ZMB
Zentrum für Medizinische Biotechnologie
Center of Medical Biotechnology

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

Rational targeting of Taspase1 nuclear import and biological function by supramolecular ligands

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften

Dr. rer. nat.

der Fakultät für Biologie

an der Universität Duisburg-Essen

CRC 1093

vorgelegt von

Alexander Höing

aus Borken

Februar 2022

Die der vorliegenden Arbeit zu Grunde liegenden Experimente wurden am Zentrum für medizinische Biotechnologie in der Abteilung für Molekularbiologie II der Universität Duisburg-Essen durchgeführt.

1. Gutachter: Prof. Dr. Shirley Knauer

2. Gutachter: Prof. Dr. Markus Kaiser

3. Gutachter: Prof. Dr. Luc Brunsveldt

Vorsitzender des Prüfungsausschusses: Prof. Dr. Stefan Westermann

Tag der mündlichen Prüfung: 01.06.2022

DOI: 10.17185/duepublico/76049
URN: urn:nbn:de:hbz:465-20230724-091926-3

Publications used in this cumulative dissertation

P. Pasch[‡], **A. Höing[‡]**, S. Ueclue, M. Killa, J. Voskuhl, S. K. Knauer and L. Hartmann, PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction, *Chem. Commun.*, 2021, 57, 3091–3094.

A. Höing[‡], A. Zimmermann[‡], L. Moews, M. Killa, M. Heimann, A. Hensel, J. Voskuhl and S. K. Knauer, A Bivalent Supramolecular GCP Ligand Enables Blocking of the Taspase1/Importin α Interaction, *ChemMedChem*, 2022, 17, e202100640.

A. Höing, R. Struth, C. Beuck, N. Rafieiolhosseini, D. Hoffmann, R. H. Stauber, P. Bayer, J. Niemeyer, S. K. Knauer. Dual activity inhibition of Threonine aspartase 1 by a single supramolecular bisphosphate ligand *(under review)*.

A. Höing[‡], A. Kirupakaran[‡], C. Beuck, J. Niemeyer, P. Bayer, T. Schrader, S. K. Knauer. Recognition of a flexible protein loop in Taspase1 by multivalent supramolecular tweezers *(in preparation)*.

[‡] authors contributed equally

Further publications

J. Dubbert[‡], **A. Höing[‡]**, N. Riek, S. K. Knauer and J. Voskuhl, Supramolecular subphthalocyanine complexes-cellular uptake and phototoxicity, *Chem. Commun*., 2020, 56, 7653–7656.

J. van den Boom, A. Hensel, F. Trusch, A. Matena, S. Siemer, D. Guel, D. Docter, **A. Höing**, P. Bayer, R. H. Stauber and S. K. Knauer, The other side of the corona: nanoparticles inhibit the protease taspase1 in a size-dependent manner, *Nanoscale,* 2020, 12, 19093–19103.

A. Sowa, **A. Höing**, U. Dobrindt, S. K. Knauer, A. Galstyan and J. Voskuhl, Umbelliferone Decorated Water-soluble Zinc(II) Phthalocyanines - In Vitro Phototoxic Antimicrobial Anti-cancer Agents, *Chemistry (Weinheim an Der Bergstrasse, Germany)*, 2021, 27, 14672–14680.

T. Rick, V. Kreiling, **A. Höing**, S. Fiedler, T. Glatter, W. Steinchen, G. Hochberg, H. Bähre, R. Seifert, G. Bange, S. K. Knauer, P. L. Graumann and K. M. Thormann, A GGDEF domain serves as a spatial on-switch for a phosphodiesterase by direct interaction with a polar landmark protein (*revision submitted,* pre-print available doi: 10.1101/2021.08.12.456111)

N. Rafieiolhosseini, M. Killa, N. Tötsch, J.-N. Grad, **A. Höing**, C. Ottmann, S. K. Knauer, J. Voskuhl and D. Hoffmann, Computational model predicts protein binding sites of a luminescent ligand equipped with guanidiniocarbonyl-pyrrole groups (*in preparation*, pre-print available doi: 10.3762/bxiv.2021.11.v1).

A. Hensel, P. Stahl, L. Moews, L. König, R. Patwardhan, **A. Höing**, N. Schulze, P. Nalbant, R. H. Stauber, S. K. Knauer. Taspase1 cleavage regulates Myosin1f-induced filopodia *(submitted)*.

[‡] authors contributed equally

Table of contents

Table of contents I

List of abbreviations III

List of figures V

Zusammenfassung VI

Summary IX

1 Introduction 16

1.1 Cancer - terminology and statistics 16

1.2 Cancerogenesis 17

1.3 Proteases in cancer development 18

1.4 Taspase1 20

1.4.1 General function and oncological relevance 20

1.4.2 Structure and activity of Taspase1 21

1.4.3 The intramolecular Taspase1 activation process 24

1.4.4 Taspase1 and Importinα – translocation as pre-requisite for activation 25

1.4.5 Therapeutic approaches to interfere with Taspase1 27

1.5 Aim of this study 29

2 Publications 32

2.1 Publication I 33

2.1.1 Author contributions 33

2.1.2 Article introduction and summary 34

2.1.3 PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction 35

2.2 Publication II 40

2.2.1 Author contributions 40

2.2.2 Article introduction and summary 41

2.2.3 A Bivalent Supramolecular GCP Ligand Enables Blocking of the Taspase1/Importin α Interaction 42

2.3 Publication III 48

2.3.1 Author contributions 48

2.3.2 Article introduction and summary 49
2.3.3 Dual activity inhibition of Threonine aspartase 1 by a single supramolecular bisphosphate ligand 50
2.4 Publication VI 58
2.4.1 Author contributions 58
2.4.2 Article introduction and summary 59
2.4.3 Recognition of a flexible protein loop in Taspase1 by multivalent supramolecular tweezers 60
3 Discussion 69
3.1 Cationic binders targeting anionic hotspots 69
3.2 Anionic binders targeting cationic clusters in the NLS 72
3.3 Further approaches of interference with Taspase1 function 75
3.4 Limitations of this study 76
4 Conclusion and outlook 79
5 References 80
6 Appendix 90
6.1 Publication I: Supplementary information 90
6.2 Publication II: Supplementary information 120
6.3 Publication III: Supplementary information 157
6.4 Publication IV: Supplementary information 177
7 Curriculum vitae 193
8 Eidesstattliche Erklärung 195

List of abbreviations

µM	Micro molar
Alloc	Allyloxycarbonyl
Cbz	Benzyl chloroformate
CD	Circular dichroism
CDK2A	Cycline dependent kinase inhibitor 2A
C-terminus	Carboxy-terminus
D_2O	Deuterium oxide
FRET	Fluorescence resonance energy transfer
GCP	Guanidiniocarbonyl-pyrrole
HDX-MS	Hydrogen/deuterium exchange mass spectrometry
HMT	Histone methyltransferase
HNSCC	Head and neck squamous cell carcinoma
IC_{50}	Half maximal inhibitory concentration
Impα	Importinα
Impβ	Importinβ
K_D	Dissociation constant
kDa	Kilo dalton
K_I	Inhibitory constant
MLL1	Mixed lineage leukemia 1
MRSA	Methicillin-resistant *Staphylococcus aureus*
NES	Nuclear export signal
NLS	Nuclear localization signal
nm	Nano meter
NPC	Nuclear pore complex
NPM1	Nucleophosmin1
N-terminus	Amino-terminus
PEG	Polyethylene glycol
PLA	Proximity ligation assay
SAMC	Simulated annealing monte carlo simulation
SAXS	Small angle x-ray scattering
SPR	Surface plasmon resonance
Stp1	Serine/threonine phosphatase
Tasp/Impα-interaction	Taspase1/Importinα-interaction

Amino acids:

C	Cysteine
D	Aspartic acid
E	Glutamic acid
F	Phenylalanine
G	Glycine
I	Isoleucine
L	Leucine
M	Methionine
Q	Glutamine
R	Arginine
T	Tyrosine

V	Valine
W	Tryptophane
X	Any residue

List of figures

Figure 1: Distribution of new incidences (A) and mortalities (B) of cancer in 2020 by cancer type (worldwide, both genders, all ages) (Sung *et al.* 2021). 12

Figure 2: Development of the hallmarks of cancer. The original hallmarks of cancer proposed in 2000 (A) comprised six traits. 13

Figure 3: Schematic illustrating the distribution of human proteases among the different classes. 14

Figure 4: Exemplary functions of Taspase1 in cancer development.....16

F**igure 5**: Model of Taspase1 (A) and the Taspase1 loop (B) according to van den Boom *et al.* *19*

Figure 6: Schematic comparison of wild type Taspase1 and the circularly permutated Taspase1. 20

Figure 7: The intramolecular activation of process of Taspase1 based on the “dock and click” model by Sabiani *et al.* 21

Figure 8: Schematic illustrating the activation process and transport of Taspase1.....22

Figure 9: Model of the Taspase1 loop (A) and the corresponding amino acid sequence (B). 26

Figure 10: Schematic representation of the structure and building blocks of trivalent ligands used in this study.....55

Figure 11: Schematic representation of the structure and building blocks of the compounds used in this study.....56

Figure 12: Schematic explaining the different use of the terms “dual inhibition” or “dual action” for inhibitors in literature. 58

Zusammenfassung

In der Krebstherapie sind Chemotherapeutika noch immer die erste Wahl bei den meisten Tumorentitäten, obwohl ihre generelle Zytotoxizität im Organismus zu teils schwerwiegenden Nebenwirkungen führt. Neue Therapieansätze haben deshalb zum Ziel, Tumorzellen spezifischer anzugreifen. Allerdings ist Krebs entstehungsbedingt eine äußerst vielfältige Erkrankung, was die Entwicklung selektiver Inhibitoren erschwert (Hanahan and Weinberg 2011). Ein Protein, welches in zahlreichen Tumoren und Leukämien verstärkt exprimiert wird, ist die Taspase1 (**T**hreonin **Asp**art**ase 1**) (Hsieh *et al.* 2003; D Wünsch *et al.* 2016). Diese Protease erkennt eine konserviertes Spaltungsmotiv $Q^{3}(F/I/L/V)^{2}D^{1}/G^{1'}X^{2'}D^{3'}D^{4'}$, und zu ihren bestätigten Substraten zählen bekannte Onkogene wie MLL1 oder Transkriptionsfaktoren wie TFIIA (Bier *et al.* 2011b; Gribko *et al.* 2017; Hsieh *et al.* 2003). Während ein *Knock-out* von Taspase1 im adulten Gewebe im Allgemeinen gut vertragen wird, führt er in Tumorzellen zu einer verlangsamten Proliferation und zur Einleitung der Apoptose, was Taspase1 zu einem attraktiven Ziel für die Krebstherapie macht (Takeda *et al.* 2006; Chen *et al.* 2010). Allerdings zeigen allgemeine Protease-Inhibitoren gegenüber Taspase1 keine Wirksamkeit, und bislang sind nur wenige spezifische Hemmstoffe verfügbar, mit welchen die proteolytische Aktivität effizient beeinflusst werden kann (Chen *et al.* 2012; Johannes van den Boom *et al.* 2014; van den Boom *et al.* 2020; Hsieh *et al.* 2003). Der intrazelluläre Aktivierungsprozess der Taspase1 beinhaltet deren aktiven Import in den Zellkern, was die hierfür maßgebliche Interaktion mit Importinα (Impα) als Angriffspunkt für neuartige Inhibitionsstrategien identifiziert (Bier *et al.* 2011a; van den Boom *et al.* 2016). Das Ziel dieser Arbeit war daher die rationale Entwicklung supramolekularer Liganden, welche diese Wechselwirkung effizient unterdrücken und so möglicherweise auch die proteolytische Aktivität von Taspase1 hemmen (Bier *et al.* 2011a).

Im Fokus standen zunächst kationische Liganden, welche an Gruppen von Aminosäuren mit sauren Seitenketten binden sollten, welche das zweiteilige, basische Kernlokalisationssignal (engl. *Nuclear localization signal*, NLS) flankieren. In der Tat haben wir in einer initialen Studie den ersten Inhibitor entwickelt, welcher die Tasp/Imp-Wechselwirkung effektiv stört (Pasch *et al.* 2021). In den hierzu entwickelten trivalenten supramolekularen Liganden wurden als kationische Binder drei Guanidiniumcarbonylpyrrol(GCP)-Einheiten, wahlweise mit oder ohne Lysin substituierten Seitenketten, durch ein gemeinsames Oligomer-Rückgrad verbunden (Liganden **3G** und **3GL**). Diese wurden in einem weiteren Schritt zur erweiterten räumlichen Abschirmung der Bindestelle zusätzlich mit Polyethylenglykol (PEG) funktionalisiert (Liganden **3GP** und **3GLP**). Während für alle GCP-Oligomere mittels Oberflächenplasmonresonanz eine direkte Bindung an Taspase1 nachgewiesen werden konnte, waren nur PEGylierte Liganden in der Lage, die Wechselwirkung mit Impα in *Pulldown*-Experimenten zu unterbinden.

In einer zweiten Studie wurde wiederum die Anionen-bindende GCP-Einheit zur Entwicklung bivalenter supramolekularer Liganden herangezogen, welche zusätzlich noch mit unterschiedlichen Schutzgruppen ausgestattet wurden (Höing *et al.* 2021). So konnte der die Schutzgruppe C(bz) enthaltende, bivalente Inhibitor **2GC** identifiziert werden, welcher die nicht nur die Tasp/Impα-Wechselwirkung in *Pulldown*-Experimenten effizient inhibierte (IC_{50} = 35 µM), sondern darüber hinaus auch eine toxische Wirkung auf Taspase1-exprimierende Tumorzellen zeigte (EC_{50} = 40-70 µM). Neben der GCP-Einheit stellten sich hier also die Bivalenz sowie das Vorhandensein der C(bz)-Schutzgruppe als entscheidende Faktoren für die inhibitorische Wirkung des Liganden heraus.

Konzeptionell basierten die bislang entwickelten Inhibitoren auf der Bindung der Liganden an saure Aminosäuren wie Glutamin- oder Asparaginsäure, welche auf der Oberfläche der Taspase1 in räumlicher Nähe zum zweigeteilten, basischen NLS auf einer flexiblen Loop-Struktur verortet sind. In weiteren Untersuchungen sollte daher letzteres als relevante Zielstruktur mittels anionischer Binder direkt angesteuert werden. In einer auf diesem Konzept basierenden dritten Studie konnten wir so den ersten Inhibitor für Taspase1 beschreiben, welcher nicht nur verbesserte Bindungseigenschaften, sondern auch einen dualen Inhibitionsmechanismus aufweist. Der ebenfalls bivalente supramolekulare Ligand **11d** verwendet para-substituierte Phosphatgruppen als Bindungsmotiv (Octa-Smolin *et al.* 2018). Über virtuelle Bindungsstudien wurde die energetisch günstigste Bindestelle am NLS im Taspase1-Loop bestimmt. Über Fluoreszenzanisotropie konnte eine Bindungssaffinität für Taspase1 im nanomolaren Bereich ermittelt werden (K_D = 300 nM). Die direkte Wechselwirkung mit dem Taspase1-Loop konnte mittels Fluoreszenztitration (K_D = 3 µM) und Kernspinresonanz bestätigt werden. In *Pulldown*-Experimenten wurde eine Hemmung der Tasp/Impα-Wechselwirkung im mikromolaren Bereich (IC_{50} = 6 µM) nachgewiesen. Kolorimetrische Spaltungstests zeigten zudem, dass **11d** in vergleichbaren Konzentrationen die proteolytische Aktivität von Taspase1 unterdrückt (IC_{50} = 2 µM), was darüber hinaus in Zellexperimenten mit einem transfizierbaren Biosensor verifiziert werden konnte.

Die bislang durchgeführten Studien zeigten eindrücklich, dass anscheinend nur die mehrfache Präsentation supramolekularer Bindungsmotive eine effiziente Inhibition der Taspase1 ermöglicht. Daher wurde in einem vierten Ansatz nun systematisch der Einfluss der Multivalenz supramolekularer Liganden auf die Tasp/Impα-Wechselwirkung untersucht (Fokkens *et al.* 2005). Hierfür wurde unter Verwendung anionischer molekularer, Lysin-bindender Pinzetten eine Reihe multivalenter Binder entwickelt, welche bis zu fünf anionische Pinzetten miteinander verknüpfen. In der Tat konnte die Bindung der Liganden an das im Taspase1-Loop befindliche NLS über Fluoreszenztitration und NMR nachgewiesen werden.

Darüber hinaus bestätigten *Pulldown*-Experimente eine Korrelation zwischen der Valenz der Bindungseinheiten und der inhibitorischen Wirkung auf die Taspase1/Impα-Wechselwirkung. Diese Beobachtung konnte in kolorimetrischen Spaltungstests ebenfalls für die proteolytische Aktivität von Taspase1 bestätigt werden.

Insgesamt wurden im Rahmen dieser Arbeit in insgesamt vier Studien verschiedene supramolekulare Liganden zur funktionellen Hemmung der krebsrelevanten Protease Taspase1 entwickelt. Mehrere, voneinander unabhängige Ansätze zur rationalen Verwendung der wahlweise kationischen oder anionischen Binder erlaubten, die Wechselwirkung mit Impα als essentiellen Schritt der zellulären Taspase1-Aktivierung sowie teils auch deren proteolytische Aktivität direkt zu inhibieren. Insbesondere das systematische Vorgehen erlaubt einen Hypothesen-getriebenen Wissenstransfer. So kann möglicherweise auch die rationale Entwicklung effektiverer Inhibitoren weiterer biologisch relevanter Zielproteine hin zu neuartigen Therapieansätzen für Krebs und andere Erkrankungen gezielt vorangetrieben werden.

Summary

Traditional cancer therapy with broad-range chemotherapeutics is still the first line of defense for most forms of cancer, but their cytotoxicity causes systemic collateral damage, leading to numerous adverse reactions. To avoid this, therapies have become more target specific over time. Due to its origins, however, cancer is a vastly diverse disease. Development of specific inhibitors is therefore a crucial but demanding task (Hanahan and Weinberg 2011). One such protein, overexpressed in numerous solid and liquid forms of cancer, is Taspase1 (**T**hreonine **Asp**art**ase 1**) (Hsieh *et al.* 2003; D Wünsch *et al.* 2016). This protease recognizes a well-conserved cleavage motif $Q^{3}(F/I/L/V)^{2}D^{1}/G^{1'}X^{2'}D^{3'}D^{4'}$ (Bier *et al.* 2011b). Among its confirmed substrates are prominent oncogenes like MLL1 or transcription factors such as TFIIA (Gribko *et al.* 2017; Hsieh *et al.* 2003). While a Taspase1 knock-out is well-tolerated in healthy adult tissue, it leads to a proliferation block and apoptosis in tumor cells, which makes Taspase1 an attractive target for cancer therapy (Takeda *et al.* 2006; Chen *et al.* 2010).

However, Taspase1 is not affected by general protease inhibitors and only a few specific inhibitors for Taspase1 have been developed, which effectively inhibit the proteolytic activity (Chen *et al.* 2012; Johannes van den Boom *et al.* 2014; van den Boom *et al.* 2020; Hsieh *et al.* 2003). Intracellular activation of Taspase1 involves its nuclear translocation mediated by Importinα (Impα), which makes this pivotal interaction a potential novel target for inhibition (Bier *et al.* 2011a; van den Boom *et al.* 2016). Thus, the aim of this thesis was the rational development of supramolecular ligands to efficiently interfere with this relevant interaction, thereby potentially impeding the activation process (Bier *et al.* 2011a).

In order to achieve these goals, we first focused on cationic ligands to address clusters of acetic amino acids flanking the lysine-rich bipartite nuclear localization signal (NLS) located in the rather flexible Taspase1 loop. In an initial study, we were able to develop the first inhibitors that effectively disrupt the Tasp/Impα-interaction (Pasch *et al.* 2021). The respective trivalent supramolecular ligands utilized guanidiniocarbonyl-pyrrole (GCP) units as cationic binders. Those were connected by an oligomer backbone and substituted either with or without lysine (ligands **3G** and **3GL**). In a next step, those ligands were additionally equipped with a polyethylene glycol (PEG) steric shield to allow for efficient blocking of the binding site (ligands **3GP** and **3GLP**). Whereas binding to Taspase1 could be verified for all ligands by surface plasmon resonance (SPR), inhibition of Tasp/Impα-interaction was only evident for PEGylated ligands in pull-down assays.

In a second study, we again developed supramolecular ligands based on cationic GCP-units, which were functionalized with different protective groups (Höing *et al.* 2021). Here, the bivalent supramolecular ligand **2GC** emerged as a ligand not only able to effectively disrupt

the Tasp/Impα-interaction in pull-down assays (IC_{50} = 35 µM), but also exhibit toxic effects on Taspase1-expressing tumor cells (EC_{50} = 40-70 µM). Besides the GCP-unit, which is essential for binding, a bivalent composition as well the existence of the hydrophobic C(bz) protecting group arose as key factors for the ligand's effectiveness.

So far ligand development was conceptionally based on binding to acidic amino acids such as aspartate or glutamate, which are present in close proximity to the bipartite basic NLS within a flexible loop on the surface of Taspase1. Further investigations therefore aimed to directly address the NLS as relevant target structure for anionic binders. In a third study based on this concept, we introduced the first dual inhibitor for Taspase1 that also exhibited optimized binding characteristics. The bivalent supramolecular ligand **11d** utilized para-substituted phosphate groups to mediate binding (Octa-Smolin *et al.* 2018). Virtual docking studies predicted the NLS containing loop as energetically most favorable binding site. Fluorescence anisotropy titration determined a nanomolar binding affinity (K_D = 300 nM) for Taspase1, and binding to the loop domain was shown by fluorescence titration (K_D = 3 µM) and NMR (Nuclear magnetic resonance) titration. Pull-down assays revealed disruption of the Tasp/Impα-interaction in the micromolar range (IC_{50} = 6 µM). Moreover, a colorimetric cleavage assay revealed an inhibitory effect on the proteolytic activity in the same range (IC_{50} = 2 µM) which has been verified in cell experiments utilizing a transfectable biosensor for intracellular Taspase1 activity.

The studies conducted so far clearly emphasized that efficient inhibition of Taspase1 can only be enforced by employing multiple supramolecular binding motifs. As a direct consequence, a systemical investigation assessing the influence multivalency of supramolecular ligands has on the Tasp/Impα-interaction was conducted. Therefore, a series of multivalent ligands based on the lysine-binding molecular tweezer CLR01 with up to five binding units was examined (Fokkens *et al.* 2005). Ligand binding to the NLS-containing surface loop was confirmed by fluorescence and NMR titration experiments. Moreover, pull-down assays revealed a positive correlation of the ligands' valency with disruptive effect on the Tasp/Impα-interaction. Importantly, colorimetric cleavage assays showed that the proteolytic activity of Taspase1 is likewise inhibited.

Taken together, this thesis succeeded in developing different supramolecular ligands to functionally inhibit the oncologically relevant protease Taspase1 in four complementary studies. Several independent scientific approaches using multivalent assemblies of either cationic or anionic binders resulted in efficient interference with the Impα interaction as an essential step for the intracellular activation of Taspase1, as well as in part its proteolytic activity itself. Notably, the systematic nature of our approach allowed to directly compare and evaluate the supramolecular design strategies, facilitating a hypothesis-driven transfer of

knowledge. This might successively drive the rational development of more effective inhibitors for other biologically relevant target proteins towards novel therapeutic approaches for cancer and other disease patterns.

1 Introduction

1.1 Cancer - terminology and statistics

In colloquial speech, the terms “tumor” and “cancer” are often used synonymously. However, the term tumor simply describes an abnormal mass of cells created by uncontrolled cell division (Cooper 2000). A benign tumor, such as a birthmark or lipoma, will not spread into surrounding tissue and is easy to remove (Cooper 2000). By contrast, a malignant tumor will start to invade and destroy surrounding healthy tissue and can rapidly spread through the body by invading the blood or lymph systems in a process called metastasis (Cooper 2000). The disease resulting from a malignant tumor is termed cancer (Cooper 2000).

Cancer is among the leading causes of death worldwide. Since a malignant transformation can occur in virtually any cell and any tissue, there are great varieties of cancers, with more than 200 forms currently described (Perez-Diez *et al.* 2007; Smith 2013). The frequency and distribution of different types of cancer varies greatly by region, gender, genetic background, and environment (Sung *et al.* 2021). For example, when comparing men and women, there are 19 % more incidences and 43 % more mortalities involving cancer in 2020.

Worldwide, and among all genders and age groups, breast cancer is estimated to have the most incidences (12 %) followed by lung cancer (11 %) and colorectal cancer (10 %) while lung cancer accounts for the most deaths (18 %), followed by colorectal cancer (9 %) (World Health Organization 2021; Sung *et al.* 2021). According to the Global Cancer Observatory, the number of incidences and mortalities is increasing and it is estimated that the global cancer burden will increase by 47 % until 2040 due to demographic changes (Sung *et al.* 2021). Understanding the molecular and biochemical mechanisms that underly cancer progression

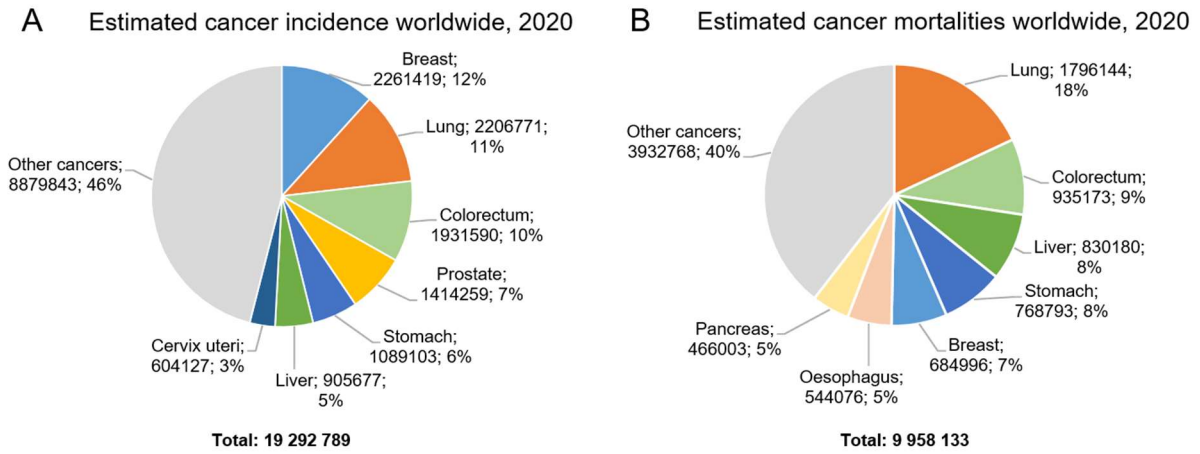


Figure 1: Distribution of new incidences (A) and mortalities (B) of cancer in 2020 by cancer type (worldwide, both genders, all ages) (Sung *et al.* 2021).

and developing new therapeutics for cancer treatment therefore remain imperative tasks in science.

1.2 Cancerogenesis

The most common theory for the origin of cancer is the somatic mutation theory proposed by Sota in 1980 (Sorsa 1980). According to this theory, cancer is a cell-based disease that originates from a single mutated cell in otherwise healthy tissue. While this is generally accepted as the dominant working theory, some other concepts exists. In 1999, for example, Sonnenschein and Soto proposed the opposing tissue field organization theory, which blames the origins of cancer development on a breakdown of tissue organization, resulting in the subsequent removal of cellular constraints (Sonnenschein and Soto 1999; Montévil and Pocheville 2017). To this day, the compatibility or contradiction of these two models is the subject of keen debates (Bedessem and Ruphy 2015; Bizzarri and Cucina 2016; Montévil and Pocheville 2017; Bedessem and Ruphy 2017).

In line with somatic mutation theory, which sees cancer as developing from an accumulation of genetic mutations in a single cell, Hanahan and Weinberg have proposed a set of common traits and capabilities in 2000 that result from different cancer genotypes. These features, acquired during multiple steps of tumorigenesis, are known as the “hallmarks of cancer” (Hanahan and Weinberg 2000). The six hallmarks originally described are the capabilities „self-sufficiency in growth signals“, „insensitivity to growth-inhibitory signals“, „evasion of programmed cell death apoptosis“, „limitless replicative potential“, „sustained angiogenesis“, and „tissue invasion and metastasis“ (Hanahan and Weinberg 2000).

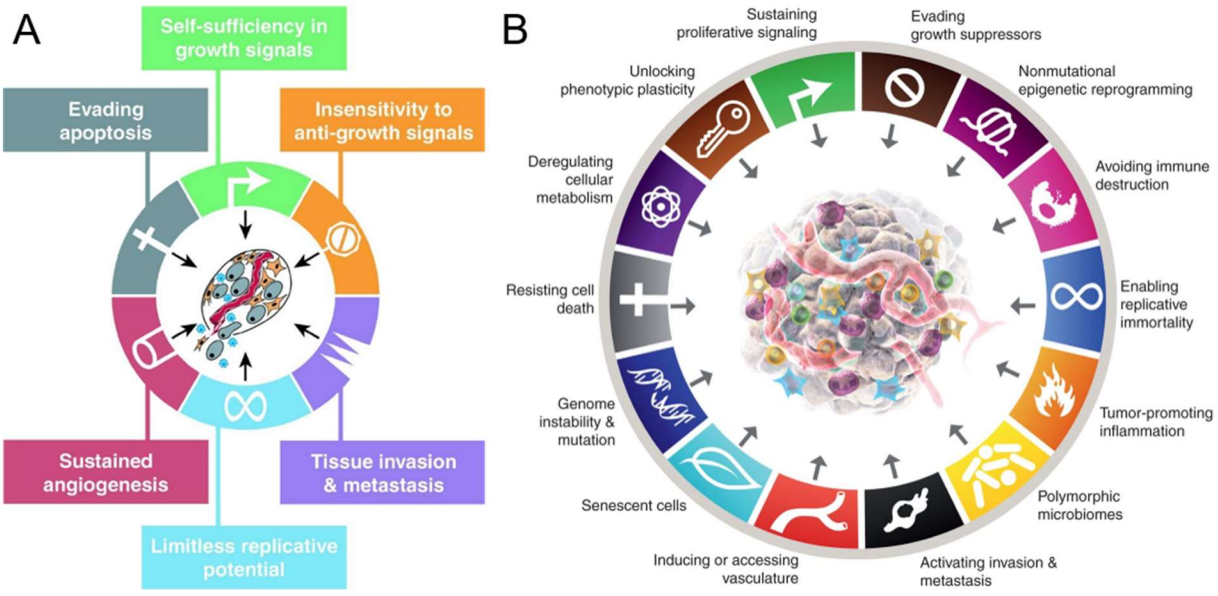


Figure 2: Development of the hallmarks of cancer. The original hallmarks of cancer proposed in 2000 (A) comprised six traits. Over the last two decades of research, more characteristics of cancer had to be added. Currently, the hallmarks of cancer comprise fourteen traits. Images taken from Hanahan and Weinberg (A) and Hanahan (B) (Hanahan and Weinberg 2000; Hanahan 2022)

Over the next two decades, researchers understanding of cancer became more complex. In 2011, Hanahan and Weinberg added the emerging hallmarks „deregulation of cellular energetics“ and „avoiding immune destruction“ as well as the enabling characteristics „genome instability“ and „tumor promoting inflammation“ (Hanahan and Weinberg 2011). By 2022, Hanahan had also added the emerging hallmarks „unlocking phenotypic plasticity“ and „senescent cells“ as well as the enabling characteristics „nonmutational epigenetic programming“ and „polymorphic microbiomes“ (Hanahan 2022). The hallmarks currently comprise 10 capabilities and four enabling characteristics (Hanahan 2022). The increase in numbers reflects the complexity and diversity of tumorigenesis (Hanahan 2022).

1.3 Proteases in cancer development

Proteases are enzymes that hydrolyze peptide bonds (Quesada *et al.* 2009). Depending on their active site, they can be categorized into five groups. The vast majority of the proteases utilize metal, serine, or cysteine to catalyze the hydrolyzation of peptide bonds while only a few proteases belong to the aspartate and threonine group (López-Otín and Matrisian 2007; Quesada *et al.* 2009).

Based on their function, they can generally be divided into exoproteases, which cleave from the termini of the peptide chain, and endopeptidases , which cleave within the peptide chain (López-Otín and Matrisian 2007). While exopeptidases are mostly involved with degradation of proteins, endopeptidases recognize cleavage sequences, enabling them to

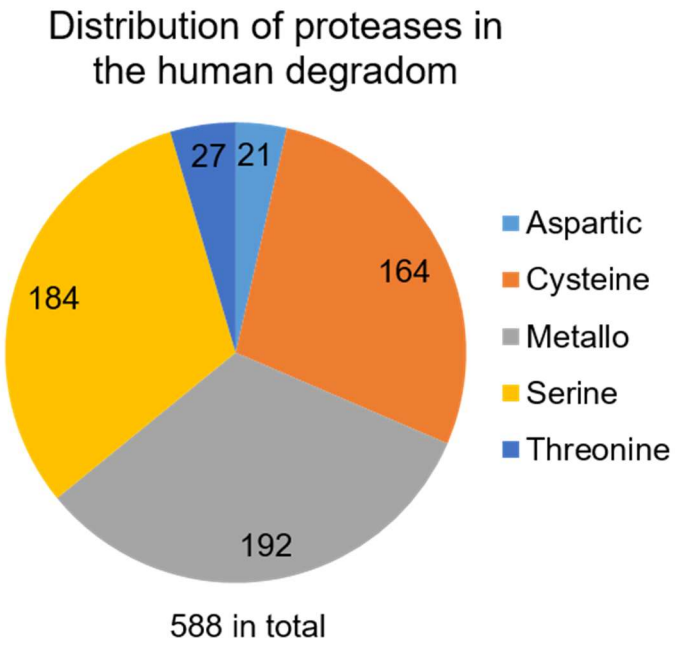


Figure 3: Schematic illustrating the distribution of human proteases among the different classes. There are currently 588 known proteases encoded in the human genome. Based on their active site, they are classified into five different groups. Numbers are based on the Degradome database (Quesada *et al.* 2009).

specifically target proteins, which leads to their involvement in regulatory processes (Wünsch *et al.* 2016).

Since proteolysis is an irreversible, post-translational modification of proteins, the dysregulation of proteases can have calamitous and pathogenic consequences for an organism (López-Otín and Bond 2008). Obviously, dysregulated proteases have also been associated with all stages of tumorigenesis and cancer progression (López-Otín and Matrisian 2007; López-Otín and Bond 2008; Mason and Joyce 2011). In reference to the hallmarks of cancer mentioned before, an obvious example for tumor-promoting proteases involved in tumor progression is resulting invasiveness (Hanahan and Weinberg 2000).

MMP7 is a member of the matrix metallo protease family, which can collectively degrade all components of the extracellular matrix (Reunanen and Kähäri 2013). It seems natural that an overexpression of MMP7 has been associated with increased invasiveness of numerous cancer cell lines (Liotta *et al.* 1980; Reunanen and Kähäri 2013). Less obviously, MMP7 is also associated with tumor survival and avoidance of apoptosis by cleaving the Fas ligand as well as the Fas receptor, normally triggering an intracellular caspase cascade that ultimately results in apoptosis of abnormal cells (Mitsiades *et al.* 2001; Susanne Strand *et al.* 2004; Almendro *et al.* 2009).

As a result of proteolytic processing, the involvement of proteases in cancerogenesis becomes more complex, if we consider the downstream effects triggered by their substrates. One example is MLL1 (Mixed lineage leukemia protein 1), a 500 kDa lysine-specific methyltransferase required for proper HOX gene expression (Hsieh *et al.* 2003; Takeda *et al.* 2006) While generally active in its uncleaved form, MLL1 contains two conserved cleavage sites for endoproteolytic processing. Upon hydrolyzation, the resulting N-terminal 320 kDa and C-terminal 180 kDa subunits of MLL1 utilize FYRN and FYRC domains to reassemble (Hsieh *et al.* 2003; Pless *et al.* 2011; Niizuma *et al.* 2021). The resulting heterodimer has an increased H3 methyl transferase activity and is stabilized (Hsieh *et al.* 2003; Takeda *et al.* 2006).

The MLL gene is often involved in translocation events resulting in the development of leukemia, a prominent example is translocation of the proto-oncogene AF4 leading to an AF4-MLL1 fusion protein (Hsieh *et al.* 2003; Bier *et al.* 2012b). While AF4 is normally targeted by E3-ligases SIAH1 and SIAH2 leading to effective degradation, the resulting fusion protein is protected upon heterodimerization of the MLL1 subunits resulting from endoproteolytic processing (Bursen *et al.* 2004; Pless *et al.* 2011). The protease responsible for cleaving MLL1 was termed Taspase1 (Hsieh *et al.* 2003).

1.4 Taspase1

1.4.1 General function and oncological relevance

Taspase1 (**T**hreonine **Asp**art**ase** 1) is a highly conserved protease that belongs to the type 2 asparaginase family (Hsieh *et al.* 2003). Within this protein family it shows the unique characteristic that it is able to function as a protease in *trans*, while it is also capable of autoproteolytic processing in *cis* (Hsieh *et al.* 2003; Khan *et al.* 2005). It uses threonine to hydrolyze peptide bonds following aspartate in a well conserved cleavage motif, first reported as $Q^{3}X^{2}D^{1}G^{1'}$, which was later refined to the consensus sequence $Q^{3}(F/I/L/V)^{2}D^{1}/G^{1'}X^{2'}D^{3'}D^{4'}$ (Hsieh *et al.* 2003; Bier *et al.* 2011b). There are more than 1400 proteins containing the general cleavage site and 27 putative targets bearing the refined cleavage site (Bier *et al.* 2011b).

Taspase1 was first described as the protease which cleaves MLL1, leading to increased stability and activity. Other confirmed substrates of Taspase1 are ubiquitous expressed TFIIAα-β (Zhou *et al.* 2006; Schrenk *et al.* 2018), the testis-specific TFIIA-like factor ALF (Zhou *et al.* 2006), UFS2 (Bier *et al.* 2011b) and the catalytic subunit of DNA polymerase ζ REV3L (Wang *et al.* 2020). Taspase1 is a central protease during embryonic development and regulates the expression of homeotic genes like HOX (Hsieh *et al.* 2003). A homozygote knockout in embryonic mice leads to 87% lethality during the first three weeks after birth and the remaining mice display reduced size, defects in liver hematopoietic stem cells and skeletal deformations, especially craniofacial malformations (Takeda *et al.* 2006; Niizuma *et al.* 2021).

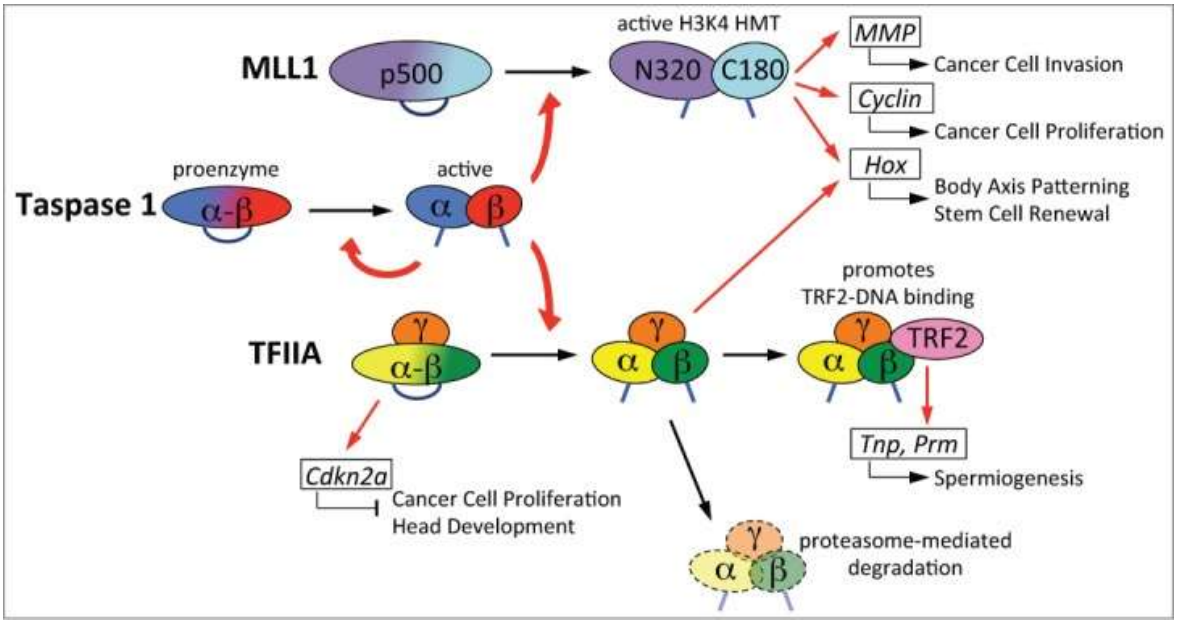


Figure 4: Exemplary functions of Taspase1 in cancer development. Active Taspase1 processes MLL1, which leads to increased Histone methyltransferase (HMT) activity and induces the expression of matrix metalloproteases (MMP), resulting in invasiveness, and Cyclins, resulting in increased proliferation. Processing fo TFIIA leads to its rapid degradation and results in down-regulation of Cyclin dependent kinase inhibitor 2a (Cdkn2a), which has inhibitory effects on the proliferation of tumor cells. Image taken from Niizuma *et al.* (Niizuma *et al.* 2015).

Also, a knockout in embryonic mouse fibroblasts leads to decreased proliferation, caused by decreased levels of cyclin-dependent kinases (Balkin *et al.* 2019).

Due to the severe consequences of a Taspase1 knockout on developmental processes, there are only two described examples of Taspase1 deletions in humans. The first was a homozygote deleterious mutant (exons 5-10) (Suleiman *et al.* 2018), while the second was a heterozygous *de novo* missense variant (V343M) (Balkin *et al.* 2019). Both phenotypes overlap with the knockout phenotype described in mice, resulting in skeletal abnormalities especially in the craniofacial area, microcephaly, and an overall decreased body size (Suleiman *et al.* 2018; Balkin *et al.* 2019). Of note, the Taspase1 knockout phenotype partially overlaps with the Wiedemann-Steiner syndrome in humans that is caused by a dysregulation of MLL1 (Suleiman *et al.* 2018). However, an induced Taspase1 knockout in seven week-old mice did not show toxic effects and it seems to be well tolerated in adult tissue, which would allow potential targeting of Taspase1 for therapy (Chen *et al.* 2012).

Since Taspase1 was originally described in the context of MLL1's involvement in leukemia, it has been associated with tumor development from the very beginning. Indeed, Taspase1 is overexpressed in numerous solid and liquid forms of cancer (Chen *et al.* 2010; Wünsch *et al.* 2016). For example, it has been shown to promote proliferation and migration in gastric cancer cells by upregulation of the p-AKT/AKT as well as epithelial-mesenchymal transition pathways (Wan *et al.* 2021). It also mediates down-regulation of Cycline dependent kinase inhibitor 2A (CDK2A) in Head and neck squamous cell carcinoma (HNSCC) by TFIIA (Gribko *et al.* 2017), as well as tumor progression in HER2-positive breast cancer by MLL1 (Dong *et al.* 2014). However, Taspase1 overexpression alone or in combination with the prominent oncogenes MYC or RAS is not sufficient to promote a malignant transformation (Chen *et al.* 2010). Taspase1 has therefore been classified as a "non-oncogene-addiction protease", since it contributes to tumor development without being an oncogene itself and Taspase1-overexpressing tumor cells become sensitive towards Taspase1 deprivation (Chen *et al.* 2010; Wünsch *et al.* 2016). As a consequence, Taspase1 has been recognized as a valuable biomarker for cancer, but more importantly, as a highly attractive target for future cancer therapeutics (Chen *et al.* 2010; Dong *et al.* 2014; Wan *et al.* 2021).

1.4.2 Structure and activity of Taspase1

There are different descriptions of Taspase1's oligomeric state, which partially results from the autoproteolytic event. While inactive Taspase1 is a monomer, and a dimer would be a homodimer, active Taspase1 is a heterodimer after autoproteolysis and a dimer would be a heterotetramer. To avoid confusion and to unify the different descriptions used in literature, I will differentiate between "active" and "inactive" Taspase1 and give the number subunits in this thesis, e. g. "active αβ heterodimer" or "inactive $(\alpha\beta)_2$ homodimer".

The first investigation of Taspase1 structure was conducted by Khan *et al.* in 2005 using crystallization (Khan *et al.* 2005). This study revealed an ongoing problem with structural investigations. Since recombinant wild type Taspase1 undergoes partial autoproteolysis during purification, this results in a heterogeneous protein population, which is not suited for structural investigations (Khan *et al.* 2005). To avoid this, they utilized a T234V mutant incapable of autoproteolyis to investigate the structure of inactive Taspase1. They also used a mixture of the two separately expressed subunits (α residues 1-206 and β residues 234-420) to investigate the structure of active Taspase1 (Khan *et al.* 2005).

Crystallization revealed dimerization of two Taspase1 units, which resulted in an $(\alpha\beta)_2$ homodimer for the inactive $Taspase1_{T234V}$ and an $\alpha_2\beta_2$ heterotetramer for the active $Taspase1_{\alpha1\text{-}206/\beta234\text{-}420}$ (Khan *et al.* 2005). Since the active sites were not in close proximity to the interaction site, the authors exclude cooperative cleavage (Khan *et al.* 2005). The crystal structures contained several disordered regions and in all resulting models a structure directly preceding the active site T234 was missing, which the authors attribute to the flexibility of this structure (Khan *et al.* 2005).

Contrary to this first study of the Taspase1 structure, Bier *et al.* were able to show that Taspase1 predominantly exists as an active αβ heterodimer in a cellular environment and only has a weak tendency to form the $\alpha_2\beta_2$ heterotetramer complex (Bier *et al.* 2012a). As an explanation, they proposed a dependency on the protein concentration, which was higher in crystallization experiments (Bier *et al.* 2012a). They also showed that the αβ heterodimers are enzymatically active after autoproteolysis and that permanent maintenance of the $\alpha_2\beta_2$ heterotetrameric complex observed in the crystal structure of active Taspase1 is not required for protease activity (Khan *et al.* 2005; Bier *et al.* 2012a).

Van den Boom *et al*. investigated the missing structure in previous crystal structures, which directly preceded the active site (Khan *et al.* 2005; van den Boom *et al.* 2016). Because of the conserved structure among the members of the type 2 asparaginases, and the shared necessity of autoproteolytic activation, they investigated the unresolved structures of six type 2 asparaginases *in silico,* using homology modelling, sequence-based secondary structure predictions, and molecular dynamics. This resulted in the prediction of a helix-turn-helix motif for Taspase1, which they termed the Taspase1 loop (van den Boom *et al.* 2016). They validated the results *in vitro*, utilizing nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopy (van Boom 2015; van den Boom *et al.* 2016). Investigation of an isolated Taspase1 loop (G178-D233) confirmed the predicted high content of helical structures and a turn motif between L200-D210 (van den Boom *et al.* 2016). The resulting loop model was added to the crystal structures of Taspase1, resulting in an enhanced Taspase1 model (van den Boom *et al.* 2016). NMR with wild type Taspase1 also indicated that the loop is indeed

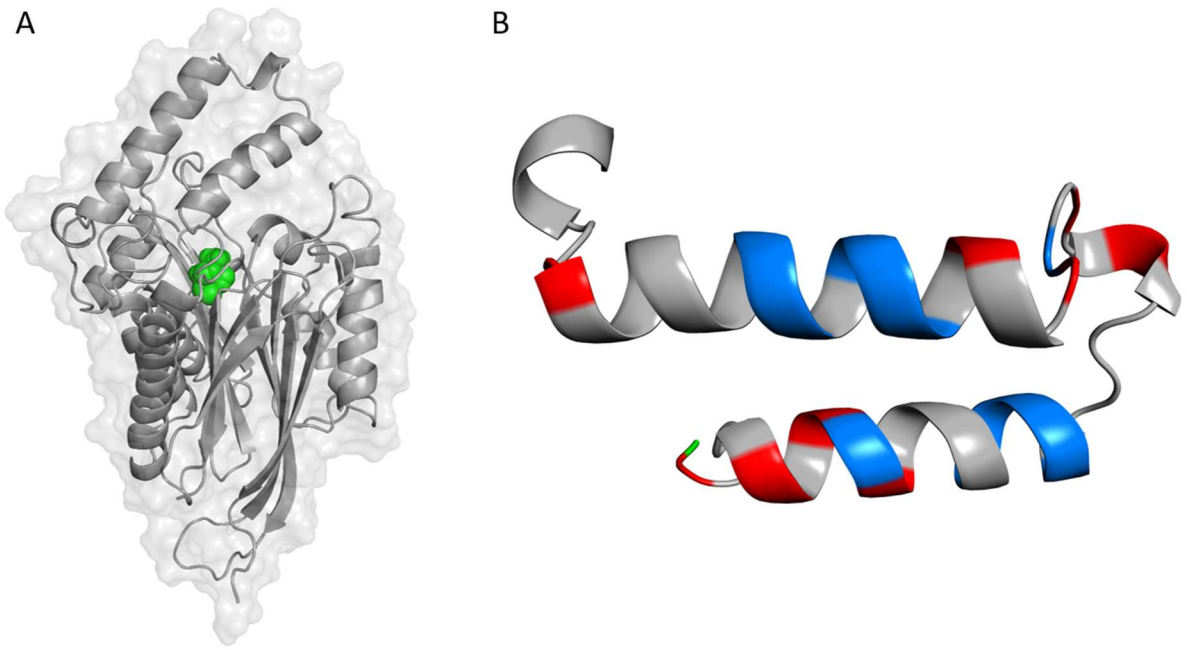


Figure 5: Model of Taspase1 (A) and the Taspase1 loop (B) according to van den Boom *et al.* The loop directly precedes the active site T234 (green) and contains clusters of basic (blue) or acetic (red) amino acids (van den Boom *et al.* 2016).

a dynamic structure that flexibility of freshly purified and therefore mostly inactive Taspase1 in this region increases over time. This reflects autoproteolysis and release of the loop preceding the active site (Khan *et al.* 2005; van den Boom *et al.* 2016).

The most recent study of the Taspase1 structure was conducted by Nagaratnam *et al.* in 2021 and focused on crystallization of the Taspase1 loop (Nagaratnam *et al.* 2021). Here, they used a circularly permutated Taspase1 mutant with both subunits in reverse order, resulting in a βα monomer, instead of the wild type αβ monomer, expressed as a single chain protein with both subunits coupled by a peptide linker (Nagaratnam *et al.* 2021). As a result, the Taspase1 variant has the active site T234, which would normally only be accessible after autoproteolyis, located at the N-terminus. Also, the Taspase1 loop, which would directly precede the active site in the wild type αβ monomer, is located at the C-terminus (Nagaratnam *et al.* 2021). The crystal structure generated shows Taspase1 as a $(\beta\alpha)_2$ homodimer and revealed that in this altered version of Taspase1, the C-terminal loop containing the NLS is no longer a flexible helix-turn-helix structure, but a long and rigid helix. They therefore proposed the term “long helical fragment” instead of Taspase1 loop (Nagaratnam *et al.* 2021).

Nagaratnam *et al.* also describe how a shortening of the C-terminal loop results in decreased Taspase1 activity. They conclude that the originally crystallized active Taspase1$_{\alpha 1\text{-}206/\beta 234\text{-}420}$ formerly noted by Khan *et al.*, which also had parts of the loop removed for better crystallization, must have also been a permanently inactive mutant, making this the first real crystal structure of active Taspase1 (Khan *et al.* 2005; Nagaratnam *et al.* 2021). Of note, the enzymatic activity of Taspase1$_{\alpha 1\text{-}206/\beta 234\text{-}420}$ was demonstrated beforehand in separate studies

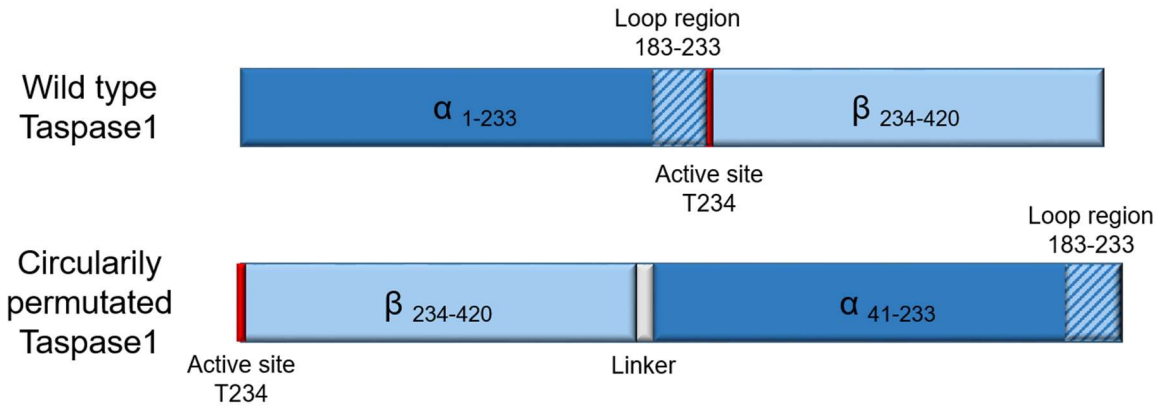


Figure 6: Schematic comparison of wild type Taspase1 and the circularly permutated Taspase1. Inactive (αβ) Taspase1 is a single chain peptide with loop and active site positioned in the center. The circularly permutated Taspase1 is a (βa) single chain peptide with the active site at the N-terminus and the loop at the C-terminus. Based on Nagaratnam *et al.* (Nagaratnam *et al.* 2021).

by Khan *et al.* and van den Boom *et al.* (Khan *et al.* 2005; van den Boom *et al.* 2020). Finally, Nagaratnam *et al.* investigated the higher order of circularly permutated Taspase1$_{\beta 234\text{-}420/\alpha 41\text{-}233}$ via dynamic light scattering. They report a dependency of the oligomeric state on protein concentration, even describing $(\beta\alpha)_3$ homohexamers, which would explain the former irregularities observed (Khan *et al.* 2005; Nagaratnam *et al.* 2021).

1.4.3 The intramolecular Taspase1 activation process

Like the other members of the type 2 asparaginase family, Taspase1 is synthesized as an inactive proenzyme and therefore has to undergo autoproteolytic processing to obtain its proteolytic function (Hsieh *et al.* 2003). *In silico* studies based on homology modelling with other members of the type 2 asparaginase family, and *in vitro* studies focused on site directed mutagenesis of Taspase1, have revealed that dimerization is needed for the occurrence of autoproteolysis, revolving around an intrinsic, serine-protease like function of Taspase1 (Sabiani *et al.* 2015). Each Taspase1 monomer contains a docking head in the α subunit (central amino acids: W173, L146, I160, C163) as well as a docking site in the β subunit (central amino acids: R262, E295, R299) (Sabiani *et al.* 2015). Exchange of these central amino acids abolishes dimerization and subsequent autoproteolytic processing (Sabiani *et al.* 2015). Binding of the docking head to the docking zone of an opposing proenzyme leads to minor conformational changes, which ultimately causes the induction of an intramolecular cleavage event (Sabiani *et al.* 2015) The modelling suggests that the mechanism is based on the induction of a movement of R262 towards E295, formally bound to R299 (Sabiani *et al.* 2015). This movement repositions S291 closer to the later active site, where it hydrolyses the peptide bond between D233 and T234, releasing the active site for substrate binding (Sabiani *et al.* 2015). The activation process is referred to as a “dock and click” mechanism, which activates an intrinsic serine protease-like function of Taspase1 (Sabiani *et al.* 2015). While the mechanism is a proposed model and has to be further validated, exchange of either S291,

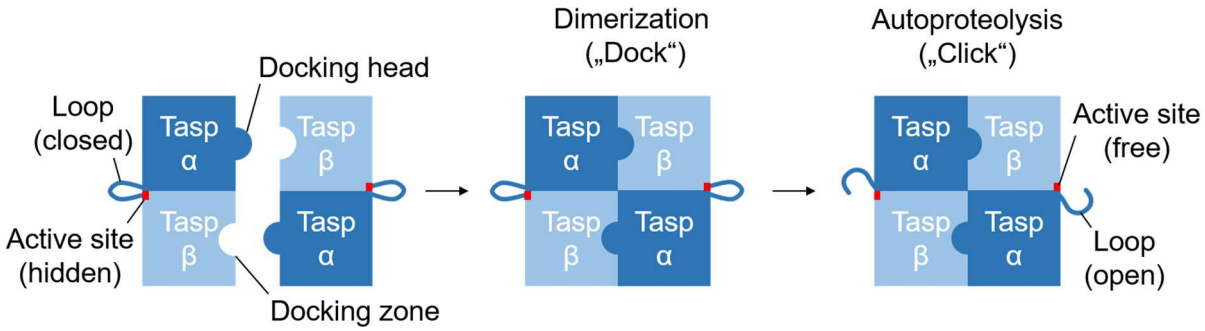


Figure 7: The intramolecular activation of process of Taspase1 based on the “dock and click” model by Sabiani *et al.* The α subunit of Taspase1 contains a docking head and the β subunit contains a docking zone. Dimerization of two inactive Taspase1 monomers („dock”) leads to a movement of S291 towards the active site, resulting in autoproteolysis („click”) that frees the active site. Based on Sabiani *et al.* (Sabiani *et al.* 2015).

D233 or T234 abolishes autoproteolysis while not interfering with the dimerization itself (Sabiani *et al.* 2015).

1.4.4 Taspase1 and Importinα – translocation as pre-requisite for activation

In theory, autoproteolysis by dimerization of two inactive Taspase1 monomers should be possible in the cytoplasm. However, for reasons unknown, Taspase1 is not autoproteolytically processed while it is located in the cytoplasm and remains in its proenzymatic state (Bier *et al.* 2011a). This observation has led to another prerequisite for Taspase1 activation in eukaryotic cells, namely its nuclear translocation by interaction with Importinα (Impα), which is also known as also known as Karyopherinα2 (Bier *et al.* 2011a).

Impα is a 58 kDa protein and belongs to the Importin or Karyopherin superfamily of proteins (Miyamoto *et al.* 2016). These proteins transfer proteins with the respective target sequence, known as NLS, from the cytoplasm through the nuclear pore complex (NPC) into the nucleus (Miyamoto *et al.* 2016). Among the members of this family are many adapter proteins that recognize different target sequences, which allows for selectively targeting a multitude of different proteins with only a few transport carriers (Lange *et al.* 2007; Miyamoto *et al.* 2016). Impα is one of these adapter proteins, and consists of three domains (Miyamoto *et al.* 2016). On the N-terminus, there is a binding domain for the transport carrier Importinβ (Impβ), which in turn transports the bound Impα/cargo-complex into the nucleus via the NPC (Miyamoto *et al.* 2016). This domain also has an autoinhibitory function to avoid the nuclear import of unloaded Impα (Miyamoto *et al.* 2016). If no cargo is bound, Impα binds this domain and is therefore unavailable for Impβ (Miyamoto *et al.* 2016). Binding of cargo proteins releases this domain and the loaded Impα can interact with Impβ (Miyamoto *et al.* 2016).

On the C-terminus, there is the binding region for nuclear export proteins to retrieve unloaded Impα from the nucleus (Miyamoto *et al.* 2016). The major part of the protein is the central cargo binding domain, which consists of 10 helical armadillo repeats (Miyamoto *et al.* 2016). The

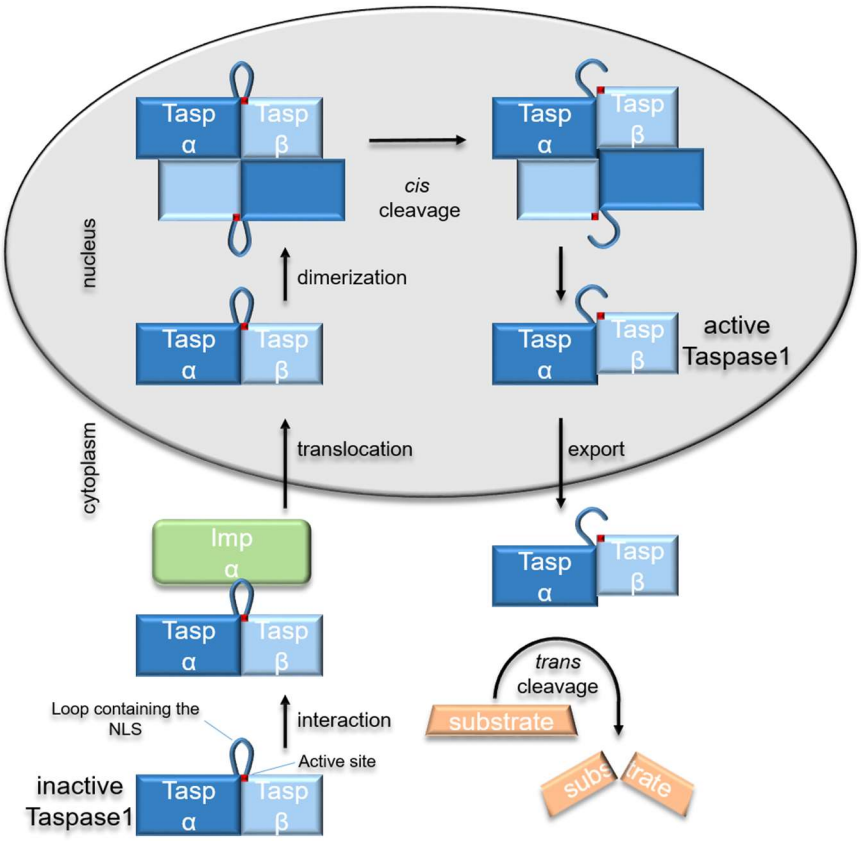


Figure 8: Schematic illustrating the activation process and transport of Taspase1. Taspase1 interacts with Impα, which in turn binds to Impβ and translocates the protein complex into the nucleus. There Taspase1 is activated by autoproteolytic processing following dimerization and active Taspase1 accululates at the nucleoli, where it interacts with NPM1. Due to the interaction, Taspase1 can hitchhike on NPM1 NES and get transient access to the cytoplasm, where it is bound by Impα again.

armadillo repeats 2-4 form the major binding site, while armadillo sites 6-8 form the minor sites (Xu *et al.* 2010; Miyamoto *et al.* 2016). These two distinct binding sites enable Impα to bind proteins with a monopartite NLS as well as proteins with a bipartite NLS (Miyamoto *et al.* 2016).

Prior to the structural investigation of the Taspase1 loop, Bier *et al.* revealed that the Taspase1 loop G178-D233 contains a bipartite NLS consisting of basic amino acid clusters (197KRNKRKLELAERVDTDFMQLKKRR220) (Bier *et al.* 2011a). This NLS is highly conserved among all known Taspase1 variants.(Bier *et al.* 2011a) *Bier et al.* first dissected the function of the region, later known as the Taspase1 loop, and described the importance of the area, which had been neglected as an unresolved structure until then, for the activation of Taspase1 (Bier *et al.* 2011a). Microinjection of GST-GFP fusion proteins with different fragments of the Taspase1 loop showed that both parts of the Taspase1 NLS are necessary for effective nuclear import (Bier *et al.* 2011a). They also showed that the NLS serves as an interaction domain for Impα and that alteration of the NLS sequence contains Taspase1 in the cytoplasm with no autoproteolytic processing (Bier *et al.* 2011a). Taspase1 is bound to adaptor protein Impα via its bipartite NLS and the protein complex is transferred into the nucleus through the NPC by transport carrier Impβ (Bier *et al.* 2011a). In the nucleus, Impβ binds to

RAN-GTP, leading to dissociation of the ternary protein complex and release of Taspase1 (Bier *et al.* 2011a; Christiansen and Dyrskjøt 2013). Impα is bound by export carrier CAS and Ran-GTP, which leads to its nuclear export and subsequent release following RAN-GTP hydrolyzation in the cytoplasm (Miyamoto *et al.* 2016). Inside the nucleus, two inactive αβ Taspase1 proenzymes form a homodimer, leading to the induction of the autoproteolytic activation of both proteins (Bier *et al.* 2011a; Sabiani *et al.* 2015).

Most of the active Taspase1 resides within the nucleus and shows a distinct accumulation at the nucleoli (Bier *et al.* 2011a). This accumulation is dependent on Taspase1 activity, as permanently inactive Taspase1 mutants will be located in the nucleus, but not accumulate at the nucleoli (Bier *et al.* 2011a). There have been numerous discussions about the contradiction between the nucleolar localization of active Taspase1 and the cytoplasmic localization of many substrates, as Taspase1 does not seem to contain any form of nuclear export signal (NES) (Bier *et al.* 2011a). Active Taspase1 is an interaction partner of Nucleophosmin1 (NPM1), which is located at the nucleoli, but also contains a weak NES, leading to a minor export into the cytoplasm (Bier *et al.* 2011a). Ectopic expression of a pathogenic NPM1 mutant with a much stronger NES leads to localization of fluorescent Taspase1-fusion proteins into the cytoplasm (Bier *et al.* 2011a). So by interaction with NPM1, Taspase1 can hitchhike on its NES, gaining transient access to the cytoplasm in its active state before it is re-imported (Bier *et al.* 2011a).

1.4.5 Therapeutic approaches to interfere with Taspase1

Since Taspase1 plays an important role in tumorigenesis, it has been the subject of drug design. Several studies have investigated the use of general protease inhibitors against Taspase1, but they did not affect it (Lee *et al.* 2009; Knauer *et al.* 2011; Chen *et al.* 2012). In addition, Bortezomib, an inhibitor targeting the β subunit of the 20 S proteasome, the only other protease in mammals that uses a N-terminal threonine as an active site, has been shown to have no effect against Taspase1 (Chen *et al.* 2012; Zheng *et al.* 2015).

The first rational approach for specific Taspase1 inhibitors focused on the design of substrate-derived inhibitors (Chen *et al.* 2010). The researchers were limited to *in vitro* studies, that used cleavage assays for Taspase1 activity. Vinyl sulfones showed the best, but still modest effects with an IC_{50} of 29.4 µM (Chen *et al.* 2010). Later, van den Boom *et al.* resumed this strategy and succeeded in advancing the substrate-derived protease inhibitors (van den Boom *et al.* 2014). Their class of inhibitors focused on a peptidyl succinimidyl peptide motif and took into account the cleavage mechanism of Taspase1, which had been proposed to hydrolyze its substrate via a succinimide-hydrate intermediate (van den Boom *et al.* 2014). The best compound showed an IC_{50} of 3.6 µM in a Taspase1 activity assay based on fluorescence

resonance energy transfer (FRET), resulting in the first Taspase1 inhibitor that could be utilized at physiological concentrations (van den Boom *et al.* 2014).

