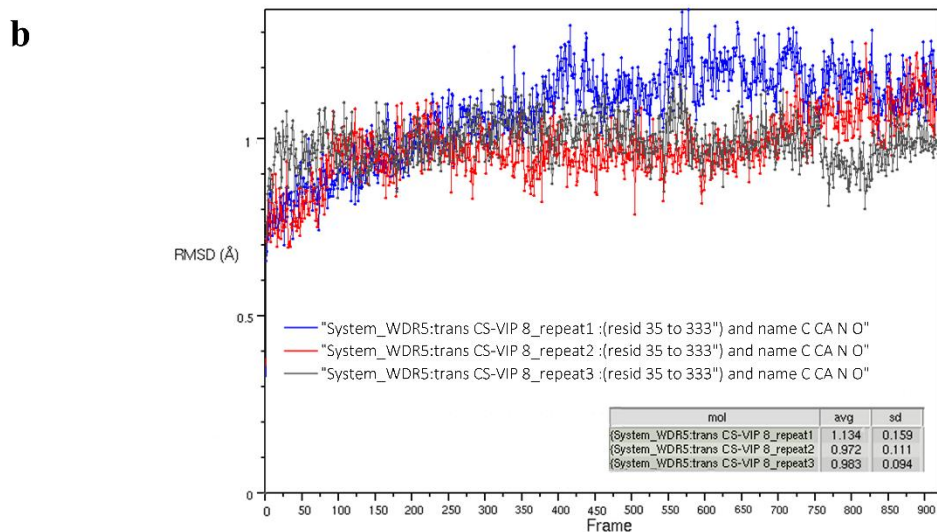
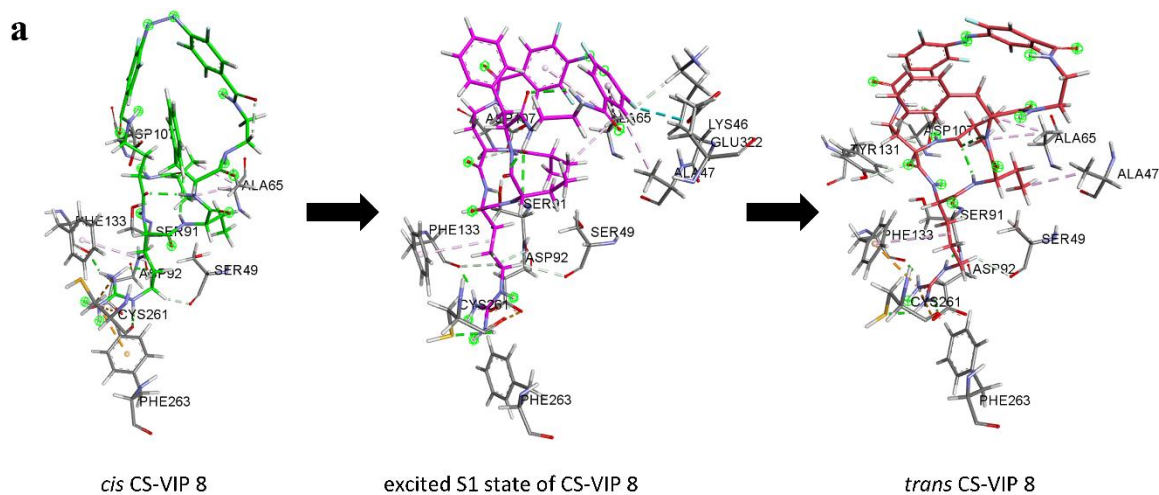


Figure S40. Diabatic switching potential simulations of WDR5:CS-VIP 8. **a**, Overall protocol overview and comparison of interaction modes of *cis*-CS-VIP 8 co-crystal structure (left – green sticks), modelled excited S1 state of WDR5:CS-VIP 8 (middle – magenta sticks) and modelled WDR5:*trans*-CS-VIP 8 (right – red sticks). **B**, Backbone RMSD fluctuations detected in three repeats of MD simulations of the *trans*-CS-VIP 8 system. Average RMSD values and standard deviations are presented in grey boxes.

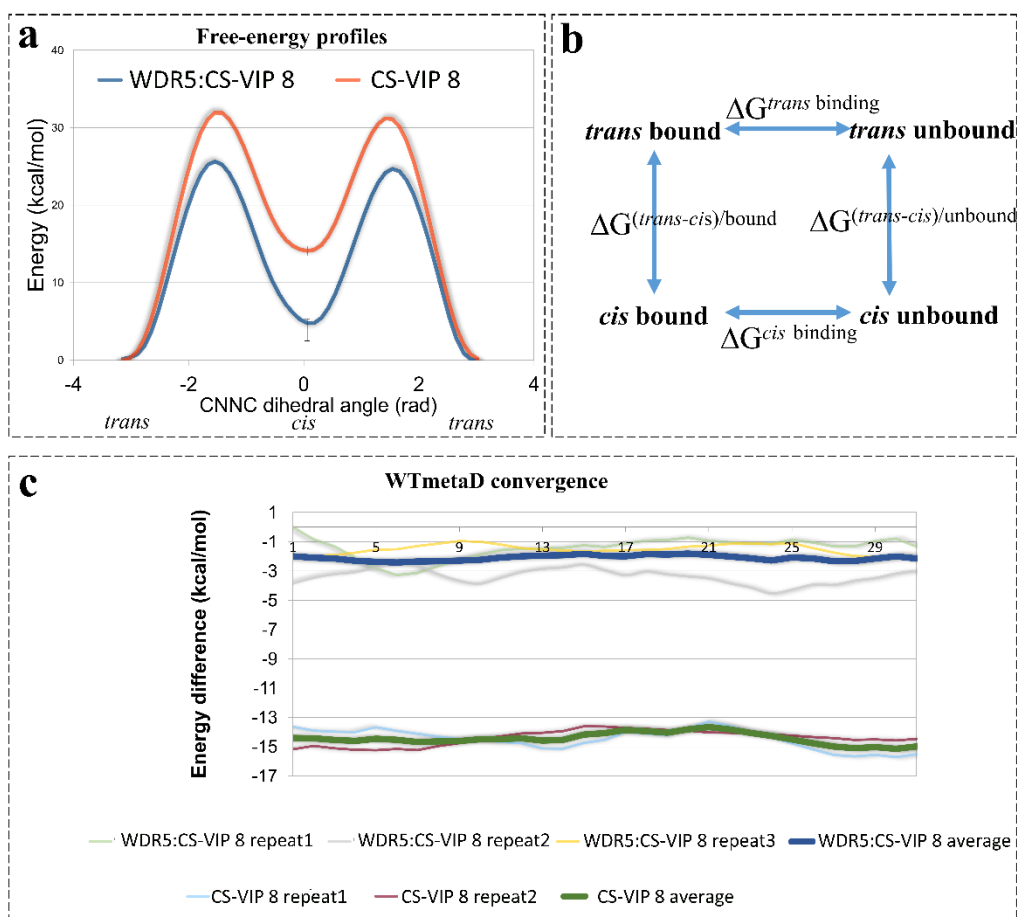


Figure S41. Represented results from WT-metaD approach. **a**, Two representatives of free energy profiles associated with *cis/trans* isomerization of WDR5:CS-VIP 8 complex (blue line) and free CS-VIP 8 (orange line) plotted against azobenzene's CNNC torsional angle. Error bar lines indicate standard deviation of measurements of free energy difference between *cis* and *trans* state across all replicas. **b**, Thermodynamic cycle used for calculation of $\Delta\Delta G^{(trans-cis)_{binding}}$ ($\Delta\Delta G^{(trans-cis)_{binding}} = \Delta G^{(trans-cis)_{unbound}} - \Delta G^{(trans-cis)_{bound}} = -14.09 - (-4.75) = -9.35$ kcal/mol). **c**, Convergence of free energy difference between *cis* and *trans* states for WDR5:CS-VIP 8 complex (blue line) and free CS-VIP 8 (orange line) in last 40 ns of WT-metaD runs.

2.12.2 Supplementary note N2

SASA calculation:

In order to gain atomistic insight into dynamical behavior of WDR5:*cis*-CS-VIP 8 complex observed during HDX, molecular dynamic (MD) simulations were performed on four different systems: System **A**: apo WDR5; System **B**: WDR5:*cis*-CS-VIP 8; System **C**: MLL1 peptide-WDR5-RbBP5 peptide; System **D**: *cis*-CS-VIP 8-WDR5-RbBP5 peptide.

HDX experiments indicated WDR5 residues 220-230 as significant for the dynamical behavior of WDR5 alone, as well as MLL1 complex. While these residues are directly involved in interaction with RbBP5, result of HDX performed on MLL1 complex are expected. In order to corroborate the potential mechanism of deuterium exchange involved in WDR5 structures alone outside the binding pocket, differences in solvent accessible surface areas (SASA) calculated on hydrogens from peptide bonds of

WDR5 for systems A and B were compared. Results revealed almost the same regions as detected by HDX outside the MLL1 binding site (Figure S42). However, the trend of the observed change was in contrast to experimental data for some of the residues. Reasons for these deviations could be found in a recent review regarding HDX/MD coupling.¹⁰

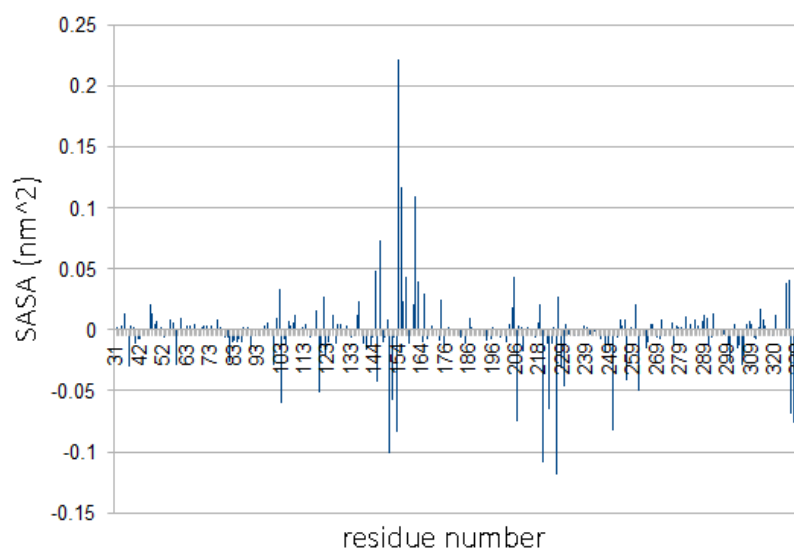


Figure S42. Difference in SASA changes observed through MD trajectory for System B - System A (y axis) versus residue ID (x axis).

2.12.3 Supplementary note N3

Perturbation-response scanning (PRS):

PRS analysis gained popularity in the last few years in analysis of allosteric influences that each residue has on all other residues during external force perturbation, such as ligand binding. Results of the PRS are usually interpreted through sensor and effector residue profiles. Effector residues are the ones, whose perturbation results in “conformational” changes in distant sites of the proteins (sensor residues).^{11–13}

Covariance matrices generated from MD simulations were used for PRS analysis. Results of PRS analysis performed on the systems A (apo WDR5, see Methods) and B (WDR5:*cis*-CS-VIP **8**, see Methods) expressed through B-factors indicated that the most of the effectors are mapped in the MLL1 binding site and the “main pore” of the WDR5, while most of the sensors are mapped on and around RbBP5 binding site. Pairwise comparison of effector profiles (Figure S43 A and C) indicates that strong communication in System B is engaged through the central pore of WDR5 as a result of the ligand binding. Sensor profiles indicate slight changes around RbBP5 binding site (Figure S43 D and C). These results were a first indication of allosteric communication between MLL1 and RbBP5 sites in WDR5. Additional analysis through residues network interactions formalism (**Supplementary note N4**) was performed in order to gain detailed insights into mechanisms behind allosteric signal transduction.

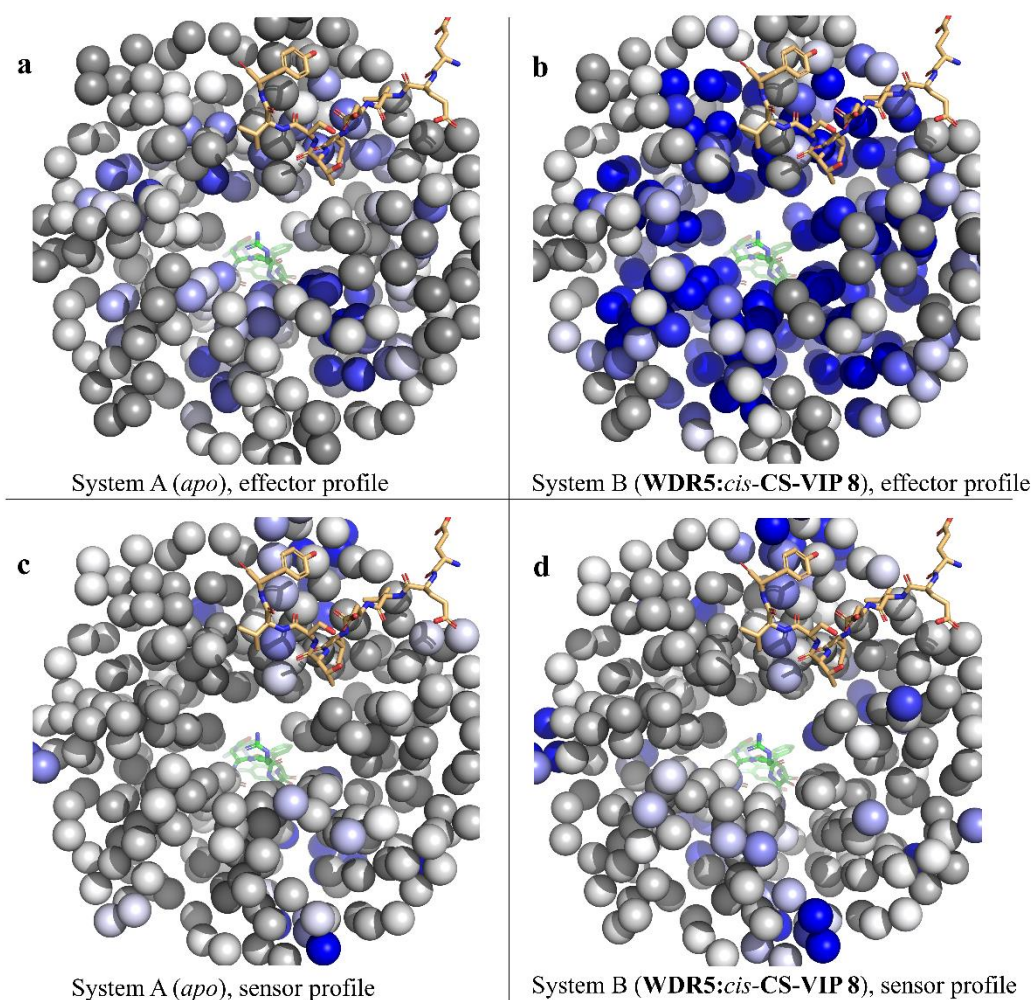


Figure S43. Effector residues (a and c) and sensor residues (b and d) profiles expressed as B-factor on C α atoms for System A and System B. RbBP5 peptide (tan sticks) and *cis-CS-VIP 8* (green sticks) are depicted in all figures as indicators of RbBP5 and MLL1 binding sites. Colour scale grey-white-blue indicates residues with low B-values as grey, and blue with high B-values. Colouring range was 0-0.2.

2.12.4 Supplementary note N4

A. Principal component analysis (PCA) and conformational analysis of obtained trajectories:

Results of the PCA analysis imply different conformational behavior of systems during the course of MD simulations. The largest difference was observed between systems A-B (A: apo WDR5, B: WDR5:*cis-CS-VIP 8*; see Methods) and systems C-D (C: MLL1 peptide-WDR5-RbBP5 peptide; D: *cis-CS-VIP 8*-WDR5-RbBP5 peptide; see Methods). While systems C and D are well localized on the plane of PC1 and PC2 projections, indicating conformationally more stable systems trapped in deeper free-energy basins, systems A and B are more delocalized and only partially overlapped (Figure S41). However, during RMSD analysis and visualization of the trajectories for systems C and D, partial unbinding of the RbBP5 peptide was detected, pointing on differences in conformational behavior in WDR5 induced by different ligands present in MLL1 site of WDR5 (Figure S44). This is in accordance

with differences between systems C and D observed in PCA analysis. RMSD calculations indicated that all trajectories obtained for Systems C and D were well converged (Figure S44). According to the comparison of cross-correlation matrices of systems C and D, larger anti-correlated motions in system D around residues 220-230 (RbBP5 binding site) further argue allosteric effects of *cis*-CS-VIP 8. Taken together, PCA and conformational analyses revealed that allosteric effects of *cis*-CS-VIP 8 resulted in different dynamical behavior, compared to allosteric effects of MLL1 peptide, which led to the partial unbinding of RbBP5 (Figure S45). These findings suggest that inhibition of MLL1 complex by competition with MLL1 on WDR5 may not be the only mechanism involved in disruption of MLL1 complex and further encourage usage of MD which can reveal different dynamical properties introduced by ligand binding.

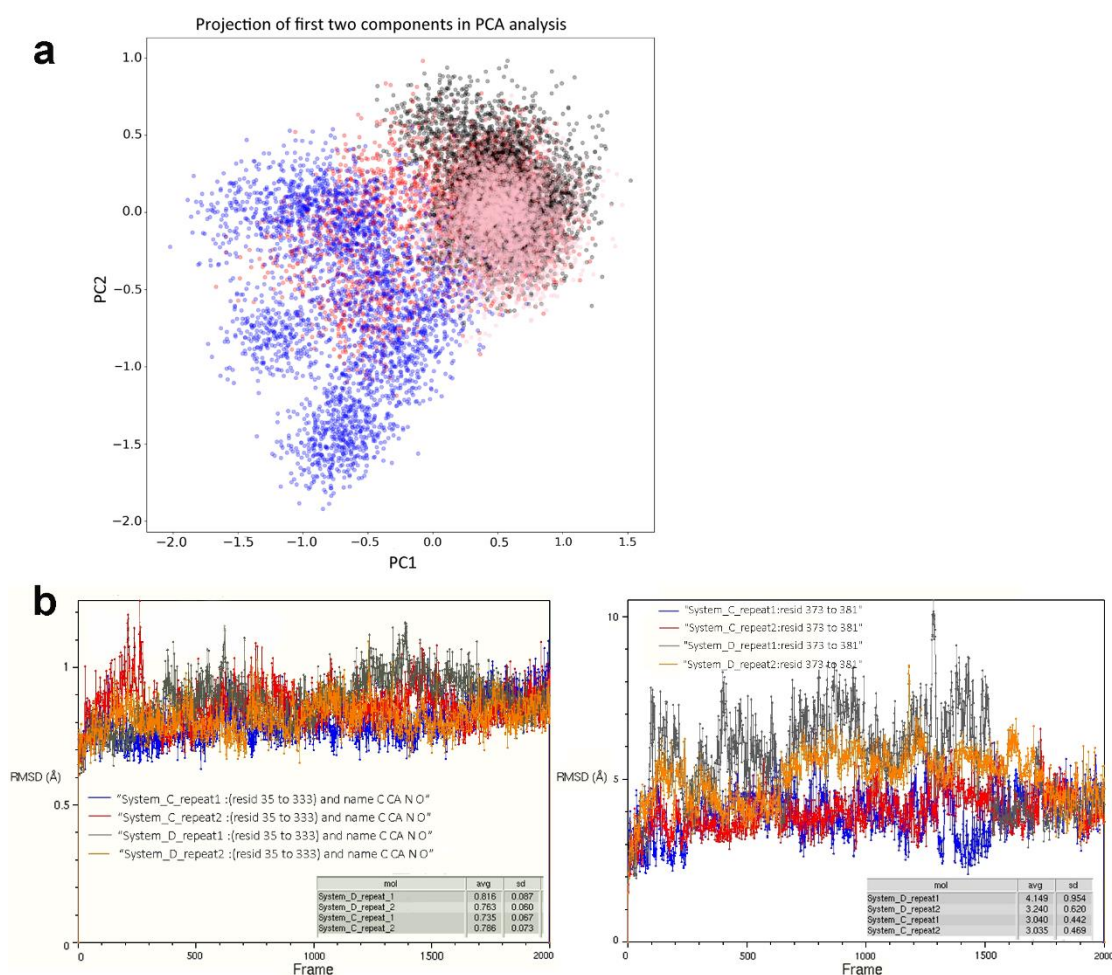


Figure S44. PCA blot and RMSD fluctuation detected in all repeats of MD simulations. **a**, Results of the PCA analysis (red- System A, blue- System B, pink- System C, black- System D); **b**, RMSD of the Systems C and D calculated for the backbone atoms of the whole WDR5 (left) and for the RbBP5 residues interacting with residues 220-230 of the WDR5 (right). Average RMSD values and standard deviations are presented in grey boxes. Systems C and D were simulated in two replicas, each 200 ns.

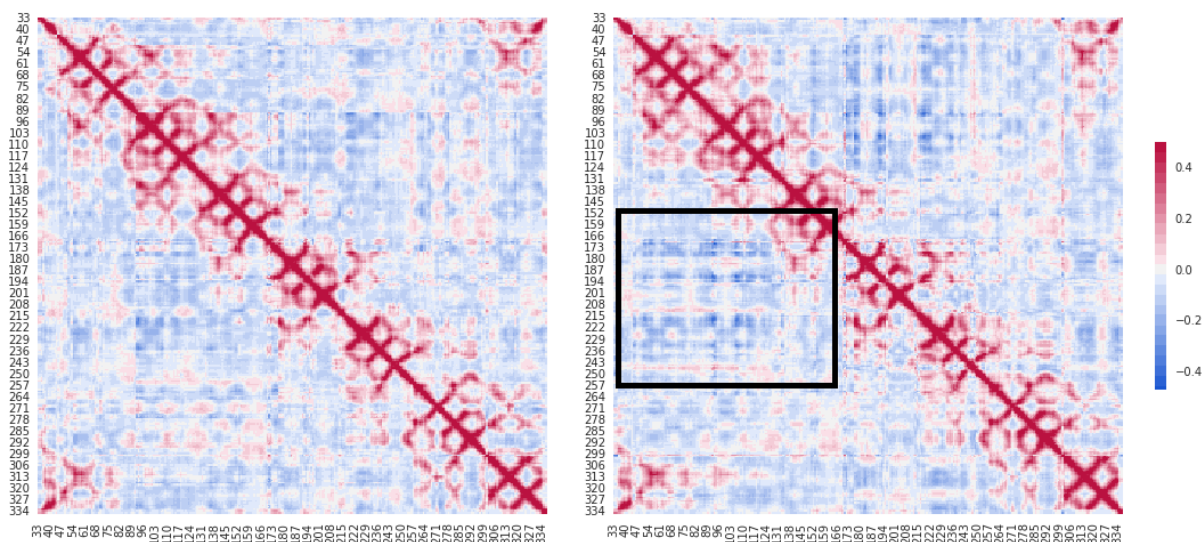


Figure S45. Cross-correlation matrices. For the system C (left) and the system D (right). Colouring scheme indicates anti-correlated motions in blue. The most prominent anti-correlated motions are labelled in the black square.

RMSD inspection of trajectories obtained for systems A and B also indicated well-converged systems (Figure S44), and further validated our simulation approach. In addition to PCA analysis, significantly lower level of conformational fluctuations expressed through average RMSD values and standard deviations (Figure S44 and Figure S46) in systems C and D, compared to systems A and B, was detected, indicating that binding of RbBP5 strongly affects dynamics of WDR5.

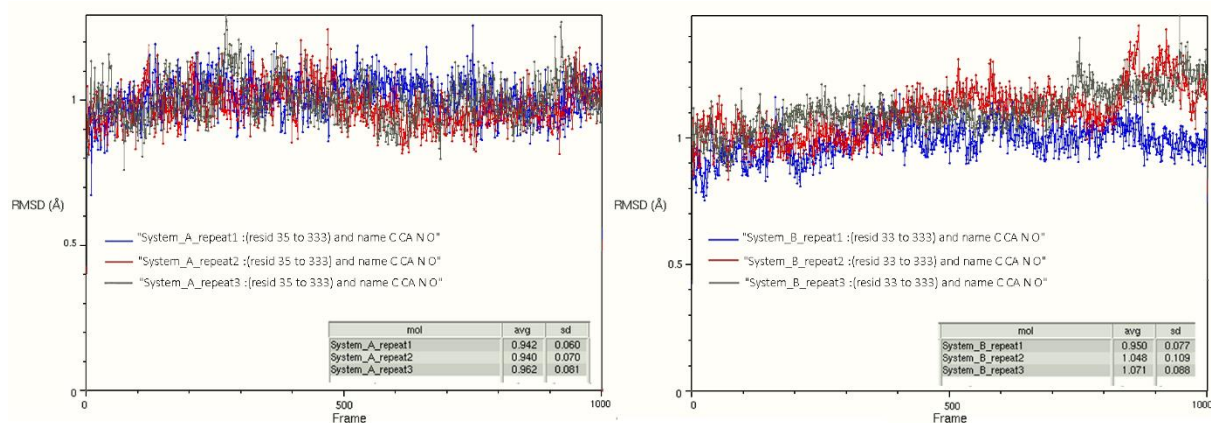


Figure S46. RMSD fluctuation detected in all repeats of MD simulations of the Systems. System A (left) and system B (right) calculated for the backbone atoms of the whole WDR5. Average RMSD values and standard deviations are presented in grey boxes. Systems A and B were simulated in three replicas, 100 ns each.

B. Network analysis:

Network analysis was utilized in order to gain deeper insight into conformational behavior observed during MD simulations (see above) and to corroborate results of PRS (see **Supplementary Note N3**). This analysis revealed the most significant communication pathways between distant sites of WDR5. Betweenness-centrality (BC) measurement was used to detect nodes that have great influence on what

flows in the network and have more control over the network. In all of simulated systems, Lys52, Asp92 and Lys221 were detected as the most important communication hubs in WDR5 (Figure S47). Because of their spatial localization (central pore of WDR5 (Figure 5d), these nodes were proposed to mediate in allosteric communication between MLL1 and RbBP5 binding sites on WDR5. According to the shortest path analysis, residue's Asp92 direct interaction with arginine from **CS-VIP 8** or MLL1 peptide is important for initiating communication through central pore. Signal is then propagated through salt bridge between Asp92 and Lys52, and further transmitted on the other side of the central pore (Figure 5d). Signal transmission to the other side of central pore and communication with RbBP5 binding site (residues 220-230 and 240-250) is facilitated by formation of water-mediated hydrogen bonding between Lys52 and Lys221. Additional analysis of shortest path revealed alternative, less important, pathway connecting residue Tyr191 with RbBP5 binding site across surface of WDR5 (Figure 5d).

In order to further investigate how the presence or absence of the inhibitor and/or RbBP5 affect the structure of residues interaction networks, and consequentially find the cause of observed differences in dynamical behaviour identified through PCA and conformational analysis, we performed pairwise comparisons of obtained networks in the terms of rewiring (Table S7). Comparing system A (apo WDR5) and B (WDR5:*cis*-**CS-VIP 8**), all of the highest BC nodes were detected to go significant change which is quantified in terms of higher values of D_n -score (rewiring metric). Namely, binding of the ligand induced significant rewiring of the interaction network, which was manifested through large D_n -scores. Additional analysis and visualization of trajectory confirmed that in apo system there is a significant difference/change in information flow compared to systems with ligands caused in part by uncoordinated motions of Asp92, which would be otherwise stabilized through direct interaction with ligands.

Comparing systems C and D, the largest rewiring was detected for residues around MLL1 binding site (Table S7) which indicates that differences in interaction patterns of **CS-VIP 8** and MLL1 peptide caused the observed differences in conformational behaviour of WDR5. Interestingly, detected rewiring of residue Asp92, as one of the most important residues for allosteric signal transduction through central pore of WDR5, could explain observed partial unbinding of RbBP5 during System D simulation. However, in terms of rewiring metrics, differences between systems C and D were less prominent compared to A/B (lower values of D_n -scores). Taken together, these findings indicate that, instead of obvious "conformational change" (which is in agreement with X-ray structures reported so far) introduced by inhibitor binding, subtle shifts in interaction patterns and rewiring of the residues interaction network of WDR5 play major role in allosteric communication between MLL1 binding site and RbBP5 binding site.

Additionally, rewiring analysis indicated that residues included in pathway connecting residue Tyr191 with RbBP5 binding site (Figure 5d) (Asp192, Arg196) are also affected by presence or absence of different ligands in MLL1 binding site (Table S7). Importance of Tyr191 in HMT activity of MLL1

complex was previously experimentally recognized through mutational analysis which is in accordance to simulation results.¹⁴

Overall, our simulations support the hypothesis that stabilization of MLL1 complex depends on WDR5's central pore communication between MLL1 binding site and RbBP5 binding site, initiated by arginine interaction with MLL1 binding site. This stabilization is facilitated by communication between Tyr191 and RbBP5 binding site.

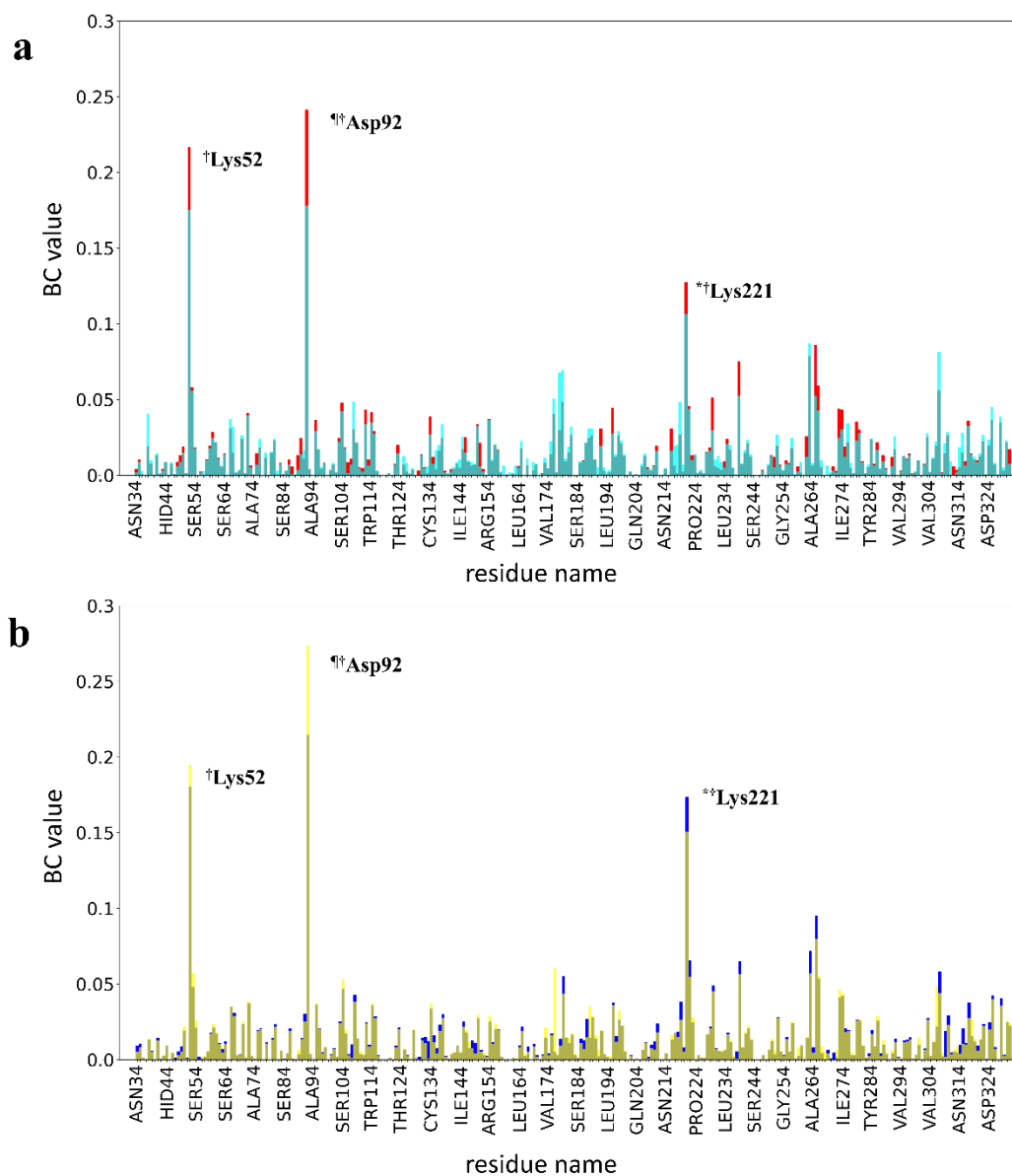


Figure S47. Pairwise comparison of BC values. Calculated for: **a**, System A (red bars) and system B (aqua bars). **b**, System C (yellow bars) and System D (blue bars). Overlapping regions in bar plot are indicated in teal (**a**) and mint green (**b**). Residues connected with the highest BC values are labelled and marked with symbols: *indicates residue from RbBP5 binding site; † indicates residues from central pore of WDR5; ‡ indicates residue from MLL1 binding site.