In 2012, Chen *et al.* proposed a small molecule Taspase1 inhibitor, as a screening of the NCI Diversity set library with a biosensor for intracellular Taspase1 activity enabled the description of biarsonic compound NSC48300 as a Taspase1 inhibitor (Chen *et al.* 2012). The inhibitor showed decent effects in a FRET-based cleavage assay and a K_I of 4.22 µM was determined (Chen *et al.* 2012). Toxicity of the compound in different tumor cell lines correlated with the Taspase1 expression levels (Chen *et al.* 2012). At higher concentrations, the compound displayed nonspecific toxicity in cellular experiments, which the authors attribute to the arsenic acid or possible interference with other proteins (Chen *et al.* 2012). Of note, the compound was originally described as an anti-angiogenesis inhibitor targeting Autotaxin, a protein involved in tumor invasiveness, and was found to be 100-fold more effective inhibiting this protein, resulting in a shutdown of the lipophosphatic acid production (Saunders *et al.* 2008). Determination of the compound's toxicity in the mouse model revealed a decrease in LDL, white blood cells and hemoglobin (Chen *et al.* 2012). Despite these side effects, a xenograft model with breast and brain cancer cell lines showed a decreased tumor growth after compound treatment (Chen *et al.* 2012). However, this compound became the controversial focus of further discussions, since neither the effects on Taspase1 activity in a cellular environment could be reproduced by other groups, nor did the compound inhibit recombinantly purified Taspase1 in *in vitro* assays (Wünsch *et al.* 2012; van Boom 2015).

There have also been genetic approaches to inhibit Taspase1 by enforced dimerization of active and inactive Taspase1, based on the assumption that Taspase1 has to maintain the dimeric state after autoproteolysis to be enzymatically active (Bier *et al.* 2012a). Overexpression of inactive Taspase1 should therefore lead to decreased activity caused by dimerization of inactive and wild type (Bier *et al.* 2012a). However, the intramolecular activation process was not investigated at the time of the study, so it was not known that the permanently inactive Taspase1 could still induce autoproteolytic activation of the wild type Taspase1 upon dimerization (Bier *et al.* 2012a). While the enforced dimerization itself did not yield an inhibitory approach for Tasapse1, error tracking led to the valuable observation that assembly of two active Taspase1 monomers is not a necessity for enzymatic activity after autoproteolytic processing (Bier *et al.* 2012a).

Later genetic approaches utilized this knowledge to generate a dominantly-negative Taspase1 double mutant (Sabiani *et al.* 2015). The first mutation rendered the Taspase1 monomer itself incapable of autoproteolysis (S291A), while the second mutation located in the docking head (C163E) prevented the conformational changes in the docking zone and thus the resulting autoproteolysis of the opposed Taspase1 monomer upon dimerization (Sabiani *et al.* 2015).

Co-IP revealed decreased autoproteolytic processing of Taspase1 in the presence of the double mutant as well as decreased cleavage of an intracellular biosensor when ectopically expressed and compared with wild type Taspase1 (Sabiani *et al.* 2015).

Another inhibition study focused on the adsorption of Taspase1 by amorphous silica nanoparticles, resulting in the formation of a nanoparticle protein corona (van den Boom *et al.* 2020). Microscale thermophoresis and fluorescence anisotropy titration revealed EC_{50} in the nano- to picomolar range, depending on the diameter of the nanoparticles (8-125 nm), as did a FRET-based cleavage assay for Taspase1 activity (van den Boom *et al.* 2020). While adsorption by the nanoparticles is not specific for Taspase1 *per se*, it was shown that neither lactate dehydrogenase nor chymotrypsin nor proteinase K were inhibited with similar effectiveness (van den Boom *et al.* 2020). Also, while the effect on Taspase1 activity was decreased in HeLa cell lysates, the nanoparticles still inhibited Taspase1 activity by 50 % in comparison to 90 % with recombinantly purified Taspase1 (van Boom 2015; van den Boom *et al.* 2020). Finally, significant effects could also be demonstrated in HeLa cells transfected with a biosensor for intracellular Taspase1 activity, that localized in the nucleus upon Taspase1 cleavage (Knauer *et al.* 2011; Bier *et al.* 2011b; van den Boom *et al.* 2020).

The most recent Taspase1 inhibitor to be discovered was found by screening of FDA-approved drugs, and was hypothized to interfere with Taspase1 dimerization (Luciano *et al.* 2021). Closantel sodium, usually employed as a anthelmintic in veterinary medicine, was found to be a allosteric inhibitor of Taspase1 (Luciano *et al.* 2021). A FRET-based cleavage assay for Taspase1 activity revealed an IC50 of 1.6 µM for *in vitro* generated Taspase1 and an IC50 of 3.9 µM for recombinantly purified Taspase1 from *E. coli*. (Luciano *et al.* 2021). The authors attribute this difference to partial autoproteolysis during the purification and conclude that the inhibitor interferes with the dimerization process, thereby targeting the proenzymatic Taspase1 (Luciano *et al.* 2021). Indeed, they show that autoproteolysis of freshly synthesized *in vitro* generated Taspase1 is decreased by Closantel sodium in a concentration dependent manner (Luciano *et al.* 2021).

1.5 Aim of this study

Taspase1 is a non-oncogene addiction protease overexpressed in a range of liquid and solid tumor cell lines and has been associated with the maintenance of an oncogenic state (Hsieh *et al.* 2003; Bier *et al.* 2011b; Niizuma *et al.* 2015; Wünsch *et al.* 2016). Numerous tumor cell lines are sensitive to an inhibition of Taspase1 as opposed to healthy adult tissue, where a knock-down of Taspase1 is well-tolerated (Takeda *et al.* 2006; Chen *et al.* 2012). Targeted inhibition of Taspase1 by rationally designed ligands could lead to a novel therapeutic approach without the side effects of chemo- or radiotherapy.

For the development of novel inhibitors, three potential areas on the Taspase1 surface can be targeted to interfere with its function: (i) Targeting of the active site to prevent substrate processing by Taspase1, (ii) targeting of the dimerization interface to interfere with autoproteolysis needed for intramolecular activation, and (iii) targeting of the Taspase1 loop containing the NLS, resulting in a disruption of the Taspase1/Importinα-interaction (Tasp/Impα-interaction) and cytoplasmic containment of proenzymatic Taspase1.

Targeting of the active site and the dimerization interface by inhibitors has already been achieved by rational approaches or screening (Lee *et al.* 2009; Chen *et al.* 2012; van den Boom *et al.* 2014; Luciano *et al.* 2021). However, there have been no studies that target the interaction between Taspase1 and Impα. Since alteration of the NLS lead interferes with the nuclear import and activation of Taspase1 in cellular experiments, the same results should be achievable by blocking the NLS with supramolecular ligands (Bier *et al.* 2011a; van den Boom *et al.* 2016).

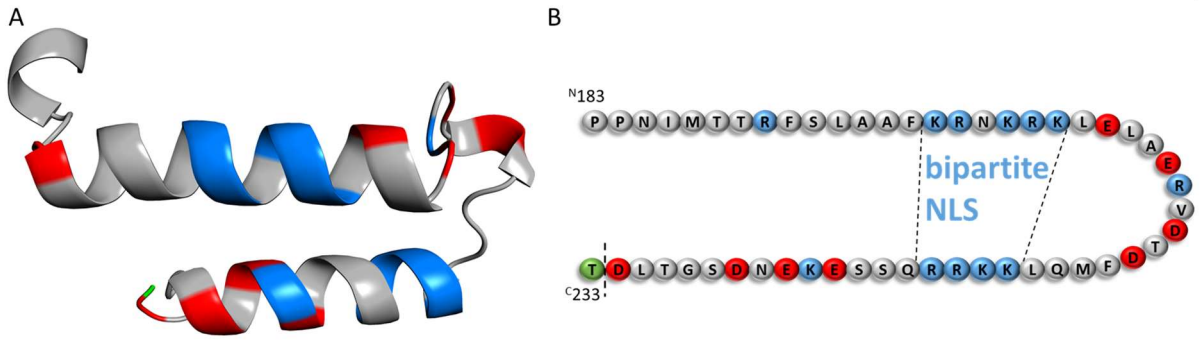


Figure 9: Model of the Taspase1 loop (A) and the corresponding amino acid sequence (B). The loop directly preceding the active site T234 (green) contains clusters of basic (blue) and acetic (red) amino acids that could be targeted by supramolecular ligands. Based on van den Boom *et al.* (van den Boom *et al.* 2016).

By definition, the inhibition strategies mentioned above can each only target parts of the intracellular Taspase1 population, leading to ongoing debates about what constitutes a potent Taspase1 inhibitor (van den Boom *et al.* 2014; Luciano *et al.* 2021). Interference with substrate binding will not disrupt the activation process, so there is a continuous replenishment. Impeding the dimerization process leading to autoproteolysis or the interaction with Impα targets the activation process, but already activated Taspase1 will still be functional until it is degraded. However, the Taspase1 loop (G178-D233) containing the NLS ends prior to the later active site T234 (van den Boom *et al.* 2016). Targeting this area could result in the first dual-action inhibitor for Taspase1, disrupting the activation process of proenzymatic Taspase1 via the interaction with Impα, while blocking substrate binding in the already activated Taspase1.

The overall aim of this study is therefore to use of rationally designed supramolecular ligands as inhibitors, which modulate the Taspase1 function either by direct inhibition of the enzymatic activity or by disruption of the Tasp/Impα-interaction, thereby interfering with the activation

process. Articles I and II focus on strategies involving cationic binders, while articles III and IV focus on anionic binders.

2 Publications

2.1 Publication I

2.1.1 Author contributions

Title:

PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction

Authors:

Peter Pasch[‡], Alexander Höing[‡], Serap Ueclue, Matthias Killa, Jens Voskuhl, Shirley K. Knauer* and Laura Hartmann*

[‡]authors contributed equally *corresponding authors

Contributions:

Conception: 15 %

Experimental work: 33 %

(Cloning, protein purification, pull-down assay, SDS-PAGE, western blotting, ITC, cell culture, MTS assay)

Data analysis: 50 %

(See experimental work)

Writing the manuscript: 50 %

(50 % of the first draft; Figures 3, S1, S25, S26, S27, S28, S29)

Revising the manuscript: 25 %

(Multiple rounds of refinement)

PhD student	Principal investigator

2.1.2 Article introduction and summary

The following paragraph is based on the article “PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction” (Pasch *et al.* 2021).

Until this study, the use of supramolecular ligands for the disruption of the Tasp/Impα-interaction was only hypothetical. There were no references in the literature and most assays only allowed qualitative investigation of the Tasp/Impα-interaction.

Our research was therefore guided by the following questions: (i) Can the interaction between Taspase1 and Impα be targeted by supramolecular ligands in general? (ii) Can the anionic hotspots surrounding the Taspase1 loop be targeted by cationic binders to disrupt the interaction with Impα? (iii) Can multivalency be used to increase the binding affinity of the ligands? (iv) Can a steric shield be utilized to increase the disruptive effect of the ligand by steric inhibition?

To investigate this, we developed trivalent cationic ligands equipped with a guanidiniocarbonyl-pyrrole (GCP) binding motif, a non-natural arginine mimetic. We used Surface plasmon resonance (SPR) to investigate binding to Taspase1 and a novel pull-down setup utilizing recombinant proteins to investigate effects on the interaction with Impα.

Indeed, we managed to disrupt the interaction between Taspase1 and Impα utilizing these ligands, resulting in the first described inhibitors of the Tasp/Impα-interaction, **3GP** and **3GLP**, answering question (i). Modelling suggested binding to anionic hotspots surrounding the Taspase1 loop (ii). While binding of monovalent control ligands to Taspase1 could not be observed, all trivalent ligands bound to Taspase1 (iii). At the same time, while binding to Taspase1 could also be observed by PEG-less control ligands, an effective disruption of the Tasp/Impα-interaction in the micromolar range could only be achieved by ligands equipped with PEG as a steric shield (iv). While the compounds were limited to *in vitro* studies, they served as a valuable reference for further studies.

2.1.3 PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction

ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2021, **57**, 3091

Received 28th October 2020,
Accepted 11th February 2021

DOI: 10.1039/d0cc07139k

rsc.li/chemcomm

PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction†

Peter Pasch,‡[a] Alexander Höing,‡[b] Serap Ueclue,[a] Matthias Killa,[c] Jens Voskuhl,[c] Shirley K. Knauer*[b] and Laura Hartmann*[a]

A novel strategy to inhibit the oncologically relevant protease Taspase1 is explored by developing PEGylated macromolecular ligands presenting the supramolecular binding motif guanidiniocarbonylpyrrole (GCP). Taspase1 requires interaction of its nuclear localization signal (NLS) with import receptor Importin α. We show the synthesis and effective interference of PEGylated multivalent macromolecular ligands with Taspase1–Importin α-complex formation.

Proteins are an important class of biomacromolecules and their interactions play key roles in almost every process of a living organism. Understanding and manipulating protein interactions offers the opportunity to treat or fight diseases.[1] Many protein–protein interactions rely on so-called multivalent binding events where multiple sites of the proteins have to interact simultaneously in order to create a strong binding.[2,3] Accordingly, synthetic molecules to interfere with protein binding often are multivalent constructs as well, consisting of a synthetic scaffold presenting multiple binding units. The design of multivalent molecules is as diverse as the protein targets they address – one important class of scaffolds being polymers due to their synthetic ease and variability.[3,4]

Today, polymer chemistry offers a new tool: the synthesis of sequence-controlled macromolecules. Different synthetic strategies have been introduced for the synthesis of sequence-controlled polymers and give access to multifunctional macromolecules with high levels of structural and thereby potentially also functional control.[5] We have developed the so-called solid phase polymer synthesis, where we employ standard peptide chemistry and tailor-made non-natural building blocks to generate sequence-defined macromolecules presenting different binding units such as carbohydrates, peptides or catechols. We have successfully demonstrated that through control over the monomer sequence and thereby parameters such as the number and position of binding units, architecture and conformation of the macromolecule, new and improved modulators of protein interactions are accessible.[6]

In this work, we extend on our previous concept with a non-natural supramolecular binding motif, the guanidiniocarbonylpyrrole (GCP) motif,[7] an arginine mimetic that binds oxoanions *via* a hydrogen-bond-assisted ion pairing[8] and shows significantly stronger affinity than natural amino acids.[9] Our goal is to create macromolecular inhibitors of Taspase1 protease (Fig. 1). From proliferation to differentiation right up to apoptosis, almost every cellular process is regulated by or involves proteases.[10,11] One of the 28 threonine proteases encoded in the human genome is the tumor-relevant Threonine aspartase 1 (Taspase1).[10,12] It is usually expressed during embryonic development, but it is re-expressed in many tumor cell lines, and a knockout decreases proliferation and promotes apoptosis in correlation to their potential drug target in tumor therapy. Surprisingly, earlier studies revealed that Taspase1 is not affected by general former Taspase1 expression levels[13,14] making Tasapse1 a protease inhibitor.[12,14,15] Previous studies focused on the enzymatic activity of Taspase1.[13,15–17] In this study, we aim at a different inhibition mechanism for Taspase1 by targeting functionally relevant interactions with the import receptor Importin α.[18,19] While Taspase1 effectively cleaves other pro-enzymes as a heterodimer consisting of the subunits α (25 kDa) and β (20 kDa), Taspase1 itself is also expressed as an inactive α/β-monomer (45 kDa) and undergoes autoproteolytic activation.[12,20,21] Autoproteolysis is supposed to take place inside the nucleus where Taspase1 is transported by interaction of its bipartite nuclear localization signal (NLS) located in the Taspase1 α-subunit (Fig. 1B) with Importin α.[18,19,21] In the nucleus, the Taspase1 monomer undergoes

[a] *Department for Organic Chemistry and Macromolecular Chemistry, Heinrich Heine University Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany. E-mail: laura.hartmann@hhu.de*
[b] *Department for Molecular Biology II, Center of Medical Biotechnology (ZMB) University Duisburg-Essen, Universitätsstrasse 5, Essen, 45117, Germany. E-mail: shirley.knauer@uni-due.de*
[c] *Faculty of chemistry (Organic Chemistry), University of Duisburg-Essen, Universitätsstraße 7, 45141 Essen, Germany*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0cc07139k

‡ These authors contributed equally.

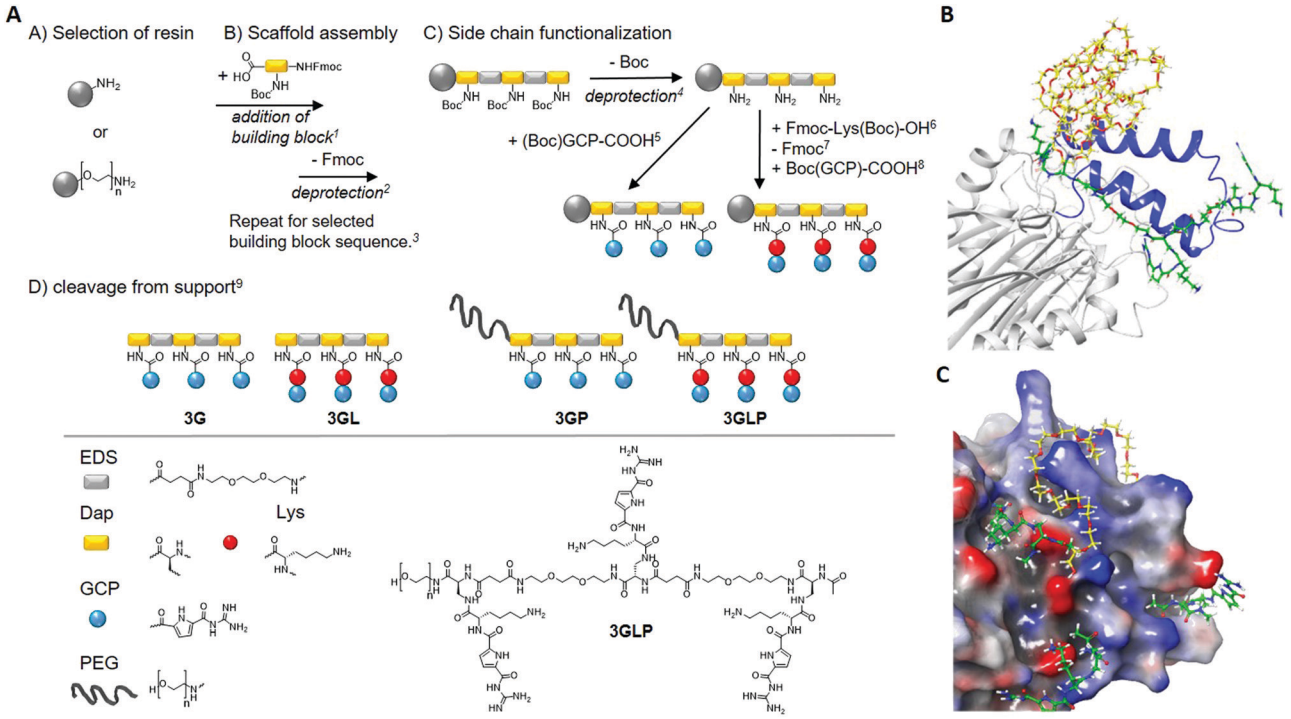


Fig. 1 Synthesis of GCP macromolecules using solid phase polymer synthesis (left) and a model of a supramolecular ligand blocking the NLS of Taspase1 (right). (A) Reaction conditions: [1]5 eq. building block, 5 eq. PyBOP, 10 eq. DIPEA in DMF, 90 min, [2]25v% piperidine in DMF, 20 min, [3]Next Fmoc-deprotection and acetylation of N-terminus (Ac_2O, 20 min), [4]4 M HCl in dioxane, 20 min (on-resin cleavage of Boc), [5]5 eq. (Boc)GCP-COOH, 5 eq. PyBOP, 10 eq. DIPEA in DMF, 90 min (double coupling), [6]5 eq. Fmoc-Lys(Boc)-OH, 5 eq. PyBOP, 10 eq. DIPEA in DMF, 90 min (double coupling), [7]25v% piperidine in DMF, 20 min, [8]5 eq. (Boc)GCP-COOH, 5 eq. PyBOP, 10 eq. DIPEA in DMF, 90 min (double coupling), [9]TentaGel® S RAM: 5% triisopropylsilane, 95% TFA, 90 min, TentaGel® PAP: TFA/thioanisole (95 : 5), 24 hours. PEG chain n = 70 (MW = 3106.7 g mol^{-1}). (B) Model of a supramolecular ligand blocking the NLS of Taspase1. (C) Schematic illustration of **3GLP** (green) as it addresses carboxylates near the loop (blue) while the PEG (yellow) masks the cationic loop. (bottom): the carboxylates (red) are addressed with GCP (green) and the cationic loop (blue) is masked by the PEG chain (yellow). The model is based on the literature.[28]

autoproteolysis, and the two subunits reassemble to form the active Taspase1 heterodimer.[12,21] Thus, interaction of the Taspase1 NLS with Importin α is pivotal for activation. Here, we aim at developing ligands that effectively block the NLS and thereby inhibit Importin α-complex formation as the first step of Taspase1 activation. Since basic amino acid clusters constituting the bipartite NLS of Taspase1 are flanked by multiple anionic amino acids such as aspartic and glutamic acids (see Fig. 1C), we envision that macromolecules presenting multiple oxo-anion binding motifs, GCP, should allow for binding to this site of the protein. It was previously shown that multivalent GCP ligands allow for the design of high affinity ligands by addressing multiple binding sites within a protein structure and can be used for stabilization of protein–protein complexes.[22] Here we now want to realize both high affinity binding to the NLS and at the same time effective inhibition of binding to Importin α. We rationalize that in order to achieve both, we require two features of the macromolecular ligand – one segment presenting multiple GCP motifs able to address anionic amino acids in the NLS domain but that do not mediate binding with Importin α, and a second segment, ideally non-binding and sterically demanding to shield the NLS domain from any further interaction. This design is thus based on the general concept of sterical shielding for multivalent ligands to achieve inhibition.[3,23]

For the first segment, we employ the previously established synthesis of sequence-controlled macromolecules (Fig. 1A).[24,25] *Via* stepwise addition on a solid support, a monodisperse, sequence-controlled scaffold is assembled and used for site selective attachment of GCP motifs.[26] Here we used a previously developed EDS building block (4-((2-(2-(2-aminoethoxy)ethoxy)ethyl)-amino)-4-oxobutanoic) introducing hydrophilic ethylene glycol units within the backbone.[25] Fmoc-Dap(Boc)-OH (N_α-Fmoc-N_β-Boc-L-2,3-diaminopropionic acid) was applied for attachment of GCP on the side chains: Dap side chains were deprotected on the solid support cleaving the Boc protecting groups and releasing primary amines for further functionalization with carboxylated GCP-derivative (see the ESI†). In order to further increase the affinity of GCP towards anionic amino acids, lysine as a cationic amino acid was added next to the GCP side chain by including an additional Fmoc-Lys(Boc)-OH during side chain assembly. Two different macromolecules were synthesized introducing three GCP side chains (**3G**) as well as Lys-GCP side chains (**3GL**). Model calculations suggest that multiple amino acids could be addressed *via* trivalent GCP macromolecules with one EDS as a spacer in between the binding motifs (Fig. 1). As our second segment in order to create GCP macromolecule inhibitors, we chose poly(ethylene glycol) (PEG) of 3 kDa that can be easily installed by starting the solid phase assembly from a PEG-preloaded resin giving PEGylated GCP

macromolecules **3GP** and **3GLP**. PEG is well known as a so-called stealth polymer to minimize non-specific interaction with proteins and to act as a steric shield blocking protein–protein interactions.[27] All macromolecules were cleaved off the resin, purified by preparative HPLC, isolated by freeze drying with relative purities >95% (as determined by RP-HPLC) and further characterized by ^{1}H-NMR, UHRMS and MALDI-ToF analysis (see the ESI†).

First, we looked at the direct binding of our ligands to Taspase1 by successfully setting up a surface plasmon resonance (SPR) assay (see the ESI†). Applying isothermal calorimetry was not successful at this time (see the ESI†). Monovalent macromolecules presenting only one GCP unit showed no binding and thus were omitted from any further testing (see the ESI†). For the trivalent macromolecules, binding is in the μM range as was expected based on previous GCP ligands (Fig. 2).[29]

We observe a clear increase in binding upon introduction of the lysine residues next to the GCP unit, indicating an increase in affinity through the additional cationic moieties. Surprisingly, for **3G** we see an increase in binding upon introduction of the PEG block (**3GP**) which might be attributed to the higher molecular weight of this ligand and slower diffusion, as PEG itself showed no binding (see the ESI†). However, we did not see such increase for **3GLP**.

We next performed an *in vitro* pull-down assay to investigate the proposed inhibitory effect on the interaction between Taspase1 and Importin α. For this, we used recombinant GST-Importin α protein bound to a GSH matrix and added Taspase1-His pre-incubated with the respective ligands (see the ESI†). Unbound protein was removed and the Tasapse1 bound to the matrix *via* its interaction with Importin α eluted. The samples were then analyzed by SDS-PAGE and Western Blotting. To validate our working hypothesis that only the PEGylated compounds will disrupt the protein interaction, we first compared the ligands (**3GP**, **3GLP**) and the controls (PEG alone, the non-PEGylated ligands **3G** and **3GL**) directly (Fig. 3). Indeed, the interaction between Taspase1 and Importin α was effectively disrupted by the pre-incubation of Taspase1 with the PEGylated GCP-ligands. Interestingly, ligands missing the PEG stealth block failed to interfere with Taspase1–Importin α complex formation

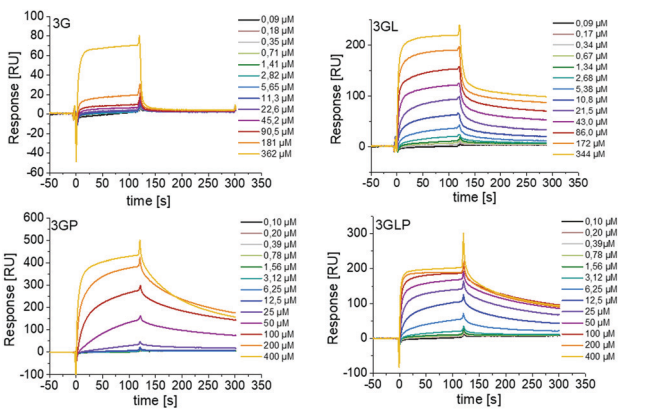


Fig. 2 SPR sensograms showing direct binding of trivalent GCP-ligands with and without PEG to immobilized Taspase1. Monovalent GCP-ligands showed no binding (see the ESI†).

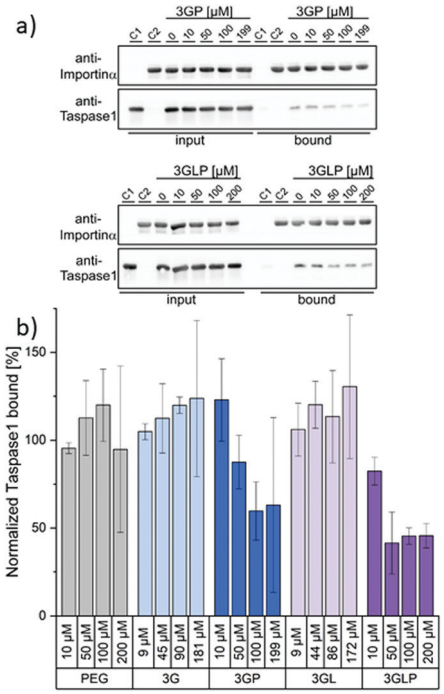


Fig. 3 The interaction between Taspase1 and Importin α is effectively disrupted by PEGylated GCP-ligands: (a) Exemplary blots from pull-down assays with increasing concentrations of the PEGylated GCP-ligands **3GP** and **3GLP** (see the ESI† for other samples). Column-bound protein fraction after the pull-down in comparison to the input initially present on the column. Controls: only Taspase 1 (C1) or GST-Importin α (C2) were added to the column. (b) Densiometric quantification of pull-down assays, comprising the mean of three replicates ± standard deviation.

and so did PEG itself. This suggests that the GCP-motif guides the ligand to Taspase1, but is not able to shield the NLS directly, in line with our model (Fig. 1). Furthermore, the PEG block itself does not affect the interaction and therefore does not ensnare the NLS while unguided. Thus, the introduction of the stealth block PEG to the guiding GCP-block is a necessary step for the ligands to act as inhibitors. Densiometric quantification of Western Blot analysis revealed a slightly increased effect of the lysine-containing ligand **3GLP** (40% Taspase bound) compared to **3GP** (55% Taspase1 bound) (see the ESI†). To further compare the PEGylated ligands that differ in their binding motifs (GCP and GCP plus lysine) **3GP** and **3GLP** were tested at different ligand concentrations ranging from 0 μM to 200 μM in the pull-down assay and the results were again quantified (Fig. 3). **3GP** effectively hampered the interaction already at 50 μM (87% Taspase1 bound). Increasing the concentration to more than 100 μM (59% Taspase1 bound) was not able to additionally fortify its effect. As seen in the direct binding study *via* SPR, addition of lysine next to the GCP motifs increased the apparent affinity. We hypothesize that this, when combined with the PEG segment, should also give more efficient inhibitors. Indeed, **3GLP** was effective already at 10 μM (82% Taspase1 bound), and its potency reached a limit at a concentration of 50 μM (41% Taspase1 bound). Importantly, we did not observe a comparable effect when using the PEG control or the non-PEGylated ligands

(see the ESI†), nor was the binding of Importin α to the column affected by the PEGylated ligand (see the ESI†).

In conclusion, we explored the concept of multivalent ligands for sterical shielding and developed structure-guided PEGylated sequence-controlled macromolecules for Taspase1 using GCP as a binding motif. We further showed that these ligands could effectively be used to disrupt the functionally relevant interaction with Importin α in a concentration-dependent manner, thereby exploiting a novel inhibition mechanism for this protease. Future studies will include investigations concerning the potential selectivity of the ligands as well as their potential for cellular studies. As a first prerequisite, the ligands were tested in a cell viability assay and showed no toxicity (see the ESI†).

The authors acknowledge the CRC 1093 funded by the ''Deutsche Forschungsgemeinschaft''

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 H. Lu, Q. Zhou, J. He, Z. Jiang, C. Peng, R. Tong and J. Shi, *Signal Transduction Targeted Ther.*, 2020, **5**, 213.
2 (*a*) E. Mahon and M. Barboiu, *Org. Biomol. Chem.*, 2015, **13**, 10590–10599; (*b*) C. T. Varner, T. Rosen, J. T. Martin and R. S. Kane, *Biomacromolecules*, 2015, **16**(1), 43–55.
3 M. Mammen, S.-K. Choi and G. M. Whitesides, *Angew. Chem., Int. Ed.*, 1998, **37**, 2754–2794.
4 (*a*) C. Fasting, C. A. Schalley, M. Weber, O. Seitz, S. Hecht, B. Koksch, J. Dernedde, C. Graf, E.-W. Knapp and R. Haag, *Angew. Chem., Int. Ed.*, 2012, **51**, 10472–10498; (*b*) C. R. Becer, *Macromol. Rapid Commun.*, 2012, **33**, 742–752.
5 (*a*) S. C. Solleder, R. V. Schneider, K. S. Wetzel, A. C. Boukis and M. A. R. Meier, *Macromol. Rapid Commun.*, 2017, **38**, 1600711; (*b*) S. A. Hill, C. Gerke and L. Hartmann, *Chem. – Asian J.*, 2018, **13**, 3611; (*c*) J. F. Lutz, M. Ouchi, D. R. Liu and M. Sawamoto, *Science*, 2013, **341**(6146), 1238149.
6 (*a*) L. Soria-Martinez, S. Bauer, M. Giesler, S. Schelhaas, J. Materlik, K. Janus, P. Pierzyna, M. Becker, N. L. Snyder, L. Hartmann and M. Schelhaas, *J. Am. Chem. Soc.*, 2020, **142**(11), 5252–5265; (*b*) S. Boden, F. Reise, J. Kania, T. K. Lindhorst and L. Hartmann, *Macromol. Biosci.*, 2019, **19**, 1800425; (*c*) L. Fischer, R. C. Steffens, T. J. Paul and L. Hartmann, *Polym. Chem.*, 2020, **11**, 6091–6096; (*d*) T. Freichel, V. Heine, D. Laaf, E. E. Mackintosh, S. Sarafova, L. Elling, N. L. Snyder and L. Hartmann, *Macromol. Biosci.*, 2020, **20**, 2000163.
7 (*a*) C. Schmuck and M. J. Schwegmann, *J. Am. Chem. Soc.*, 2005, **127**, 3373–3379; (*b*) D. Maity, A. Gigante, P. A. Sánchez-Murcia, E. Sijbesma, M. Li, D. Bier, S. Mosel, S. Knauer, C. Ottmann and C. Schmuck, *Org. Biomol. Chem.*, 2019, **17**, 4359–4363.
8 (*a*) C. Schmuck, *Coord. Chem. Rev.*, 2006, **250**, 3053–3067; (*b*) J. Matić, F. Šupljika, T. Tandarić, M. Dukši, P. Piotrowski, R. Vianello, A. Brozovic, I. Piantanida, C. Schmuck and M. R. Stojković, *Int. J. Biol. Macromol.*, 2019, **134**, 422–434.
9 (*a*) C. Schmuck and M. Heil, *ChemBioChem*, 2003, **4**(11), 1232–1238; (*b*) X. Liu, K. Wang, M. Externbrink, J. Niemeyer, M. Giese and X. Y. Hu, *Chin. Chem. Lett.*, 2020, **31**(5), 1239–1242.
10 C. López-Otín and L. M. Matrisian, *Nat. Rev. Cancer*, 2007, **7**, 800–808.
11 C. López-Otín and J. S. J. Bond, *Biol. Chem.*, 2008, **283**, 30433–30437.
12 J. J.-D. Hsieh, E. H.-Y. Cheng and S. J. Korsmeyer, *Cell*, 2003, **115**, 293–303.
13 D. Y. Chen, Y. Lee, B. A. van Tine, A. C. Searleman, T. D. Westergard, H. Liu, H.-C. Tu, S. Takeda, Y. Dong and D. R. Piwnica-Worms, *Cancer Res.*, 2012, **72**, 736–746.
14 D. Y. Chen, H. Liu, S. Takeda, H.-C. Tu, S. Sasagawa, B. A. van Tine, D. Lu, E. H.-Y. Cheng and J. J.-D. Hsieh, *Cancer Res.*, 2010, **70**, 5358–5367.
15 S. K. Knauer, V. Fetz, J. Rabenstein, S. Friedl, B. Hofmann, S. Sabiani, E. Schröder, L. Kunst, E. Proschak and E. Thines, *PLoS One*, 2011, **6**, 18253.
16 J. T. Lee, D. Y. Chen, Z. Yang, A. D. Ramos, J. J.-D. Hsieh and M. Bogyo, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 5086–5090.
17 J. van den Boom, M. Mamić, D. Baccelliere, S. Zweerink, F. Kaschani, S. Knauer, P. Bayer and M. Kaiser, *ChemBioChem*, 2014, **15**, 2233–2237.
18 Y. Miyamoto, K. Yamada and Y. Yoneda, *J. Biochem.*, 2016, **160**, 2728–2735.
19 C. Bier, S. K. Knauer, D. Docter, G. Schneider, O. H. Kramer and R. H. Stauber, *Traffic*, 2011, **12**, 703–714.
20 D. Wünsch, A. Hahlbrock, S. Jung, T. Schirmeister, J. van den Boom, O. Schilling, S. K. Knauer and R. H. Stauber, *Oncogene*, 2016, **35**, 3351–3364.
21 A. Corbett, K. Mills, B. A. Lange, M. Stewart, E. Devine and J. C. J. Lange, *Biol. Chem.*, 2007, **282**, 5101–5105.
22 L. Bartsch, M. Bartel, A. Gigante, J. Iglesias-Fernández, Y. B. Ruiz-Blanco, C. Beuck, J. Briels, N. Toetsch, P. Bayer, E. Sanchez-Garcia, C. Ottmann and C. Schmuck, *ChemBioChem*, 2019, **20**(23), 2921–2926.
23 F. Jacobi, D. Wilms, T. Seiler, T. Queckbörner, M. Tabatabai, L. Hartmann and S. Schmidt, *Biomacromolecules*, 2020, **21**(12), 4850–4856.
24 (*a*) F. Wojcik, S. Mosca and L. Hartmann, *J. Org. Chem.*, 2012, **77**(9), 4226–4234; (*b*) D. Ponader, F. Wojcik, F. Beceren-Braun, J. Dernedde and L. Hartmann, *Biomacromolecules*, 2012, **13**(6), 1845–1852.
25 C. Gerke, M. F. Ebbesen, D. Jansen, S. Boden, T. Freichel and L. Hartmann, *Biomacromolecules*, 2017, **18**, 787–796.
26 C. Schmuck, V. Bickert, M. Merschky, L. Geiger, D. Rupprecht, J. Dudaczek, P. Wich, T. Rehm and U. Machon, *Eur. J. Org. Chem.*, 2008, 324–329.
27 (*a*) S. Schöttler, G. Becker, S. Winzen, T. Steinbach, K. Mohr, K. Landfester, V. Mailänder and F. R. Wurm, *Nat. Nanotechnol.*, 2016, **11**, 372–377; (*b*) S. De Santis, R. Chiaraluce, V. Consalvi, F. Novelli, M. Petrosino, P. Punzi, F. Sciubba, C. Giordano, G. Masci and A. Scipioni, *ChemPlusChem*, 2017, **82**, 241–250; (*c*) S. Akocak, M. R. Alam, A. M. Shabana, R. Kishore, K. Sanku, D. Vullo, H. Thompson, E. R. Swenson, C. T. Supuran and M. A. J. Ilies, *Med. Chem.*, 2016, **59**, 5077–5088; (*d*) K. Chitphet, S. M. Geary, C. H. F. Chan, A. L. Simons, G. J. Weiner and A. K. Salem, *Biomater. Sci. Eng.*, 2020, **6**(5), 2659–2667.
28 J. van den Boom, F. Trusch, L. Hoppstock, C. Beuck and P. Bayer, *PLoS One*, 2016, **11**(3), 1–13.
29 C. Schmuck and V. Bickert, *J. Org. Chem.*, 2007, **72**(18), 6832–6839.

2.2 Publication II

2.2.1 Author contributions

Title:

A Bivalent Supramolecular GCP Ligand Enables Blocking of the Taspase1/Importin α Interaction

Authors:

Alexander Höing[‡], Alexander Zimmermann[‡], Lisa Moews, Matthias Killa, Marius Heimann, Astrid Hensel, Jens Voskuhl*, Shirley K. Knauer*

[‡]authors contributed equally *corresponding authors

Contributions:

Conception: 10 %

Experimental work: 45 %

(Protein purification, pull-down assay, SDS-PAGE, western blotting, cell culture, MTS assay)

Data analysis: 45 %

(See experimental work)

Writing the manuscript: 50 %

(Half of the first draft; Figures 1, 3, 5, 6, S26, S27, S28, S32)

Revising the manuscript: 25 %

(Multiple rounds of refinement)

PhD student Principal investigator

2.2.2 Article introduction and summary

The following excerpt is based on the article “A Bivalent Supramolecular GCP Ligand Enables Blocking of the Taspase1/Importin α Interaction” (Höing *et al.* 2021).

With the first inhibitors of the Tasp/Impα-interaction described by Pasch *et al.* we set a reference (Pasch *et al.* 2021). Utilization of the GCP motif to address anionic hotspots surrounding the Taspase1 NLS had proven successful, but due to the use of an oligomeric backbone and the steric shield, the ligands were rather large and could not be used for further cellular studies (Pasch *et al.* 2021).

The next study focused on improving the inhibition strategy utilizing the cationic binding motif GCP. We addressed the following questions: (i) Can we generate a compound that can be taken up by cells and which shows effects in a cellular environment? (ii) Can clusters of anionic amino acids in the loop directly flanking the NLS be utilized for a more disruptive effect? (iii) Can a more effective disruptor of the Tasp/Impα-interaction target also Taspase1 activity directly? (iv) Which components of the ligands are important for effectiveness?

We generated the cationic bivalent GCP-based inhibitor **2GC**. Modelling suggested binding to anionic amino acids in the Taspase1 loop between both parts of the NLS. In pull-down assays, the compound disrupted the Tasp/Impα-interaction and was more effective than previous compounds, displaying a concentration dependency with an IC_{50} of 35 µM, which answers question (ii).(Pasch *et al.* 2021) The compounds had toxic effects on tumor cells as revealed by an MTS assay. Depending on the cell line, the EC_{50} values determined ranged from 30-70 µM (i).

While the compound also affected Taspase1 activity in a lysate-based cleavage assay, unphysiologically high concentrations of **2GC** were needed, which was therefore not a potent inhibitor of Taspase1 activity (iii). Using various control compounds, each with single components changed. The GCP motif, bivalency and a hydrophobic protection group were determined to be key factors (iv). Any modification of those components abolished the ligands effect in all assays used.

2.2.3 A Bivalent Supramolecular GCP Ligand Enables Blocking of the Taspase1/Importin α Interaction

A Bivalent Supramolecular GCP Ligand Enables Blocking of the Taspase1/Importin α Interaction

Alexander Höing[+],[a] Alexander Zimmermann[+],[b] Lisa Moews,[a] Matthias Killa,[b] Marius Heimann,[b] Astrid Hensel,[a] Jens Voskuhl,*[b] and Shirley K. Knauer*[a]

In memory of Prof. Dr. Carsten Schmuck (1968–2019)

Taspase1 is a unique protease not only pivotal for embryonic development but also implicated in leukemia as well as solid tumors. As such, it is a promising target in cancer therapy, although only a limited number of Taspase1 inhibitors lacking general applicability are currently available. Here we present a bivalent guanidiniocarbonyl-pyrrole (GCP)-containing supramolecular ligand that is capable of disrupting the essential interaction between Taspase1 and its cognate import receptor Importin α in a concentration-dependent manner *in vitro* with an IC_{50} of 35 μM. Here, size of the bivalent *vs* the monovalent construct as well as its derivation with an aromatic cbz-group arose as critical determinants for efficient interference of **2GC**. This was also evident when we investigated the effects in different tumor cell lines, resulting in comparable EC_{50} values (~40–70 μM). Of note, in higher concentrations, **2GC** also interfered with Taspase1's proteolytic activity. We thus believe to set the stage for a novel class of Taspase1 inhibitors targeting a pivotal protein-protein interaction prerequisite for its cancer-associated proteolytic function.

Due to its vastly diverse nature, cancer remains one of the most challenging diseases in the history of humankind. In 2018, 18.1 million people were diagnosed with cancer and it was the cause of death for 9.6 million.[1] While classical treatments involves surgery, radiotherapy and/or chemotherapy, the last becomes progressively limited due to the emergence of resistances.[2–3] Therefore, the development of new therapeutic approaches and novel anti-cancer drugs still remains an imperative task. There are many different proteins that are promising targets in anti-cancer therapies.[4–6] One of these is the protease Taspase1 (Threonine aspartase 1), a protein normally involved in embryonic developmental processes.[7–8] It is widely absent in adult, differentiated tissues, but re-expressed in many tumor cell lines. Although Taspase1 alone is not sufficient to transform cells, tumors become increasingly dependent on its presence.[9] Taspase1 is therefore classified as a "non-oncogene addiction protease". It was initially reported as the protease responsible for cleavage of the Mixed Lineage Leukemia protein (MLL), and its oncogenic fusion proteins.[7] Subsequently, more and more oncologically relevant proteins were identified as Taspase1 targets, including e.g. TFIIA (Transcription factor IIA) and USF2 (Upstream stimulatory factor 2), and the unconventional myosin Myo1F.[10–11] As a Taspase1 knock-out is moreover well tolerated in normal adult tissue, it is regarded as an immensely attractive drug target.[12]

In the last decades, several approaches have been presented or proposed to interfere with its enzymatic activity. Relevant strategies comprised substrate analogues, nanoparticles, as well as enforced dimerization of its two subunits.[13–15]

Nevertheless, none of those inhibitors has yet reached the clinics. Although Taspase1, together with the proteasome, belongs to the rather small class of threonine proteases, its catalytic activity is neither affected by common protease inhibitors nor by proteasome inhibitors.[7, 16]

As already indicated, Taspase1 is a very unique protease belonging to the type 2 asparaginase family of enzymes.[7] All members of this family share the ability to be autocatalytically processed in *cis*, but Taspase1 is the only family member that functions as a protease, cleaving other substrates by recognizing a conserved peptide motif with an aspartate at the P_1 position.[7,17–18] Referring to its rather complex activation process (Figure 1), Taspase1 is initially expressed as an inactive α/β-monomer (50 kDa).[7,17] Autoproteolysis into the two subunits α (28 kDa) and β (22 kDa) results in a proteolytically active heterodimer subsequently enabling cleavage of target proteins in *trans*.

Of note, mutation of the catalytic nucleophile, Thr234, not only results in loss of *cis*-activity and thus precludes formation of the two subunits, but also completely abolishes Taspase1's proteolytic function in *trans*.[7]

[a] *A. Höing,[+] L. Moews, Dr. A. Hensel, Prof. S. K. Knauer*
Institute for Molecular Biology II
Center for Medical Biotechnology (ZMB)
University of Duisburg-Essen
Universitätsstrasse 5, 45117 Essen (Germany)
E-mail: shirley.knauer@uni-due.de
[b] *A. Zimmermann,[+] M. Killa, M. Heimann, Prof. J. Voskuhl*
Faculty of Chemistry (Organic Chemistry) and CENIDE
University of Duisburg Essen
Universitätsstrasse 7, 45141 Essen (Germany)
E-mail: jens.voskuhl@uni-due.de
[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cmdc.202100640

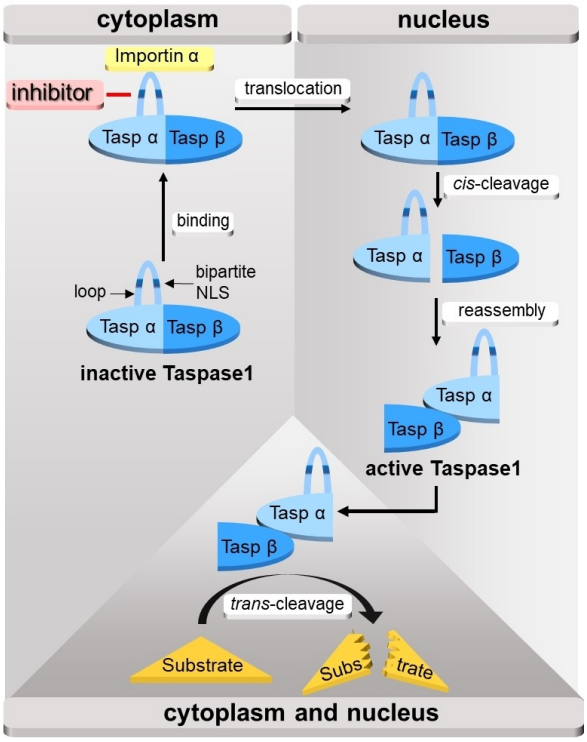


Figure 1. Cellular activation of Taspase1. The inactive Taspase1 proenzyme is synthesized in the cytoplasm, where it interacts with Importin α and is translocated into the nucleus. Here, Taspase1 is autoproteolytically processed into the α- and the β-subunit, re-assembling into a heterodimer enabling cleavage of cellular substrates in *trans*. Inhibition of this pivotal interaction (red line) should thus indirectly interfere with Taspase1's proteolytic function.

Moreover, the N-terminal part of the protein representing the α-subunit of Taspase1 contains a flexible loop region (aa178–233), consisting of two alpha helices.[19] These helices harbor a bipartite nuclear localization signal (NLS). As such, the latter comprises two basic amino acid clusters (^{197}KRNKRK^{202}, ^{217}KKRR^{220}) located in close proximity on the neighboring helices and thus constitute one single surface-accessible basic cluster.[17, 20]

This can be recognized by the adaptor protein Importin α, which might additionally recruit the carrier protein Importin β or solely transport Taspase1 into the nucleus.[20–21] Importantly, effective autoproteolysis into the two subunits *in vivo* requires a functionally intact NLS to efficiently interact with Importin α.[17] The Taspase1/Importin α interaction is thus regarded as an essential prerequisite to ensure full proteolytic activation.

Therefore, we aimed to develop cell-permeable molecules which target the respective protein binding interface using a structure-guided approach. Ligand design was based on the Schmuck binding motif guanidiniocarbonyl-pyrrole (GCP), generally suited for a wide range of applications in biomedical research.[22] It is used as protein recognition and modulation element but also serves as a pivotal component in transfection vectors[23–26] and as a building block in supramolecular polymers, gels and nanostructures.[22,27] Indeed, due to its function as a synthetic, and in comparison to its natural analogue arginine physiologically stable and thus superior recognition unit for oxo-anions,[24] the Schmuck binding motif is an ideal moiety to address protein surfaces in general, and in particular the rather flexible and completely surface-exposed loop region of Taspase1.[7,19,28–29] Here, a polycationic motif was chosen to primarily address negatively charged amino acids such as aspartic acid and glutamic acid present in this region. The Taspase1 loop adopts a helix-turn-helix conformation. Here, the amino acid sequence constituting the turn element indeed is the most exposed and accessible part of the loop. This turn region is rich in negatively charged amino acids, as well as a second surface-exposed stretch of negatively charged aspartic acid and glutamic acid in direct vicinity of the second helix.

To target both regions simultaneously, we decided to place two GCP units in a tandem arrangement to be tested in comprehensive biological assays (Figure 2). The compounds were synthesized by SPPS (Solid Phase Peptide Synthesis). The bivalent **2G** derivatives were obtained by dimerizing the corresponding monovalent **1G** derivatives with a 1,8-diaminooctane spacer connecting the (unprotected) lysins at first position of the G derivatives *via* two amide bonds (see Supporting Information for details, Figure S1, Table S1). Moreover, the protecting group of the second lysine was varied during SPPS to deduct potential effects of small structural changes. By introducing protecting groups like alloc (**A**) and cbz (**C**), the affinity of the structures with respect to hydrophobic amino acids such as valine or phenylalanine should be increased. The resulting ligands **2GA** and **2GC** might thus reveal an enhanced disruptive potency.

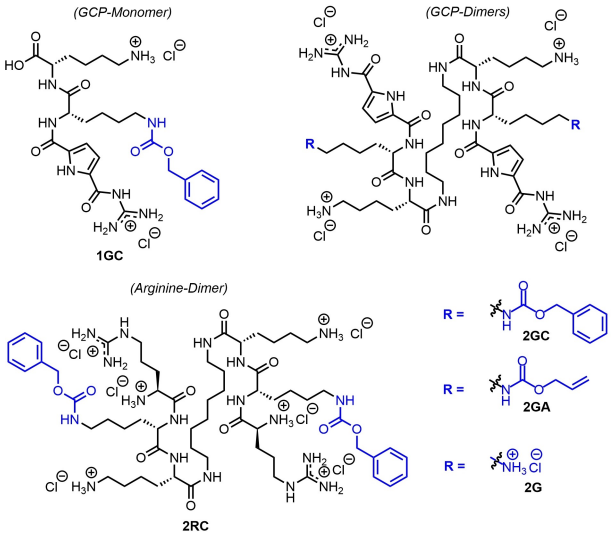


Figure 2. All compounds used in this study were synthesized by SPPS. **G** (GCP-containing-binding-unit) consist of two lysines and one GCP at the N-terminus (see Supporting Information for details, Figure S1–S22, Table S1). **G** was derivatized at the second lysine with a **C**(bz) or **A**(lloc) protecting group to generate the precursors for **2GC** and **2GA**. The bivalent compounds contain **2G** in a symmetric/palindromic arrangement with the sequence GCP-K(protected)-K-Spacer-K-K(protected)-GCP, harbouring the two unprotected lysines in the center. **2RC** represents a non-GCP containing bivalent control analogously equipped with two arginines (**R**) with two **C**(bz) protective groups.

First, we analysed the effects of our bivalent compounds on the Taspase1/Importin α interaction, utilizing a customized *in vitro* pull-down assay. Here, Importin α was recombinantly expressed with a N-terminal glutathione S-transferase (GST) affinity tag and immobilized on a glutathione sepharose-column to test for retention and thus binding of recombinant Taspase1-His protein subsequently applied to the column (see Supporting Information for details, Figure S26). Since wildtype Taspase1 partially undergoes autoproteolysis during protein purification, we used an inactive Taspase1 mutant unable to cleave in *cis* or *trans* (D233A/T234A).

Hereby, we assured homogeneity of the protein population required for an optimal reproducibility and robustness of our *in vitro* assay. Preceding pre-incubation of Taspase1 with either 100 μM of each compound or increasing concentrations of compound **2GC** (up to 100 μM) allowed to test the compound's inhibitory potential. Proteins were subsequently analysed by SDS-PAGE and immunoblotting with specific antibodies. Interestingly, **2GC** most efficiently interfered with the Taspase1/Importin α interaction (Figure 3A). Quantification of the pull-down data revealed an approx. 3-fold reduction; only 37% Taspase1 bound to column-immobilised Importin α compared to the untreated control. In contrast, the underivatized compound **2G** as well as the alloc derivative **2GA** showed no effect on the interaction, although both compounds differ from **2GC** only in the protecting group. Of note, routinely performed immunoblots of the unbound fractions revealed that column binding of Importin α was not affected by the ligands (Supporting Information Figure S27). Even more importantly, we synthesized the dimeric control compound **2RC**, bearing arginine (R) residues as generic cationic groups instead of the GCP moieties (see Supporting Information for details, Figure S2). In contrast to **2GC**, the arginine control **2RC** was not able to efficiently inhibit the interaction with Importin α (Supporting Information Figure S28).

Next, we performed molecular docking studies to shed more light on the binding mechanism of **2GC** (Figure 4, see Supporting Information for details, Figure S23–S25, Table S2). In contrast to a plethora of potential interactions of **2GC** and Taspase1 *via* hydrogen bonds, the hydrophobic cbz protecting group obviously does not take part in hydrogen bonding or electrostatic interactions with the protein but might be involved in hydrophobic interactions. However, **2GC** is suggested to

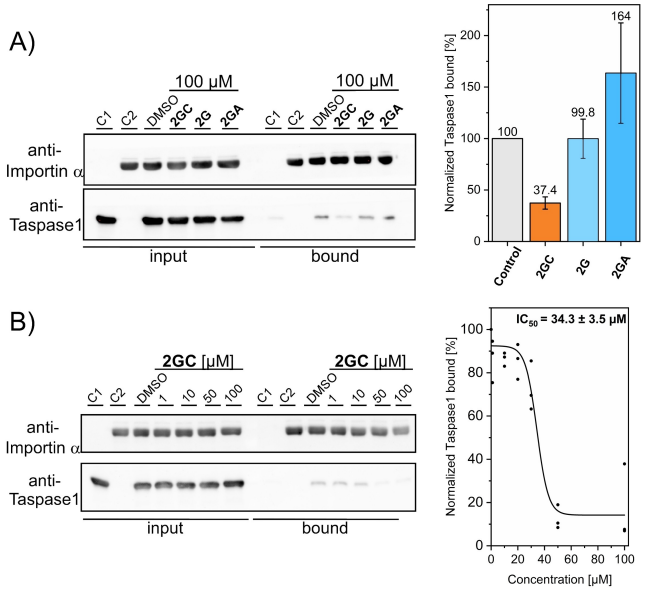


Figure 3. **2GC** effectively disrupts the interaction between Taspase1 and Importin α in a concentration-dependent manner. A) A specific pull-down setup allows to directly compare the effect of Taspase1 pre-incubation with 100 μM compound (**2GC**, **2G**, **2GA**). Controls include either only Taspase1 (C1), GST-Importin α (C2) or DMSO treatment. Quantification of results comprises the mean of three replicates ± standard deviation. B) By utilizing different concentrations of **2GC**, the pull-down assay reveals an IC_{50} of 34 ± 3.5 μM. Controls include only Taspase1 (C1) or GST-Importin α (C2). The results are the mean of three replicates ± standard deviation.

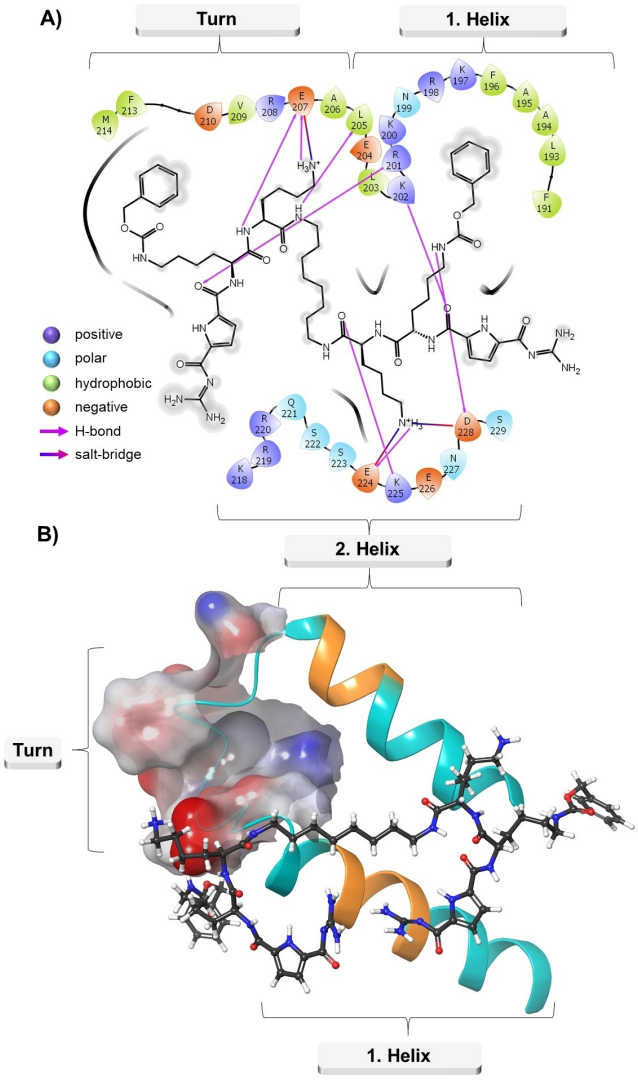


Figure 4. Modell of the interaction between **2GC** and the Taspase1 loop. A) LID of **2GC** and the loop sequence 189–229 aa. B) 3D model of **2GC** interacting with the loop. The latter consists of a helix(189–205 aa)-turn(206–215 aa)-helix(216–229 aa) motif, where basic amino acid clusters in both helices constitute the bipartite NLS (orange) that interacts with Importin α. The surface of the turn is rich in negative charged amino acids (red).

cover a large portion of the Taspase1 loop necessary for the interaction with the import receptor and thus might convey efficient steric shielding (Figure 4, Supporting Information Figure S25A/B).

We assume that the cbz protecting group might reveal an additional repulsive effect on Importin α, thus contributing to the observed inhibition of its interaction with Taspase1. Of note, although the docking scores of the GCP-containing compounds **2GC**, **2GA** and **1GC** were in the same order of magnitude, the score of **2RC** was less negative (Supporting Information Table S2), indicative for a decreased stabilizing energy. Comparative docking of **2RC** and **2GC** however revealed that similar areas were populated by the compounds, although **2GC** binds tighter and in a more closed conformation (Supporting Information Figure S24). This could hind towards a beneficial effect of the GCP unit compared to arginine at the same position.

We were subsequently focusing on **2GC** as the so far most effective compound. However, an exact quantification of pull-down experiments is not trivial and rather allows to determine an order of magnitude instead of discrete binding parameters. However, by rationally adapting the concentration range we acquired sufficient data points for a robust fit and could finally determine an IC_{50} of 34 ± 3.5 μM for its disruptive effect observed in our pull-down setup (Figure 3B). To further underscore our bivalent design concept, we also compared **2GC** to its monovalent counterpart **1GC**. As hypothesized, the prominent effect of **2GC** (quantification revealed 23% Taspase1 bound) could not be retained using the equally derivatized monovalent building block **1GC** (79% Taspase1 bound as revealed by quantification) in the pull-down assay (Figure 5). This strongly indicates that the molecular surface size is indeed important to mediate efficient interference with the interaction between Taspase1 and Importin α. Interestingly, molecular docking studies demonstrated that **1GC** is also able to interact with different Taspase1 amino acid residues (Supporting Information Figure S25D). However, **1GC** is not supposed to interact with the bipartite NLS inside the loop that is necessary for the interaction with Importin α, explaining its impaired potency.

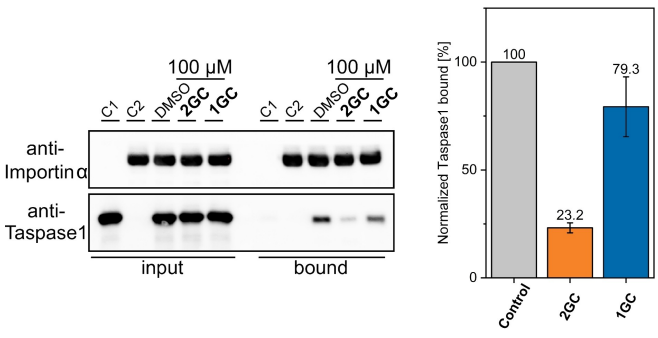


Figure 5. Only the bivalent but not the monovalent compound allows to efficiently interfere with the Taspase1/Importin α interaction. In our pull-down setup, pre-incubation Taspase1 with **2GC** hampers binding to column-bound Importin α in contrast to **1GC**. Controls include either only Taspase1 (C1), GST-Importin α (C2) or DMSO treatment (DMSO). Quantification of results comprises the mean of three replicates ± standard deviation.

The 3D model also indicates that **1GC** might not be able to shield an area of sufficient dimension to efficiently interfere with the Taspase1/Importin α interaction (Supporting Information Figure S25C).

As the residues targeted by **2GC** are in close proximity to Taspase1's active site, we further aimed to analyse whether **2GC** also affects its proteolytic activity. Therefore, we newly established a robust, *semi-in vitro* Taspase1 substrate cleavage assay (see Supporting Information for details, Figure S29). Here, we used 293T cells which express only neglectable amounts of endogenous Taspase1 to only rely on the activity of defined amounts of recombinant, fully active Taspase1-His to the cell lysates. As a confirmed Taspase1 substrate, we decided for the transcription factor USF2 (Upstream stimulatory factor 2),[10] which was overexpressed in cell culture. Respective cell lysates were incubated with recombinant Taspase1-His in the absence or presence of 500 μM of each compound for 4 h and 6 h. Indeed, immunoblot analysis revealed an inhibitory effect of **2GC** on Taspase1-mediated USF2 cleavage, which could not be evidenced for compounds **1GC**, **2GA**, **2G** (Supporting Information Figure S30A) or **2RC** (Supporting Information Figure S31), irrespective of the incubation time. Next, we stepwise decreased the concentration of **2GC** from 500 μM down to 100 μM. However, **2GC** was only effective up to 400 μM (Supporting Information Figure S30B). As in this assay the use of cell lysates supersedes the necessity of nuclear translocation as a prerequisite for Taspase1 activation, the observed effect could not be attributed to the compound's ability to interfere with Importin α binding. Moreover, the high concentration of the compound needed to affect Taspase1's proteolytic activity rather indicates the occupation of a neighbouring, least favoured region in the flexible loop. However, we next indeed investigated the effect of our compounds *in vivo* using two different Taspase1-expressing tumor cell lines, namely the cervical carcinoma cell line HeLa and the lung cancer cell line A549. Cells were incubated with different compound concentrations for 24 h, and toxicity was determined utilizing a colorimetric MTS assay for the quantification of viable cells. Indeed, **2GC** decreased the viability of HeLa ($EC_{50} = 69.9 \pm 1.8$ μM) and A549 ($EC_{50} = 40.9 \pm 8.2$ μM) cells (Figure 6A, B, Supporting Information Figure S32A). In contrast, the compounds **2GA**, **2G**, **1GC** and **2RC** had a rather negligible *in vivo* effect, even when applied in concentrations of 100 μM or even above (Figure 6, Supporting Information Figure S32B–E). Although this is still no airtight proof that Taspase1 is indeed responsible for the observed effect, these results are congruent with those achieved in the pull-down assays (Figure 3A, B, Figure 5).

In sum, our results demonstrate the feasibility of targeting the Taspase1-Importin α interaction with symmetry-based GCP-containing ligands. Ligand docking simulations by molecular force field calculations indicate that **2GC** might act as a symmetric clamp grasping the Taspase1 loop at its turning point and thus shield the NLS by steric hindrance. This results in an effective disruption of the Taspase1/Importin α interaction substantiated by *in vitro* pull-down assays.

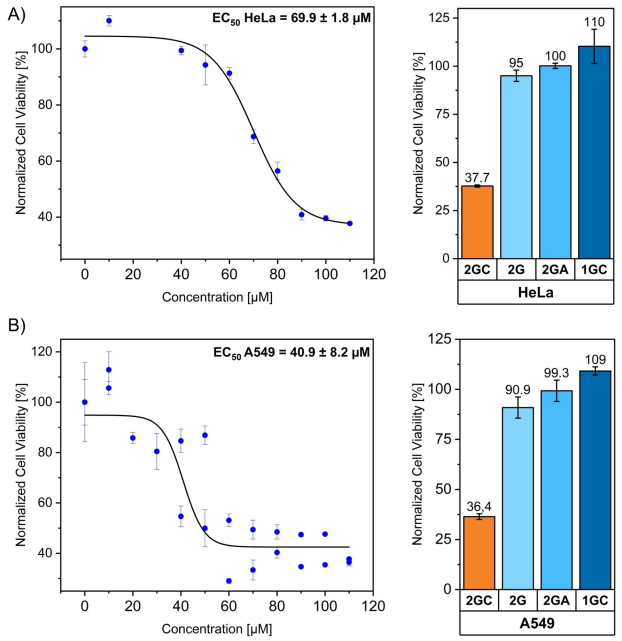


Figure 6. 2GC affects the viability of tumor cells. A) Determination of the EC_{50} value of **2GC** in HeLa cells ($EC_{50} = 69.9 \pm 1.8$ µM). In contrast to **2GC**, **2G**, **2GA** and **1GC** do not impair tumor cell viability even at the maximal concentration of 110 µM. Each data point is the mean of three replicates ± standard deviation. B) Determination of the EC_{50} value of **2GC** in A549 cells ($EC_{50} = 40.9 \pm 8.2$ µM). In contrast to **2GC**, **2G**, **2GA** and **1GC** do not impair tumor cell viability even at the maximal concentration of 110 µM. To allow a more robust curve fitting, results of two experiments were combined. Each data point is the mean of three replicates ± standard deviation.

Moreover, a *semi*-in vitro Taspase1 substrate cleavage assay also showed that the enzyme activity of Taspase1 is affected by **2GC**, albeit in rather high micromolar concentrations. Finally, we could demonstrate an anti-proliferative effect of **2GC** in Taspase1-expressing tumor cell lines.

However, further studies are now required to deeply investigate the binding kinetics and the mechanism underlying the effect on cell viability.

In conclusion, we developed a bivalent supramolecular GCP ligand that effectively targets the interaction between Taspase1 and Importin α, which is essential for its proteolytic activation. This now sets the stage for the development of a novel class of inhibitors targeting this therapeutically relevant protease.

Acknowledgements

This work was supported by the DFG collaborative research centre1093 (CRC 1093 – Projects A1, A10 and B5). Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: Taspase1 · Importin α · protein-interaction · nuclear localization signal (NLS) · protease · substrate cleavage assay · SPPS · supramolecular inhibition

[1] World Health Organization: Regional Office for Europe, *World Cancer Report: Cancer Research for Cancer Prevention*, IARC, **2020**.
[2] F. McCormick, *Nat. Rev. Cancer.* **2001**, *1*, 130–141.
[3] N. Vasan, J. Baselga, D. M. Hyman, *Nature* **2019**, *575*, 299–309.
[4] J. L. Delgado, C.-M. Hsieh, N.-L. Chan, H. Hiasa, *Biochem. J.* **2018**, *475*, 373–398.
[5] Q. Li, W. Xu, *Curr. Med. Chem. Anti-Cancer Agents* **2005**, *5*, 53–63.
[6] B. Kumar, S. Singh, I. Skvortsova, V. Kumar, *Curr. Med. Chem.* **2017**, *24*, 4729–4752.
[7] J. J.-D. Hsieh, E. H.-Y. Cheng, S. J. Korsmeyer, *Cell* **2003**, *115*, 293–303.
[8] D. Y. Chen, Y. Lee, B. A. van Tine, A. C. Searleman, T. D. Westergard, H. Liu, H.-C. Tu, S. Takeda, Y. Dong, D. R. Piwnica-Worms, K. J. Oh, S. J. Korsmeyer, A. Hermone, R. Gussio, R. H. Shoemaker, E. H.-Y. Cheng, J. J.-D. Hsieh, *Cancer Res.* **2012**, *72*, 736–746.
[9] D. Y. Chen, H. Liu, S. Takeda, H.-C. Tu, S. Sasagawa, B. A. van Tine, D. Lu, E. H.-Y. Cheng, J. J.-D. Hsieh, *Cancer Res.* **2010**, *70*, 5358–5367.
[10] C. Bier, S. K. Knauer, A. Klapthor, A. Schweitzer, A. Rekik, O. H. Kramer, R. Marschalek, R. H. Stauber, *J. Biol. Chem.* **2011**, *286*, 3007–3017.
[11] S. Takeda, S. Sasagawa, T. Oyama, A. C. Searleman, T. D. Westergard, E. H. Cheng, J. J. Hsieh, *J. Clin. Invest.* **2015**, *125*, 1203–1214.
[12] S. Takeda, D. Y. Chen, T. D. Westergard, J. K. Fisher, J. A. Rubens, S. Sasagawa, J. T. Kan, S. J. Korsmeyer, E. H.-Y. Cheng, J. J.-D. Hsieh, *Genes Dev.* **2006**, *20*, 2397–2409.
[13] J. van den Boom, A. Hensel, F. Trusch, A. Matena, S. Siemer, D. Guel, D. Docter, A. Hoing, P. Bayer, R. H. Stauber, S. K. Knauer, *Nanoscale* **2020**, *12*, 19093–19103.
[14] J. van den Boom, M. Mamic, D. Baccelliere, S. Zweerink, F. Kaschani, S. Knauer, P. Bayer, M. Kaiser, *ChemBioChem* **2014**, *15*, 2233–2237.
[15] C. Bier, S. K. Knauer, D. Wünsch, L. Kunst, S. Scheiding, M. Kaiser, C. Ottmann, O. H. Krämer, R. H. Stauber, *FASEB J.* **2012**, *26*, 3421–3429.
[16] M. Bochtler, L. Ditzel, M. Groll, C. Hartmann, R. Huber, *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 295–317.
[17] C. Bier, S. K. Knauer, D. Docter, G. Schneider, O. H. Kramer, R. H. Stauber, *Traffic* **2011**, *12*, 703–714.
[18] D. Wünsch, A. Hahlbrock, S. Jung, T. Schirmeister, J. van den Boom, O. Schilling, S. K. Knauer, R. H. Stauber, *Oncogene* **2016**, *35*, 3351–3364.
[19] J. van den Boom, F. Trusch, L. Hoppstock, C. Beuck, P. Bayer, *PLoS One* **2016**, *11*, e0151431, 1–13.
[20] A. Lange, R. E. Mills, C. J. Lange, M. Stewart, S. E. Devine, A. H. Corbett, *J. Biol. Chem.* **2007**, *282*, 5101–5105.
[21] Y. Miyamoto, K. Yamada, Y. Yoneda, *J. Biochem.* **2016**, *160*, 69–75.
[22] M. Giese, J. Niemeyer, J. Voskuhl, *ChemPlusChem* **2020**, *85*, 985–997.
[23] C. Vallet, D. Aschmann, C. Beuck, M. Killa, A. Meiners, M. Mertel, M. Ehlers, P. Bayer, C. Schmuck, M. Giese, S. K. Knauer, *Angew. Chem. Int. Ed.* **2020**, *59*, 5567–5571; *Angew. Chem.* **2020**, *132*, 5614–5619.
[24] L. Bartsch, M. Bartel, A. Gigante, J. Iglesias-Fernández, Y. B. Ruiz-Blanco, C. Beuck, J. Briels, N. Toetsch, P. Bayer, E. Sanchez-Garcia, C. Ottmann, C. Schmuck, *ChemBioChem* **2019**, *20*, 2921–2926.
[25] M. Li, S. Schlesiger, S. K. Knauer, C. Schmuck, *Angew. Chem. Int. Ed.* **2015**, *54*, 2941–2944; *Angew. Chem.* **2015**, *127*, 2984–2987.
[26] D. Maity, A. Gigante, P. A. Sanchez-Murcia, E. Sijbesma, M. Li, D. Bier, S. Mosel, S. Knauer, C. Ottmann, C. Schmuck, *Org. Biomol. Chem.* **2019**, *17*, 4359–4363.
[27] G. Gröger, W. Meyer-Zaika, C. Böttcher, F. Gröhn, C. Ruthard, C. Schmuck, *J. Am. Chem. Soc.* **2011**, *133*, 8961–8971.
[28] C. Schmuck, *Chem. Commun.* **1999**, 843–844.
[29] C. Schmuck, L. Geiger, *Curr. Org. Chem.* **2003**, *7*, 1485–1502.

Manuscript received: October 1, 2021
Accepted manuscript online: October 8, 2021
Version of record online: October 19, 2021

2.3 Publication III

2.3.1 Author contributions

Title:

Dual activity inhibition of Threonine aspartase 1 by a single supramolecular bisphosphate ligand

Authors:

Alexander Höing, Robin Struth, Christine Beuck, Neda Rafieiolhosseini, Daniel Hoffmann, Roland. H. Stauber, Peter Bayer, Jochen Niemeyer*, Shirley K. Knauer*

*corresponding authors

Contributions:

Conception: 35 %

(Started independent cooperation, developed strategies and assays)

Experimental work: 90 %

(Cloning, protein purification, pull-down assay, SDS-PAGE, western blotting, cell culture, MTS assay, fluorescence anisotropy, fluorescence titration, colorimetric cleavage assay, FRET-based cleavage assay, cellular biosensor assay)

Data analysis: 90 %

(See experimental work)

Writing the manuscript: 30 %

(Complete first draft; Figures 1, 3, 4, 6, 7, 8, S1, S5, S6, S7, S8, S10, S11; Tables S2, S3)

PhD student	Principal investigator

2.3.2 Article introduction and summary

Previous studies focus on the use of cationic ligands to inhibit the Tasp/Impα-interaction. A more straightforward strategy would be to directly target the Taspase1 NLS, which consists of two basic clusters. So the next study focuses on the utilization of anionic supramolecular binders to interfere with Taspase1 function.

In this article, we are guided by the questions: (i) Can we directly target the cationic clusters, of which the NLS consists, with anionic supramolecular ligands to inhibit the Tasp/Impα-interaction? (ii) Is this strategy more effective than using cationic binders, as in previous studies? (iii) Can we visualize the intracellular effect of the ligands on Taspase1 activity? (iv) Can we target two key functions of Taspase1 at the same time, resulting in a dual-action inhibitor?

We investigated the BINOL-derived bivalent anionic binder **11d,** whose para-substituted binding domains are based on phosphoric acid (Octa-Smolin *et al.* 2018). Bivalency and para-substitution of the binding units arose as critical determinant's of the compounds effect. Modelling with Epitopsy and SAMC determined that the energetically most favorable binding position was indeed located at the Taspase1 loop.

Binding to Taspase1 and binding to the loop could be confirmed by fluorescence anisotropy, NMR spectroscopy and fluorescence titration, answering our first question (i). Indeed, the compounds inhibitory effect on the Tasp/Impα-interaction surpassed all previous cationic ligands with an IC50 of 6 µM (ii). A novel colorimetric cleavage assay was utilized to assess the ligands effect on proteolytic activity of Taspase1 *in vitro*, revealing an IC50 of 2 µM (iv). Effect on proteolytic activity in the presence of other proteins could be confirmed by a modified version of an established FRET-assay (Hsieh *et al.* 2003). Finally, the effects of **11d** on the proteolytic activity of Taspase1 could be shown in HeLa cells utilizing a biosensor for intracellular Taspase1 activity (iii) (Knauer *et al.* 2011; Bier *et al.* 2011b).

2.3.3 Dual activity inhibition of Threonine aspartase 1 by a single supramolecular bisphosphate ligand

Dual activity inhibition of Threonine aspartase 1 by a single supramolecular bisphosphate ligand

Alexander Höing[a], Robin Struth[b], Christine Beuck[c], Neda Rafieiolhosseini[d], Daniel Hoffmann[d], Roland H. Stauber[e], Peter Bayer[c], Jochen Niemeyer*[b], Shirley K. Knauer*[a]

[a]Molecular Biology II, Center of Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

[b]Organic Chemistry, Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Universitätsstrasse 7, 45141 Essen, Germany

[c]Structural and Medicinal Biochemistry, Center for Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

[d]Bioinformatics and Computational Biophysics, Center for Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

[e]Molecular and Cellular Oncology/ENT, University Medical Center Mainz (UMM), Langenbeckstrasse 1, 55101 Mainz, Germany

Keywords: dual activity inhibitor; therapy resistance; protease; Importin α; supramolecular chemistry; protein-protein-interaction; biosensor

ABSTRACT: Therapy resistance remains an ongoing quest for the clinics and rational drug design. To minimize resistance, the application of dual-active chemicals that simultaneously inhibit independent functions in disease-relevant proteins is a highly promising strategy. The unique protease Threonine aspartase1 is involved in various cancer types but resistant to general protease inhibitors. We hypothesized that addressing basic residues in its bipartite nuclear localization signal (NLS) on a flexible loop by precise supramolecular ligands affects critical steps in protease activity. We report the design, biological mechanism, and therapeutic effects of a potent dual inhibitor. A series of anionic ligands featuring two or three phosphate groups was synthesized and tested for their inhibitory potential. The most active bivalent inhibitor **11d** selectively bound to Threonine aspartase1 with high affinity (K_D=300nM) and disrupted its interaction with Importin α (IC_{50}=6μM). NMR-studies and computational analyses identified a basic NLS cluster (^{220}KKRR^{223}) as the contact region for **11d**, required for Importinα binding, nuclear entry, and function. Biochemical cell-free assays revealed that **11d** additionally inhibited the protease's catalytic substrate *trans*-cleavage activity, most likely by interfering with the loop's molecular flexibility. Notably, functional assays and confocal microscopy comprehensively demonstrated that **11d** inhibited Threonine aspartase1 also in cancer cell models *in vivo.* In summary, we introduce a selective supramolecular bisphosphate ligand allowing dual inhibition of Threonine aspartase1 by targeting simultaneously its protein-protein interaction-based activation and enzymatic function. Our discovery demonstrates that chemical interference with independent key functions in disease-relevant proteins by a single inhibitor can be achieved.

INTRODUCTION

'Hit once better hit twice!' This saying from the clinical bedside is accepted when treating viral (e.g., COVID) or bacterial infections as well as live threatening diseases, such as cancer.[1-3] Hence, combination therapies are practiced to improve treatment success and minimize the complications of therapy resistances.[1-3] Thus, modern chemistry in basic and applied research seeks to expand our treatment repertoire by investigating chemical structures that target multiple independent functions in disease relevant proteins, ideally by a single inhibitor.[2-3] So far, compounds affecting multiple (unrelated) proteins by 'off target' effects or inhibiting protein function by binding to different sites as well as hybrid drugs combining pharmacophores simultaneously targeting various protein functions have been explored.[2, 4-7] The concept of 'two for the price of one' is clearly under intense investigation.[8] For example, it has been shown that inhibiting two different enzymes in inflammatory pathways with a single peptide reduced inflammation.[9] However, proven examples of molecules targeting independent functions in proteins by binding to a single site not only *in vitro* but also *in vivo* are to the best of our knowledge missing to date.

To thus expand the repertoire of such dual-active compounds, a *trans*-approach combining synthetic and analytical chemistry with a profound knowledge of disease-relevant molecular structures and functions is obligatory. Moreover, as shown by others and us, precise supramolecular chemistry now allows to target and selectively inhibit small, functionally pivotal protein domains.[10-12] Based on our previous work[13-20], we here choose the cancer-relevant protease **T**hreonine **asp**art**ase 1** (Taspase 1) as a model for the development of dual active inhibitors for following reasons:

For one, proteases are central for life and protease deregulation is often associated with a variety of diseases.[21-22] As shown by others and us, Taspase 1 is not only critically involved in the regulation of cellular development but also in liquid and solid malignancies.[15, 23-24] Although it belongs to the group of threonine proteases, Taspase 1 is a unique enzyme. In contrast to the exclusively *cis*-active type 2 asparaginases, only Taspase 1 is also able to cleave other substrates in *trans* by hydrolyzing its target proteins at conserved $Q^3[FILV]^2D^1{\downarrow}G^1X^{2'}D^{3'}D^{4'}$ motifs.[18, 25-27] Thus, the discovery of Taspase 1 founded a new class of endopeptidases that utilize the N-terminal threonine (Thr^{234}, Figure 1) of its mature β-subunit as the active site (Figure 1). Besides the proto-oncogene Mixed Lineage Leukemia (MLL),[25] other essential proteins such as the precursor of the general transcription factor IIA (TFIIA)[18, 28-30] or the Upstream Stimulatory Factor 2 (USF2)[18] could be identified as *bona fide* Taspase 1 targets in the protease's degradome.[18] Since Taspase 1 is normally expressed mainly during embryonic development, interference with its activity would not affect healthy adult tissue.[23, 27, 31] In contrast, liquid as well as solid tumor cells have been shown to re-express Taspase 1.[23, 27, 31] This increased dependency on the protease's activity correlates with increased proliferation and reduced apoptosis, as shown by RNAi depletion studies.[24, 31] Therefore, Taspase 1 is classified as a

'non-oncogene addiction' protease.[23-24] However, the full repertoire of physiological or pathological pathways regulated in humans by Taspase 1 still remains to be understood.

Second, although Taspase 1 belongs to the group of threonine proteases, such as the proteasome, it is not affected by general protease inhibitors.[23-24] To date, some more or less specific inhibitors have been suggested, including ubiquitous Cl^- anions.[13, 16, 24-25, 31] However, their mode of action is often not understood, they need to be used in high concentrations, and completely fail to inhibit Taspase 1 in living systems.[13, 16, 24-25, 31]. As no effective *in vivo* Taspase 1 inhibitors are available, this caveat not only hampers the dissection of Taspase 1's disease mechanisms, but also precludes the assessment/translation of its clinical relevance.[9, 14-17] Hence, Taspase 1 not only represents a highly relevant disease target but remains a challenging biochemical model for the design of novel inhibitors.[16, 32]

Third, we discovered that Taspase 1 has to undergo a distinct, multistep activation process to execute its pathobiological cleavage activity (see SI, Figure S1).[19, 23] The protease is expressed as an inactive α/ß-proenzyme (45 kDa) in the cytoplasm and depends on active nuclear transport for its autoproteolytic activation, prerequisite for *trans*-cleavage of its degradome.[19, 23] Active nuclear import is mediated by Taspase 1's bipartite nuclear localization signal (NLS) consisting of two distinct basic amino acid clusters. [19, 23] These are located on two neighboring α-helices arranged on a surface-exposed loop in the α-subunit (Figure 1), mediating binding to the nuclear transport receptor Importin α.[19] Of note, this interacting loop behaves flexible in solution as shown by NMR analyses[33] although some but not all crystal structures indicate more rigid helical structures.[26, 34] Inside the nucleus, autoproteolysis results in the formation of the two subunits α (25 kDa) and ß (20 kDa), which reassemble into the active heterodimer capable of cleaving substrates.[19, 25] Thus, the Taspase 1/Importin α-axis is prerequisite for intracellular protease activation, and besides the catalytically active Thr^{234}, represents an alternative target for chemical intervention strategies inside cells (see SI, Figure S1, and Figure 1).

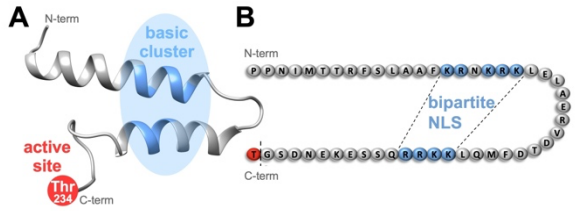


Figure 1: In Taspase 1, the basic clusters (blue) constituting the bipartite NLS are arranged on two neighboring α-helices on a surface-exposed loop, close to the catalytically active threonine (Thr^{234}).[33] Proposed tertiary structure (A) and amino acid sequence (B) including residues 183-233 constituting the loop are shown. NLS-targeting ligands may not only interfere with Importin α-binding but also with the loops' flexibility and the protease's catalytic activity. N-/C-termini are indicated.

Hence, based on our knowledge and the available structural information,[26, 33] we hypothesized that compounds with multiple anionic binding sites may specifically address the surface-exposed bipartite NLS-loop in Taspase 1. Besides inhibition of the critical protein interaction with Importin α, such ligands may restrict the enzyme's molecular motion and thus, additionally inhibit its catalytic activity (Figure 1). Thus, such structures might function as 'dual inhibitors', adressing two independent enzyme functions by binding to a single site.

As the first 'proof of principle' for such a dual active inhibitor, we here report a potent bisphosphate ligand simultaneously acting on two pivotal protein functions.

RESULTS AND DISCUSSION

Synthesis and characterization of NLS-targeting ligands

To rationally address the basic, positively-charged NLS domains, we exploited three anionic phosphate-based ligands **11d**/**e**/**f** (Figure 2). These compounds were previously reported to allow the fluorescence-based detection of cationic amino-sugars, but have not yet been tested for their interaction with cationic protein-residues.[35-37] Among these ligands, **11d**/**e** are regio-isomers featuring two phosphate-units, while ligand **11f** posesses three phosphates of the type $ArOP(O)(OH)_2$. Accordingly, these ligands can possess four (**11d**/**e**) or six negative charges (**11f**), respectively, potentially allowing strong and selective binding to the positively charged NLS-domain. In addition, the extended aromatic backbone of these compounds allows their fluorescence-based detection *in vitro* as well as *in vivo* , with the best fluorescence intensities observed for the *para*-derivative **11d** (blue fluorescence, λ_{em} = 453 nm for λ_{ex} = 380 or 405 nm; see SI, Figure S2).

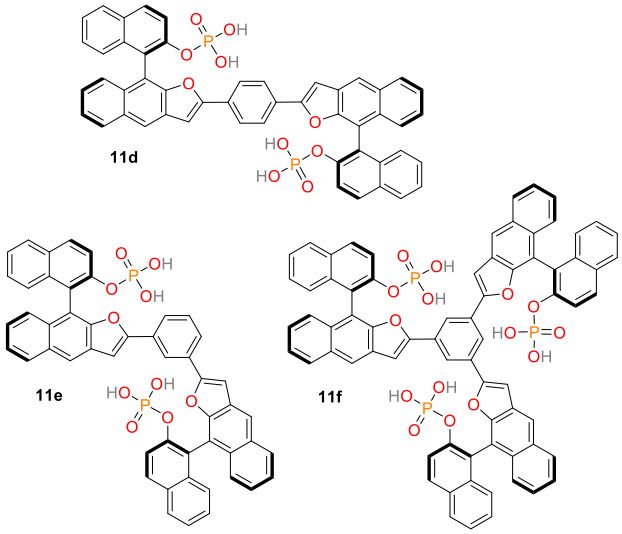


Figure 2: Structures of phosphate-based ligands **11d**/**e**/**f**.

11d is a potent inhibitor of the Taspase 1/Importin α interaction

Next, we used a customized biochemical pull-down assay to analyse the effects of our ligands on the Taspase 1/Importin α interaction (Figure 3, see SI for details).[14, 32] As this interaction crucially relies on the basic NLS clusters present in the Taspase 1loop, the assay serves as a reliable and direct readout to support our ligand binding hypothesis. Briefly, recombinant Importin α is expressed with an N-terminal glutathione S-transferase (GST) affinity tag and immobilized on a glutathione sepharose-column. Subsequently, binding of recombinant Taspase 1-His protein in the presence or absence of ligands allows to quantify the compounds inhibitory potential. Of note, since wildtype Taspase 1 partially undergoes autoproteolysis during protein purification, we here used a proteolytically inactive Taspase 1 mutant ensuring a stable concentration of the proenzyme for robust and reliable detection. As shown in Fig. 4, **11d** was the most potent inhibitor of the Taspase 1/Importin α interaction, reducing the amount of bound Taspase 1 by more than 93% at a concentration of 100 µM (Figure 3). In contrast, compound **11e** was less effective (reduction by 50%), and the trivalent derivate **11f** did not reduce binding. Therefore, we decided to focus on **11d** as the most promising candidate for further characterization.

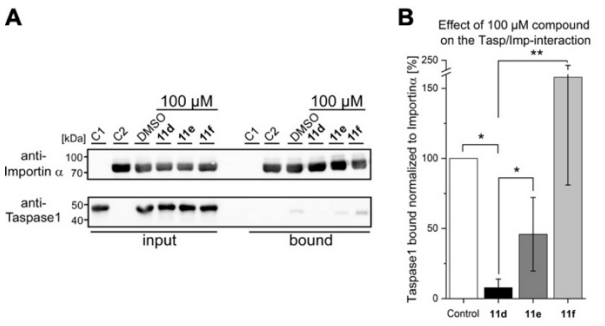


Figure 3: The interaction between Taspase 1 and Importin α is effectively disrupted by **11d**. A) Pull-down assay with 100 µM of each phosphate-based ligand. Controls: Taspase 1-His (C1) or GST-Importin α (C2) alone were added to the column, and a DMSO-treated sample served as reference. B) Densitometric quantification of pull-down assays, comprising the mean of three replicates ± standard deviation (*p < 0.05, **p < 0,01), compared to the DMSO control.

To increase our resolution, we also performed pull-down assays with low micromolar **11d** concentrations (2-8 µM; Figure 4A). Notably, **11d** significantly interfered with the Taspase 1/Importin α interaction already at a concentration of 6 µM, and densitometric quantification of the respective immunoblots resulted in an IC_{50} of 6.10 ± 0.27 µM (Figure 4, see SI, Figure S6). As an important control, the binding of GST-Importin α to the matrix was not affected by the ligand itself (see SI, Figure S5).

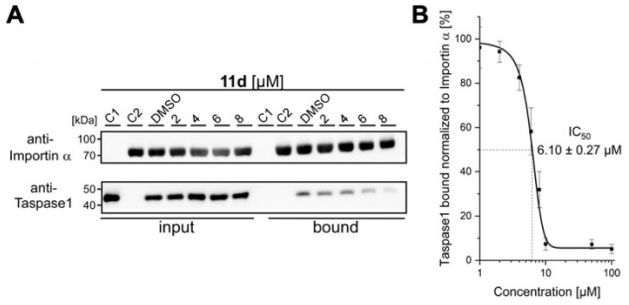


Figure 4: **11d** is a very potent disruptor of the Taspase 1/Importin α interaction. A) Pull-down assay with decreasing concentrations of **11d** between 8 and 2 µM. Controls: Taspase 1-His or GST-Importin α alone were added to the column; a DMSO-treated sample served as reference. B) Densitometric quantification of pull-down assays reveals a half maximal inhibitory concentration of **11d** in the low micromolar range (IC_{50} = 6.10 µM ± 0.27 µM). Quantification comprises the mean of three replicates ± standard deviation (see SI).

11d is a Taspase 1 loop binder

Our assumption that **11d** might specifically recognize the NLS domain was comprehensively supported by independent evidence: First, we employed *in silico* molecular modelling (see SI). Using a recently developed tool for scanning the surface of proteins with fragments of large ligands called *Epitopsy* followed by a *Simulated Annealing Monte Carlo Simulation* (SAMC), we identified the energetically most favorable binding site for **11d** on the surface of Taspase 1, indeed located at the loop contacting the basic NLS residues (Figure 5; see SI, Figures S3 and S4).

Second, to also experimentally prove the direct binding of **11d** to the protease, we used a quantitative fluorescence anisotropy assay. Here, full-length Taspase 1 was titrated to a constant concentration of the fluorescent ligand **11d**, which revealed a strong binding K_D in the nanomolar range (K_D = 300 ± 50 nM) (see SI, Figure S7).

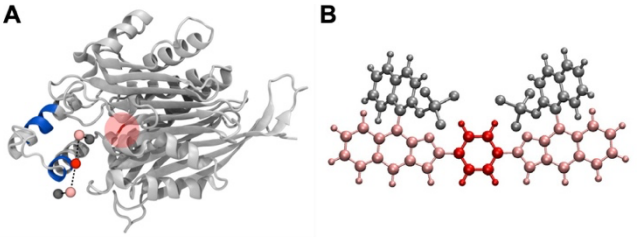


Figure 5: A) SAMC simulations predict the energetically most favorable binding site of **11d** on Taspase 1, with the ligand represented as a bead-spring model. Blue, bipartite NLS; highlighted red tube: active site. B). Ligand **11d** consisting of five groups represented with different colors. In our bead-spring model, each group is represented by one bead.

Third, to further narrow down **11d**'s binding site, we coupled an N-terminal FAM-label to the NLS peptide loop (aa S_{181}-D_{233}) and performed fluorescence titration experiments. A decrease in fluorescence due to compound binding in close proximity to the label was observed (K_D = 3 µM), in contrast to the negative control (FAM-label without peptide; see SI, Figure S8).

Fourth, NMR titration experiments with the isolated Taspase 1 NLS loop and **11d** demonstrated a change in relative signal intensities and thus, additionally verified binding of **11d** to the NLS loop (see SI, Figure S9).

Dual activity: 11d also inhibits the proteolytic activity of Taspase 1

After we convincingly demonstrated that **11d** efficiently disrupts the Taspase 1/Importin α interaction (activity #1), we next tested our hypothesis that binding of **11d** to the loop region also affects the protease's proteolytic activity.

To verify this second effect (activity #2), we employed a novel biochemical cleavage assay using a recombinant fusion protein harboring the MLL CS2 substrate cleavage site (see SI). As illustrated in Figure 6A, addition of recombinant proteolytically active Taspase 1 resulted in substrate cleavage. Importantly, this proteolytic processing was efficiently inhibited by **11d** in a concentration-dependent manner (Figure 6A). **11d** turned out to act as a highly potent inhibitor, already preventing cleavage at low micromolar concentrations as revealed by densitometric quantification ($IC_{50} \approx 2$ µM) (Figure 6B).

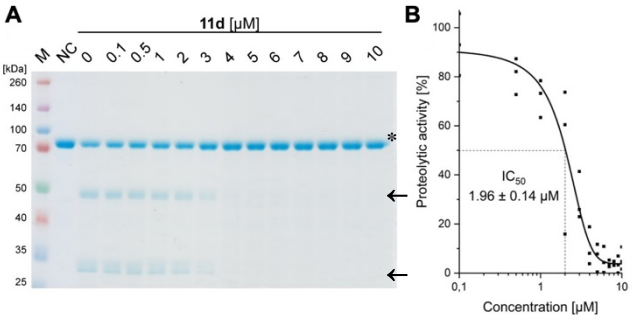


Figure 6: **11d** interferes with Taspase 1's proteolytic activity. A) Cleavage of a recombinant substrate by Taspase 1 was assessed via SDS-PAGE and Coomassie staining after 90 min reaction time in absence and presence of **11d** (0.1 - 10 µM). Inhibition of proteolysis by **11d** was already evident at a concentration between 3 and 4 µM. A DMSO-treated sample served as reference (NC, negative control). B) Densitometric quantification of the uncleaved substrate (asterisk) and its cleaved products (arrows) verifies an IC50 in the low micromolar range (1.96 ± 0.14 µM).

To independently substantiate the dual inhibition mechanism, we additionally established a FRET-based activity assay. Using a MLL CS2 cleavage site peptide coupled to a quencher and a fluorophore, proteolytic cleavage by Taspase 1 lead to an increase in fluorescence (see SI, Figure S10). In contrast, addition of **11d** efficiently inhibited the proteolytic activity of Taspase 1. To next mimic a cellular environment in which other proteins may act as competitors, we added different amounts of cell lysates to our assay. Importantly, **11d** was still able to inhibit Taspase 1's proteolytic activity to 50 % even in the presence of a 30-fold excess of cellular proteins, setting the stage for subsequent assays in living cancer cells.

Dual inhibition *in vivo*: 11d suppresses Taspase 1 function also in a live cellular environment

An absolute requirement for an inhibitor's *in vivo* activity and potential therapeutic relevance in drug development is its access to the relevant target protein inside cells. Notably, most suggested Taspase 1 inhibitors reported to date failed to convincingly take this key hurdle.[16,23,38] Hence, we first controlled the cellular uptake of **11d** utilizing its autofluorescence. Live cell microscopy confirmed that **11d** efficiently entered Taspase 1 expressing cervical carcinoma cells (Figure 7) and was detectable already 1 h after treatment, indicative for its intracellular bioavailability.

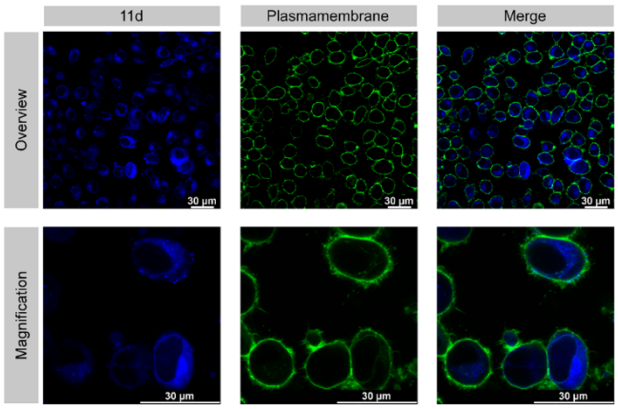


Figure 7: 11d is cell permeable. Cells were incubated with 50 µM **11d** for 1 h, and the plasma membrane was stained with CellBrite® (green). Confocal microscopy detects the blue-fluorescent compound **11d** inside HeLa Kyoto cells. Scale bars, 30 µm.

Next, we tested the inhibitory activity of **11d** in tumor cells using a cell-based *in vivo* biosensor assay. Our biosensor fusion protein ($Tasp_{BS}$) consists of a red fluorescent (mCherry) protein coupled to a nuclear localization sequence (NLS) and a dominant nuclear export signal (NES) separated by the *bona fide* MLL CS2 Taspase 1 cleavage site (Figure 8A). $Tasp_{BS}$ shuttles between the nucleus and the cytoplasm. As the NES is dominant over the NLS, the uncleaved biosensor's steady-state localization is predominantly cytoplasmic (see SI Figure S11). In contrast, cleavage by Taspase 1 results in the loss of the NES triggering its nuclear accumulation (Figure 8A). Experimentally, the red-fluorescent $Tasp_{BS}$ accumulated in the nucleus in Taspase 1 co-expressing HeLa Kyoto cells (Figure 8B, lower panel). In contrast, treatment with **11d** effectively inhibited Taspase 1-mediated biosensor cleavage inside cells (>10-fold; Figure 8C/D) as demonstrated by the $Tasp_{BS}$'s exclusive cytoplasmic localization (Figure 8B/C, upper panel). Results were quantified by randomized image quantification (Figure 6C, and SI, Figure S11 and Table S3), clearly characterizing **11d** as an unique highly potent cell-permeable inhibitor of this protease.

Clearly, future work needs to thoroughly investigate **11d's** effectivity and bioavailability in tumor models with improved complexity. Besides the rational engineering of pharmacophores, delivery by nano-based drug formulations might be explored.[39-40] However, convincing Taspase 1 specific small animal tumor models are not yet available.[23] Thus, we do not wish to postulate that **11d** is 'ready' for clinical translation, which may be considered a potential limitation of our work.

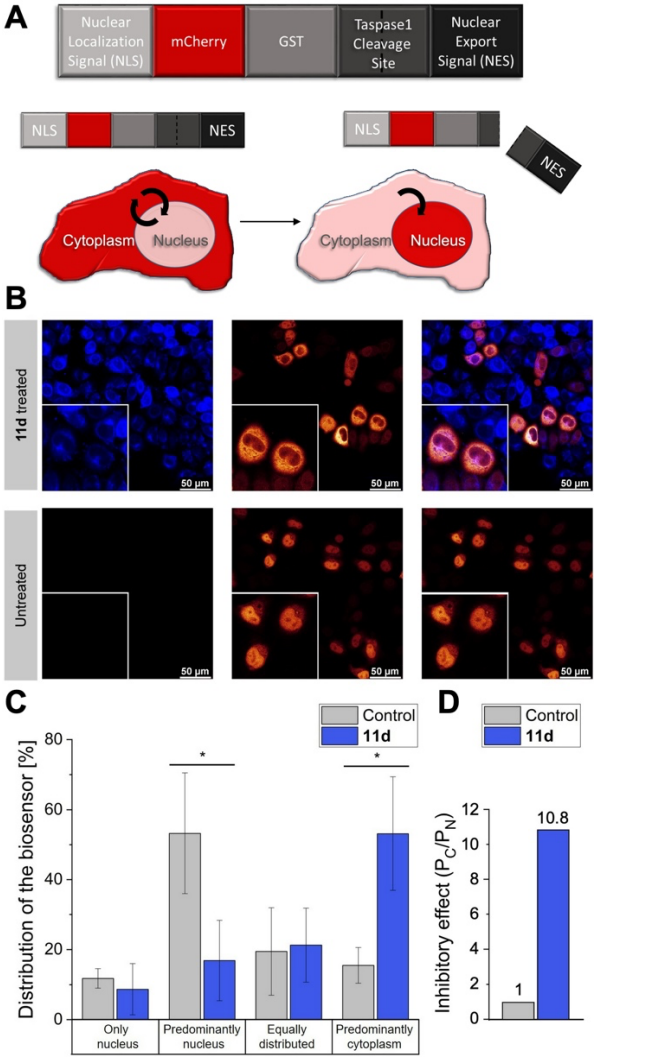


Figure 8: 11d interferes with Taspase 1 proteolytic activity in living cells. A) Modular structure of the autofluorescent Taspase 1-biosensor ($Tasp_{BS}$). Taspase 1-mediated cleavage results in the biosensor's nuclear accumulation. B) Confocal microscopy of living, $Tasp_{BS}$-expressing HeLa Kyoto cells, treated with 50 µM **11d** for 24 h (upper panel) or DMSO control (lower panel). Scale bars, 50 µm; frames indicate representative close-ups. C) Quantification of $Tasp_{BS}$'s intracellular localization. Microscopic images were acquired in at least 100 cells and randomized for localization assignment (see SI, Table S3). Results are the mean of triplicates ± standard deviation (*$p < 0.5$). D) Inhibitory effect of **11d** as represented by the ratio of the mean percentage of cells in the category "Predominantly cytoplasm" (P_C) vs. "Predominately nucleus" (P_N) normalized to 1 for the DMSO control.

CONCLUSION

Based on a comprehensive *in silico*, analytical, biochemical, and *in vivo* approach, we demonstrated that the bisphosphate ligand **11d** is a dual active inhibitor targeting the cancer-relevant protease, Taspase 1. Our conclusion is based on independent experimental evidence: For one, molecular modeling combined with NMR and binding experiments confirmed **11d**'s selectivity for the basic cluster (^{220}KKRR^{223}) in Taspase 1's NLS. Second, **11d**'s high NLS binding affinity prevented Taspase 1's interaction with Importin α interfering with nuclear entry and protease activation (activity #1). Third, biochemical cell-free assays showed that **11d** additionally affected the protease's catalytic cleavage function, most likely by restricting the molecular flexibility of the bound loop (activity #2; Figure 1). Fourth, **11d** efficiently inhibited Taspase 1 also in a cancer-relevant cell model *in vivo*.

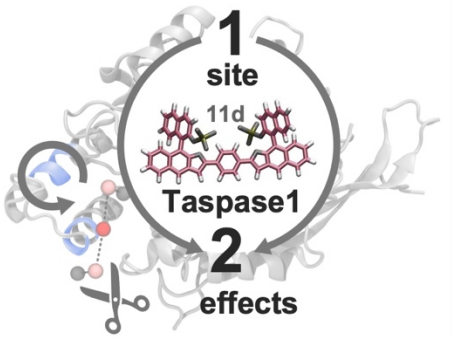
In summary, we could show that by binding to a single site, compound **11d** simultaneously inhibits Taspase 1's protein-protein-interaction-based activation and its enzymatic function both *in vitro* and *in vivo*. This 'proof of principle' study thereby introduces the concept that dual activity inhibition through a single ligand by binding to a single site is indeed possible. Hence, 'targeting two for the price of one' by addressing multiple independent functions in disease-relevant proteins via a single inhibitor should no longer be considered a 'too ambitious' goal thereby stimulating the design of similar dual inhibitors for other targets.

ASSOCIATED CONTENT

Supporting Information. Experimental details regarding synthesis and biological assays. This material is available free of charge via the Internet at http://pubs.acs.org.".

TABLE OF CONTENT

TOC. "Targeting two for the price of one": A selective supramolecular bisphosphate ligand allows dual inhibition of Threonine aspartase 1 by simultaneously addressing independent key functions - its protein-protein interaction-based activation and enzymatic function.



AUTHOR INFORMATION

Corresponding Authors

*shirley.knauer@uni-due.de
*jochen.niemeyer@uni-due.de

ACKNOWLEDGMENT

The authors acknowledge the CRC 1093 funded by the "Deutsche Forschungsgemeinschaft'. We acknowledge the use of the imaging equipment and the support of the "Imaging Center Campus Essen" (ICCE). Instrument Leica TCS SP8X FALCON was obtained through DFG funding (Major Research Instrumentation Program as per Art. 91b GG, INST 20876/294-1 FUGG).

REFERENCES

(1) Baby, K.; Maity, S.; Mehta, C. H.; Suresh, A.; Nayak, U. Y.; Nayak, Y. SARS-CoV-2 entry inhibitors by dual targeting TMPRSS2 and ACE2: An in silico drug repurposing study. *Eur J Pharmacol* **2021,** *896*, 173922.

(2) Stankovic, T.; Dinic, J.; Podolski-Renic, A.; Musso, L.; Buric, S. S.; Dallavalle, S.; Pesic, M. Dual Inhibitors as a New Challenge for Cancer Multidrug Resistance Treatment. *Curr Med Chem* **2019,** *26* (33), 6074-6106.

(3) Szumilak, M.; Wiktorowska-Owczarek, A.; Stanczak, A. Hybrid Drugs-A Strategy for Overcoming Anticancer Drug Resistance? *Molecules* **2021,** *26* (9).

(4) Flemming, A. Bivalent mTOR inhibitors — the next generation. *Nature Reviews Drug Discovery* **2016,** *15* (7), 455-455.

(5) Harrison, C. A double strike at kinases. *Nature Reviews Drug Discovery* **2010,** *9* (4), 271-271.

(6) Cohen, P.; Cross, D.; Jänne, P. A. Kinase drug discovery 20 years after imatinib: progress and future directions. *Nature Reviews Drug Discovery* **2021,** *20* (7), 551-569.

(7) Romero, D. Dual targeting to defeat resistance. *Nature Reviews Clinical Oncology* **2017,** *14* (6), 328-328.

(8) Brazil, M. Targeting two for the price of one. *Nature Reviews Drug Discovery* **2003,** *2* (3), 170-171.

(9) Sohn, J.; Kim, T. I.; Yoon, Y. H.; Kim, J. Y.; Kim, S. Y. Novel transglutaminase inhibitors reverse the inflammation of allergic conjunctivitis. *J Clin Invest* **2003,** *111* (1), 121-8.

(10) Vallet, C.; Aschmann, D.; Beuck, C.; Killa, M.; Meiners, A.; Mertel, M.; Ehlers, M.; Bayer, P.; Schmuck, C.; Giese, M.; Knauer, S. K. Functional Disruption of the Cancer-Relevant Interaction between Survivin and Histone H3 with a Guanidiniocarbonyl Pyrrole Ligand. *Angew Chem Int Ed Engl* **2020,** *59* (14), 5567-5571.

(11) Meiners, A.; Backer, S.; Hadrovic, I.; Heid, C.; Beuck, C.; Ruiz-Blanco, Y. B.; Mieres-Perez, J.; Porschke, M.; Grad, J. N.; Vallet, C.; Hoffmann, D.; Bayer, P.; Sanchez-Garcia, E.; Schrader, T.; Knauer, S. K. Specific inhibition of the Survivin-CRM1 interaction by peptide-modified molecular tweezers. *Nat Commun* **2021,** *12* (1), 1505.

(12) Uhlenheuer, D. A.; Petkau, K.; Brunsveld, L. Combining supramolecular chemistry with biology. *Chem Soc Rev* **2010,** *39* (8), 2817-26.

(13) Knauer, S. K.; Fetz, V.; Rabenstein, J.; Friedl, S.; Hofmann, B.; Sabiani, S.; Schröder, E.; Kunst, L.; Proschak, E.; Thines, E.; Kindler, T.; Schneider, G.; Marschalek, R.; Stauber, R. H.; Bier, C. Bioassays to Monitor Taspase1 Function for the Identification of Pharmacogenetic Inhibitors. *PloS one* **2011,** *6* (5).

(14) Höing, A.; Zimmermann, A.; Moews, L.; Killa, M.; Heimann, M.; Hensel, A.; Voskuhl, J.; Knauer, S. K. A bivalent supramolecular GCP-ligand enables blocking of the Taspase1/Importin α interaction. *ChemMedChem* **2021**.

(15) Gribko, A.; Hahlbrock, A.; Strieth, S.; Becker, S.; Hagemann, J.; Deichelbohrer, M.; Hildebrandt, A.; Habtemichael, N.; Wünsch, D. Disease-relevant signalling-pathways in head and neck cancer: Taspase1's proteolytic activity fine-tunes TFIIA function. *Scientific Reports* **2017,** *7*.

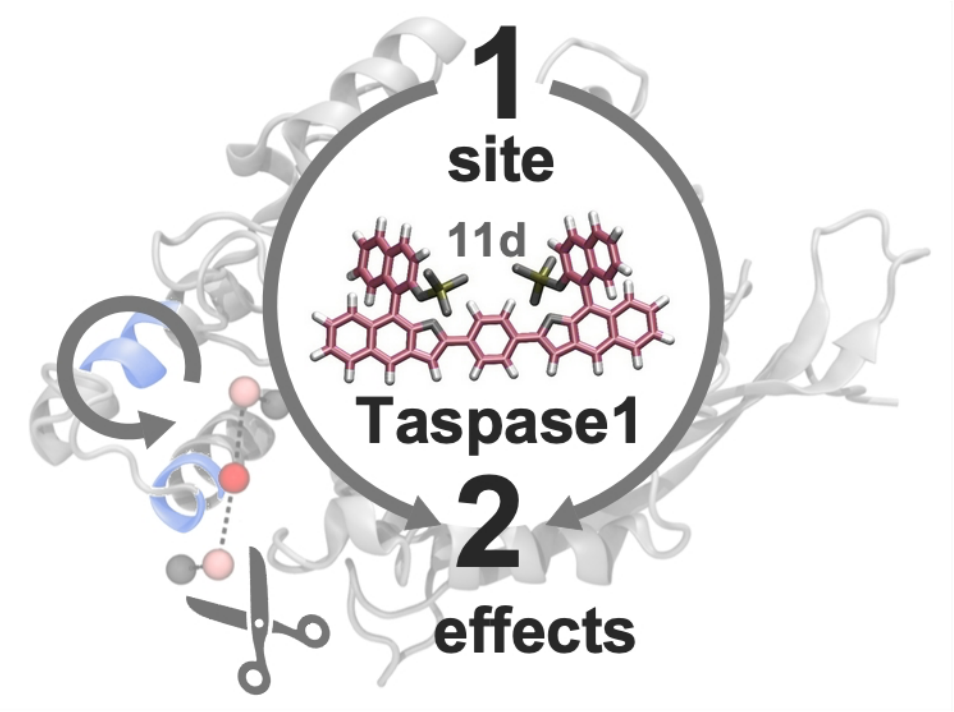
(16) Boom, J. v. d.; Mamić, M.; Baccelliere, D.; Zweerink, S.; Kaschani, F.; Knauer, S.; Bayer, P.; Kaiser, M. Peptidyl Succinimidyl Peptides as Taspase 1 Inhibitors. *ChemBioChem* **2014,** *15* (15), 2233–2237.

(17) Bier, C.; Knauer, S. K.; Wünsch, D.; Kunst, L.; Scheiding, S.; Kaiser, M.; Ottmann, C.; Krämer, O. H.; Stauber, R. H. Allosteric inhibition of Taspase1's pathobiological activity by enforced dimerization in vivo. *The FASEB Journal* **2012,** *26* (8), 3421–3429.

(18) Bier, C.; Knauer, S. K.; Klapthor, A.; Schweitzer, A.; Rekik, A.; Krämer, O. H.; Marschalek, R.; Stauber, R. H. Cell-based Analysis of

Structure-Function Activity of Threonine Aspartase 1*. *The Journal of Biological Chemistry* **2011,** *286* (4), 3007–3017.
(19) Bier, C.; Knauer, S. K.; Docter, D.; Schneider, G.; Kramer, O. H.; Stauber, R. H. The importin-alpha/nucleophosmin switch controls taspase1 protease function. *Traffic (Copenhagen, Denmark)* **2011,** *12* (6), 703–714.
(20) Bier, C.; Hecht, R.; Kunst, L.; Scheiding, S.; Wunsch, D.; Goesswein, D.; Schneider, G.; Kramer, O. H.; Knauer, S. K.; Stauber, R. H. Overexpression of the catalytically impaired Taspase1 T234V or Taspase1 D233A variants does not have a dominant negative effect in T(4;11) leukemia cells. *PloS one* **2012,** *7* (5), e34142.
(21) López-Otín, C.; Bond, J. S. Proteases: Multifunctional Enzymes in Life and Diseases. *The Journal of Biological Chemistry* **2008,** *283* (45), 30433–30437.
(22) Bond, J. S. Proteases: History, discovery, and roles in health and disease. *The Journal of Biological Chemistry* **2019,** *294* (5), 1643–1651.
(23) Wünsch, D.; Hahlbrock, A.; Jung, S.; Schirmeister, T.; Boom, J. v. d.; Schilling, O.; Knauer, S. K.; Stauber, R. H. Taspase1: a 'misunderstood' protease with translational cancer relevance. *Oncogene* **2016,** *35* (26), 3351–3364.
(24) Chen, D. Y.; Liu, H.; Takeda, S.; Tu, H.-C.; Sasagawa, S.; van Tine, B. A.; Lu, D.; Cheng, E. H.-Y.; Hsieh, J. J.-D. Taspase1 functions as a non-oncogene addiction protease that coordinates cancer cell proliferation and apoptosis. *Cancer research* **2010,** *70* (13), 5358–5367.
(25) 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* **2003,** *115* (3), 293–303.
(26) Khan, J. A.; Dunn, B. M.; Tong, L. Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. *Structure (London, England : 1993)* **2005,** *13* (10), 1443–1452.
(27) Takeda, S.; Chen, D. Y.; Westergard, T. D.; Fisher, J. K.; Rubens, J. A.; Sasagawa, S.; Kan, J. T.; Korsmeyer, S. J.; Cheng, E. H.-Y.; Hsieh, J. J.-D. Proteolysis of MLL family proteins is essential for Taspase1-orchestrated cell cycle progression. *Genes & Development* **2006,** *20* (17), 2397–2409.
(28) Zhou, H.; Spicuglia, S.; Hsieh, J. J.-D.; Mitsiou, D. J.; Høiby, T.; Veenstra, G. J. C.; Korsmeyer, S. J.; Stunnenberg, H. G. Uncleaved TFIIA Is a Substrate for Taspase 1 and Active in Transcription. *Molecular and Cellular Biology* **2006,** *26* (7), 2728-2735.
(29) Takeda, S.; Sasagawa, S.; Oyama, T.; Searleman, A. C.; Westergard, T. D.; Cheng, E. H.; Hsieh, J. J. Taspase1-dependent TFIIA cleavage coordinates head morphogenesis by limiting Cdkn2a locus transcription. *The Journal of Clinical Investigation* **2015,** *125* (3), 1203–1214.
(30) Schrenk, C.; Fetz, V.; Vallet, C.; Heiselmayer, C.; Schroder, E.; Hensel, A.; Hahlbrock, A.; Wunsch, D.; Goesswein, D.; Bier, C.; Habtemichael, N.; Schneider, G.; Stauber, R. H.; Knauer, S. K. TFIIA transcriptional activity is controlled by a 'cleave-and-run' Exportin-1/Taspase 1-switch. *J Mol Cell Biol* **2018,** *10* (1), 33-47.
(31) Chen, D. Y.; Lee, Y.; van Tine, B. A.; Searleman, A. C.; Westergard, T. D.; Liu, H.; Tu, H.-C.; Takeda, S.; Dong, Y.; Piwnica-Worms, D. R.; Oh, K. J.; Korsmeyer, S. J.; Hermone, A.; Gussio, R.; Shoemaker, R. H.; Cheng, E. H.-Y.; Hsieh, J. J.-D. A Pharmacological Inhibitor of the Protease Taspase1 Effectively Inhibits Breast and Brain Tumor Growth. *Cancer research* **2012,** *72* (3), 736–746.
(32) Pasch, P.; Höing, A.; Ueclue, S.; Killa, M.; Voskuhl, J.; Knauer, S. K.; Hartmann, L. PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction. *Chemical Communications* **2021,** *57* (25), 3091–3094.
(33) van den Boom, J.; Trusch, F.; Hoppstock, L.; Beuck, C.; Bayer, P. Structural Characterization of the Loop at the Alpha-Subunit C-Terminus of the Mixed Lineage Leukemia Protein Activating Protease Taspase1. *PloS one* **2016,** *11* (3), e0151431.
(34) Nagaratnam, N.; Delker, S. L.; Jernigan, R.; Edwards, T. E.; Snider, J.; Thifault, D.; Williams, D.; Nannenga, B. L.; Stofega, M.; Sambucetti, L.; Hsieh, J. J.; Flint, A. J.; Fromme, P.; Martin-Garcia, J. M. Structural insights into the function of the catalytically active human Taspase1. *Structure* **2021,** *29* (8), 873-885.e5.
(35) Octa-Smolin, F.; Mitra, R.; Thiele, M.; Daniliuc, C. G.; Stegemann, L.; Strassert, C.; Niemeyer, J. Rigidly Tethered Bis-phosphoric Acids: Generation of Tunable Chiral Fluorescent Frameworks and Unexpected Selectivity for the Detection of Ferric Ions. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2017,** *23* (42), 10058–10067.
(36) Octa-Smolin, F.; van der Vight, F.; Yadav, R.; Bhangu, J.; Soloviova, K.; Wolper, C.; Daniliuc, C. G.; Strassert, C. A.; Somnitz, H.; Jansen, G.; Niemeyer, J. Synthesis of Furan-Annelated BINOL Derivatives: Acid-Catalyzed Cyclization Induces Partial Racemization. *J Org Chem* **2018,** *83* (23), 14568-14587.
(37) Yadav, R.; Kwamen, C.; Niemeyer, J. Development of Fluorescent Chemosensors for Amino-Sugars. *Israel Journal of Chemistry* **2021,** *61*.
(38) Lee, J. T.; Chen, D. Y.; Yang, Z.; Ramos, A. D.; Hsieh, J. J.; Bogyo, M. Design, syntheses, and evaluation of Taspase1 inhibitors. *Bioorg Med Chem Lett* **2009,** *19* (17), 5086-90.
(39) Siemer, S.; Bauer, T. A.; Scholz, P.; Breder, C.; Fenaroli, F.; Harms, G.; Dietrich, D.; Dietrich, J.; Rosenauer, C.; Barz, M.; Becker, S.; Strieth, S.; Reinhardt, C.; Fauth, T.; Hagemann, J.; Stauber, R. H. Targeting Cancer Chemotherapy Resistance by Precision Medicine-Driven Nanoparticle-Formulated Cisplatin. *ACS Nano* **2021**.
(40) van der Meer, S. B.; Hadrovic, I.; Meiners, A.; Loza, K.; Heggen, M.; Knauer, S. K.; Bayer, P.; Schrader, T.; Beuck, C.; Epple, M. New Tools to Probe the Protein Surface: Ultrasmall Gold Nanoparticles Carry Amino Acid Binders. *J Phys Chem B* **2021,** *125* (1), 115-127.

For Table of Contents Only:



TOC: "Targeting wo for the price of one": A selective supramolecular bisphosphate ligand allows dual inhibition of Threonine aspartase 1 by simultaneously addressing independent key functions - its protein-protein interaction-based activation and enzymatic function.

2.4 Publication VI

2.4.1 Author contributions

Title:

Recognition of a flexible protein loop in Taspase1 by multivalent supramolecular tweezers

Authors:

Alexander Höing[‡], Abbna Kirupakaran[‡], Christine Beuck, Jochen Niemeyer, Peter Bayer, Shirley K. Knauer*, Thomas Schrader*

[‡]authors contributed equally *corresponding authors

Contributions:

Conception: 20 %

Experimental work: 45 %

(Protein purification, pull-down assay, SDS-PAGE, western blotting, fluorescence titration, colorimetric cleavage assay)

Data analysis: 50 %

(See experimental work)

Writing the manuscript: 50 %

(50 % of the first draft; Figures 1, 3, 4, 5, 7, S4, S5, S6, S7)

PhD student Principal investigator

2.4.2 Article introduction and summary

All previous studies reveale that the multivalency of supramolecular ligands is a key factor for effectiveness. **3GP** and **3GLP** are trivalent, **2GC** and **11d** bivalent binders, while respective monovalent controls never proved effective (Octa-Smolin *et al.* 2018; Pasch *et al.* 2021; Höing *et al.* 2021). We therefore designed a comparative study on the influence of multivalency on binding affinity and inhibitory effects with Taspase1, using the molecular tweezer CLR01, a well-investigated anionic binder, as an example (Fokkens *et al.* 2005).

In this study, we are guided by the following questions: (i) Does an increase in valency result in a more specific targeting of the Taspase1 NLS, which consists of densely packed basic clusters? (ii) Will multivalent tweezer molecules target different amino acids in the Taspase1 loop? (iii) Does the increase in valency result in a higher overall binding affinity, namely multiavidity? (iv) Will multivalent tweezer molecules differ in the effects on the Taspase1 activity?

We utilized polyanionic molecular tweezers with a varying number of up to five binding units. Indeed, the tweezer with a higher valency had greater effects on the Tasp/Impα-interaction in the pull-down assay, in answer to question (i). While NMR spectroscopy provided proof of binding to the Tasapse1 loop for all tweezer molecules, there were no distinguishable patterns of targeted amino acids in the loop(ii). Binding to the loop was also confirmed by fluorescence titration with a FAM-labeled Taspase1 loop, but the binding affinities of all tweezer molecules were in the same range (iii). A colorimetric cleavage assay for Taspase1 activity revealed increased effects on the proteolytic activity in correlation to the valency, but the IC_{50} vales determined were in the same range of 2-3 µM (iv).

2.4.3 Recognition of a flexible protein loop in Taspase1 by multivalent supramolecular tweezers

RESEARCH ARTICLE

Recognition of a flexible protein loop in Taspase 1 by multivalent supramolecular tweezers

Alexander Höing[a+], Abbna Kirupakaran[b+], Christine Beuck[c], Marius Pörschke[c], Felix Niemeyer[b], Peter Bayer[c], Thomas Schrader[b*], Shirley K. Knauer[a*]

[a] Molecular Biology II, Faculty of Biology, Center of Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany
E-mail: shirley.knauer@uni-due.de
[b] Institute of Organic Chemistry I, Faculty of Chemistry, University of Duisburg-Essen, Essen, Universitätsstrasse 7, 45141 Essen, Germany
Email: thomas.schrader@uni-due.de
[c] Structural and Medicinal Biochemistry, Center of Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

[+] These authors contributed equally.
[*] Corresponding authors.

Supporting information for this article is given via a link at the end of the document.

Abstract: Many protein-protein interactions utilize a multivalent display of epitopes for binding to enhance molecular affinity and specificity. Imitating this natural concept, we here report the sophisticated design of multivalent supramolecular tweezers that allow to address lysine and arginine clusters on a flexible protein surface loop. The unique protease Taspase 1 is not only involved in cancer development, but also characterized by a basic bipartite nuclear localization signal (NLS) interacting with Importin α pivotal for proteolytic activation. Newly established synthesis routes enabled us to covalently fuse several tweezer molecules into multivalent ligands. The resulting bi- up to pentavalent constructs were then systematically compared in comprehensive biochemical assays. Indeed, the stepwise increase in valency was robustly reflected by the ligands' gradually enhanced potency to disrupt the interaction of Taspase 1 with Importin α, correlated with higher binding affinity and inhibition of proteolytic activity.

Introduction

The more, the better: Multivalency, also known as polyvalency, is a common concept used in nature where proteins often form multimeric architectures.[1] Here, synergistic combinations of several cooperative binding events allow to compensate for weak individual non-covalent interactions (Fig. 1A). Multivalent binding is thus significantly more resilient than individual bonds between the equivalent number of monovalent ligands and can drastically enhance the binding affinity between interacting species. Unlike covalent bonds, multivalent interactions are also reversible through the sequential separation of individual bonds, rendering the respective physiological interactions enormously dynamic. Consequently, multivalency plays a pivotal role in a plethora of biological processes involving interactions between cells, or with organisms like bacteria or viruses but also between different types of molecules.

A prominent example are protein-protein interactions (PPIs) which modulate the communication between more than 200,000 human proteins encoded in our genome. Rational modulation of PPIs not only provides insight into biological mechanisms but also allows to interfere with pathological processes.[2-3]

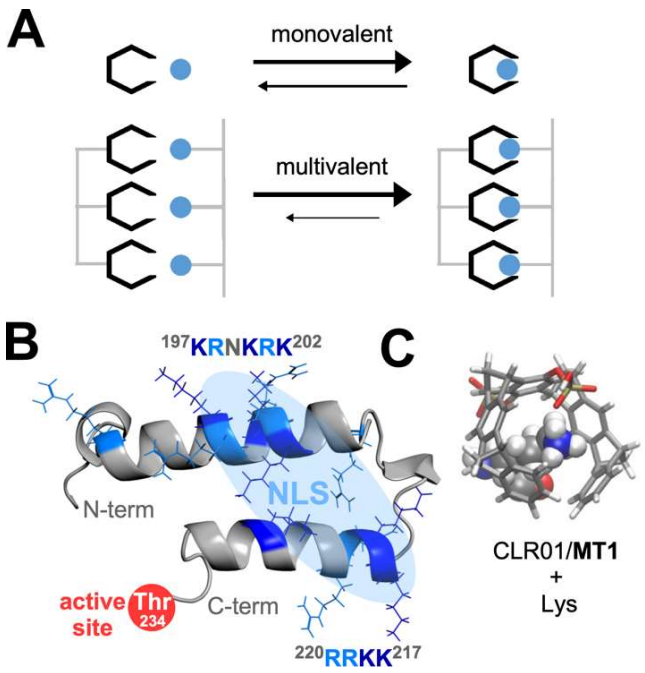


Figure 1. The concept of multivalency was applied to enhance binding of supramolecular tweezers to surface-exposed lysine and arginine residues in the Taspase 1 loop. A. Schematic representation of mono- versus multivalent binding. **B.** NMR solution structure of the Taspase 1 loop, comprising residues 183-233.[4] The bipartite NLS (highlighted in blue) consists of two basic clusters on two neighboring helices. Side chains of lysines (dark blue) and arginines (light blue), N-/C-termini and the position of the catalytically active site in the context of the full-length protein are indicated. **C.** Stick representation of the structure of the monovalent tweezer CLR01/MT1 and its engagement of a lysine side chain.

Many natural proteins contain flexible loops which explore a well-defined complementary binding site in partner proteins and often undergo a disorder-to-order-transition when they dock on.[5-8] For example, cargo proteins insert their flexible loop motif into the large open binding cleft of 14-3-3 adaptor proteins and are thus transported to their destination where they dissociate again.[9-12]

The same is true for export and import receptors between the cell nucleus and the cytosol.[13] In all cases, affinities remain moderate in the medium micromolar K_D regime and allow facile release of the cargo protein.[14-15] Deliberate interference with such processes by external agents requires the ability to recognize a flexible protein loop. This is highly challenging due to the inherent conformational dynamics and only very few artificial ligands have been reported until today which solve this problem.[16-19] Quite often, however, such protein loops contain clustered charged amino acid residues which exert a powerful electrostatic attraction towards the oppositely charged interface on the partner protein.[20-22] This provides a handle for the supramolecular chemist: single host molecules for the charged residues could be covalently fused into a multivalent protein ligand with mobile arms which adapt to the varying distance of their guests on distinct flexible protein loops.

As regulators of many biological processes, proteases are essential for cell viability and their deregulation is often associated with various pathologies including cancer.[23-24] In particular, the unique threonine protease Taspase 1 (**T**hreonine **asp**art**ase 1**) is not only implicated in the regulation of cellular development but also in the development and progression of leukemias and solid tumors.[25-27] In the group of threonine proteases, only Taspase 1 is able to cleave other substrates in *trans* by recognizing a conserved $Q^3[FILV]^2D^1{\downarrow}G^1X'^{2'}D^{3'}D^{4'}$ motif.[28-31] Thus, Taspase 1 is an endopeptidase that utilizes the N-terminal threonine (Thr^{234}) of its mature β-subunit as the active site. Besides the Mixed Lineage Leukemia protein MLL,[28] the general transcription factor IIA (TFIIA)[27, 31-34] and the Upstream Stimulatory Factor 2 (USF2)[31] could all be identified as Taspase 1 target proteins.[31] Importantly, Taspase 1 is preferentially expressed during embryonic development, and thus pharmacological intervention should not affect healthy adult cells.[25, 30, 35-37] In contrast, tumor cells have been demonstrated to upregulate expression of Taspase 1 to promote their proliferation and counteract apoptotic cell death.[25-26, 30, 35]. This classifies Taspase 1 as a 'non-oncogene addiction' protease.[25-26] However, the full repertoire of physiological and pathological targets and pathways affected by Taspase 1 is still not fully understood.

Although Taspase 1 belongs to the family of threonine proteases it is not affected by general protease inhibitors.[25-26] Some more or less specific inhibitors have been suggested for Taspase 1, but their mode of action is often not understood or they need to be used in high concentrations.[26, 28, 35, 38-39]. This clearly impedes the molecular dissection of Taspase 1's (patho)biologic functions and hence the full exploitation of its therapeutic potential.[27, 36, 39-41] Until today, Taspase 1 is not only a promising target for intervention but also a challenging model for rational drug design.[39, 42]

Mechanistically, Taspase 1 has to undergo a rather complex multistep activation process to fully propagate its proteolytic activity.[25, 43] In the cytoplasm, it is initially expressed as an inactive proenzyme (45 kDa). Subsequently, it depends on active import into the nucleus to allow autoproteolytic activation, prerequisite for *trans*-cleavage of its cellular substrates.[25, 43] Nuclear access relies on Taspase 1's bipartite nuclear localization signal (NLS, aa ^{197}KRNKRK-x_{14}-KKRR^{220}). [25, 43] The basic amino acid clusters are arranged on two neighboring α-helices constituting a surface-exposed loop (Fig. 1B). It belongs to the α-subunit and interacts with the intracellular transport receptor Importin α.[43] Of note, this surface-exposed loop behaves flexible in solution as shown by NMR analyses[4] although some but not all crystal structures indicate more rigid helical structures.[29, 44] Inside the nucleus, autoproteolysis precedes the formation of two subunits, α (25 kDa) and ß (20 kDa), which reassemble into the active heterodimer to finally propagate substrate cleavage.[28, 43] Thus, the Taspase 1/Importin α interaction is essential for intracellular protease activation; moreover, the catalytically active center (Thr^{234}) is located at the base of the loop, and represents an attractive target itself for chemical intervention.

In our attempt to block both Taspase 1 functions simultaneously, we decided to address the densely packed basic amino acid clusters of Taspase 1's NLS presented on the surface-exposed loop with tailored supramolecular ligands. A newly established synthesis route now enabled us to use our well-known lysine and arginine tweezers to construct multivalent ligands with flexible arms (Fig. 2). These should allow to reach several guests at the same time and maintain conformational flexibility in the complex to minimize entropic costs. The resulting architectures - resembling a kraken - should ideally bridge both α-helices and act as a conformational clamp. A versatile bottom-up design of our ligand portfolio allowed to systematically investigate the influence of multivalency and spacer flexibility on ligand performance in comprehensive biochemical assays.