Table S7. Results of pairwise comparison of network models between different systems. N-terminal and C-terminal residues were omitted from analysis.

Systems A/B ^a		Systems C/D ^a	
Resid ID	D _n -Score	Resid ID	D _n -Score
His255	5.099489	Asp150	1.503232
His86	4.70898	Asp211	1.344799
His128	4.273141	Asp92 ^b	1.042917
Asp108 ^b	3.338209	Asp107 ^b	1.01655
Asn180	3.094606	Asn287	1.010103
Arg196 ^c	2.796166	Phe133 ^b	1.006474
Ser106 ^b	2.642996	Ser190 ^b	1.004777
Lys126	2.556442	Ser49 ^b	1.002467
Lys221 ^d	2.502327	Pro173 ^b	1.001611
Asp150	2.337324	Asn130 ^b	1.001167
His310	2.329052	Val48 ^b	1.0005
His170	2.285121	Leu194	0.572672
Asn265	1.916545	Asn265	0.559351
Pro311	1.824915	Arg196 ^c	0.549544
His178	1.703283	Asp192 ^c	0.542799
Asp92 ^b	1.680356	Glu313	0.536194
Glu292	1.65801	Ser91 ^b	0.524734
Lys162	1.631025	Asp66 ^b	0.517267
Lys52 ^b	1.628334	Ser64 ^b	0.515594
Ser175 ^b	1.592773	Ser148 ^b	0.514068
Leu167	1.575877	Ser148 ^b	0.514068
Leu111	1.571174		
Asn257	1.497048		
Lys239 ^d	1.48971		
Asp157	1.458518		
Ser148 ^b	1.310101		
Arg154	1.302673		
Thr237	1.261176		
Thr110	1.256864		
Trp156	1.234355		
Thr233	1.166917		
Tyr252 ^b	1.15467		
Leu194	1.145334		
Asp192 ^c	1.140995		
Glu58	1.121784		
Ile83	1.091371		
Ser190 ^{b,c}	1.091004		
Asp107 ^b	1.060644		
Pro173 ^b	1.0554		
Asn130 ^b	1.026998		

^a Due to larger differences in dynamical behaviour between systems A and B, different cut-off values were used: for systems A and B cut-off was 1; for Systems C and D cut-off was 0.5.

^b Residues included in MLL1 binding site.

^c Residues included in alternative pathway connecting MLL1 and RbBP5 binding sites.

^d Residues included in RbBP5 binding site.

5 References

1. Morris, M. C., Depollier, J., Heitz, F. & Divita, G. A Peptide Carrier for the Delivery of Biologically Active Proteins into Mammalian Cells. *Nat. Biotechnol.* **19**, 1173–1176 (2001).
2. Kabsch, W. Xds. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
3. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
4. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 1204–1214 (2013).
5. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
6. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
7. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. Ligplot - a Program To Generate Schematic Diagrams of Protein Ligand Interactions. *Protein Eng.* **8**, 127–134 (1995).
8. Biswas, M. & Burghardt, I. Azobenzene photoisomerization-induced destabilization of B-DNA. *Biophys. J.* **107**, 932–940 (2014).
9. Gaspari, R., Prota, A. E., Bargsten, K., Cavalli, A. & Steinmetz, M. O. Structural Basis of cis- and trans-Combretastatin Binding to Tubulin. *Chem* **2**, 102–113 (2017).
10. Huang, L., So, P. K. & Yao, Z. P. Protein dynamics revealed by hydrogen/deuterium exchange mass spectrometry: Correlation between experiments and simulation. *Rapid Commun. Mass Spectrom.* **33**, 83–89 (2019).
11. Atilgan, C. & Atilgan, A. R. Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. *PLoS Comput. Biol.* **5**, (2009).
12. General, I. J. *et al.* ATPase Subdomain IA Is a Mediator of Interdomain Allostery in Hsp70 Molecular Chaperones. *PLoS Comput. Biol.* **10**, (2014).
13. Dutta, A. *et al.* Cooperative Dynamics of Intact AMPA and NMDA Glutamate Receptors: Similarities and Subfamily-Specific Differences. *Structure* **23**, 1692–1704 (2015).
14. Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 (2006).

4.4.1 Author Contribution

The idea of achieving photoinhibition of MLL1 in zebrafish by using photoswitchable peptides was envisioned by Jun.-Prof. Dr. Olalla Vázquez. I, Lea Albert, synthesized the *o*F₄Azo photoswitch as well as the whole peptide library. I characterized the photoisomerization, expressed proteins (MLL1, WDR5, RbBP5, Ash2L, DPY30) and reconstituted the MLL1 complex. Furthermore, I conducted the FP-based, HMT-, cell-based, pull-down and in vivo zebrafish assays, as well as I performed crystal structure- and HDX interpretation and contributed to write the manuscript. Laura Werel and Prof. Dr. Lars-Oliver Essen solved crystal structures. Dr. Wieland Steinchen performed HDX analysis and interpretation and contributed to write the manuscript. Prof. Dr. Gert Bange performed HDX interpretation. Dr. Jatin Nagpal designed and initiated me in the in vivo experiments, which I performed myself. Prof. Dr. Soojin Ryu designed and interpreted the in vivo experiments. Dr. Katarina Nikolić, Nemanja Djoković and Dušan Ružić designed, conducted and interpreted MD simulations. Prof. Dr. Lars-Oliver Essen and Dr. Antoine Royant performed the UV-vis measurements of co-crystals at the cryobench. Prof. Dr. Lars-Oliver Essen further contributed to write the manuscript. Malte Hoffarth expressed WDR5, synthesized the *o*F₄Azo-moiety, which was used for incorporation into the initially evaluated CS-VIPs library, and set up preliminary co-crystal experiments. Initial introduction into the handling of suspension cell lines was provided to me by Dr. Jing Xu in the laboratory of Prof. Yali Dou. Jun. Prof. Olalla Vázquez conceived of the idea for the study, designed, directed and interpreted experiments, wrote the manuscript, and had the role of corresponding author, i.e., she was leading and/or involved in all the 14-role taxonomy for author.¹

Marburg,

Jun. Prof. Dr. Olalla Vázquez

Lea Albert

5 Conclusions and Outlook

The manuscripts of the cumulative part of this thesis address the utilization of light-controllable peptides, containing molecular photoswitches, for modulation of biological functions. While manuscript A presents a detailed overview about the state of the art of photoresponsive peptides used to modulate cellular processes, and puts this thesis into perspective, the remaining three manuscripts (B, C, D) report our contribution to this research field. With the work of this thesis, we provided proof of principle of a novel strategy for external control of epigenetic states. Thus, by targeting the protein-protein interaction (PPI) between the histone methyltransferase MLL1 and the protein WDR5, MLL1 methylation activity is reversibly inhibited. This was achieved by incorporation of the classical azobenzene 4-[(4'-aminomethyl)phenylazo]benzoic acid (AMPB) into the WDR5-targeting MLL1-peptide. To the best of our knowledge, we delivered the first example of photoresponsive peptides capable of controlling the activity of histone methyltransferases, and verified that these peptidomimetics act as optoepigenetic inhibitors via regulating the expression levels of MLL1-target genes. Once we corroborated the potential of photopharmacological inhibition of MLL1, we further advanced the photochemical properties of our PPI modulators by the design of novel visible-light switchable azobenzene amino acid building blocks: a cyclic bridged azobenzene (cAzo) and a tetra-*ortho*-fluoroazobenzene (*oF*₄Azo) moiety. By incorporation of these novel visible-light responsive switches into the same linear WDR5-targeting MLL1-peptide, we provided novel visible-light responsive peptides that could be applied in different contexts. In our case, although photoconversion of the *oF*₄Azo switch was amplified compared to the classical AMPB moiety, the improved photochromic properties did not lead to an increased affinity window between isomeric states. This finding emphasizes the challenges to design effective light-controllable PPI inhibitors and downplays the need of quantitative photoconversion to yield high differences in a biological environment. In an attempt to maximize the difference of the biological output through a different approach, we introduced the *oF*₄Azo switch into a WDR5-targeting cyclic peptide, based on the MM-401 scaffold. Thus, we synthesized a library of 14 conformationally strained visible-light photoswitches (CS-VIPs). Our best hit compound displayed drastically increased differences in WDR5-binding affinity between isomeric states (cyclic *oF*₄Azo peptide: 10-fold vs. linear *oF*₄Azo peptide: 3-fold and linear AMPB peptide: 5-fold difference according to FP-based assays). These results corroborate the hypothesis that utilization of conformationally strained peptides as scaffolds lead to increased functional differences upon illumination compared to their unstructured linear analogues. In addition, our cyclic probe shed light on the action mechanism of WDR5-MLL1 inhibitors via both hydrogen-deuterium exchange (HDX) experiments and MD-calculations. Thus, WDR5-binding of the CS-VIP led to total dissociation of MLL1 from the core complex. Furthermore, conformational/allosteric changes of the WDR5 β -blades upon binding of the cyclic inhibitor have been revealed, which caused altered interaction modes of WDR5 to RbBP5. Different RbBP5 positioning might promote the disability of the RbBP5-Ash2L dimer to stabilize the closed form of MLL1 SET1 domain, thereby contributing to MLL1 inhibition. Ultimately, we expanded the control of MLL1 complex disruption with our cyclic photoswitchable tool, to an *in vivo* system. Thus, for the first time, we demonstrated the optochemical control of hematopoietic function utilizing zebrafish as a model organism, thereby highlighting the potential of optoepigenetics *in vivo*. We are convinced that the content of this thesis will inspire other researchers in the use of visible-light photoswitches to advance the understanding of epigenetic mechanisms at the molecular level.

It could be envisioned to further apply our novel CS-VIP in other sophisticated biological contexts, similar to MM-401, which was used to reprogram cells to naive pluripotency.² Our probe would have

the advantage that it could work as a COMET (“chemo-optical modulation of epigenetically regulated transcription”) probe according to the publication of Reis et al.³ with the purpose of achieving high spatio-temporal control of the epigenome. To characterize inhibitor-induced changes in genome-wide H3K4 methylation, chromatin immuno-precipitation followed by illumina-based next-generations sequencing (ChIP-seq) could be carried out. Ultimately, an improved cell-permeable CS-VIP could be employed to study the effects on neural stem cell (NSC) development. Just recently, Delgado et al.⁴ postulated that maintenance of NSCs positional identity requires an epigenetic memory, dependent on MLL1. Loss of ventral identity, caused by MLL1 inhibition, resulted in neuron generation with the characteristics of dorsal NSCs.⁴ In this study, MM-401 was employed as MLL1 inhibitor, upon whose application, NSCs stopped producing neurons. To re-establish neuron production, MM-401 needed be washed out of the cell culture. Utilization of a light-switchable probe would have the benefit of spatio-temporal control, avoiding additional steps as washing and, ultimately, allowing in vivo applications.

One limitation of our best CS-VIP is the poor cellular uptake and the need of a carrier for biological function. Although macrocyclic peptides generally have great potential to target PPIs due to their larger sizes compared to small molecules, most of them are impermeable to the cell membrane, which complicates their application for intracellular targets.⁵ Next to grafting the cargo onto cell penetrating peptides (CPPs) or introducing positive charge to the side chains via e.g. guanylation, ideally the “Lipinskis rule of 5” (Ro5)⁶ should be followed to enable cell permeability. Besides keeping the molecular weight ideally below 500 g/mol, hydrophobicity would have to be carefully adjusted (octanol-water coefficient (LogP) should not exceed 5) and the number of hydrogen bonds would have to be reduced, e.g. via N^a-methylation.⁵ However, there are several orally available peptidic drugs, which do not obey the Ro5⁷ such as e.g. the immunosuppressive alisporivir.⁸ It has been hypothesized that a chameleon-like behavior – being able to change shape and polarity due to intramolecular hydrogen bonding, depending on solvent polarity – of such molecules is the cause for these Ro5 outliers.^{8,9} A statistical review revealed, that drug-like peptides mostly do not exceed 1200 g/mol and exert LogP values from -5 to 8.⁸ In addition, solubility of the compound plays an important role and affects oral bioavailability.⁹ Although our best CS-VIP obeys these extended rules (molecular weight ~925 g/mol, LogP ~2.78), the size is still considered quite big, which makes it challenging to achieve a cell-permeable peptide that is able to engage the chameleon-like behavior while at the same time maintaining high target affinity. In addition, compound solubility should be increased. Thus, a careful and extensive amino acid scan would need to be applied in order to balance these different effects.

Besides that, the photoswitchable moiety of the inhibitor could be further improved to allow isomerization with red or even near-infrared light leading to deeper penetration potential to allow in vivo studies in mammals. Several studies already demonstrated the possibility of (near-infra)red-light isomerization of azobenzene photoswitches by introduction of bulky electron-rich substituents, as methoxy- chloro- or bromo-substituents in *ortho* position.¹⁰ In addition, combinations of bromo-/methoxy-,¹¹ as well as the substitution on *o*F₄Azo of two, out of the four, fluorines by chloride enables red-light isomerization.¹² Near-infrared isomerization is achieved for azobenzenes that exhibit methoxy substituents at all four *ortho* and two *meta* positions via fusion to dioxanes and methoxygroups.¹³ However, these photoswitches, yet, display other limitations. Although, in comparison to azobenzene *para*-substitution, the *ortho*-substitution already markedly slows down thermal *cis* to *trans* relaxation,¹⁴ they still display relatively short half-lives of the *cis* isomer (range of several minutes to hours).¹⁰ In addition, intracellular stability against glutathione (GSH) is not yet given for every of the proposed red-light switching azobenzenes, which is an essential element for cellular applications. Therefore, the “perfect” molecular photoswitch has, yet, to be developed.

6 Further Developments

In addition to the results represented in the above displayed manuscripts of the cumulative part (section 4), several other experiments have been performed during this doctoral thesis. Supplementary performed experiments will be detailed in the following sections.

6.1 Additional Cell Viability Assays

In connection with the submitted manuscript D (MD) in the cumulative part of section 4.4, the effect of cyclic peptide candidate **MD-CS-VIP 8** on cell viability of MOLM-13 was evaluated in more detail. As illustrated in manuscript D, quite high concentrations had to be utilized, in order to observe an efficient repression of cell viability and significant differences between isomeric states (**MD-CS-VIP 8** for 3 d: 200 μ M *trans*: 37%, *cis*: 0%, 100 μ M *trans*: 57%, *cis*: 26%; see manuscript D). Use of such high concentrations though, could indicate a different toxicity mechanism than the specific MLL1 inhibition. Therefore, we performed a control experiment, where cell viability of MOLM-13 cells was already assessed after 3 h incubation. As the toxicity mechanism through inhibition of MLL1 would take longer than 3 h to show an effect on cell viability, evaluation after such a short incubation time will yield hints whether or not the observed toxicity after 3 days is only caused by binding of **MD-CS-VIP 8** to WDR5. Toxicity due to MLL1 inhibition would be implied, when an increase in toxicity from 3 h to 3 days would be detected. In our experiments we observed already high toxicity of **MD-CS-VIP 8** after 3 h, (Figure F-1, **MD-CS-VIP 8** for 3 h: 200 μ M *trans*: 50 %, *cis*: 5 %, 100 μ M *trans*: 57 %, *cis*: 48 %). However, at 100 μ M no great difference between isomers was yet detectable and the *trans* isomer did not increase its toxicity from 3 h to 3 days, while the *cis* counterpart did, suggesting an interplay of different toxicity mechanisms. This confirmed our hypothesis that only the *cis* isomer is able to repress cell viability in an MLL1 dependent manner, while toxicity of the *trans* isomer results from different mechanisms.

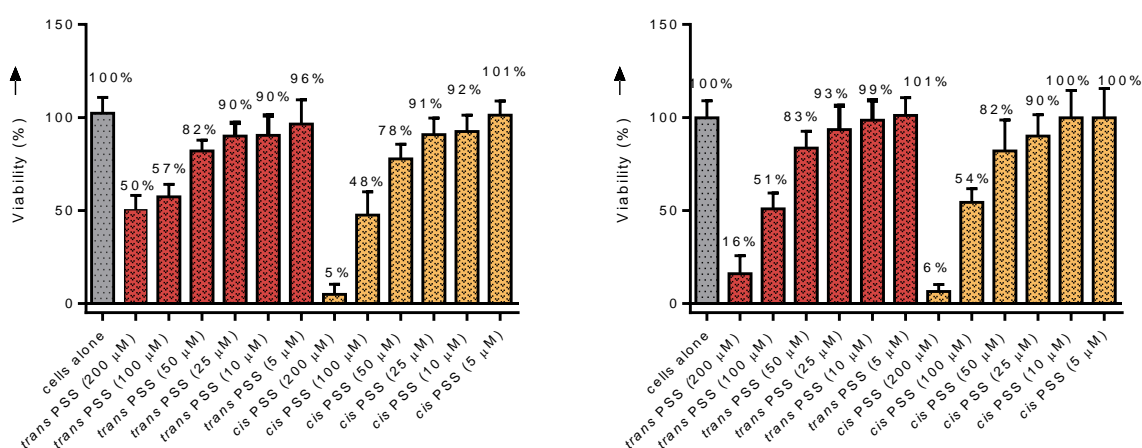


Figure F-1. Effect on MOLM-13 cell viability after 3 h incubation. Left: Incubation of cells with **MD-CS-VIP 8**. Right: Incubation of cells with **MD-CS-VIP 14**.

In manuscript D, we further utilized the negative control non-WDR5 binding peptide **MD-CS-VIP 14** containing the σ F₄Azo and exchanged stereo information compared to **MD-CS-VIP 8** as well as the L-Arg was exchanged by a D-Lys. This D-Lys variant **MD-CS-VIP 14** was applied to rule out that the observed effects only result from the toxicity of the σ F₄Azo moiety itself (**MD-CS-VIP 14** for 3 d: 200 μ M

trans: 9%, *cis*: 1%, 100 μ M *trans*: 42%, *cis*: 40%; see manuscript D). We, too, evaluated cell viability after 3 h incubation with **MD-CS-VIP 14**, which demonstrated that, despite the high toxicity of the negative control, no significant difference between isomeric states could be observed (Figure F-1, **MD-CS-VIP 14** for 3 h: 200 μ M *trans*: 16%, *cis*: 6%, 100 μ M *trans*: 51%, *cis*: 54%). In addition, toxicity of **MD-CS-VIP 14** did not increase from 3 h to 3 days. This implies that toxicity of the negative control peptide **MD-CS-VIP 14** is only caused by non-MLL1 dependent effects like membrane permeabilization. This corroborates that only the *cis* isomeric state of **MD-CS-VIP 8** is able to affect cell viability in a MLL1-dependent fashion. Experiments, where we included an additional cell washing step after cargo transfection, verified this proposal (see discussion in manuscript D).

Along these lines, the effects of **MD-CS-VIP 8** and **MD-CS-VIP 14** on cell viability on the human leukemic cell line, U937 (cell type: histocytic lymphoma) were assessed, too. After verifying that both DMSO and pep-1 (**MD-15**) controls hardly affect viability (Figure F-2), **MD-CS-VIP 8** and the negative control: the D-Lys variant **MD-CS-VIP 14**, were evaluated after 3 h and 3 days. As displayed in Figure F-3, **MD-CS-VIP 8** only exhibited very low toxicity, even at high concentrations. In contrast, the D-Lys variant **MD-CS-VIP 14** (Figure F-4) presented high toxicity of the *cis* PSS at 200 μ M concentration, while its *trans* PSS displayed increased toxicity down to 100 μ M in comparison to its own *cis* PSS and to both isomeric PSSs of **MD-CS-VIP 8**. As it has been demonstrated that the D-Lys variant **MD-CS-VIP 14** is not binding to WDR5 (see MD, section 4.4), this observed toxicity must result from a different mechanism than MLL1 inhibition. One explanation could be enhanced membrane permeabilization of the *trans* isomer of D-Lys variant **MD-CS-VIP 14** leading to decreased cell viability. In addition, for both CS-VIPs, toxicity did not increase from 3 h to 3 days, indicating that cell viability is not repressed in a MLL1-dependent manner. Thus, compared to the MV4-11 and MOLM-13 cells, U937 cells were not as suitable for our system. As it was demonstrated, that the carrier efficiency of pep-1 **MD-15** is altering for different cell types (MDCK, HEK293, HeLa, and Cos-7),¹⁵ this might be an explanation for the observed toxicity variations. Of note, these preliminary experiments using U937 cells have only been performed once, and it would need to be repeated as three independent experiments, at least, to draw accurate conclusions. Experimentally, cell assays were performed as described in the methods section and supporting information of the manuscript D in section 4.4.

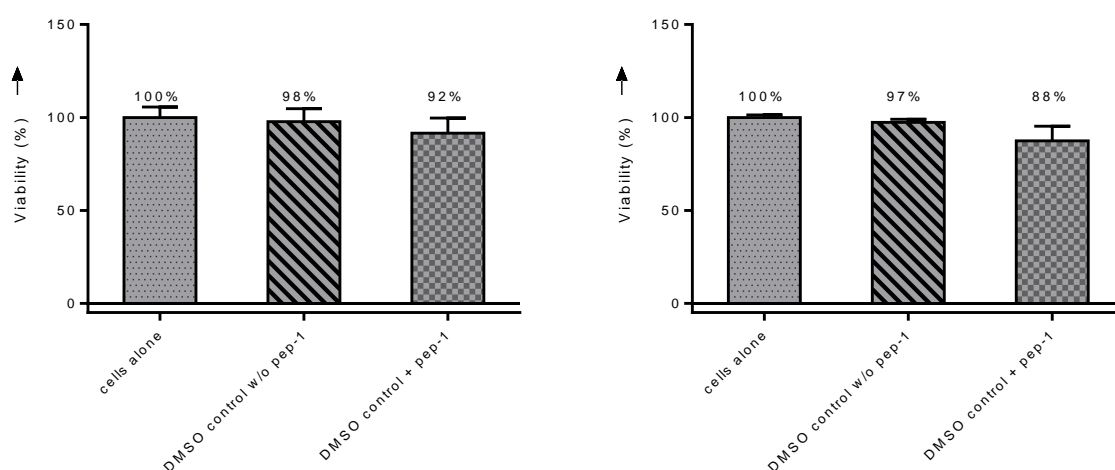


Figure F-2. U937 cell viability of DMSO/pep-1 controls. Left: Cell viability after 3h. Right: Cell viability after 3 days.

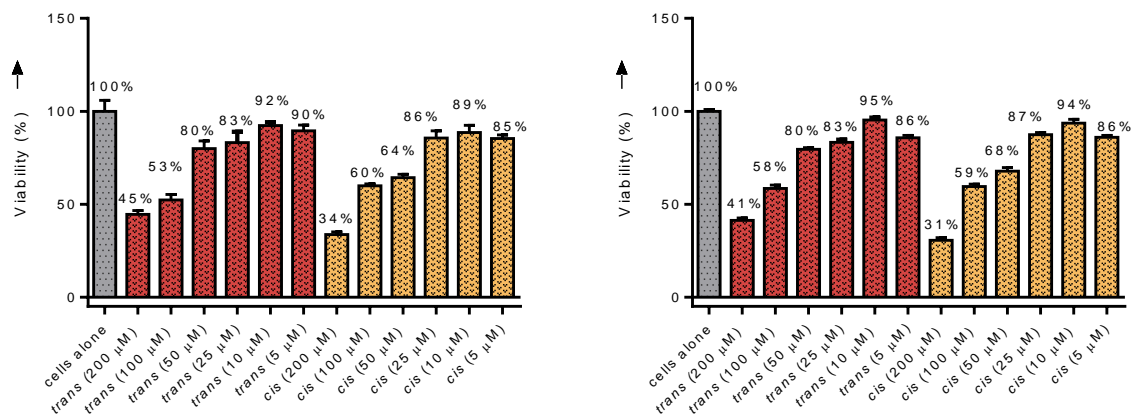


Figure F-3. Effect on U937 cell viability of cyclic peptide **MD-CS-VIP 8**. Left: Cell viability after 3h. Right: Cell viability after 3 days.

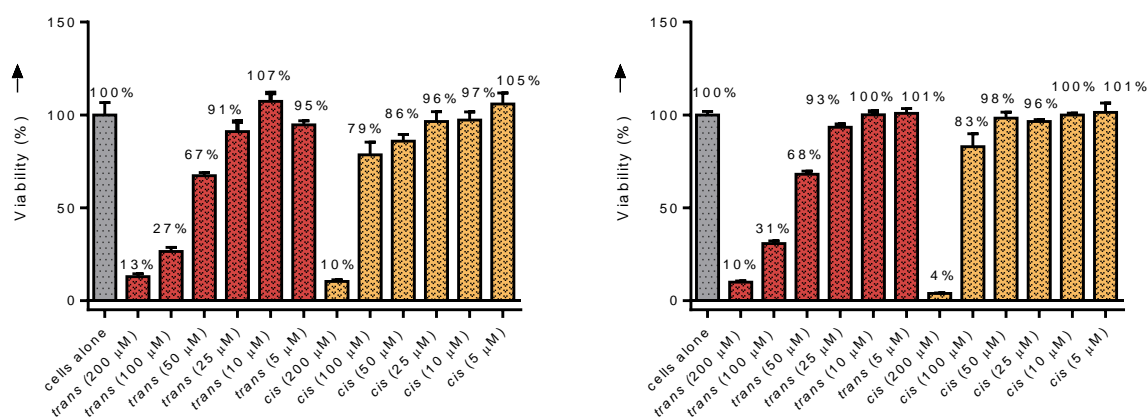
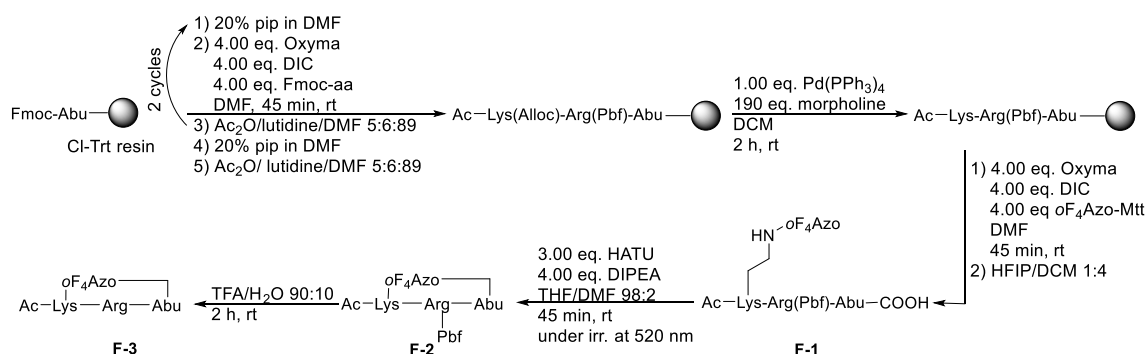


Figure F-4. Effect on U937 cell viability of cyclic D-Lys variant **MD-CS-VIP 14**. Left: Cell viability after 3h. Right: Cell viability after 3 days.

6.2 Synthesis of Additional Cyclic CS-VIP

In connection with the submitted manuscript D (MD) in the cumulative part of section 4.4, another CS-VIP peptide was synthesized (**F-3**). This peptide was prepared utilizing easily available L-amino acids, instead of the later employed D-amino acids. Synthesis was performed as in the Methods part of MD (Scheme F-1).



Scheme F-1. Synthesis route of cyclic peptide **F-3**.

In addition, competitive fluorescence polarization (FP)-based assays were performed too (Table F-1 and Figure F-5). The evaluation of binding affinities of both, *trans* and *cis* isomeric PSS of **F-3X**, to WDR5 displayed lower binding than the *oF*₄Azo incorporated best candidate cyclic peptide **MD-CS-VIP 8 (F-3)**: (K_i (*trans*) = 92.0 nM; K_i (*cis*) = 125.0 nM) vs. **MD-CS-VIP 8**: K_i (*trans*) = 79.0 nM; K_i (*cis*) = 8.00 nM), as well as decreased differences of the affinity window between both PSSs of the cyclic peptide **F-3** (ratio **MD-CS-VIP 8** = 9.88-fold; ratio **F-3** = 1.36-fold).

Table F-1. IC₅₀ and K_i values of **F-3** to WDR5. Mean values of 2 independent experiments.

peptide	IC ₅₀ [μM] (PSS 405 nm)	IC ₅₀ [μM] (PSS 520 nm)	K _i [nM] (PSS 405 nm)	K _i [nM] (PSS 520 nm)	Ratio
F-3	5.90 ± 0.31	8.03 ± 0.25	92.0 ± 5.00	125.0 ± 4.00	1.36

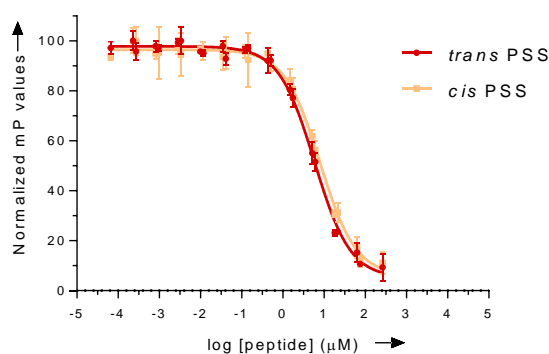
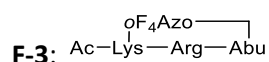


Figure F-5. Fluorescence polarization-based assays of both isomeric PSSs of **F-3**, displaying dose-response curves of their binding to WDR5. Mean of two independent experiments.

Experimental Procedure

Synthesis of Cyclic Peptide **F-3**:



For the lineal precursor of **F-3**, Arg(Pbf) was employed. Cyclization was performed in a 15.7 μmol scale as detailed in MD. After purification, the 1 × TFA salt product (0.8 mg, 0.911 μmol, 6%) was obtained as an orange solid. **HPLC**: t_R = 14.18 min (*trans*); 16.37 min (*cis*). Purity = 99%. Formula: C₃₃H₄₁F₄N₁₁O₆. **HRMS-ESI⁺ (m/z)**: [M+H]⁺ calcd.: 764.3250; found: 764.3254.