Results and Discussion

Design and synthesis of multivalent lysine tweezers

For the synthesis of multivalent molecular tweezers, we chose a central benzene core which could be equipped with a desired number of carboxylates and esterified with terminal alkynols. As a second reaction partner, a monofunctionalized tweezer derivative with one terminal azide group was required. In a short sequence, employing the well-known TCA (trichloroacetonitrile) strategy the hydroquinone tweezer was therefore manipulated into a diphosphate tweezer carrying one azidoalkyl ester moiety. The reverse approach involves template molecules with multiple azides and was abandoned to avoid explosions. Initially, azidopropanol was attached to the tweezer; however massive upfield shifts of the ester arm pointed to unwanted self-inclusion inside the tweezer cavity and all attempts to perform a click reaction failed. The problem could be solved by shortening the spacer and using azidoethanol instead (SI, Fig. S1-S3). Now click reactions proceeded smoothly in good yields, but these were limited to dimers. After extensive optimization, an iterative protocol was developed which involves stepwise performance of each coupling step in one pot. It is imperative that only after the first coupling is complete, the next equivalent of azidotweezer is added together with the full reagent cocktail (Fig. 2A, see SI for details). Thus, five consecutive coupling steps lead to the pentameric tweezer in good yield. We explain this necessity with the consumption of reagents even under inert conditions. Similar observations have been reported by other groups.[42] With this iterative approach, the full series from bivalent to pentavalent tweezers became accessible in high purity and good yields (Fig. 2, and see SI). Molecular weights of the largest compounds are in the low kDa range and approach the size of a polypeptide (20 - 40 mers). Importantly, even the large members of the new supramolecular ligand family remain very well soluble in aqueous buffer and furnish clean molecular ion peaks in the ESI-MS as well as sharp signals in the ^{1}H NMR spectrum (see SI).

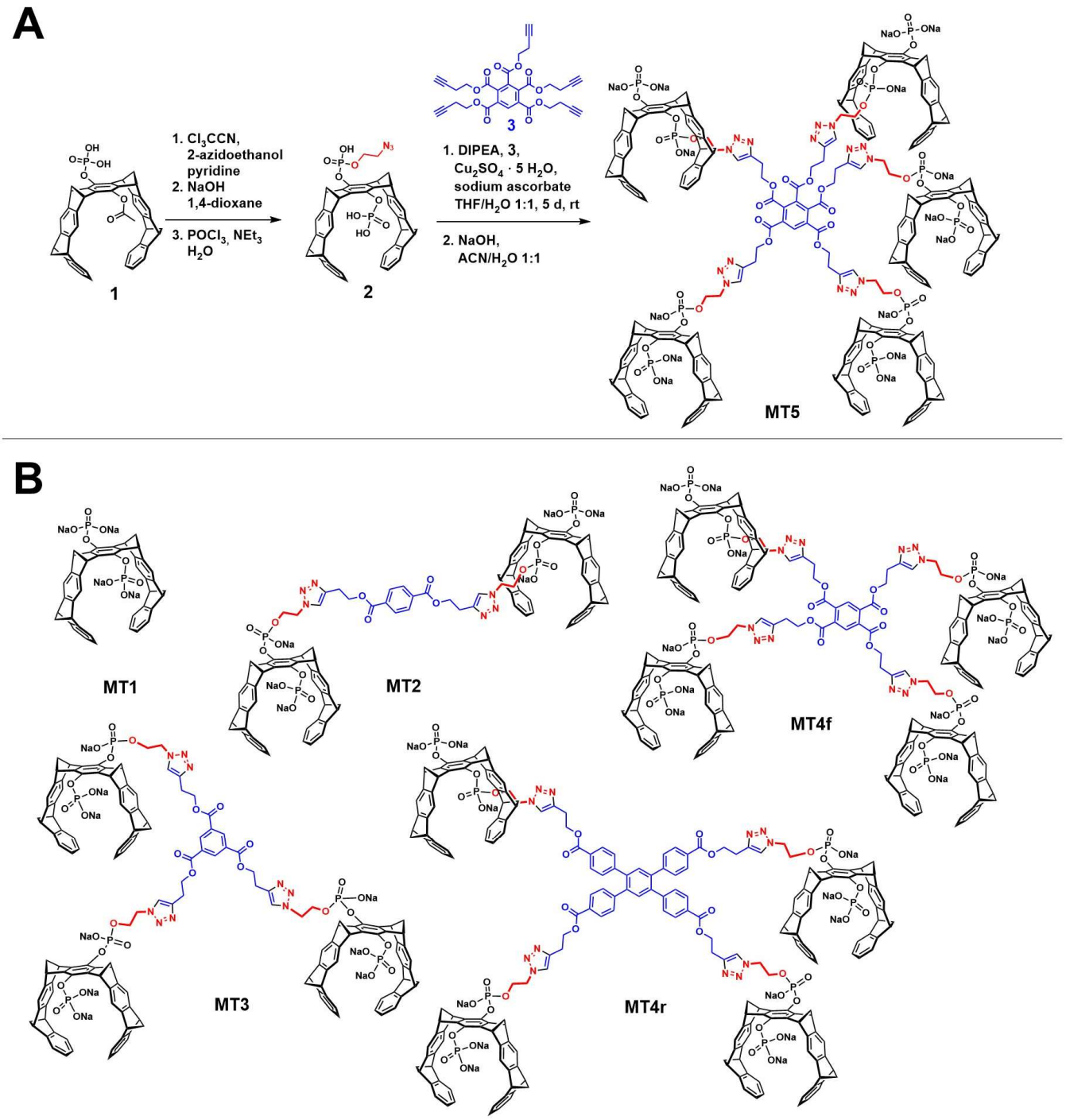


Figure 2. Synthetic strategy leading to a library of multivalent molecular tweezers with mobile connecting arms. A. Introduction of one azidoethylphosphate ester arm on monoacetoxy monophosphate tweezer **1** (TCA coupling) followed by Click chemistry with 1,2,3,4,5-penta-3-butyn-1-yl-1,2,3,4,5-benzenepentacarboxylate **3** furnishes pentavalent tweezer **MT5**. **B.** Monovalent (MT1) to tetravalent (MT4f; f = flexible) tweezers possess identical flexible spacers, whereas MT4r (r = rigid) is a rigid analogue with a larger pentaphenyl core unit.

Together with the absence of upfield shifted aromatic signals of the tweezer sidewalls this excludes aggregation and self-inclusion of the tweezer units, most likely because each of them carries three negative charges.

Multivalency gradually increases the tweezers potential to interfere with the Taspase 1/Importin α interaction

A customized biochemical pull-down assay was utilized to evaluate the effects of our multivalent molecular tweezers on the Taspase 1/Importin α interaction (Fig. 3 and 4, see SI for details).[41-42] Because Importin binding crucially relies on the basic NLS clusters present in the Taspase 1 loop, the assay serves as a reliable and direct readout for tweezer binding to the responsible lysine and arginine residues. Briefly, recombinant Importin α is expressed with an N-terminal glutathione S-transferase (GST) affinity tag and immobilized on a glutathione sepharose column. Consequently, binding of recombinant Taspase 1-His protein in the presence or absence

of tweezers allows to quantify their inhibitory potential. Since wildtype Taspase 1 partially undergoes autoproteolysis during protein purification, we used a catalytically inactive mutant (Taspase $1_{233A/234A}$) ensuring a stable proenzyme concentration and thus a robust and reliable detection.

As predicted, increasing valency gradually enhanced the potency of the tweezers to interfere with the Taspase 1/Importin α interaction, particularly in the lower micromolar range. In contrast to the monovalent construct, multivalent tweezers were already active at concentrations of 10 µM, with the flexible tetravalent fMT4 significantly reducing the amount of Taspase 1 (2.2 µM) bound to Importin α (2.5 µM) by more than 80 % (Fig. 3, and see SI, Fig. S4-S5). The gradual increase correlating with ligand valency is comprehensively visualized in a heatmap representation summarizing the pull-down data for MT1 to MT4f (Fig. 3B). Indeed, the supremacy of multivalent ligands becomes most evident at low ligand concentrations where enhanced binding affinities have the highest impact.

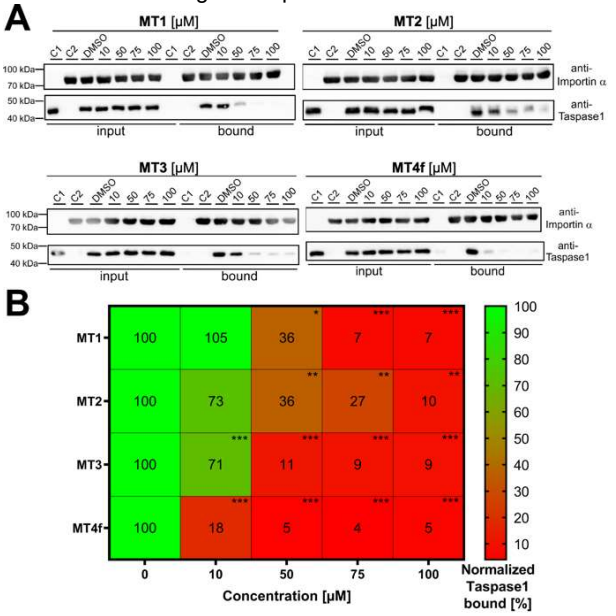


Figure 3. Multivalency increases the tweezers' potential to interfere with the Taspase 1/Importin α interaction. A. Pull-down assays were performed with increasing concentrations (10-100 µM) of indicated tweezers MT1, MT2, MT3, and MT4f. Controls: Taspase 1-His (C1) or GST-Importin α (C2) alone were added to the column, and a DMSO-treated sample served as reference. **B.** Heatmap representation of the densitometric quantification of pull-down assays including tweezer molecules MT1, MT2, MT3 and MT4f. Results are given as mean of three replicates including standard deviations (*p < 0.05, **p < 0,01) relative to the DMSO control. See ESI for individual bar charts.

For the most potent tetra- and pentavalent tweezers concentrations effective in the pull-down assays were lowered further to 2µM (Fig. 4). While the tetravalent tweezers MT4f and MT4r reduced the amount of bound Taspase 1 by 27 % and 41 %, respectively, MT5 induced a decrease of more than 60 %. Thus, a general trend was observed for all tweezer generations, and higher valency was always accompanied by a steadily increasing potential to disrupt the essential protein-protein interaction responsible for Taspase 1 import.

The difference between both tetravalent tweezers MT4f and MT4r may reflect rigidity, but also sheer size which makes it easier to span both α-helices.

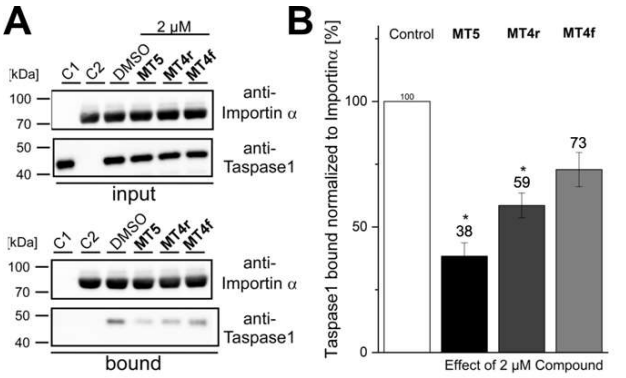


Figure 4. High valency significantly decreases the effective inhibitory concentration of molecular tweezers. A. Pull-down assays of Taspase 1 and Importin α were performed with 2 µM of multivalent tweezers MT4f, MT4r and MT5. Controls: Taspase 1-His (C1) or GST-Importin α (C2) alone were added to the column, and a DMSO-treated sample served as reference. **B.** Densitometric quantification of pull-down assays, comprising the mean of three replicates ± standard deviation (*p < 0.05, **p < 0,01), compared to the DMSO control.

Multivalent tweezers directly address the Taspase 1 loop

To characterize the direct interaction between multivalent tweezers and surface-exposed basic residues in Taspase 1's NLS domain we coupled an N-terminal FAM-label to the NLS-bearing loop peptide (aa S_{181}-D_{233}) and performed fluorescence titration experiments. Here, ligand binding is expected to quench the fluorescent signal of the loop peptide. Addition of mono- as well as multivalent ligands to the isolated labeled peptide loop indeed resulted in a significant decrease in their fluorescence intensity (Fig. 5). However, mono- and divalent ligands MT1 and MT2 did not reach saturation even up to 300 µM. By contrast, ligands MT3, MT4f, MT4r and MT5 with higher valency produced K_D values in the low micromolar range between 4 and 19 µM, indicative of potent binding. The free FAM label without attached peptide was used as a negative control, and data were normalized to account for possible dilution effects. Surprisingly, the pentamer is not the best binder, possibly due to the lower preorganization of the free helix-turn-helix motif compared to the context full-length protein.

Fluorescence titrations thus corroborated our tweezer molecules as direct loop binders; further insight was sought from structural biology. Specifically, we carried out 2D NMR experiments with the helix-turn-helix motif and successive tweezer generations to identify their contact regions on the Taspase 1 loop (Fig. 6). HSQC spectra of 1:1 complexes indeed revealed distinct changes in relative signal intensities of regions which contain the bipartite NLS with several lysine and arginine residues. In contrast to largely unspecific binding of the small monovalent MT1, known as camouflage,[45] tweezer molecules with higher valency showed rather specific association to both helices, evidenced by massive line broadening resulting in strong signal intensity decrease. Presumably, effects of additional intramolecular interactions in the vicinity of this specific basic cluster and the neighboring

turn motif contribute to this selective interaction. These experiments verify the direct interaction of our multivalent tweezer molecules with the regions containing several well accessible basic amino acids of the NLS loop.

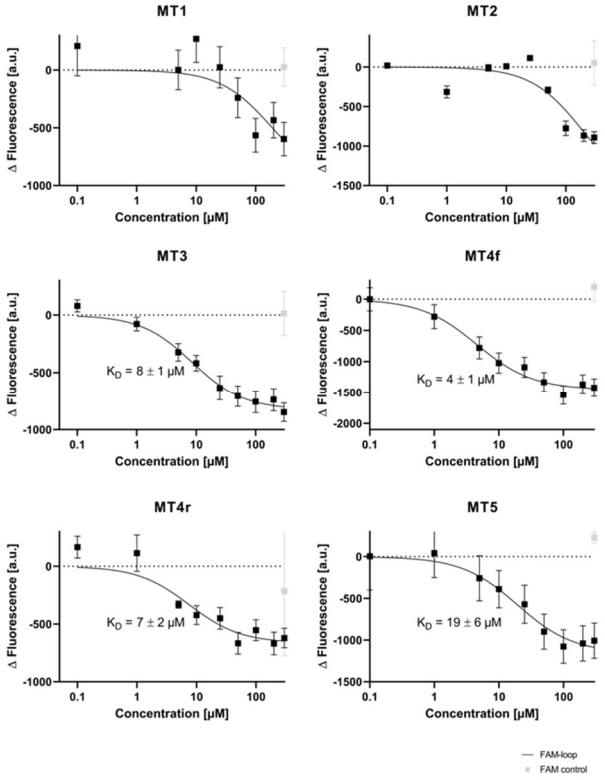


Figure 5. Multivalent supramolecular tweezers directly address the Taspase 1 loop. Fluorescence titrations of a fluorophore-labeled Taspase 1 loop containing the bivalent NLS with multivalent tweezers displayed significant fluorescence quenching. FAM without the linked peptide served as a control. The results are the mean of three replicates ± standard deviation. K_D values are given as fit ± standard error of the mean.

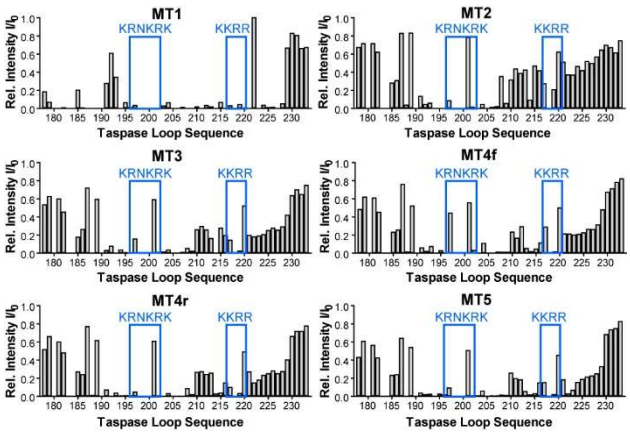


Figure 6. Protein NMR spectroscopy identifies the NLS regions in the Taspase 1 loop as binding sites of multivalent tweezers. B) The relative signal intensities of $^{1}H,^{15}N$-HSQC spectra of ^{15}N-Taspase 1 $loop_{178-233}$ (250-300 μM) with one equivalent of indicated tweezers compared to the spectrum without tweezers show a decrease within the regions that contain Taspase 1's bipartite NLS (blue frames) indicating ligand binding.

Multivalent tweezers also inhibit the proteolytic activity of Taspase 1

Since the active site of Taspase 1 is directly located at the base of the loop harboring the bipartite NLS, it may be expected that large multivalent ligands binding to this signal might also affect the protease's proteolytic activity. In principle the presence of a sterically demanding binding partner may hinder or even prevent substrate binding and subsequent cleavage.

We thus employed a recently developed biochemical cleavage assay which uses a recombinant fusion protein harboring the MLL CS2 substrate cleavage site (see SI). Indeed, addition of recombinant, proteolytically active Taspase 1 (200 nM) resulted in substrate cleavage (Fig. 7A, and see SI, Fig. S7). While addition of the mono- and divalent ligands had only marginal effects, tweezers of higher valency substantially interfered with Taspase 1's proteolytic activity in a concentration-dependent manner. Again, inhibition improved gradually proportional to ligand valency; it could be comprehensively visualized in a heatmap representation summarizing the cleavage assay data for MT1 to MT5 (Fig. 7B). Importantly, the supremacy of the tweezer pentamer becomes again apparent at lower concentrations where enhanced binding affinities have the strongest impact.

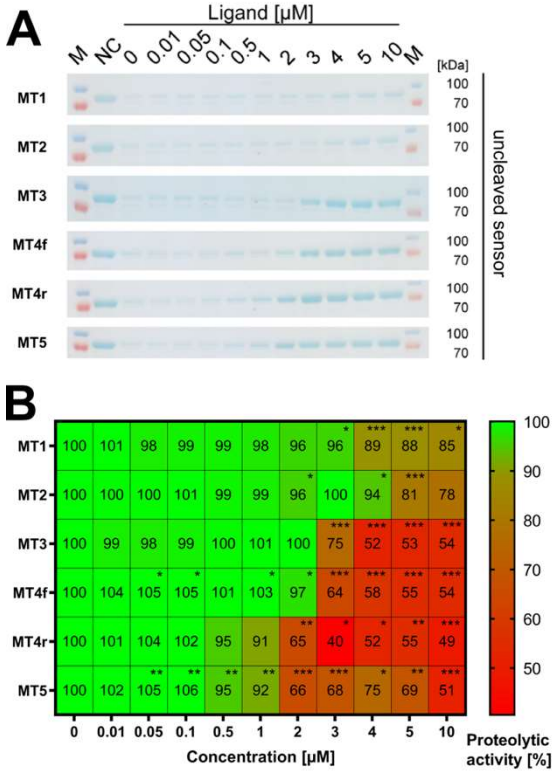


Figure 7. Multivalent supramolecular tweezers interfere with Taspase 1's proteolytic activity. A. Cleavage of a recombinant substrate by Taspase 1 was assessed via SDS-PAGE and Coomassie staining after 90 min reaction time in presence of indicated multivalent tweezers (0.1 - 10 μM). Inhibition of proteolysis was evaluated by the amount of uncleaved sensor protein. A sample without Taspase 1 served as negative control (NC). M, molecular weight marker. **B.** Heatmap representation of the densiometric quantification of Coomassie gels of indicated tweezers. Results are given as mean of three replicates including standard deviations (*p < 0.05, **p < 0,01) relative to the control. See ESI for individual bar charts.

The results from our preceding experiments were finally employed *in silico* to obtain a deeper imagination of the interaction between multivalent tweezers and the flexible loop of Taspase 1 (see SI for details). Molecular modeling was performed with a focus on the most potent pentavalent construct MT5. After manual docking of all tweezer units on both Taspase 1 helices, extensive MD-simulations were run both on the isolated loop as well as on full-length Taspase 1 (PBS buffer, discrete solvent). They revealed fast dissociation of two tweezer units from their complexed lysine residue already after 10 ns (Fig. 8). This was, however, followed by an elongated period without significant changes indicating a highly stable binding mode by three tweezer arms in both cases. Time scales were 300 ns for full-length Taspase 1, and even 1500 ns for the isolated loop. Importantly, the large oligomer easily spanned and thus linked both basic clusters on the two helices right until the end and constituted a dense network of side chains and receptor arms. In addition, tweezer phosphates were found to participate in external electrostatic interactions with other basic residues present in the loop. Of note, the NMR solution structure of the Taspase 1 loop was solved for the isolated loop peptide and then modelled onto the crystal structure of full-length Taspase 1 where the loop was disordered and thus invisible.[4, 29] A more recent crystal structure shows the loop as one long extended helix,[44] indicating the possibility of multiple conformers. The NMR signals of the loop residues in the full-length protein closely overlap with the signals of the isolated loop, indicating that the helix-turn-helix motif is also present.[4] However, there might be a less structured and more flexible population present that might have been missed in the NMR structure due to lack of inter-residue NOE restraints, and the fraction of this population might differ between the isolated loop peptide and full-length protein. Although the extended simulations reached similar final structures, molecular recognition of a free loop peptide may be substantially more difficult to achieve as a preceding refolding and disorder-to-order transition is energy-costly.

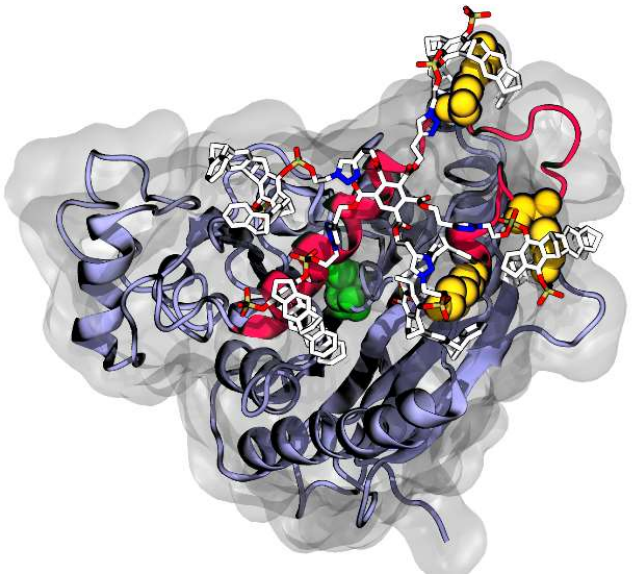


Figure 8. Molecular modelling of a kraken-like multivalent tweezer on the Taspase 1 loop. Snapshot of MD simulations placing the pentavalent tweezer MT5 on full-length Taspase 1.

Conclusion

Taken together, a novel elaborate synthesis route enabled us to develop a series of multivalent tweezer molecules with gradually increasing valency. Those were subjected to comprehensive (bio)chemical characterization regarding their inhibitory potential to target basic clusters in the Taspase 1 NLS loop. It could be demonstrated that multivalency gradually enhances the tweezers potential to interfere with the Taspase 1/Importin α interaction by direct binding to the helix-turn-helix motif. At the same time, multiavid binding also affected Taspase1's proteolytic activity. Systematic experiments involving the full-length protease corroborated a gradual increase in potency from the mono- to the pentavalent ligand revealing IC_{50}-values of less than 2 µM. Earlier approaches using sequence-controlled macromolecules[42] or bivalent guanidiniocarbonyl-pyrrole (GCP)-containing supramolecular ligands[41] were substantially less effective in targeting the Taspase 1 loop and especially in interfering with proteolytic activity.

The strict correlation between valency and inhibitory power found for the flexible protein loop in Taspase 1 is sharply contrasted by the preferred recognition of compactly folded protein regions with defined distances of single lysine residues. In these scenarios, monomers or small oligomers with rigid linkers are superior to their larger counterparts, as shown for the critical interface on microtubule-binding NDC80 and the segregase p97.[46-49] The fact that multivalency of tweezer molecules also allows to gradually increase their inhibitory potential towards the proteolytic activity of full-length Taspase 1 implies that bulky, multivalent ligands resembling a kraken most likely hinder substrate access to the active center, where they outperform smaller ligands.[41] In the future, the nature of the flexible arm will be optimized with respect to length and rigidity. Finally, introduction of an element of target specificity into our multivalent ligand design, e.g., by utilizing a distinct peptide sequence,[50] might allow to achieve selective and powerful complexation of a single flexible protein loop. As such, our study might now stimulate the rational development of multiavid inhibitors also for other disease-relevant proteins.

Acknowledgements

This work was supported by the DFG collaborative research centre1093.

Conflict of interest

The authors declare no conflict of interest.

Keywords: lysine tweezers • multivalency • protease • protein-protein interactions • supramolecular chemistry

References

[1] M. Mammen, S. K. Choi, G. M. Whitesides, *Angew Chem Int Ed Engl* **1998**, *37*, 2754-2794.
[2] O. Keskin, N. Tuncbag, A. Gursoy, *Chem Rev* **2016**, *116*, 4884-4909.
[3] R. Kubota, I. Hamachi, *Chem Soc Rev* **2015**, *44*, 4454-4471.
[4] J. van den Boom, F. Trusch, L. Hoppstock, C. Beuck, P. Bayer, *PLoS One* **2016**, *11*, e0151431.
[5] P. Mendoza-Espinosa, V. García-González, A. Moreno, R. Castillo, J. Mas-Oliva, *Mol Cell Biochem* **2009**, *330*, 105-120.
[6] K. Moritsugu, T. Terada, A. Kidera, *Journal of the American Chemical Society* **2012**, *134*, 7094-7101.
[7] G. M. Verkhivker, D. Bouzida, D. K. Gehlhaar, P. A. Rejto, S. T. Freer, P. W. Rose, *Proceedings of the National Academy of Sciences* **2003**, *100*, 5148-5153.
[8] X. Pang, H. X. Zhou, *Biophys J* **2015**, *109*, 1706-1715.
[9] N. N. Sluchanko, D. M. Bustos, *Prog Mol Biol Transl Sci* **2019**, *166*, 19-61.
[10] M. Kacirova, J. Novacek, P. Man, V. Obsilova, T. Obsil, *Biophysical journal* **2017**, *112*, 1339-1349.
[11] M. Uhart, D. Bustos, *Frontiers in Genetics* **2014**, *5*.
[12] K. V. Tugaeva, D. I. Kalacheva, R. B. Cooley, S. V. Strelkov, N. N. Sluchanko, *Scientific Reports* **2019**, *9*, 15007.
[13] J. M. Wubben, S. C. Atkinson, N. A. Borg, in *Cells, Vol. 9*, **2020**.
[14] M. Fuxreiter, *International Journal of Molecular Sciences* **2020**, *21*, 8615.
[15] V. N. Uversky, *Curr Pharm Des* **2013**, *19*, 4191-4213.
[16] M. L. Danielson, M. A. Lill, *Proteins* **2012**, *80*, 246-260.
[17] E. S. Kempner, *FEBS Lett* **1993**, *326*, 4-10.
[18] M. Amaral, D. B. Kokh, J. Bomke, A. Wegener, H. P. Buchstaller, H. M. Eggenweiler, P. Matias, C. Sirrenberg, R. C. Wade, M. Frech, *Nature Communications* **2017**, *8*, 2276.
[19] P. Tuffery, P. Derreumaux, *Journal of The Royal Society Interface* **2012**, *9*, 20-33.
[20] J. Gavenonis, B. A. Sheneman, T. R. Siegert, M. R. Eshelman, J. A. Kritzer, *Nature chemical biology* **2014**, *10*, 716-722.
[21] T. R. Siegert, M. J. Bird, K. M. Makwana, J. A. Kritzer, *Journal of the American Chemical Society* **2016**, *138*, 12876-12884.
[22] Z. Y. Zhu, S. Karlin, *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 8350-8355.
[23] C. López-Otín, J. S. Bond, *The Journal of Biological Chemistry* **2008**, *283*, 30433–30437.
[24] J. S. Bond, *The Journal of Biological Chemistry* **2019**, *294*, 1643–1651.
[25] D. Wünsch, A. Hahlbrock, S. Jung, T. Schirmeister, J. v. d. Boom, O. Schilling, S. K. Knauer, R. H. Stauber, *Oncogene* **2016**, *35*, 3351–3364.
[26] D. Y. Chen, H. Liu, S. Takeda, H.-C. Tu, S. Sasagawa, B. A. van Tine, D. Lu, E. H.-Y. Cheng, J. J.-D. Hsieh, *Cancer research* **2010**, *70*, 5358–5367.
[27] A. Gribko, A. Hahlbrock, S. Strieth, S. Becker, J. Hagemann, M. Deichelbohrer, A. Hildebrandt, N. Habtemichael, D. Wünsch, *Scientific Reports* **2017**, *7*.
[28] J. J.-D. Hsieh, E. H.-Y. Cheng, S. J. Korsmeyer, *Cell* **2003**, *115*, 293–303.
[29] J. A. Khan, B. M. Dunn, L. Tong, *Structure (London, England : 1993)* **2005**, *13*, 1443–1452.
[30] S. Takeda, D. Y. Chen, T. D. Westergard, J. K. Fisher, J. A. Rubens, S. Sasagawa, J. T. Kan, S. J. Korsmeyer, E. H.-Y. Cheng, J. J.-D. Hsieh, *Genes & Development* **2006**, *20*, 2397–2409.
[31] C. Bier, S. K. Knauer, A. Klapthor, A. Schweitzer, A. Rekik, O. H. Krämer, R. Marschalek, R. H. Stauber, *The Journal of Biological Chemistry* **2011**, *286*, 3007–3017.
[32] H. Zhou, S. Spicuglia, J. J.-D. Hsieh, D. J. Mitsiou, T. Høiby, G. J. C. Veenstra, S. J. Korsmeyer, H. G. Stunnenberg, *Molecular and Cellular Biology* **2006**, *26*, 2728-2735.
[33] S. Takeda, S. Sasagawa, T. Oyama, A. C. Searleman, T. D. Westergard, E. H. Cheng, J. J. Hsieh, *The Journal of Clinical Investigation* **2015**, *125*, 1203–1214.
[34] C. Schrenk, V. Fetz, C. Vallet, C. Heiselmayer, E. Schroder, A. Hensel, A. Hahlbrock, D. Wunsch, D. Goesswein, C. Bier, N. Habtemichael, G. Schneider, R. H. Stauber, S. K. Knauer, *J Mol Cell Biol* **2018**, *10*, 33-47.
[35] D. Y. Chen, Y. Lee, B. A. van Tine, A. C. Searleman, T. D. Westergard, H. Liu, H.-C. Tu, S. Takeda, Y. Dong, D. R. Piwnica-Worms, K. J. Oh, S. J. Korsmeyer, A. Hermone, R. Gussio, R. H. Shoemaker, E. H.-Y. Cheng, J. J.-D. Hsieh, *Cancer research* **2012**, *72*, 736–746.
[36] C. Bier, S. K. Knauer, D. Wünsch, L. Kunst, S. Scheiding, M. Kaiser, C. Ottmann, O. H. Krämer, R. H. Stauber, *The FASEB Journal* **2012**, *26*, 3421–3429.
[37] C. Bier, R. Hecht, L. Kunst, S. Scheiding, D. Wunsch, D. Goesswein, G. Schneider, O. H. Kramer, S. K. Knauer, R. H. Stauber, *PloS one* **2012**, *7*, e34142.
[38] S. K. Knauer, V. Fetz, J. Rabenstein, S. Friedl, B. Hofmann, S. Sabiani, E. Schröder, L. Kunst, E. Proschak, E. Thines, T. Kindler, G. Schneider, R. Marschalek, R. H. Stauber, C. Bier, *PloS one* **2011**, *6*.
[39] J. v. d. Boom, M. Mamić, D. Baccelliere, S. Zweerink, F. Kaschani, S. Knauer, P. Bayer, M. Kaiser, *ChemBioChem* **2014**, *15*, 2233–2237.
[40] J. Sohn, T. I. Kim, Y. H. Yoon, J. Y. Kim, S. Y. Kim, *J Clin Invest* **2003**, *111*, 121-128.
[41] A. Höing, A. Zimmermann, L. Moews, M. Killa, M. Heimann, A. Hensel, J. Voskuhl, S. K. Knauer, *ChemMedChem* **2021**.
[42] P. Pasch, A. Höing, S. Ueclue, M. Killa, J. Voskuhl, S. K. Knauer, L. Hartmann, *Chemical Communications* **2021**, *57*, 3091–3094.
[43] C. Bier, S. K. Knauer, D. Docter, G. Schneider, O. H. Kramer, R. H. Stauber, *Traffic (Copenhagen, Denmark)* **2011**, *12*, 703–714.
[44] N. Nagaratnam, S. L. Delker, R. Jernigan, T. E. Edwards, J. Snider, D. Thifault, D. Williams, B. L. Nannenga, M. Stofega, L. Sambucetti, J. J. Hsieh, A. J. Flint, P. Fromme, J. M. Martin-Garcia, *Structure* **2021**, *29*, 873-885.e875.
[45] M. Mallon, S. Dutt, T. Schrader, P. B. Crowley, *Chembiochem* **2016**, *17*, 774-783.
[46] F. Trusch, K. Kowski, K. Bravo-Rodriguez, C. Beuck, A. Sowislok, B. Wettig, A. Matena, E. Sanchez-Garcia, H. Meyer, T. Schrader, P. Bayer, *Chem Commun (Camb)* **2016**, *52*, 14141-14144.
[47] V. A. Volkov, P. J. Huis in 't Veld, M. Dogterom, A. Musacchio, *eLife* **2018**, *7*, e36764.
[48] K. Suvarna, K. Honda, M. Muroi, Y. Kondoh, H. Osada, N. Watanabe, *J Biol Chem* **2019**, *294*, 2988-2996.
[49] M. R. Hoq, F. S. Vago, K. Li, M. Kovaliov, R. J. Nicholas, D. M. Huryn, P. Wipf, W. Jiang, D. H. Thompson, *ACS Nano* **2021**, *15*, 8376-8385.
[50] A. Meiners, S. Backer, I. Hadrovic, C. Heid, C. Beuck, Y. B. Ruiz-Blanco, J. Mieres-Perez, M. Porschke, J. N. Grad, C. Vallet, D. Hoffmann, P. Bayer, E. Sanchez-Garcia, T. Schrader, S. K. Knauer, *Nat Commun* **2021**, *12*, 1505.

Entry for the Table of Contents

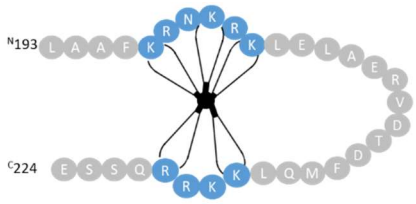


Table of Contents abstract:

The more the better: Multivalent, kraken-like supramolecular tweezers allow the recognition of a flexible protein loop in the oncologically relevant protease Taspase 1. Novel synthesis routes enabled the development of a toolbox of ligands with increasing valency that allowed not only to interfere with the Taspase 1 NLS/Importin α interaction by direct loop binding but also to affect its proteolytic activity.

3 Discussion

The disruption of the Tasp/Impα-interaction as a potential inhibition strategy for the Taspase1 activation process was first hypothesized by Bier *et al*. (Bier *et al.* 2011a). Using microscopic and immunologic methods it was shown that mutated versions of Taspase1 with a functionless NLS remained in the cytoplasm as inactive monomers (Bier *et al.* 2011a). The idea of a novel inhibition mechanism was further revisited by van den Boom *et al.* in their study about the Taspase1 loop's secondary structure (van den Boom *et al.* 2016). At that time, only a few Taspase1 inhibitors had been described and all of them aimed to inhibit the proteolytic activity of Taspase1 directly, though none of them resulted in a therapeutic product (Chen *et al.* 2012; van den Boom *et al.* 2014).

The goal of this project is the development and characterization of rationally designed supramolecular ligands to modulate Taspase1 function: either classically, by direct inhibition of the Taspase1 activity, or by using a proposed novel strategy, that involves inhibition of the Tasp/Impα-interaction, resulting in the disruption of the Taspase1 activation process.

3.1 Cationic binders targeting anionic hotspots

The first effective rational ligands, which affected the Tasp/Impα-interaction, were the macromolecular trivalent ligands **3GP** and **3GLP**, which utilized the cationic GCP-motif (Schmuck 1999). Both ligands consisted of 3 cationic GCP-groups (**G**) connected by an oligomer backbone of defined length and were equipped with a PEG. **3GLP** was supported by 3 lysines (**L**) that were positioned next to the GCP as additional binding units. While the cationic groups should serve as the binding units targeting the Taspase1 surface, the PEG was supposed to sterically block the binding site by its sheer size and dimensions.

Modelling predicted binding of the ligands to anionic clusters surrounding the Taspase1 loop. Binding of **3GP**, **3GLP,** and the unPEGgylated control compounds **3G** and **3GL** to the Taspase1 surface was investigated by SPR. To prevent effects of autoproteolysis, or even the loss of subunits during measurement, we utilized a permanently inactive monomeric mutant. While neither PEG nor monovalent compounds variants bound to Taspase1$_{D233A/T234A}$, we could give qualitative proof of binding to Taspase1$_{D233A/T234A}$ for all trivalent compounds independent of their PEGylation status.

Next, we set out to investigate the proposed effects of the PEGylated ligands **3GP** and **3GLP** on the Tasp/Impα-interaction. As expected, PEG alone did not show an effect on the Tasp/Impα-interaction without the cationic binding units. Also, the unPEGylated trivalent ligands **3G** and **3GL** showed no or only neglectable effects on the Tasp/Impα-interaction,

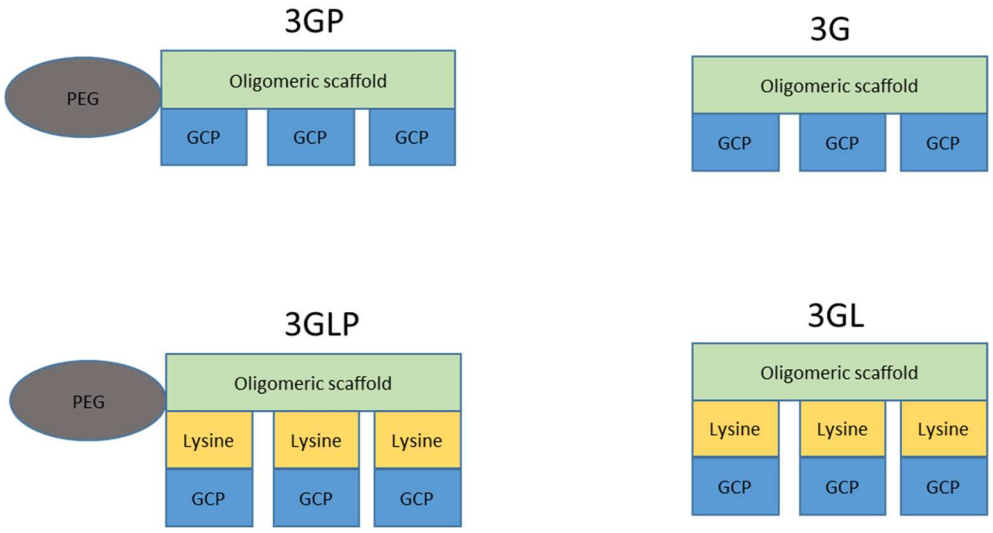


Figure 10: Schematic representation of the structure and building blocks of trivalent ligands used in this study. All compounds are trivalent (**3**) and utilize GCP (**G**) as a binding motif. **3GP** and **3GLP** are also equipped with a PEGylation (**P**). **3GLP** and **3GL** also contain lysines (**L**) as additional cationic binders.

thereby underlining the necessity of the steric shield. As previously proposed, **3GLP** was shown to be superior over **3GP** due to the introduction of additional cationic binders. With this, we validated the rational design of our compounds and described the first disruptor of the Tasp/Impα-interaction that binds to Taspase1.

Initially we hypothesized that the cationic groups might interact electrostatically with the negatively charged cell membrane to allow cellular uptake, as the GCP motif has been used to develop powerful transfection agents (Kuchelmeister *et al.* 2012; Samanta *et al.* 2016; Giese *et al.* 2020). We investigated the compounds' effects in tumor cells, but observed no effects on cell viability. Since cellular uptake decreases rapidly for structures with a molecular weight above 1 kDa, we attributed this to the compounds molecular weight (**3GLP** 4.8 kDa, **3GP** 4.4 kDa) (Matsson and Kihlberg 2017). Strategies for optimization could include downsizing of the steric shield, e.g. by reducing the size of the employed PEGylation, or variations of the oligomer backbones with altered length and rigidity. Other options for future investigations could comprise partial exchange of binding units, which might result in multiavid ligands with altered characteristics and behavior, or the addition of fluorescent groups, which would enable the investigation of cellular uptake.

Since GCP units have proven to be reliable as binding units for Taspase1, we set out to generate further supramolecular Taspase1 ligands based on this motif (Schmuck 1999). To surpass the limitations of the previous cationic ligands, the compounds were supposed to be less bulky and smaller to make them cell permeable. We generated the cationic supramolecular binder **2GC**, a bivalent (**2**) GCP-based (**G**) ligand that contains a hydrophobic cbz protecting group (**C**), along with several control compounds, each with one of the building blocks modified. **1GC** was a monovalent (**1**) variant of **2GC**. We exchanged the cbz group (**C**)

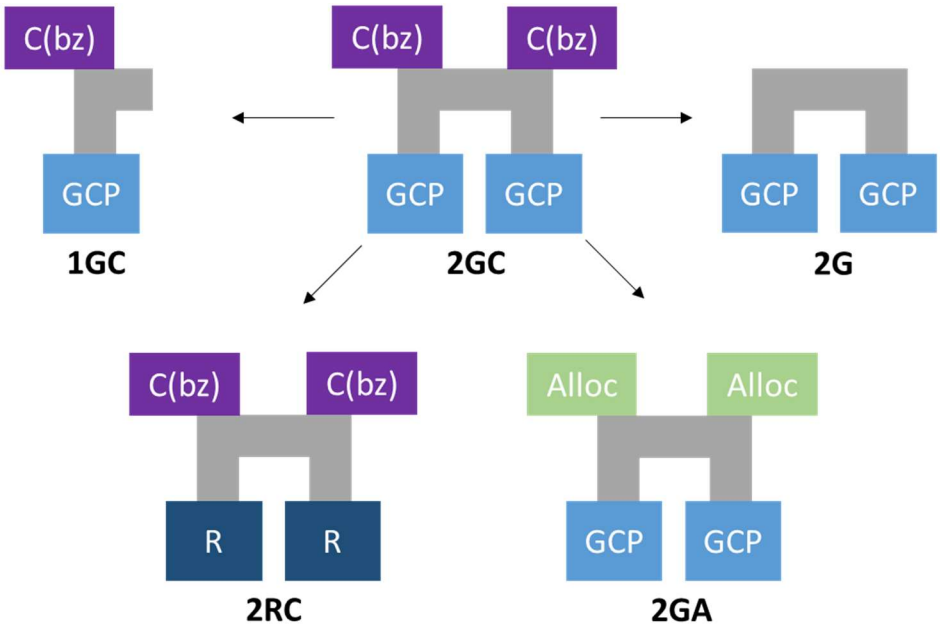


Figure 11: Schematic representation of the structure and building blocks of the compounds used in this study. The compounds are either monovalent (**1**) or bilvalent (**2**) and utilize either GCP (**G**) or arginine (**R**) as binding units. Some compounds are also equipped with hydrophobic protecting groups cbz (**C**) or alloc (**A**).

for a similar alloc protecting group (**A**) in **2GA** and completely removed the protective group in **2G**. Finally, we exchanged the unit GCP for another cationic binder, namely arginine (**R**), resulting in **2RC**.

As mentioned before, **2GC** is less bulky than previous compounds (1.7 kDa). Modelling suggested binding of **2GC** to acetic clusters directly flanking the NLS in the Taspase1 loop, resulting in steric shielding of the loop without the need for formerly used PEGylation (Pasch *et al.* 2021). As expected, pull-down studies revealed that **2GC** was a potent inhibitor of the Tasp/Impα-interaction (IC_{50} = 35 µM), while none of the control compounds did show inhibitory effects (Pasch *et al.* 2021). Since the suggested binding site in the loop is in close proximity to the active site of Taspase1, we examined possible effects on the proteolytic activity of active Taspase1 (van den Boom *et al.* 2016). However, **2GC** only showed effects in unphysiologically high concentrations and was no potent inhibitor of proteolytic activity, but more of a minor sterical block for substrate binding.

Next, we investigated the compounds effect on the viability of tumor cells (Chen *et al.* 2010; Knauer *et al.* 2011). Indeed, **2GC** showed toxic effects in Taspase1-expressing tumor cell lines (EC_{50} HeLa = 70 µM, EC_{50} A549 = 40 µM). It was notable that once again any alteration of the compounds structure diminished all effects observed.

The novel Taspase1 inhibitor **2GC** was an improvement on the previous compounds **3GLP** and **3GP** in many aspects. First of all, while **3GLP** and **3GP** did had no effect on tumor cell viability, **2GC** was toxic for the Taspase1-expressing tumor cell lines HeLa and A549. Secondly, **2GC** exhibited a lower minimum inhibitory concentration in pull-down assays. While

neither **3GLP** nor **3GP** showed any effect at 10 µM, **2GC** showed first effects in the range of 1-10 µM. Finally, the compound was more effective in pull-down assays regarding the maximum effect. While **3GLP** and **3GP** were able to decrease the Tasp/Impα-interaction to approximately 40-50 %, **2GC** decreased the Taspase1 binding to 12 % at maximum. Virtual docking studies helped understanding the differences we observed. While **3GP** and **3GLP** supposedly target acetic clusters in close proximity around the Taspase1 loop, **2GC** was predicted to bind acetic clusters inside the loop. We therefore decided to directly target the cationic NLS inside the Taspase1 loop.

3.2 Anionic binders targeting cationic clusters in the NLS

We set out to directly target the bipartite NLS in the Taspase1 loop with anionic supramolecular ligands. The BINOL-based bivalent anionic compound **11d** possessed two para-substituted phosphate groups and was synthesized from a precursor molecule using a cyclization process. (Octa-Smolin *et al.* 2018). This straightforward synthesis process was published in 2018 without application (Octa-Smolin *et al.* 2018). We focused on **11d** and two comparative compounds. The bivalent ligand **11e** had the two binding groups meta-substituted, while **11f** was a trivalent variant (Octa-Smolin *et al.* 2018).

Indeed, pull-down studies revealed an IC_{50} in the low micromolar range, and that **11d** has a significantly greater inhibitory effect on the Tasp/Impα-interaction when compared with **11e** and **11f**. *In silico* docking studies with Epitopsy and simulated annealing monte carlo simulation (SAMC) determined the energetically most favorable position for **11d** at the NLS in the Taspase1 loop (Grad *et al.* 2018; Rafieiolhosseini *et al.* 2021). We investigated binding affinity for Taspase1 (K_D = 300 nM) and for a FAM-labled Taspase1 loop (K_D = 3 µM). The differences are clearly caused by the increased structural flexibility of the loop peptide compared to the loop of the full-length protein, which is far more limited in its conformation and thus less accessible. We also attempted to map the binding site in the Taspase1 loop utilizing NMR spectroscopy, which unfortunately did not work out, but at least gave another qualitative proof of binding. The indicated binding site at the Taspase1 loop not only covers the NLS, but is also in close proximity to the Taspase1 active site. Significant effects on the proteolytic activity of wild type Taspase1 were observed *in vitro* by colorimetric cleavage assay (IC_{50} = 2 µM), in presence of other proteins by competitive FRET-assay and in tumor cells by utilization of a transfectable biosensor (Knauer *et al.* 2011; Bier *et al.* 2011b). **11d** was therefore classified as the first Taspase1 inhibitor with a dual inhibition mechanism.

The terms “dual inhibition mechanism” or “dual action inhibitor” have been used loosely in literature and are not clearly defined. In most cases, they describe inhibitors that can target two proteins at the same time (Shenoy *et al.* 2011; Poongavanam *et al.* 2014; Dumas *et al.*

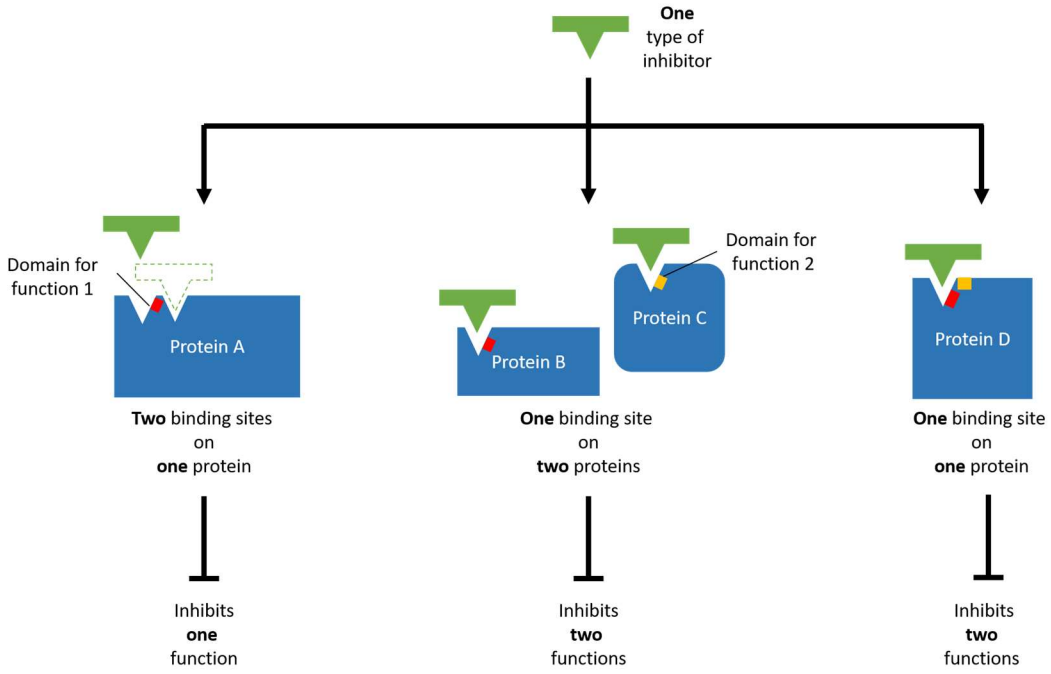


Figure 12: Schematic explaining the different use of the terms "dual inhibition" or "dual action" for inhibitors in literature. The ligand can either bind to two sites of the same protein and interfere with the same function or it can also bind sites in two distinct proteins resulting in the inhibition of two different functions or it can target one site on a protein and interfere with two different functions.

2021). For example, Gilteritinib inhibits tyrosine kinases FLT3 and AXL, both involved in leukemia (Dumas *et al.* 2021). The terms are also used to describe inhibitors that bind two different positions on the same protein leading to inhibition of the same function (Liu *et al.* 1999; Herbring *et al.* 2016; Yang *et al.* 2020). For example, Corilagin targets two different binding sites in the bacterial serine/threonine phosphatase (Stp1), which is a potential drug target in methicillin-resistant *Staphylococcus aureus (*MRSA), but each binding event results in a decrease of its enzymatic activity (Yang *et al.* 2020). However, to our knowledge, there are no inhibitors described that target two distinctive key functions of the same protein with a single molecule, as displayed by **11d**.

Supramolecular bivalent ligand **11d** surpassed previous compounds in many ways. **11d** already showed significant effects at 4 µM and the IC_{50} was determined to be as low as 6.1 µM, which is 6-7-fold lower than the IC_{50} determined for **2GC,** making **11d** the so far most effective inhibitor of the Tasp/Impα-interaction (Höing *et al.* 2021). The fluorescent properties enabled us to directly track the uptake and the ligand's location inside of cells, which has never been done for Taspase1 inhibitors before. It is also the first Taspase1 inhibitor with a dual inhibition mechanism, inhibiting the interaction with Impα, thereby interfering with its translocation and the subsequent activation process, as well as inhibiting the proteolytic activity of active Taspase1. **11d** proved to be as potent as the recently described proteolytic inhibitor Closantel sodium (IC_{50} = 1.6µM), the only Taspase1 inhibitor that is effective in a physiological range, or the most potent of the peptidyl succinimidyl peptides (IC_{50} = 3.6µM) (van den Boom *et al.* 2014; Luciano *et al.* 2021). Notably, **11d** might even surpass Closantel sodium's effect on Taspase1

activity, since the IC_{50} given by the authors differs between *in vitro* generated cell free Taspase1 (IC_{50} = 1.6µM) and recombinantly purified wild type Taspase1 from *E. coli* (IC_{50} = 3.9 µM), which will be more active than cell free Taspase1, due to partial autoproteolytic activation during the purification process (Luciano *et al.* 2021).

In our recent studies, the multivalency of supramolecular ligands always turned out to be a key factor for effective targeting of Taspase1. **3GP** and **3GLP** were trivalent, **2GC** and **11d** bivalent ligands and monovalent control compounds never allowed binding of Taspase1 or interference with the Tasp/Impα-interaction (Octa-Smolin *et al.* 2018; Pasch *et al.* 2021; Höing *et al.* 2021). We therefore conducted an investigation of the impact multivalency has in supramolecular ligands with Taspase1 and anionic ligands as examples.

The loop of Taspase1 contains two densely packed clusters of basic amino acids. A simple approach for more selective targeting could be achieved by conjugating multiple anionic binders in close proximity, thereby leading to more specific targeting and an increased multiavidity. Similar approaches have shown that increased multivalency allows more effective targeting of viral receptors when compared with monovalent inhibitors *in silico* (Schubertová *et al.* 2017). We set out to test this strategy with multivalent ligands based on the molecular tweezer, a potent lysine binder often used in multiavid supramolecular inhibitors (van der Meer *et al.* 2021; Meiners *et al.* 2021). The resulting multivalent ligands had up to five binding units (**MT1**-**MT5**). As expected, the disruptive effect on the Tasp/Impα-interaction in pull-down studies increased with the valency of the multivalent tweezers and we identified pentavalent **MT5** as the most potent disruptor. In addition, we investigated effects on the proteolytic activity of wild type Taspase1 and again, the effects correlated with the valency of the tweezer molecules with IC_{50} values in the same range (2-3 µM). Of note, the effect on wild type Taspase1 was limited to 45 % proteolytic activity for all tweezers.

We attempted to map the binding sites in the Taspase1 loop utilizing NMR spectroscopy, which did not reveal distinct binding sites, but did provide a qualitative proof of binding for all multivalent tweezers. We instead investigated a potential increase in binding affinity by fluorescence titration with a FAM-labeled Taspase1 loop. The effective range for the molecules clearly varied in accordance with their valency, since we could not determine a binding affinity for **MT1** and **MT2** in the range observed, even with a 300-fold excess of Taspase1 loop. The tweezers **MT3**, **MT4f**, and **MT4r** showed K_D values in a similar range (K_D = 2-3 µM), while **MT5** even showed a slight decrease in binding affinity (K_D = 19 µM).

Judging from our evaluation, not all inhibitory abilities increase on correlation with the valency. For example, we expected an increased binding affinity, since multiavidity of the multivalent ligands is the accumulation of individual binding affinities. This can be explained, hence every binding event restricts future binding events within the conformational and dimensional

limitations of the molecule leading to a maximum of simultaneously binding tweezers. As a result, unbound binding units do not further increase the binding affinity and therefore do not contribute to multiavidity. Molecular dynamic simulations and different binding studies, for example SPR, will be needed to further elucidate this saturational behavior . Also, the tweezers did not seem to be as potent in regard to Taspase1 activity. While **11d** had an IC_{50} of 2 µM, inhibiting Taspase1 activity almost completely at maximum effect (2.4 % substrate cleaved), the effects of all multivalent tweezers on proteolytic activity were limited (approximately 45 % substrate cleaved) in the same range. We assume that the multivalent tweezers did not completely block the active site of Taspase1, but limited the access of substrate as a sterical hindrance.

Regarding the Tasp/Impα-interaction the multivalent tweezers indeed do surpass **11d** in some aspects. **11d** reached the maximum effect at lower concentrations (10 µM 7% Taspase1 binding, **MT4f** 50 µM 17%). However, the multivalent tweezers **MT4r** and **MT5** already show significant effects at a lower concentration of 2 µM (**MT4r** 59 %, **MT5** 38 % Taspase1 binding), while **11d** was barely effective at this concentration (82 % Taspase1 binding). While the effects on proteolytic activity were limited, the multivalent tweezers might also be classified as novel dual action inhibitors for Taspase1 and a combinatory approach employing different platforms or strategies could be investigated by future studies.

3.3 Further approaches of interference with Taspase1 function

There are other potential inhibition strategies for Taspase1, which could be exploited or combined with the stated inhibitors. In another potential strategy nanoparticles could be used, which have successfully been used by van den Boom *et al.* before to interfere with Taspase1 activity (van den Boom *et al.* 2020). While recent silica nanoparticles studies focused only on adsorption, which is not a specific targeting strategy *per se*, but proved to be effective for Taspase1, future studies could incorporate supramolecular binders to enhance protein recognition and lead to increased specificity towards Taspase1 (van den Boom *et al.* 2020). This method has been extensively investigated for Survivin, a member of the “inhibitor of apoptosis” family, with numerous inhibitors (Wang *et al.* 2016; Jenkins *et al.* 2016; van der Meer *et al.* 2021). The most recent example utilized ultrasmall gold nanoparticles coated with anionic binders to increase binding affinity for Survivin and interfere with its function (van der Meer *et al.* 2021). This approach could be transferred to Taspase1 in combination with effective supramolecular ligands such as **11d**, resulting in multiavid nanoparticles with specific protein recognition and increased binding affinity by an interplay of supramolecular interaction and adsorption. Nanoparticles, which can be tagged with fluorophores to track cellular uptake and

potential intracellular effects, could also be coated by a mixed population of different Taspase1 binders to achieve multiple accumulative binding events on each protein bound.

Another popular strategy for specific protein targeting, which could be applied to Taspase1, is the development of synthetic peptides as high-affinity binders to target proteins with high specificity (Xiao *et al.* 2019; Chen *et al.* 2020b; Haugaard-Kedström *et al.* 2021). These peptides can even be generated without prior knowledge about a protein structure or computational analysis by identification of a low-affinity lead sequence and subsequent screening of point-variations (Greving *et al.* 2010). The variations with the highest affinity are then combined and the additivity of binding energies results in peptides with a high-affinity for the target. For example, this approach has resulted in a peptide targeting inflammation-associated TNFα with nanomolar affinity (Greving *et al.* 2010). Regioselective peptides can also be generated *in silico* by combination of molecular dynamics modeling and free energy calculations, which has resulted in peptides against tumor-relevant HER2, again with nanomolar affinity (Geng *et al.* 2015).

These approaches could potentially be used to target Taspase1. Depending on the strategy, different domains on the Taspase1 surface could be addressed. Peptides targeting the NLS of Taspase1 could be derived from the interface of Impα, thus shutting its nuclear import, while peptides targeting the docking head could be derived from the docking zone, thus interfering with the dimerization. The active site could be targeted with peptides mimicking the cleavage motif, which has already been demonstrated by the development of the peptidyl succinimidyl peptides (van den Boom *et al.* 2014). The peptides generated from any of those approaches could be combined with additional binding units, which has been demonstrated for Survivin (Meiners *et al.* 2021). A short peptide derived from the dimerization interface equipped with an anionic binder, resulted in an increased binding affinity, when compared with the peptide alone, and specific targeting of Survivin's NES (Meiners *et al.* 2021).

3.4 Limitations of this study

The overall structure and higher order of Taspase1 are controversial topics of debate in literature. Initial crystal structures suggested the formation of a stable dimer complex for inactive as well as active Taspase1 (Khan *et al.* 2005). However, pull-downs and gel filtrations revealed that Taspase1 is mostly found as an active αβ heterodimer in a cellular environment (Bier *et al.* 2012a). As a result, several groups have reported influence of protein concentration as well as puffer conditions on the oligomeric state of Taspase1, as the circularly permeabilized Taspase1 even seems to adapt a hexameric $(\beta\alpha)_3$ structure according to the particle size, determined by dynamic light scattering (Bier *et al.* 2012a; Nagaratnam *et al.* 2021). Also, other results suggest that the active $\alpha_2\beta_2$Taspase1 complex is not as stable as initially assumed and

its maintenance after autoproteolytic processing is not a requirement for Taspase1 activity (Bier *et al.* 2012a).

Since some hypotheses seem to oppose each other and since all currently available models of Taspase1 structure are based on mutated versions, each with different parts or aspects missing, respective models must be carefully chosen depending on study design to make reliable predictions. In this study, virtual docking and other *in silico* investigations were based on the model proposed by van den Boom *et al*., which displays the loop as a helix-turn-helix structure (Khan *et al.* 2005; van den Boom *et al.* 2016). Recently, a new model proposed by Nagaratnam *et al.* provided an alternative model for the loop in active Taspase1 structure, based on a circularly permutated version, which clashed with many previous studies (Nagaratnam *et al.* 2021). In the crystal structure model of circularly permutated Taspase1 by Nagaratnam *et al.* the loop was present as a rigid “long helical fragment” at the C-terminus (Nagaratnam *et al.* 2021). While this is certainly intriguing and might provide insight into loop dynamics in active Taspase1, it was not a suitable model for our studies, since inactive wild type Taspase1 is expressed as a single chain monomer with the loop not located at the C-terminus, but between the later subunits (Nagaratnam *et al.* 2021). Since both ends of the loop are attached to Taspase1 in close proximity, the formation of the rigid “long helical fragment” in inactive Taspase1 is unlikely to happen when compared to the helix-turn-helix structure proposed by van den Boom *et al*. (van den Boom *et al.* 2016; Nagaratnam *et al.* 2021).

In addition to that, Nagaratnam *et al.* also described loss of proteolytic activity by shortening of the “long helical fragment” and claimed that well-established active Taspase1 mutants without the loop must have been inactive, although these loop-less variants have been shown to be permanently active (Khan *et al.* 2005; van Boom 2015; van den Boom *et al.* 2020). The “long helical fragment” and the effects of its shortening might be an artefact caused by circularization. We therefore utilized the Taspase1 model with the loop as a helix-turn-helix structure for *in silico* experiments like virtual docking.

These *in silico* studies have provided insight into potential binding sites or binding dynamics. However, one area, which needs to be investigated further, is the experimental mapping of ligand binding sites on Taspase1. Crystallization could be utilized to determine a precise binding position of a ligand’s position. However crystallization of Taspase1 has been known to require high salt concentrations for stability, which might interfere with binding of supramolecular ligands (Khan *et al.* 2005; van Boom 2015). Another problem would be the flexibility of the loop, which makes it a missing structure in all crystal structure models of inactive Taspase1 (Khan *et al.* 2005). Therefore, ligands targeting the Taspase1 loop could not easily be investigated by crystallization.

NMR spectroscopy has been utilized to map the binding site of a ligand for small proteins like Survivin before and we utilized NMR to investigate binding positions of supramolecular ligands in the Taspase1 loop, but NMR is also limited to investigation of the loop as a peptide, since full-length Taspase1 is too big and leads to a high background signal (van den Boom *et al.* 2016; Vallet *et al.* 2020; Meiners *et al.* 2021). A potential technology suitable for mapping of ligand binding sites in full-length Taspase1 might be hydrogen/deuterium exchange mass spectrometry (HDX-MS) (Marciano *et al.* 2014; Hodge *et al.* 2020). In general, a protein is incubated in deuterium oxide (D_2O, also known as heavy water), leading to a rapid exchange of surface-exposed protons for deuterium (Marciano *et al.* 2014). However, areas covered by protein folding or interactions are not accessible and will be exposed by subsequent proteolytic digestion and mass spectrometry (Marciano *et al.* 2014). This method has long been established to investigate antibody/antigen-interactions or ligand/receptor-binding, but has also been used to investigate binding of inhibitors or small molecules in the recent years (Huzil *et al.* 2008; Chalmers *et al.* 2011; Lewallen *et al.* 2014).

Another highly attractive method, which is rarely used in the context of drug development is small angle x-ray scattering (SAXS) (Tuukkanen and Svergun 2014; Chen *et al.* 2020a). In recent years, SAXS has been used to investigate dynamic changes in protein structure as well as ligand binding, which makes it a good choice for investigating dynamic changes upon ligand binding in the Taspase1 loop (Tuukkanen and Svergun 2014). For example, multivalent ligands could act as cross-linkers of the two opposing helices in the Taspase1 loop or act analogously to a stapler by fixing the loop to the core of Taspase1, which would affect the flexibility and dynamic of the Taspase1 loop and could potentially interfere with binding by Impα (Bier *et al.* 2011a). These changes in protein dynamic could be investigated by SAXS, which complements well with other structural methods such as crystallography or NMR (Kikhney and Svergun 2015).

4 Conclusion and outlook

In this thesis, we investigated different strategies to interfere with the function of the cancer-relevant protease Taspase1 by using of rationally designed supramolecular ligands.

Different strategies have been devised for interfering with Taspase1 function, but so far no study has focused on the interaction with Impα, which is essential for Taspase1 activation. We therefore tried to exploit this novel inhibition mechanism and utilized different cationic or anionic supramolecular ligands as potential inhibitors. We succeeded in describing the trivalent cationic ligands **3GP** and **3GLP** based on the GCP motif as the first inhibitors that bind to Taspase1 and inhibit the interaction with Impα (Pasch *et al.* 2021). We also characterized the bivalent cationic ligand **2GC**, which was shown to be a more potent inhibitor of the Tasp/Impα-interaction and also showed toxic effects in Taspase1-expressing tumor cells (Höing *et al.* 2021).

Bivalent anionic supramolecular ligand **11d** is the first Taspase1 inhibitor to display a dual inhibition mechanism for Taspase1 (Octa-Smolin *et al.* 2018). While it is the most potent inhibitor of the Tasp/Impα-interaction so far, thereby targeting the activation process, **11d** also effectively inhibits the proteolytic activity of already active Taspase1 and binds to Taspase1 in the higher nanomolar range. Significant effects could also be shown in a cellular environment, where treatment with **11d** leads to less digestion of an intracellular biosensor for Taspase1 activity (Knauer *et al.* 2011; Bier *et al.* 2011b).