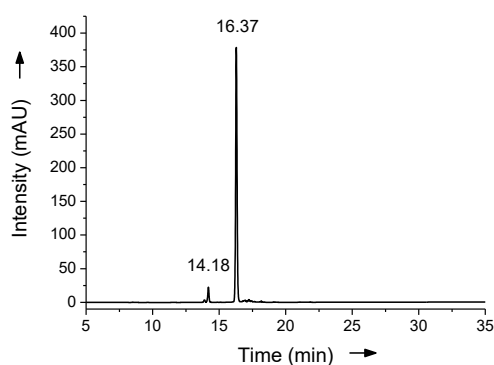
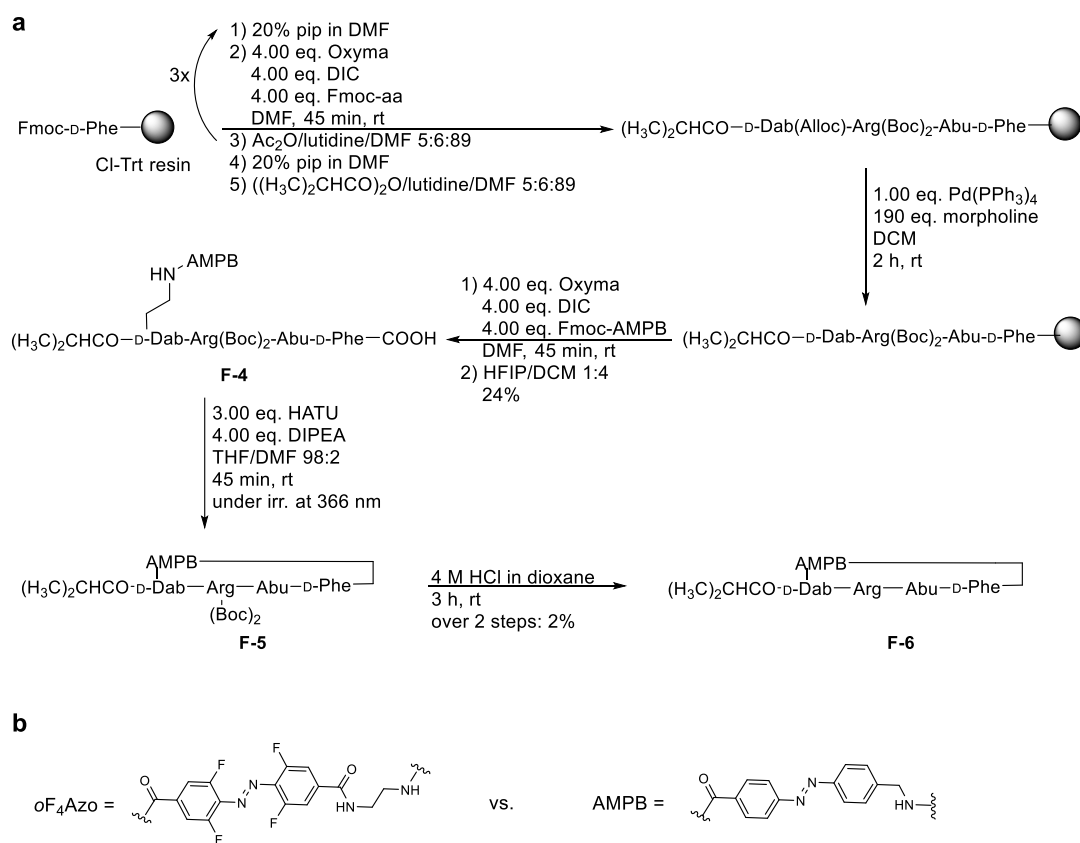


Figure F-6. HPLC chromatogram of purified compound **F-3**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm.

6.3 Classical Azobenzene AMPB Introduced into Cyclic Peptide

In the submitted manuscript D (MD) in the cumulative part of section 4.4, it was demonstrated that the cyclic peptide **MD-CS-VIP 8**, carrying the *o*F₄Azo switch was the most promising candidate for in vivo photoinhibition of MLL1. To directly compare the *o*F₄Azo and AMPB ([4-aminomethyl]phenylazobenzoic acid) photoswitches, the latter **MB-23** was also introduced into the cyclic peptide backbone. The synthesis of the AMPB-containing cyclic peptide **F-6** was performed as in the Methods part of MD (Scheme F-2a). The only difference was, that irradiation to the *cis* PSS, before cyclization, was performed at 366 nm for the AMPB cyclic peptide.

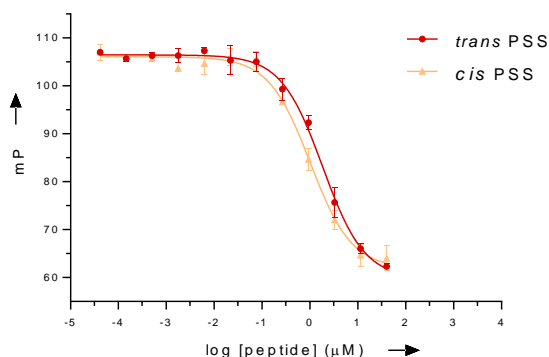
The evaluation of preliminary — i.e. only performed once — binding affinities of cyclic AMPB peptide **F-6**, to WDR5 using fluorescence polarization (FP)-based assays displayed lower interaction than for the *o*F₄Azo incorporated cyclic peptide **MD-CS-VIP 8** (AMPB peptide **F-6**: (K_i (*trans*) = 27.9 nM; K_i (*cis*) = 14.9 nM) vs. *o*F₄Azo peptide **MD-CS-VIP 8**: K_i (*trans*) = 79.0 nM; K_i (*cis*) = 8.00 nM), as well as drastically decreased differences of the affinity window between both PSSs (ratio **MD-CS-VIP 8** = 9.88-fold; ratio **F-6** = 1.87-fold; Table F-2, Figure F-7). This might partly be due to the different linker lengths of the two respective azobenzene moieties (Scheme F-2b). As discussed in manuscript C (section 4.3), the linker lengths of the azobenzene moiety has a significant influence on binding affinity. While for linear analogues, the shorter linker of the AMPB moiety is advantageous over the longer *o*F₄Azo linker, for the cyclic derivatives the trend is inverted. As the effect of linker length increases significantly when incorporated in a cyclic system, the difference in binding affinity is also amplified.



Scheme F-2. a, Synthesis route of AMPB incorporated cyclic peptide **F-6**. b, Direct comparison between *o*F₄Azo and AMPB structure and linker length.

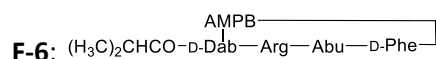
Table F-2. IC₅₀ and K_i values of **F-6** to WDR5. Values from one independent preliminary experiment.

peptide	IC ₅₀ [μM] (PSS 405 nm)	IC ₅₀ [μM] (PSS 520 nm)	K _i [nM] (PSS 405 nm)	K _i [nM] (PSS 520 nm)	Ratio
F-6	1.84	1.01	27.9	14.9	1.87

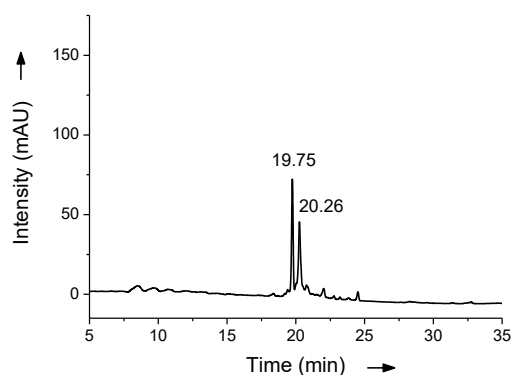
**Figure F-7.** Fluorescence polarization-based assays of both isomeric PSSs of cyclic AMPB peptide **F-6**, displaying dose-response curves of their binding to WDR5. Preliminary experiment, which was performed once.

Experimental Procedure

Synthesis of Cyclic AMPB Peptide:



Cyclization was performed in a 4.88 μmol scale as detailed in manuscript D, and the Arg(Boc)₂ side chain was deprotected using 4 M HCl in dioxane. After final purification, the product as 1 × TFA salt (0.10 mg, 0.110 μmol, 2%) was obtained as an orange solid. **HPLC:** $t_R = 19.75$ min (*trans*), 20.26 min (*cis*). Purity = 92%. Formula: C₄₁H₅₃N₁₁O₆. **HRMS-ESI⁺ (m/z):** [M+H]⁺ calcd.: 796.4253; found: 796.4249.

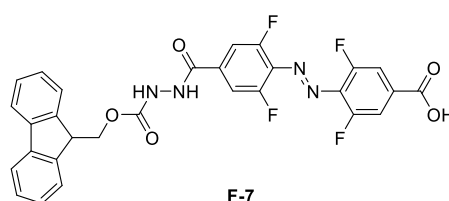
**Figure F-8.** HPLC chromatogram of purified compound **F-6**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-75% B monitored at 220 nm.

6.4 Synthesis of Fmoc-Hydrazine-*o*F₄Azobenzene

In manuscript C (MC) in the cumulative part of section 4.3, the synthesis of the Mtt-*o*F₄Azo building block **MC-2** was described. In addition to such Mtt-ethylenediamine derivative, the Fmoc-hydrazine analogue **F-7**, was synthesized too. However, due to the instability of the *o*F₄Azo moiety against Fmoc-deprotection conditions (section 4.3), only the Mtt-monomer was utilized for solid phase peptide synthesis (SPPS). The Fmoc-hydrazine-*o*F₄Azo **F-7** was sent to the collaboration partner, Prof. Dr. Luis Álvarez Cienfuegos Rodríguez to explore its abilities in formation of hydrogels as well as the possibility to photocontrol such process. Synthesis of the Fmoc-hydrazine-*o*F₄Azo **F-7** building block was performed according to the synthesis procedure for **MC-2**, described in manuscript C, up to the diacid-*o*F₄azobenzene **MC-9**. **MC-9** was then coupled to Fmoc-hydrazine, as described below.

Experimental Procedure

Synthesis of Fmoc-hydrazine-*o*F₄Azo **F-7**:



The diacid-*o*F₄azobenzene **MC-9** (465 mg, 1.36 mmol, 1.50 eq.) was suspended in THF (35.0 mL). HATU (345 mg, 0.907 mmol, 1.00 eq.) and DIPEA (316 μ L, 1.81 mmol, 2.00 eq.) were added and the solution was stirred at r.t. After 15 min, Fmoc-hydrazine (344 mg, 1.09 mmol, 1.20 eq.) in THF (2.00 mL) was added and the solution was stirred for 3 h at rt. Afterwards, the solvent was removed under reduced pressure and the crude was purified by flash column chromatography (DCM/MeOH 45:1) to yield the desired product **F-7** (86.8 mg, 0.150 mmol, 77%) as a red solid. To obtain the pure NMR spectra, a small amount was further purified by RP-HPLC (5-95% MeCN in 30 min). **TLC**: R_f = 0.09 (DCM/MeOH 9:1). **¹H-NMR (300 MHz, DMSO, δ)**: 13.91 (s, 1H, OH), 10.71 (s, 1H, NH), 9.61 (s, 1H, NH), 8.00-7.90 (m, 2H, 2 \times CH_{ar}), 7.83-7.80 (m, 3H, 3 \times CH_{ar}), 7.77-7.68 (m, 3H, 3 \times CH_{ar}), 7.45-7.34 (m, 2H, 2 \times CH_{ar}), 4.43 (d, 2H, ³J = 4.4 Hz, *o*CH₂CH), 4.30 (t, 3H, ³J = 4.3 Hz, CHCH₂O) ppm. **¹³C-NMR (75 M, Hz, DMSO, δ)**: 164.6(CO), 162.8 (CO), 156.1 (CO), 155.5 (C_{ar}F), 153.4 (C_{ar}F), 151.5 (C_{ar}F), 149.5 (C_{ar}F), 143.6 (2 \times C_{q,ar,Fmoc}), 140.8 (2 \times C_{q,ar,Fmoc}), 136.2 (C_{q,ar}), 134.9 (C_{q,ar}), 132.9 (C_{q,ar}), 132.4 (C_{q,ar}), 127.8 (2 \times C_{ar,FmocH}), 127.2 (2 \times C_{ar,FmocH}), 125.3 (2 \times C_{ar,FmocH}), 120.2 (2 \times C_{ar,FmocH}), 114.1 (C_{ar}HCF), 113.9 (C_{ar}HCF), 112.6 (C_{ar}HCF), 112.3 (C_{ar}HCF), 66.3 (CHCH₂O), 46.5 (CHCH₂O) ppm. **HRMS-ESI⁺ (m/z)**: calcd. for [M+Na]⁺ C₂₉H₁₈N₄F₄O₅Na, 601.1106; found, 601.1113.

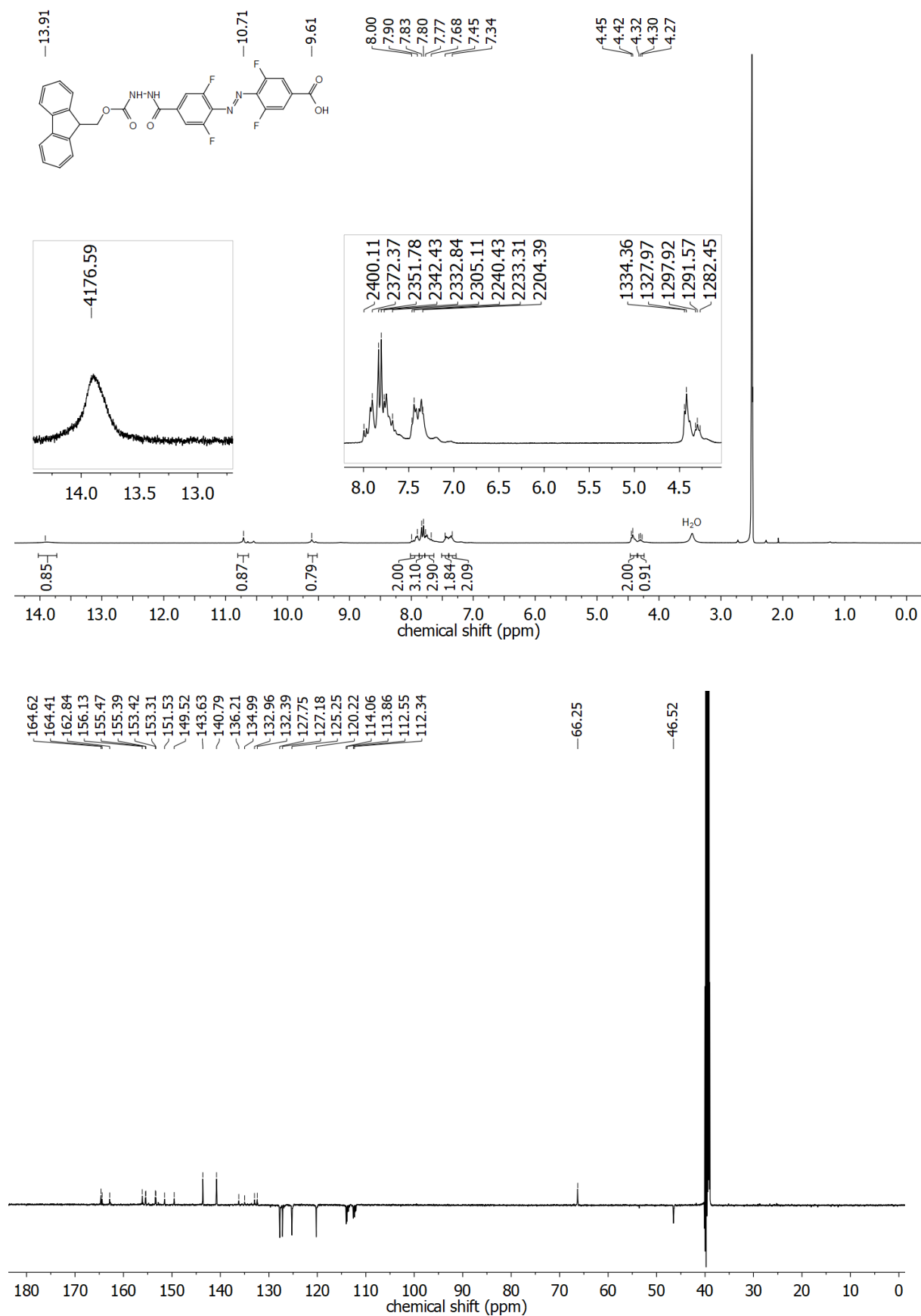


Figure F-9. Top: ¹H-NMR spectrum of the compound **F-7**. Insert: zoom-in of signals in the aromatic region to highlight the shifts in Hz units. Bottom: ¹³C-APT-NMR spectrum of compound **F-7**.

6.5 Cryo-Electron Microscopy (Cryo-EM) of the MLL1 Complex

Cryo-electron microscopy (cryo-EM) is a technique developed in the 1970's to visualize molecular assemblies, which are too flexible and large to be analyzed by X-ray crystallography.¹⁶ The advantage of cryo-EM over X-ray is that it investigates structural properties in frozen solutions. In addition to the fact that the need of sample crystallization is overcome in order to obtain structural insights, this means that the macromolecule of interest can be "trapped" at different stages of their conformational states to study dynamic interactions.¹⁶ Besides, much lower concentrations of complex are required for cryo-EM measurements. In 2017, J. Dubochet, J. Frank and R. Henderson have been awarded with the Nobel Prize for their work in developing cryo-EM.¹⁷

Technically, the specimen of interest is applied to an EM grid, which is frozen in liquid ethane in a liquid nitrogen bath. Due to the rapid temperature decrease, water turns into vitreous ice where crystal formation is neglected and disruption of the molecules avoided. Using a transmission electron microscope, a beam of electrons is applied to the frozen sample and the emerging scattered electrons pass through a lens in order to create a magnified image on the detector, from which their structures can be worked out.¹⁶ Before the specimen are applied to cryo-EM measurements, most often they are initially screened with the negative-staining technique at rt. There, the sample is mixed with a 1-2% solution of a heavy metal salt and applied to a carbon film on a copper-grid. This leads to background staining of molecules, outlining their shape, which yields high contrast in the EM image.¹⁶

The greatest limitation for cryo-EM is its relatively low spatial resolution (10-20 Å¹⁶) compared to X-ray crystallography. Although, during the last decade, higher resolution images have been obtained using cryo-EM.¹⁸ Thus, the structure of the ribosome has been solved, not only by X-ray,¹⁹⁻²² but also using cryo-EM up to a resolution of 2.5 Å, where water molecules, as well as all 60 rRNA modification sites could be identified, rendering this cryo-EM map suitable for drug design.²³ Still, such high resolutions are not always achievable, wherefore the interplay between both techniques, cryo-EM and X-ray, is desirable. When, for example, the structure of a macromolecular complex, that is not suitable for X-ray crystallography, is going to be determined by cryo-EM, it is advantageous if the X-ray structures of all the components are known to atomic resolution. This allows the modelling of such high-resolution components into the obtained low-resolution cryo-EM map, permitting deduction of atomic interactions.¹⁶

One important requirement, to be able to obtain structural information using cryo-EM, is that the sample solutions containing the specimen are of high homogeneity. To increase sample quality, the GraFix procedure can be applied to purify and stabilize macromolecular complexes.²⁴ There, a glycerol gradient is used, where the complex of interest is centrifuged into an increasing concentration of a chemical fixation reagent, such as glutaraldehyde, leading to cross-linking of the complex. Using cross-linking gradients is essential, because direct exposure of the macromolecule to the cross-linking agent could result in intermolecular fixation, increasing heterogeneity instead of reducing it. The centrifugal force applied by using GraFix is sufficient to disrupt aggregations in order to expose single complexes to the cross-linking agent, leading to mostly intramolecular fixation. Utilization of the GraFix protocol was demonstrated to stabilize macromolecules, prevent aggregation and reduce heterogeneity problems, thereby improving structure determination by single-particle cryo-EM.²⁴

During this thesis, we were performing first experiments in collaboration with Simone Prinz from the Max Planck Institute (MPI) of Biophysics Frankfurt of the group of Prof. Kühlbrandt. We obtained preliminary results of negative stained samples of the 5-mer MLL1 complex (MLL1, WDR5, RbBP5,

Ash2L, DPY30) imaged on an electron microscope (EM). Experimental details are explained below. After expression and purification of the single proteins of the MLL1 complex, as well as reconstitution and size exclusion chromatography (SEC) of the 5-mer complex, purity was confirmed by SDS-gel analysis, depicted in Figure F-10, and 5-mer complex containing fractions were utilized for further evaluation.

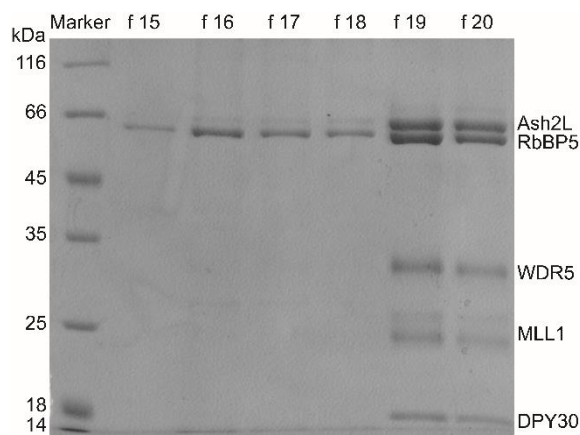


Figure F-10. SDS-PAGE displaying different fractions of 5-mer MLL1 complex after size exclusion chromatography. Fractions f 19 and f 20 contained the full 5-mer MLL1 complex.

During this thesis, different experiments using negative stained samples to visualize the 5-mer complex with an electron microscope have been performed. They are summarized in Table F-3 and images are illustrated in Figure F-11 and Figure F-12.

In the first experiment (Table F-3, Experiment 1a), directly after size exclusion chromatography of the 5-mer complex, samples were visualized using the negative stain protocol on an electron microscope (Figure F-11, left). Although some of the particles displayed sizes that could fit the ~200 kDa 5-mer complex, they lacked homogeneity. In order to improve particle quality, the GraFix procedure was applied (Table F-3, Experiment 1b), wherefore six gradients were loaded with 200 pmol of complex, each. After centrifugation and fractionation of the gradients, the MLL1 complex-containing fractions were, again, negative stained and visualized on the electron microscope (Figure F-11, right). The additional GraFix purification significantly enhanced homogeneity to yield particles more consistent in size and structure. In order to be able to use those samples not only for the negative stain protocol but also for cryo-electron microscopy (cryo-EM) imaging, the glycerol, which is accumulated in the complex solutions after the GraFix procedure, needs to be removed. Thus, in a second experiment, after initial SEC (Table F-3, Experiment 2a) and GraFix (Table F-3, Experiment 2b), a buffer exchange using small concentrator spin columns was performed (Table F-3, Experiment 2c) to dispose the glycerol. Negative stained images, however, demonstrated the loss of homogeneity due to the concentration process (Figure F-12, left). To circumvent the destructive buffer exchange using the concentrator spin columns, in a third experiment, the glycerol was discharged by running another analytical size exclusion chromatography (Table F-3, Experiment 3c) of the complex containing fractions obtained from the GraFix procedure. This, in turn, improved particle quality, although homogeneity suitable for cryo-EM specimen preparation and imaging was not reached (Figure F-12, right).

At that point the cryo-EM experiments were stopped due to both, time reasons and the fact that the cryo-EM map of the MLL1 complex was, at that time, published by Xue et al.²⁵ and Park et al.²⁶ This is discussed in the introduction of this thesis.

Table F-3. Negative stain experiments performed during this thesis. Table indicates different purification methods and quality outcome of particles. Experiments with the same number were all prepared from the same purified complex and performed on the same / following day.

Experiment	Purification of 5-mer complex	Quality of negative stained particles
1a	SEC	too heterogeneous
1b	SEC + GraFix	quite homogeneous
2a	SEC	too heterogeneous
2b	SEC + GraFix	quite homogeneous
2c	SEC + GraFix + buffer exchange using concentrator spin columns	homogeneity destroyed by buffer exchange using spin columns
3a	SEC	too heterogeneous
3b	SEC + GraFix	quite homogeneous
3c	SEC + GraFix + buffer exchange using SEC	quite homogeneous

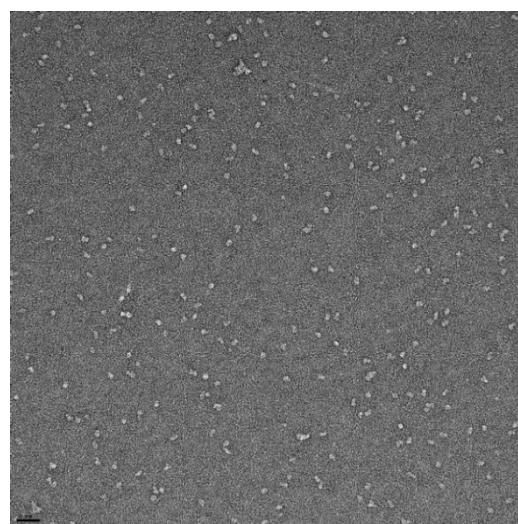
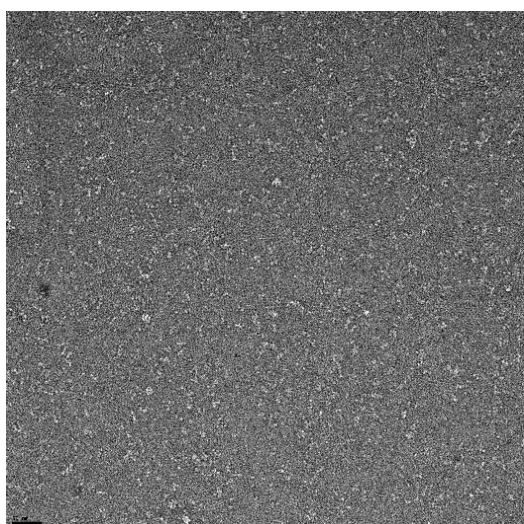


Figure F-11. Negative stained images to visualize 5-mer complex. Left: After initial SEC (1a). Right: After additional GraFix purification (1b).

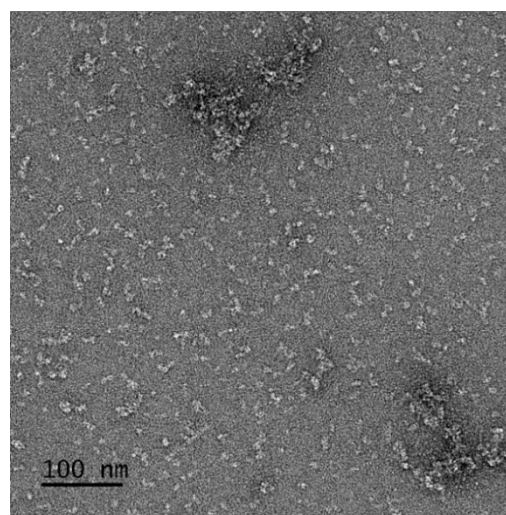
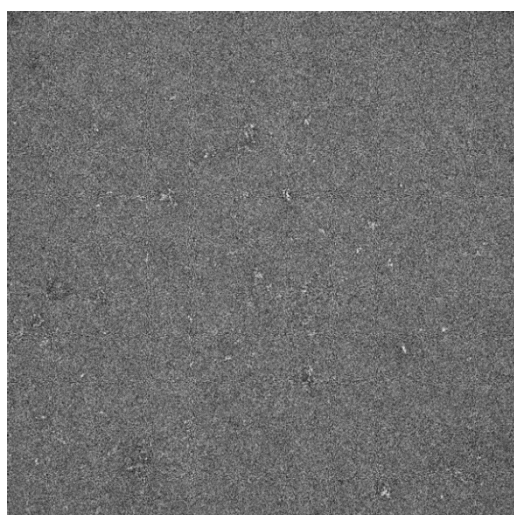


Figure F-12. Negative stain images to visualize 5-mer complex after SEC and GraFix followed by, Left: Buffer exchange using concentrator spin columns (2c). Right: Second SEC to discharge the glycerol from the sample (3c).

Experimental Procedures

Protein Expression and Purification:

The expression and purification of the single proteins for the MLL1 complex (MLL1, WDR5, RbBP5, Ash2L, DPY30) was performed according to the supporting information of the manuscripts C, detailed in sections 4.3.

Reconstitution of MLL1-complex:

To reconstitute the 5-mer MLL1 complex, specific equivalents of the single proteins (1.00 eq. Ash2L, 1.00 eq. RbBP5, 1.50 eq. MLL1, 1.50 eq. WDR5, 2.00 eq. DPY30) were mixed together and incubated for 30 min on ice. Afterwards, the complex was purified via size exclusion chromatography (SEC) on an Äkta system using an analytical Superdex 200 column (CV = 24 mL, settings: 280 nm, flow: 0.5 mL/min, buffer: 25 mM Hepes, 150 mM NaCl, 1 mM DTT, pH = 8.0). SDS-gel analysis was used to verify complex purity.

GraFix procedure:

To enhance the stability of the MLL1 complex, and thus the homogeneity, the complex was chemically fixed and additionally purified using the GraFix procedure. For that, 2.00 mL of buffer A (10% glycerol, 20 mM HEPES, 100 mM NaCl, 2 mM CaCl₂) were added to a centrifugation tube and 2.00 mL of buffer B (30% glycerol, 20 mM HEPES, 100 mM NaCl, 2 mM CaCl₂, 0.2% glutaraldehyde) were underlaid. The gradient was generated using a gradient mixer (Biocomp Gradient Master; settings: 1.10 min, 83°, 22 rpm). After cooling the gradients to 4 °C, 200 pmol (in max. 200 µL) of MLL1 complex were added to each tube, followed by centrifugation (40000 rpm, 15 h, 4 °C, no brake, slow start). Then, the centrifuged gradients were fractionated in 100 µL steps and analyzed via negative stain on the electron microscope. To remove the glycerol from the fractions containing the MLL1 complex, they were, again, purified via analytical size exclusion chromatography on an Ettan system using an analytical Superdex 200 Increase 5/50 GL column (CV = 3 mL, settings: 280 nm, flow: 0.15 mL/min, buffer: 25 mM HEPES, 150 mM NaCl, 1 mM DTT, pH = 8.0). For that, the excess of glutaraldehyde was inactivated by the addition of glycine to a final concentration of 80 mM beforehand.

Specimen Preparation and EM imaging of negative stained samples:

For initial negative stained images of the MLL1 complex, the respective solutions were diluted to ~20 µg/µL. 3 µL of such solution was added onto small grids (400 mesh copper grids coated with a continuous carbon film, SPI supplies) and incubated for 30 sec. Uranylformate (2% aq. solution, electron microscopy science) was added (3 × 3 µL) for background staining and images were taken at rt. on an electron microscope (Tecnai Spirit, FEI Thermo Fisher with LaB6-Cathode, Cs 2, 2 mm (spherical aberration), acceleration voltage: 120 kV). Images were recorded at a magnification of 42 kx and a defocus value of -1.5 µm on a Rio camera (CMOS, 4kx4K, Gatan; imaging software: GMS 3, Digital Micrograph version 03.30.2017, Gatan).

6.6 Synthesis of WDR5 Targeting PROTACS

The proteolysis-targeting chimera (PROTAC) technique was developed by Sakamoto et al. in 2001.²⁷ In general PROTACs are heterobifunctional molecules that induce target protein degradation. They consist of one end, which recruits the E3 ligase complex and another that binds the target protein. Monofunctional PROTACs, also referred to as “molecular glues”,²⁸ are molecules that are able to connect the protein of interest (POI) with the E3 ligase.^{28,29} Once the POI is brought into close proximity of the E3 ligase and upon ternary complex formation, the ligase ubiquitinates the POI leading to subsequent degradation of the target protein by the ubiquitin-proteasome system (UPS) (Figure F-13). Advantageously, PROTACs act catalytically and are able to facilitate several cycles of POI degradation.

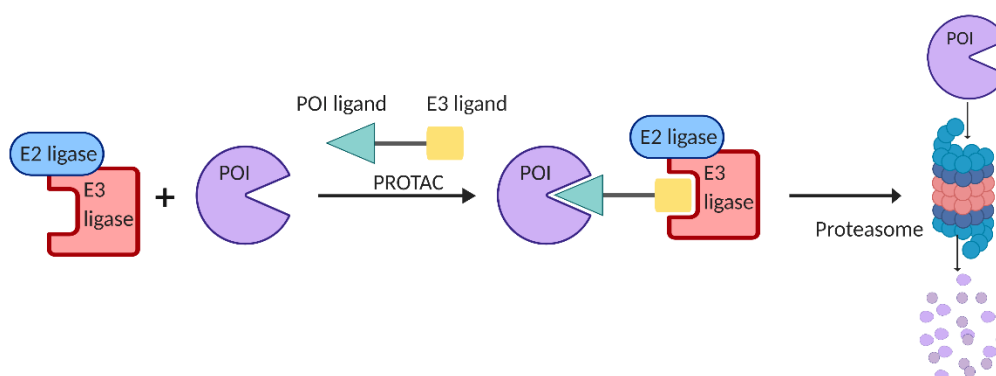


Figure F-13. Schematic representation of the ubiquitin-proteasome system (UPS). The PROTAC enables formation of a ternary complex, bringing the protein of interest (POI) into close proximity to the E3 ligase, enabling ubiquitination leading to its degradation via hydrolysis by the proteasome.