We further investigated the inhibitory effects of multivalency with anionic molecular tweezers and Taspase1 as an example (Fokkens *et al.* 2005). A set of multivalent tweezers with up to five binding units was utilized to investigate effects on Taspase1 inhibition and binding affinity. The effect of valency was most prominent regarding the Tasp/Impα-interaction and the disruptive effect increased with the valency of the ligands. The proteolytic activity of Taspase1 was also affected in correlation to the valency of the ligands, but the effects on active Taspase1 seemed to be limited. Binding studies revealed a similar K_D for binding of the Taspase1 loop for the tweezers **MT3**, **MT4f**, and **MT4r**. Binding affinity was decreased for **MT5** and could not be determined for **MT1** and **MT2**.

The information gained by utilization of these supramolecular ligands will lead to rational design of future Taspase1 inhibitors. Future studies could investigate the combinatory potential of the supramolecular ligands described to create synergistic effects. From the methodological perspective, we hope to add HDX-MS and SAXS to our arsenal to elucidate binding positions and mechanisms of the supramolecular ligands. Further studies should also focus on quantification of the Tasp/Impα-interaction in cells, potential methods could include Proximity ligation assay (PLA).

5 References

Almendro, Vanessa; Ametller, Elisabet; Garcia-Recio, Susana; Collazo, Olga; Casas, Ignasi; Auge, Josep M. et al. (2009): The role of MMP7 and its cross-talk with the FAS/FASL system during the acquisition of chemoresistance to oxaliplatin. In *PloS one* 4 (3), e4728. DOI: 10.1371/journal.pone.0004728.

Balkin, Daniel M.; Poranki, Menitha; Forester, Craig M.; Dorsey, Morna J.; Slavotinek, Anne; Pomerantz, Jason H. (2019): TASP1 mutation in a female with craniofacial anomalies, anterior segment dysgenesis, congenital immunodeficiency and macrocytic anemia. In *Molecular genetics & genomic medicine* 7 (9), e818. DOI: 10.1002/mgg3.818.

Bedessem, Baptiste; Ruphy, Stéphanie (2015): SMT or TOFT? How the two main theories of carcinogenesis are made (artificially) incompatible. In *Acta biotheoretica* 63 (3), pp. 257–267. DOI: 10.1007/s10441-015-9252-1.

Bedessem, Baptiste; Ruphy, Stphanie (2017): SMT and TOFT Integrable After All: A Reply to Bizzarri and Cucina. In *Acta biotheoretica* 65 (1), pp. 81–85. DOI: 10.1007/s10441-016-9286-z.

Bier, C.; Knauer, S. K.; Wünsch, D.; Kunst, L.; Scheiding, Sabine; Kaiser, Markus et al. (2012a): Allosteric inhibition of Taspase1′s pathobiological activity by enforced dimerization in vivo. In *The FASEB Journal* 26 (8), pp. 3421–3429. DOI: 10.1096/fj.11-202432.

Bier, Carolin; Hecht, Rouven; Kunst, Lena; Scheiding, Sabine; Wunsch, Desiree; Goesswein, Dorothee et al. (2012b): Overexpression of the catalytically impaired Taspase1 T234V or Taspase1 D233A variants does not have a dominant negative effect in T(4;11) leukemia cells. In *PloS one* 7 (5), e34142. DOI: 10.1371/journal.pone.0034142.

Bier, Carolin; Knauer, Shirley K.; Docter, Dominic; Schneider, Gunter; Kramer, Oliver H.; Stauber, Roland H. (2011a): The importin-alpha/nucleophosmin switch controls taspase1 protease function. In *Traffic (Copenhagen, Denmark)* 12 (6), pp. 703–714. DOI: 10.1111/j.1600-0854.2011.01191.x.

Bier, Carolin; Knauer, Shirley K.; Klapthor, Alexander; Schweitzer, Andrea; Rekik, Alexander; Krämer, Oliver H. et al. (2011b): Cell-based Analysis of Structure-Function Activity of Threonine Aspartase 1*. In *The Journal of Biological Chemistry* 286 (4), pp. 3007–3017. DOI: 10.1074/jbc.M110.161646.

Bizzarri, Mariano; Cucina, Alessandra (2016): SMT and TOFT: Why and How They are Opposite and Incompatible Paradigms. In *Acta biotheoretica* 64 (3), pp. 221–239. DOI: 10.1007/s10441-016-9281-4.

Bursen, Adelheid; Moritz, Sven; Gaussmann, Anne; Moritz, Sören; Dingermann, Theo; Marschalek, Rolf (2004): Interaction of AF4 wild-type and AF4.MLL fusion protein with SIAH proteins: indication for t(4;11) pathobiology? In *Oncogene* 23 (37), pp. 6237–6249. DOI: 10.1038/sj.onc.1207837.

Chalmers, Michael J.; Busby, Scott A.; Pascal, Bruce D.; West, Graham M.; Griffin, Patrick R. (2011): Differential hydrogen/deuterium exchange mass spectrometry analysis of protein-ligand interactions. In *Expert review of proteomics* 8 (1), pp. 43–59. DOI: 10.1586/epr.10.109.

Chen, David Y.; Lee, Yishan; van Tine, Brian A.; Searleman, Adam C.; Westergard, Todd D.; Liu, Han et al. (2012): A Pharmacological Inhibitor of the Protease Taspase1 Effectively Inhibits Breast and Brain Tumor Growth. In *Cancer research* 72 (3), pp. 736–746. DOI: 10.1158/0008-5472.CAN-11-2584.

Chen, David Y.; Liu, Han; Takeda, Shugaku; Tu, Ho-Chou; Sasagawa, Satoru; van Tine, Brian A. et al. (2010): Taspase1 functions as a non-oncogene addiction protease that coordinates cancer cell proliferation and apoptosis. In *Cancer research* 70 (13), pp. 5358–5367. DOI: 10.1158/0008-5472.CAN-10-0027.

Chen, Po-chia; Masiewicz, Pawel; Perez, Kathryn; Hennig, Janosch (2020a): Structure-based screening of binding affinities via small-angle X-ray scattering. In *IUCrJ* 7 (Pt 4), pp. 644–655. DOI: 10.1107/S2052252520004169.

Chen, Qiuqiang; Jia, Gang; Zhao, Xiaolei; Bao, Ying; Zhang, Yu; Ozkan, Cengiz et al. (2020b): Novel Survivin Peptides Screened With Computer Algorithm Induce Cytotoxic T Lymphocytes With Higher Cytotoxic Efficiency to Cancer Cells. In *Front. Mol. Biosci.* 7, p. 570003. DOI: 10.3389/fmolb.2020.570003.

Christiansen, Anders; Dyrskjøt, Lars (2013): The functional role of the novel biomarker karyopherin α 2 (KPNA2) in cancer. In *Cancer letters* 331 (1), pp. 18–23. DOI: 10.1016/j.canlet.2012.12.013.

Cooper, Geoffrey M. (Ed.) (2000): The Cell: A Molecular Approach. 2nd edition: Sinauer Associates.

Dong, Yiyu; van Tine, Brian A.; Oyama, Toshinao; Wang, Patricia I.; Cheng, Emily H.; Hsieh, James J. (2014): Taspase1 cleaves MLL1 to activate cyclin E for HER2/neu breast tumorigenesis. In *Cell Res* 24 (11), pp. 1354–1366. DOI: 10.1038/cr.2014.129.

Dumas, Pierre-Yves; Villacreces, Arnaud; Guitart, Amélie V.; El-habhab, Ali; Massara, Layal; Mansier, Olivier et al. (2021): Dual Inhibition of FLT3 and AXL by Gilteritinib Overcomes Hematopoietic Niche-Driven Resistance Mechanisms in FLT3-ITD Acute Myeloid Leukemia. In *Clin Cancer Res* 27 (21), pp. 6012–6025. DOI: 10.1158/1078-0432.CCR-20-3114.

Fokkens, Michael; Schrader, Thomas; Klärner, Frank-Gerrit (2005): A molecular tweezer for lysine and arginine. In *J. Am. Chem. Soc.* 127 (41), pp. 14415–14421. DOI: 10.1021/ja052806a.

Geng, Lingling; Wang, Zihua; Yang, Xiaoliang; Li, Dan; Lian, Wenxi; Xiang, Zhichu et al. (2015): Structure-based Design of Peptides with High Affinity and Specificity to HER2 Positive Tumors. In *Theranostics* 5 (10), pp. 1154–1165. DOI: 10.7150/thno.12398.

Giese, Michael; Niemeyer, Jochen; Voskuhl, Jens (2020): Guanidiniocarbonyl-Pyrroles (GCP) - 20 Years of the Schmuck Binding Motif. In *ChemPlusChem* 85 (5), pp. 985–997. DOI: 10.1002/cplu.202000142.

Grad, Jean-Noël; Gigante, Alba; Wilms, Christoph; Dybowski, Jan Nikolaj; Ohl, Ludwig; Ottmann, Christian et al. (2018): Locating Large, Flexible Ligands on Proteins. In *Journal of chemical information and modeling* 58 (2), pp. 315–327. DOI: 10.1021/acs.jcim.7b00413.

Greving, Matthew P.; Belcher, Paul E.; Diehnelt, Chris W.; Gonzalez-Moa, Maria J.; Emery, Jack; Fu, Jinglin et al. (2010): Thermodynamic additivity of sequence variations: an algorithm for creating high affinity peptides without large libraries or structural information. In *PloS one* 5 (11), e15432. DOI: 10.1371/journal.pone.0015432.

Gribko, Alena; Hahlbrock, Angelina; Strieth, Sebastian; Becker, Sven; Hagemann, Jan; Deichelbohrer, Max et al. (2017): Disease-relevant signalling-pathways in head and neck cancer: Taspase1's proteolytic activity fine-tunes TFIIA function. In *Scientific Reports* 7. DOI: 10.1038/s41598-017-14814-x.

Hanahan, Douglas (2022): Hallmarks of Cancer: New Dimensions. In *Cancer Discov* 12 (1), pp. 31–46. DOI: 10.1158/2159-8290.CD-21-1059.

Hanahan, Douglas; Weinberg, Robert A. (2000): The Hallmarks of Cancer. In *Cell* 100 (1), pp. 57–70. DOI: 10.1016/S0092-8674(00)81683-9.

Hanahan, Douglas; Weinberg, Robert A. (2011): Hallmarks of cancer: the next generation. In *Cell* 144 (5), pp. 646–674. DOI: 10.1016/j.cell.2011.02.013.

Haugaard-Kedström, Linda M.; Clemmensen, Louise S.; Sereikaite, Vita; Jin, Zeyu; Fernandes, Eduardo F. A.; Wind, Bianca et al. (2021): A High-Affinity Peptide Ligand Targeting Syntenin Inhibits Glioblastoma. In *Journal of Medicinal Chemistry* 64 (3), pp. 1423–1434. DOI: 10.1021/acs.jmedchem.0c00382.

Herbring, Valentina; Bäucker, Anja; Trowitzsch, Simon; Tampé, Robert (2016): A dual inhibition mechanism of herpesviral ICP47 arresting a conformationally thermostable TAP complex. In *Scientific Reports* 6, p. 36907. DOI: 10.1038/srep36907.

Hodge, Edgar A.; Benhaim, Mark A.; Lee, Kelly K. (2020): Bridging protein structure, dynamics, and function using hydrogen/deuterium-exchange mass spectrometry. In *Protein Science : A Publication of the Protein Society* 29 (4), pp. 843–855. DOI: 10.1002/pro.3790.

Höing, Alexander; Zimmermann, Alexander; Moews, Lisa; Killa, Matthias; Heimann, Marius; Hensel, Astrid et al. (2021): A bivalent supramolecular GCP-ligand enables blocking of the Taspase1/Importin α interaction. In *ChemMedChem*. DOI: 10.1002/cmdc.202100640.

Hsieh, James J-D; Cheng, Emily H-Y; Korsmeyer, Stanley J. (2003): Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. In *Cell* 115 (3), pp. 293–303. DOI: 10.1016/s0092-8674(03)00816-x.

Huzil, J. Torin; Chik, John K.; Slysz, Gordon W.; Freedman, Holly; Tuszynski, Jack; Taylor, Richard E. et al. (2008): A unique mode of microtubule stabilization induced by peloruside A. In *Journal of molecular biology* 378 (5), pp. 1016–1030. DOI: 10.1016/j.jmb.2008.03.026.

Jenkins, Ragini; Bandera, Yuriy P.; Daniele, Michael A.; Ledford, LeAnna L.; Tietje, Ashlee; Kelso, Andrew A. et al. (2016): Sequestering survivin to functionalized nanoparticles: a strategy to enhance apoptosis in cancer cells. In *Biomaterials science* 4 (4), pp. 614–626. DOI: 10.1039/c5bm00580a.

Khan, Javed A.; Dunn, Ben M.; Tong, Liang (2005): Crystal structure of human Taspase1, a crucial protease regulating the function of MLL. In *Structure (London, England : 1993)* 13 (10), pp. 1443–1452. DOI: 10.1016/j.str.2005.07.006.

Kikhney, Alexey G.; Svergun, Dmitri I. (2015): A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. In *FEBS letters* 589 (19 Pt A), pp. 2570–2577. DOI: 10.1016/j.febslet.2015.08.027.

Knauer, Shirley K.; Fetz, Verena; Rabenstein, Jens; Friedl, Sandra; Hofmann, Bettina; Sabiani, Samaneh et al. (2011): Bioassays to Monitor Taspase1 Function for the Identification of Pharmacogenetic Inhibitors. In *PloS one* 6 (5). DOI: 10.1371/journal.pone.0018253.

Kuchelmeister, Hannes Y.; Gutschmidt, Aljona; Tillmann, Sarah; Knauer, Shirley; Schmuck, Carsten (2012): Efficient gene delivery into cells by a surprisingly small three-armed peptide ligand. In *Chem. Sci.* 3 (4), p. 996. DOI: 10.1039/C2SC01002J.

Lange, A.; Mills, R. E.; Lange, C. J.; Stewart, M.; Devine, S. E.; Corbett, A. H. (2007): Classical Nuclear Localization Signals: Definition, Function, and Interaction With Importin Alpha. In *The Journal of Biological Chemistry* 282 (8). DOI: 10.1074/jbc.R600026200.

Lee, Jeong Tae; Chen, David Y.; Yang, Zhimou; Ramos, Alexander D.; Hsieh, James J.-D.; Bogyo, Matthew (2009): Design, syntheses, and evaluation of Taspase1 inhibitors. In

Bioorganic & medicinal chemistry letters 19 (17), pp. 5086–5090. DOI: 10.1016/j.bmcl.2009.07.045.

Lewallen, Daniel M.; Sreelatha, Anju; Dharmarajan, Venkatasubramanian; Madoux, Franck; Chase, Peter; Griffin, Patrick R. et al. (2014): Inhibiting AMPylation: a novel screen to identify the first small molecule inhibitors of protein AMPylation. In *ACS chemical biology* 9 (2), pp. 433–442. DOI: 10.1021/cb4006886.

Liotta, L. A.; Tryggvason, K.; Garbisa, S.; Hart, I.; Foltz, C. M.; Shafie, S. (1980): Metastatic potential correlates with enzymatic degradation of basement membrane collagen. In *Nature* 284 (5751), pp. 67–68. DOI: 10.1038/284067a0.

Liu, X.; Kim, C. S.; Kurbanov, F. T.; Honzatko, R. B.; Fromm, H. J. (1999): Dual mechanisms for glucose 6-phosphate inhibition of human brain hexokinase. In *The Journal of Biological Chemistry* 274 (44), pp. 31155–31159. DOI: 10.1074/jbc.274.44.31155.

López-Otín, Carlos; Bond, Judith S. (2008): Proteases: Multifunctional Enzymes in Life and Diseases. In *The Journal of Biological Chemistry* 283 (45), pp. 30433–30437. DOI: 10.1074/jbc.R800035200.

López-Otín, Carlos; Matrisian, Lynn M. (2007): Emerging roles of proteases in tumour suppression. In *Nat Rev Cancer* 7 (10), pp. 800–808.

Luciano, Vanessa; Proschak, Ewgenij; Langer, Julian D.; Knapp, Stefan; Heering, Jan; Marschalek, Rolf (2021): Closantel is an allosteric inhibitor of human Taspase1. In *iScience* 24 (12), p. 103524. DOI: 10.1016/j.isci.2021.103524.

Marciano, David P.; Dharmarajan, Venkatasubramanian; Griffin, Patrick R. (2014): HDX-MS guided drug discovery: small molecules and biopharmaceuticals. In *Current opinion in structural biology* 28, pp. 105–111. DOI: 10.1016/j.sbi.2014.08.007.

Mason, Steven D.; Joyce, Johanna A. (2011): Proteolytic networks in cancer. In *Trends in Cell Biology* 21 (4), pp. 228–237. DOI: 10.1016/j.tcb.2010.12.002.

Matsson, Pär; Kihlberg, Jan (2017): How Big Is Too Big for Cell Permeability? In *Journal of Medicinal Chemistry* 60 (5), pp. 1662–1664. DOI: 10.1021/acs.jmedchem.7b00237.

Meiners, Annika; Bäcker, Sandra; Hadrović, Inesa; Heid, Christian; Beuck, Christine; Ruiz-Blanco, Yasser B. et al. (2021): Specific inhibition of the Survivin-CRM1 interaction by peptide-modified molecular tweezers. In *Nat Commun* 12 (1), p. 1505. DOI: 10.1038/s41467-021-21753-9.

Mitsiades, N.; Yu, W. H.; Poulaki, V.; Tsokos, M.; Stamenkovic, I. (2001): Matrix metalloproteinase-7-mediated cleavage of Fas ligand protects tumor cells from chemotherapeutic drug cytotoxicity. In *Cancer research* 61 (2), pp. 577–581.

Miyamoto, Yoichi; Yamada, Kohji; Yoneda, Yoshihiro (2016): Importin α: a key molecule in nuclear transport and non-transport functions. In *J Biochem* 160 (2), pp. 69–75. DOI: 10.1093/jb/mvw036.

Montévil, Maël; Pocheville, Arnaud (2017): The Hitchhiker's Guide to the Cancer Galaxy. How two critics missed their destination. With assistance of Sapienza Università di Roma, Bizzarri Mariano, Raffaella Marandola.

Nagaratnam, Nirupa; Delker, Silvia L.; Jernigan, Rebecca; Edwards, Thomas E.; Snider, Janey; Thifault, Darren et al. (2021): Structural insights into the function of the catalytically active human Taspase1. In *Structure (London, England : 1993)* 29 (8), 873-885.e5. DOI: 10.1016/j.str.2021.03.008.

Niizuma, Hidetaka; Cheng, Emily H.; Hsieh, James J. (2015): Taspase 1: A protease with many biological surprises. In *Molecular & Cellular Oncology* 2 (4), e999513. DOI: 10.1080/23723556.2014.999513.

Niizuma, Hidetaka; Searleman, Adam C.; Takeda, Shugaku; Armstrong, Scott A.; Park, Christopher Y.; Cheng, Emily H.; Hsieh, James J. (2021): Taspase1 orchestrates fetal liver hematopoietic stem cell and vertebrae fates by cleaving TFIIA. In *JCI Insight* 6 (15). DOI: 10.1172/jci.insight.149382.

Octa-Smolin, Frescilia; van der Vight, Felix; Yadav, Rohan; Bhangu, Jasmine; Soloviova, Kateryna; Wölper, Christoph et al. (2018): Synthesis of Furan-Annelated BINOL Derivatives: Acid-Catalyzed Cyclization Induces Partial Racemization. In *The Journal of organic chemistry* 83 (23), pp. 14568–14587. DOI: 10.1021/acs.joc.8b02353.

Pasch, Peter; Höing, Alexander; Ueclue, Serap; Killa, Matthias; Voskuhl, Jens; Knauer, Shirley K.; Hartmann, Laura (2021): PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction. In *Chem. Commun.* 57 (25), pp. 3091–3094. DOI: 10.1039/D0CC07139K.

Perez-Diez, Ainhoa; Morgun, Andrey; Shulzhenko, Natalia (2007): Microarrays for cancer diagnosis and classification. In *Advances in experimental medicine and biology* 593, pp. 74–85. DOI: 10.1007/978-0-387-39978-2_8.

Pless, B.; Oehm, C.; Knauer, S.; Stauber, R. H.; Dingermann, T.; Marschalek, R. (2011): The heterodimerization domains of MLL-FYRN and FYRC--are potential target structures in t(4;11) leukemia. In *Leukemia* 25 (4), pp. 663–670. DOI: 10.1038/leu.2010.308.

Poongavanam, Vasanthanathan; Narayana Moorthy, N. S. Hari; Kongsted, Jacob (2014): Dual mechanism of HIV-1 integrase and RNase H inhibition by diketo derivatives – a computational study. In *RSC Adv* 4 (73), pp. 38672–38681. DOI: 10.1039/C4RA05728G.

Quesada, Víctor; Ordóñez, Gonzalo R.; Sánchez, Luis M.; Puente, Xose S.; López-Otín, Carlos (2009): The Degradome database: mammalian proteases and diseases of proteolysis. In *Nucleic acids research* 37 (Database issue), D239-43. DOI: 10.1093/nar/gkn570.

Rafieiolhosseini, Neda; Killa, Matthias; Tötsch, Niklas; Grad, Jean-Noël; Höing, Alexander; Ottmann, Christian; Knauer, Shirley K.; Voskuhl, Jens; Hoffmann, Daniel (2021): Computational model predicts protein binding sites of a luminescent ligand equipped with guanidiniocarbonyl-pyrrole groups.

Reunanen, Niina; Kähäri, Veli Matti (2013): Matrix Metalloproteinases in Cancer Cell Invasion. In Niina Reunanen, VeliMatti Kähäri (Eds.): Madame Curie Bioscience Database [Internet]: Landes Bioscience.

Sabiani, Samaneh; Geppert, Tim; Engelbrecht, Christian; Kowarz, Eric; Schneider, Gisbert; Marschalek, Rolf (2015): Unraveling the Activation Mechanism of Taspase1 which Controls the Oncogenic AF4-MLL Fusion Protein. In *EBioMedicine* 2 (5), pp. 386–395. DOI: 10.1016/j.ebiom.2015.04.009.

Samanta, Krishnananda; Jana, Poulami; Bäcker, Sandra; Knauer, Shirley; Schmuck, Carsten (2016): Guanidiniocarbonyl pyrrole (GCP) conjugated PAMAM-G2, a highly efficient vector for gene delivery: the importance of DNA condensation. In *Chem. Commun.* 52 (84), pp. 12446–12449. DOI: 10.1039/C6CC06404C.

Saunders, Lauren P.; Ouellette, Amy; Bandle, Russ; Chang, William Chozen; Zhou, Hongwen; Misra, Raj N. et al. (2008): Identification of small-molecule inhibitors of autotaxin that inhibit melanoma cell migration and invasion. In *Molecular cancer therapeutics* 7 (10), pp. 3352–3362. DOI: 10.1158/1535-7163.MCT-08-0463.

Schmuck, Carsten (1999): Side chain selective binding of N-acetyl-α-amino acid carboxylates by a 2-(guanidiniocarbonyl)pyrrole receptor in aqueous solvents. In *Chem. Commun.* 0 (9), pp. 843–844. DOI: 10.1039/A901126I.

Schrenk, C.; Fetz, V.; Vallet, C.; Heiselmayer, C.; Schröder, E.; Hensel, A. et al. (2018): TFIIA Transcriptional Activity Is Controlled by a 'Cleave-And-Run' Exportin-1/Taspase 1-switch. In *Journal of molecular cell biology* 10 (1). DOI: 10.1093/jmcb/mjx025.

Schubertová, Veronika; Martinez-Veracoechea, Francisco J.; Vácha, Robert (2017): Design of Multivalent Inhibitors for Preventing Cellular Uptake. In *Sci Rep* 7 (1), p. 11689. DOI: 10.1038/s41598-017-11735-7.

Shenoy, Rajesh T.; Thangamani, Saravanan; Velazquez-Campoy, Adrian; Ho, Bow; Ding, Jeak Ling; Sivaraman, J. (2011): Structural basis for dual-inhibition mechanism of a non-

classical Kazal-type serine protease inhibitor from horseshoe crab in complex with subtilisin. In *PloS one* 6 (4), e18838. DOI: 10.1371/journal.pone.0018838.

Smith, Caitlin (2013): Cancer shows strength through diversity. In *Nature* 499 (7459), pp. 505–508. DOI: 10.1038/499505a.

Sonnenschein, C.; Soto, A. M. (1999): The society of cells. Cancer and control of cell proliferation. Oxford: BIOS Scientific.

Sorsa, M. (1980): Somatic mutation theory. In *Journal of toxicology and environmental health* 6 (5-6), pp. 977–982. DOI: 10.1080/15287398009529919.

Suleiman, J.; Mundt, M.; Sampath, S.; El-Hattab, A. W. (2018): TASP1 is deleted in an infant with developmental delay, microcephaly, distinctive facial features, and multiple congenital anomalies. In *Clinical genetics* 94 (1), pp. 170–173. DOI: 10.1111/cge.13258.

Sung, Hyuna; Ferlay, Jacques; Siegel, Rebecca L.; Laversanne, Mathieu; Soerjomataram, Isabelle; Jemal, Ahmedin; Bray, Freddie (2021): Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. In *CA: A Cancer Journal for Clinicians* 71 (3), pp. 209–249. DOI: 10.3322/caac.21660.

Susanne Strand; Petra Vollmer; Lothar van den Abeelen; Daniela Gottfried; Vijay Alla; Hans Heid et al. (2004): Cleavage of CD95 by matrix metalloproteinase-7 induces apoptosis resistance in tumour cells. In *Oncogene* 23 (20), pp. 3732–3736. DOI: 10.1038/sj.onc.1207387.

Takeda, Shugaku; Chen, David Y.; Westergard, Todd D.; Fisher, Jill K.; Rubens, Jeffrey A.; Sasagawa, Satoru et al. (2006): Proteolysis of MLL family proteins is essential for Taspase1-orchestrated cell cycle progression. In *Genes & Development* 20 (17), pp. 2397–2409. DOI: 10.1101/gad.1449406.

Tuukkanen, Anne T.; Svergun, Dmitri I. (2014): Weak protein-ligand interactions studied by small-angle X-ray scattering. In *The FEBS journal* 281 (8), pp. 1974–1987. DOI: 10.1111/febs.12772.

Vallet, Cecilia; Aschmann, Dennis; Beuck, Christine; Killa, Matthias; Meiners, Annika; Mertel, Marcel et al. (2020): Funktionelle Inhibition der krebsrelevanten Interaktion von Survivin und Histon H3 mit einem Guanidiniumcarbonylpyrrol-Liganden. In *Angew. Chem.* 132 (14), pp. 5614–5619. DOI: 10.1002/ange.201915400.

van Boom, Johannes den (2015): Novel inhibitors for the protease Taspase1. With assistance of Peter Bayer. Available online at https://duepublico2.uni-due.de/receive/duepublico_mods_00035785.

van den Boom, J.; Mamić, M.; Baccelliere, D.; Zweerink, S.; Kaschani, F.; Knauer, S. K. et al. (2014): Peptidyl Succinimidyl Peptides as Taspase 1 Inhibitors. In *ChemBioChem* 15 (15), pp. 2233–2237. DOI: 10.1002/cbic.201402108.

van den Boom, Johannes; Hensel, Astrid; Trusch, Franziska; Matena, Anja; Siemer, Svenja; Guel, Désirée et al. (2020): The other side of the corona: nanoparticles inhibit the protease taspase1 in a size-dependent manner. In *Nanoscale*. DOI: 10.1039/D0NR01631D.

van den Boom, Johannes; Trusch, Franziska; Hoppstock, Lukas; Beuck, Christine; Bayer, Peter (2016): Structural Characterization of the Loop at the Alpha-Subunit C-Terminus of the Mixed Lineage Leukemia Protein Activating Protease Taspase1. In *PloS one* 11 (3), e0151431. DOI: 10.1371/journal.pone.0151431.

van der Meer, Selina Beatrice; Hadrovic, Inesa; Meiners, Annika; Loza, Kateryna; Heggen, Marc; Knauer, Shirley K. et al. (2021): New Tools to Probe the Protein Surface: Ultrasmall Gold Nanoparticles Carry Amino Acid Binders. In *The journal of physical chemistry. B* 125 (1), pp. 115–127. DOI: 10.1021/acs.jpcb.0c09846.

Wan, Xue-Mei; Zhou, Xue-Lei; Du, Yong-Jun; Shen, Hui; Yang, Zhengxia; Ding, Yanhong (2021): TASP1 Promotes Proliferation and Migration in Gastric Cancer via EMT and AKT/P-AKT Pathway. In *Journal of Immunology Research* 2021, p. 5521325. DOI: 10.1155/2021/5521325.

Wang, Fengting; Li, Pan; Shao, Yuan; Li, Yanyan; Zhang, Kai; Li, Miaomiao et al. (2020): Site-specific proteolytic cleavage prevents ubiquitination and degradation of human REV3L, the catalytic subunit of DNA polymerase ζ. In *Nucleic acids research* 48 (7), pp. 3619–3637. DOI: 10.1093/nar/gkaa096.

Wang, Zhongqiu; Tong, Mingmin; Chen, Xiao; Hu, Shouyou; Yang, Zhijian; Zhang, Yu et al. (2016): Survivin-targeted nanoparticles for pancreatic tumor imaging in mouse model. In *Nanomedicine : nanotechnology, biology, and medicine* 12 (6), pp. 1651–1661. DOI: 10.1016/j.nano.2016.02.008.

World Health Organization (2021): Press release N° 294. World Cancer Day: Breast cancer overtakes lung cancer in terms of number of new cancer cases worldwide. IARC showcases key research projects to address breast cancer. Available online at https://www.iarc.who.int/wp-content/uploads/2021/02/pr294_E.pdf.

Wünsch, D.; Fetz, V.; Heider, D.; Tenzer, S.; Bier, C.; Kunst, L. et al. (2012): Chemico-genetic strategies to inhibit the leukemic potential of threonine aspartase-1. In *Blood Cancer Journal* 2 (6), e77. DOI: 10.1038/bcj.2012.22.

Wünsch, D.; Hahlbrock, A.; Jung, S.; Schirmeister, T.; van den Boom, J.; Schilling, O. et al. (2016): Taspase1: a 'misunderstood' protease with translational cancer relevance. In *Oncogene* 35 (26), pp. 3351–3364. DOI: 10.1038/onc.2015.436.

Xiao, Wenwu; Ma, Weijie; Wei, Sixi; Li, Qianping; Liu, Ruiwu; Carney, Randy P. et al. (2019): High-affinity peptide ligand LXY30 for targeting α3β1 integrin in non-small cell lung cancer. In *J Hematol Oncol* 12 (1), p. 56. DOI: 10.1186/s13045-019-0740-7.

Xu, Darui; Farmer, Alicia; Chook, Yuh Min (2010): Recognition of nuclear targeting signals by Karyopherin-β proteins. In *Current opinion in structural biology* 20 (6), pp. 782–790. DOI: 10.1016/j.sbi.2010.09.008.

Yang, Yanan; Wang, Xiyan; Gao, Yawen; Niu, Xiaodi (2020): Insight into the Dual Inhibition Mechanism of Corilagin against MRSA Serine/Threonine Phosphatase (Stp1) by Molecular Modeling. In *ACS Omega* 5 (51), pp. 32959–32968. DOI: 10.1021/acsomega.0c03955.

Zheng, R. P.; WANG, W. E.I.; Wei, C. D. (2015): Bortezomib inhibits cell proliferation in prostate cancer. In *Experimental and Therapeutic Medicine* 10 (3), pp. 1219–1223. DOI: 10.3892/etm.2015.2617.

Zhou, H.; Spicuglia, S.; Hsieh, J.; Mitsiou, D.; Høiby, T.; Veenstra, G. et al. (2006): Uncleaved TFIIA Is a Substrate for Taspase 1 and Active in Transcription. In *Molecular and cellular biology* 26 (7). DOI: 10.1128/MCB.26.7.2728-2735.2006.

6 Appendix

6.1 Publication I: Supplementary information

Supporting Information

PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction

Peter Pasch[a†], Alexander Höing[b†], Serap Ueclue[a], Matthias Killa[c], Jens Voskuhl[c], Shirley K. Knauer[b*] and Laura Hartmann[a*]

a. Department for Organic Chemistry and Macromolecular Chemistry, Heinrich Heine University Düsseldorf, Universitätsstraße 1, Düsseldorf 40225, Germany.
b. Department for Molecular Biology II, Center of Medical Biotechnology (ZMB), University Duisburg-Essen, Universitätsstrasse 5, Essen, 45117, Germany.
c. Faculty of Chemistry (Organic Chemistry), University of Duisburg-Essen, Universitätsstraße 7, 45117 Essen, Germany.

†These authors contributed equally.
* Corresponding authors.

Materials:

Diethyl ether (with BHT as inhibitor, ≥ 99.8%), triisopropylsilane (TIPS) (98%), concentrated hydrochloric acid (pa), acetic anhydride (pa) and formic acid (pa) were purchased from Sigma Aldrich. N,N-Diisopropylethylamine (DIPEA) (≥ 99%) was purchased from Carl Roth. N,N-Dimethylformamide (DMF) (99.8%, for peptide synthesis), piperidine (99%) were obtained from Acros Organics. Dichloromethane (DCM) (99.99%), ethyl acetate (analytical reagent grade) and 1,4-dioxane (analytic reagent grade) were purchased from Fisher Scientific. Acetonitrile was purchased from AppliChem. Trifluoroacetic acid (TFA) (99%), (benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), and triethylsilane (analytical reagent grade) were purchased from Fluorochem. TentaGel® S RAM (Rink Amide) and TentaGel® PAP resins (loading: 0.23 mmol / g) were purchased from RAPP Polymer. N_α-Fmoc-N_ε-Boc-L-lysine (≥98.0%) was purchased from Iris Biotech. N_α-Fmoc-N_β-Boc-L-2,3-diaminopropionic acid (≥98.0%) was purchased from TCI. Polyethylene glycol 3000 was purchased from Merck.

Analytical Methods:

Preparative Reversed Phase- High Pressure Liquid Chromatography (prep RP-HPLC)

An Agilent 1260 Infinity device was used to purify the oligo(amidoamines), which is coupled to a variable wavelength detector (VWD) (set to 214 nm) and an automated fraction collector. The RP HPLC column, CAPCELL PAK C18 (20 x 250 mm, 5 µm), was used. The mobile phases A and B were H_2O and acetonitrile, each containing 0.1 vol% formic acid. The flow rate was set at 15 ml/min.

Reversed Phase- High Pressure Liquid Chromatography- Mass Spectrometry (RP-HPLC-MS)/Electron Spray Ionization- Mass Spectrometry (ESI-MS)

RP-HPLC-MS was carried out on an Agilent 1260 Infinity instrument coupled to a variable wavelength detector (VWD) (set to 214 nm) and a 6120 Quadrupole LC/MS containing an Electrospray Ionization (ESI) source (operation mode positive, m/z range from 200 to 2000). A MZ-AquaPerfect C18 (3.0 × 50 mm, 3 µm) RP column from Mz-Analysentechnik was used. As eluent system water/acetonitrile containing 0.1 vol% formic acid was applied. The mobile phases A and B were: System A) H_2O/acetonitrile (95/5, *v*/*v*); System B) H_2O / acetonitrile (5/95, v/v). The samples were analyzed at a flow rate of 0.4 ml/min using a linear gradient,

starting with 100% of system A) and reaching 100% system B) within 30 min. The temperature of the column room was set to 40 °C. All purities were determined using the OpenLab ChemStation software for LC/MS from Agilent Technologies.

Electron Spray Ionization- Mass Spectrometry (ESI-MS) measurements were performed with the above mentioned ESI source and quadrupole detector.

Ultra High Resolution - Mass Spectrometry (UHR-MS)

UHR-MS measurements were performed with a Bruker UHR-QTOF maXis 4G instrument with a direct inlet via syringe pump, an ESI source and a quadrupole followed by a Time of Flight (QTOF) mass analyzer.

Matrix-Assisted Laser Desorption Ionization-Time of Flight–Mass Spectrometry

(MALDI-TOF-MS) Compounds were detected using a Bruker MALDI-TOF Ultraflex I system with 2,5-dihydroxybenzoic acid (DHB) as matrix. The matrix to compound ratio of was 10:1. Spectra were acquired for reflector mode for a m/z range 2000-20000. The reflector mode was calibrated using a protein mixture.

Nuclear Magnetic Resonance Spectroscopy (NMR)

The ^{1}H-NMR spectra were recorded on a Bruker Avanace III 600 (600 MHz). These spectra were evaluated according to the following scheme: (frequency in MHz, deuterated solvent), chemical shift in ppm (multiplicity, coupling constant, integral, signal assignment). The chemical shift is given in relation to the ^{1}H signals of the deuterated solvents used (D_2O: 4.79 ppm). The multiplicities of the signals were abbreviated as follows: s (singlet), d (doublet), t (triplet), m (multiplet).

Freeze dryer

The final oligomers were lyophilized with an Alpha 1-4 LD plus instrument from Martin Christ Freeze Dryers GmbH. The drying method was set to -40 °C and 0.1 mbar.

Docking

Maestro 11.5 Schroedinger was used for the images.

DAPGCP, DAPLysGCP and PEG700 were used for the docking. The molecules were prepared with LigPrep. A model of the Taspase1 crystal structure extended by a NMR based structure of the loop (Taspase1_40-420_van_den_Boom [1][2]) was used for the grids.

A grid around the amino acids Arg190/201 Lys 225/218 with a size of 36 A (Loop), Asp233 with a size of 15 A, Asp337 with a size of 15 A and Glu207 with a size of 10 A were generated with glide grid generator.

The prepared molecules and grids were used for Docking. The method was XP (extra precise) and the sampling was flexible. The following conditions were chosen: sample nitrogen inversions, bias sampling of torsions for amides.

Visualisation of Ligand and Loop

The three DapLysGCPs were dragged to the Asp233, Asp337 and Glu207. The PEG3000 was coiled by hand and put at the loop. After that a minimization was performed.

Cloning

The plasmid for the inactive Taspase1$_{D233/T234A}$ mutant was generated as previously described [3].

The gene for Importinα was amplified from a "pc3DNA-Importinα-HA" plasmid and the ends modified via PCR (Forward primer: CAGGGGCCCTCCACCAACGAGAATGCTAAT, Reverse primer: TTCGGATCCTTAGAGAAAGTTAAAGGTCCC). The gene, now with overhangs for ApaI/BamHI digestion, was cloned in a blunt pJET1.2 vector (Thermo Fisher) according to the CloneJET PCR cloning kit (Thermo Fisher). After transfection of *E. coli* NEB-10ß (New England BioLabs), the plasmid was amplified using the NucleoBond Xtra Midi kit (Macherey-Nagel). The sequence for Importinα was then ApaI/BamHI cloned into a modified pET-41b vector containing an N-terminal GST tag and a PreScission protease cleavage site (GeneArt). The plasmid was then again amplified using *E. coli* NEB10-ß and isolated with the NucleoBond Xtra Midi kit. The sequence was verified by sequencing.

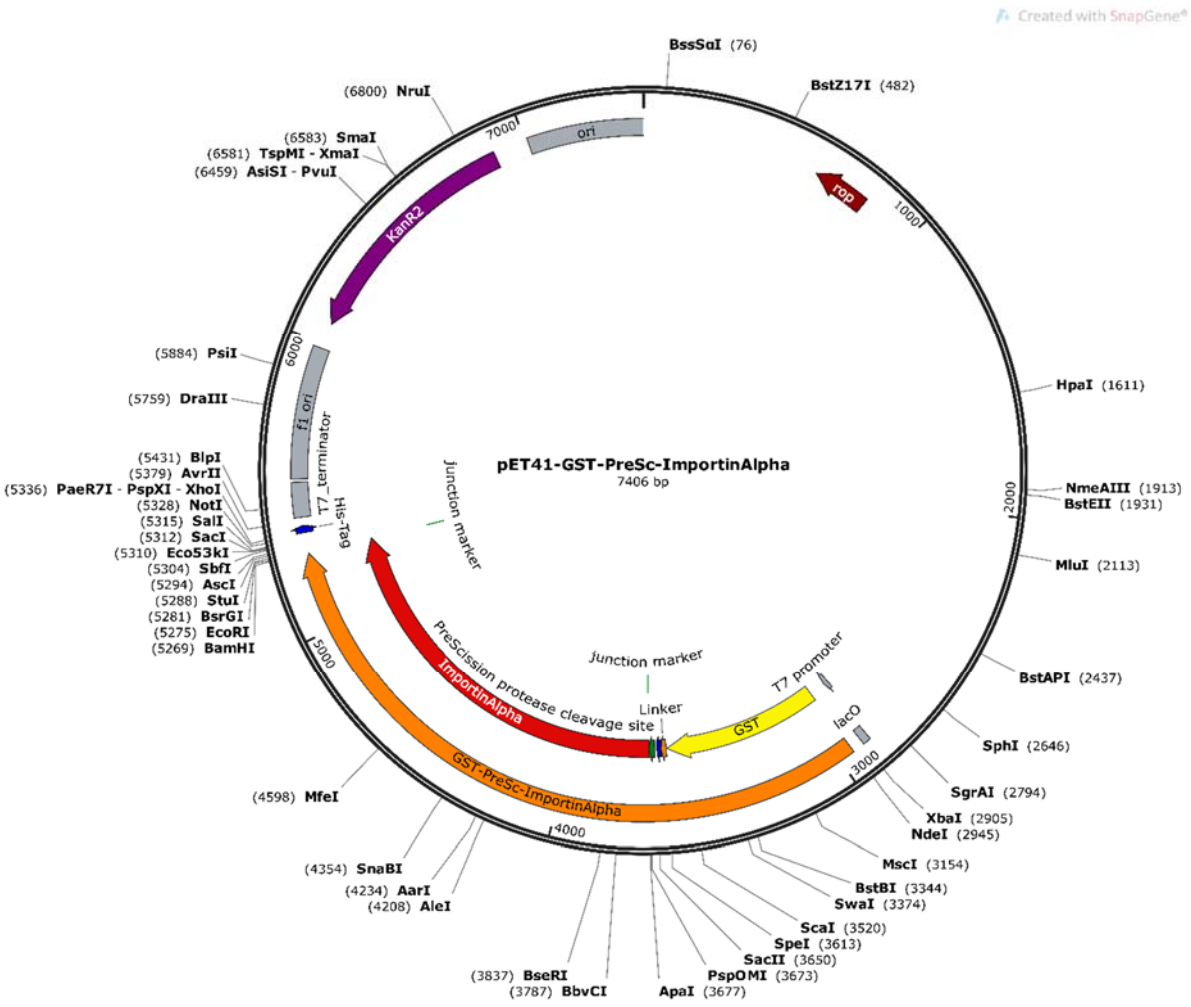


Figure S 1: Map of the plasmide pET41-GST-PreScn-Importinα generated by cloning. The map was created using "Gene Construction Kit" (Texco Biosoftware) and visualized with "Snap Gene Viewer" (GSL Biotech).

Purification of recombinant proteins

pET22-Taspase1$_{D233A/T234}$-His was expressed in *E. coli* BL21 (DE3). The cells were lysed using ultrasonic sheering and enzymatic lysis by lysozyme. The protein was purified using the His tag for affinity chromatography with a HisTrap FF (GE Healthcare). After imidazole elution the Tasapse1-His containing fractions were pooled and loaded onto a Superdex 200 HiLoad16/600 column (GE Healthcare) for size exclusion chromatography.

pET41-GST-PreScn-Importinα was expressed in *E. coli* BL21 (DE3), the cells were lysed using sonication and enzymatic lysis by lysozyme and the soluble fraction obtained with centrifugation and filtration. The protein was purified using the GST tag for affinity chromatography with a GSTrap 4B (GE Healthcare). After glutathion elution, the GST-Importinα containing fractions were pooled and loaded onto a Superdex 200 HiLoad16/600 column (GE Healthcare) for size exclusion chromatography. GST-Importinα containing fractions were pooled, frozen in liquid nitrogen and stored at – 20°C.

Pull-down assay

For this assay, all solutions were prepared with Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) containing 0,1 % (v/v) Triton X-100 (Carl Roth) and 1 mM DTT (Carl Roth) (PBST), all incubation steps were carried out at 4 °C to preserve the proteins, all centrifugation steps were carried out at 400 xG and samples taken for later analysis were mixed with 5x sample buffer and heated to 95 °C for 5 Min. 50 µM Glutathione Sepharose 4 B (Merck) were transferred to a Spin Column (IBA Lifescience), equilibrated with 500 µL PBST followed by centrifugation. 500 µL 2,5 µM GST-Importinα were added to the column, a sample for the "Input" fraction was taken and the column then incubated for 2 h on a rotator. Unbound protein was then removed by three washing steps with PBST followed by centrifugation. 500 µL 2,2 µM inactive Taspase1-His with the respective concentration of compound were pre-incubated for 1 h on a rotator and a sample for the "Input" fraction was taken. The free binding sites on the column were blocked with 1 % (w/v) BSA (Carl Roth) in PBST for 30 Min on a rotator. The blocking solution was removed from the column by centrifugation for 1 Min. After that, the inactive Taspase1-His pre-incubated with the compound was added to the column and allowed to bind for 1 h on a rotator. A sample from the "Unbound" fraction was taken and unbound protein was then removed by three washing steps with PBST followed by centrifugation for 1 Min. 500 µL 1x sample Buffer were added to the column and heated to 95 °C for 10 Min. The proteins were eluted by centrifugation for 2 Min.

SDS-PAGE and Western Blotting

For these assays, we used the standard recipes for SDS-PAGE according to Laemmli [4] and for Western Blotting according to Towbin [5]. For SDS-PAGE, Tris-glycine gels with 10 % (v/v) acrylamide in the stacking gel and 4 %(v/v) acrylamide in the separating gel were cast according. For the electrophoresis, we used the TetraCell system (BioRad) set to 200 V for 45 Min. The proteins were then transferred to a nitrocellulose membrane using a wet blot tank (Peqlab) set to 360 mA for 90 Min at 4 °C. To detect the different proteins, the membrane was first reversibly stained with Ponceau S (AppliChem) and then cut between the protein bands according to the was cut according to the Spectra Multicolor Broad Range Protein Ladder (Thermo Fisher). Free binding sites were blocked with 5 % (w/v) powdered milk (Carl Roth) in Tris buffered saline with Tween-20 (TBST) (Carl Roth) for 30 Min at room temperature. After that, the membranes were incubated with the respective primary antibodies rabbit anti-Taspase1 1:2000 (sc-85945,Santa Cruz) and mouse anti-Karyopherinα2 1:1000 (sc-55538, Santa Cruz) in 5 % (w/v) powdered milk in TBST for 1 h at room temperature. Unbound antibodies were removed by three washing steps with TBST. The membranes were incubated with the respective secondary antibodies donkey anti-rabbit HPR-coupled 1:10000 (NA934,

GE Healthcare) and sheep anti-mouse HPR-coupled 1:10000 (NXA931, GE Healthcare) in 5 % w/v) powdered milk with TBST for 1 h. Unbound antibodies were removed by four washing steps in TBST. For the detection of chemiluminescence, we used Pierce ECL Plus Western Blotting Substrate (Thermo Fisher) and the ChemidocImaging System (BioRad).

The signal was quantified with “Fiji” [6]. If necessary, the signal of Taspase1 in the eluted fraction was corrected for the Taspase1 stuck to the column without Importinα. To correct possible loading differences the signal of Taspase1 in the eluted fraction was then normalized for the signal of Importinα in the eluted fractions. The data was evaluated using “Origin2019” (OriginLab).

Toxicity Assay

1 x 10^4 cells were cultured in Corning 96 Well microplates (Sigma-Aldrich) in 100 µl Dulbecco’s modified eagle medium (DMEM) (Thermo Fisher Scientific) supplied with 10 % (v/v) fetal calf serum (FCS) (Life Technologies GmbH), Antibiotic-Antimycotic (Life Technologies GmbH) and the respective compound concentration. The cells were then cultivated at 37 °C and 5 % CO_2 for 24 h. After that, the compound-containing medium was removed, and cells were washed with PBS once. 100 µL fresh DMEM with 10 % FCS and Antibiotic-Antimycotic were added to each well. After the Following addition of 20 µL Cell Titer Aqueous One (Promega), absorption at 490 nm was recorded with the plate reader Promega Glow Max (Promega) after 30 min of incubation. Since the compounds were dissolved in water, the results were then normalized to a water treated control to correct for the dilution of the media. The data are the mean of at least three replicates ± standard deviation.

Isothermal titration calorimetry (ITC)

ITC was performed with MicroCal ITC2000 (Malvern Pananalytical). The proteins were rebuffered five times into a tenfold volume of assay buffer and then concentrated using Vivaspin 6, 10000 MWCO (Sartorius) at 4900 xg and 4 °C. The rebuffered samples were degassed immediately before ITC with MicroCal ThermoVac (Malvern Pananalytical). The experimental setup was 18 injections of 2.0 µL with 1-1.5 mM ligand to 200 µL 70-80 µM inactive Taspase1-His with an injection time of 4 s and 180 s spacing between each injection at 25 °C and with constant stirring at 750 rpm. The first injection set to 0.4 µL to remove air and mixed reactants from the tip. For each experiment, we performed a ligand-to-buffer (LtB) titration as well as a buffer-to-protein (BtP) titration with the same experimental setup to correct for possible heat of dilution introduced by the ligand or protein. The data was analysed using MicroCal Analysis (OriginLab). The first injection peak was discarded and the isotherms of the LtB and BtP controls were subtracted from the experimental isotherm.

At this time, ITC experiments were not successful (see titration curves below). At this point we attribute this to the challenge in stabilizing Taspase1 at high concentrations which requires high ionic strength, which in turn is expected to affect interaction of the GCP units.

Surface Plasmon Resonance (SPR)

The measurements were performed on a Biacore X100 from GE Healthcare Life Sciences, Uppsala, Sweden. The sensograms were recorded with the Biacore X100 Control Software and evaluated with the Biacore X100 Evaluation Software.

For the measurements a C1 sensor chip from GE Healthcare Life Science was used. Before immobilization the sensor chip surface was activated by twofold injection of 0.1 M glycin-NaOH + 0.3 % Triton X 100, pH 12 and followed by washing with HBS-P^+ buffer. Taspase1 was immobilized on the sensor chip surface on flow cell 2 via carbodiimide chemistry by the wizard

template for immobilization. Therefore, a 126 µM stock solution of Taspase1 in PBS buffer was diluted in 10 mM acetate buffer (pH 5.5, GE Healthcare) to get a final protein concentration of 1.26 µM. For flow cell 2 an immobilization level of 1391,6 RU was reached. The flow cell 1 was blocked by a solution of ethanolamine (1 M, pH 8.5, GE Healthcare) and an immobilization level of 1.6 RU was obtained. As running buffer HBS-P$^+$ buffer (pH 7.4, GE Healthcare) at a flow rate of 5 µl min^{-1} was used.

After immobilization the system was primed with the running buffer and two startup cycles were performed. The PEGylated and non-PEGylated GCP macromolecules were injected in concentrations of 0.10 – 400 µM in HBS-P$^+$ buffer with a dilution factor of 2. A flow rate at 5 µL min^{-1} and the contact and dissociation time were 120 s, respectively 180 s were used. After each measurement, the sensor chip was regenerated by injection of 0.1 M arginine in HBS-P$^+$ buffer at a flow rate of 5 µL min^{-1} with a contact time of 90 s to ensure that all sample was washed out and to achieve a stable baseline for the following measurements. For each sample the measurements were repeated three times.

To test for reproducibility, the assay was performed on a second chip prepared as described above giving an immobilization level of 1282,6 RU for flow cell 2 and an immobilization level of 144,6 RU for flow cell 1. Due to different immobilization levels, different absolute values are determined that are, however, in the same range [µM] and show the same trend for the different ligands as observed for the first chip. These measurements were repeated two times.

Measurements for monovalent ligands as well as only PEG were performed on the second chip and repeated two times.

Macromolecule Synthesis:

Synthesis EDS and GCP

(4-((2-(2-(2-aminoethoxy)ethoxy)ethyl)- amino)-4-oxobutanoic)[7] as well as N-Boc-protected 5-(guanidinocarbonyl)-1H-pyrrole-2-carboxylic acid (GCP) as triethylamine salt [8] were synthesized according to literature procedures. The free acid of the GCP was obtained by crystallization from methanol.

General

Oligomer synthesis were carried out manually in 10 ml polypropylene syringe reactors with a polyethylene frit and a Luer stopper from Multisyntech GmbH. All oligomers were synthesized on the TentaGel® S RAM (Rink Amide) or TentaGel® PAP resin with a loading of 0.23 mmol/g. Batch size of all oligomers was 0.15 mM.

Fmoc cleavage

The resin was swollen in DCM for 30 min and subsequently washed three times with DMF. Secondly the Fmoc protecting group of the resin as well as from the coupled building blocks or amino acids was cleaved by means of a 25% solution of piperidine in DMF achieving an amine end group. The deprotection was carried out twice for 20 min. Afterwards the resin was washed 10 times with DMF.

Coupling protocol

First the resin was swollen in DCM for 30 min and then washed three times with DMF. The Fmoc protecting group has to be removed before further couplings!

For the building block, amino acid or GCP (5 eq.), 5 eq. PyBOP and 10 eq. DIPEA were dissolved in 5 ml DMF, drawn into the reactor syringe and shaken for 90 min, followed by washing ten times with DMF. A double coupling, adding fresh building block and coupling reagents, was performed each time the GCP motif was coupled.

Capping of N-terminal primary amine

With scaffold completion of the oligomer the N-terminal amine group was acetylated with 8 ml of acetic anhydride, shaking for 20 min. After that, the resin was washed 5 times with DMF.

Boc-Cleavage

For Boc-deprotection, 6 ml of a 4 M HCl in dioxane solution (2 ml HCL conc. and 4 ml dioxane) was drawn into the reactor syringe and shaken for 10 min. Afterwards the reaction mixture was washed 3 times with dioxane and again 6 ml fresh 4 M HCl dioxane solution was drawn into the syringe and shaken for 20 min. Subsequently the solution was removed and the resin washed three times alternately with dioxane and DCM. To neutralize the resin, a 10 volume percent ice-cold DIPEA DCM solution was drawn up twice and shaken for 10 min. Last the resin was washed alternately three times with dioxane and DCM and finally 10 times with DMF.

Cleavage from solid phase

The oligomers were cleaved from the TentaGel® S RAM resin by drawing up a solution of 5 vol% triisopropylsilane (TIPS) and 95 vol% TFA into the syringe and shaking for 1.5 hours. The TentaGel® PAP cleavage was achieved with TFA/thioanisole (95/5, *v*/*v*) for 24 hours at room temperature.

Afterwards the solution was placed in ice cooled diethyl ether. The resulting precipitate was centrifuged off and the supernatant was decanted off. The pellet was washed 3 times with diethyl ether.

The product was dried and dissolved in MilliQ water. The entire solution was collected in a falcon tube and freeze-dried to isolate the product. Subsequently, the products were purified by means of preparative HPLC. Due to the purification by preparative HPLC and the added 0.1 vol% formic acid in the mobile phases, the structures are present as formates. The number of formates was quantified by ^{1}H NMR spectra. Further information can be found in the respective ^{1}H NMR data.

Analytical data for macromolecules:

*Macromolecule **3G***

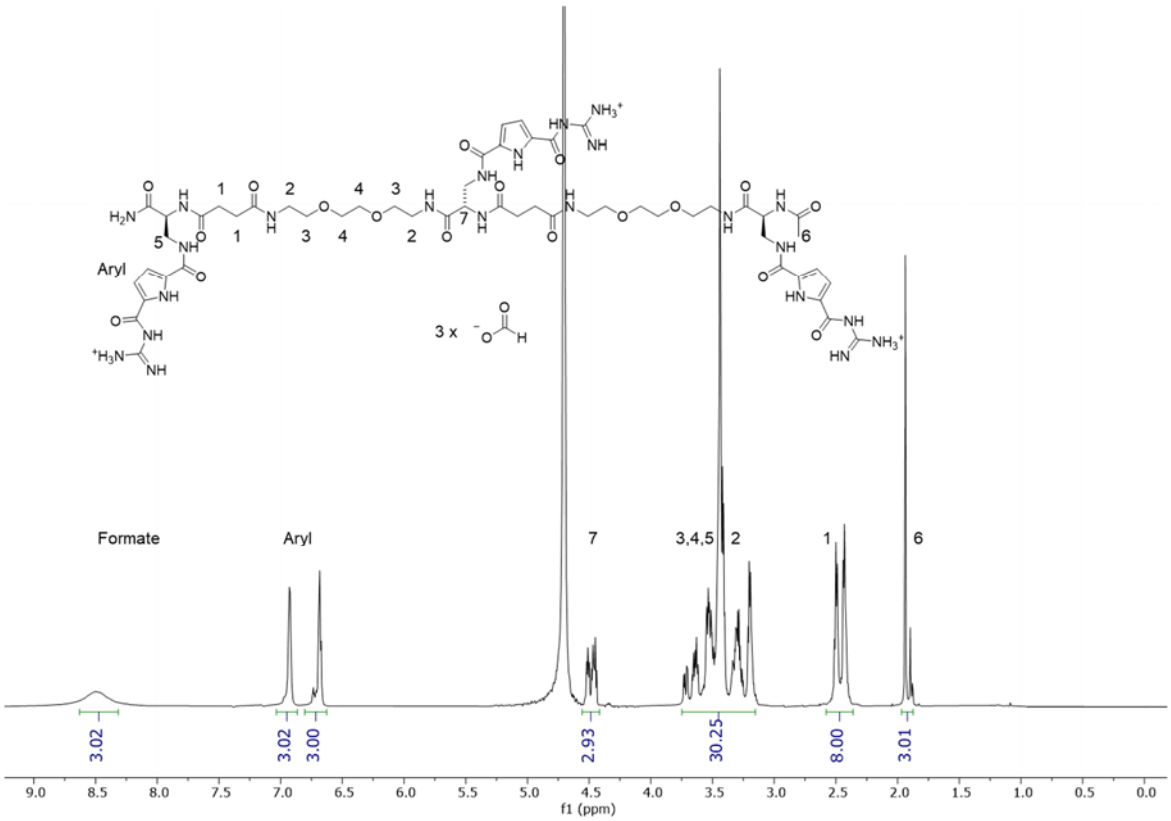


*Figure S 2: 600 MHz ^{1}H NMR spectrum of **3G** as formate salt in D_2O at 25°C.*

1H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.25 to 8.75 signal of the formate (s, 3H, **3G** is present with three formate anions), 7.09-6.85 (m, 3H, Ar-H), 6.79-6.24 (m, 3H, Ar-H), 4.50-4.40 (m, 3H, H7), 3.51-3.16 (m, 30H, H2-H5), 2.63-2.37 (m, 8H, H1), 1.94 (s, 3H, H6, second small signal cannot be assigned).

The effective molar mass for **3G** with three formates is 1450.4 g/mol.

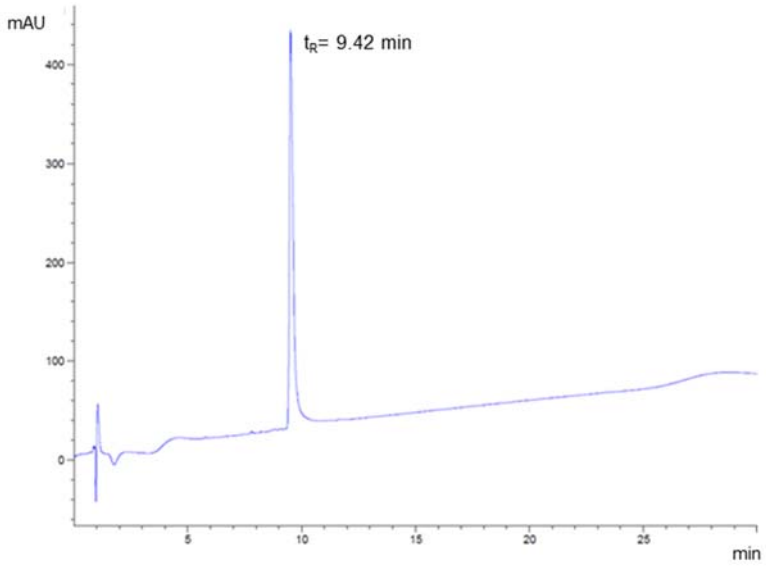


*Figure S 3: **3G** detected with relative purities >95% by RP-HPLC analysis (linear gradient from 5 – 50 vol% eluent H_20/acetonitrile) in 30 min at 40 °C.*

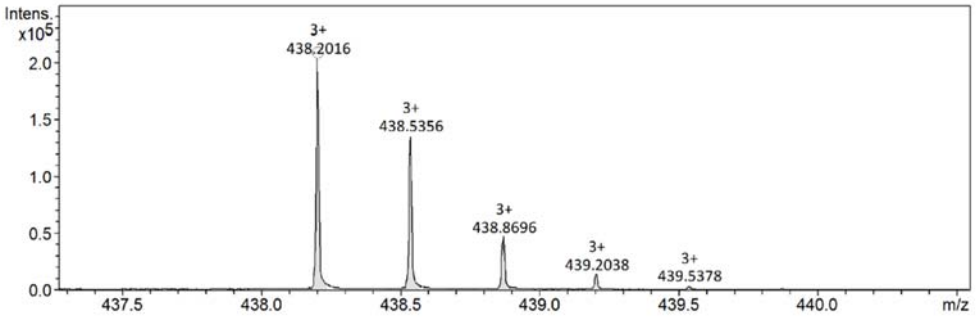


*Figure S 4: HR-ESI-MS of **3G**.*

HR-ESI-MS: for $C_{52}H_{80}N_{23}O_{18}$ m/z $[M+3H]^{3+}$ calcd.: 438.2012, found: 438.2016, mass accuracy -0.9 ppm.

*Macromolecule **3GL***

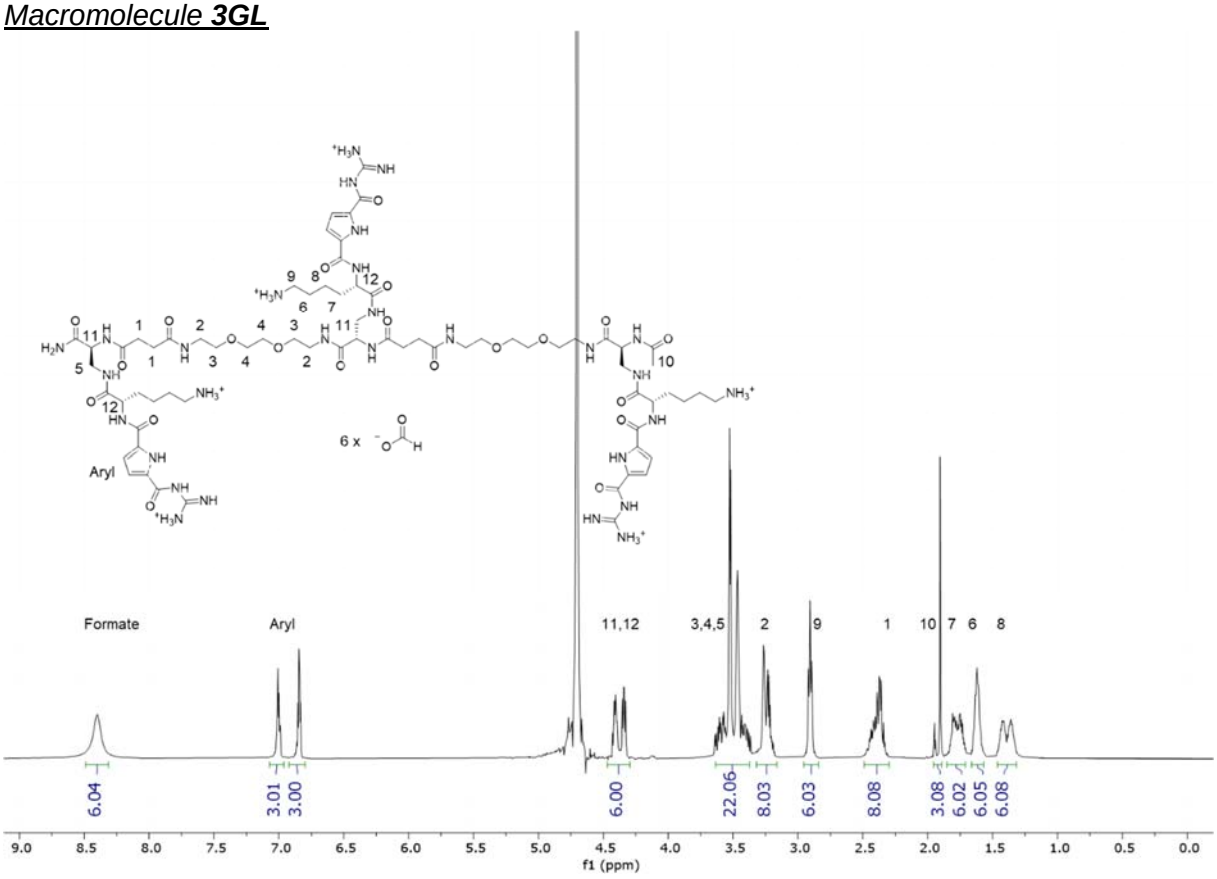


*Figure S 5: 600 MHz ^{1}H NMR spectrum of **3GL** as formate salt in D_2O at 25°C*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.30 to 8.50 signal of the formate (s, 6H, **3GL** is present with six formate anions), 7.12-6.92 (m, 3H, Ar-H), 6.87-6.79 (m, 3H, Ar-H), 4.48-4.27 (m, 6H, H11, H12), 3.69-3.38 (m, 22H, H3-H5), 3.33-3.19 (m, 8H, H2), 2.91-2.82 (m, 6H, H9), 2.49-2.27 (m, 8H, H1), 1.91 (s, 3H, H10, second small signal cannot be assigned), 1.82-1.71 (m, 6H, H7), 1.68-1.59 (m, 6H, H6), 1.45-1.31 (m, 6H, H8).

The effective molar mass for **3GL** with six formates is 1971.7 g/mol.

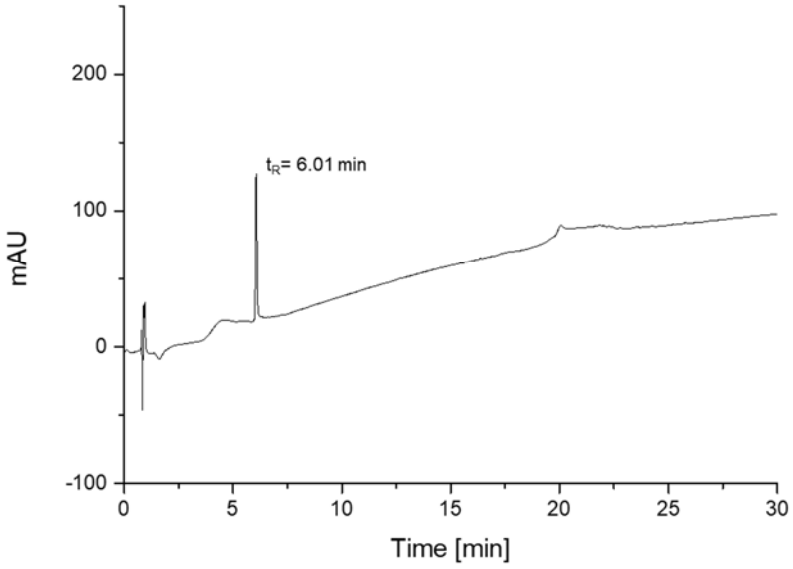


*Figure S 6: **3GL** detected with relative purities >95% by RP-HPLC analysis (linear gradient from 5 – 50 vol% eluent H_20/acetonitrile) in 30 min at 40 °C.*

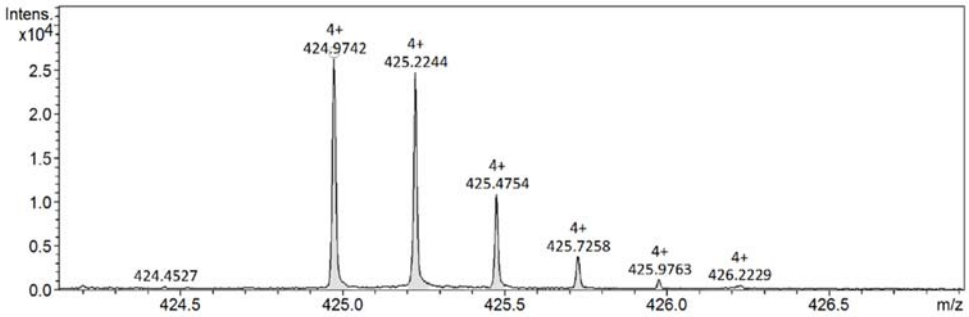


*Figure S 7: HR-ESI-MS of **3GL**.*

HR-ESI-MS: for $C_{70}H_{117}N_{29}O_{21}$ m/z $[M+4H]^{4+}$ calcd.: 424.9739, found: 424.9742, mass accuracy -0.6 ppm.

*Macromolecule **3GP***

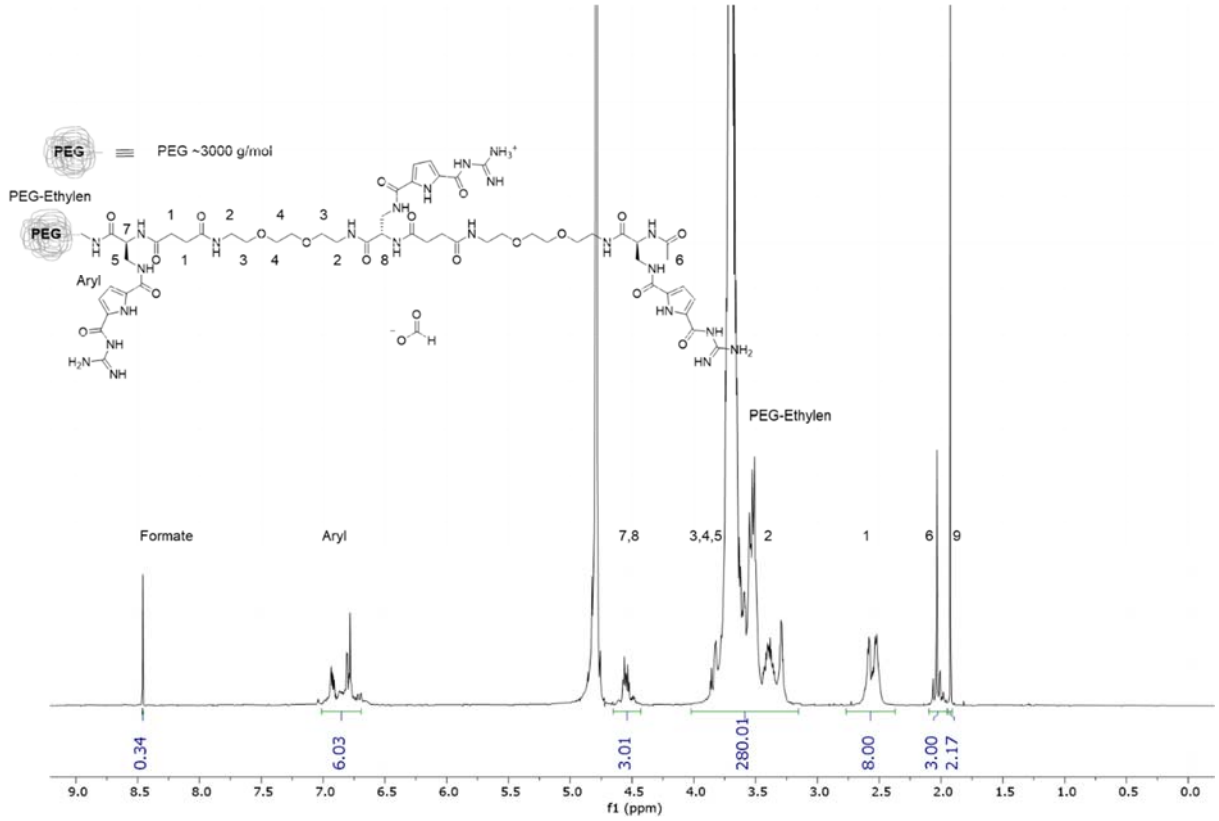


*Figure S 8: 600 MHz ^{1}H NMR spectrum of **3GP** as formate salt in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.40 to 8.50 signal of the formate (s, 0.34H, **3GP** is present at an average of 0.34 formate anion per molecule), 7.08-6.65 (m, 6H, Ar-H), 4.70-4.39 (m, 3H, H7, H8), 4.06-3.19 (m, 280H, HPEG, H2-H5), 2.78-2.36 (m, 8H, H1), 2.11-1.97 (m, 3H, H6), 1.94 (s, 2H, H9, signal of the end group of the PEG chain).

The effective molar mass for **3GP** with formate is 4413.4 g/mol.

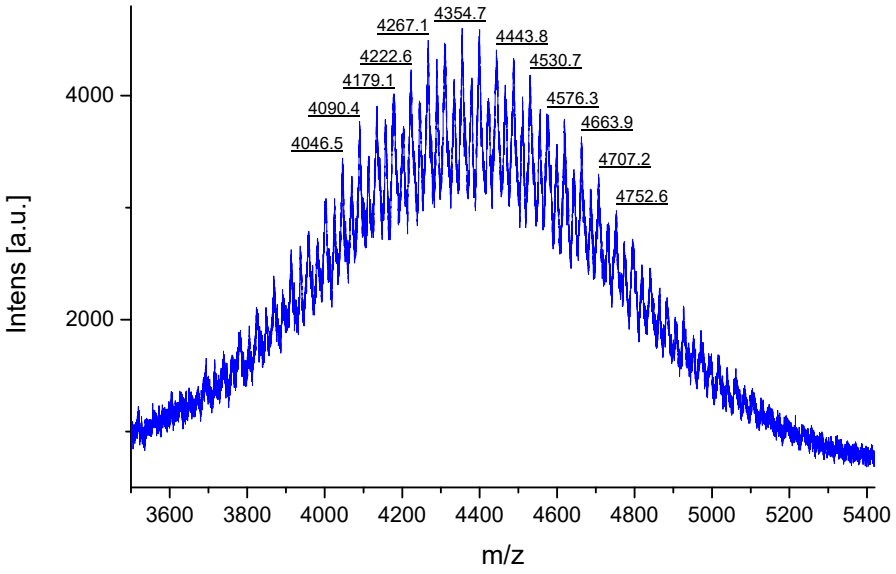


*Figure S 9: MALDI-TOF-MS of **3GP** in a m/z range using DHB as matrix in a compound to matrix ratio of 1:5.*

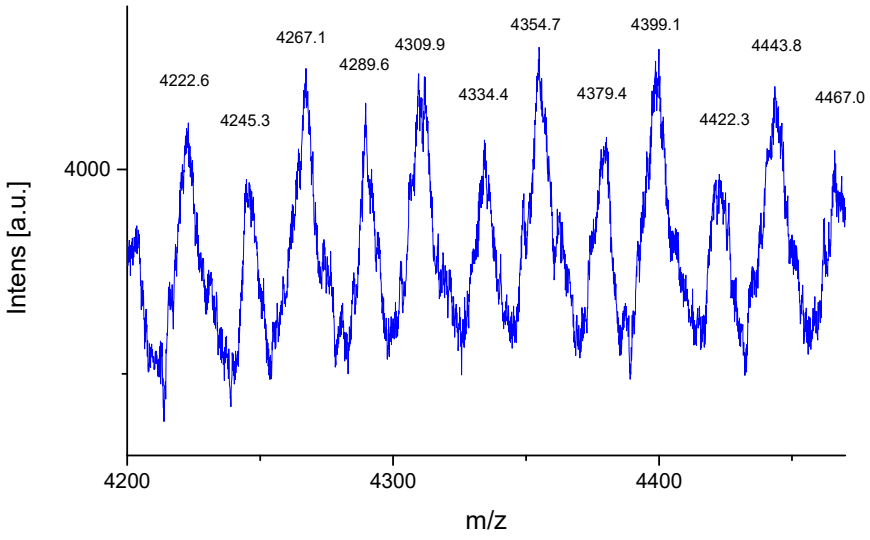


*Figure S 10: Detailed view on MALDI-TOF-MS of **3GP**, focusing on PEG-repeating units (every second signal corresponds to one PEG unit more, intermediate signal corresponds to one additional sodium ion).*

Mass analysis MALDI-TOF-MS of **3GP**: m/z found 4399.1- 3106.8 (PEG-Part) = 1292.3

(m/z calcd. Oligomer Part: 1311.6).

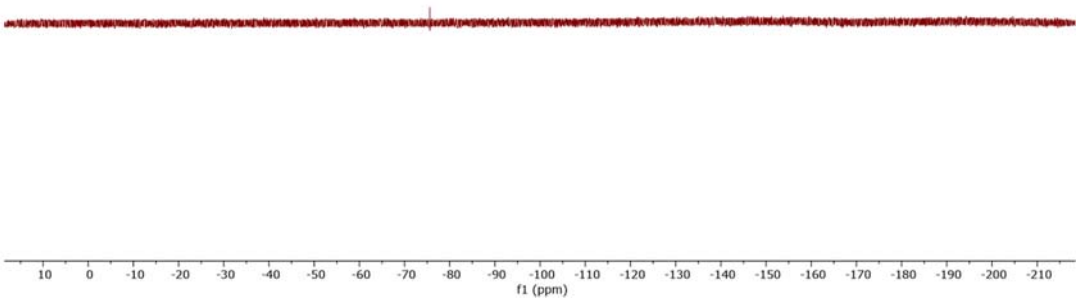


*Figure S 11: Exemplary 600 MHz ^{19}F NMR spectrum of **3GP** in D_2O at 25°C showing that no TFA counterions are present.*

*Macromolecule **3GLP***

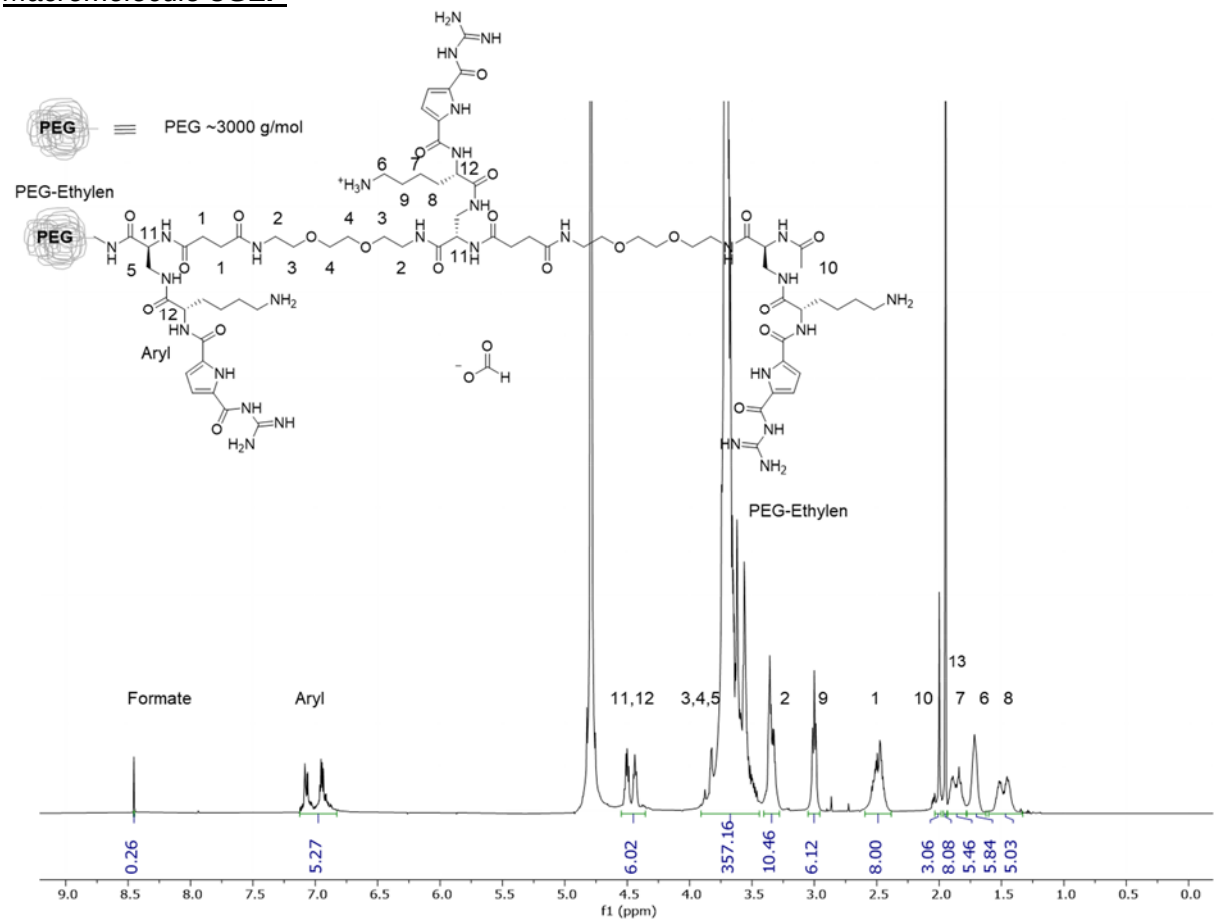


*Figure S 12: 600 MHz ^{1}H NMR spectrum of **3GLP** as formate salt in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.40 to 8.50 signal of the formate (s, 0.26H, **3GLP** is present at an average of 0.26 formate anion per molecule), 7.18-6.79 (m, 6H, Ar-H), 4.57-4.32 (m, 6H, H11, H12), 3.82-3.46 (m, 357H, HPEG, H3-H5), 3.41-3.27 (m, 10H, H2), 3.08-2.90 (m, 6H, H9), 2.61-2.31 (m, 8H, H1), 2.04 (m, 3H, H10). 2.00-1.93 (m, H13, signal of the end group of the PEG chain) 1.90-1.78 (m, 6H, H7), 1.76-1.62 (m, 6H, H6), 1.59-1.30 (m, 6H, H8).

The effective molar mass for **3GLP** with formate is 4818.8 g/mol.

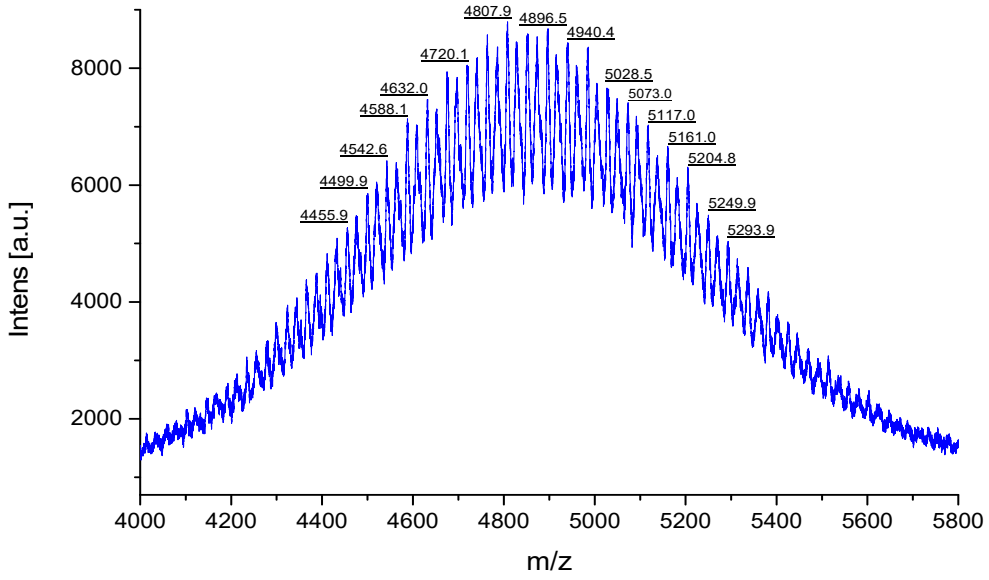


*Figure S 13: MALDI-TOF-MS of **3GLP** in a m/z range using DHB as matrix in a compound to matrix ratio of 1:10.*

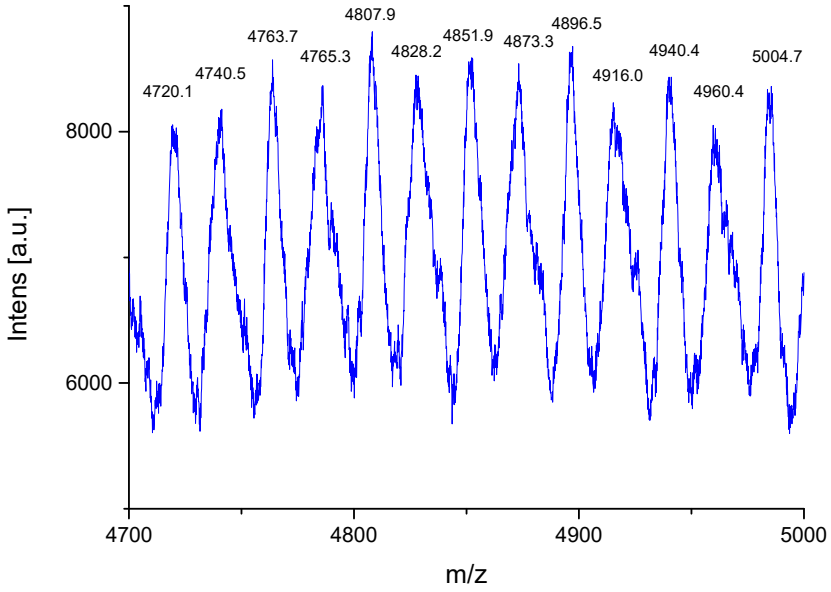


*Figure S 14: Detailed view on MALDI-TOF-MS of **3GLP**, focusing on PEG-repeating units (every second signal corresponds to one PEG unit more, intermediate signal corresponds to one additional sodium ion).*

Mass analysis MALDI-TOF-MS of **3GLP**: m/z found 4807.9- 3106.8 (PEG-Part) = 1701.2

(m/z calcd. Oligomer Part: 1695.7).

*Macromolecule **G***

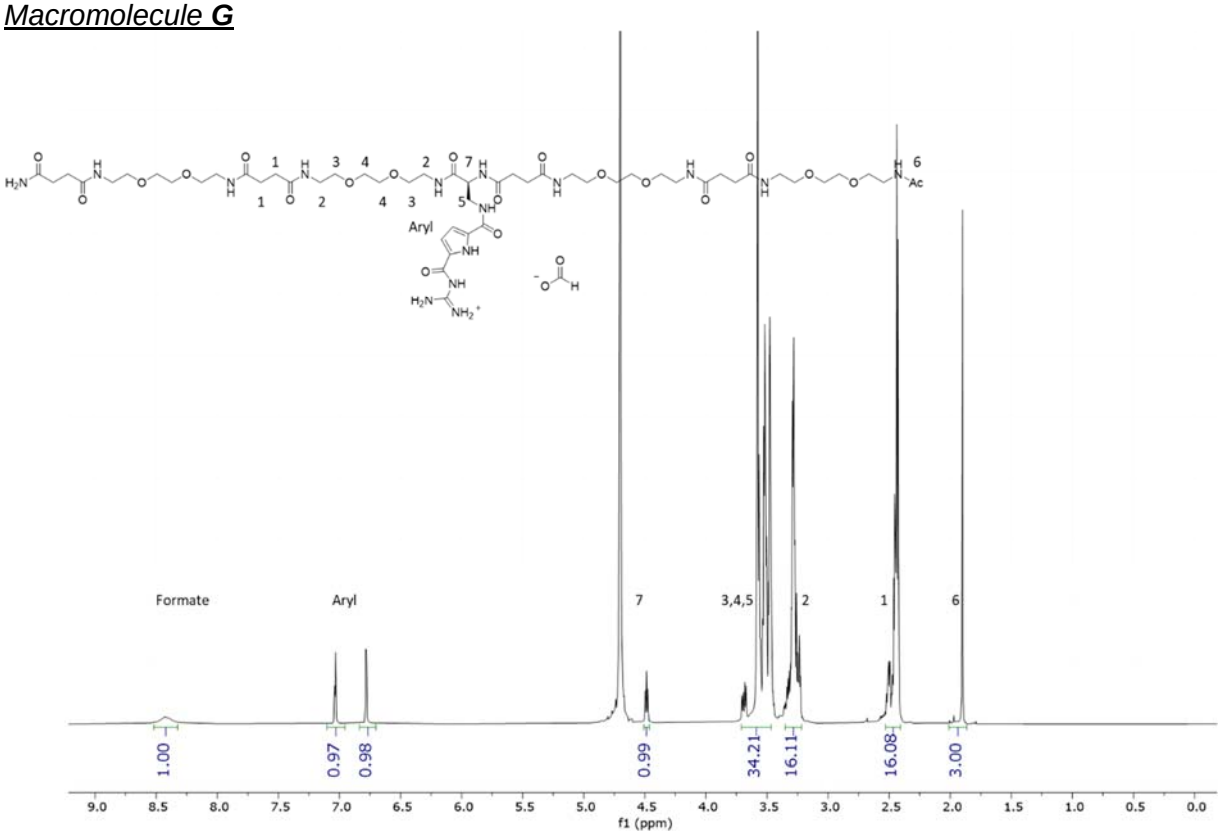


*Figure S 15: 600 MHz ^{1}H NMR spectrum of **G** as formate salt in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.30 to 8.50 signal of the formate (s, 1H, **G** is present with one formate anion), 7.08 (s, 1H, Ar-H), 6.78 (s, 1H, Ar-H), 4.52-4.44 (m, 1H, H7), 3.69-3.41 (m, 34H, H3-H5), 3.35-3.23 (m,16H, H2), 2.56-2.39 (m, 16H, H1), 1.92 (s, 3H, H6).

The effective molar mass for **G** with one formate is 1291.42 g/mol.

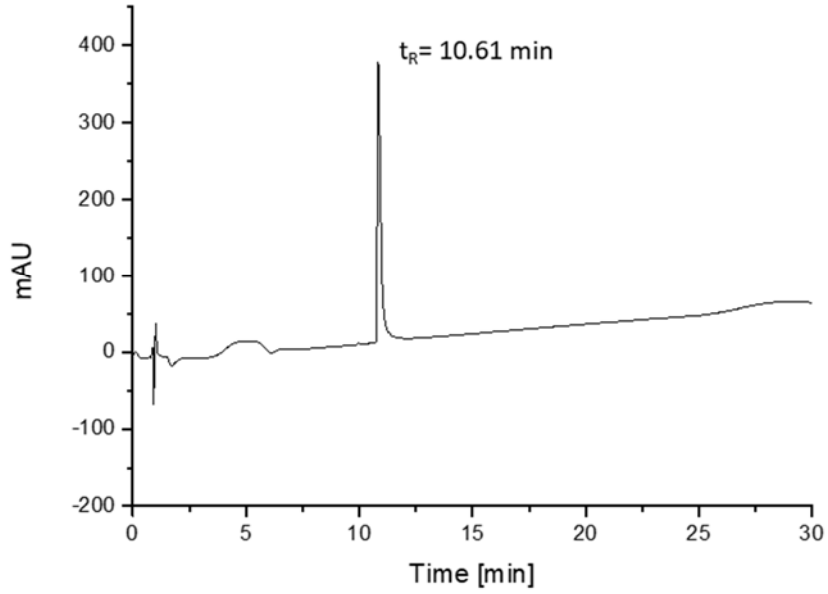


*Figure S 16: **G** detected with relative purities >95% by RP-HPLC analysis (linear gradient from 5 – 50 vol% eluent H_20/acetonitrile) in 30 min at 40 °C.*

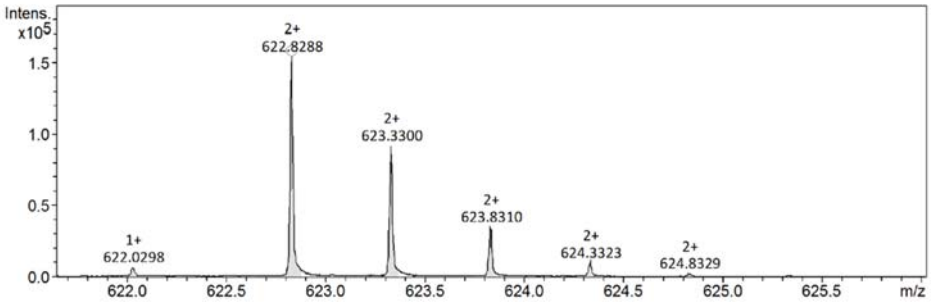


*Figure S 17: HR-ESI-MS of **G**.*

HR-ESI-MS: for $C_{52}H_{99}N_{15}O_{20}$ m/z $[M+2H]^{2+}$ calcd.: 622.8277, found: 622.8288, mass accuracy -1.7 ppm.

*Macromolecule **GL***

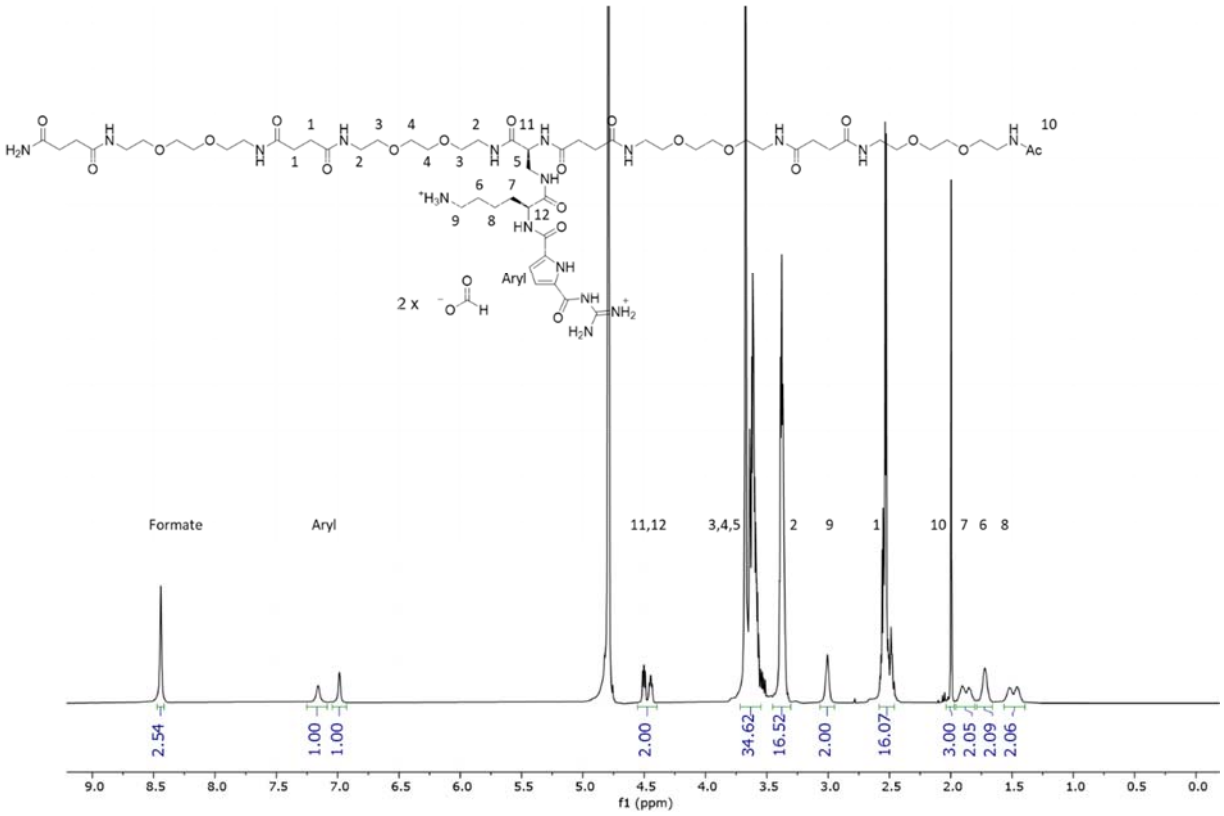


*Figure S 18: 600 MHz ^{1}H NMR spectrum of **GL** as formate salt in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = In the range from 8.40 to 8.50 signal of the formate (s, 2H, **GL** is present with two formate anions) amide, guanidino functionalities occur, 7.17 (s, 1H, Ar-H), 7.01 (s, 1H, Ar-H), 4.56-4.40 (m, 2H, H11, H12), 3.73-3.54 (m, 34H, H3-H5), 3.36-2.27 (m,16H, H2), 3.07-2.95 (m, 2H), H9), 2.59-2.44 (m, 16H, H1), 2.01 (s, 3H, H10), 1.93-1.80 (m, 2H, H7), 1.76 (s, 2H, H6), 1.57-1.40 (m, 2H, H8).

The effective molar mass for **GL** with two formates is 1457.7 g/mol.

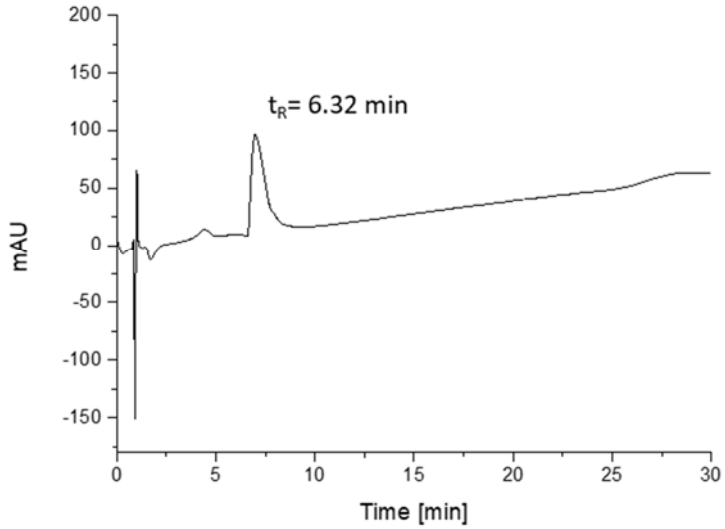


*Figure S 19: **GL** detected with relative purities >95% by RP-HPLC analysis (linear gradient from 5 – 50 vol% eluent H_20/acetonitrile) in 30 min at 40 °C.*

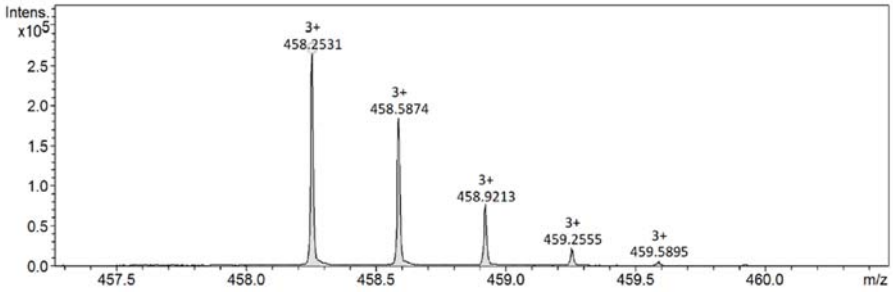


*Figure S 20: HR-ESI-MS of **1GL**.*

HR-ESI-MS: for $C_{58}H_{101}N_{17}O_{21}$ m/z $[M+3H]^{2+}$ calcd.: 458.2525, found: 458.2531, mass accuracy -1.2 ppm.

PEG

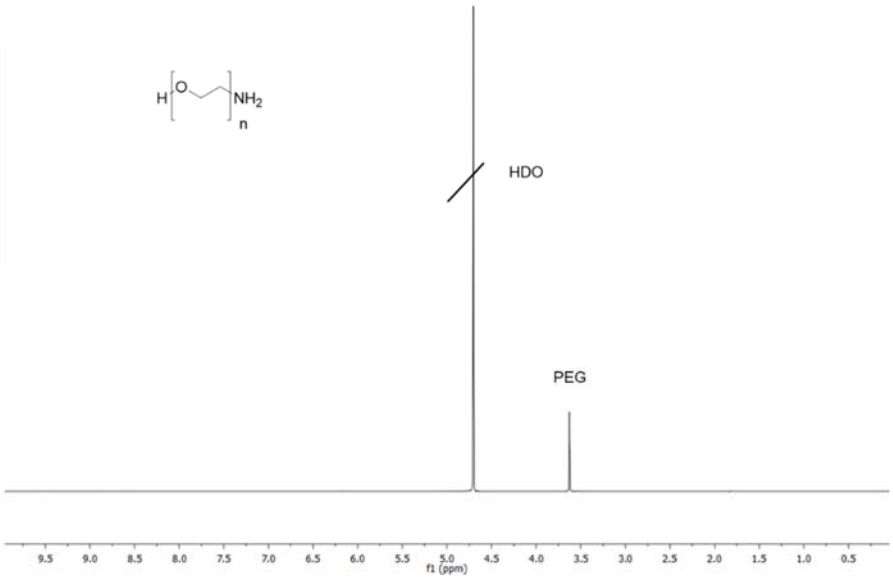


*Figure S 21: 600 MHz ^{1}H NMR spectrum of **PEG** in D_2O at 25°C.*

^{1}H-NMR (600 MHz, D_2O, 25°C): δ (ppm) = 3.64 (s, HPEG).

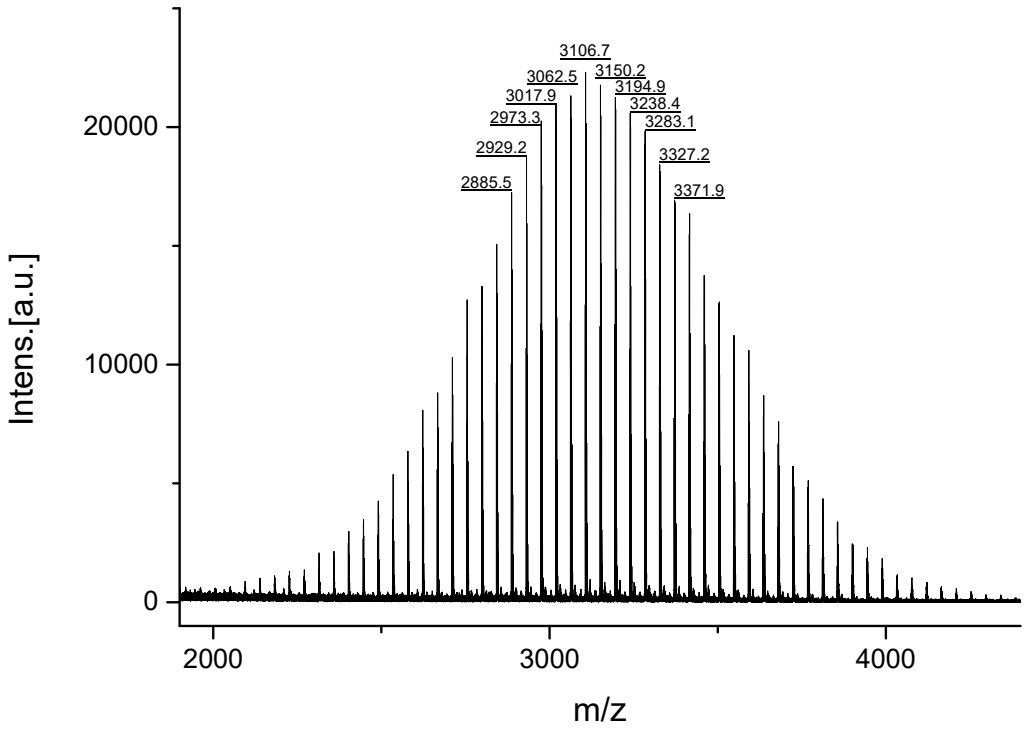


*Figure S 22: MALDI-TOF-MS of **PEG** in a m/z range using DHB as matrix in a compound to matrix ratio of 1:10. The value 3106.7 g/mol corresponds to 70 polyethylene glycol repeating units.*

Binding to Taspase1: SPR Assay

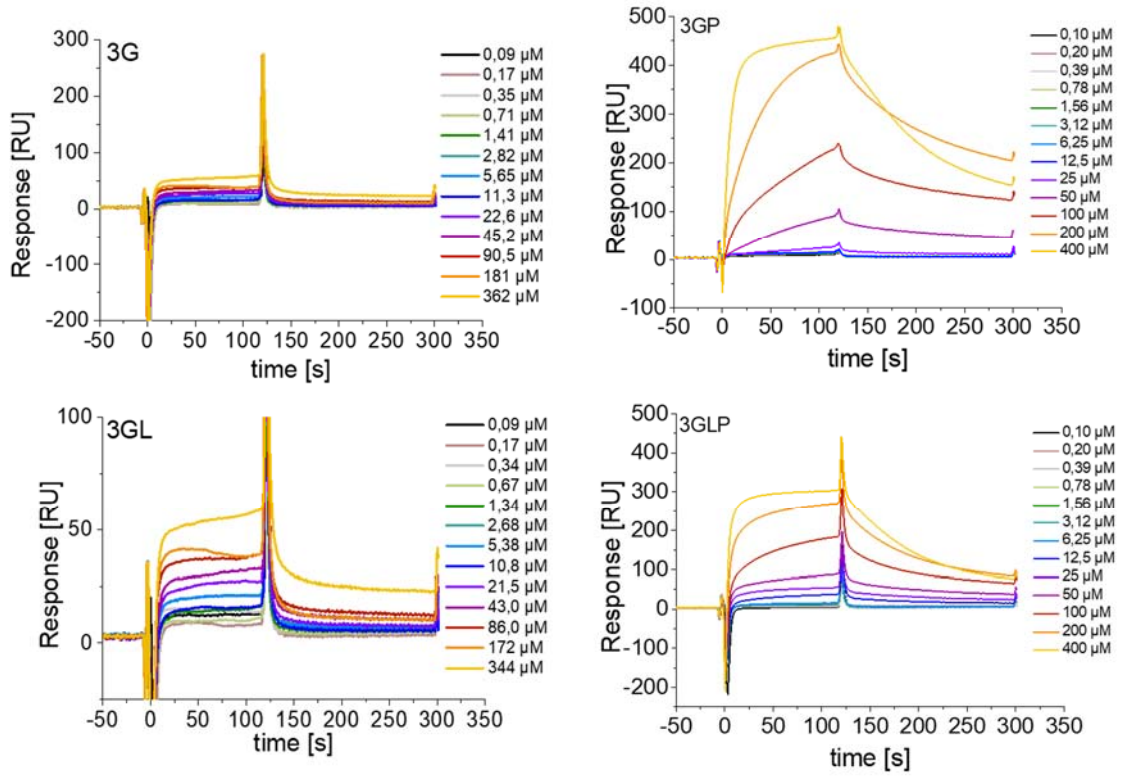


Figure S 23: SPR-Sensograms for PEGylated und non-PEGylated GCP macromolecules for the second C1 sensor chip. Each measurement was repeated two times.

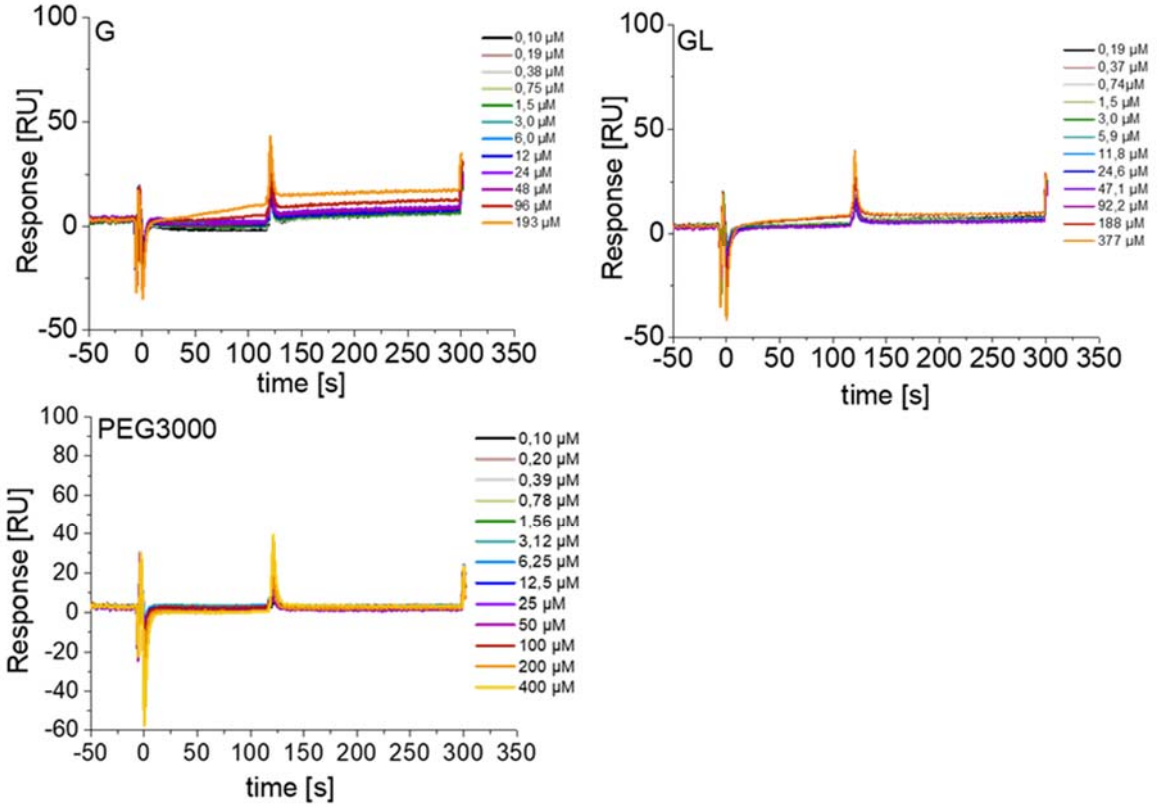


Figure S 24: SPR-Sensograms for monovalent non-PEGylated GCP macromolecules (top) and PEG3000 (bottom). The measurements showed no binding to Taspase1. Each measurement was repeated two times on the second C1 sensor chip.

Binding to Taspase1: Pull-down assay

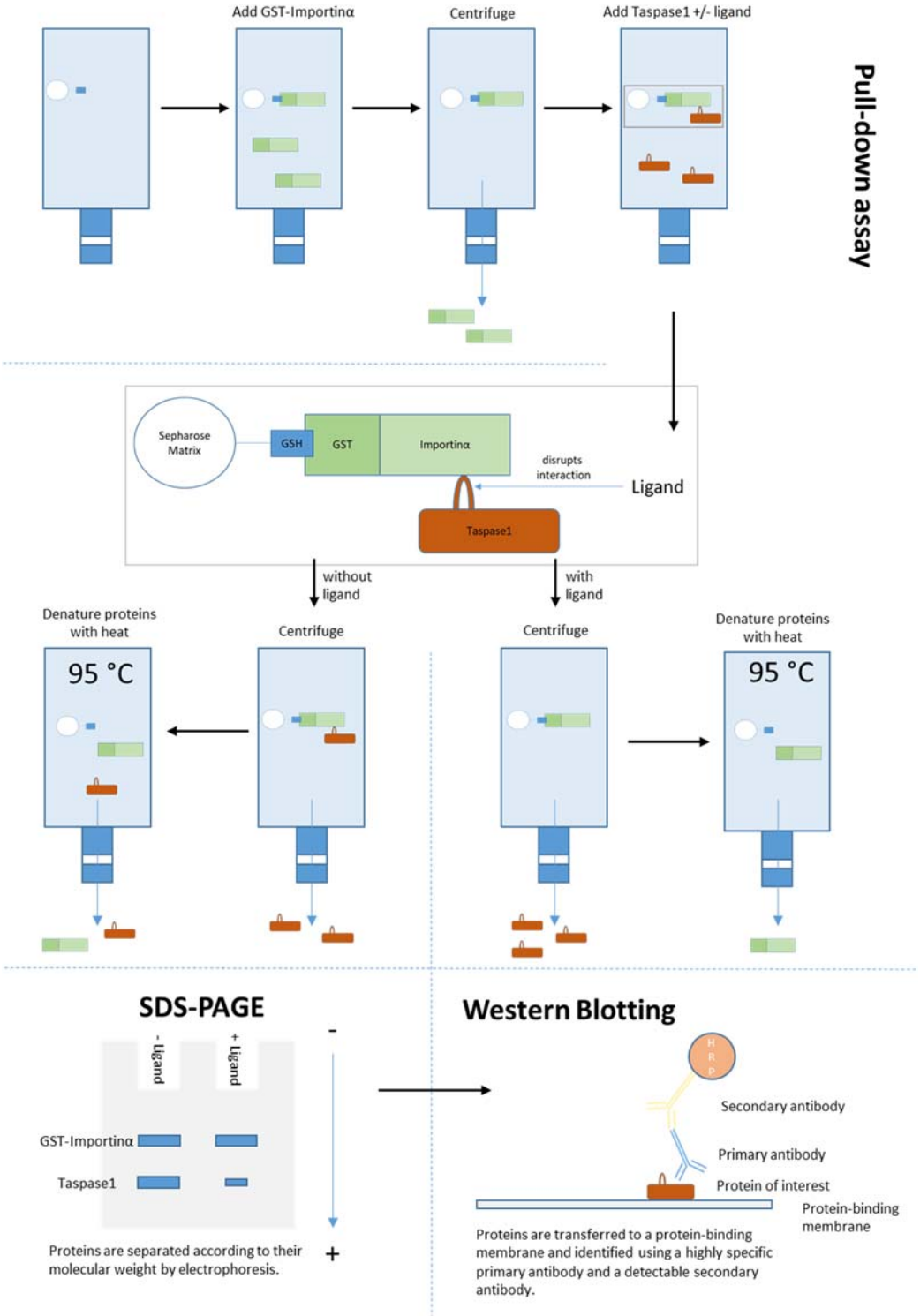


Figure S 25: Workflow of the modified pull-down assay. A spin column was used to fix GST-Importin α on a Sepharose matrix coated with glutathione. First, GST was allowed to bind to glutathione with high affinity, and unbound protein was removed by centrifugation. Then, Taspase1-His was pre-incubated with ligand or left untreated as indicated, subsequently added to the column, and unbound protein was again removed by centrifugation. Next, a buffer containing ionic detergents as well as reducing agents was applied to the column and heated to 95 °C to denature and thus dissociate all protein from the matrix. Finally, the proteins were separated according to their molecular weight by SDS page and analyzed by Western Blot analysis for quantification.

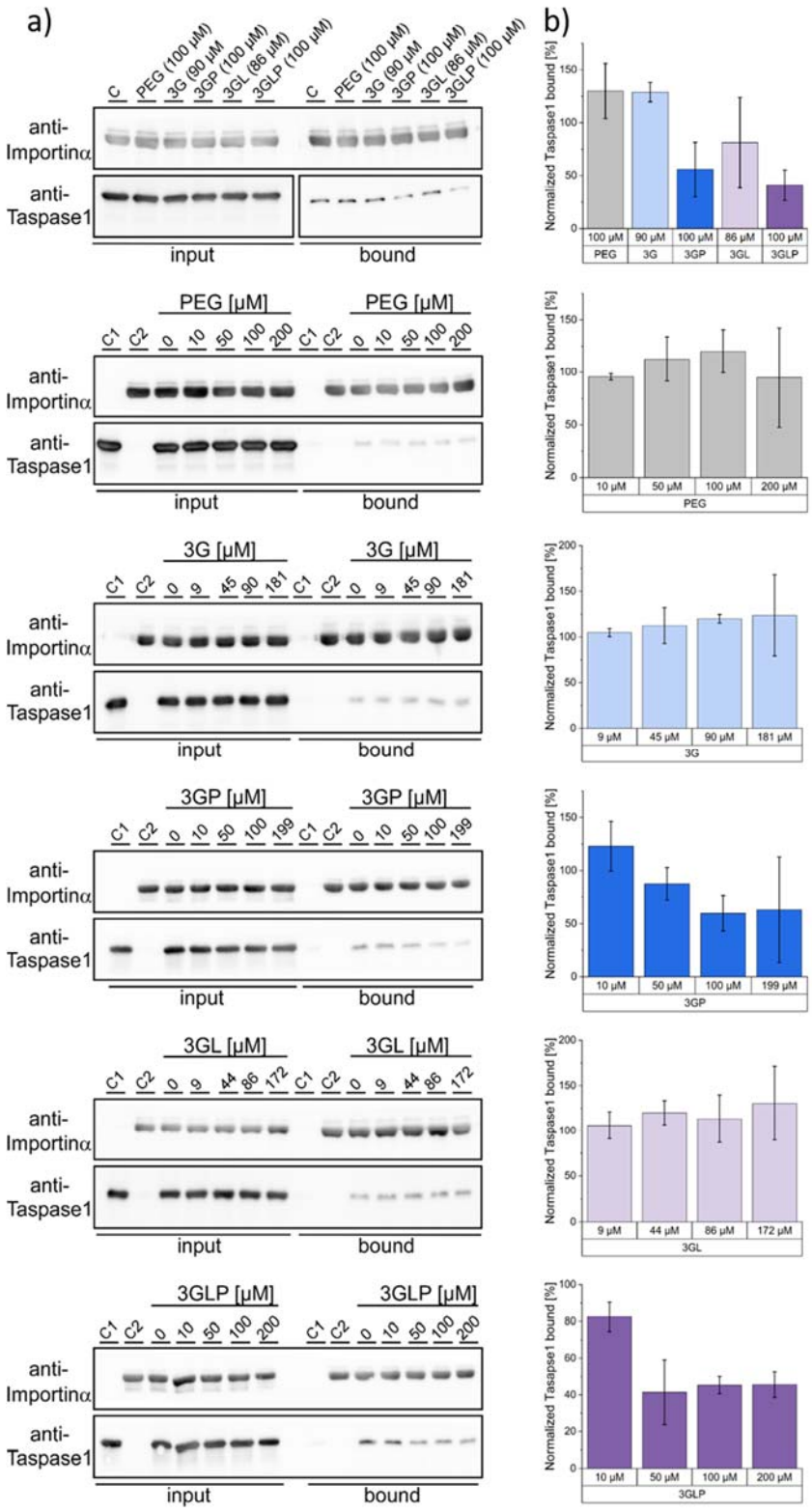


Figure S 26: a) Representative Western Blot analyses from the pull-down assays performed with the different compounds. The input fraction contains samples of the complete protein preparation added to the column, the bound fraction contains the respective portion bound to the column. The latter comprises GST-Importin α directly associated with the column and Taspase1 indirectly bound via its interaction with Importinα. C = Untreated control, C1 = Control with only Taspase1, C2 = Control with only Importin α. b) Densiometric quantification of the respective pull-down assays, comprising three replicates ± standard deviation. ***Please note:*** *originally concentrations were calculated not considering the counterions present in the structures. This was corrected leading to the here shown concentrations.*

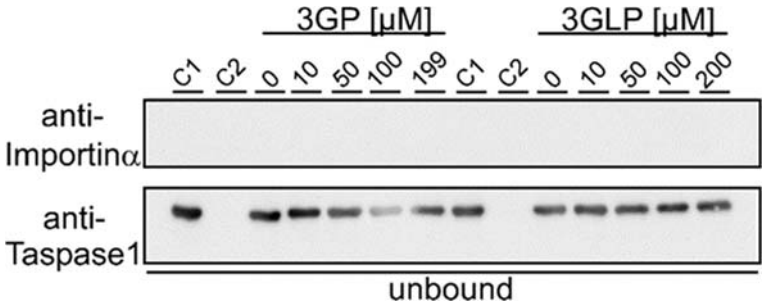


Figure S 27: Binding of Importin α to the column was not affected by the PEGylated ligands during the assay. Western Blot of the unbound fraction after incubation of the Importin α-loaded column with Taspase1 in the presence of the indicated ligands. C1 = Control with only Taspase1, C2 = Control with only Importin α.

Binding to Taspase1: ITC measurements

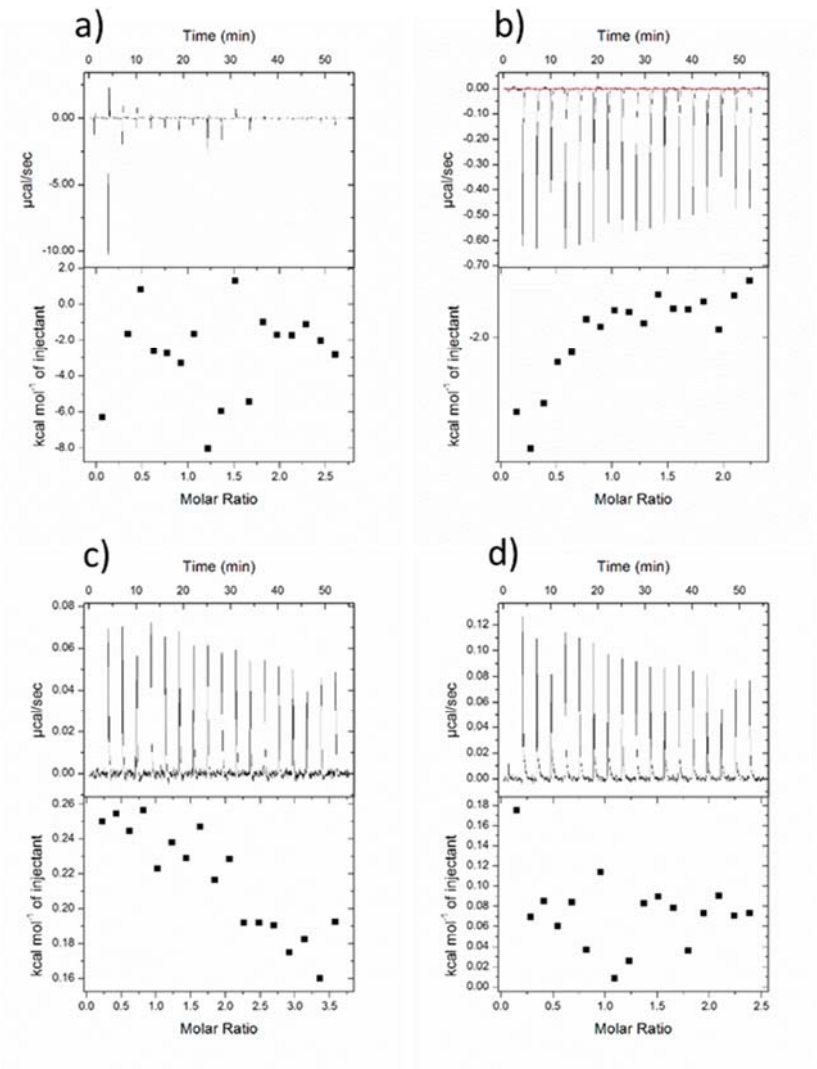


Figure S 28: Test of different ITC conditions with inactive Taspase1 and the respective ligands. a) 1 mM GLP to 70 µM inactive Taspase1, buffer: 150 mM NaCl, 50 mM NaH_2PO_4, pH 7,4 . b) 1 mM GLP to 80 µM inactive Taspase1, buffer: 10 % (w/v) Sucrose, 50 mM NaH_2PO_4, pH 7,4. c) 1,5 mM GLP to 75 µM inactive Taspase1, buffer: 10 % (w/v) Sucrose, 50 mM NaH_2PO_4, pH 7,4. d) 1 mM GP to 75 µM inactive Taspase1, buffer: 10 % (w/v) Sucrose, 50 mM NaH_2PO_4, pH 7,4. e) 1 mM GL to 75 µM inactive Taspase1, buffer: 10 % (w/v) Sucrose, 50 mM NaH_2PO_4, pH 7,4.

Toxicity study

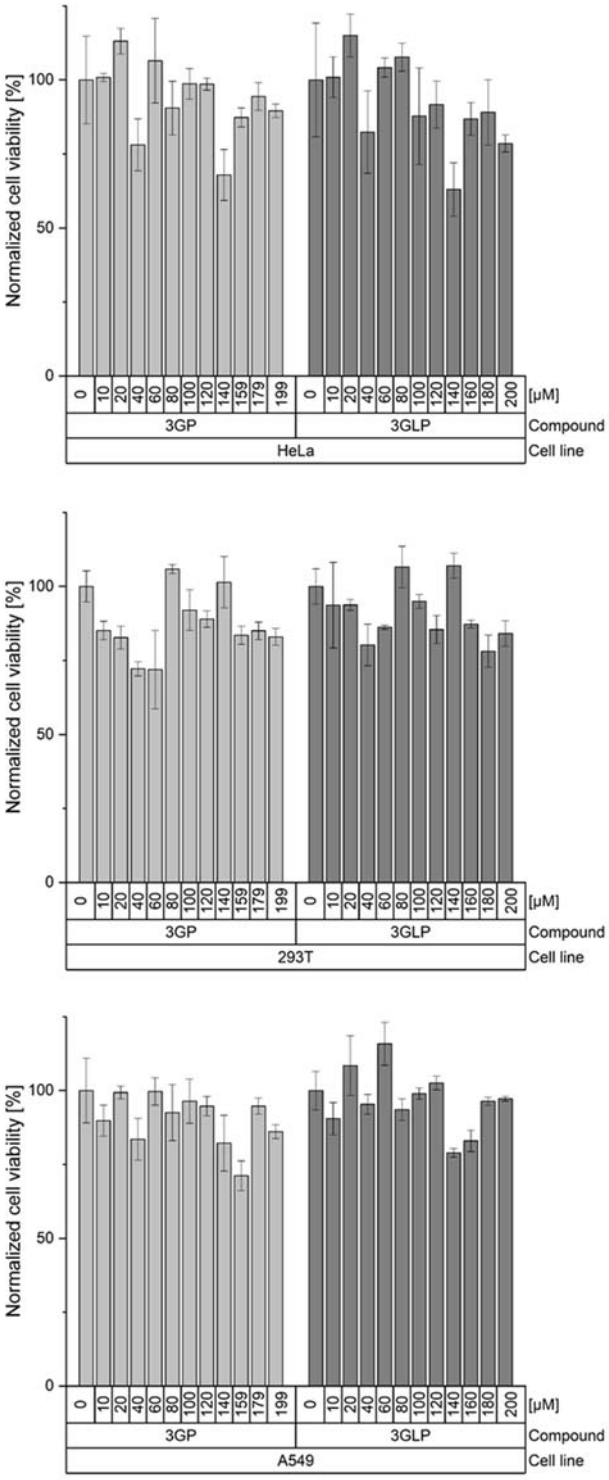


Figure S 29: The compounds do not affect the cell viability of various tumor cell lines. 293T (a), A549 (b) and HeLa (c) cells were cultivated in cell culture medium supplied with the respective concentrations of compound for 24 h. After that, we performed an MTS assay to determine the cell viability. The data points are the mean of triplicates ± standard deviation. ***Please note:*** *originally concentrations were calculated not considering the counterions present in the structures. This was corrected leading to the here shown concentrations.*

References:

[1] J. van den Boom, F. Trusch, L. Hoppstock, C. Beuck, P. Bayer, *PloS one* **2016**, *11* (3), 1-13.

[2] J. A. Khan, B. M. Dunn, L. Tong. *Structure*. **2005**, 13(10), 1443-1452.

[3] J. van den Boom, A. Hensel, F. Trusch, A. Matena, S. Siemer, D. Guel, D. Docter, A. Höing, P. Bayer, R. H. Stauber, S. K. Knauer, *Nanoscale* **2020**, *12*, 19093-19103.

[4] S. Ramakrishnan, K. N. Sulochana, R. Punitham, K. Arunagiri, *Glycoconj. J*. **1996**, *13*, 519-523.

[5] H. Towbin, T. Staehelin, J. Gordon, *PNAS* **1979,** *76* (9), 4350-4354.

[6] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, A. Cardona, *Nat. Methods* **2012**, *9* (7), 676-682.

[7] M. F. Ebbesen, C. Gerke, P. Hartwig, L. Hartmann, *Polym. Chem.* **2016***, 7* (46)*,* 7086-7093.

[8] C. Schmuck, V. Bickert, M. Merschky, L.Geiger, D. Rupprecht, J. Dudaczek, P. Wich, T. Rehm, U. Machon, *Eur. J. Org. Chem*. **2008**, *2*, 324-329.

6.2 Publication II: Supplementary information

ChemMedChem

Supporting Information

A Bivalent Supramolecular GCP Ligand Enables Blocking of the Taspase1/Importin α Interaction

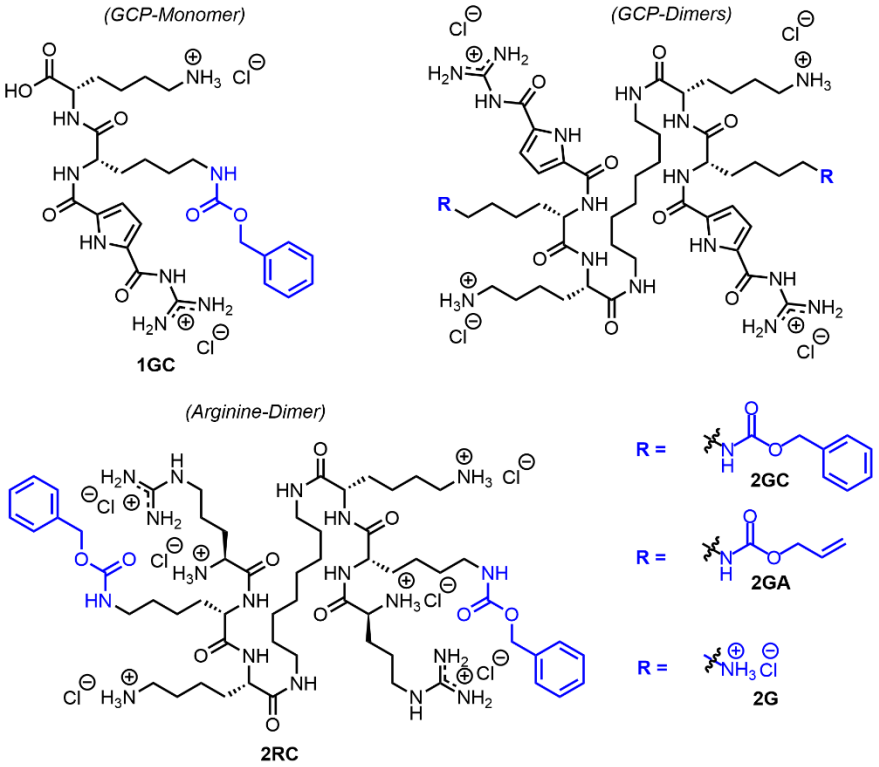
Alexander Höing[+], Alexander Zimmermann[+], Lisa Moews, Matthias Killa, Marius Heimann, Astrid Hensel, Jens Voskuhl,* and Shirley K. Knauer*

Table of Contents **Page**

General information	*2*
Synthesis	*3*
Mass spectra of the final compounds	*15*
NMR spectra	*17*
HPLC chromatograms	*22*
Molecular docking	*24*
Biological assays	*27*
References	*35*

General information

All chemicals were purchased either from Fluorochem, Sigma Aldrich or ABCR and used without further purification. The reactions were carried out using dried solvents and under inert gas atmosphere (argon). Reactions were monitored by thin-layer chromatography (TLC), which was performed on 0.2 mm Macherey-Nagel ALUGRAM precoated silica gel aluminum sheets. Spots were visualized using basic $KMnO_4$ solution or by an UV-hand-lamp (254 and 365 nm). Column chromatography was carried out on silica gel 60 (0.063 - 0.2 mm, Merck). Analytical HPLC chromatograms were performed with a system from Dionex and reversed phase column: YMC-ODS-AQ (150 mm, Ø = 3 mm, particle size = 5 µm, pore size = 12 nm). Preparative reversed phase chromatography was performed with the MPLC system "SPOT Liquid Chromatography FLASH" of Armen Instruments. The column contained YMC*Gel ODS A RP18 material (50 g, Ø = 25 mm particle size = 50 µm, pore size = 12 nm). Preparative column chromatography was performed on Silica 60 M (0.04 - 0.063 mm). The NMR spectra were recorded on a Bruker Avance HD 600 [^{1}H: 600.13 MHz, ^{13}C: 150.90 MHz] or Bruker Avance Neo 400 [^{1}H: 400.13 MHz, ^{13}C: 100.61 MHz] spectrometer. All measurements were performed at room temperature, using d_1-chloroform, d_6-DMSO or d_4-MeOD as solvents. The chemical shifts are referenced relative to the residual proton signals of the solvents in the ^{1}H-NMR spectrum ($CDCl_3$: δ = 7.24 ppm, d_6-DMSO: δ = 2.50 ppm, d_4-MeOD: δ = 3.31 ppm) or relative to the solvent signal in the ^{13}C-NMR spectrum ($CDCl_3$: δ = 77.16 ppm, d_6-DMSO: δ = 39.51 ppm, d_4-MeOD: δ = 49.0 ppm). Coupling constants (*J*) are reported in Hertz (Hz). In the NMR reports: "-" indicates a single bond, "=" a double bond and ">" two single bonds. High resolution mass spectra were measured on a Bruker maXis 4G UHR-TOF.



Scheme 1: Compounds synthesized and investigated in this study.