The UPS is a highly regulated signaling cascade, responsible for the controlled degradation of a wide variety of proteins that are being hydrolyzed by the proteasome upon ubiquitination.³⁰ Currently, the von Hippel-Lindua (VHL), Cereblon (CRBN), MDM2 and clAP1 are the predominant E3 ligases used in the PROTAC technology. The ligands **F-8/9/10/11** that specifically recognize these E3 ligases are depicted in Figure F-14.³¹

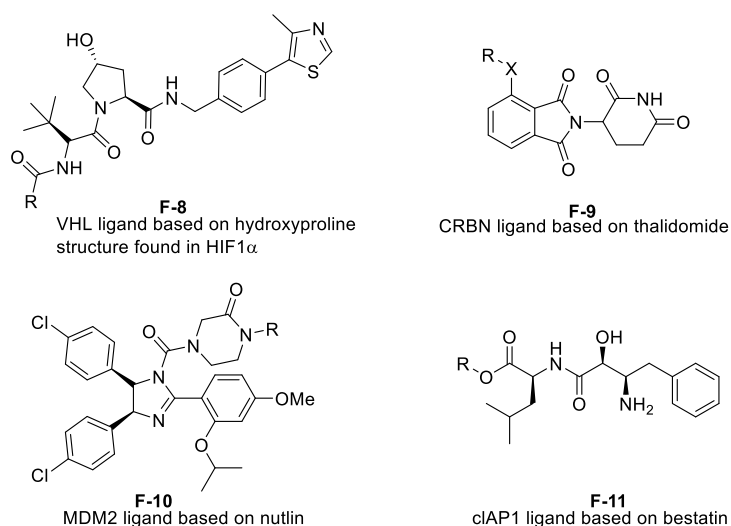
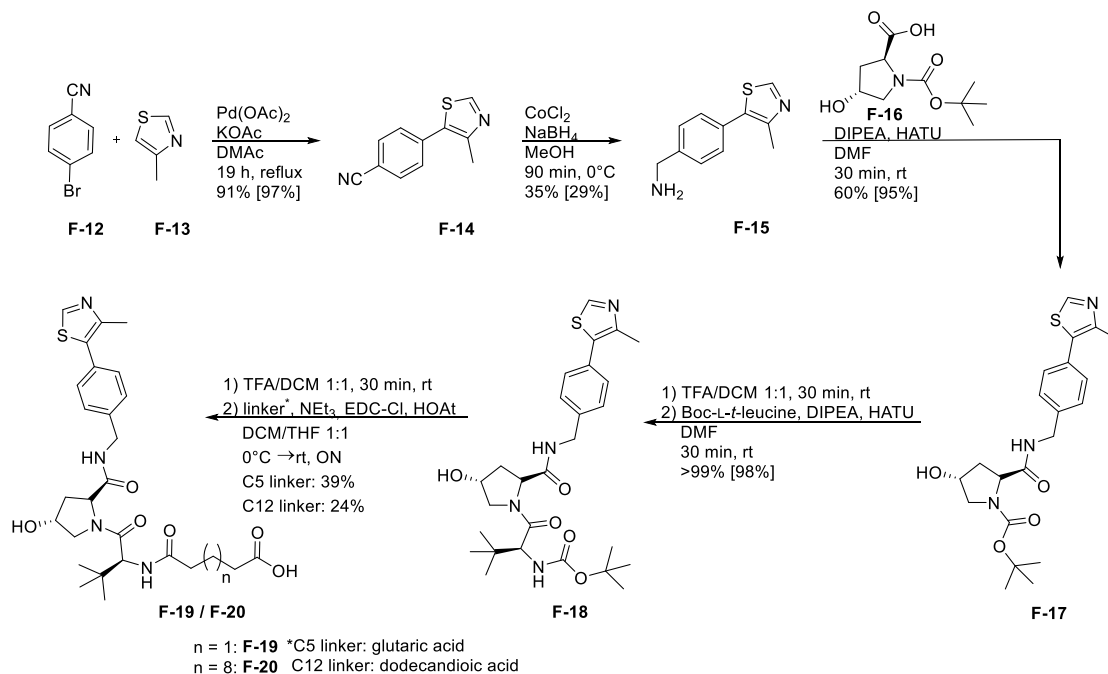


Figure F-14. Structures of E3 ligase ligands, that can be utilized in the PROTAC technique to specifically degrade proteins of interest (POIs).

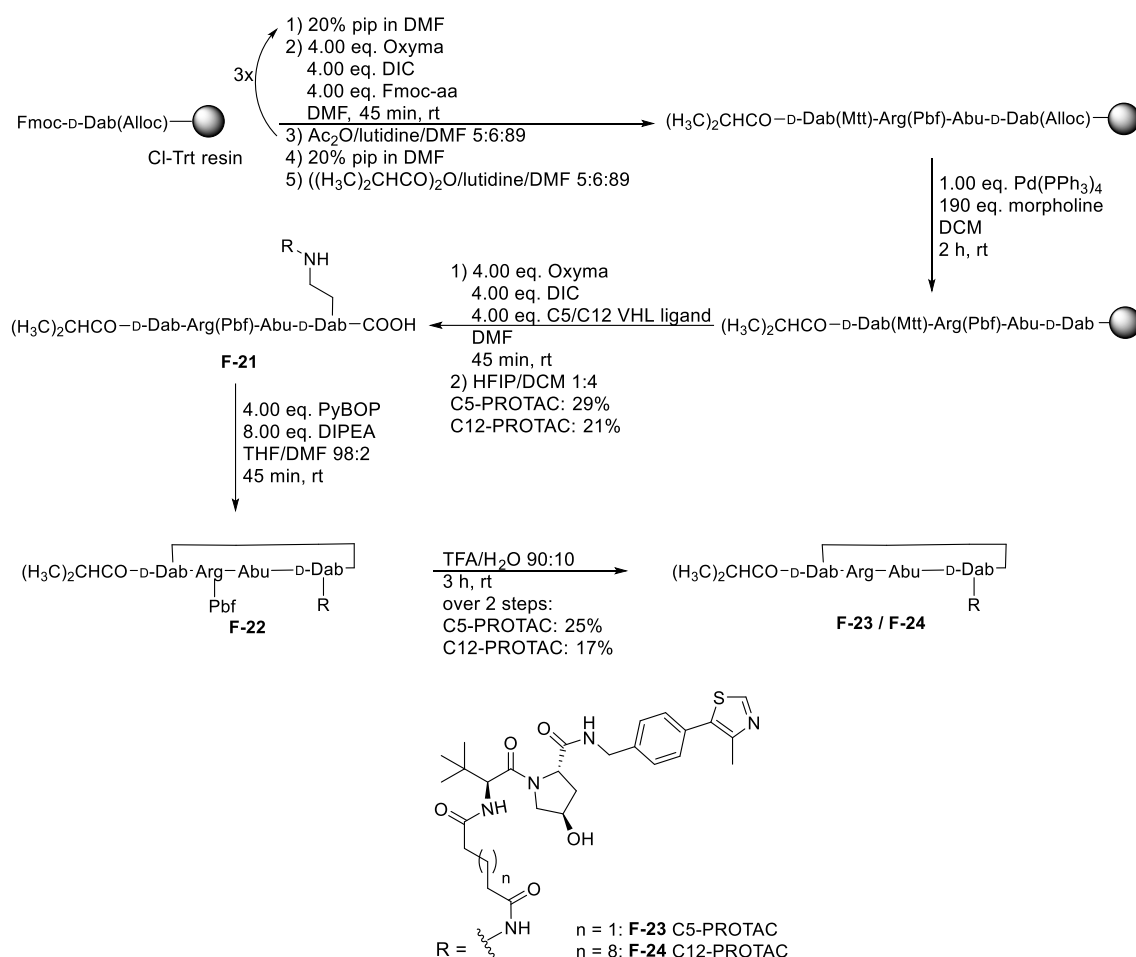
Both, peptide and small molecule based PROTACs hold great potential to target a wide range of proteins and have been successfully utilized to target proteins such as estrogen receptor α ,^{32–36} estrogen-related receptors α ³⁵, HaloTag7 fusion proteins,³⁷ BET-proteins,^{38–40} cellular retinoic acid binding proteins,⁴¹ α -synuclein protein,⁴² oncoprotein CREPT,⁴³ and the hypoxia inducible factors (HIF).^{44,45} Still, to date PROTACs are mainly based on small molecules, probably due to the drawbacks of peptides regarding physiological properties, as low intracellular stability and poor cell permeability.³⁶ On the contrary, peptide-based PROTACs offer great advantages over small molecules as they could more easily target PPIs, their higher specificity and low toxicity.⁴⁶

We now aimed at applying the PROTAC technique to WDR5, as not only it is known that WDR5 is crucial to maintain MLL1 complex formation⁴⁷ but it also has been demonstrated, that the knockout of WDR5 decreases the HMT-activity of MLL1.^{47,48} Furthermore, WDR5 knockdown in MLL-AF9 cells led to reduced transformation capability of the cells, similar to *ml1* deletion.⁴⁹ Thus, WDR5 is a promising target for PROTAC development. Off note, WDR proteins are able to bind ubiquitin in their β -propeller regions,⁵⁰ and WDR5 was proven to be a substrate of the Cullin4B-Ring (CRL4B)-E3 ligase.⁵¹ Therefore, we hypothesized that the PROTAC technique would allow specific WDR5 degradation by enabling recruitment of a E3 ligase. Thus, during this thesis, in a side project, peptide-based PROTACs targeting WDR5 and the VHL E3 ligase were developed, and first biological evaluations were carried out. We intended at utilizing PROTACs based on the macrocyclic WDR5 binding peptide MM-401.⁵² There, the VHL ligand should be coupled to an additionally incorporated D-diaminobutyric acid (D-Dab) side chain. In order to introduce the VHL ligand onto an amine bearing amino acid side chain via solid phase peptide synthesis (SPPS), the ligand was designed to include a carboxylic acid linker. Thus, two VHL ligands with different linker lengths (C5- **F-19** and C12-linker **F-20**) were synthesized (Scheme F-3).



Scheme F-3. Synthesis route for the C5/ C12-VHL-ligands **F-19** and **F-20**. Literature yields⁴⁴ are indicated in square brackets.

The linear precursor peptides were prepared following the standard Fmoc-based SPPS procedure (Scheme F-4). After Alloc-deprotection of the D-Dab side chain, the respective VHL ligands **F-19** and **F-20** were introduced into peptide **F-21** via standard conditions, followed by peptide cyclization. Final side chain deprotection yielded the desired PROTACs **F-23** and **F-24** (Scheme F-4).



Scheme F-4. Synthesis route of VHL-ligand incorporated cyclic peptides **F-23** (C5-PROTAC) and **F-24** (C12-PROTAC).

To evaluate the binding affinity of the two novel PROTACs, fluorescence polarization (FP)-based assays were performed (Table F-4 and Figure F-15) according to the procedure described in manuscripts B-C (section 4). Both cyclic PROTACs displayed high affinities to WDR5 in the low nM range. The C12-PROTAC **F-24** presented slightly better binding than its C5 analogue **F-23** (**F-23**: $K_i = 18.5$ nM; **F-24**: $K_i = 7.00$ nM). Compared to the linear AMPB peptide **MB-7** as well as the cyclic *o*F₄Azo compound **MD-CS-VIP 8**, binding of C12 is similar to the affinity of both *cis* isomeric states (**MB-7 trans**: $K_i = 1.25$ nM; *cis*: $K_i = 6.50$ nM; manuscript B; **MD-CS-VIP 8 cis**: $K_i = 8.00$ nM; *trans*: $K_i = 79.0$ nM; manuscript D). Therefore, it was a promising candidate for further biological evaluation.

Table F-4. IC₅₀, K_i values of cyclic VHL-PROTACs to WDR5. ^aExperiments have been performed with practical course students.

Peptide	Exp.	IC ₅₀ [μM]	K_i [nM]
F-23 (C5)	1	1.089	16.2
	2 ^a	1.060	15.7
	3 ^a	1.558	23.5
	mean	1.24 ± 0.28	18.5 ± 4.4
	F-24 (C12)	1 ^a	0.429
	2 ^a	0.628	8.92
	3 ^a	0.462	6.30
	mean	0.506 ± 0.106	7.00 ± 1.68

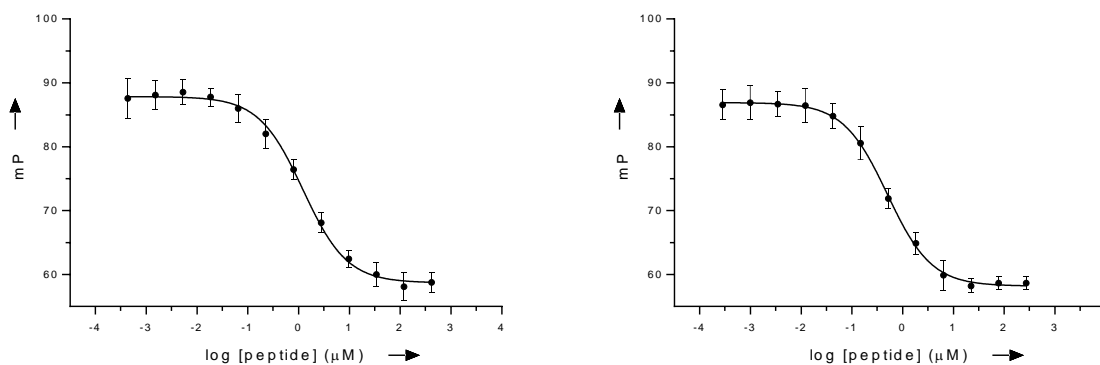


Figure F-15. Fluorescence polarization-based assays of VHL-PROTACs. Graphs represent mean values of three independent measurements Left: Peptide **F-23** (C5-PROTAC). Right: Peptide **F-24** (C12-PROTAC).

To test the cellular effects of the cyclic PROTACs, cell viability assays using human leukemia MV4-11 cells were performed as described in the manuscript D in section 4.4. However, compared to the cyclic peptides in manuscript D, a lower cell density (1×10^4 cells/mL instead of 2×10^4 cells/mL) was used and instead of using the Cell Titer-Glo[®] assay, the read out was performed using resazaurin fluorescence-based assays.⁵³ The effects on cell viability of peptides **F-23** and **F-24** are displayed in Figure F-15.

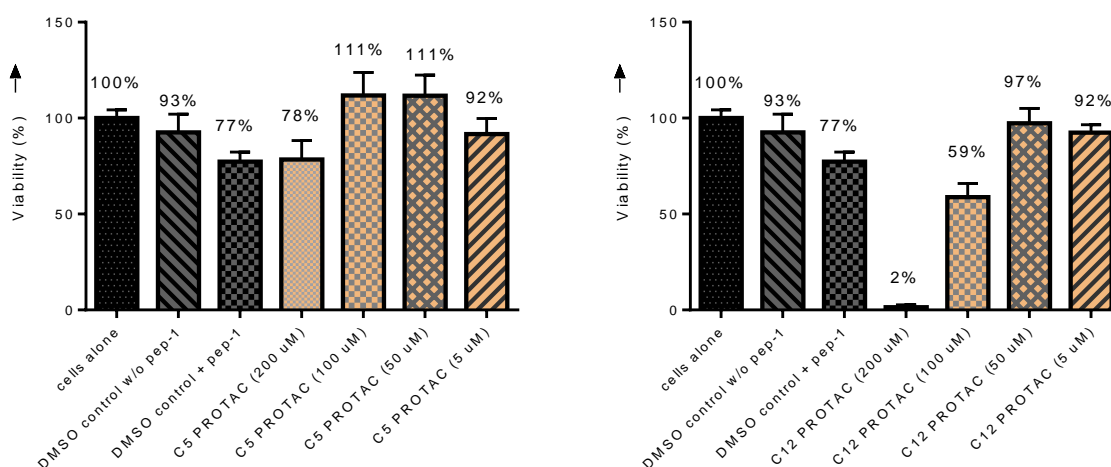


Figure F-16. MV4-11 cell viability after 3 day incubation with VHL-PROTAC cyclic peptides. Left: Peptide **F-23** (C5-PROTAC). Right: Peptide **F-24** (C12-PROTAC). Mean values of two independent measurements.

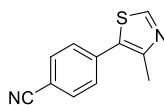
As can be deduced from the graphs in Figure F-15, the C5-PROTAC did not display any significant toxicity up to 200 μ M concentration, which is supposedly due to poor cellular uptake. The C12-PROTAC, which presented higher binding affinity to WDR5 than the C5-PROTAC in previous FP-assays, also exhibited greater effects on cell viability. Thus, it was demonstrated to totally repress cell viability applying 200 μ M concentration but displayed moderate cytotoxicity using 100 μ M, while toxicity for lower concentrations was neglectable. In order to being able to properly compare these results to the ones obtained using the cyclic σ F₄Azo peptide **MD-CS-VIP 8**, the experiments would need to be repeated with the same cell density, which exceeds the timeline of this thesis. However, for the cyclic σ F₄Azo peptide **MD-CS-VIP 8**, a dose-dependent trend in cell viability of MV4-11 cells could be observed after treatment with **MD-CS-VIP 8** for 3 days (see manuscript D, SI, Fig. S22), which failed for

the C5-PROTAC. This suggests that the cyclic PROTACs display very low abilities for cellular uptake, even lower than the *o*F₄Azo peptide **MD-CS-VIP 8**, which is a common obstacle for cyclic peptides generally.⁵ To further elucidate the toxicity mechanism, additional experiments would need to be performed, such as the 3 h viability evaluation and the cell washing experiments described in section 6.1 and in manuscript D in section 4.4.

To summarize, utilization of the PROTAC technique represents an alternative approach to inhibit protein function by inducing specific POI degradation using the UPS. In this thesis, first preliminary tests have been performed. We verified that novel cyclic PROTACs can bind to WDR5 in the low nM-range, comparable to our published/submitted linear and cyclic azobenzene containing peptides. The evaluation of these compounds on MV4-11 cells displayed no toxicity for the C5-PROTAC and for the C12-PROTAC only at very high concentrations (200 μM). Thus, in the future novel PROTACs, which display improved cell permeability would need to be designed and synthesized in order to obtain improved intracellular effects. In addition, biological immunoblotting assays to test, whether or not the POI (WDR5) is, indeed, recruited to the E3 ligase and, thus, degraded, would need to be performed. In any case, it was postulated recently that WDR5 is not believed to induce proteasomal degradation, making it unsuitable for PROTAC discovery.⁵⁴ Moreover, it was demonstrated that WDR5 is a highly conserved, essential regulator of protein synthesis and gene expression, which performs multiple functions in the context of chromatin.⁵⁵ It was revealed that WDR5 degradation leads to different biological outputs than inhibition of its *Win* site. Thus, WDR5 degradation leads to greater effects on neuronal differentiation and sensory precipitation genes, while *Win* site blockage displays broader effects on genes linked to the cell cycle.⁵⁵ In order to achieve more selective biological outputs, inhibition of the WDR5 *Win* site might provide the superior alternative, as WDR5 degradation leads to broader effects than *Win* site inhibition.⁵⁵

Experimental Procedures

Synthesis of VHL-Ligands:



F-14

Following the procedure of Galdeano et al.⁴⁴, 4-bromobenzonitrile **F-12** (995 mg, 5.50 mmol, 1.00 eq.) was dissolved in Dimethylacetamide (5.5 mL) and Pd(OAc)₂ (1.23 mg, 0.0055 mmol, 0.1 mol%), KOAc (1.08 g, 11.0 mmol, 2.00 eq.), and 4-methylthiazole **F-13** (1.09 g, 11.0 mmol, 2.00 eq.) were added. After stirring the solution for 19 h at 150 °C, water was added and the product was extracted with CH₂Cl₂ (3 × 50 mL). The combined organic layers were washed with brine dried over MgSO₄ and the solvent was removed under reduced pressure to yield the **F-14** as a yellow solid crude product (1.00 g, 4.99 mmol, ~90%), which was used without further purification. **TLC:** R_f = 0.89 (DCM:MeOH 1:5). **¹H NMR (300 MHz, CDCl₃, δ):** 8.75 (s, 1H, SCH_{ar}), 7.72 (d, ³J = 7.7 Hz, 2H, 2 × CH_{ar}), 7.56 (d, ³J = 7.6 Hz, 2H, 2 × CH_{ar}), 2.56 (s, 3H, CH₃). **¹³C NMR (75 M, Hz, CDCl₃, δ):** 151.7 (S_{Car}H), 150.2 (C_{ar,q}), 137.0 (C_{ar,q}), 132.6 (2 × C_{ar}H), 130.2 (CN), 129.9 (2 × C_{ar}H), 118.6 (C_{ar,q}), 111.7 (C_{ar,q}), 16.5 (CH₃). **HRMS-ESI⁺ (m/z):** calcd. for [M+H]⁺ C₁₁H₈N₂S₁H, 201.0481; found, 201.0480.

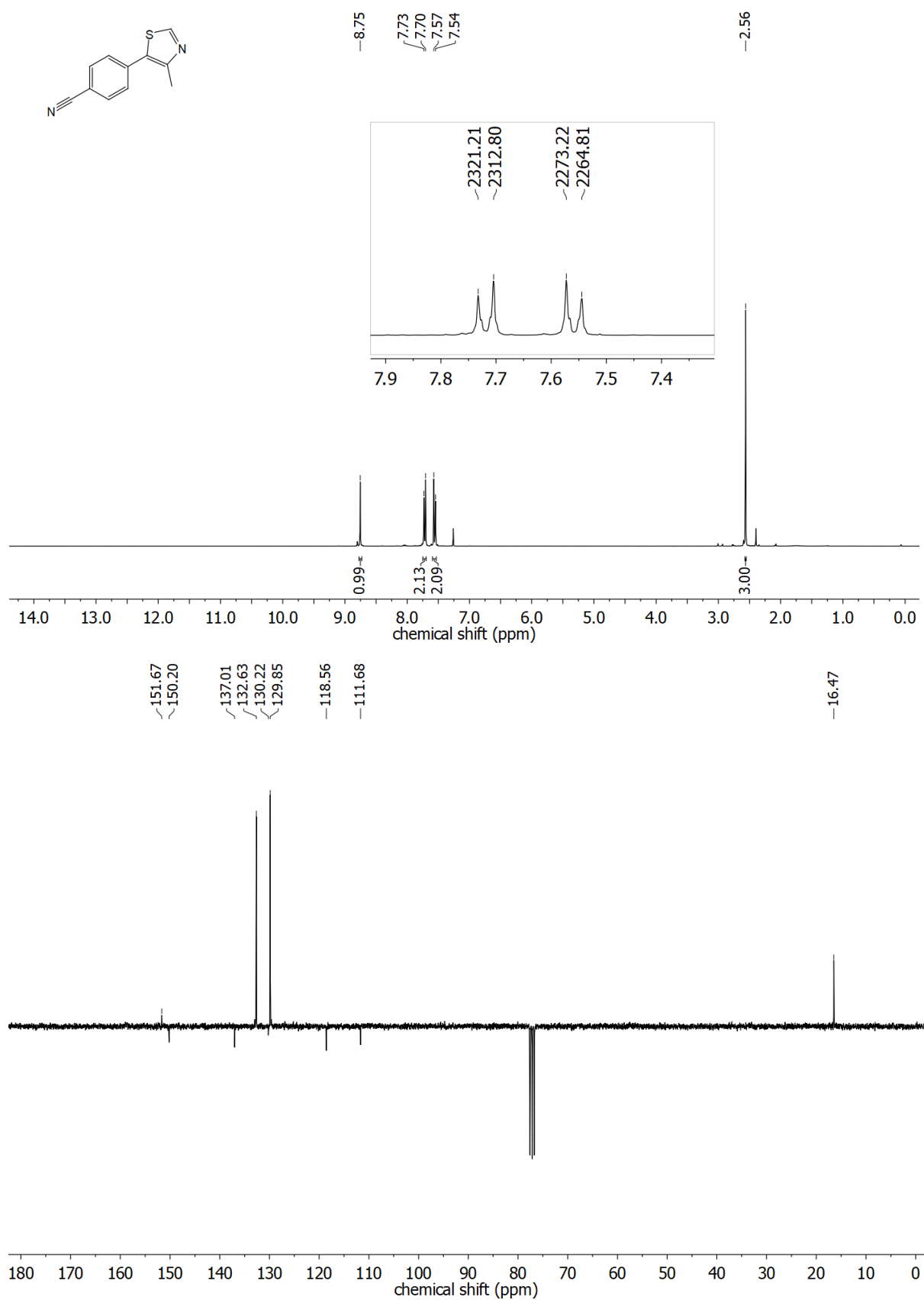
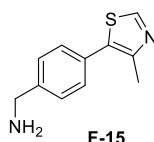


Figure F-17. Top: $^1\text{H-NMR}$ spectrum of the compound F-14. Insert: zoom-in of signals in the aromatic region to highlight the shifts in Hz units for calculating J -constants. Bottom: $^{13}\text{C-APT-NMR}$ spectrum of compound F-14.



Following the procedure of Galdeano et al.⁴⁴, compound **F-14** (400 mg, 2.00 mmol, 1.00 eq.) was dissolved in dry MeOH (22.0 mL) in a Schlenk flask under N₂-atmosphere and CoCl₂ (390 mg, 3.0 mmol, 1.50 eq.) was added. After cooling to 0 °C, NaBH₄ (378 mg, 10.0 mmol, 5.00 eq.), was added portion-wise and the solution stirred for 90 min at 0 °C. The reaction was quenched by the addition of 0.5 M NH₃ in water (20 mL) and the product was extracted with chloroform (5 × 20 mL). The combined organic layers were washed with brine and dried over MgSO₄. After removing the solvent under reduced pressure, the crude was purified via flash column chromatography (CH₂Cl₂/ MeOH (0.5 M NH₃) 19:1), to yield the product **F-15** (143 mg, 0.700 mmol, 35%) as a yellow oil. The characterization is in agreement with the literature.⁴⁵ **TLC**: R_f = 0.40 (DCM:MeOH 1:5). **¹H NMR (300 MHz, CDCl₃, δ)**: 8.67 (s, 1H, SCH_{ar}), 7.43-7.36 (m, 4H, 4 × CH_{ar}), 3.92 (s, 2H, NH₂), 2.53 (s, 3H, CH₃), 2.17 (s, 2H, C_{ar}CH₂). **¹³C NMR (75 MHz, CDCl₃, δ)**: 150.2 (SC_{ar}H), 148.6 (C_{ar,q}), 143.2 (C_{ar,q}), 131.9 (C_{ar,q}), 130.6 (C_{ar,q}), 129.6 (2 × C_{ar}H), 127.6 (2 × C_{ar}H), 46.2 (C_{ar}CH₂), 16.2 (CH₃). **HRMS-ESI⁺ (m/z)**: calcd. for [M+H]⁺ C₁₁H₁₂N₂S₁H, 205.0805; found: 205.0795.

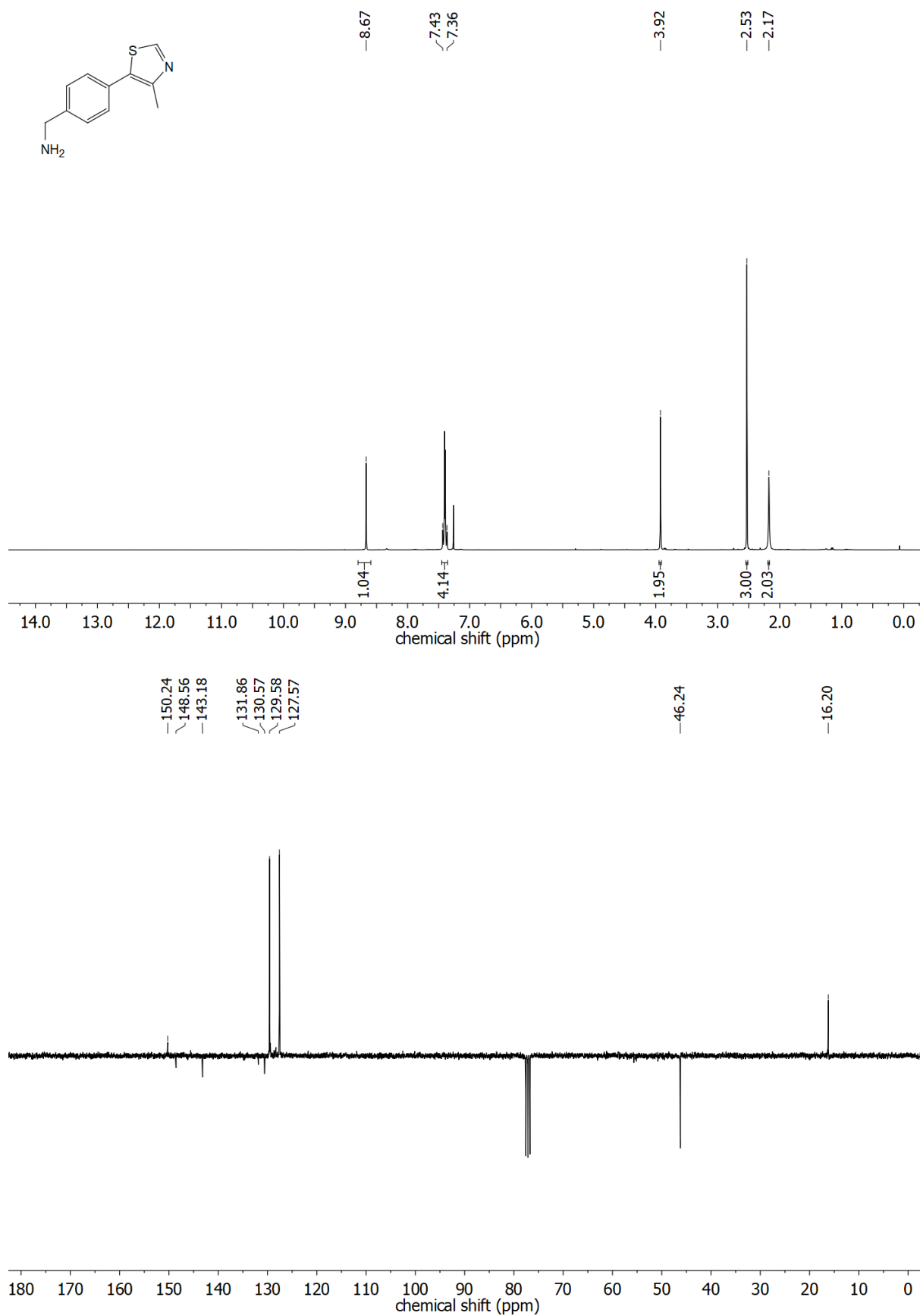
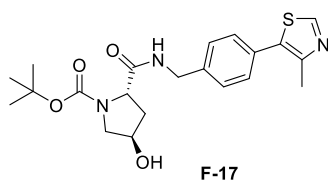
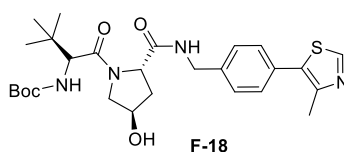


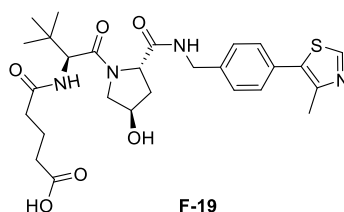
Figure F-18. Top: ¹H-NMR spectrum of the compound F-15. Bottom: ¹³C-APT-NMR spectrum of compound F-15.