Synthesis

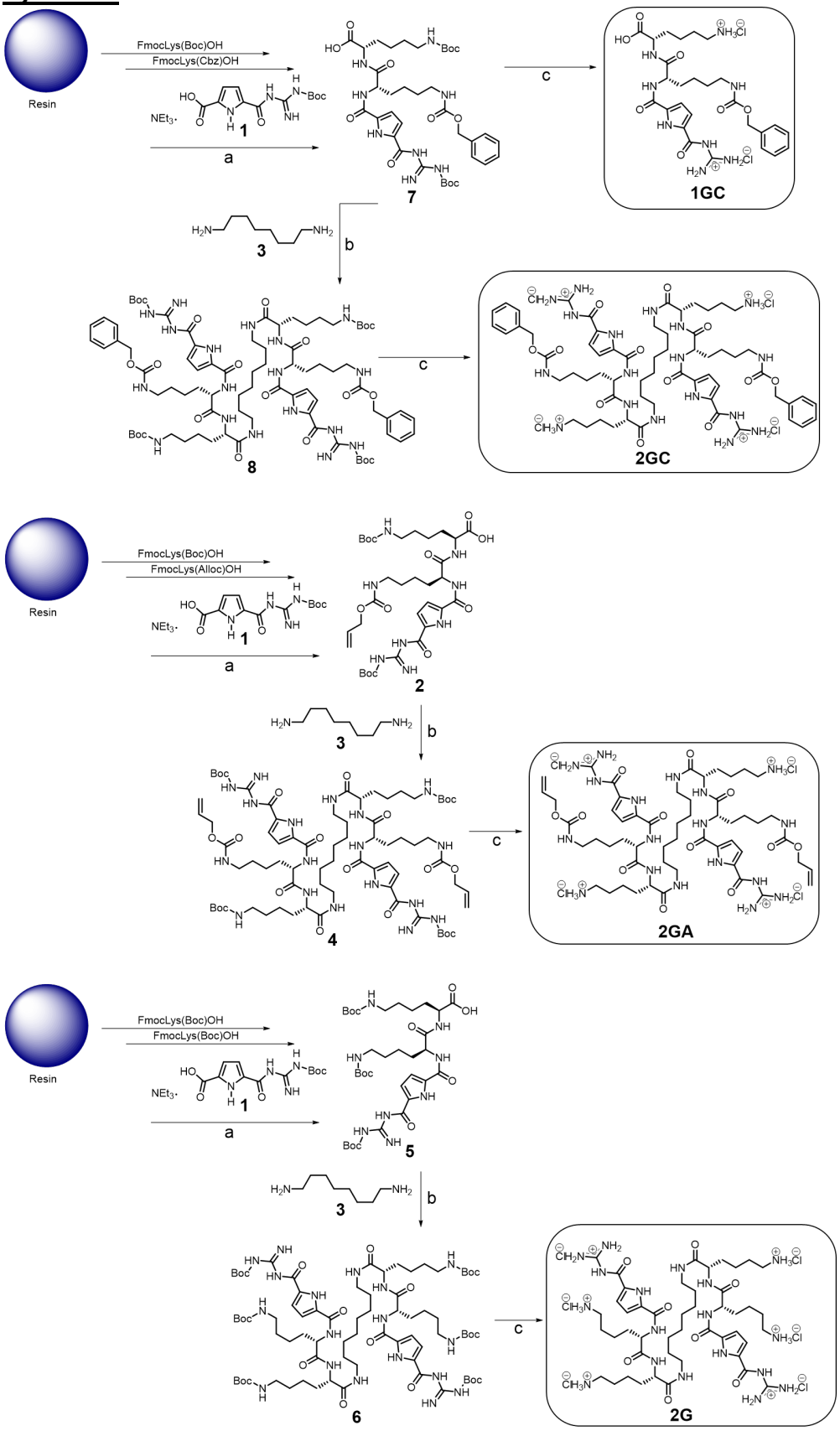


Figure S1: Synthesis route of **1GC**, **2GA**, **2G** and **2GC**. Reagents and conditions: **a)** 1.) Resin, DMF, RT, 2.) Fmoc(Boc)LysOH, DIPEA, PyBOP, DMF, RT, 3 h, 3.) 1:4 piperidine:DMF, RT, 0.5 h, 4.) Fmoc(Boc/Cbz/Alloc)Lys-OH, DIPEA, PyBOP, DMF, RT, o. n.; 5.) 1:4 piperidine:DMF, RT, 0.5 h, 6.) **1**, DIPEA, PyBOP, DMF, RT, 3 h, 7) DCM:TFA:TIS:H_2O 94:1:2.5:2.5, RT, 20 min.; **b) 3**, PyBOP, DIPEA, DCM, RT 16 h; **c)** 1.) TFA:DCM 1:1 RT, 1 h, 2.) HCl.

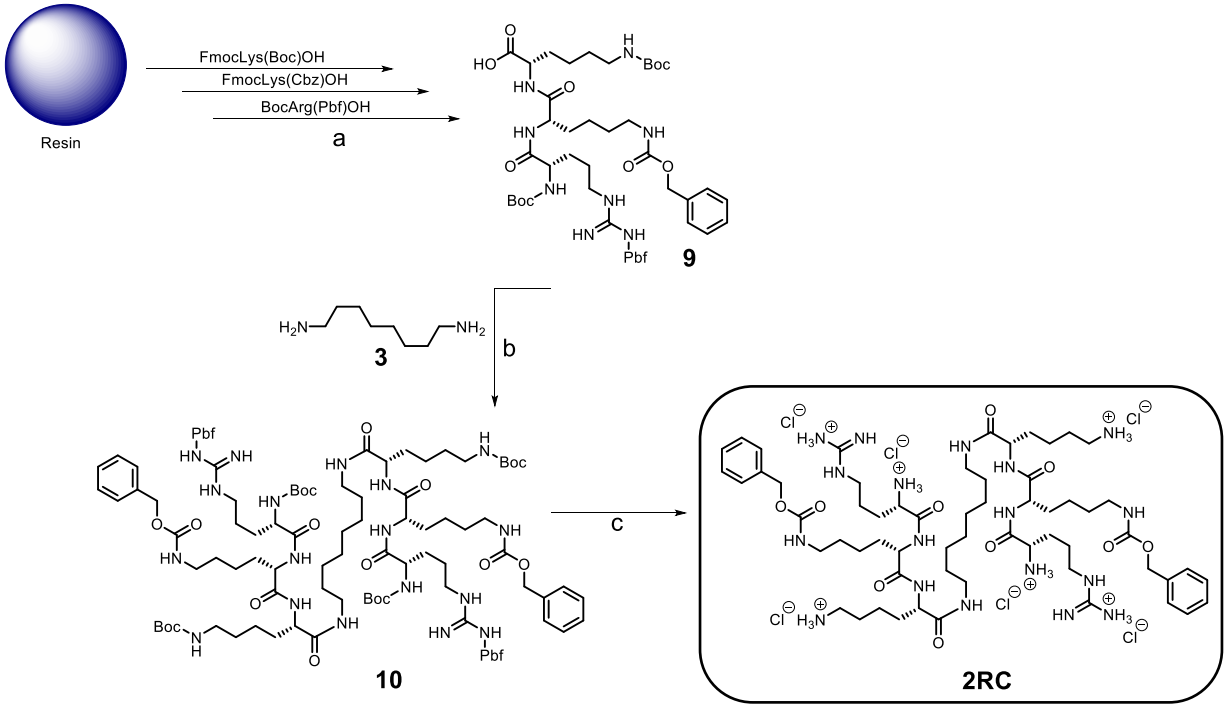


Figure S2: Synthesis route of **2RC**. Reagents and conditions: **a)** 1.) Resin, DMF, RT, 2.) Fmoc(Boc)LysOH, Et_3N, PyBOP, DMF, RT, 4 h, 3.) 1:4 piperidine:DMF, RT, 2 x 20 min, 4.) Fmoc(Cbz)Lys-OH, Et_3N, PyBOP, DMF, RT, o. n.; 5.) 1:4 piperidine:DMF, RT, 2 x 20 min, 6.) BocArg(Pbf)-OH, Et_3N, HATU, DMF, RT, 4 h, 7) DCM:TFA:TIS:H_2O 94:1:2.5:2.5, RT, 2 x 20 min.; **b) 3**, PyBOP, Et_3N, DMF, RT 19 h; **c)** 1.) TFA:DCM 1:1 RT, 1 h, 2.) 5 x 0,05 M HCl lyophilisation.

General Procedure A (GP-A)

Table S1: General Procedure of Solid Phase Peptide Synthesis (SPPS).

1. Swelling	Resin	20 ml DMF	2 h
2. Coupling	1. amino acid, DIPEA Washing Washing	In swelling solution 20 ml MeOH 20 ml DMF	overnight 10 min **3 x** 5 min
3. Fmoc deprotection	Piperidine in DMF 20% Piperidine in DMF 20% Washing	20 ml 20 ml 20 ml DMF	20 min 20 min **5 x** 5 min
4. Coupling	2. amino acid, PyBOP, DIPEA Washing	20 ml DMF 20 ml DMF	3 h **3 x** 5 min
5. Fmoc deprotection	Piperidine in DMF 20% Piperidine in DMF 20% Washing	20 ml 20 ml 20 ml DMF	20 min 20 min **5 x** 5 min
6. Coupling	Boc-GCP-OH, PyBOP, DIPEA	20 ml DMF	overnight
7. Washing	DMF DCM	20 ml 20 ml	**3 x** 5 min **3 x** 5 min
8. Cleavage	DCM/TFA/TIS/H_2O (94:1:2.5:2.5) Washing	10 ml 20 ml DCM	20 min **3 x** fast
9. Work up	Evaporation Disperse Precipitate	 20 ml DCM 20 ml Et_2O	Fast **2 x** **1x**

In a 100 ml SPPS vessel, 1.00 eq. 2-chlorotrityl resin (loading 1.55 mmol/g) was suspended in 20 ml DMF and swollen at room RT (room temperature) for 2 h. Then 2.00 - 3.00 eq. of the first amino acid and 6.00 eq. DIPEA were added and at RT shaken overnight. The solvent was suctioned off. The resin was washed with 20 ml MeOH for 10 min and then three times with 20 ml DMF for 5 min. Fmoc deprotection took place with 20 ml of a 1:4 piperidine:DMF mixture. The resin was treated with this solution twice, for 20 minutes. Subsequently the resin was washed five times for 5 min. with 20 ml DMF. 2.00-3.00 eq. of the 2nd amino acid, 2.10 - 3.00 eq. PyBOP and 6.00 eq. DIPEA were added to the resin and dissolved with 20 ml DMF. After the solution was shaken for 3 h, the solution was suctioned off. The Fmoc protective group was then cleaved analogously to the first amino acid and washed five times, for 5 min. with 20 ml of DMF. Then 2.00 - 3.00 eq. Boc-GCP-OH together with 2.10 - 3.00 eq. PyBOP and 6.00 eq. DIPEA were added to the resin and dissolved in 20 ml DMF. The reaction mixture was shaken overnight and then suctioned off. The resin was washed three times, 5 minutes each, with 20 ml of DMF and 20 ml of DCM and then dried. 10 ml of a 94:1:2.5:2.5 DCM:TFA:TIS:H_2O mixture was added to the resin and shaken for 20 minutes. The resin was then washed three times with 20 ml of DCM. The filtrate was freed from the solvent under reduced pressure and dispersed twice with 20 ml of DCM. The product was then suspended in 20 ml Et_2O and filtered. No further purification of the product was conducted.

General Procedure B (GP-B)

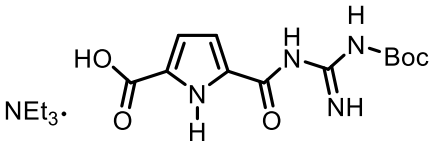
1.00 eq. of a primary amine, 2.20 eq. of a carboxylic acid, 2.50 eq. PyBOP and 5.00 eq. DIPEA were solved in DCM. The reaction mixture was stirred at RT for 16 h. The solution became cloudy during the reaction time. The precipitate was filtered and washed with DCM and dried on a rotary evaporator.

General Procedure C (GP-C)

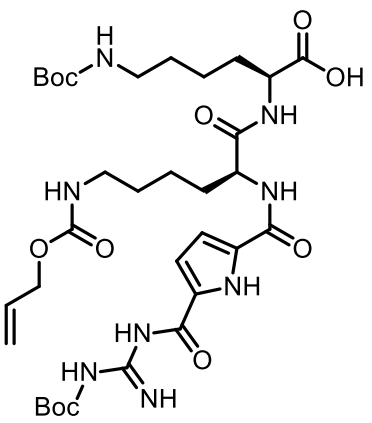
10 ml of a 1:1 TFA:DCM mixture was added to the Boc-protected compound and stirred at RT for 1 h. Afterwards the crude product was precipitated with Et_2O and filtered. The crude product was dried on a rotary evaporator and purified *via* MPLC. After purification, 150 ml of a 3 M hydrochloric acid solution was added to the product. The solvent was distilled off to dryness on a rotary evaporator. This procedure was repeated once. The product was taken up in water and lyophilized.

Compound 1

The synthesis of **1** was carried out according to a known literature procedure. All spectra obtained were in accordance with those reported beforehand.[1]

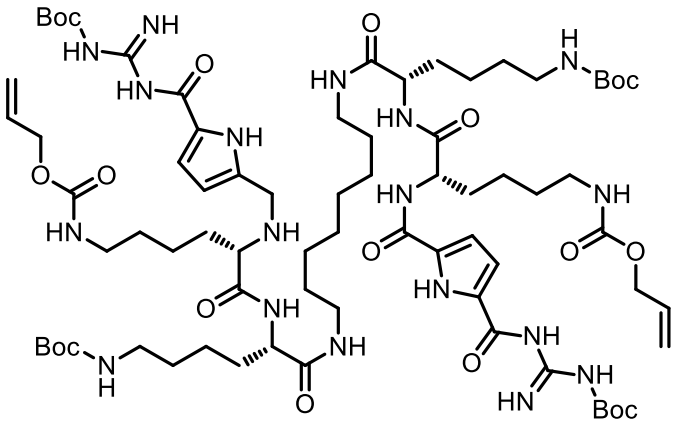


Compound 2



Compound **2** was synthesized according to **GP-A**. 1.00 g (1.55 mmol/g, 1,00 eq.) 2-chlorotrityl resin, 2.18 g (4.65 mmol, 3.00 eq.) Fmoc-Lys(Boc)-OH, 2.10 g (4.64 mmol, 2.99 eq.) Fmoc-Lys(Alloc)-OH, 1.85 g (4.64 mol, 2.99 eq.) Boc-GCP-OH, 2x 2.42 g (4.65 mmol, 3.00 eq.) PyBOP and 3x 1.60 ml (9.19 mmol, 5.93 eq.) DIPEA were used during the SPPS. 1.05 g (1.43 mmol, 92% based on the resin loading) of **2** as white solid were isolated. **RF** = 0.68 (MeOH:DCM 1:9). **mp.:** 127°C (decomposition).
^{1}H-NMR (400 MHz, DMSO-d_6) δ [ppm] = 11.65 (s, 2H, -N**H**-), 9.32 (s, 1H, -N**H**-), 8.57 (s, 1H, -N**H**-), 8.40 (d, *J* = 8.0 Hz, 1H, -N**H**-), 8.19 (d, *J* = 7.6 Hz, 1H, -N**H**-), 7.17 (t, *J* = 5.6 Hz, 1H, -N**H**-), 6.83 (s, 2H, =C**H**-), 6.76 (t, *J* = 5.5 Hz, 1H, -N**H**-), 5.99 – 5.75 (m, 1H, =C**H**-), 5.29 – 5.19 (m, 1H, C=C**H$_2$**), 5.14 (dd, *J* = 10.5, 1.6 Hz, 1H, C=C**H$_2$**), 4.45 (dd, *J* = 14.5, 7.0 Hz, 3H, O-C**H$_2$**- + >C**H**-), 4.12 (dd, *J* = 13.0, 8.2 Hz, 1H, >C**H**-), 3.04 – 2.91 (m, 2H, -C**H$_2$**-), 2.90 – 2.81 (m, 2H, -C**H$_2$**-), 1.80 – 1.65 (m, 2H, -C**H$_2$**-), 1.65 – 1.52 (m, 2H, -C**H$_2$**-), 1.46 (s, 9H, -C**H$_3$**), 1.43 – 1.26 (m, 17H, -C**H$_3$** + -C**H$_2$**-). **^{13}C-NMR** (101 MHz, DMSO-d_6) δ [ppm] = 173.6 (s, -**C**(O)OH), 171.9 (s, >**C**=O), 159.4 (s, >**C**=O), 158.5 (s, >**C**=O), 155.9 (s, >**C**=O), 155.5 (s, >**C**=O), 133.9 (s, =**C**H-), 116.8 (s, C=**C**H$_2$), 113.6 (s, =**C**H-), 112.9 (s, =**C**H-), 77.4 (s, -**C**(CH$_3$)$_3$), 64.1 (s, O-**C**H$_2$-), 52.5 (s, >**C**H-), 51.9 (s, >**C**H-), 31.7 (s, -**C**H$_2$-), 30.6 (s, -**C**H$_2$-), 29.2 (s, -**C**H$_2$-), 29.1 (s, -**C**H$_2$-), 28.3 (s, -**C**H$_3$), 27.8 (s, -**C**H$_3$), 22.9 (s, -**C**H$_2$-), 22.8 (s, -**C**H$_2$-). **IR (ATR):** [cm^{-1}] = 3315 (NH), 3192 (NH), 3091 (CH arom.), 2978 (CH aliph.), 2933 (CH aliph.), 2866 (CH aliph.), 1687 (C=O). **ESI-HRMS** m/z (%): 737 $[M+H]^+$ (100), 269 $[M-C_{10}H_{14}O_2]^{2+}$ (52), 759 $[M+Na]^+$ (9), 637 $[M-C_5H_7O_2]^+$ (4). Calculated for $C_{33}H_{53}N_8O_{11}$ $[M+H]^+$: 737.3828; found: 737.3836.

Compound 4



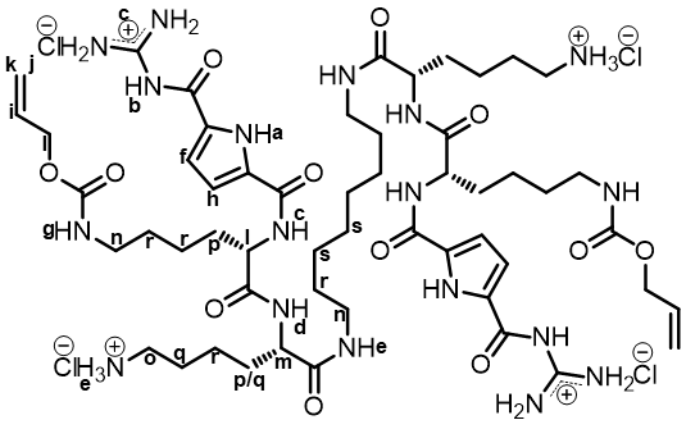
0.10 g (0.69 mmol, 1.00 eq.) **3**, 1.12 g (1.53 mmol, 2.20 eq.) **2**, 0.90 g (1.73 mmol, 2.49 eq.) PyBOP and 0.60 ml (3.44 mmol, 4.97 eq.) DIPEA in 35 ml DCM were implemented according to **GP-B**. The crude product was purified by column chromatography. Column: V(SiO_2) = 300 ml, Ø = 4 cm, eluent = 1:19 MeOH:DCM. 0.76 g (0.48 mmol, 69%) of slightly yellow solid were isolated. **RF** = 0.62 (MeOH/DCM 1:9). **mp.:** 125°C (decomposition).
^{1}H-NMR (400 MHz, DMSO-d_6) δ [ppm] = 11.64 (s, 2H, -N**H**-), 10.88 (s, 2H, -N**H**-), 9.33 (s, 2H, -N**H**-), 8.58 (s, 2H, -N**H**-), 8.41 (d, *J* = 7.2 Hz, 2H, -N**H**-), 7.96 (d, *J* = 7.8 Hz, 2H, -N**H**-), 7.76 (s, 2H, -N**H**-), 7.16 (t, *J* = 5.0 Hz, 2H, -N**H**-), 6.83 (s, 4H, =C**H**-), 6.72 (s, 2H, -N**H**-), 5.87 (ddt, *J* = 16.0, 10.5, 5.3 Hz, 2H, =C**H**-), 5.24 (d, *J* = 17.2 Hz, 2H, C=C**H$_2$**), 5.13 (d, *J* = 10.4 Hz, 2H, C=C**H$_2$**), 4. 54 – 4.29 (m, 6H, -C**H$_2$**- + >C**H**-), 4.25 – 4.06 (m, 2H, >C**H**-), 3.18 – 2.90 (m, 8H, -C**H$_2$**-), 2.90 – 2.76 (m, 4H, -C**H$_2$**-), 1.78 – 1.55 (m, 6H, -C**H$_2$**- + -C**H$_2$**-), 1.53 – 1.42 (m, 20H, -C**H$_3$** + -C**H$_2$**-), 1.44 – 1.27 (m, 34H, -C**H$_3$** + -C**H$_2$**-), 1.26 – 1.11 (m, 12H, -C**H$_2$**-). **^{13}C-NMR** (101 MHz, DMSO-d_6) δ [ppm] = 171.6 (s,

>**C**=O), 171.2 (s, >**C**=O), 159.6 (s, >**C**=O), 158.5 (s, >**C**=O), 155.9 (s, >**C**=O), 155.6 (s, >**C**=O), 133.9 (s, =**C**H-), 116.8 (s, C=**C**H_2), 113.7 (s, =**C**H-), 113.0 (s, =**C**H-), 77.3 (s, -**C**$(CH_3)_3$), 64.1 (s, O-**C**H_2-), 53.1 (s, >**C**H-), 52.6 (s, >**C**H-), 38.5 (s, -**C**H_2-), 31.8 (s, -**C**H_2-), 31.5 (s, -**C**H_2-), 29.3 (s, -**C**H_2-), 29.2 (s, -**C**H_2-), 29.0 (s, -**C**H_2-), 28.8 (s, -**C**H_2-), 28.3 (s, -**C**H_3), 27.8 (s, -**C**H_3), 26.4 (s, -**C**H_2-), 23.0 (s, -**C**H_2-), 22.8 (s, -**C**H_2-). **IR (ATR):** [cm^{-1}] = 3292 (NH), 3081 (CH arom.), 2978 (CH aliph.), 2931 (CH aliph.), 2860 (CH aliph.), 1689 (C=O). **ESI-HRMS** m/z (%): 791 $[M+2H]^{2+}$ (100), 741 $[M-C_5H_7O_2]^{2+}$ (18), 1582 $[M+H]^+$ (5). Calculated for $C_{74}H_{121}N_{18}O_{20}$ $[M+H]^+$: 1581.8999; found: 1581.8830.

Compound 2GA

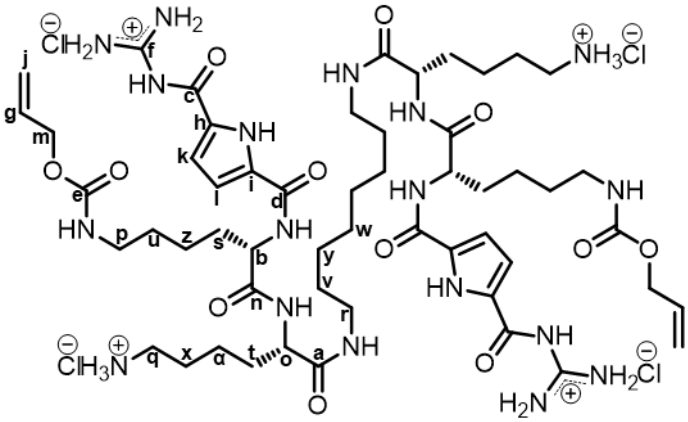
1.00 g (0.63 mmol, 1.00 eq.) **4** were implemented according to **GP-C**. The product was purified *via* MPLC (gradient: 40:60 → 70:30 MeOH:H_2O, 100 min.). 0.18 g (0.11 mmol, 17%) of a white solid were isolated. **R_t** HPLC RP18 (Gradient 10:90 → 100:0 MeOH:H_2O, 30 min.) = 19.4 min. **mp.:** 92 - 95°C.

^{1}H-NMR (400 MHz, DMSO-d_6) δ [ppm] = 12.54 (s, 2H) **a**-H, 12.15 (s, 2H) **b**-H, 8.98 – 8.28 (m, 10H) **c**-H, 8.14 (d, *J* = 8.1 Hz, 2H) **d**-H, 8.04 – 7.72 (m, 8H) **e**-H, 7.59 (s, 2H) **f**-H, 7.19 (t, *J* = 5.6 Hz, 2H) **g**-H, 6.90 (d, *J* = 2.7 Hz, 2H) **h**-H, 5.95 – 5.79 (m, 2H) **i**-H, 5.24 (dd, *J* = 17.2, 1.6 Hz, 2H) **j**-H, 5.14 (dd, *J* = 10.5, 1.4 Hz, 2H) **k**-H, 4.51 – 4.33 (m, 6H) **l**-H, 4.25 – 4.09 (m, 2H) **m**-H, 3.12 – 2.85 (m, 8H) **n**-H, 2.81 – 2.63 (m, 4H) **o**-H, 1.80 – 1.59 (m, 6H) **p**-H, 1.59 – 1.46 (m, 6H) **q**-H, 1.46 – 1.25 (m, 16H) **r**-H, 1.21 (s, 8H) **s**-H.



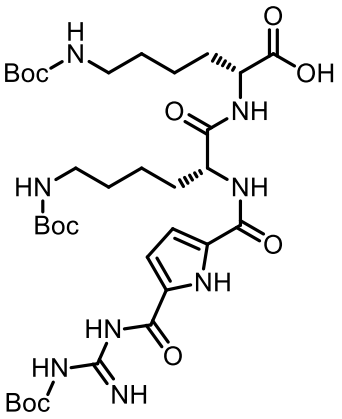
---

^{13}C-NMR (101 MHz, DMSO-d_6) δ [ppm] = ^{13}C-NMR (101 MHz, DMSO) δ [ppm] = 171.5 (s) **a**, 171.2 (s) **b**, 159.8 (s) **c**, 159.0 (s) **d**, 155.9 (s) **e**, 155.6 (s) **f**, 133.9 (s) **g**, 132.3 (s) **h**, 125.6 (s) **i**, 116.8 (s) **j**, 115.8 (s) **k**, 113.6 (s) **l**, 64.1 (s) **m**, 53.2 (s) **n**, 52.4 (s) **o**, 40.2 (s) **p**, 38.5 (s) **q**, 38.4 (s) **r**, 31.5 (s) **s**, 31.4 (s) **t**, 29.2 (s) **u**, 29.0 (s) **v**, 28.7 (s) **w**, 26.5 (s) **x**, 26.3 (s) **y**, 22.9 (s) **z**, 22.3 (s) **α**.



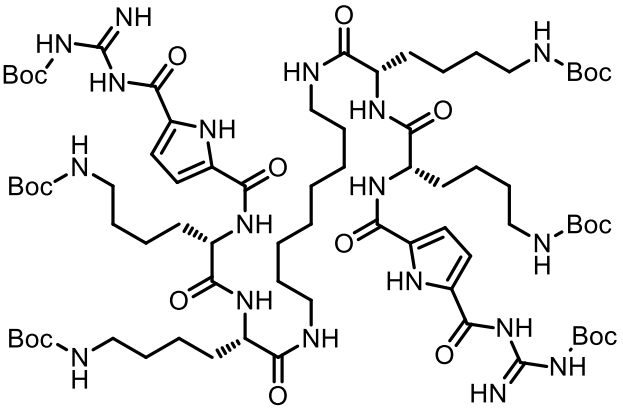
IR (ATR): [cm^{-1}] = 3280 (NH), 3099 (NH), 2929 (CH aliph.), 2860 (CH aliph.), 1693 (C=O). **ESI-HRMS** m/z (%): 395 $[M+3H]^{3+}$ (100), 296 $[M+4H]^{4+}$ (32), 591 $[M+2H]^{2+}$ (24), 1182 $[M+H]^+$ (1). Calculated for $C_{54}H_{90}N_{18}O_{12}$ $[M+2H]^{2+}$: 591.3487; found: 591.3496.

Compound 5



Compound **5** was synthesized according to **GP-A**. 1.00 g (1.55 mmol/g, 1.00 eq.) 2-chlorotrityl resin, 2x 1.45 g (3.09 mmol, 2.00 eq.) Fmoc-Lys(Boc)-OH, 1.24 g (3.11 mol, 2.01 eq.) Boc-GCP-OH, 2x 1.69 g (3.25 mmol, 2.10 eq.) PyBOP and 3x 1.60 ml (9.19 mmol, 5,93 eq.) DIPEA were used during the SPPS. 1.12 g (1.49 mmol, 96% based on the resin loading) of **5** as white solid were isolated. **RF** = 0.59 (MeOH/DCM 1:9). **mp.:** 187 - 190°C
^{1}H-NMR (400 MHz, DMSO-d_6) δ [ppm] = 12.32 – 11.17 (m, 2H, -N**H**-), 9.37 (s, 1H, -N**H**-), 8.56 (s, 1H, -N**H**-), 8.41 (d, *J* = 8.0 Hz, 1H, -N**H**-), 8.20 (d, *J* = 7.6 Hz, 1H, -N**H**-), 6.88 (s, 1H, =C**H**-), 6.86 – 6.81 (m, 1H, =C**H**-), 6.80 – 6.70 (m, 2H, -N**H**-), 4.53 – 4.38 (m, 1H, >C**H**-), 4.20 – 4.06 (m, 1H, >C**H**-), 2.97 – 2.78 (m, 4H, -C**H$_2$**-), 1.77 – 1.65 (m, 2H, -C**H$_2$**-), 1.64 – 1.52 (m, 2H, -C**H$_2$**-),1.46 (s, 9H, -C**H$_3$**), 1.42 – 1.27 (m, 26H, -C**H$_3$** + -C**H$_2$**-). **^{13}C NMR** (101 MHz, DMSO-d_6) δ [ppm] = 173.5 (s, -**C**(O)OH), 172.0 (s, >**C**=O), 159.3 (s, >**C**=O), 155.6 (s, >**C**=O), 113.9 (s, =**C**H-), 113.0 (s, =**C**H-), 77.3 (s, -**C**$(CH_3)_3$), 52.5 (s, >**C**H-), 51.9 (s, >**C**H-), 31.7 (s, -C**H$_2$**-), 30.6 (s, -C**H$_2$**-), 29.3 (s, -C**H$_2$**-), 29.1 (s, -C**H$_2$**-), 28.3 (s, -C**H$_3$**), 27.8 (s, -C**H$_3$**), 22.9 (s, -C**H$_2$**-), 22.8 (s, -C**H$_2$**-).
IR (ATR): [cm^{-1}] = 3313 (NH), 3195 (NH), 3086 (CH arom.), 2976 (CH aliph.), 2931 (CH aliph.), 2866 (CH aliph.), 1685 (C=O). **ESI-HRMS** m/z (%): 753 $[M+H]^+$ (100), 227 $[M\text{-}C_{15}H_{22}O_6]^{2+}$ (26), 653 $[M\text{-}C_5H_7O_2]^+$ (6), 453 $[M\text{-}C_{15}H_{23}O_6]^+$ (3), 553 $[M\text{-}C_{10}H_{15}O_4]^+$ (1). Calculated for $C_{35}H_{57}N_8O_{11}$ $[M+H]^+$: 753.4141; found: 753.4160.

Compound 6

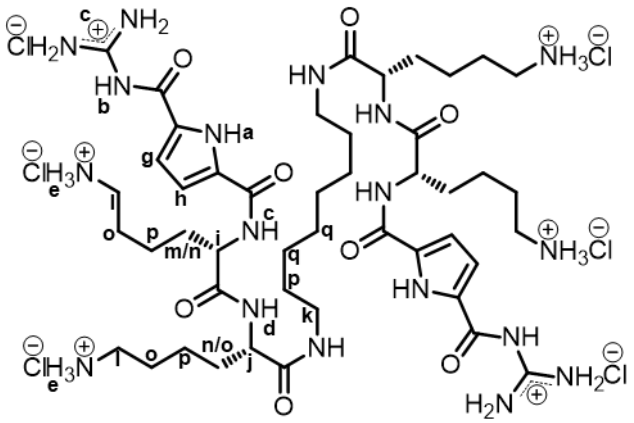


0.14 g (1.94 mmol, 0.97 eq.) **3**, 1.54 g (2.05 mmol, 2.10 eq.) **5**, 1.27 g (2.44 mmol, 2.50 eq.) PyBOP and 0.85 ml (4.88 mmol, 5.01 eq.) DIPEA in 60 ml DCM were implemented according to **GP-B**. No further purification was performed. 0.84 g (0.52 mmol, 54%) of a white solid was isolated. **RF** = 0.49 (MeOH/DCM 1:9). **mp.:** 140°C (decomposition).
^{1}H-NMR (400 MHz, DMSO-d_6) δ [ppm] = 11.08 (s, 4H, -N**H**-), 9.33 (s, 2H, -N**H**-), 8.77 – 8.34 (m, 4H, -N**H**-), 8.01 (d, *J* = 7.9 Hz, 2H, -N**H**-), 7.83 – 7.74 (m, 2H, -N**H**-), 6.84 – 6.78 (m, 4H, =C**H**-), 6.78 – 6.66 (m, 4H, -N**H**-), 4.40 – 4.29 (m, 2H, >C**H**-), 4.22 – 4.06 (m, 2H, >C**H**-), 3.09 – 2.93 (m, 4H, -C**H$_2$**-), 2.93 – 2.75 (m, 8H, -C**H$_2$**-), 1.75 – 1.55 (m, 6H, -C**H$_2$**- + -C**H$_2$**-), 1.53 – 1.42 (m, 20H, -C**H$_2$**- + -C**H$_3$**), 1.41 – 1.09 (m, 64H, -C**H$_3$** + -C**H$_2$**-).
^{13}C NMR (101 MHz, DMSO-d_6) δ [ppm] = 171.7 (s, >**C**=O), 171.2 (s, >**C**=O), 165.8 (s, >**C**=O), 158.5 (s, >**C**=O), 155.5 (s, >**C**=O), 113.6 (s, =**C**H-), 113.0 (s, =**C**H-), 77.3 (s, -**C**$(CH_3)_3$), 53.2 (s, >**C**H-), 52.6 (s, >**C**H-), 38.5 (s, -**C**H_2-), 31.8 (s, -**C**H_2-), 31.5 (s, -**C**H_2-), 29.3 (s, -**C**H_2-), 29.2 (s, -**C**H_2-), 29.0 (s, -**C**H_2-), 28.7 (s, -**C**H_2-), 28.3 (s, -**C**H_3), 27.8 (s, -**C**H_3), 26.3 (s, -**C**H_2-), 23.0 (s, -**C**H_2-), 22.7 (s, -**C**H_2-). **IR (ATR):** [cm^{-1}] = 3386 (NH), 3284 (NH), 3082 (CH arom.), 2978 (CH aliph.), 2931 (CH aliph.), 2860 (CH aliph.), 1687 (C=O). **ESI-HRMS** m/z (%): 807 $[M+H]^{2+}$ (100), 372 $[M\text{-}C_{25}H_{39}O_{10}]^{3+}$ (13), 757 $[M\text{-}C_5H_7O_2]^{2+}$ (12), 707 $[M\text{-}C_{10}H_{15}O_2]^{2+}$ (1). Calculated for $C_{76}H_{130}N_{18}O_{20}$ $[M+2H]^{2+}$: 807.4849; found: 807.4875.

Compound 2G

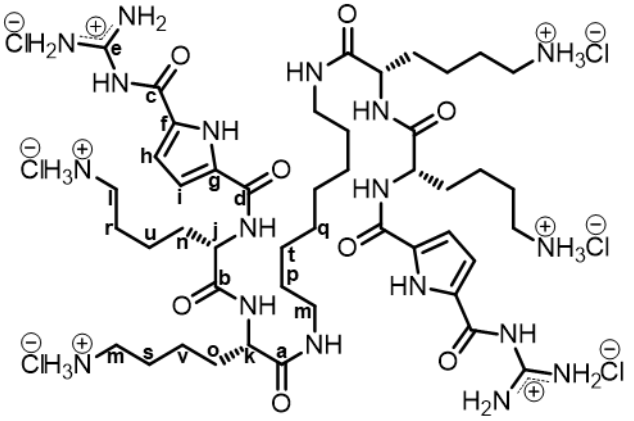
0.12 g (0.07 mmol, 1.00 eq.) **6** were implemented according to **GP-C**. The product was purified *via* MPLC (gradient: 20:80 → 50:50 MeOH:H_2O, 100 min.). 0.03 g (0.02 mmol, 24%) of a white solid were isolated. **R_t** HPLC RP18 (Gradient 10:90 → 100:0 MeOH:H_2O, 30 min.) = 12.5 min. **mp.:** 92 - 95°C.

^{1}H-NMR (600 MHz, DMSO-d_6) δ [ppm] = 12.55 (s, 2H) **a**-H, 12.20 (s, 2H) **b**-H, 8.81 – 8.49 (m, 10H) **c**-H, 8.19 (d, *J* = 8.1 Hz, 2H) **d**-H, 7.98 (s, 12H) **e**-H, 7.93 (t, *J* = 5.4 Hz, 2H) **f**-H, 7.65 – 7.59 (m, 2H) **g**-H, 6.93 (dd, *J* = 3.8, 2.2 Hz, 2H) **h**-H, 4.47 – 4.39 (m, 2H) **i**-H, 4.21 – 4.16 (m, 2H) **j**-H, 3.09 – 2.96 (m, 4H) **k**-H, 2.79 – 2.70 (m, 8H) **l**-H, 1.81 – 1.73 (m, 2H) **m**-H, 1.72 – 1.61 (m, 6H) **n**-H, 1.61 – 1.50 (m, 10H) **o**-H, 1.44 – 1.27 (m, 12H) **p**-H, 1.21 (s, 8H) **q**-H.



---

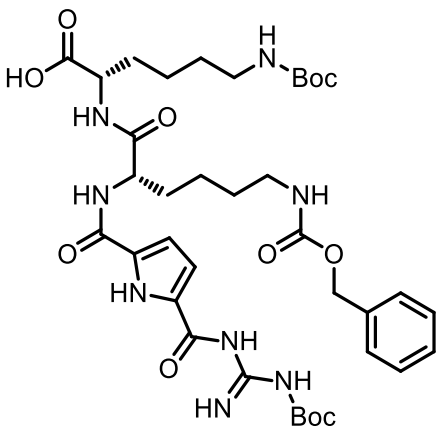
^{13}C-NMR (151 MHz, DMSO-d_6) δ [ppm] = 171.4 (s) **a**, 171.2 (s) **b**, 159.7 (s) **c**, 159.1 (s) **d**, 155.6 (s) **e**, 132.4 (s) **f**, 125.6 (s) **g**, 115.9 (s) **h**, 113.6 (s) **i**, 53.0 (s) **j**, 52.5 (s) **k**, 38.5 (s) **l**, 38.4 (s) **m**, 31.3 (s) **n**, 31.1 (s) **o**, 29.0 (s) **p**, 28.7 (s) **q**, 26.5 (s) **r**, 26.4 (s) **s**, 26.3 (s) **t**, 22.5 (s) **u**, 22.3 (s) **v**.



IR (ATR): [cm^{-1}] = 3311 (NH), 3235 (NH), 3045 (CH arom.), 2929 (CH aliph.), 2856 (CH aliph.), 1687 (C=O). **ESI-HRMS** m/z (%): 254 $[M+4H]^{4+}$ (100), 339 $[M+3H]^{3+}$ (98), 507 $[M+2H]^{2+}$ (20). Calculated for $C_{46}H_{82}N_{18}O_8$ $[M+2H]^{2+}$: 507.3276; found: 507.3278.

Compound 7

Compound **7** was synthesized according to **GP-A**. 1.00 g (1.55 mmol/g, 1.00 eq.) 2-chlorotrityl resin, 2.28 g (4.87 mmol, 3.14 eq.) Fmoc-Lys(Boc)-OH, 2.38 g (4.74 mmol, 3.06 eq.) Fmoc-Lys(Cbz)-OH, 1.85 g (4.64 mol, 3.00 eq.) Boc-GCP-OH, 2x 2.42. g (4.65 mmol, 3.00 eq.) PyBOP and 3x 1.60 ml (9.19 mmol, 5.93 eq.) DIPEA were used during the SPPS. 1.11 g (1.41 mmol, 91% based on the resin loading) of **7** as white solid were isolated. **RF** = 0.64 (MeOH:DCM 1:9). **mp.:** 90°C (decomposition).



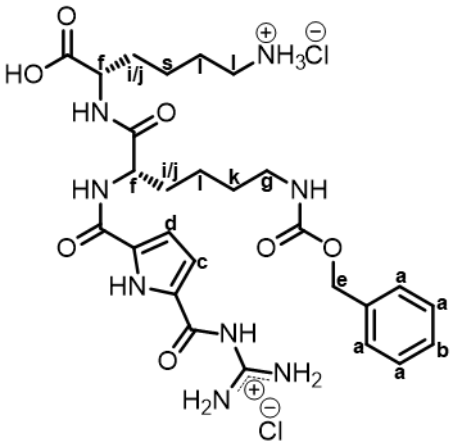
^{1}H-NMR (400 MHz, DMSO-d_6) δ [ppm] = 11.65 (s, 2H, -N**H**-), 9.32 (s, 1H, -N**H**-), 8.57 (s, 1H, -N**H**-), 8.41 (d, *J* = 8.0 Hz, 1H, -N**H**-), 8.19 (d, *J* = 7.6 Hz, 1H, -N**H**-), 7.39 –

7.26 (m, 5H, =C**H**-), 7.23 (t, *J* = 5.6 Hz, 1H, -N**H**-), 6.89 – 6.79 (m, 2H, =C**H**-), 6.76 (t, *J* = 5.5 Hz, 1H, -N**H**-), 5.01 (s, 2H, O-C**H₂**-), 4.46 (td, *J* = 8.7, 5.1 Hz, 1H, >C**H**-), 4.17 – 4.06 (m, 1H, >C**H**-), 3.03 – 2.92 (m, 2H, -C**H₂**-), 2.92 – 2.82 (m, 2H, -C**H₂**-), 1.78 – 1.64 (m, 2H, -C**H₂**-), 1.64 – 1.52 (m, 2H, -C**H₂**-), 1.46 (s, 9H, -C**H₃**), 1.43 – 1.24 (m, 17H, -C**H₃** + -C**H₂**-). **^{13}C-NMR** (101 MHz, DMSO-d_6) δ [ppm] = 173.6 (s, -**C**(O)OH), 171.9 (s, >**C**=O), 159.4 (s, >**C**=O), 158.5 (s, >**C**=O), 156.0 (s, >**C**=O), 155.6 (s, >**C**=O), 137.3 (s, >**C**=C), 128.3 (s, =**C**H-), 127.7 (s, =**C**H-), 113.7 (s, =**C**H-), 112.9 (s, =**C**H-), 77.4 (s, -**C**$(CH_3)_3$), 65.1 (s, O-**C**H_2-), 52.5 (s, >**C**H-), 51.9 (s, >**C**H-), 31.7 (s, -**C**H_2-), 30.7 (s, -**C**H_2-), 29.2 (s, -**C**H_2-), 29.1 (s, -**C**H_2-), 28.3 (s, -**C**H_3), 27.8 (s, -**C**H_3), 22.9 (s, -**C**H_2-), 22.8 (s, -**C**H_2-).
IR (ATR): [cm^{-1}] = 3315 (NH), 3182 (NH), 3086 (CH arom.), 2978 (CH aliph.), 2933 (CH aliph.), 2866 (CH aliph.), 1687 (C=O). **ESI-HRMS** m/z (%): 787 $[M+H]^+$ (100), 687 $[M-C_5H_7O_2]^+$ (17), 453 $[M-C_{18}H_{21}O_4]^+$ (12). Calculated for $C_{37}H_{55}N_8O_{11}$ $[M+H]^+$: 787.3985; found: 787.3978.

Compound 1GC

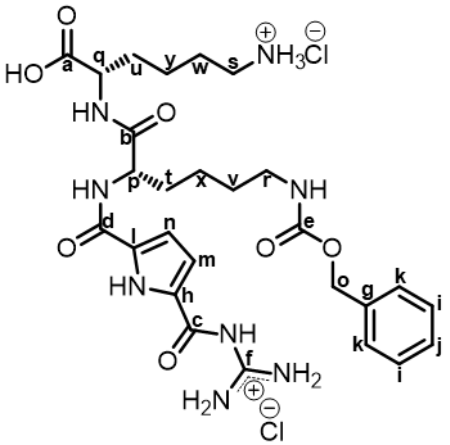
1.00 g (1.27 mmol, 1.00 eq.) **7** were implemented according to **GP-C**. The product was cleaned *via* MPLC (gradient: 30:70 → 60:40 MeOH:H_2O (acidified with 1‰ TFA), 80 min.). 0.13 g (0.18 mmol, 15%) of a white solid was isolated. **R_t** HPLC RP18 (Gradient 10:90 → 100 MeOH:H_2O (acidified with 1‰ TFA), 30 min.) = 17.5 min. **mp.:** 170°C (decomposition).

^{1}H-NMR (600 MHz, MeOD-d_4) δ [ppm] = 7.34 – 7.30 (m, 4H) **a**-H, 7.30 – 7.25 (m, 1H) **b**-H, 7.24 (d, *J* = 4.1 Hz, 1H) **c**-H, 7.00 (d, *J* = 4.1 Hz, 1H) **d**-H, 5.04 (s, 2H) **e**-H, 4.52 – 4.42 (m, 2H) **f**-H, 3.14 (t, *J* = 6.7 Hz, 2H) **g**-H, 2.93 (t, *J* = 7.5 Hz, 2H) **h**-H, 2.01 – 1.87 (m, 2H) **i**-H, 1.87 – 1.73 (m, 2H) **j**-H, 1.73 – 1.62 (m, 2H) **k**-H, 1.61 – 1.44 (m, 6H) **l**-H.



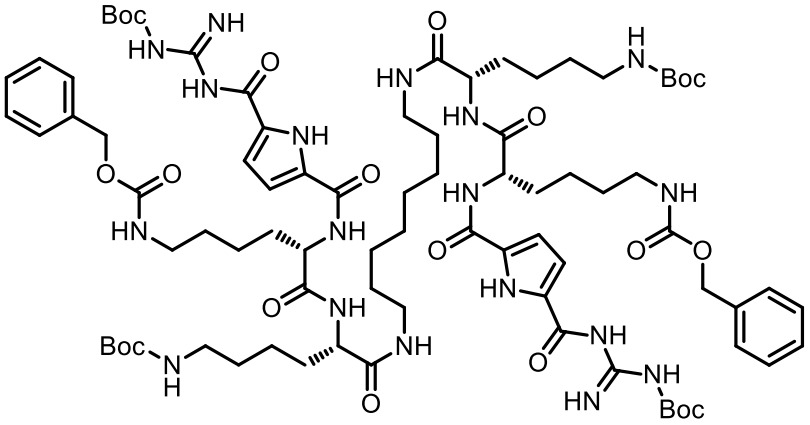
---

^{13}C-NMR (151 MHz, MeOD-d_4) δ [ppm] = 175.1 (s) **a**, 174.8 (s) **b**, 162.2 (s) **c**, 161.5 (s) **d**, 159.2 (s) **e**, 157.4 (s) **f**, 138.6 (s) **g**, 133.5 (s) **h**, 129.6 (s) **i**, 129.1 (s) **j**, 128.8 (s) **k**, 127.3 (s) **l**, 116.5 (s) **m**, 114.3 (s) **n**, 67.5 (s) **o**, 55.4 (s) **p**, 53.1 (s) **q**, 41.6 (s) **r**, 40.7 (s) **s**, 32.6 (s) **t**, 32.1 (s) **u**, 30.7 (s) **v**, 27.9 (s) **w**, 24.3 (s) **x**, 23.9 (s) **y**.



IR (ATR): [cm^{-1}] = 3321 (NH), 3074 (CH arom.), 2943 (CH aliph.), 2873 (CH aliph.), 1662 (C=O). **ESI-HRMS** m/z (%): 294 $[M+2H]^{2+}$ (100), 453 $[M+C_8H_7O_2]^+$ (49), 587 $[M+H]^+$ (32). Calculated for $C_{27}H_{39}N_8O_7$ $[M+H]^+$: 587.2936; found: 587.2943

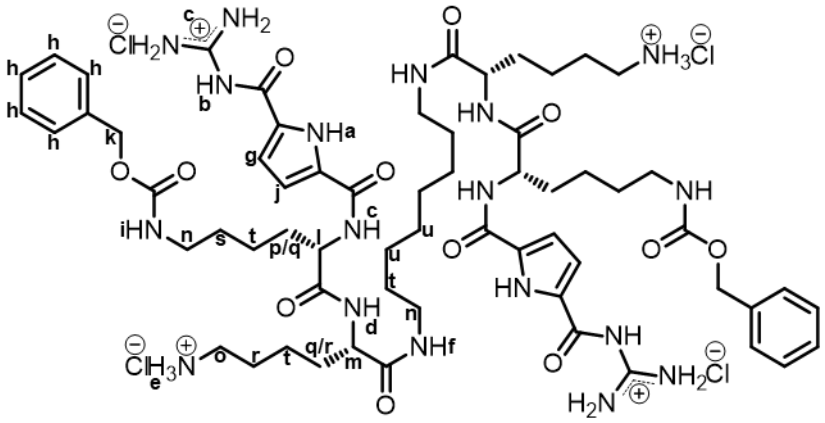
Compound 8



0.28 g (1.94 mmol, 1.00 eq.) **3**, 3.21 g (4.08 mmol, 2.10 eq.) **7**, 2.52 g (4.84 mmol, 2.50 eq.) PyBOP and 1.00 ml (5.74 mmol, 2.96 eq.) DIPEA in 60 ml DCM were implemented according to **GP-B**. No further purification was performed. 1.72 g (1.02 mmol, 53%) of a white solid was isolated. **RF** = 0.59 (MeOH:DCM 1:9). **mp.:** 149°C (decomposition).
^{1}H-NMR (400 MHz, DMSO-d_6) δ [ppm] = 11.72 (s, 2H, -N**H**-), 10.88 (s, 2H, -N**H**-), 9.34 (s, 2H, -N**H**-), 8.51 (s, 2H, -N**H**-), 8.41 (d, *J* = 7.4 Hz, 2H, -N**H**-), 7.96 (d, *J* = 7.9 Hz, 2H, -N**H**-), 7.76 (s, 2H, -N**H**-), 7.40 – 7.25 (m, 10H, =C**H**-), 7.23 (t, *J* = 5.4 Hz, 2H, -N**H**-), 6.83 (s, 4H, =**C**H-), 6.72 (t, *J* = 5.2 Hz, 2H, -N**H**-), 4.99 (s, 4H, O-C**H$_2$**-), 4.47 – 4.31 (m, 2H, >C**H**-), 4.20 – 4.07 (m, 2H, >C**H**-), 3.12 – 2.91 (m, 8H, -C**H$_2$**-), 2.91 – 2.78 (m, 4H, -C**H$_2$**-), 1.81 – 1.56 (m, 6H, -C**H$_2$**- + -C**H$_2$**-), 1.56 – 1.44 (m, 20H, -C**H$_2$**- + -C**H$_3$**), 1.43 – 1.26 (m, 34H, -C**H$_3$** + -C**H$_2$**-), 1.21 (s, 12H, -C**H$_2$**-). **^{13}C-NMR** (101 MHz, DMSO-d_6) δ [ppm] = 171.6 (s, >**C**=O), 171.2 (s, >**C**=O), 159.6 (s, >**C**=O), 158.5 (s, >**C**=O), 156.1 (s, >**C**=O), 155.6 (s, >**C**=O), 137.3 (s, >**C**=C), 128.3 (s, =**C**H-), 127.7 (s, =**C**H-), 113.0 (s, =**C**H-), 77.3 (s, -**C**$(CH_3)_3$), 65.1 (s, O-**C**H_2-), 53.1 (s, >**C**H-), 52.6 (s, >**C**H-), 40.2 (s, -**C**H_2-), 38.5 (s, -**C**H_2-), 31.8 (s, -**C**H_2-), 31.5 (s, -**C**H_2-), 29.3 (s, -**C**H_2-), 29.2 (s, -**C**H_2-), 29.0 (s, -**C**H_2-), 28.8 (s, -**C**H_2-), 28.3 (s, -**C**H_3), 27.8 (s, -**C**H_3), 26.4 (s, -**C**H_2-), 23.0 (s, -**C**H_2-), 22.8 (s, -**C**H_2-). **IR (ATR):** [cm^{-1}] = 3298 (NH), 3066 (CH arom.), 2974 (CH aliph.), 2931 (CH aliph.), 2862 (CH aliph.), 1693 (C=O). **ESI-HRMS** m/z (%): 841 $[M+2H]^+$ (100), 1682 $[M+H]^+$ (3). Calculated for $C_{82}H_{125}N_{18}O_{20}$ $[M+H]^+$: 1681.9312; found: 1681.9323.

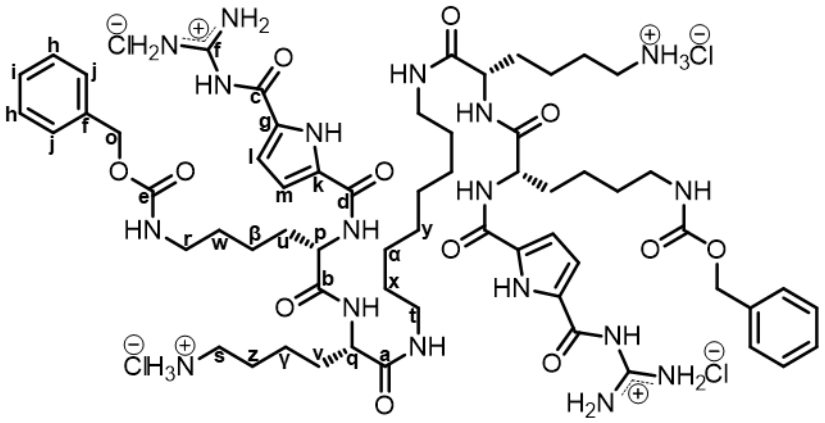
Compound 2GC

0.20 g (0.12 mmol, 1.00 eq.) **8** were implemented according to **GP-C**. The product was purified *via* MPLC (gradient: 40:60 → 70:30 MeOH:H_2O, 100 min.). 0.09 g (0.06 mmol, 44%) of a white solid were isolated. **R_t** HPLC RP18 (Gradient 10:90 → 100:0 MeOH:H_2O, 30 min.) = 21.6 min. **mp.:** 156 - 159°C.



^{1}H-NMR (600 MHz, DMSO-d_6) δ [ppm] = 12.54 (s, 2H) **a**-H, 12.14 (s, 2H) **b**-H, 8.78 – 8.43 (m, 10H) **c**-H, 8.14 (d, *J* = 8.1 Hz, 2H) **d**-H, 7.88 (s, 6H) **e**-H, 7.84 (t, *J* = 5.4 Hz, 2H) **f**-H, 7.59 (s, 2H) **g**-H, 7.38 – 7.27 (m, 10H) **h**-H, 7.25 (t, *J* = 5.7 Hz, 2H) **i**-H, 6.92 – 6.86 (m, 2H) **j**-H, 5.00 (s, 4H) **k**-H, 4.46 – 4.36 (m, 2H) **l**-H, 4.22 – 4.15 (m, 2H) **m**-H, 3.09 – 2.94 (m, 8H) **n**-H, 2.76 – 2.68 (m, 4H) **o**-H, 1.76 – 1.69 (m, 2H) **p**-H, 1.69 – 1.59 (m, 4H) **q**-H, 1.56 – 1.47 (m, 6H) **r**-H, 1.47 – 1.39 (m, 4H) **s**-H, 1.38 – 1.25 (m, 12H) **t**-H, 1.23 – 1.13 (m, 8H) **u**-H.

^{13}C-NMR (151 MHz, DMSO-d_6) δ [ppm] = 171.5 (s) **a**, 171.2 (s) **b**, 159.8 (s) **c**, 159.0 (s) **d**, 156.1 (s) **e**, 155.6 (s) **f**, 137.3 (s) **g**, 132.3 (s) **h**, 128.3 (s) **i**, 127.7 (s) **j**, 127.6 (s) **k**, 125.6 (s) **l**, 115.8 (s) **m**, 113.6 (s) **n**, 65.1 (s) **o**, 53.2 (s) **p**, 52.4 (s) **q**, 40.2 (s) **r**, 38.5 (s) **s**, 38.4 (s) **t**, 31.5 (s) **u**, 31.4 (s) **v**, 29.2 (s) **w**, 29.0 (s) **x**, 28.7 (s) **y**, 26.5 (s) **z**, 26.3 (s) **α**, 22.9 (s) **β**, 22.3 (s) **γ**.

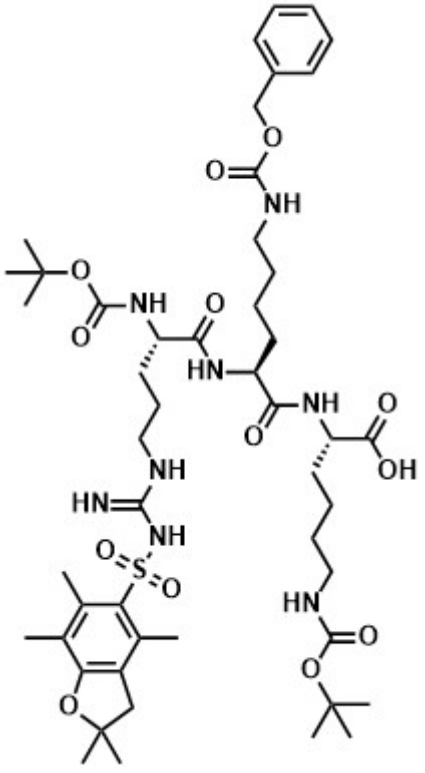


IR (ATR): [cm^{-1}] = 3275 (NH), 3091 (CH arom.), 3066 (CH arom.), 3035 (CH arom.), 2927 (CH aliph.), 2858 (CH aliph.), 1687 (C=O).

ESI-HRMS m/z (%): 428 $[M+3H]^{3+}$ (100), 383 $[(M+3H)-C_8H_6O_2]^{3+}$ (53), 641 $[M+2H]^{2+}$ (29), 507 $[(M+2H)-C_{16}H_{12}O_4]^{2+}$ (9), 321 $[M+4H]^{4+}$ (5), 1182 $[M+H]^{+}$ (2). Calculated for $C_{62}H_{94}N_{18}O_{12}$ $[M+2H]^{2+}$: 641.3644; found: 641.3593.

Compound 9

Compound **9** was synthesised according to **GP-A**. 1.00 g (1.55 mmol, 1.00 eq.) 2-chlorotrytril resin was used for the synthesis and 1.00 mL triethylamine (0.7 mmol, 4.62 eq.) were added for every coupling step. For the first step 2.28 g (4.74 mmol, 3.06 eq.) Fmoc-Lys(Boc)-OH, for the second step 2.38 g (4.64 mmol, 3.06 eq.) Fmoc-Lys(Z)-OH with 2.42 g (4.65 mmol, 3.00 eq.) PyBOP and for the last step 2.44 g (4.64 mmol, 3.00 eq.) Boc-Arg(Pbf)-OH and 1.77 g (4.65 mmol, 3.00 eq.) HATU were coupled on the resin. Yield: 0.32 g (0.32 mmol 20% based on the resin loading) The product was isolated without further purification. **RT** = 0.8 (MeOH:DCM 1:9). **mp.:** 114°C (decomposition).



^{1}H-NMR (400 MHz, $CDCl_3$) δ [ppm] = 8.04 - 8.35 (m, 1 H, -N**H**-), 7.45 - 8.01 (m, 3, -N**H**-), 7.21 - 7.37 (m, 6 H, -N**H**-=C**H**$_{arom.}$-), 6.09 - 6.74 (m, 1 H, =N**H**), 5.42 - 6.04 (m, 2 H, -N**H**-), 5.05 (br s, 2 H, -C**H**$_{2\ benzyl}$-), 4.04 - 4.57 (m, 3 H, C-**H**$_α$), 2.77 - 3.37 (m, 8 H, -C**H**$_2$-), 2.42 - 2.62 (m, 4 H, -C**H**$_2$-) 2.03 - 2.11 (m, 2 H, -C**H**$_2$-), 1.65 - 1.94 (m, 6 H, -C**H**$_2$-), 1.29 - 1.63 (m, 24 H, -C**H**$_3$), (t, *J* =7.02 Hz, 6 H, -C**H**$_3$).

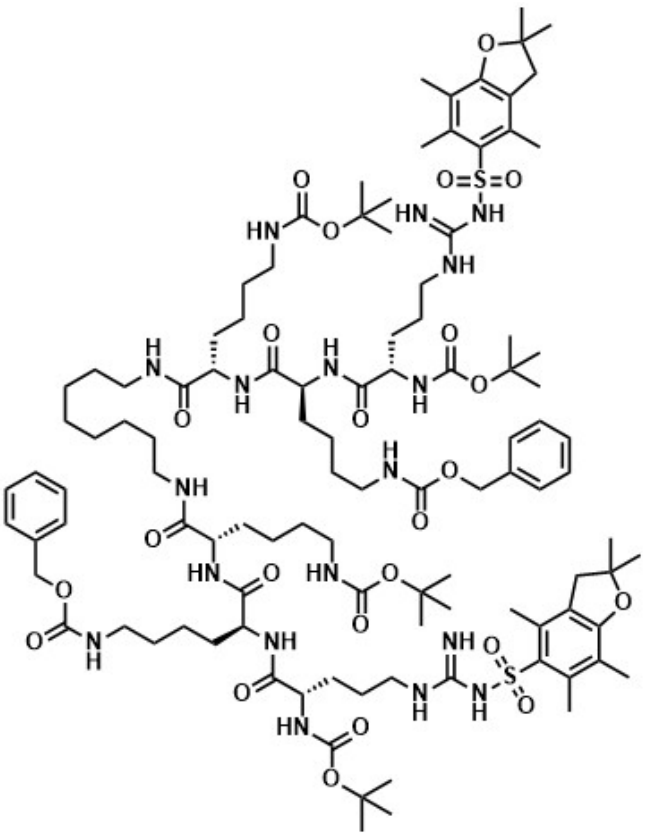
^{13}C-NMR (101 MHz, $CDCl_3$) δ [ppm] = 148.3 (s, -**C**=O), 128.6 (s, =**C**H$_{arom}$-), 128.1 (s, =**C**H$_{arom}$-), 128.1 (s, =**C**H$_{arom}$-), 77.4 (s, **C**H), 66.0 (s, -**C**H$_{2\ benzyl}$-), 40.5 (s, -**C**H$_2$-), 29.3 (s, -**C**H$_2$-), 28.7 (s, -**C**H$_2$-), 28.7 (s, -**C**H$_2$-), 28.6 (s, **C**H$_2$), 28.5 (s, -**C**H$_2$-), 15.4 (s, -**C**H$_3$), 12.6 (s, -**C**H$_3$).

IR (ATR): [cm^{-1}] = 3329 (NH), 3066 (CH arom.), 2935 (CH aliph.), 2868 (CH aliph.), 1654 (N-C=O), 1540 (O-C=O).

ESI-HRMS m/z (%): 1017 $[M+H]^{+}$ (100), 1040 $[M+Na]^{+}$ (51). Calculated for $C_{49}H_{76}N_8O_{13}S$ $[M+H]^{+}$: 1017.5325; found: 1017.5329.

Compound 10

11.4 mg octanediamine (0.10 mmol, 1 eq.) 200.0 mg of Compound **9** (0.20 mmol, 2 eq.), 102.0 mg of PyBOP (0.20 mmol, 2 eq.) and 0.4 mL Et_3N (2.89 mmol, 30 eq.) were dissolved in 10.0 mL DMF and stirred for 19 h. The solvent was added to 100.0 mL of ice cold water. The precipitate was filtered off and dried under vacuum. Yield: 83.0 mg (39.0 µmol, yield 39%) of the off-white product was obtained without further purification. **RF** = 0.68 (MeOH:DCM 1:9). **mp.:** 125°C (decomposition).



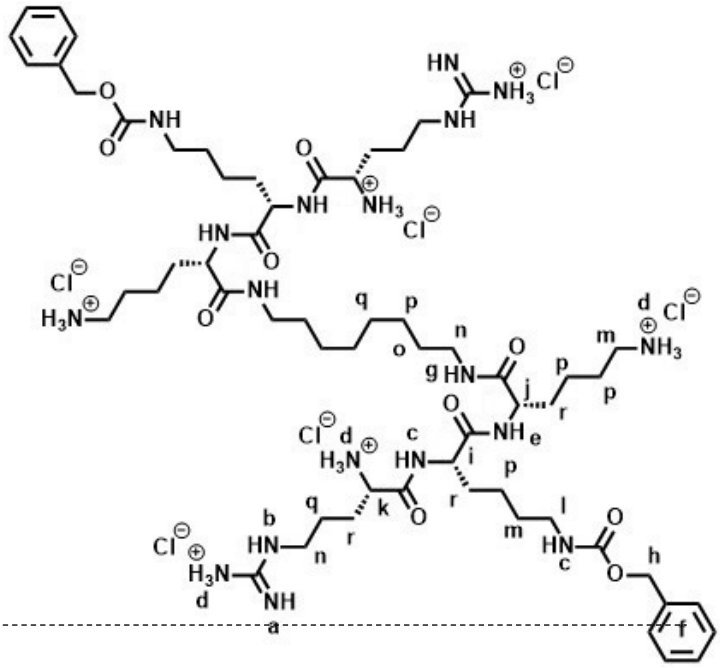
^{1}H-NMR (400 MHz, DMSO-d_6) δ [ppm] = 7.66 - 8.10 (m, 6 H, -N**H**-), 7.24 - 7.40 (m, 14 H, C-H_{arom}, -N**H**-), 7.12 - 7.24 (m, 2 H, -N**H**-), 6.80 - 7.03 (m, 2 H, -N**H**-), 6.57 - 6.80 (m, 2 H, -N**H**-), 6.19 - 6.56 (m, 2 H, -N**H**-), 5.76 (s, 2 H, =N**H**), 4.92 - 5.07 (m, 8 H, -C**H**$_2$-), 4.09 - 4.32 (m, 6 H, -C**H**$_α$-), 2.89 - 3.09 (m, 22 H, -C**H**$_2$-), 2.47 (s, 6 H, -C**H**$_3$), 2.42 (s, 6 H, -C**H**$_3$), 1.43 - 1.68 (m, 18 H, -C**H**$_2$-), 1.29 - 1.43 (m, 62 H, -C**H**$_3$, -C**H**$_2$-), 1.09 - 1.28 (m, 16 H, -C**H**$_2$-).