Following the procedure of Galdeano et al.⁴⁴, compound **F-15** (97.8 mg, 0.479 mmol, 1.00 eq.) was dissolved in DMF (4.00 mL) and N-Boc-*trans*-4-hydroxy-L-proline **F-16** (111 mg, 0.479 mmol, 1.00 eq.) and DIPEA (332 μ L, 1.90 mmol, 4.00 eq.) were added and stirred for 5 min. Then, HATU (378 mg, 10.0 mmol, 5.00 eq.), was added and the solution stirred for 30 min at rt. After the addition of water (10 mL), the product was extracted with EtOAc (3 \times 20 mL) and the combined organic layers were washed with brine and dried over MgSO₄. After removing the solvent under reduced pressure, the crude **F-17** (121 mg, 0.290 mmol, 60%) was directly used for the next step without purification. NMR analysis was performed on the final compounds **F-19** and **F-20**. TLC: R_f = 0.40 (DCM:MeOH 9:1). HRMS-ESI⁺ (m/z): calcd. for [M+H (w/o Boc)]⁺ C₁₆H₁₉N₃O₂S₁H, 318.1271; found: 318.1269.



Following the procedure of Galdeano et al.⁴⁴, compound **F-17** (120 mg, 0.290 mmol, 1.00 eq.) was dissolved in TFA/CH₂Cl₂ 1:1 (3.00 mL) and stirred for 30 min. The solvent was removed under reduced pressure to yield the Boc-deprotected intermediate, which was then dissolved in DMF (2.40 mL). Boc-L-*tert*-leucine (67.1 mg, 0.290 mmol, 1.00 eq.) and DIEPA (202 μ L, 1.16 mmol, 4.00 eq.) were added and stirred for 5 min. HATU (121 mg, 0.319 mmol, 1.10 eq.) was added and the mixture stirred for 30 min. After the addition of water, the product was extracted with EtOAc (3 \times 20 mL). The combined organic layers were washed with brine and dried over MgSO₄. The crude **F-18** (154 mg, 0.290 mmol, >99%) was directly used for the next step without purification. NMR analysis was performed on the final compounds **F-19** and **F-20**. TLC: R_f = 0.18 (DCM:MeOH 9:1). HRMS-ESI⁺ (m/z): calcd. for [M+Na]⁺ C₂₇H₃₈N₄O₅S₁Na, 553.2455; found: 553.2445.



Compound **F-18** (154 mg, 0.29 mmol, 1.00 eq.) was dissolved in TFA/CH₂Cl₂ (3.00 mL) and stirred for 30 min. The solvent was removed under reduced pressure to yield the Boc-protected crude, which was directly dissolved in CH₂Cl₂/THF 1:1 (4.00 mL). Glutaric acid (192 mg, 1.45 mmol, 5.00 eq.) was dissolved in CH₂Cl₂/THF 1:1 (25.0 mL) and cooled to 0 °C. After the addition of NEt₃ (143 μL, 1.03 mmol, 3.55 eq.), HOAt (43.4 mg, 0.319 mmol, 1.10 eq.) and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimidohydrochloride (EDC×Cl) (61.2 mg, 0.319 mmol, 1.10 eq.), the Boc-protected crude was added to the mixture and stirred for 1 h at 0°C, followed by stirring at rt for 3 h. The solvent was removed under reduced pressure and the crude was purified by flash column chromatography (CH₂Cl₂/MeOH 19:1 + 1% formic acid) to yield the product **F-19** as a yellow oil (72.4 mg, 0.133 mmol, 46%). **TLC:** R_f = 0.34 (DCM:MeOH 9:1 + 0.1% formic acid). **¹H-NMR (300 MHz, MeOD, δ):** 8.87 (s, 1H, SCH_{ar}), 7.48-7.40 (m, 4H, 2 × CH_{ar}), 4.63-4.51 (m, 4H, 2 × CH₂), 4.38-4.33 (m, 1H, CH), 3.93-3.89 (m, 1H, CH), 3.81-3.78 (m, 1H, CH), 2.47 (s, 3H, CH₃), 2.34-2.30 (m, 4H, 2 × CH₂), 2.22-2.05 (m, 2H, CH₂), 1.91-1.86 (m, 2H, CH₂), 1.04 (s, 9H, 3 × CH₃). **¹³C NMR (75 MHz, CDCl₃, δ):** 176.6 (CO), 174.4 (CO), 172.0 (CO), 171.8 (CO), 151.5 (SC_{ar}H), 147.5 (C_{ar}), 138.4 (C_{ar}), 132.4 (C_{ar}), 130.2 (C_{ar}), 129.5 (2 × C_{ar}H), 127.9 (2 × C_{ar}H), 70.3 (CH), 59.6 (CH), 57.9 (CH), 57.4 (CH₂), 43.2 (CH₂), 37.1 (CH₂), 35.6 (C_q), 34.8 (CH₂), 33.0 (CH₂), 26.4 (3 × CH₃), 20.8 (CH₂), 15.4 (CH₃). **HRMS-ESI⁺ (m/z):** calcd. for [M+Na]⁺ C₂₇H₃₆N₄S₁O₆Na, 567.2248; found: 567.2246. **HPLC:** t_R = 15.66 min.

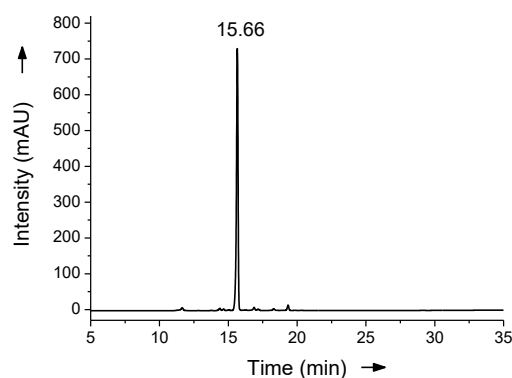


Figure F-19. HPLC chromatogram of purified compound **F-19**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm.

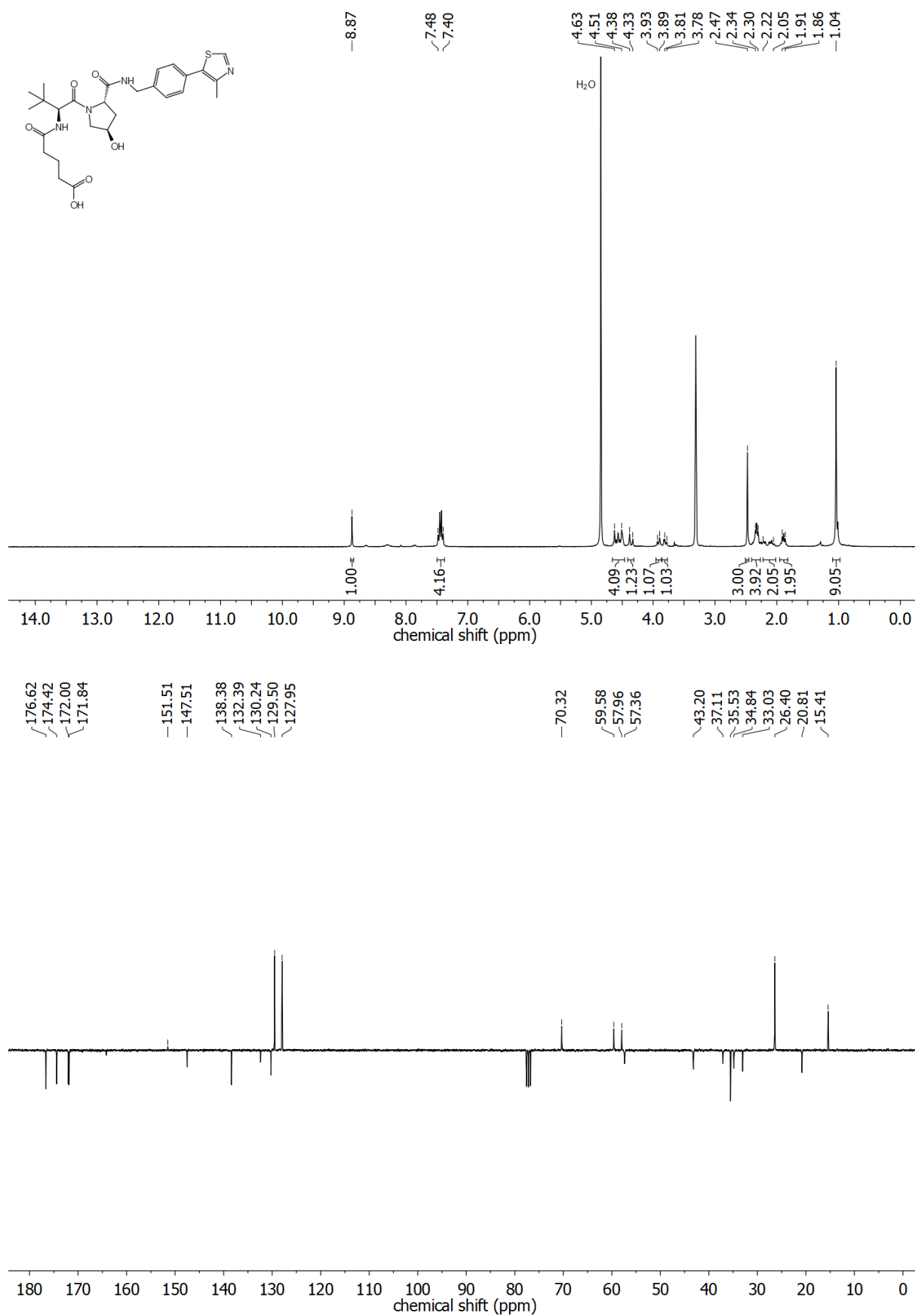
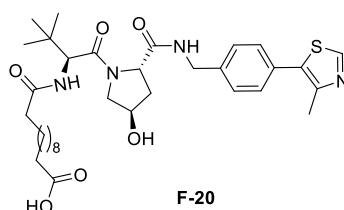


Figure F-20. Top: ¹H-NMR spectrum of the compound F-19. Bottom: ¹³C-APT-NMR spectrum of compound F-19.



Compound **F-18** (690 mg, 1.30 mmol, 1.00 eq.) was dissolved in TFA/CH₂Cl₂ (3.00 mL) and stirred for 30 min. The solvent was removed under reduced pressure to yield the Boc-protected crude, which was directly dissolved in CH₂Cl₂/THF 1:1 (10.0 mL). Dodecanedioic acid (1.50 g, 6.50 mmol, 5.00 eq.) was dissolved in CH₂Cl₂/THF 1:1 (120 mL) and cooled to 0 °C. After the addition of NEt₃ (640 μL, 4.62 mmol, 3.55 eq.), HOAt (190 mg, 1.43 mmol, 1.10 eq.) and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimid-hydrochloride (EDC×Cl) (270 mg, 1.43 mmol, 1.10 eq.), the Boc-protected crude was added to the mixture and stirred for 1 h at 0 °C, followed by stirring at rt for 3 h. The solvent was removed under reduced pressure and the crude was purified by flash column chromatography (CH₂Cl₂/MeOH 19:1 + 1% formic acid) to yield the product **F-20** as a yellow oil (200 mg, 0.311 mmol, 24%). **TLC:** R_f = 0.56 (DCM:MeOH 9:1 + 1% formic acid). **¹H-NMR (300 MHz, CDCl₃, δ):** 8.71 (s, 1H, SCH_{ar}), 7.36 (s, 4H, 2 × CH_{ar}), 4.69-4.54 (m, 4H, 2 × CH₂), 4.33-4.26 (m, 2H, CH₂), 3.68-3.65 (m, 1H, CH), 2.50 (s, 3H, CH₃), 2.41 (m, 1H, CH), 2.31-2.18 (m, 4H, 2 × CH₂), 2.00 (s, 1H, CH), 1.58 (m, 4H, 2 × CH₂), 1.26 (m, 12H, 6 × CH₂), 0.94 (s, 9H, 3 × CH₃). **¹³C-NMR (75 MHz, CDCl₃, δ):** 177.6 (CO), 175.3 (CO), 172.7 (CO), 171.5 (CO), 158.7 (SC_{ar}H), 150.9 (C_{ar}), 148.9 (C_{ar}), 147.6 (C_{ar}), 138.7 (C_{ar}), 130.1 (2 × C_{ar}H), 128.6 (2 × C_{ar}H), 70.6 (CH), 59.8 (CH), 58.2 (CH), 57.7 (CH₂), 43.8 (CH₂), 36.6 (CH₂), 35.7 (C_q), 34.3 (CH₂), 29.3 (CH₂), 28.9 (CH₂), 28.7 (CH₂), 28.6 (CH₂), 28.5 (CH₂), 26.9 (3 × CH₃), 25.7 (CH₂), 24.9 (CH₂), 16.5 (CH₃). **HRMS-ESI⁺ (m/z):** calcd. for [M+Na]⁺ C₃₄H₅₀N₄O₆S₁Na, 665.3343; found: 665.3335. **HPLC:** t_R = 21.10 min.

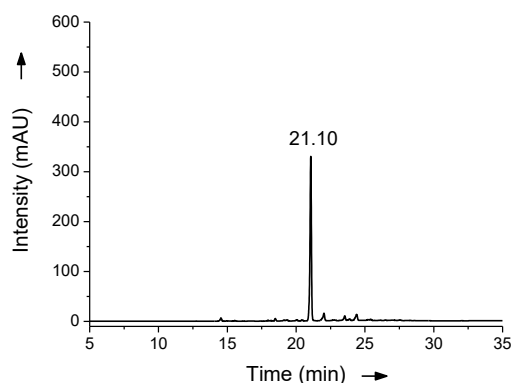


Figure F-21. HPLC chromatogram of purified compound **F-20**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm.

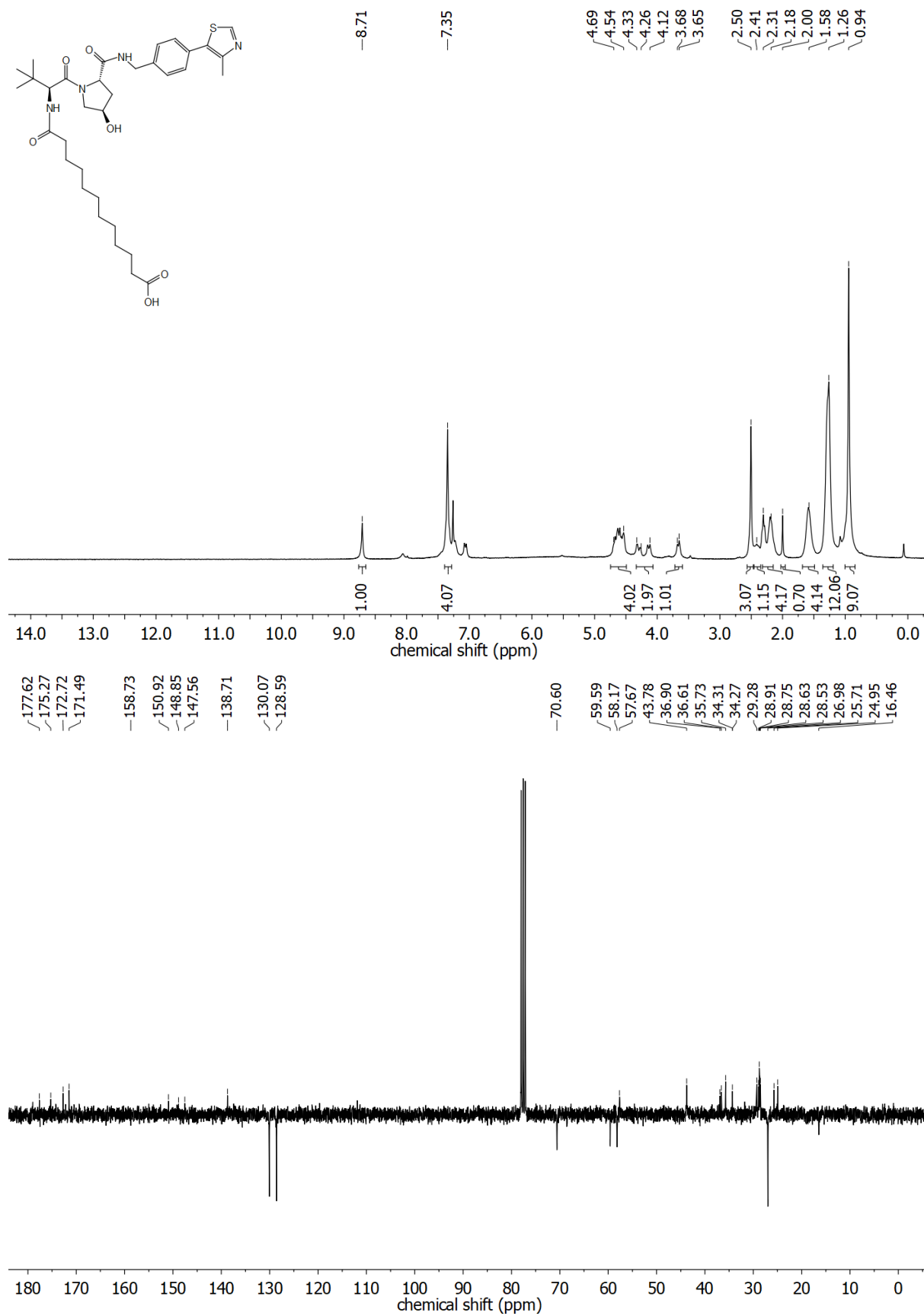
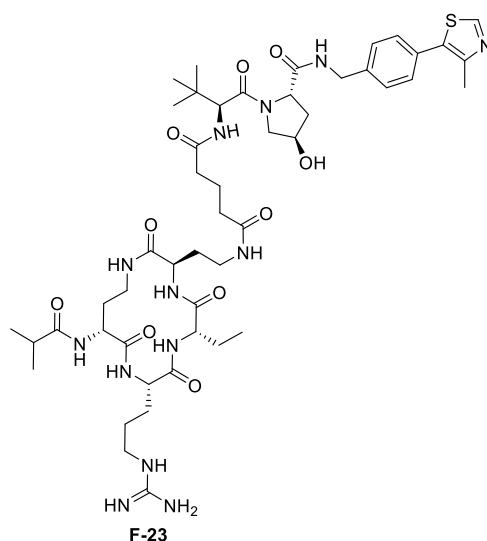


Figure F-22. Top: ¹H-NMR spectrum of the compound F-20. Bottom: ¹³C-APT-NMR spectrum of compound F-20.

Peptide Synthesis:

The synthesis of the linear peptides was performed manually according to the standard Fmoc-solid-phase methodology using Oxyma/DIC as coupling agents and Cl-Trt resin, which was manually loaded with the first amino acid D-Dab(Alloc), following the protocol described above in the manuscripts B-C in section 4. After coupling of the last N-terminal amino acid D-Dab(Mtt), their N-termini were acetylated with isobutyric anhydride, according to the standard capping procedure. The VHL-ligand was incorporated on the side chain of the corresponding Alloc-protected amino acid, after orthogonal deprotection⁵⁶ applying 4.00 eq. VHL-ligand ([VHL-ligand] = 0.35 M), 4.00 eq. PyBOP, 8.00 eq. DIPEA in DMF for 45 min. The final linear precursor peptides were cleaved from the resin with hexafluoroisopropanol HFIP/DCM (1:4) for 3 h, and purified using a preparative RP-HPLC (eluent A: H₂O + 0.1% TFA, eluent B: MeCN + 0.1% TFA, gradient of 5-95% B). Intramolecular C-terminal end to N-terminal D-Dab side chain cyclization was performed. The corresponding peptide (1.00 eq.) was dissolved in DMF ([peptide] = 75 mM), and diluted with THF to an approximated 1.5 mM solution in THF/DMF (98:2). HATU (3.00 eq.) and DIPEA (4.00 eq.) were dissolved in THF containing 2% DMF and added to the peptide solution reaching a final peptide concentration of 1 mM. The mixture was stirred for 45 min, followed by evaporation of the solvent. The crude was purified via preparative RP-HPLC (gradient: 5-95% B). After freeze-drying, the arginine side chain protecting group (Pbf) was removed by treating the peptide with TFA/H₂O 90:10 (535 μ L / 1 mg peptide) for 2 h. Final purification via semi-preparative RP-HPLC (gradient: 5-95% B) yielded the products as light yellow solids. Below, the characterization data for the peptides are displayed.



Cyclization was performed in a 11.0 μ mol scale as detailed above. After final purification, the product (2.90 mg, 2.79 μ mol, 25%) was obtained as a yellow solid. **HPLC:** t_R = 15.99 min. Purity = 96%. Formula: C₄₉H₇₅N₁₃O₁₀S. **HRMS-ESI⁺ (m/z):** [M+H]⁺ calcd.: 1038.5553; found: 1038.5569.

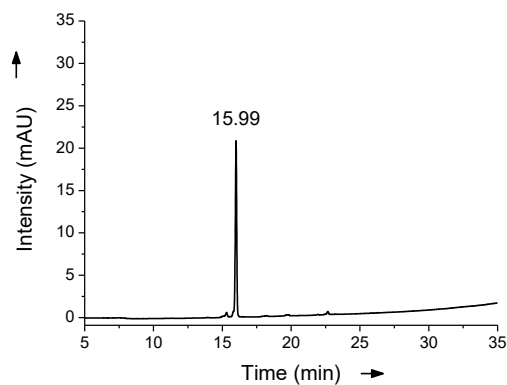
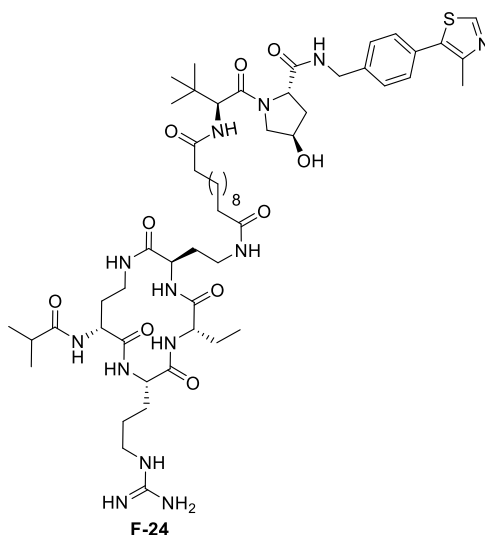


Figure F-23. HPLC chromatogram of purified compound **F-23**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm.



Cyclization was performed in a 7.90 μmol scale as detailed above. After final purification, the product (1.50 mg, 1.32 μmol , 17%) was obtained as a yellow solid. **HPLC**: $t_R = 19.32$ min. Purity = 97%. Formula: $\text{C}_{56}\text{H}_{89}\text{N}_{13}\text{O}_{10}\text{S}$. **HRMS-ESI⁺ (m/z)**: $[\text{M}+\text{H}]^+$ calcd.: 1136.6649; found: 1136.6661.

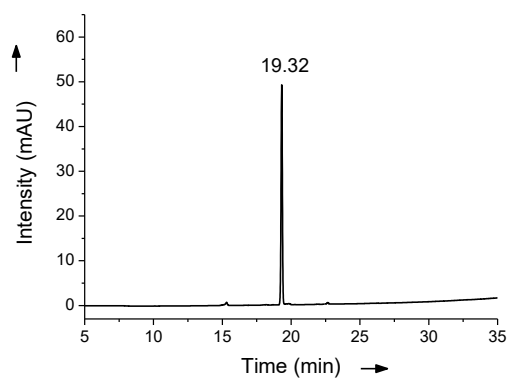


Figure F-24. HPLC chromatogram of purified compound **F-24**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm.

6.7 Synthesis of Additional Peptides to Target Protein-Protein Interactions

During this thesis, a number of additional peptides for targeting different PPIs have been designed and synthesized for collaboration projects. For example, the most straightforward collaboration was, where the linear peptides **MB-7**, **MB-16** and **MB-18** from the manuscript B (MB) were sent to Johanna Girant (group of Prof. Dr. Birchmeier, MDC Berlin) to test their effect on Ls174T and DLD1 human colon cancer cells. However, as for MM-401, no effects were observed. The other projects are explained in detail below.

6.7.1 MLL2 Complex

MLL2 (also known as KMT2D), belongs to the family of SET1-histone methyl-transferases. Similar to MLL1, MLL2 is associating in a core complex with WDR5, RbBP5 and Ash2L, which is enhancing enzymatic activity of the SET domain. The *Win* motif of MLL2 (GCARSEPKILT; $K_d = 0.16 \mu\text{M}$) binds to the Arg-binding cavity of WDR5, comparably to the *Win* motif of MLL1 (GCSARAEVHLRK; $K_d = 1.13 \mu\text{M}$).⁵⁷ In contrast to MLL1, it was demonstrated that WDR5 is not essential for MLL2 methylation activity (relative activity of MLL2 without WDR5: 80%),⁵⁸ and that the interaction of MLL2 to the RbBP5-Ash2L dimer is sufficient.^{58,59}

It is known that mutations in the *kmt2d* gene, are the major cause of Kabuki syndrome.⁶⁰ The Kabuki syndrome is a haploinsufficient congenital multi-organ malfunction syndrome, which belongs to the human developmental disorders based on disturbed epigenetic regulation of gene expression during embryonic morphogenesis. Kabuki patients feature a typical facial gestalt, intellectual disability, skeletal findings, dermatoglyphic anomalies and short stature.⁶¹ Schwenty-Lara et al.⁶² found that *kmt2d* is expressed in *Xenopus laevis* neural crest (NC) cells at all stages, and loss of *kmt2d* function leads to severe impairment of NC formation and cell migration. Furthermore, they demonstrated formation of Kabuki-like craniofacial symptoms in *Xenopus* upon loss of *kmt2d* function via morpholino-mediated *kmt2d* knockdown. They verified that hypoplastic heart defects in Kabuki syndrome patients are recapitulated by *kmt2d* loss of function in *Xenopus*, postulating *Xenopus* morphants as a tool for elucidation of congenital heart defects associated with Kabuki syndrome.⁶²

During this thesis, peptides **F-25-27** based on the MLL2 *Win* motif (Figure F-25), were synthesized for our collaboration partners, Janina Schwenty-Lara and Denise Nehl (AG Borchers, Biology Department, Philipps-University Marburg), for testing in vivo effects in *Xenopus*. As peptides should be injected directly into embryos in the two- or four cell state of *Xenopus*, intrinsic uptake is not required. Still, an octaarginine-containing MLL2 *Win* motif as well as a FAM labelled analogue were synthesized too (Figure F-25). The capacity of these peptides to lead to observable effects in *Xenopus* phenotype comparable to *kmt2d* knockdown was evaluated.

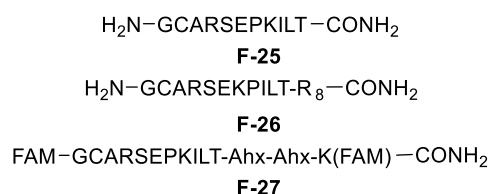


Figure F-25. MLL2 *Win* based peptides synthesized for testing their effects in *Xenopus*.

Our collaborators did not detect any phenotype changes in *Xenopus* development upon treatment with these peptides. These observations are in agreement with the fact that WDR5 is not required for MLL2 enzymatic activity. As expected, the MLL1-based peptide **MB-7** did not yield any effect neither.

Experimental Procedures

Peptide Synthesis:

The synthesis of the linear peptides was performed manually according to the standard Fmoc-solid-phase methodology using rink amide resin following the protocol described above in the manuscripts B-C in section 4. After final deprotection and cleavage, purification was performed via semi-preparative RP-HPLC (eluent A: water + 0.1% TFA, B: MeCN + 0.1% TFA; gradient: 5-40% B). Characterization data of final peptides are displayed below.

F-25: H₂N-Gly-Cys-Ala-Arg-Ser-Glu-Pro-Lys-Ile-Leu-Thr-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail (94% TFA/ 1% TIS/ 2.5% DODT/ 2.5% H₂O). After purification, the product as a 3 × TFA salt (4.7 mg, 3.10 μmol, 31%) was obtained as a white solid. **HPLC:** t_R = 14.59 min. Purity = 95%. Formula: C₄₉H₈₈N₁₆O₁₅S. **HRMS-ESI⁺ (m/z):** [M+H]⁺ calcd.: 1173.6409; found: 1173.6403.

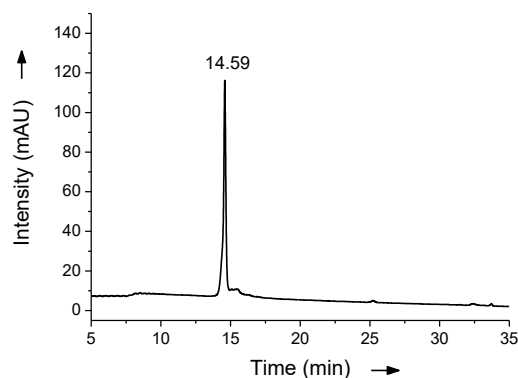


Figure F-26. HPLC chromatogram of purified peptide **F-25**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm.

F-26: H₂N-Gly-Cys-Ala-Arg-Ser-Glu-Pro-Lys-Ile-Leu-Thr-(Arg)₈-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail (94% TFA/ 1% TIS/ 2.5% DODT/ 2.5% H₂O). After purification, the product as a 11 × TFA salt (1.40 mg, 1.116 μmol, 11%) was obtained as a white solid. **HPLC:** t_R = 16.76 min. Purity = 98%. Formula: C₉₇H₁₈₄N₄₈O₂₃S. **HRMS-ESI⁺ (m/z):** [M+3H]³⁺ calcd.: 808.1548; found: 808.1558.

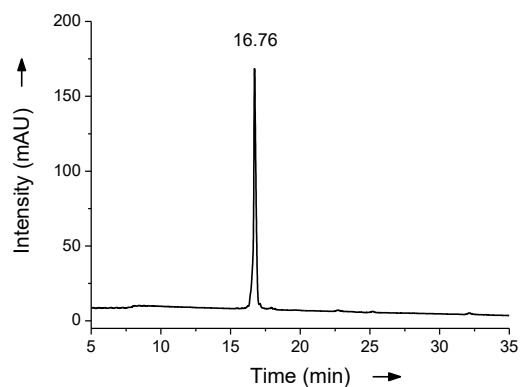


Figure F-27. HPLC chromatogram of purified peptide **F-26**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm..

F-27: FAM-Gly-Cys-Ala-Arg-Ser-Glu-Pro-Lys-Ile-Leu-Thr-Ahx-Ahx-Lys(FAM)-CONH₂; The synthesis was performed in a 10 μ mol scale. The C-terminal lysine was introduced as Fmoc-Lys(Alloc)-OH. After coupling the whole sequence, the Alloc group was selectively removed in presence of standard conditions.^{63,64} Pd(PPh₃)₄ (1.00 eq.), morpholine (190 eq.) and a solution of DCM with 2% water (final volume: 1.5 mL) were added to the resin under nitrogen atmosphere, and the mixture was shaken for 2 h. The resin was filtered off and washed with DMF (3 x 2.0 mL), a solution of sodium diethyldithiocarbamate (2 x 2.0 mL), DMF (3 x 2.0 mL) and CH₂Cl₂ (2 x 2.0 mL). Afterwards, 5,6-carboxyfluorescein (FAM) was coupled to the selectively deprotected lysine side chain, as well as to the N-terminus according to the standard coupling conditions (4.00 eq. Oxyma, DIC and 5-carboxyfluorescein in DMF). For the final cleavage step, the resin (scale = 10 μ mol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail (94% TFA/ 1% TIS/ 2.5% DODT/ 2.5% H₂O). After purification, the product as a 2 x TFA salt (0.4 mg, 0.162 μ mol, 2%) was obtained as a yellow solid. **HPLC:** t_R = 30.17 min. Purity = 98%. Formula: C₁₀₉H₁₄₂N₂₀O₃₀S. **HRMS-ESI⁺ (m/z):** [M+2H]²⁺ calcd.: 1122.5033; found: 1122.5059.

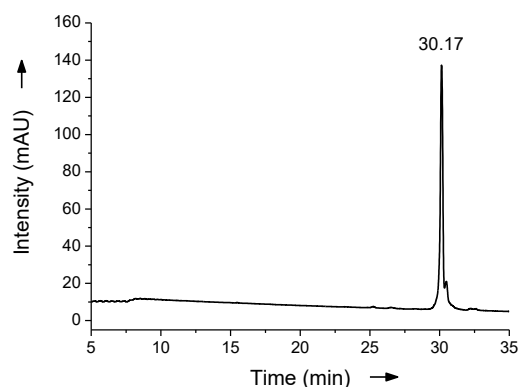


Figure F-28. HPLC chromatogram of purified peptide **F-27**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.

6.7.2 NuRD Complex

The nucleosome remodeling and deacetylase (NuRD) complex is involved in chromatin organization,⁶⁵ gene transcription,^{66,67} genomic stability and developmental signalling.^{68,69} It is the only known chromatin remodeler that combines ATPase and histone deacetylase functions. Such dual enzymatic activity is proposed to be important for efficient formation of the densely packed silent heterochromatin. NuRD consists of the following seven subunits (Figure F-29): 1) CHD3/4, an ATPase that utilizes energy derived from hydrolysis of ATP for DNA sliding and repositioning of nucleosomes;⁷⁰ 2) HDAC1/2 that deacetylate lysine side chains of histones or other proteins;⁶⁵ 3) MBD2/3 are required for complex formation and bind to methylated DNA;⁷¹⁻⁷³ 4) RbBP4/7 that act as histone chaperons and bind to the H4, thereby altering H4 conformation, thought to be essential in the ATP dependent chromatin remodeling process;⁷⁴⁻⁷⁶ 5) MTA1/2/3 that read histone tails and promoters and represent the most regulated genes in human cancer;⁷⁷⁻⁷⁹ 6) p66 α/β , a GATA zinc finger involved in binding unmodified histone tails,^{80,81} 7) Doc1, a tumor suppressor whose role is unclear until now.

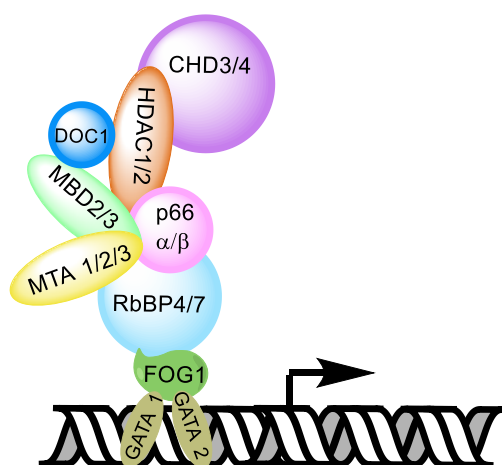


Figure F-29. Different subunits of the NuRD complex, interacting with GATA transcription factors through FOG1 motif.

Since aberrant activities of the NuRD complex are associated with human disease, including cancer⁸² and aging,⁸³ the NuRD components represent potential targets for therapeutic interventions. As a promising target protein, we chose RbBP4, which also belongs to the WD40 protein class as WDR5 (section 1.2), since it not only binds to the NuRD component MTA1 ($K_d = 33$ nM; $K_i = 308$ nM)⁸⁴ at the H4 binding site,⁸⁵ but also to the transcriptional regulator FOG-1 (friend of GATA; $K_d = 488$ nM; $K_i = 6.8$ μ M)⁸⁴ at the H3 binding site. FOG-1, in turn, interacts with the transcription factor GATA at specific genomic loci, promoting transcriptional repression of downstream targets.⁸⁶ Notably, only the 15 amino acid N-terminus of FOG-1 (MSRRKQSNPRQIKRSL) is required for NuRD binding via RbBP4/7 ($K_i = 1.54$ μ M)⁷⁶ and MTA1 (pull down experiments only).^{76,86} This, so called, FOG-repression motif can be found in various transcriptional repressors such as SALL1/4, BCL11A/B, Evi3, EBFAZ and Zeb-2, which are able to interact with NuRD concordantly.⁸⁶ The interaction between RbBP4 and MTA1 was identified to be mediated by a short α -helical peptide. It was demonstrated that the MTA1 residues 656-686 (KRAARR) resemble the RbBP4 binding motif of histone H4, leading to displacement of H4.⁸⁵ MTA1 and FOG-1 binding to RbBP4 do not hamper each other. Indeed an approach to crosslink two peptides, one bearing the MTA1 and one the FOG-1 motif, could be envisioned for further modulation of NuRD functions.

During this thesis, as a side project, various peptides based on the FOG-1 repression motif or on the MTA1 peptide were synthesized in order to be used for FP-based binding studies to RbBP4 in the future as well as for our collaboration partners of the lab of Prof. Dr. Alexander Brehm (Institute of Molecular Biology and Tumor Research (IMT), Philipps-University Marburg), and Prof. Dr. Sjaak Philipsen (Dept. of Cell Biology, Erasmus MC, Rotterdam), as part of the TRR81 (Transregional Collaborative Research Centre 81) project, to study their effects on NuRD function. The synthesized peptides are displayed in Figure F-30.

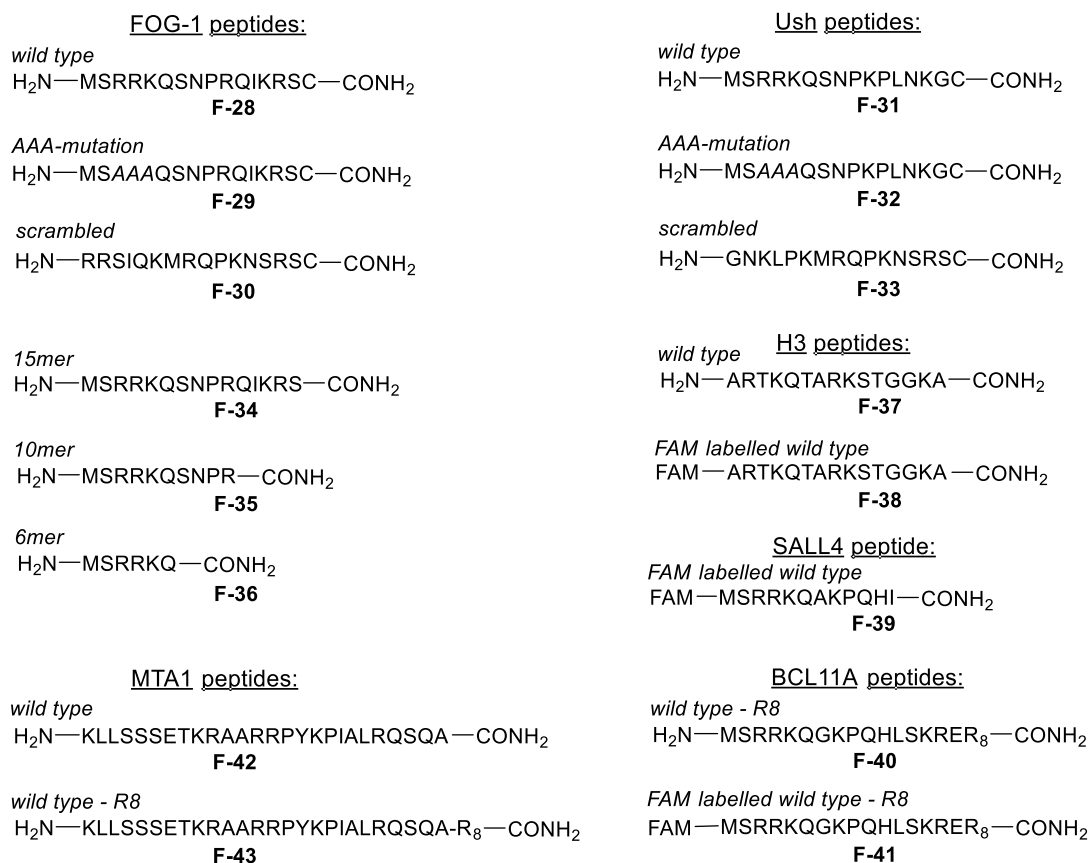


Figure F-30. Peptide sequences synthesized during this thesis for targeting the NuRD complex.

In order to explore FOG1-derived peptides with improved binding affinities to the NuRD complex, specifically to the RbBP4 protein, it would be beneficial to have an easy *in vitro* assay at hand that allows fast high throughput screening of peptide libraries. Thus, we aimed to apply fluorescence polarization (FP)-based assays to measure affinities of peptides containing the FOG repression motif to RbBP4. Initially, we performed preliminary FP-based binding experiments utilizing nuclear cellular extracts of S2 cells, which were proven via SDS analysis to contain the NuRD components. In order to determine the binding constants (K_d) to the protein of interest, it is necessary to have fluorescently labelled peptides (in the following labelled either with carboxyfluorescein (FAM, **F-44**) or fluorescein-isothiocyanate (FITC, **F-45**), see Figure F-31) to measure their saturation curve. Afterwards, such fluorescently labelled peptide can be utilized as a tracer for competitive FP-based assays to determine inhibition constants (K_i) of a library of potential, non-labelled, inhibitors.

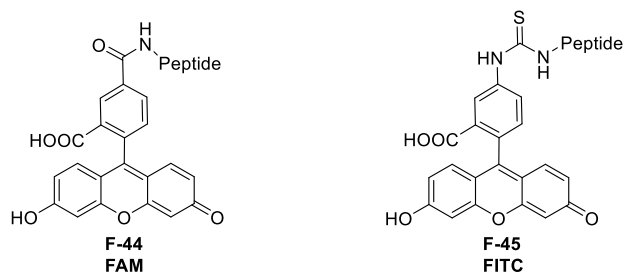


Figure F-31. Structures of FAM **F-44** and FITC **F-45**, used to label peptides, turning them into tracers suitable for FP-experiments.

In the context of RbBP4, there have been precedents of FP-based assays using the purified protein. Both, saturation FP experiments, as well as competitive based ones have been explored by different groups (Table F-5). Thus, Abbey et al.⁸⁴ determined the affinities of several RbBP4 binding peptides and performed optimizations of saturation FP experiments with RbBP4 and different tracer peptides, labelled with FITC. (Table F-5, Entries 1-3). At the same time, Moody et al.⁸⁷ developed a competitive FP-based assay, using FAM-H3 as the tracer (ARTKQTARKSTGGKAPRKQLA-K(FAM)); $K_d = 0.84 \pm 0.08 \mu\text{M}$), to evaluate the inhibition constants of BCL11A and SALL4 to RbBP4 (Table F-5, Entries 4-7).

Table F-5. Summarized results of saturation and competitive FP-experiments using purified RbBP4. *for saturation experiments, protein dilutions are utilized, while for competitive assays, the protein concentration is fixed.

Entry	Tracer peptide Sequence	Buffer	Highest / fixed* protein conc. (μM)	Fixed tracer conc. (nM)	Dynamic range (mP)	K_d / K_i (μM) to RbBP4
1	MTA1 ₆₅₆₋₆₈₆ : FITC-DV FYMATEETRKI RKLLSSE TKRAARRPYK	50 mM Sodium phosphate, pH 7.5, 150 mM KCl, 0.01% NP-40	0.15	10	0-140	0.033 \pm 0.0013
2	H3 ₁₋₂₁ : FITC-ARTKQTARKSTGG KAPRKQLA	10 mM Tris pH 8.0, 0.005% NP-40	3.0	5.0	0-80	1.4 \pm 0.2
3	FOG1 ₁₋₁₅ : FITC- MSRRKQSNPRQIKRS	50 mM Tris, pH 7.5, 0.005% NP-40, 2 Mm DTT	0.3	5.0	0-50	0.488 \pm 0.245
4	BCL11A ₂₋₁₂ : SRRKQGK PQHL	50 mM Tris pH 7.5, 50 mM NaCl, 0.1% glycerol	1.0	20	50-120	0.27
5	SALL4 ₂₋₁₂ : SRRKQAK PQHI	50 mM Tris pH 7.5, 50 mM NaCl, 0.1% glycerol	1.0	20	50-120	0.09

6	SALL4 ₁₋₁₀ : MSRRKQAKPQ	50 mM Tris pH 7.5, 50 mM NaCl, 0.1% glycerol	1.0	20	50-120	14.6
7	BCL11A ₂₋₁₂ ; <i>scr.</i> (sequence not displayed in publication)	50 mM Tris pH 7.5, 50 mM NaCl, 0.1% glycerol	1.0	20	50-120	n.c.

Those examples all utilized the purified RbBP4 protein for their in vitro FP-based assays. However, we firstly aimed at applying nuclear cell extracts containing the NuRD components, including RbBP4 protein, wherefore further adjustments had to be carried out. Some successful attempts, not based on the RbBP4 protein though, to utilize FP-based methods on crude cell extracts to obtain saturation binding curves have been reported (Table F-6, Entries 1-2).⁸⁸⁻⁹⁰ In these examples, however, binding of the tracer to the protein of interest is in the nanomolar range, which is in contrast to peptides containing the FOG-repression motif, where binding to RbBP4 is in the micromolar range.

In addition, there are also precedents of FP-based assays utilizing nuclear cell extracts and FOG-containing peptides, however, no saturation experiments but only competitive FP-assays were performed. There, SNU-398 nuclear cell extracts and FITC-SALL4 (FITC-MSRKQAKPQHI-CONH₂) as a tracer at 1 μ M were applied. This allowed to determine IC₅₀ values of wt-SALL4 and demonstrated that it is, in contrast to the scrambled sequence, binding to the NuRD complex (Table F-6, Entries 3-4).⁹¹

Table F-6. Summarized results of saturation FP-experiments using crude cell extracts. *for saturation experiments, protein dilutions are utilized, while for competitive assays, the protein concentration is fixed.

Entry	Peptide Sequence	Buffer	Highest / fixed* extract conc. (μ M)	Fixed tracer conc. (nM)	Dynamic range (mP)	IC ₅₀ in cell extracts
1	FAM-labelled SAHA (HDAC inhibitor) ⁹⁰	50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM KCl, 1 mM MgCl ₂ , 0.25 mg/mL BGG	200 μ g/mL HeLa extracts	2.0	0-0.9 (normalized)	HeLa: 248 μ g/mL K _d purified HDAC: 2.0 nM
2	FAM labelled c-myc peptide: ⁸⁹ EQKILISEEDLGGC-FAM	PBS-buffer + 0.2 mg/mL BSA	0-200 nM c-myc antibody	2.5	50-170	10 nM
3	SALL4 _{wt} : MSRKQAKPQHI	Not specified	from 2 x 10 ⁷ SNU-398 cells	1.0	20-350	0.5 μ M
4	SALL4 _{scr.} (sequence not displayed in publication)	Not specified	from 2 x 10 ⁷ SNU-398 cells	1.0	350 (no curve)	n.c.

Due to the complexity of RBbP4 (or p55, the *drosophila* analogue) protein expression in insect cells, we initially explored the possibility of using nuclear cell extracts to determine the K_d value of tracer peptides FAM-SALL4 and FAM-H3. For this purpose, we utilized S2 / S2-cas9 nuclear cell extracts, as well as cell extracts from MTA1-Flag expressing cells in order to obtain a higher local concentration of the NuRD components in the extracts (prepared by Jonathan Lenz, group of Prof. Dr. Brehm) and, finally, Mel cell extracts (prepared by Anna Hsieh, AG Philipson). Different concentrations of the extracts in different buffers were employed (all conditions, sequences and results are summarized in Table F-7). The blots of the obtained results are presented in Figure F-32 to Figure F-34.

Table F-7. Comparison of the results obtained from our saturation FP-experiments. Buffer C: 20 mM Hepes pH = 7.9, 420 mM NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA; is the buffer in which the nuclear cell extracts are usually obtained. FP buffer: 0.1 M phosphate pH = 6.5, 25 mM KCl, 0.01% Triton; is the buffer usually utilized in our laboratory for FP-based assays.

Peptide Sequence	Buffer	Highest extract conc. (mg/mL)	Fixed tracer conc. (nM)	Dynamic range (mP)	IC ₅₀ (μM) to RBbP4
F-39 , FAM-SALL4A: FAM-MSRRKQAKPQHI	Buffer C	2.96 (S2-wt cells)	20	50	n.c.
F-39 , FAM-SALL4A: FAM-MSRRKQAKPQHI	Buffer C	3.06 (S2-cas9 cells)	20	50	n.c.
F-39 , FAM-SALL4A: FAM-MSRRKQAKPQHI	FP-Buffer	2.01 (S2-wt cells)	20	50-65	n.c.
F-39 , FAM-SALL4A: FAM-MSRRKQAKPQHI	FP-Buffer	1.99 (S2-cas9 cells)	20	50-65	n.c.
F-38 , FAM-H3: FAM-ARTQTARKSTGGKA	Buffer C	2.96 (S2-wt cells)	20	50	n.c.
F-38 , FAM-H3: FAM-ARTQTARKSTGGKA	Buffer C	3.06 (S2-cas9 cells)	20	50	n.c.
F-38 , FAM-H3: FAM-ARTQTARKSTGGKA	FP-Buffer	2.01 (S2-wt cells)	20	50-65	n.c.
F-38 , FAM-H3: FAM-ARTQTARKSTGGKA	FP-Buffer	1.99 (S2-cas9 cells)	20	50-65	n.c.
F-39 , FAM-SALL4A: FAM-MSRRKQAKPQHI	FP-Buffer	Conc. not known (MTA1-Flag)	20	50	n.c.
F-39 , FAM-SALL4A: FAM-MSRRKQAKPQHI	FP-Buffer	4.18 (Mel cells)	20	45-75	n.c.
F-38 , FAM-H3: FAM-ARTQTARKSTGGKA	FP-Buffer	4.18 (Mel cells)	20	45	n.c.

S2 and S2 cas9 nuclear cell extracts in buffer C (20 mM HEPES pH = 7.9, 420 mM NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA) did not show any increase of mP values with increasing extract concentration for either of the tracer peptides (Figure F-32). However, cellular extracts in FP-buffer (0.1 M phosphate pH = 6.5, 25 mM KCl, 0.01% Triton) presented the start of a slight increase of mP values (Figure F-32). Therefore, for further studies we focused on cell extracts obtained in the FP-buffer. Surprisingly, utilization of MTA1-Flag expressing nuclear cell extracts did not lead to an increased binding (Figure F-33). However, utilization of Mel extracts in FP-buffer, and the SALL4 peptide as a tracer, displayed the highest increase in mP values (Figure F-34). The higher dynamic range of SALL4 over the H3-tracer (Figure F-34), results from the higher binding affinity to RbBP3 of SALL4 compared to H3. Although these results were promising, since the beginning of the sigmoidal curve in the dose-response inhibition experiments was observed, the concentration of RbBP4 protein inside the nuclear cell extracts was too low to reach the upper saturation plateau.

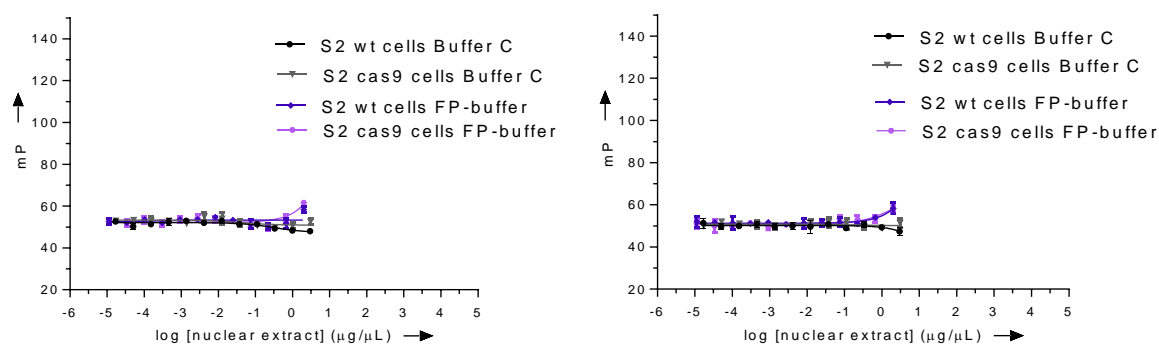


Figure F-32. FP-based saturation experiments using two different S2 cell extracts in different buffers: S2 wt in Buffer C ($c = 2.96 \mu\text{g}/\mu\text{L}$); S2 cas9, Buffer C ($c = 3.06 \mu\text{g}/\mu\text{L}$); S2 wt, FP buffer ($c = 2.01 \mu\text{g}/\mu\text{L}$); S2 cas9, FP buffer ($c = 1.99 \mu\text{g}/\mu\text{L}$). Left: FAM-SALL4 tracer **F-39**. Right: FAM-H3 tracer **F-38**.

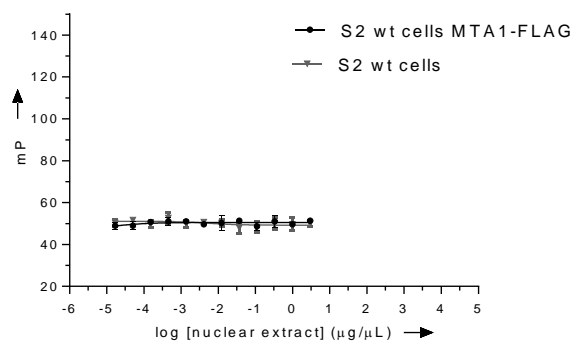


Figure F-33. FP-based saturation experiments using FAM-SALL4 tracer and extracts of MTA1-Flag expressing cells to enrich NuRD components (FP-buffer).

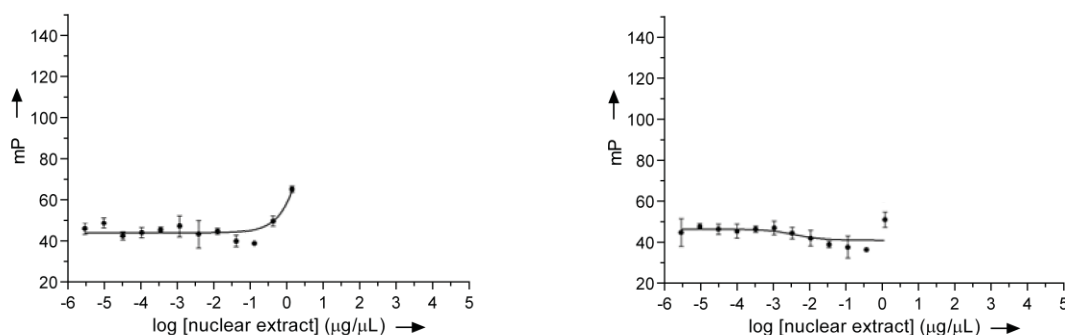


Figure F-34. FP-based saturation experiments using Mel cell extracts (FP-buffer ($c = 4.18 \mu\text{g}/\mu\text{L}$)). Left: FAM-SALL4 tracer **F-39**. Right: FAM-H3 tracer **F-38**.

It is known that the requirements on the protein and ligand concentrations in a FP-based assay, depend on the binding affinity. Thus, the lower the binding affinity of a given peptide to the protein of interest, the higher is the required protein concentration and the lower should be the concentration of the tracer (pictured in Figure F-35).⁹² For competitive FP-based assays, typically a fixed tracer concentration near the K_d value is chosen and a protein concentration, which results in between 50-80% of the bound ligand (Figure F-35).⁹²

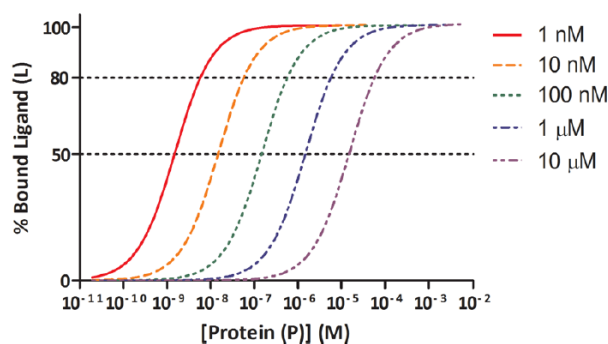


Figure F-35. Exemplified binding saturation curves for different K_d values. The higher the K_d value, the higher the required protein concentration, and the lower the optimal fixed tracer concentration.⁹²

Due to the fact that the binding affinities of peptides containing the FOG repression motif to RbBP4 are in the low micromolar range (K_d : FOG = 0.488 μM ;⁸⁴ H3 = 1.4 μM ;⁸⁴ K_i : BCL11A = 0.268 μM ;⁸⁷ SALL4 = 0.09 μM ⁸⁷), very high concentrations of RbBP4 protein are necessary to reach the saturation plateau of the binding curve. It seems that such high concentration cannot be reached by using nuclear cell extracts, which might be a reason why Gao et al.⁹¹ only performed competitive FP-based assays with cell extracts instead of saturation binding experiments.⁹¹ Since for us, it is crucial to have the saturation binding curve, in order to provide the binding constant (K_d) of a suitable tracer to RbBP4, which is necessary to evaluate not only the IC_{50} but also the K_i values of a library of optimized peptides upon their inhibition function, it was necessary to obtain the purified RbBP4 protein. Successful expression of p55, the drosophila analogue to RbBP4, has been performed and FP-experiments with such are at the moment ongoing in the Vázquez group.

As soon as this initial screening will reveal first hit compounds, those could be further optimized to increase their binding affinity. For example, it could be taken advantage of the fact that the MTA1 peptide engages a helical conformation upon binding, by utilizing stapled peptidomimetics to increase binding affinity, cellular uptake and resistance against proteases. Just recently Waldmann et al.⁹³ published bicyclic peptide inhibitors targeting the PPI between RbBP4 and MTA-1 in the low nanomolar range. They applied structure-based design to optimize binding affinity of a parental linear MTA-1 peptide. Their findings of increased affinity upon exchange of polar amides by hydrophobic aromatic linkers allowing mono- and bicyclization via cysteine alkylation, should be considered in order to further stimulate binding affinity. Moreover, the PROTAC technique could be applied by linking a thalidomide moiety to such a parental peptide to enable specific degradation of the POI to further study and modulate NuRD function. The NuRD component SALL4 has already been reported to be a substrate of the thalidomide-mediated formation of Cereblon E3 ligase complex formation^{94,95} and its, thereby, induced degradation has been reported.⁹⁶ Furthermore, thalidomide-based PROTACS have been utilized to induce degradation of histone deacetylases HDAC 1-3 in HEK293T cells.⁹⁷

The next steps, after FP-screening, would be to perform pull-down and co-immunoprecipitation assays to evaluate the ability to target the NuRD complex via disrupting its PPIs. Furthermore, FAM labelled analogue peptides could be synthesized to verify cellular uptake via confocal microscopy, and finally, the biological effects of the peptides would need to be evaluated in vitro and in vivo.

Jonathan Lenz (group of Prof. Dr. Brehm, IMT, Philipps-University Marburg) was already performing initial pull-down peptide competition assays using nuclear extracts from S2 cells containing the NuRD subunits and a GST-fusion protein comprising the N-terminal (residues: 1-45) mFOG-1. The ability of Ush (FOG-1 analogue in *drosophila*) and FOG-1 peptides to compete for NuRD binding with the GST-mFOG-1 peptide was assessed. The results are represented in Figure F-36 and demonstrate that both peptides, Ush and FOG-1, are able to bind the NuRD complex, while the mutated and scrambled sequences did not show any affinity, as can be deduced by the fully intact pull-down potential upon addition of 50 × the concentration of the Ush scrambled and mutated peptides. Notably, the FOG-1 was effective at 10 × lower concentration than its Ush analogue, encouraging further developments into the utilization of FOG-1 peptides instead of Ush ones. These results of our synthesized peptides were included in a manuscript of the group of Prof. Dr. Alexander Brehm (first author: Jonathan Lenz; doi: <https://doi.org/10.1101/2020.06.10.143701>), about the multi-zinc finger protein Ush, which maintains multipotency of stem cell-like hemocyte progenitors during *drosophila* hematopoiesis.⁹⁸ It was demonstrated that Ush controls the expression of gene classes encoding proteins relevant to stem cell-like functions and differentiation. It was further discovered that Ush and NuRD shape the proliferation, metabolism and hemocyte-related functions of S2 cells via maintaining an extensive gene expression programme. As S2 cells are derived from embryonic hemocytes this indicates that Ush and NuRD hold gene regulatory roles during embryonic hematopoiesis. Specifically, Ush and NuRD subunits were required to restrict enhancer activity to cells of the prosterior signaling center in L3 larva lymph glands. Thus, verifying that Ush and NuRD are able to regulate enhancer activity during hemocyte differentiation in vivo.

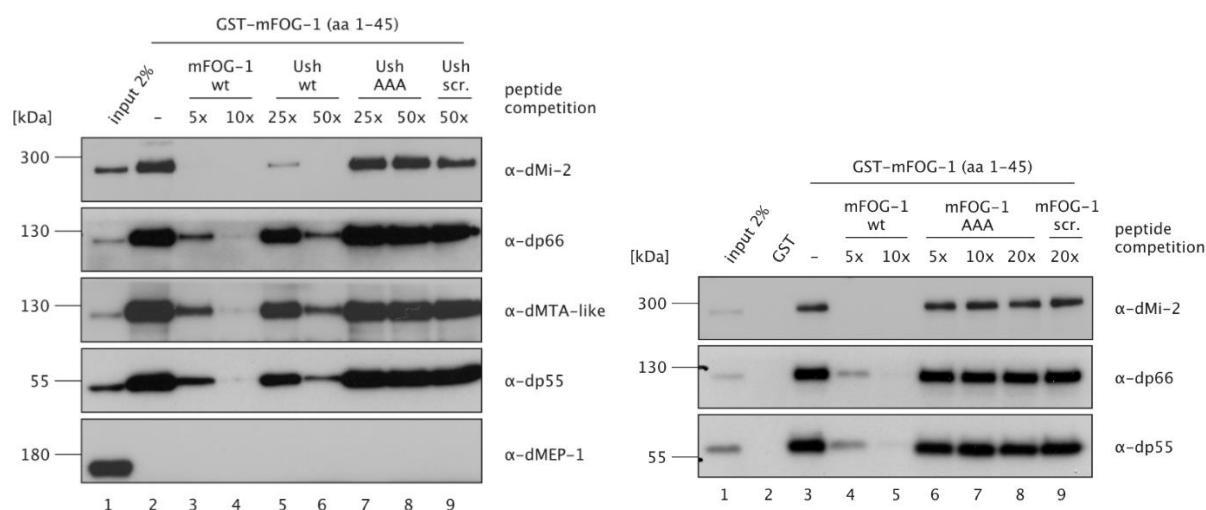


Figure F-36. Pull-down assay performed by Jonathan Lenz (AG Brehm) to test ability of synthesized FOG-1 and Ush peptides to compete against GST-FOG-1 in binding to NuRD complex. Excess of peptide competitors over GST-mFOG ranged between 5 × and 50 × as illustrated. wt = wild type sequence, AAA = triple Ala mutation, scr: scrambled sequence. Left: mFOG-1 wt **F-28**, Ush wt **F-31**, Ush AAA **F-32**, Ush scr. **F-33** sequences. Right: FOG-1 wt **F-28**, FOG-1 AAA **F-29**, FOG-1 scr. **F-30** sequences.

Experimental Procedures

Peptide Synthesis: The synthesis of the linear peptides was performed manually according to the standard Fmoc-solid-phase methodology using rink amide resin following the protocol described above in the manuscripts B-C in section 4. After final deprotection and cleavage (cleavage cocktail: 94% TFA/ 1% TIS/ 2.5% DODT/ 2.5% H₂O), purification was performed via semi-preparative RP-HPLC (eluent: water + 0.1% TFA, B: MeCN + 0.1% TFA; gradient: 5-40% B, except peptide **F-36**: 0-30% B). In the following, the characterization data of the final peptides are displayed.

F-28: *FOG-1 wt*: H₂N-Met-Ser-Arg-Arg-Lys-Gln-Ser-Asn-Pro-Arg-Gln-Ile-Lys-Arg-Ser-Cys-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 7 × TFA salt (8.40 mg, 3.03 μmol, 30%) was obtained as a white solid. **HPLC:** t_R = 10.58 min. Purity = 98%. Formula: C₇₈H₁₄₄N₃₄O₂₂S₂. **HRMS-ESI⁺ (m/z):** [M+3H]³⁺ calcd.: 658.6946; found: 658.6942.

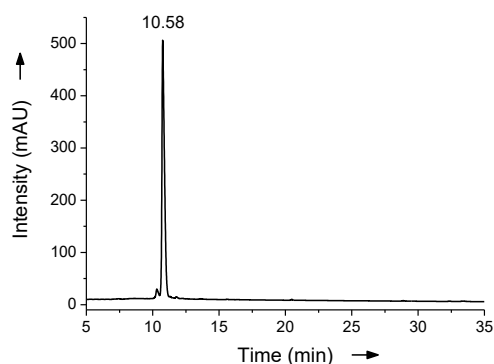


Figure F-37. HPLC chromatogram of purified peptide **F-28**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm.

F-29: *FOG-1 AAA-mut*: H₂N-Met-Ser-Ala-Ala-Ala-Gln-Ser-Asn-Pro-Arg-Gln-Ile-Lys-Arg-Ser-Cys-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 4 × TFA salt (9.00 mg, 4.09 μmol, 41%) was obtained as a white solid. **HPLC:** t_R = 12.99 min. Purity = 91%. Formula: C₆₉H₁₂₃N₂₇O₂₂S₂. **HRMS-ESI⁺ (m/z):** [M+2H]²⁺ calcd.: 873.9461; found: 873.9451.

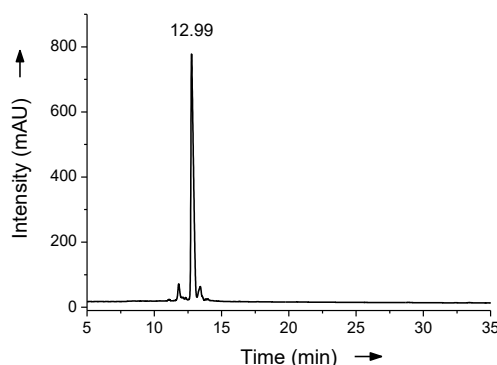


Figure F-38. HPLC chromatogram of purified peptide **F-29**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm.

F-30: *FOG-1 scr*: H₂N-Arg-Arg-Ser-Ile-Gln-Lys-Met-Arg-Gln-Pro-Lys-Asn-Ser-Arg-Ser-Cys-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 7 × TFA salt (9.50 mg, 3.43 μmol, 34%) was obtained as a white solid. **HPLC:** t_R = 11.46 min. Purity = 98%. Formula: C₇₈H₁₄₄N₃₄O₂₂S₂. **HRMS-ESI⁺** (**m/z**): [M+3H]³⁺ calcd.: 658.6946; found: 658.6942.

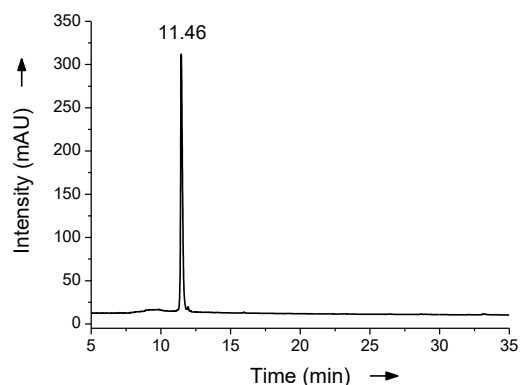


Figure F-39. HPLC chromatogram of purified peptide **F-30**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm.

F-31: *Ush wt*: H₂N-Met-Ser-Arg-Arg-Lys-Gln-Ser-Asn-Pro-Lys-Pro-Leu-Asn-Lys-Gly-Cys-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 6 × TFA salt (14.0 mg, 5.54 μmol, 55%) was obtained as a white solid. **HPLC:** t_R = 11.30 min. Purity = 93%. Formula: C₇₅H₁₃₅N₂₉O₂₁S₂. **HRMS-ESI⁺** (**m/z**): [M+2H]²⁺ calcd.: 921.9987; found: 921.9977.

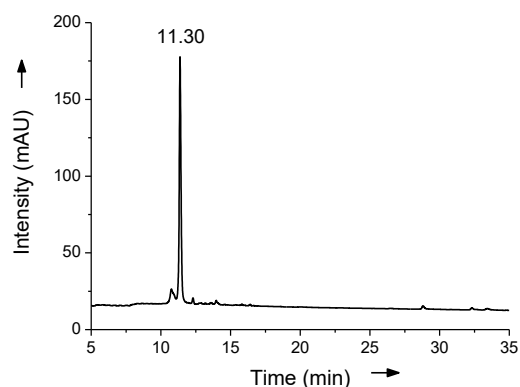


Figure F-40. HPLC chromatogram of purified peptide **F-31**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm.

F-32: *Ush AAA-mut*: H₂N-Met-Ser-Ala-Ala-Ala-Gln-Ser-Asn-Pro-Lys-Pro-Leu-Asn-Lys-Gly-Cys-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 3 × TFA salt (7.60 mg, 3.88 μmol, 39%) was obtained as a white solid. **HPLC:** t_R = 13.63 min. Purity = 91%. Formula: C₆₆H₁₁₄N₂₂O₂₁S₂. **HRMS-ESI⁺** (**m/z**): [M+2H]²⁺ calcd.: 808.4058; found: 808.4047.

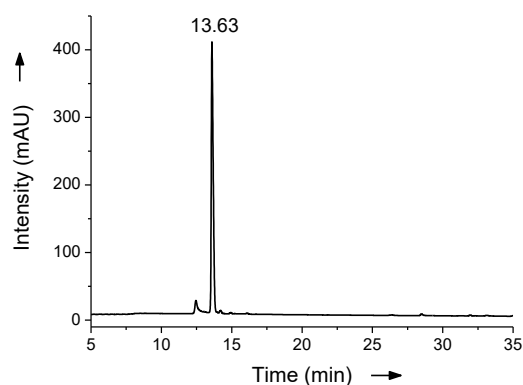


Figure F-41. HPLC chromatogram of purified peptide **F-32**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm.

F-33: *Ush scr*: H₂N-Gly-Asn-Lys-Leu-Pro-Lys-Met-Arg-Gln-Pro-Lys-Asn-Ser-Arg-Ser-Cys-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 6 × TFA salt (10.0 mg, 3.96 μmol, 40%) was obtained as a white solid. **HPLC:** t_R = 13.15 min. Purity = 98%. Formula: C₇₅H₁₃₅N₂₉O₂₁S₂. **HRMS-ESI⁺** (**m/z**): [M+2H]²⁺ calcd.: 921.9978; found: 921.9976.

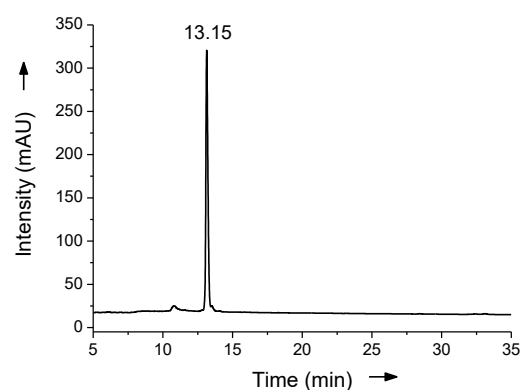


Figure F-42. HPLC chromatogram of purified peptide **F-33**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm.

F-34: *FOG-1 15mer*: H₂N-Met-Ser-Arg-Arg-Lys-Gln-Ser-Asn-Pro-Arg-Gln-Ile-Lys-Arg-Ser-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail (82.5% TFA/ 5% phenol/ 5% H₂O/ 5% thioanisole/ 2.5 DODT). After purification, the product as a 7 × TFA salt (4.80 mg, 1.80 μmol, 18%) was obtained as a white solid. **HPLC**: t_R = 18.98 min. Purity = 98%. Formula: C₇₅H₁₃₉N₃₃O₂₁S. **HRMS-ESI⁺ (m/z)**: [M+2H]²⁺ calcd.: 936.0345; found: 936.0352.

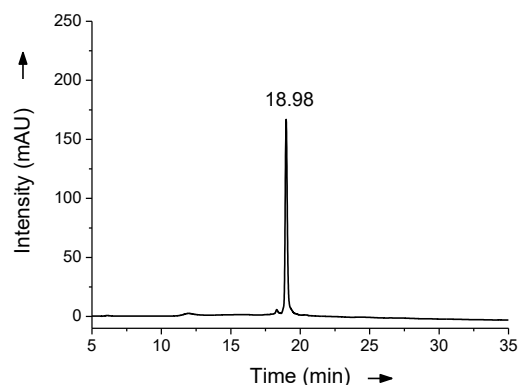


Figure F-43. HPLC chromatogram of purified peptide **F-34**. An EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 0-30% B monitored at 220 nm.

F-35: *FOG-1 10mer*: H₂N-Met-Ser-Arg-Arg-Lys-Gln-Ser-Asn-Pro-Arg-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail (82.5% TFA/ 5% phenol/ 5% H₂O/ 5% thioanisole/ 2.5% DODT). After purification, the product as a 5 × TFA salt (7.40 mg, 4.05 μmol, 40%) was obtained as a white solid. **HPLC**: t_R = 15.88 min. Purity = 98%. Formula: C₄₉H₉₁N₂₃O₁₄S₁. **HRMS-ESI⁺ (m/z)**: [M+2H]²⁺ calcd.: 629.8484; found: 629.8495.

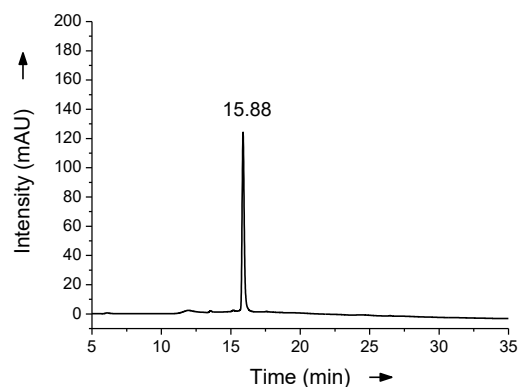


Figure F-44. HPLC chromatogram of purified peptide **F-35**. An EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 0-30% B monitored at 220 nm.

F-36: *FOG-1 6mer*: H₂N-Met-Ser-Arg-Arg-Lys-Gln-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail (82.5% TFA/ 5% phenol/ 5% H₂O/ 5% thioanisole/ 2.5% DODT). After purification, the product as a 4 × TFA salt (3.20 mg, 2.54 μmol, 25%) was obtained as a white solid. **HPLC**: t_R = 5.00 min. Purity = 98%. Formula: C₃₁H₆₁N₁₅O₈S₁. **HRMS-ESI⁺ (m/z)**: [M+H]⁺ calcd.: 804.4621; found: 804.4622.

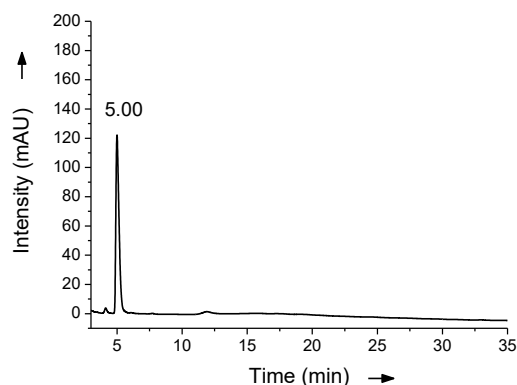


Figure F-45. HPLC chromatogram of purified peptide **F-36**. An EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 0-30% B monitored at 220 nm.

F-37: *H3 wt*: H₂N-Ala-Arg-Thr-Lys-Gln-Thr-Ala-Arg-Lys-Ser-Thr-Gly-Gly-Lys-Ala-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 6 × TFA salt (2.90 mg, 1.29 μmol, 13%) was obtained as a white solid. **HPLC**: t_R = 15.24 min. Purity = 95%. Formula: C₆₃H₁₁₈N₂₆O₂₀. **HRMS-ESI⁺ (m/z)**: [M+2H]²⁺ calcd.: 780.4581; found: 780.4582.

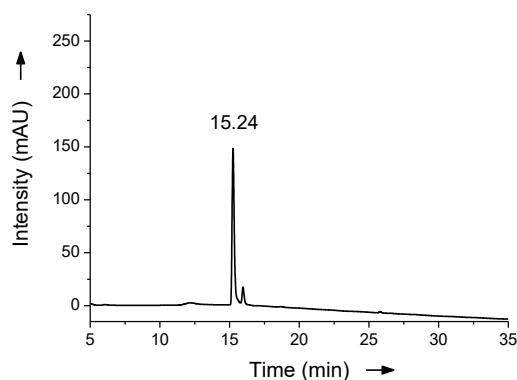


Figure F-46. HPLC chromatogram of purified peptide **F-37**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 0-30% B monitored at 220 nm.

F-38: *FAM-H3*: FAM-Ala-Arg-Thr-Lys-Gln-Thr-Ala-Arg-Lys-Ser-Thr-Gly-Gly-Lys-Ala-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 5 × TFA salt (9.20 mg, 3.70 μmol, 37%) was obtained as a white solid. **HPLC:** t_R = 12.36 min. Purity = 94%. Formula: C₈₄H₁₂₈N₂₆O₂₆. **HRMS-ESI⁺ (m/z):** [M+3H]³⁺ calcd.: 639.9904; found: 639.9905.

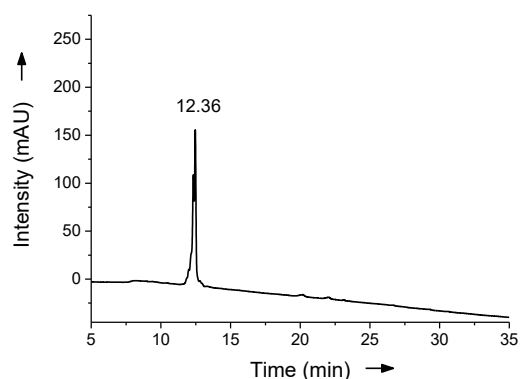


Figure F-47. HPLC chromatogram of purified peptide **F-38**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 0-30% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.

F-39: *FAM-SALL4*: FAM-Met-Ser-Arg-Arg-Lys-Gln-Ala-Lys-Pro-Gln-His-Ile-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail (82.5% TFA/ 5% phenol/ 5% H₂O/ 5% thioanisole/ 2.5% DODT). After purification, the product as a 5 × TFA salt (3.80 mg, 1.58 μmol, 16%) was obtained as a white solid. **HPLC:** t_R = 16.62 & 17.05 min (2 FAM isomers). Purity = 89% (impurity: oxidized product = oxidized methionine side chain). Formula: C₈₃H₁₂₁N₂₅O₂₁S₁. **HRMS-ESI⁺ (m/z):** [M+3H]³⁺ calcd.: 612.9701; found: 612.9708.

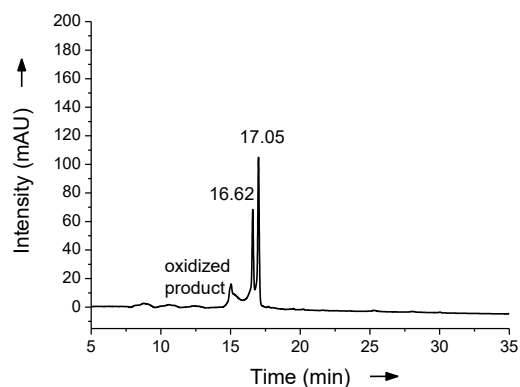


Figure F-48. HPLC chromatogram of purified peptide **F-39**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-50% B monitored at 220 nm. Oxidized product = oxidized methionine side chain. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.

F-40: *BCL11A wt-R8*: $\text{H}_2\text{N-Met-Ser-Arg-Arg-Lys-Gln-Gly-Lys-Pro-Gln-His-Leu-Ser-Lys-Arg-Glu-(Arg)}_8\text{-CONH}_2$; For the final cleavage step, the resin (scale = 10 μmol , loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 16 \times TFA salt (14.8 mg, 2.94 μmol , 29%) was obtained as a white solid. **HPLC:** $t_R = 19.22$ min. Purity = 99%. Formula: $\text{C}_{129}\text{H}_{241}\text{N}_{65}\text{O}_{30}\text{S}_1$. **HRMS-ESI⁺ (m/z):** $[\text{M}+5\text{H}]^{5+}$ calcd.: 643.5882; found: 643.5885.

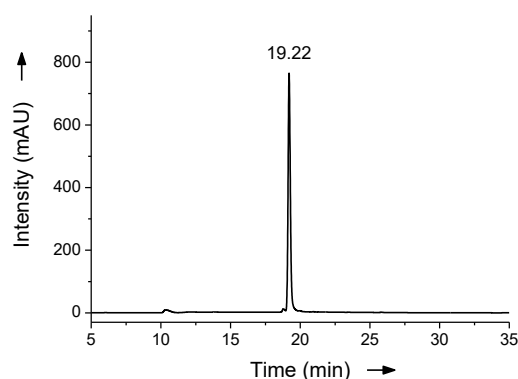


Figure F-49. HPLC chromatogram of purified peptide **F-40**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 0-30% B monitored at 220 nm.

F-41: *FAM-BCL11A-R8*: $\text{FAM-Met-Ser-Arg-Arg-Lys-Gln-Gly-Lys-Pro-Gln-His-Leu-Ser-Lys-Arg-Glu-(Arg)}_8\text{-CONH}_2$; For the final cleavage step, the resin (scale = 10 μmol , loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 15 \times TFA salt (6.00 mg, 1.14 μmol , 11%) was obtained as a white solid. **HPLC:** $t_R = 28.06$ min. Purity = 98%. Formula: $\text{C}_{150}\text{H}_{251}\text{N}_{65}\text{O}_{36}\text{S}$. **HRMS-ESI⁺ (m/z):** $[\text{M}+5\text{H}]^{5+}$ calcd.: 715.1979; found: 715.1982.

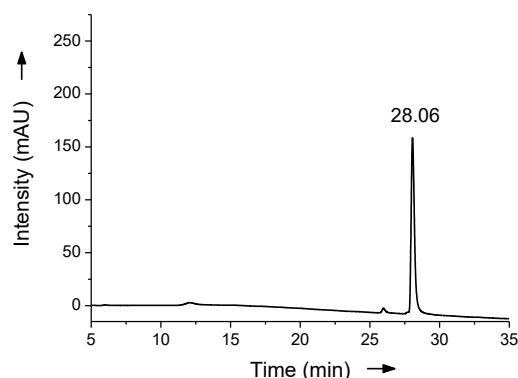


Figure F-50. HPLC chromatogram of purified peptide **F-41**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 0-30% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.

F-42: *MTA1 wt*: H₂N-Lys-Leu-Leu-Ser-Ser-Ser-Glu-Thr-Lys-Arg-Ala-Ala-Arg-Arg-Pro-Tyr-Lys-Pro-Ile-Ala-Leu-Arg-Gln-Ser-Gln-Ala-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 8 × TFA salt (1.10 mg, 0.284 μmol, 3%) was obtained as a white solid. **HPLC:** t_R = 20.53 min. Purity = 90%. Formula: C₁₂₈H₂₂₄N₄₄O₃₆. **HRMS-ESI⁺ (m/z):** [M+5H]⁵⁺ calcd.: 591.7481; found: 591.7485.

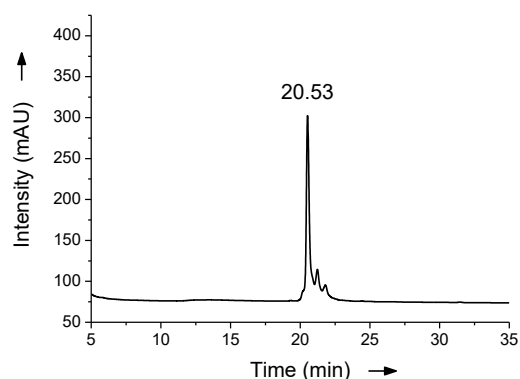


Figure F-51. HPLC chromatogram of purified peptide **F-42**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-40% B monitored at 220 nm.

F-43: *MTA1 wt-R8*: H₂N-Lys-Leu-Leu-Ser-Ser-Ser-Glu-Thr-Lys-Arg-Ala-Ala-Arg-Arg-Pro-Tyr-Lys-Pro-Ile-Ala-Leu-Arg-Gln-Ser-Gln-Ala-(Arg)₈-CONH₂; For the final cleavage step, the resin (scale = 10 μmol, loading = 0.25 mmol/g) was treated with 1.50 mL cleavage cocktail. After purification, the product as a 16 × TFA salt (3.00 mg, 0.498 μmol, 5%) was obtained as a white solid. **HPLC:** t_R = 26.76 min. Purity = 92%. Formula: C₁₇₆H₃₂₀N₇₆O₄₄. **HRMS-ESI⁺ (m/z):** [M+5H]⁵⁺ calcd.: 841.5101; found: 841.5109.

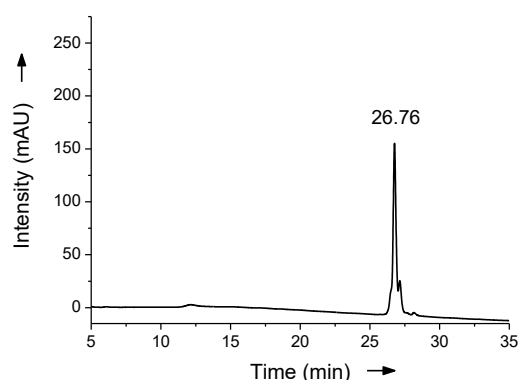


Figure F-52. HPLC chromatogram of purified peptide **F-43**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 0-30% B monitored at 220 nm.

6.7.3 tRNA-Guanine Transglycosylase Enzyme

The tRNA-guanine transglycosylase (TGT) homodimer catalyzes the base exchange of the genetically encoded guanine on the wobble position of particular tRNAs with the modified base 7-cyano-7-deaza-guanine (preQ1).⁹⁹ This can result in the development of a virulence phenotype of *Shigella*, which are bacteria that constitute the causative agent of bacillary dysentery, an acute inflammatory disease triggering the death of more than one million humans per year.¹⁰⁰ Thus, TGT represents an attractive drug target. It is known that one TGT monomer performs the catalytic base exchange reaction, while the second subunit is required to correctly position the tRNA substrate. Therefore, TGT only displays biological activity, when present in its homodimer form.^{101,102} Hence, disruption of its protein-protein interaction offers an attractive opportunity for therapeutic intervention. For a collaboration project with Dzung Nguyen (group of Prof. Dr. Klebe, Pharmacy Department, Philipps-University of Marburg), several interface-derived peptides have been designed and synthesized, which should interact with the TGT and inhibit enzymatic activity. The dimer interface is stabilized via a hotspot of four aromatic amino acids: Trp326, Tyr330, Phe92, His333, forming a well-organized hydrophobic patch.^{103,104} In comparison to the wild type enzyme, which is almost exclusively present in its dimer form, replacement of the Tyr330 by Phe and Trp326 by Glu results in a drastically loss of dimer stability.¹⁰⁴ In addition, interface properties are characterized by a loop-helix motif (loop region residues: 45-63; helical residues: 55-63). It was suggested that this motif plays a crucial role for dimer formation and stability as it is dynamically disordered in the monomeric solution, while it takes on an ordered formation upon dimerization.¹⁰³ The Klebe group performed site-directed mutations on the important aromatic amino acids, yielding in destabilized mutated variants, where the loop-helix motif adopts different conformations in the interface region.¹⁰⁵ Furthermore, they were able to demonstrate that the integrity of the aromatic hotspot cluster is essential for dimer stability. Among the studied variants, the H333D one displayed the largest structural variations compared to the wild type. While the wild type exhibits a dimer:monomer ratio of 98:2, the H333D variants ratio is proposed to be 42:58.¹⁰⁵

We aimed to develop peptides for our collaborators that would be able to disrupt the TGT homodimer, and thus, hamper TGT function. Binding affinities of our designed and synthesized peptides should then be evaluated by Dzung Nguyen using FP-based assays, wherefore the peptides had to be labelled with a FAM-tag. Our collaborators performed initial epitope- and alanine scans utilizing a peptide-array (SPOT assay technique using a peptide microarray purchased from peptides&elephants GmbH (Hennigsdorf, Germany)) on the TGT wild type, which indicated that the peptides: YIHHLIRAGE, SRAYIHHLIRAGE and ARAYIHHLIRAGE, are promising candidates. Therefore, we synthesized these three variants containing the hotspot residues Tyr330, His333 and Glu339 and a FAM-tag either at their N- or C-termini (Figure F-53). As the peptides engage a helical conformation upon binding, we additionally synthesized stapled peptide analogues **F-46-F-52** (Figure F-53) in order to probe a possibly increased binding affinity to the TGT enzyme. Preliminary FP-based assays, conducted by our collaboration partner Dzung Nguyen, utilizing the H333D TGT variant in combination with stapled peptide **F-52** revealed the most promising results. Although, the obtained dose-response curve proposed binding of **F-52** to the H333D variant (Figure F-54), the experiments revealed that the binding curve was shifting over time (Figure F-54), apparently improving binding affinity. To find an explanation for such behavior, the reducing agent TCEP was added to the peptide solution, which was reverting the process and yielded, again, lower binding affinities. This led to the hypothesis, that oxidation of peptide **F-52** over time, might yield in a more potent peptide. However, experiments of our collaborators to study this in more detail are still undergoing at the moment.

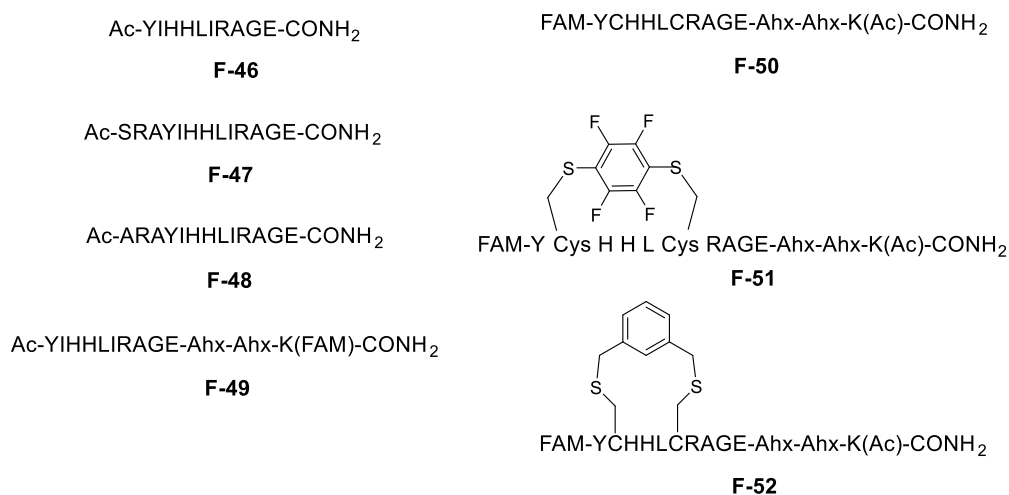


Figure F-53. Peptides **F-44-F-50**, synthesized by us for the collaboration project with the Klebe group.

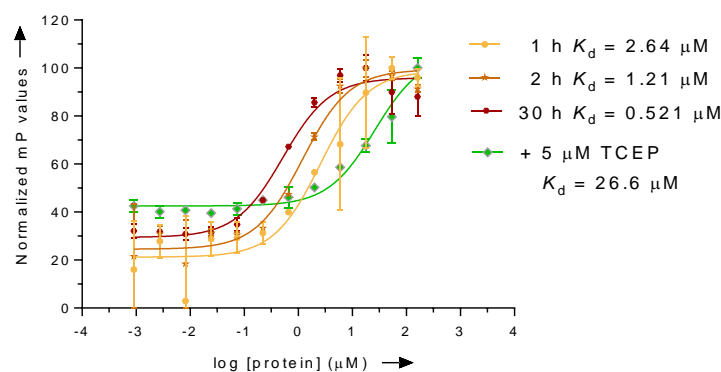


Figure F-54. Dose-response binding curve of peptide FF-50 to H333D TGT variant. mP values are recorded after the indicated time periods, and after addition of TCEP. Blot represents one independent measurement, each concentration in triplicates. Experiment has been performed by Dzung Nguyen.

Experimental Procedures:

Peptide Synthesis:

The synthesis of the linear peptides was performed manually according to the standard Fmoc-solid-phase methodology using rink amide resin following the protocol described above in the manuscripts B-C in section 4. After final deprotection and cleavage, purification was performed via semi-preparative RP-HPLC (eluent: A: water + 0.1% TFA; B: MeCN + 0.1% TFA; gradient: 5-75% MeCN). In the following, the characterization data of the final peptides are displayed.

F-46: FAM-Tyr-Ile-His-His-Leu-Ile-Arg-Ala-Gly-Glu-CONH₂; the peptide coupling was performed manually in a 10 μmol scale. For the final cleavage step, the resin was treated with 2.0 mL cleavage cocktail (90% TFA/ 5.0% DCM/ 2.5% H₂O/ 2.5% TIS (3 h)). After purification, the product as a 3 x TFA salt (7.4 mg, 3.88 μmol, 38%) was obtained as a yellow solid. t_R = 21.65 min. Purity ≥ 99%. Formula: C₇₆H₉₆N₁₈O₁₉. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 783.3622; found: 783.3629.

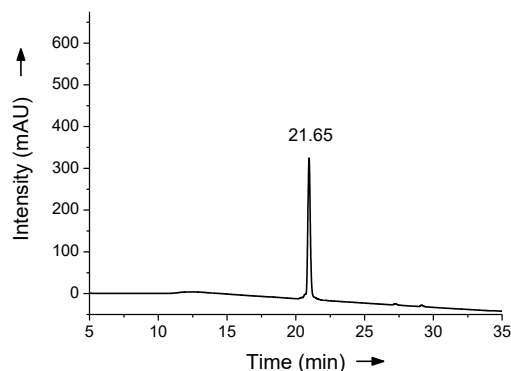


Figure F-55. HPLC chromatogram of purified peptide **F-46**. An EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-75% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.

F-47: FAM-Ser-Arg-Ala-Tyr-Ile-His-His-Leu-Ile-Arg-Ala-Gly-Glu-CONH₂; the peptide coupling was performed manually in a 10 μ mol scale. For the final cleavage step, the resin was treated with 2.0 mL cleavage cocktail (90% TFA/ 5.0% DCM/ 2.5% H₂O/ 2.5% TIS (3 h)). After purification, the product as a 4 x TFA salt (3.7 mg, 1.58 μ mol, 16%) was obtained as a yellow solid. t_R = 20.96 min. Purity \geq 99%. Formula: C₈₈H₁₁₈N₂₄O₂₃. HRMS-ESI⁺ (m/z): [M+3H]³⁺ calcd.: 627.3005; found: 627.3013.

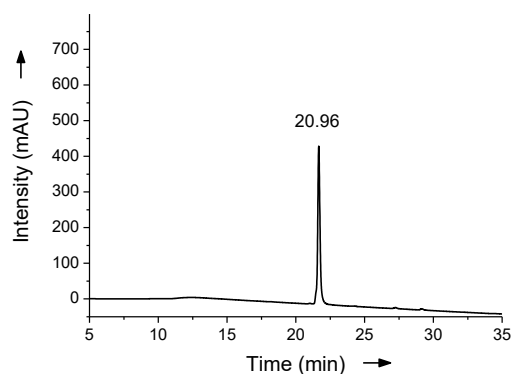


Figure F-56. HPLC chromatogram of purified peptide **F-47**. An EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-75% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.

F-48: FAM-Ala-Arg-Ala-Tyr-Ile-His-His-Leu-Ile-Arg-Ala-Gly-Glu-CONH₂; the peptide coupling was performed manually in a 10 μ mol scale. For the final cleavage step, the resin was treated with 2.0 mL cleavage cocktail (90% TFA/ 5.0% DCM/ 2.5% H₂O/ 2.5% TIS (3 h)). After purification, the product as a 4 x TFA salt (0.7 mg, 0.302 μ mol, 3%) was obtained as a yellow solid. t_R = 22.01 min. Purity = 98%. Formula: C₈₈H₁₁₈N₂₄O₂₂. HRMS-ESI⁺ (m/z): [M+3H]³⁺ calcd.: 622.9689; found: 622.9698.

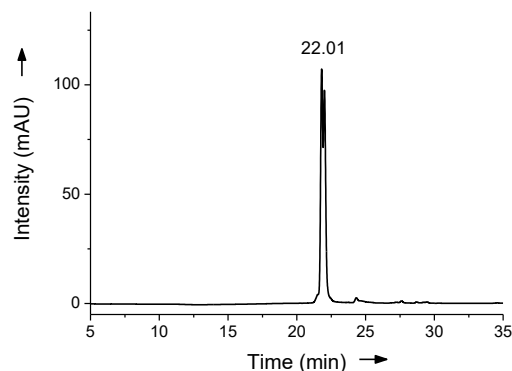


Figure F-57. HPLC chromatogram of purified peptide **F-48**. An EC 125/2 Nucleodur 100-C18 ec column (Macherey & Nagel) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-75% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.

F-49: Ac-Tyr-Ile-His-His-Leu-Ile-Arg-Ala-Gly-Glu-Ahx-Ahx-Lys(FAM)-CONH₂; the peptide coupling was performed manually in a 10 μmol scale. For the final cleavage step, the resin was treated with 2.0 mL cleavage cocktail (90% TFA/ 5.0% DCM/ 2.5% H₂O/ 2.5% TIS (3 h)). After purification, the product as a 3 x TFA salt (4.9 mg, 2.13 μmol, 21%) was obtained as a yellow solid. $t_R = 16.34$ min. Purity = 98%. Formula: C₉₆H₁₃₂N₂₂O₂₃. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 981.4991; found: 981.5010.

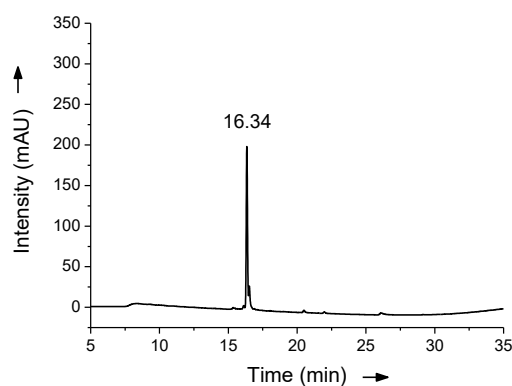


Figure F-58. HPLC chromatogram of purified peptide **F-49**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm.

F-50: FAM-Tyr-Cys-His-His-Leu-Cys-Arg-Ala-Gly-Glu-Ahx-Ahx-Lys(Ac)-CONH₂; the peptide coupling was performed manually in a 10 μmol scale. For the final cleavage step, the resin was treated with 2.0 mL cleavage cocktail (95% TFA/ 2.5% DODT/ 2.5% H₂O/ 1.0% TIS (3 h)). After purification, the product as a 3 x TFA salt (5.33 mg, 2.33 μmol, 23%) was obtained as a yellow solid. $t_R = 15.10$ min. Purity = 98%. Formula: C₉₀H₁₂₀N₂₂O₂₃S₂. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 971.4242; found: 971.4261.

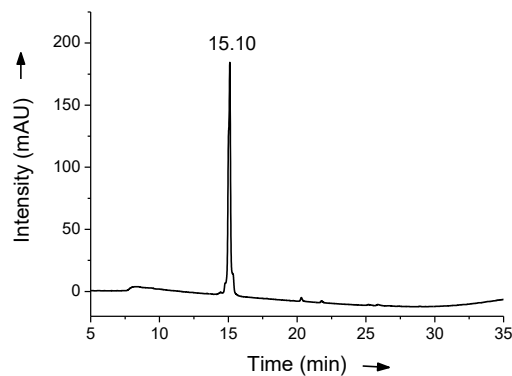
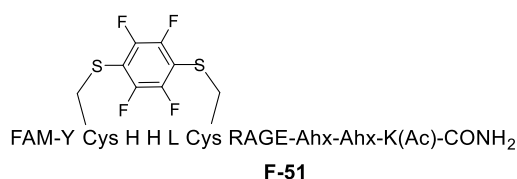


Figure F-59. HPLC chromatogram of purified peptide **F-50**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.



Following the procedure of Spokoyny et al.,¹⁰⁶ the peptide **F-50** (1.00 eq., 0.797 μ mol, 1.73 mg) was dissolved in 202 μ L of a 100 mM solution of hexafluorobenzene in DMF (25.0 eq., 19.9 μ mol). Then 159 μ L of a 50 mM solution tris(hydroxymethyl)aminomethane (Tris base) in DMF were added to the mixture and shaken for 5 days. After purification via semi-preparative RP-HPLC (detailed above), the product **F-51** as a 3 x TFA salt (0.4 mg, 0.165 μ mol, 21%) was obtained as a yellow solid. $t_R = 17.17$ min. Purity = 96%. Formula: C₉₆H₁₁₈N₂₂O₂₃S₂F₄. HRMS-ESI⁺ (m/z): [M+3H]³⁺ calcd.: 696.6112; found: 696.6120.

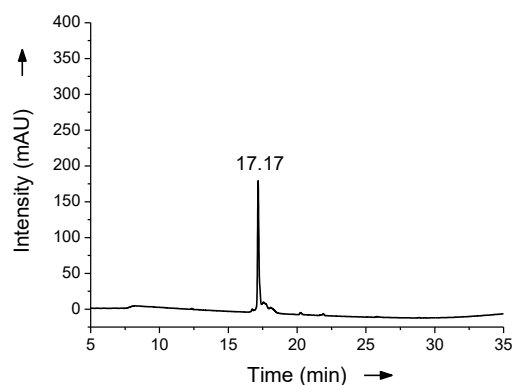
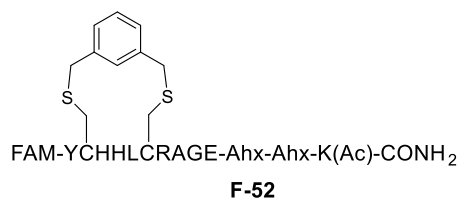


Figure F-60. HPLC chromatogram of purified peptide **F-51**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.



Following the procedure of Jo et al.,¹⁰⁷ the peptide **F-50** (1.00 eq., 0.889 μ mol, 1.93 mg) was dissolved in 740 μ L of 100 mM Tris-HCl pH = 8 and 1 μ L of TCEP (1 mM in water), was added. After 1 h, 1,3-bisbromoethylbenzene (20.0 eq., 17.78 μ mol, 4.69 mg) in 79 μ L DMSO were added and shaken for 4 h. The reaction was quenched by the addition of 16 μ L 0.6 M HCl. After purification via semi-preparative RP-HPLC (detailed above), the product **F-52** as a 3x TFA salt (0.8 mg, 0.335 μ mol, 38%) was obtained as a yellow solid. t_R = 16.36 min. Purity = 93%. Formula: C₉₈H₁₂₆N₂₂O₂₃S₂. HRMS-ESI⁺ (m/z): [M+2H]²⁺ calcd.: 1022.4477; found: 1022.4501.

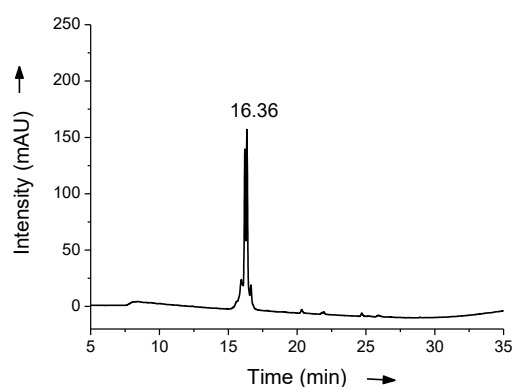


Figure F-61. HPLC chromatogram of purified peptide **F-52**. An Eclipse XBD-C18 column (Agilent Technologies) and eluents A: water + 0.05% TFA, B: MeCN + 0.03% TFA were used. Gradient: 5-95% B monitored at 220 nm. The double peak originates from the two isomers of the utilized 5,6-carboxyfluorescein.

7. References

1. Allen, L., Brand, A., Scott, J., Altman, M. & Hlava, M. Credit where credit is due. *Nature* **508**, 312–313 (2014).
2. Zhang, H. *et al.* MLL1 Inhibition Reprograms Epiblast Stem Cells to Naive Pluripotency. *Cell Stem Cell* **18**, 481–494 (2016).
3. Reis, S. a *et al.* Light-controlled modulation of gene expression by chemical optoepigenetic probes. *Nat. Chem. Biol.* **12**, 317–23 (2016).
4. Delgado, R. N. *et al.* Maintenance of neural stem cell positional identity by mixed-lineage leukemia 1. *Science* **368**, (2020).
5. Dougherty, P. G., Sahni, A. & Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chem. Rev.* **119**, 10241–10287 (2019).
6. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Acta Petrol. Sin.* **23**, 3–25 (1997).
7. Giordanetto, F. & Kihlberg, J. Macrocyclic drugs and clinical candidates: What can medicinal chemists learn from their properties? *J. Med. Chem.* **57**, 278–295 (2014).
8. Santos, G. B., Ganesan, A. & Emery, F. S. Oral Administration of Peptide-Based Drugs: Beyond Lipinski's Rule. *ChemMedChem* **11**, 2245–2251 (2016).
9. Edmondson, S. D., Yang, B. & Fallan, C. Proteolysis targeting chimeras (PROTACs) in 'beyond rule-of-five' chemical space: Recent progress and future challenges. *Bioorganic Med. Chem. Lett.* **29**, 1555–1564 (2019).
10. Woolley, G. A. *et al.* Photoswitching Azo Compounds in Vivo with Red Light. *J. Am. Chem. Soc.* **135**, 9777–9784 (2013).
11. Martínez-López, D., Babalhavaeji, A., Sampedro, D. & Andrew Woolley, G. Synthesis and characterization of bis(4-amino-2-bromo-6-methoxy)azobenzene derivatives. *Beilstein J. Org. Chem.* **15**, 3000–3008 (2019).
12. Konrad, D. B. *et al.* Computational Design and Synthesis of a Deeply Red-Shifted and Bistable Azobenzene. *J. Am. Chem. Soc.* **142**, 6538–6547 (2020).
13. Dong, M. *et al.* Near-Infrared Photoswitching of Azobenzenes under Physiological Conditions. *J. Am. Chem. Soc.* **139**, 13483–13486 (2017).
14. Samanta, S., McCormick, T. M., Schmidt, S. K., Seferos, D. S. & Woolley, G. A. Robust visible light photoswitching with ortho-thiol substituted azobenzenes. *Chem. Commun.* **49**, 10314–10316 (2013).
15. Mueller, J., Kretzschmar, I., Volkmer, R. & Boisguerin, P. Comparison of cellular uptake using 22 CPPs in 4 different cell lines. *Bioconjug. Chem.* **19**, 2363–2374 (2008).
16. Frank, J. Single-particle imaging of macromolecules by cryo-electron microscopy. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 303–319 (2002).
17. Callaway, E. Molecular-imaging pioneers scoop Nobel. *Nature* **550**, 167 (2017).
18. Frank, J. Advances in the field of single-particle cryo-electron microscopy over the last decade. *Nat. Protoc.* **12**, 209–212 (2017).
19. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **289**, 905–920 (2000).
20. Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339 (2000).

7. REFERENCES

21. Carter, A. P. *et al.* Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **407**, 340–348 (2000).
22. Schluenzen, F. *et al.* Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell* **102**, 615–623 (2000).
23. Fu, Z. *et al.* Key Intermediates in Ribosome Recycling Visualized by Time-Resolved Cryoelectron Microscopy. *Structure* **24**, 2092–2101 (2016).
24. Kastner, B. *et al.* GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55 (2008).
25. Xue, H. *et al.* Structural basis of nucleosome recognition and modification by MLL methyltransferases. *Nature* **573**, 445–449 (2019).
26. Park, S. H. *et al.* Cryo-EM structure of the human MLL1 core complex bound to the nucleosome. *Nat. Commun.* **10**, 3541–3554 (2019).
27. Sakamoto, K. M. *et al.* Protacs: Chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 8554–8559 (2001).
28. Barghout, S. H. Targeted Protein Degradation: An Emerging Therapeutic Strategy in Cancer. *Anticancer. Agents Med. Chem.* **20**, (2020).
29. Yang, J., Wang, Q., Feng, G. & Zeng, M. Significance of Selective Protein Degradation in the Development of Novel Targeted Drugs and Its Implications in Cancer Therapy. *Adv. Ther.* **3**, 1900210 (2020).
30. Ciechanover, A. The ubiquitin-proteasome proteolytic pathway. *Cell* **79**, 13–21 (1994).
31. Lohbeck, J. & Miller, A. K. Practical synthesis of a phthalimide-based Cereblon ligand to enable PROTAC development. *Bioorganic Med. Chem. Lett.* **26**, 5260–5262 (2016).
32. Sakamoto, K. M. *et al.* Development of Protacs to target cancer-promoting proteins for ubiquitination and degradation. *Mol. Cell. Proteomics* **2**, 1350–1358 (2003).
33. Bargagna-Mohan, P., Baek, S. H., Lee, H., Kim, K. & Mohan, R. Use of PROTACS as molecular probes of angiogenesis. *Bioorganic Med. Chem. Lett.* **15**, 2724–2727 (2005).
34. Okuhira, K. *et al.* Development of hybrid small molecules that induce degradation of estrogen receptor-alpha and necrotic cell death in breast cancer cells. *Cancer Science* **104**, 1492–1498 (2013).
35. Bondeson, D. P. *et al.* Catalytic in vivo protein knockdown by small-molecule PROTACs. *Nat. Chem. Biol.* **11**, 611–617 (2015).
36. Dai, Y. *et al.* Development of cell-permeable peptide-based PROTACs targeting estrogen receptor alpha. *Eur. J. Med. Chem.* **187**, 111967 (2020).
37. Buckley, D. L. *et al.* HaloPROTACS: Use of Small Molecule PROTACs to Induce Degradation of HaloTag Fusion Proteins. *ACS Chem. Biol.* **10**, 1831–1837 (2015).
38. Zengerle, M., Chan, K. H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **10**, 1770–1777 (2015).
39. Winter, G. E. *et al.* Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science (80-.)*. **348**, 1376–1381 (2015).
40. Raina, K. *et al.* PROTAC-induced BET protein degradation as a therapy for castration-resistant prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7124–7129 (2016).
41. Itoh, Y. *et al.* Development of target protein-selective degradation inducer for protein knockdown. *Bioorganic Med. Chem.* **19**, 3229–3241 (2011).

42. Au, Y. Z., Wang, T., Sigua, L. H. & Qi, J. Peptide-Based PROTAC: The Predator of Pathological Proteins. *Cell Chem. Biol.* **27**, 637–639 (2020).
43. Ma, D. *et al.* A cell-permeable peptide-based PROTAC against the oncoprotein CREPT proficiently inhibits pancreatic cancer. *Theranostics* **10**, 3708–3721 (2020).
44. Galdeano, C. *et al.* Structure-guided design and optimization of small molecules targeting the protein-protein interaction between the von hippel-lindau (VHL) E3 ubiquitin ligase and the hypoxia inducible factor (HIF) alpha subunit with in vitro nanomolar affinities. *J. Med. Chem.* **57**, 8657–8663 (2014).
45. Buckley, D. L. *et al.* Small-molecule inhibitors of the interaction between the E3 ligase VHL and HIF1 α . *Angew. Chemie - Int. Ed.* **51**, 11463–11467 (2012).
46. Jin, J. *et al.* The peptide PROTAC modality: A novel strategy for targeted protein ubiquitination. *Theranostics* **10**, 10141–10153 (2020).
47. Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat. Struct. Mol. Biol.* **13**, 713–719 (2006).
48. Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859–872 (2005).
49. Cao, F. *et al.* Targeting MLL1 H3K4 Methyltransferase Activity in Mixed-Lineage Leukemia. *Mol. Cell* **53**, 247–261 (2014).
50. Pashkova, N. *et al.* WD40 Repeat Propellers Define a Ubiquitin-Binding Domain that Regulates Turnover of F Box Proteins. *Mol. Cell* **40**, 433–443 (2010).
51. Nakagawa, T. & Xiong, Y. X-Linked Mental Retardation Gene CUL4B Targets Ubiquitylation of H3K4 Methyltransferase Component WDR5 and Regulates Neuronal Gene Expression. *Mol. Cell* **43**, 381–391 (2011).
52. Cao, F. *et al.* Targeting MLL1 H3K4 Methyltransferase Activity in Mixed-Lineage Leukemia. *Mol. Cell* **53**, 247–261 (2014).
53. Linden, G. *et al.* Conditional Singlet Oxygen Generation through a Bioorthogonal DNA-targeted Tetrazine Reaction. *Angew. Chemie - Int. Ed.* **58**, 12868–12873 (2019).
54. Schapira, M., Calabrese, M. F., Bullock, A. N. & Crews, C. M. Targeted protein degradation: expanding the toolbox. *Nat. Rev. Drug Discov.* **18**, 949–963 (2019).
55. Bryan, A. F. *et al.* WDR5 is a conserved regulator of protein synthesis gene expression. *Nucleic Acids Res.* **48**, 2924–2941 (2020).
56. Albert, L. *et al.* Modulating Protein–Protein Interactions with Visible-Light-Responsive Peptide Backbone Photoswitches. *ChemBioChem* **20**, 1417–1429 (2019).
57. Zhang, P., Lee, H., Brunzelle, J. S. & Couture, J. F. The plasticity of WDR5 peptide-binding cleft enables the binding of the SET1 family of histone methyltransferases. *Nucleic Acids Res.* **40**, 4237–4246 (2012).
58. Cao, F. *et al.* Targeting MLL1 H3K4 Methyltransferase Activity in Mixed-Lineage Leukemia. *Mol. Cell* **53**, 247–261 (2014).
59. Li, Y. *et al.* Structural basis for activity regulation of MLL family methyltransferases. *Nature* **530**, 447–452 (2016).
60. Ng, S. B. *et al.* Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790–793 (2010).
61. Niikawa, N., Matsuura, N., Fukushima, Y., Ohsawa, T. & Kajii, T. Kabuki make-up syndrome: A syndrome of mental retardation, unusual facies, large and protruding ears, and postnatal

- growth deficiency. *J. Pediatr.* **99**, 565–569 (1981).
62. Schwenty-Lara, J., Nürnberger, A. & Borchers, A. Loss of function of Kmt2d, a gene mutated in Kabuki syndrome, affects heart development in *Xenopus laevis*. *Dev. Dyn.* **248**, 465–476 (2019).
 63. Thieriet, N., Alsina, J., Giralt, E., Guib??, F. & Albericio, F. Use of Alloc-amino acids in solid-phase peptide synthesis. Tandem deprotection-coupling reactions using neutral conditions. *Tetrahedron Lett.* **38**, 7275–7278 (1997).
 64. Grieco, P., Gitu, P. M. & Hruby, V. J. Preparation of ‘side-chain-to-side-chain’ cyclic peptides by Allyl and Alloc strategy: Potential for library synthesis. *J. Pept. Res.* **57**, 250–256 (2001).
 65. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128**, 693–705 (2007).
 66. Miccio, A. *et al.* NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. *EMBO J.* **29**, 442–456 (2010).
 67. Denslow, S. A. & Wade, P. A. The human Mi-2/NuRD complex and gene regulation. *Oncogene* **26**, 5433–5438 (2007).
 68. Ahringer, J. NuRD and SIN3: Histone deacetylase complexes in development. *Trends Genet.* **16**, 351–356 (2000).
 69. McDonel, P., Costello, I. & Hendrich, B. Keeping things quiet: Roles of NuRD and Sin3 co-repressor complexes during mammalian development. *Int. J. Biochem. Cell Biol.* **41**, 108–116 (2009).
 70. Bouazoune, K. *et al.* The dMi-2 chromodomains are DNA binding modules important for ATP-dependent nucleosome mobilization Bouazoune_et_al-2002-The_EMBO_Journal.pdf. *EMBO J.* **21**, 2430–2440 (2002).
 71. Le Guezennec, X. *et al.* MBD2/NuRD and MBD3/NuRD, Two Distinct Complexes with Different. *Mol. Cell. Biol.* **26**, 843–851 (2006).
 72. Baubec, T., Ivánek, R., Lienert, F. & Schübeler, D. Methylation-dependent and -independent genomic targeting principles of the mbd protein family. *Cell* **153**, 480–492 (2013).
 73. Günther, K. *et al.* Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences. *Nucleic Acids Res.* **41**, 3010–3021 (2013).
 74. Zhang, Y. *et al.* Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* **13**, 1924–1935 (1999).
 75. Murzina, N. V. *et al.* Structural Basis for the Recognition of Histone H4 by the Histone-Chaperone RbAp46. *Structure* **16**, 1077–1085 (2008).
 76. Lejon, S. *et al.* Insights into association of the NuRD complex with FOG-1 from the crystal structure of an RbAp48-FOG-1 complex. *J. Biol. Chem.* **286**, 1196–1203 (2011).
 77. Kumar, R., Wang, R. A. & Bagheri-Yarmand, R. Emerging Roles of MTA Family Members in Human Cancers. *Semin. Oncol.* **30**, 30–37 (2003).
 78. Li, W. *et al.* Expression profile of MTA1 in adult mouse tissues. *Tissue Cell* **41**, 390–399 (2009).
 79. Li, D. Q., Pakala, S. B., Nair, S. S., Eswaran, J. & Kumar, R. Metastasis-associated protein 1/nucleosome remodeling and histone deacetylase complex in cancer. *Cancer Res.* **72**, 387–394 (2012).
 80. Feng, Q. *et al.* Identification and Functional Characterization of the p66/p68 Components of the MeCP1 Complex. *Mol. Cell. Biol.* **22**, 536–546 (2002).
 81. Brackertz, M., Boeke, J., Zhang, R. & Renkawitz, R. Two highly related p66 proteins comprise a new family of potent transcriptional repressors interacting with MBD2 and MBD3. *J. Biol. Chem.* **277**, 40958–40966 (2002).

82. Lai, A. Y. & Wade, P. A. Cancer biology and NuRD: A multifaceted chromatin remodelling complex. *Nat. Rev. Cancer* **11**, 588–596 (2011).
83. Pegoraro, G. *et al.* Ageing-related chromatin defects through loss of the NURD complex. *Nat. Cell Biol.* **11**, 1261–1267 (2009).
84. Abbey, M., Trush, V., Gibson, E. & Vedadi, M. Targeting Human Retinoblastoma Binding Protein 4 (RBBP4) and 7 (RBBP7). *bioRxiv* **4**, 303537 (2018).
85. Alqarni, S. S. M. *et al.* Insight into the architecture of the NuRD complex: Structure of the RbAp48-MTA1 subcomplex. *J. Biol. Chem.* **289**, 21844–21855 (2014).
86. Hong, W. *et al.* FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J.* **24**, 2367–2378 (2005).
87. Reed Moody, R. *et al.* Probing the interaction between the histone methyltransferase/deacetylase subunit RBBP4/7 and the transcription factor BCL11A in epigenetic complexes. *J. Biol. Chem.* **293**, 2125–2136 (2018).
88. Lin, S. *et al.* A high-throughput fluorescent polarization assay for nuclear receptor binding utilizing crude receptor extract. *Anal. Biochem.* **300**, 15–21 (2002).
89. Sun, S., Nguyen, L. H. T., Harold Ross, O., Hollis, G. F. & Wynn, R. Quantitative analysis of c-myc-tagged protein in crude cell extracts using fluorescence polarization. *Anal. Biochem.* **307**, 287–296 (2002).
90. Mazitschek, R., Patel, V., Wirth, D. F. & Clardy, J. Development of a fluorescence polarization based assay for histone deacetylase ligand discovery. *Bioorganic Med. Chem. Lett.* **18**, 2809–2812 (2008).
91. Gao, C. *et al.* Targeting transcription factor SALL4 in acute myeloid leukemia by interrupting its interaction with an epigenetic complex. *Blood* **121**, 1413–1421 (2013).
92. Hall, M. D. *et al.* and Drug Discovery : a Review. *Methods Appl. Fluoresc.* **4**, 1–41 (2017).
93. Waldmann, H. *et al.* Structure based design of bicyclic peptide inhibitors of RbAp48. *Angew. Chemie* (2020). doi:10.1002/ange.202009749
94. Donovan, K. A. *et al.* Thalidomide promotes degradation of SALL4, a transcription factor implicated in Duane radial ray syndrome. *Elife* **7**, 1–25 (2018).
95. Matyskiela, M. E. *et al.* SALL4 mediates teratogenicity as a thalidomide-dependent cereblon substrate. *Nat. Chem. Biol.* **14**, 981–987 (2018).
96. Kong, N. R. *et al.* Zinc finger protein SALL4 functions through an AT-rich motif to regulate gene expression Nikki. *bioRxiv* (2020).
97. Roatsch, M., Vogelmann, A., Herp, D., Jung, M. & Olsen, C. A. Proteolysis-Targeting Chimeras (PROTACs) Based on Macrocyclic Tetrapeptides Selectively Degrade Class I Histone Deacetylases 1–3. *ChemRxiv* (2020). doi:10.26434/chemrxiv.12416303
98. Banerjee, U., Girard, J. R., Goins, L. M. & Spratford, C. M. Drosophila as a genetic model for hematopoiesis. *Genetics* **211**, 367–417 (2019).
99. Iwata-Reuyl, D. Biosynthesis of the 7-deazaguanosine hypermodified nucleosides of transfer RNA. *Bioorg. Chem.* **31**, 24–43 (2003).
100. Sansonetti, P. J. Microbes and microbial toxins: Paradigms for microbial-mucosal interactions III. Shigellosis: From symptoms to molecular pathogenesis. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **280**, 319–323 (2001).
101. Stengl, B. *et al.* Crystal Structures of tRNA-guanine Transglycosylase (TGT) in Complex with Novel and Potent Inhibitors Unravel Pronounced Induced-fit Adaptations and Suggest Dimer

- Formation Upon Substrate Binding. *J. Mol. Biol.* **370**, 492–511 (2007).
102. Ritschel, T. *et al.* An Integrative Approach Combining Noncovalent Mass Spectrometry, Enzyme Kinetics and X-ray Crystallography to Decipher Tgt Protein-Protein and Protein-RNA Interaction. *J. Mol. Biol.* **393**, 833–847 (2009).
103. Immekus, F. *et al.* Launching spiking ligands into a protein-protein interface: A promising strategy to destabilize and break interface formation in a tRNA modifying enzyme. *ACS Chem. Biol.* **8**, 1163–1178 (2013).
104. Jakobi, S. *et al.* Hot-spot analysis to dissect the functional protein-protein interface of a tRNA-modifying enzyme. *Proteins Struct. Funct. Bioinforma.* **82**, 2713–2732 (2014).
105. Jakobi, S. *et al.* What Glues a Homodimer Together: Systematic Analysis of the Stabilizing Effect of an Aromatic Hot Spot in the Protein-Protein Interface of the tRNA-Modifying Enzyme Tgt. *ACS Chem. Biol.* **10**, 1897–1907 (2015).
106. Spokoyny, A. M. *et al.* A perfluoroaryl-cysteine SNAr chemistry approach to unprotected peptide stapling. *J. Am. Chem. Soc.* **135**, 5946–5949 (2013).
107. Jo, H. *et al.* Development of α -helical calpain probes by mimicking a natural protein-protein interaction. *J. Am. Chem. Soc.* **134**, 17704–17713 (2012).

8. Supplementary

8.1 List of Publications

- 5) J. Lenz, R. Liefke, J. Funk, S. Shoup, A. Nist, T. Stiewe, R. Schulz, Y. Tokusumi, L. Albert, H. Raifer, K. Förstemann, O. Vázquez, T. Tokusumi, N. Fossett, A. Brehm*, *Ush regulates hemocyte-specific gene expression, fatty acid metabolism and cell cycle progression and cooperates with dNuRD to orchestrate hematopoiesis*, bioRxiv 2020.06.10.143701, doi: <https://doi.org/10.1101/2020.06.10.143701>
- 4) L. Albert, W. Steinchen, L. Werel, J. Nagpal, N. Djokovic, D. Ruzic, M. Hoffarth, J. Xu, A. Royant, G. Bange, K. Nikolic, S. Ryu, Y. Dou, L.-O. Essen, O. Vázquez*, *In Vivo* Optochemical Control of Hematopoiesis by MLL1 Complex Disruption, *Nature Chemical Biology* **2020**, manuscript submitted.
- 3) L. Albert, O. Vázquez*, Photoswitchable peptides for spatiotemporal control of biological functions, *Chem. Commun.* **2019**, 55, 10192-10213.
- 2) L. Albert, A. Peñalver, N. Djokovic, L. Werel, M. Hoffarth, D. Ruzic, J. Xu, L.-O. Essen, K. Nikolic, Y. Dou, O. Vázquez*, Modulating Protein-Protein Interactions with Visible-Light Responsive Peptide Backbone Photoswitches, *ChemBioChem* **2019**, 20, 1417-1429.
- 1) L. Albert, J. Xu, R. Wan, V. Srinivasan, Y. Dou, O. Vázquez*, Controlled inhibition of methyltransferases using photoswitchable peptidomimetics: towards an epigenetic regulation of leukemia, *Chem. Sci.* **2017**, 8, 4612-4618.

8.2 Permissions for the Reprint of Publications

Chem. Sci. **2017**, *8*, 4612-4618.

Referring to the Royal Society of Chemistry, authors accepting the standard license to publish, have the sharing rights to use the accepted manuscript in submissions of grant applications, or academic requirements such as thesis or dissertations. (<https://www.rsc.org/journals-books-databases/open-access/open-access-info/#choose>, last accessed 05.012.2019)

ChemBioChem **2019**, *20*, 1417-1429.

This agreement between Philipps-University Marburg – Lea Albert ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4775281044488
License date	Feb 24, 2020
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	ChemBioChem
Licensed Content Title	Modulating Protein–Protein Interactions with Visible-Light-Responsive Peptide Backbone Photoswitches
Licensed Content Author	Olalla Vázquez, Yali Dou, Katarina Nikolic, et al
Licensed Content Date	Apr 25, 2019
Licensed Content Volume	20
Licensed Content Issue	11
Licensed Content Pages	13
Type of use	Dissertation/Thesis

Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Optoepigenetics: Photoswitchable Peptides as Tools for Light-controlled Modulation of the Histone Metyltransferase MLL1
Expected completion date	Apr 2020
Expected size (number of pages)	350
Requestor Location	Philipps-Universität Marburg Hans-Meerwein-Str 4 Marburg, Hessen 35043 Germany Attn: Philipps-Universität Marburg
Publisher Tax ID	EU826007151
Total	0.00 EUR

Chem. Commun. **2019**, *55*, 10192-10213.

Referring to the Royal Society of Chemistry, authors accepting the standard license to publish, have the sharing rights to use the accepted manuscript in submissions of grant applications, or academic requirements such as thesis or dissertations. (<https://www.rsc.org/journals-books-databases/journal-authors-reviewers/licences-copyright-permissions/#deposition-sharing>, last accessed 05.012.2019)

Nature Chemical Biology **2020**, manuscript submitted.

Referring to the Nature Research, authors keep the copyrights and are allowed to use the accepted manuscript in submissions of grant applications, or academic requirements such as thesis or dissertations. (<https://www.nature.com/nature-research/reprints-and-permissions/permissions-requests>, last accessed 05.12.2019).