^{13}C-NMR (101 MHz, DMSO-d_6) δ [ppm] = 194.2 (s, >**C**=O), 194.1 (s, >**C**=O), 194.0 (s, >**C**=O), 194.0 (s, >**C**=O),193.9 (s, >**C**=O), 133.7 (s, >**C**=C), 133.7 (s, >**C**=C), 133.6 (s, >**C**=C), 130.4 (s, =**C**H$_{arom.}$-), 130.4 (s, =**C**H$_{arom}$-), 130.32 (s, =**C**H$_{arom}$-), 130.21 (s, =**C**H$_{arom}$-), 130.11 (s, =**C**H$_{arom}$-), 123.52 (s, **C**-CH$_3$), 100.08 (s, >**C**H-), 100.0 (s, >**C**H-), 99.9 (s, >**C**H-), 69.9 (s, -**C**H$_2$-), 69.8 (s, -**C**H$_2$-), 69.7 (s, -**C**H$_2$-), 69.6 (s, -**C**H$_2$-), 67.2 (s, -**C**H$_2$-), 66.9 (s, -**C**H$_2$-), 66.5 (s, -**C**H$_2$-), 66.4 (s, -**C**H$_2$-), 66.3 (s, -**C**H$_2$-), 66.3 (s, -**C**H$_2$-), 10.6 (s, -**C**H$_3$), 9.4 (s, -**C**H$_3$), 9.3 (s, -**C**H$_3$), 7.0 (s, -**C**H$_3$), 6.1 (s, -**C**H$_3$-), 6.0 (s, -**C**H$_3$), 5.9 (s, -**C**H$_3$).

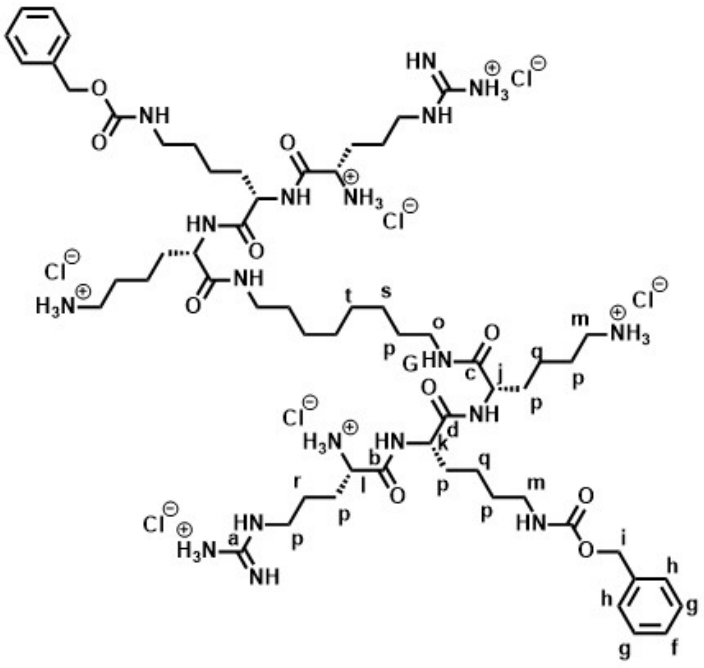
IR (ATR): [cm^{-1}] = 3309 (NH), 3064 (CH arom.), 2972 (CH aliph.), 2933 (CH aliph.), 2868 (CH aliph.), 1685 (C=O).
ESI-HRMS m/z (%): 1072 $[M+2H]^+$ (100). Calculated for $C_{106}H_{168}N_{18}O_{24}S_2$ $[M+2H]^+$: 1072.1048; found: 1072.1053.

Compound 2RC

23 mg (0.01 mmol) of compound **10** was stirred in 10 mL of a 1:1 mixture of TFA and DCM. After 1 h, the solvent was evaporated, and the residue was poured into 12 mL of cold diethylether. The precipitate was filtered off, dried under vacuum, and purified with a preparative HPLC (80/20 H_2O/ACN + 0.01% TFA over 26 min R_t = 8,5 min). 15 mg (<0,01 mmol, yield 78 %) of a white solid were obtained. 10 mg of the product was lyophilized 5-times with 0.05 M HCl to obtain the HCl-salt. **R_t** HPLC RP18 (Gradient 10:90 → 100:0 ACN: H_2O + 0,1% TFA, 30 min.) = 13.96 min. **mp.:** 184°C (decomposition).



^{1}H-NMR (600 MHz, DMSO-d_6) δ [ppm] = 8.75 - 8.80 (d, 2H) **a**-H, 8.26 - 8.30 (s, 2H) **b**-H, 8.20 (d, 2 H) **c**-H, 7.95 - 8.15 (m, 12 H) **d**-H, 7.90 (m, 2 H) **e**-H, 7.25 - 7.40 (m, 10 H) **f**-H, 7.18 (s, 2 H) **g**-H, 5.0 (s, 4 H) **h**-H, 4.30 (m, 2 H) **i**-H, 4.15 (m, 2 H) **j**-H, 3.80 (m, 2 H) **k**-H, 3.15 (m, 4 H) **l**-H, 2.95 (m, 8 H) **m**-H, 2.75 (m, 8 H) **n**-H, 1.77 (m, 4 H) **o**-H, 1.60 (m, 16 H) **p**-H, 1.35 (m, 8 H) **q**-H, 1.25 (m, 12 H) **r**-H.



^{13}C-NMR (151 MHz, DMSO-d_6) δ [ppm] = 171.44 (s) **a**, 168.83 (s) **b**, 157.48 (s) **Cc**, 156.54 (s) **d**, 167.72 (s) **e**, 129.31 (s) **f**, 129.11 (s) **g**, 128.82 (s) **g**, 128.23 (s) **h**, 128.19 (s) **h**, 65.59 (s) **i**, 53.52 (s) **j**, 52.93 (s) **k**, 51.89 (s) **l**, 46.66 (s) **m**, 31.94 (s) **o**, 29.64 (s) **p**, 29.46 (s) **p**, 29.20 (s) **p**, 28.71 (s) **p**, 26.85 (s) **q**, 26.82 (s) **q**, 24.34 (s) **r**, 23.16 (s) **s**, 22.61 (s) **t**.

IR (ATR): [cm^{-1}] = 3294 (NH), 3079 (CH arom.), 3066 (CH arom.), 3032 (CH arom.), 2972 (CH aliph.), 2887 (CH aliph.), 1684 (C=O).

ESI-HRMS m/z (%): 1238 $[M+H]^+$ (100), 1104 $[M+H-1Cbz]^+$ (87)), 970 $[M+H-2Cbz]^+$ (16). Calculated for $C_{60}H_{104}N_{18}O_{10}$ $[M+H]^+$: 1237.8256; found: 1237.8280.

Mass spectra of the final compounds

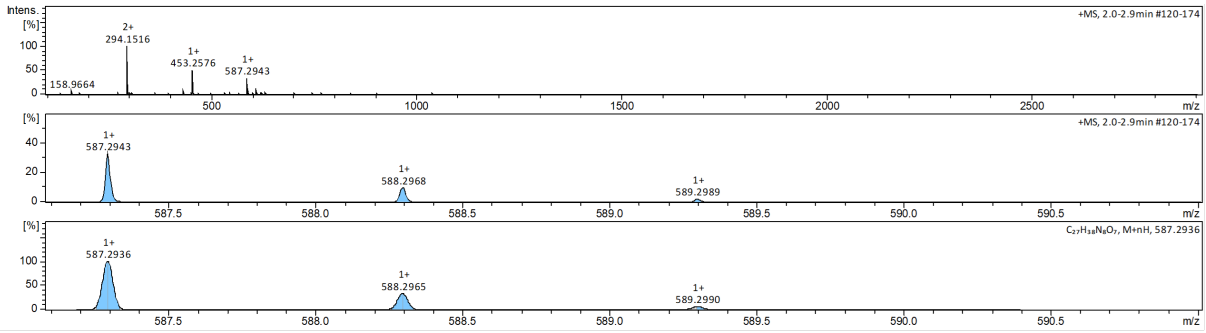


Figure S3: HR-ESI-MS spectrum of compound **1GC**. The upper panel represents the whole spectrum, whereas the spectrum shown in the middle is a close-up and the lower panel is the calculated spectrum.

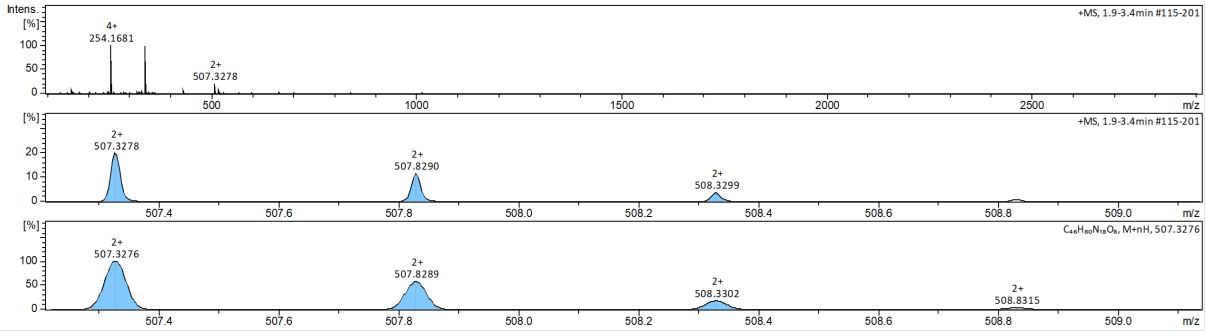


Figure S4: HR-ESI-MS spectrum of compound **2G**. The upper panel represents the whole spectrum, whereas the spectrum shown in the middle is a close-up and the lower panel is the calculated spectrum.

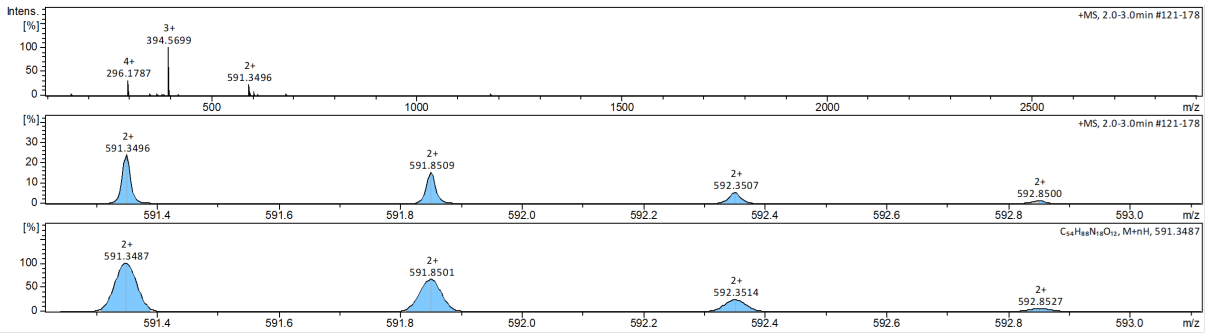


Figure S5: HR-ESI-MS spectrum of compound **2GA**. The upper panel represents the whole spectrum, whereas the spectrum shown in the middle is a close-up and the lower panel is the calculated spectrum.

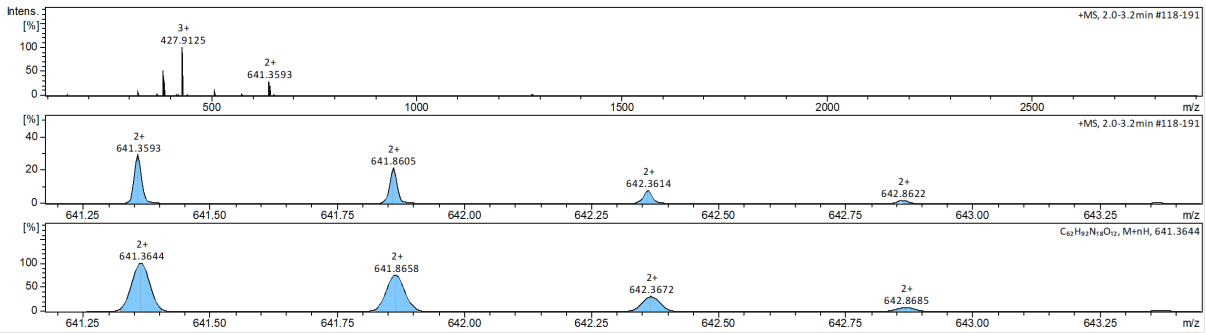


Figure S6: HR-ESI-MS spectrum of compound **2GC**. The upper panel represents the whole spectrum, whereas the spectrum shown in the middle is a close-up and the lower panel is the calculated spectrum.

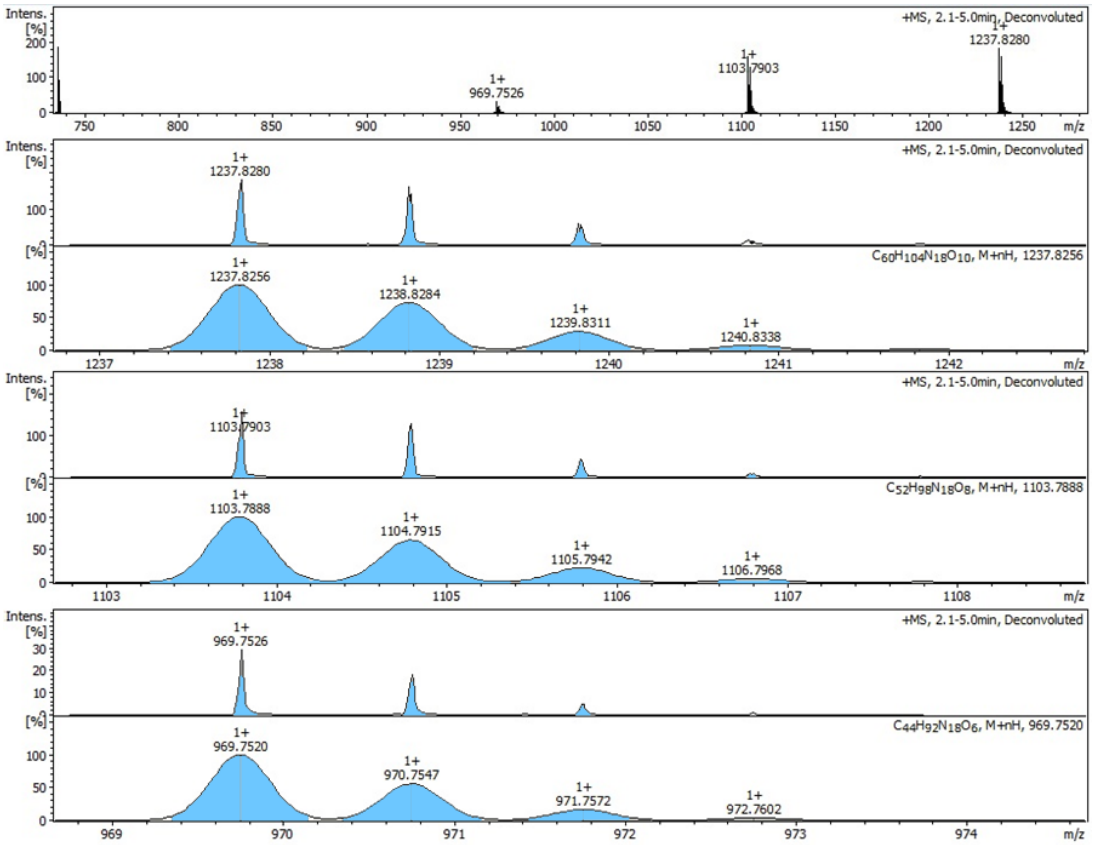


Figure S7: HR-ESI-MS spectrum of compound **2RC**. The first spectrum represents the deconvoluted spectra. Underneath this spectrum are three pairs of spectra representing the deconvoluted and simulated MS spectra of the fully Cbz-protected **2RC**, the mono Cbz-protected **2RC**, and the unprotected **2RC** respectively. The loss of Cbz groups can be attributed to the ionization process.

NMR spectra

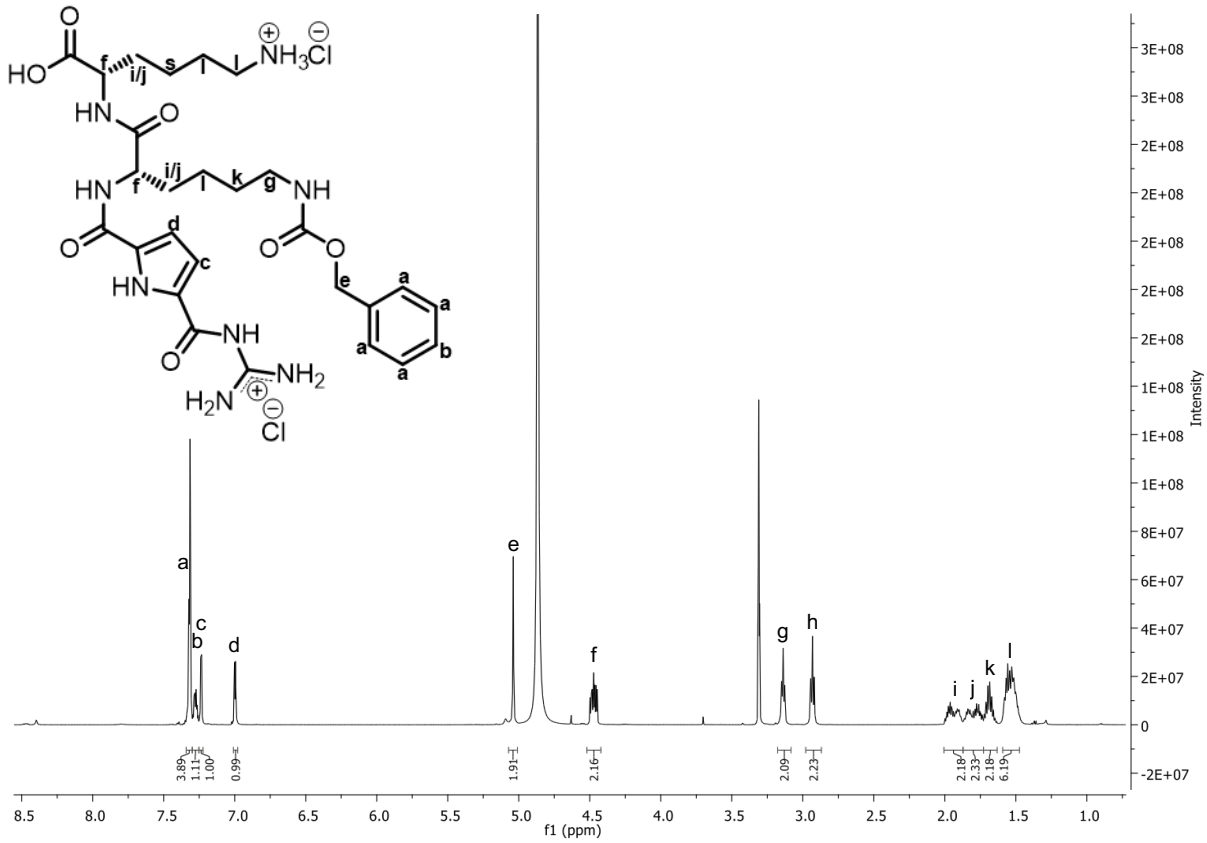


Figure S8: 1H-NMR of compound **1GC** in MeOD-d_4 (600 MHz).

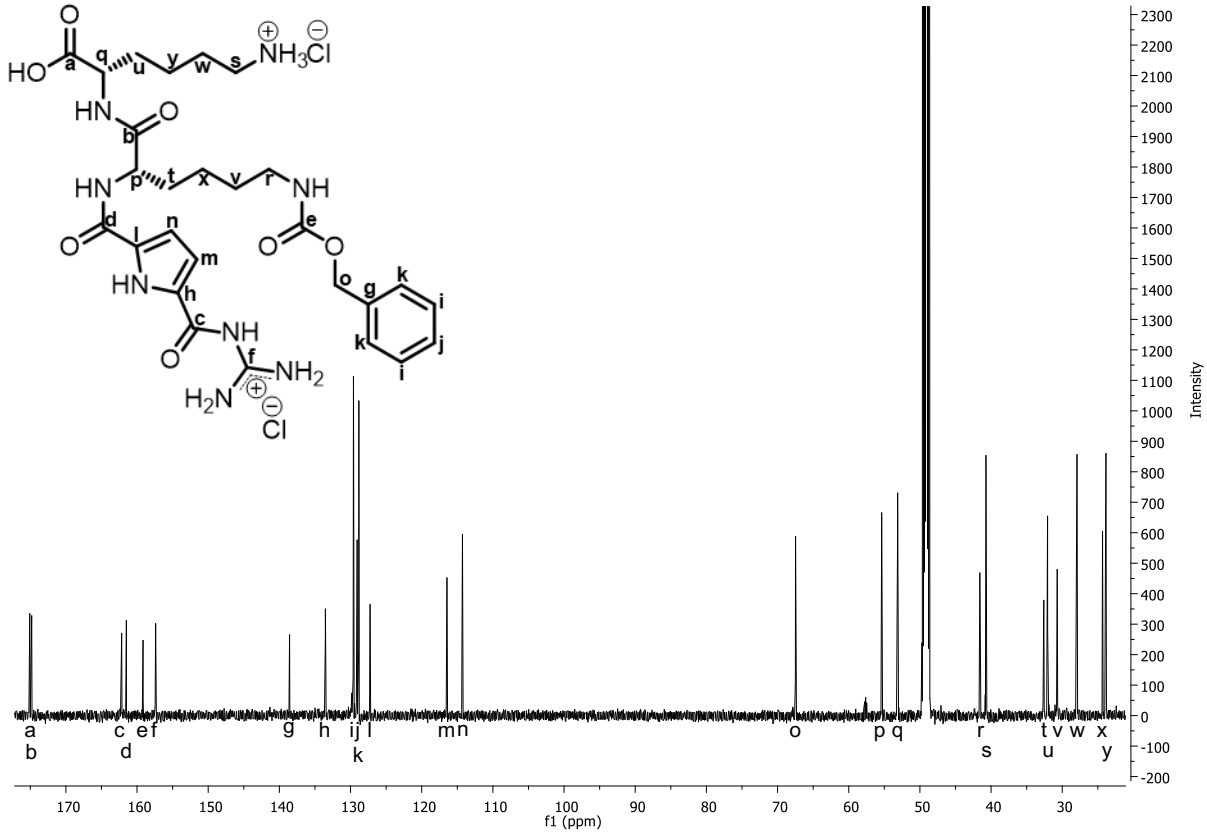


Figure S9: ^{13}C-NMR of compound **1GC** in MeOD-d_4 (151 MHz).

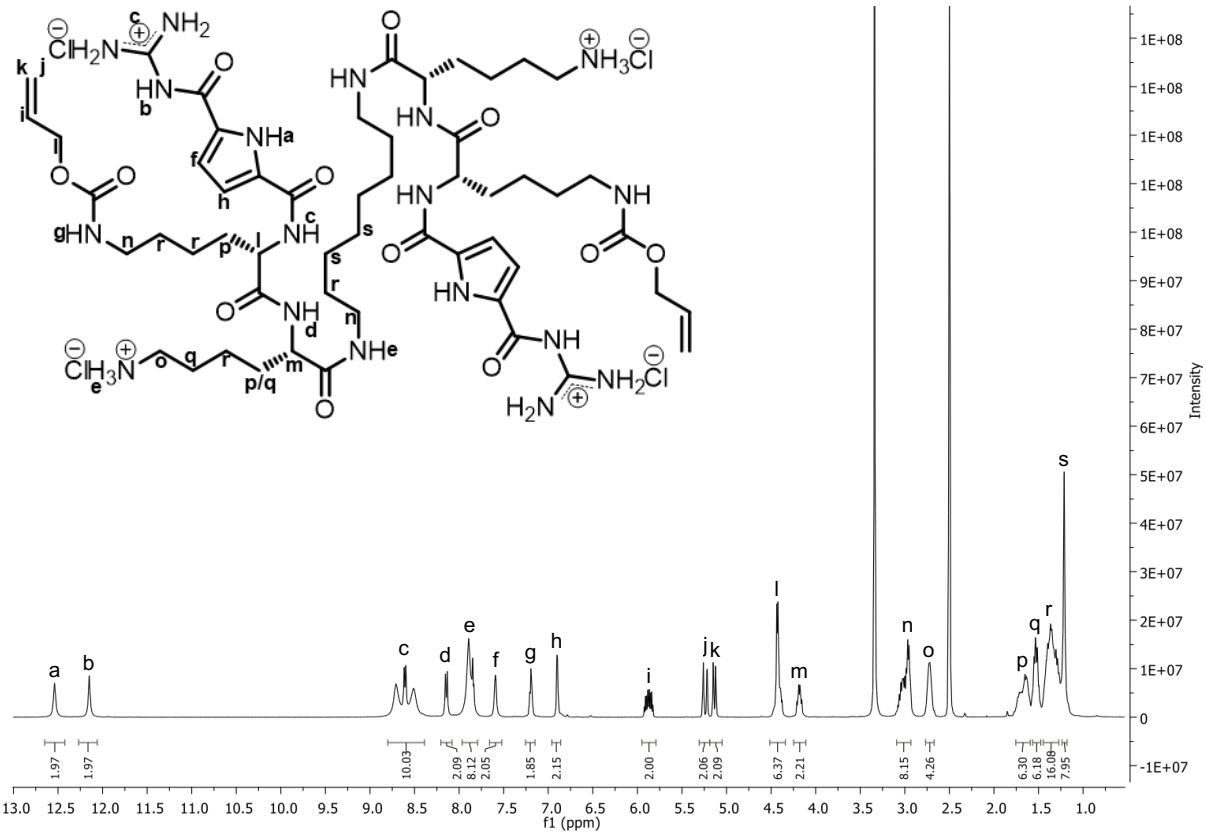


Figure S10: ^{1}H-NMR of compound **2GA** in DMSO-d_6 (400 MHz).

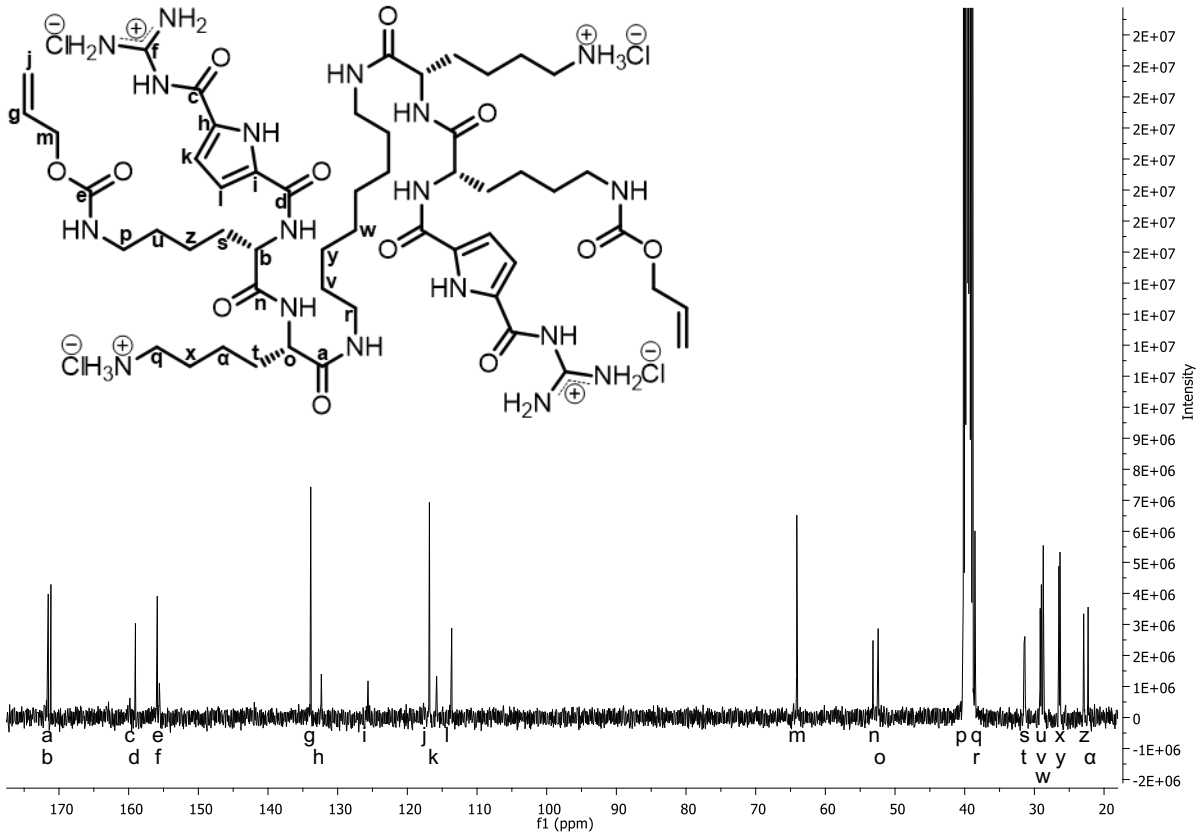


Figure S11: ^{13}C-NMR of compound **2GA** in DMSO-d_6 (101 MHz).

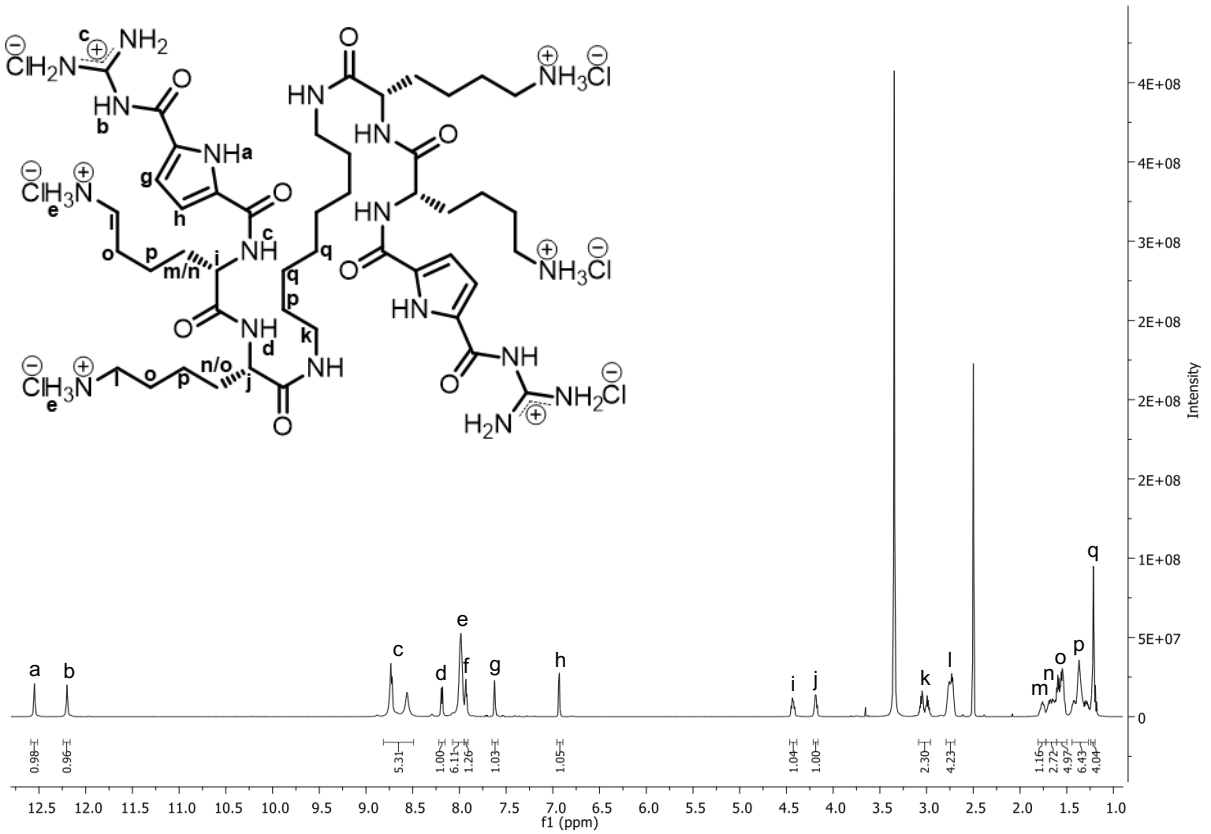


Figure S12: ^{1}H-NMR of compound **2G** in DMSO-d_6 (600 MHz).

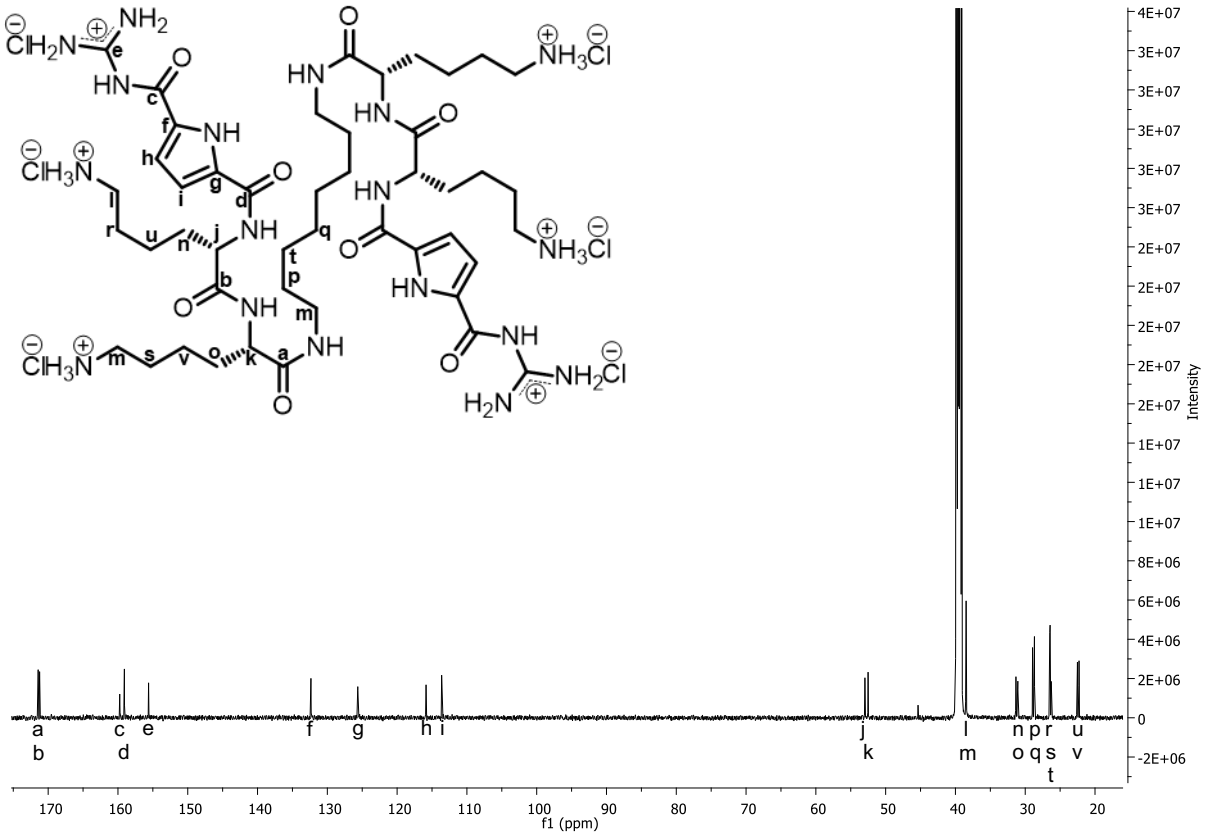


Figure S13: ^{13}C-NMR of compound **2G** in DMSO-*d6* (151 MHz).

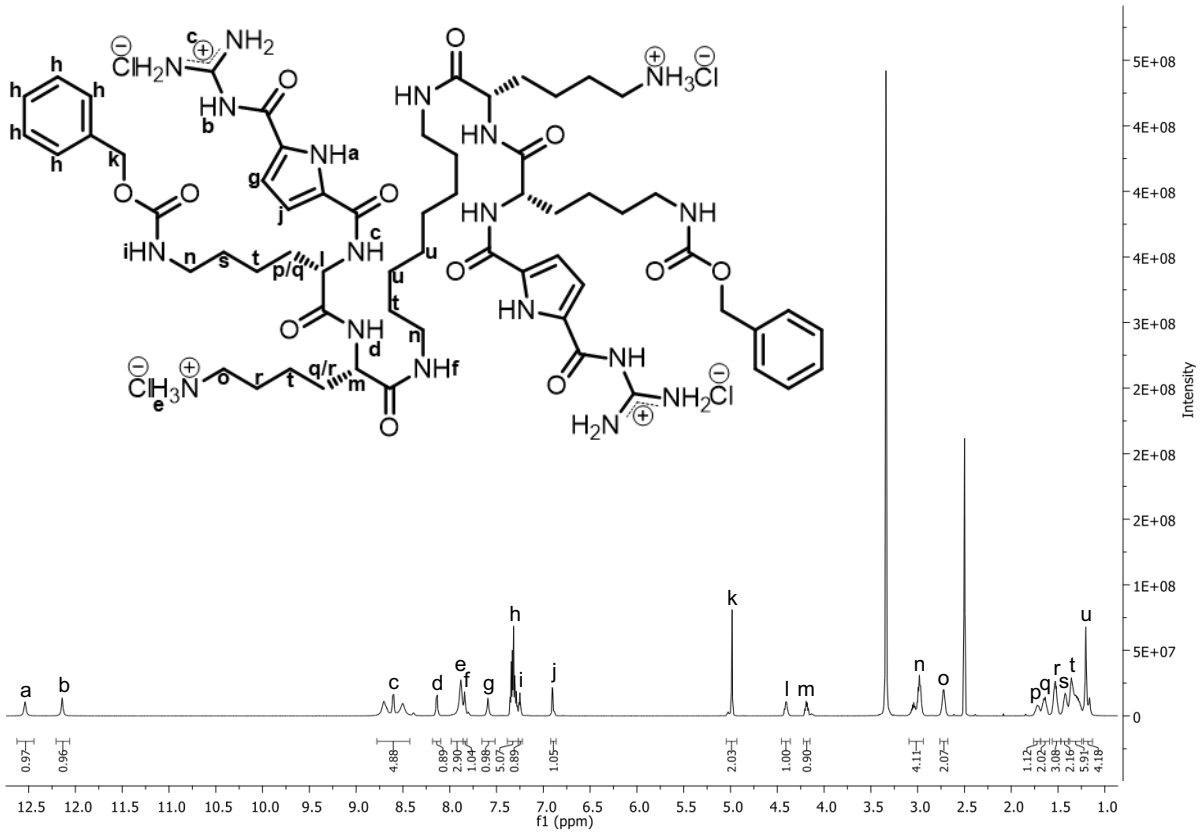


Figure S14: ^{1}H-NMR of compound **2GC** in DMSO-d_6 (600 MHz).

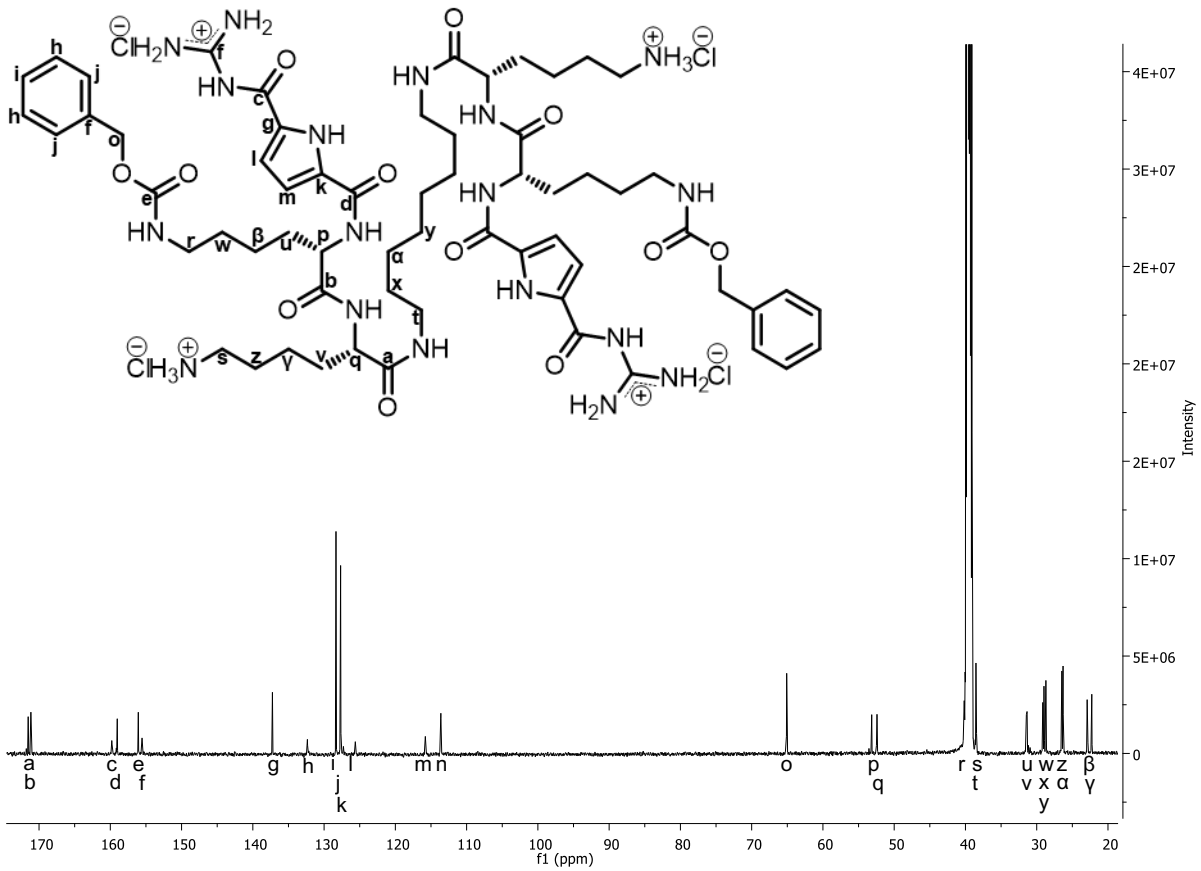


Figure S15: ^{13}C-NMR of compound **2GC** in DMSO-d_6 (151 MHz).

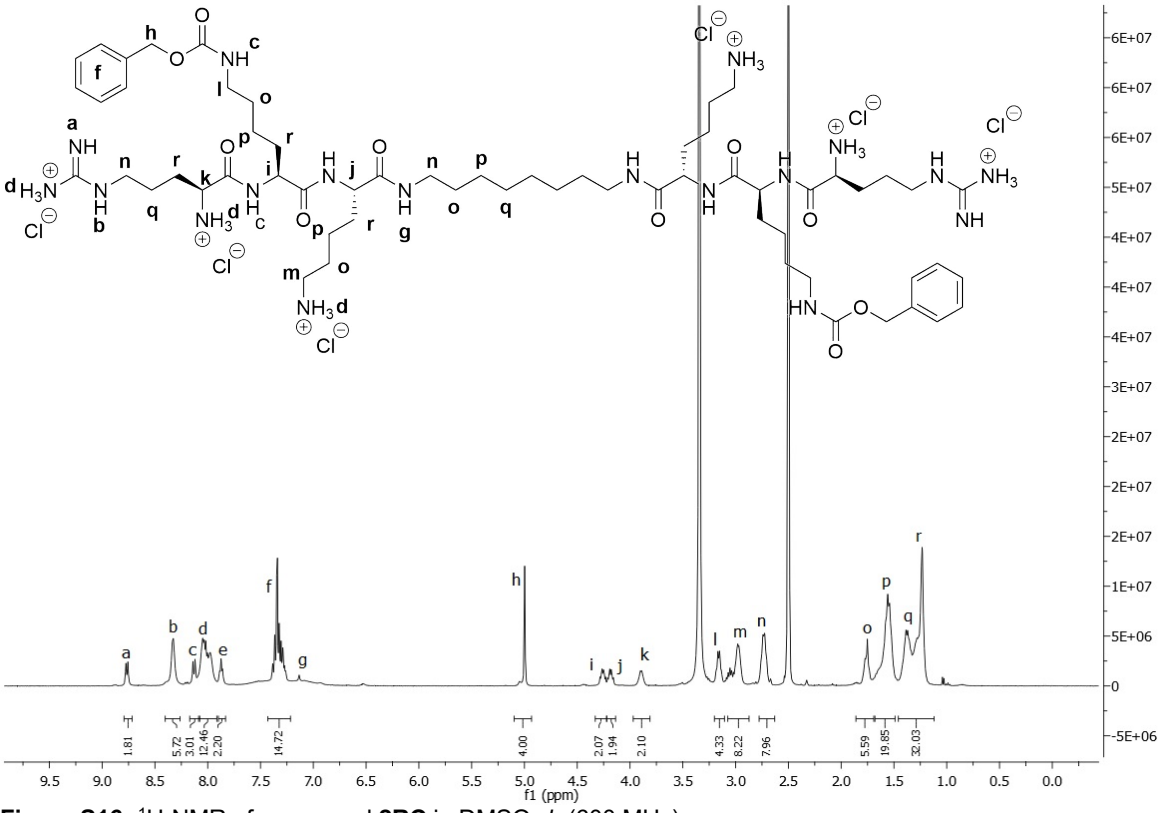


Figure S16: ^{1}H-NMR of compound **2RC** in DMSO-d_6 (600 MHz).

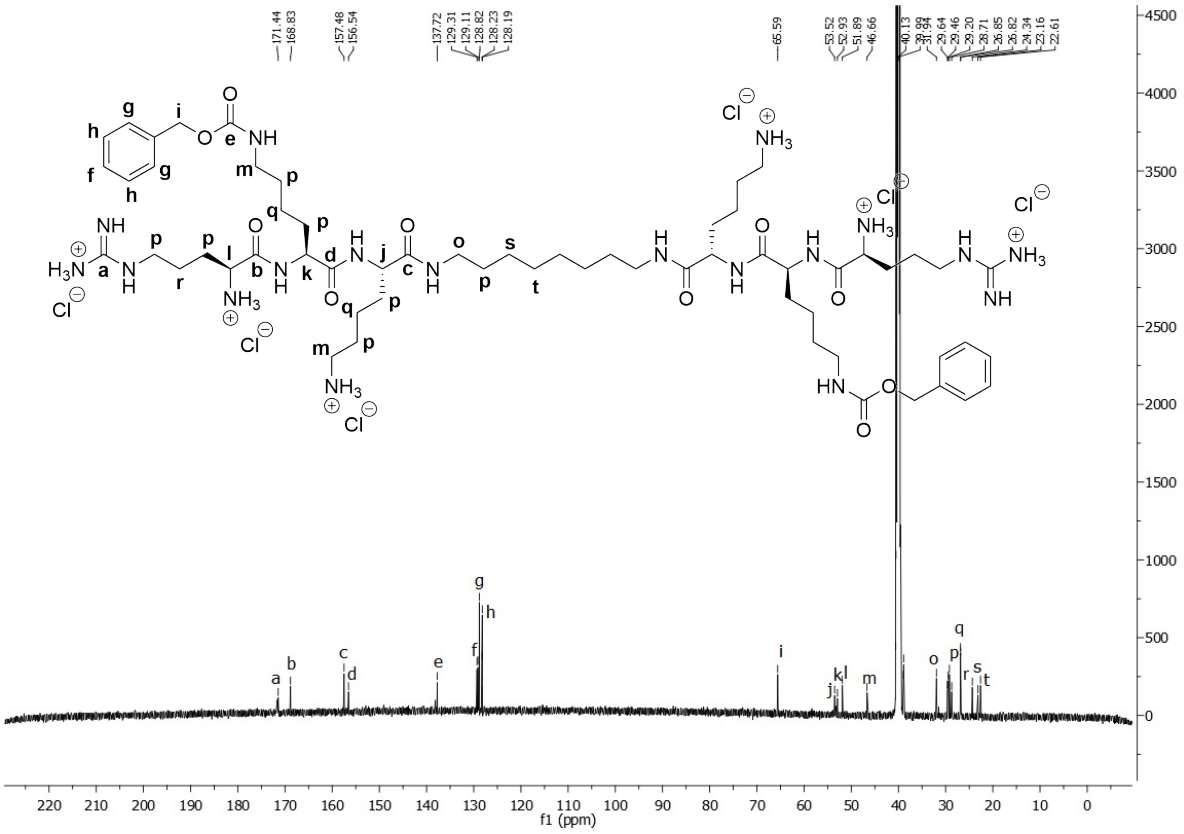


Figure S17: ^{13}C-NMR of compound **2RC** in DMSO-d_6 (151 MHz).

HPLC chromatograms

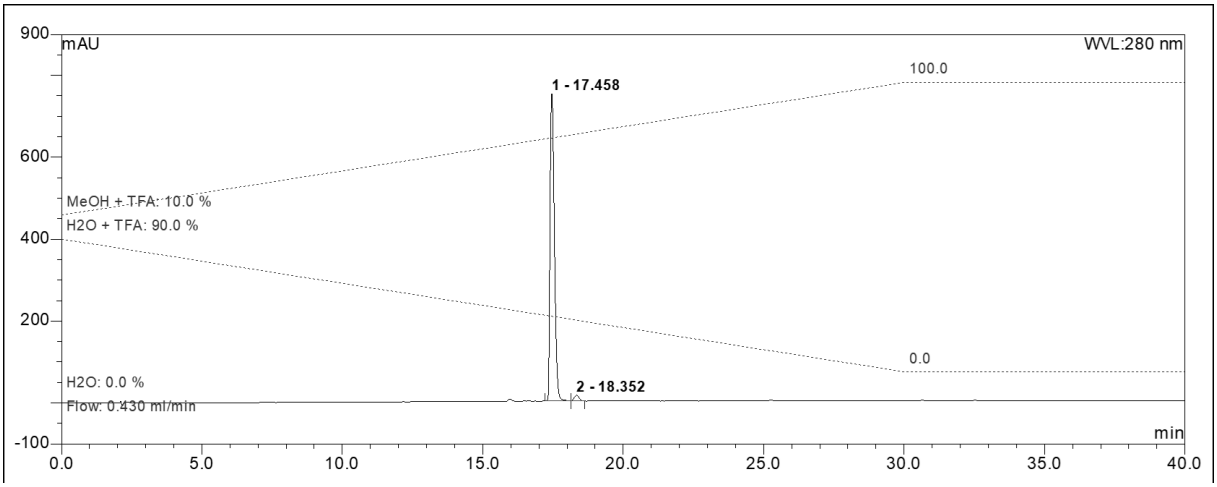


Figure S18: HPLC chromatogram of compound **1GC**.

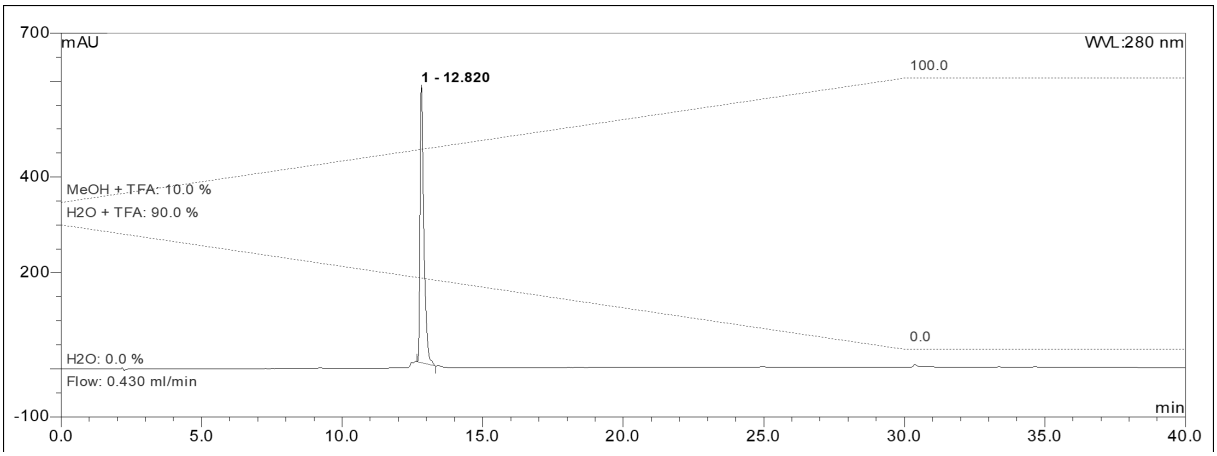


Figure S19: HPLC chromatogram of compound **2G**.

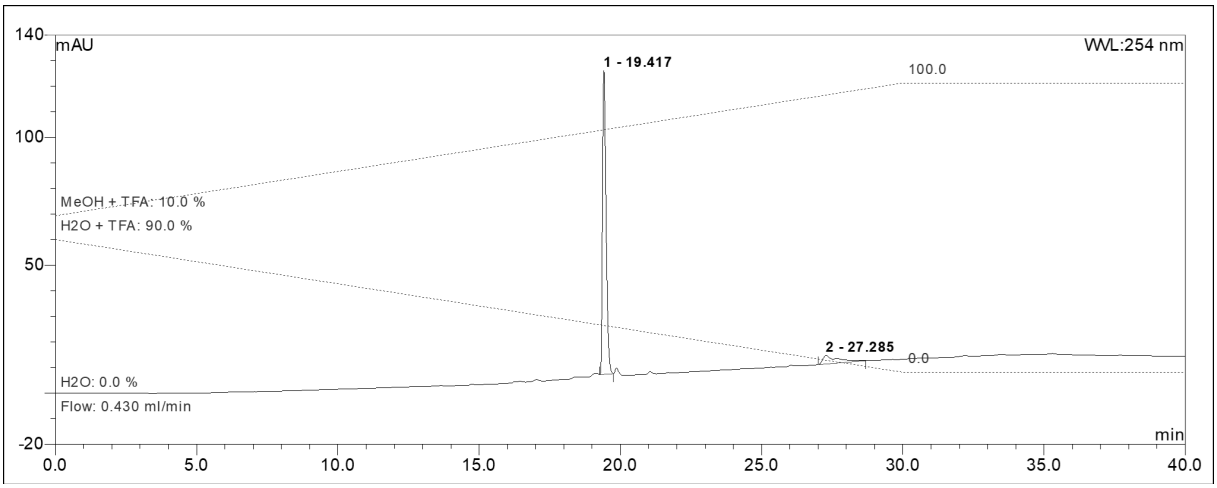


Figure S20: HPLC chromatogram of compound **2GA**.

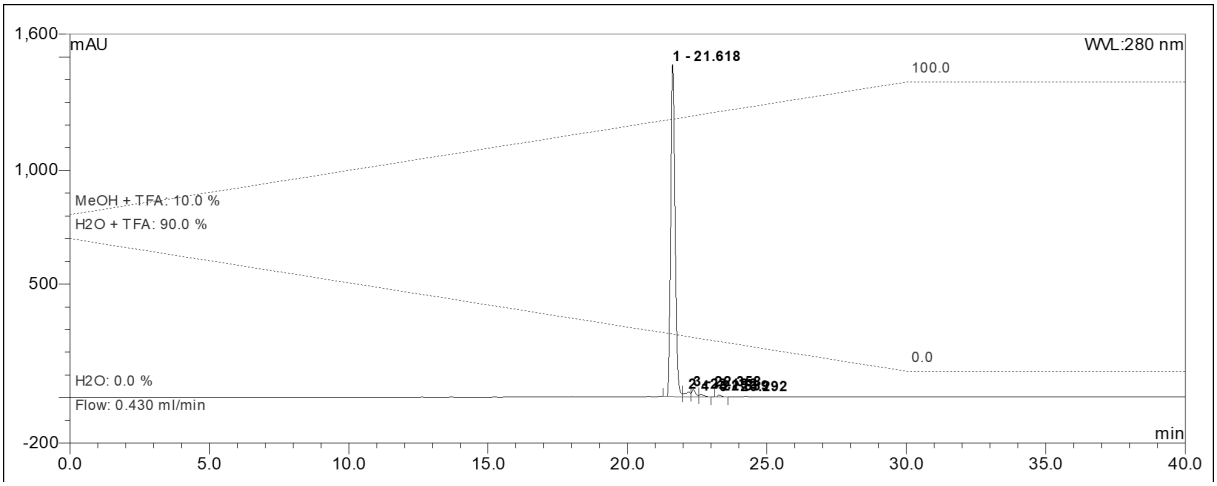


Figure S21: HPLC chromatogram of compound **2GC**.

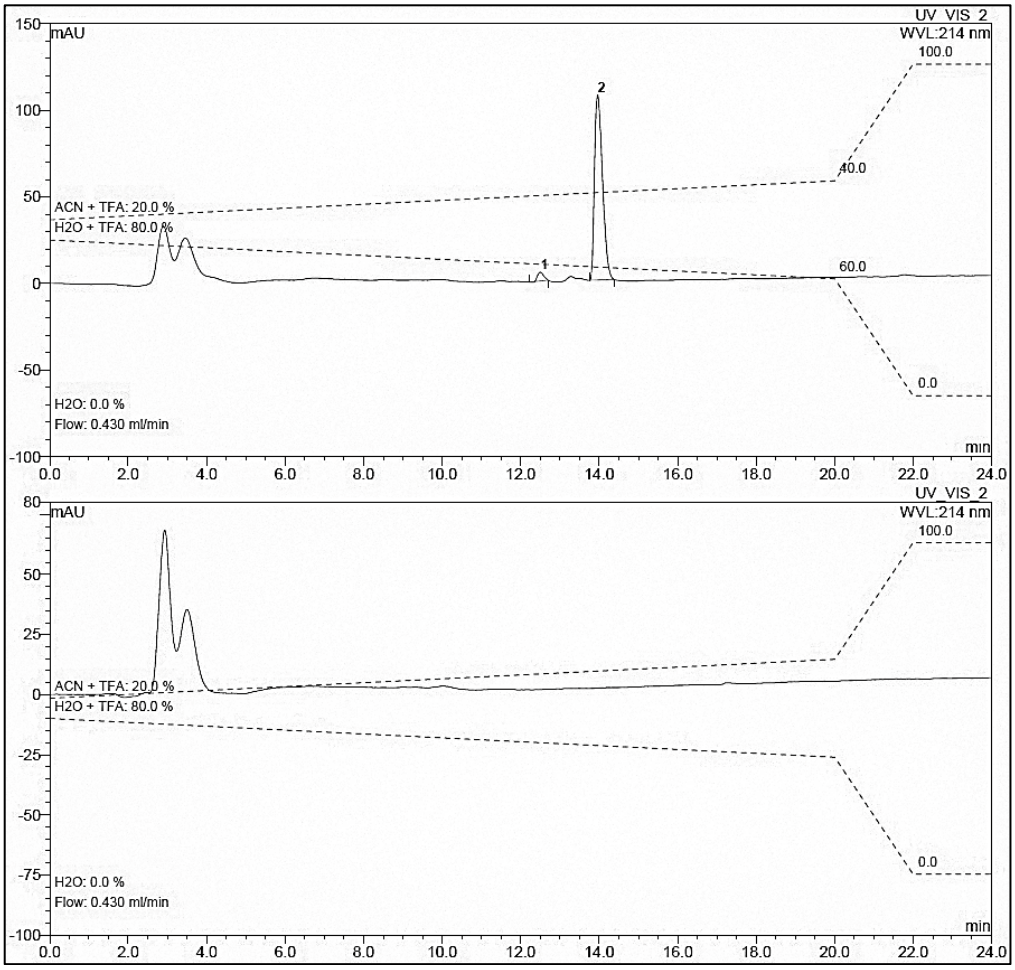


Figure S22: HPLC chromatogram of compound **2RC** (upper chromatogram) and the control chromatogram showing the solvents residuals between 2 and 4 minutes (lower chromatogram).

Molecular docking

The program Maestro 11.5 Schroedinger was used for docking. The following 4 ligands were prepared with LigPrep. Tautomers and possible states at pH =7 +/-3 were generated. The specified chiralities were retained.

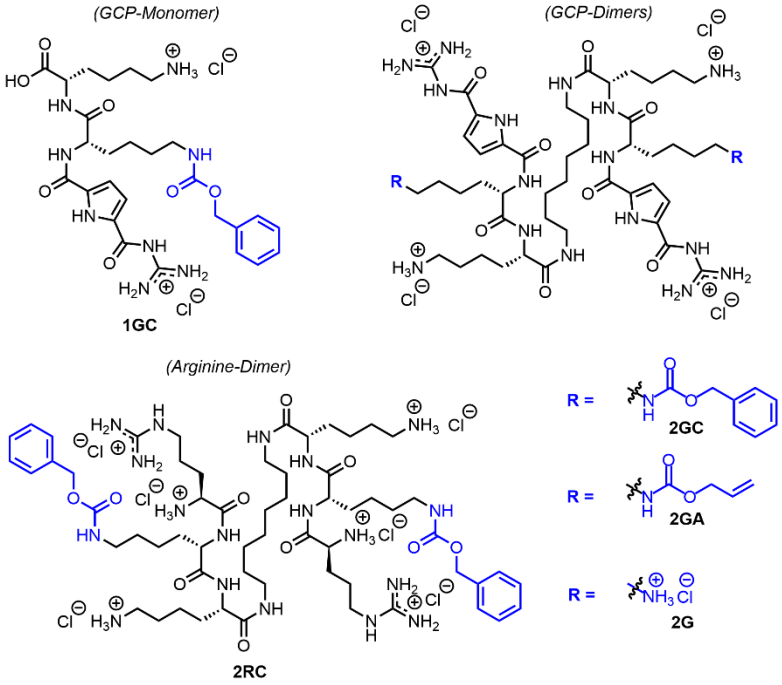


Figure S23: Compounds used for docking studies.

The crystal structure of Taspase1 was prepared with Protein Preparation.[2] Hydrogens, missing chains, and loops were added. Zero-order bonds to metals and disulfide bonds were created. Water was removed, which was beyond 5 Å away from het groups. Het states were created with the use of Epik (pH = 7+/-2).
A grid around the amino acids tyr234, arg190, asp228 and ser56 with a size of (36 Å)^3 was generated with glide grid generator (active center).

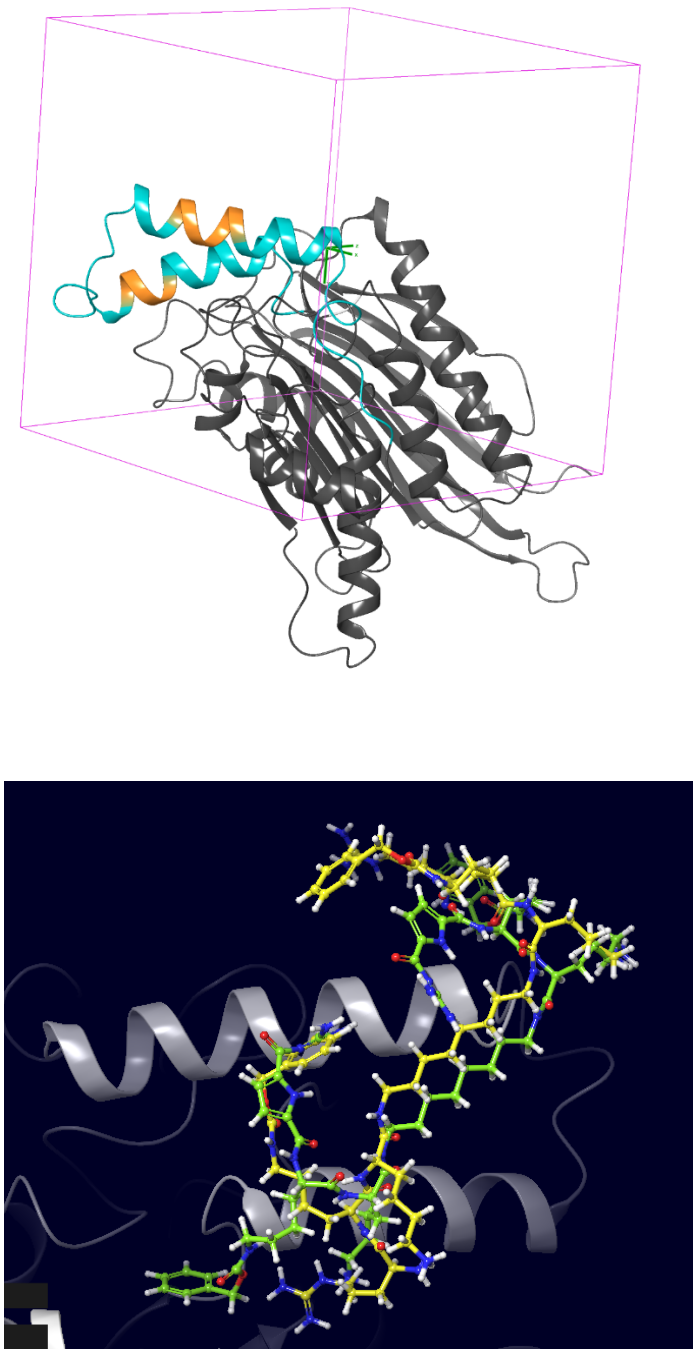


Figure S24: Top: Grid created around the active centre (36 A)^3. Amino acids 178-197, 203-216 and 221-233 coloured in turquoise and 197-202 and 217-220 in orange. Bottom: Docking of **2RC** (yellow) and **2GC** (green) demonstrates the population of the same region on the Taspase loop, although **2GC** binds tighter and in a more closed conformation.

The prepared ligands and grid were used for the ligand docking. The method was XP (extra precise). The molecule sampling was flexible. Additionally, the following conditions were chosen: sample nitrogen inversions and ring conformation, bias sampling of torsions for amides and add Epik state penalties to docking score. All protonation degrees were obtained using *Ligprep* as software.

Table S2: Docking scores calculated with Taspase1 and the corresponding ligand. Docking scores for the best conformations are indicated. The docking scores indicate how strongly a ligand can be stabilized by a certain protein in a defined volume (see pink cube, **Fig. S18**). The more negative the score, the higher the stabilizing energy. For **2G**, two additional free lysines allow for surplus electrostatic interactions. Due to its smaller size, **1GC** can access a cavity within the loop, specifically contributing to a more negative score. In contrast, **2GA**, **2GC** and **2RC** are characterized by rather comparable structures and thus docking scores. Indeed, direct comparison reveals that lack of the GCP unit in **2RC** might result in decreased binding.

Compound	Docking Score
2G	-9.004
1GC	-6.832
2GA	-6.515
2GC	-6.370
2RC	-5.634

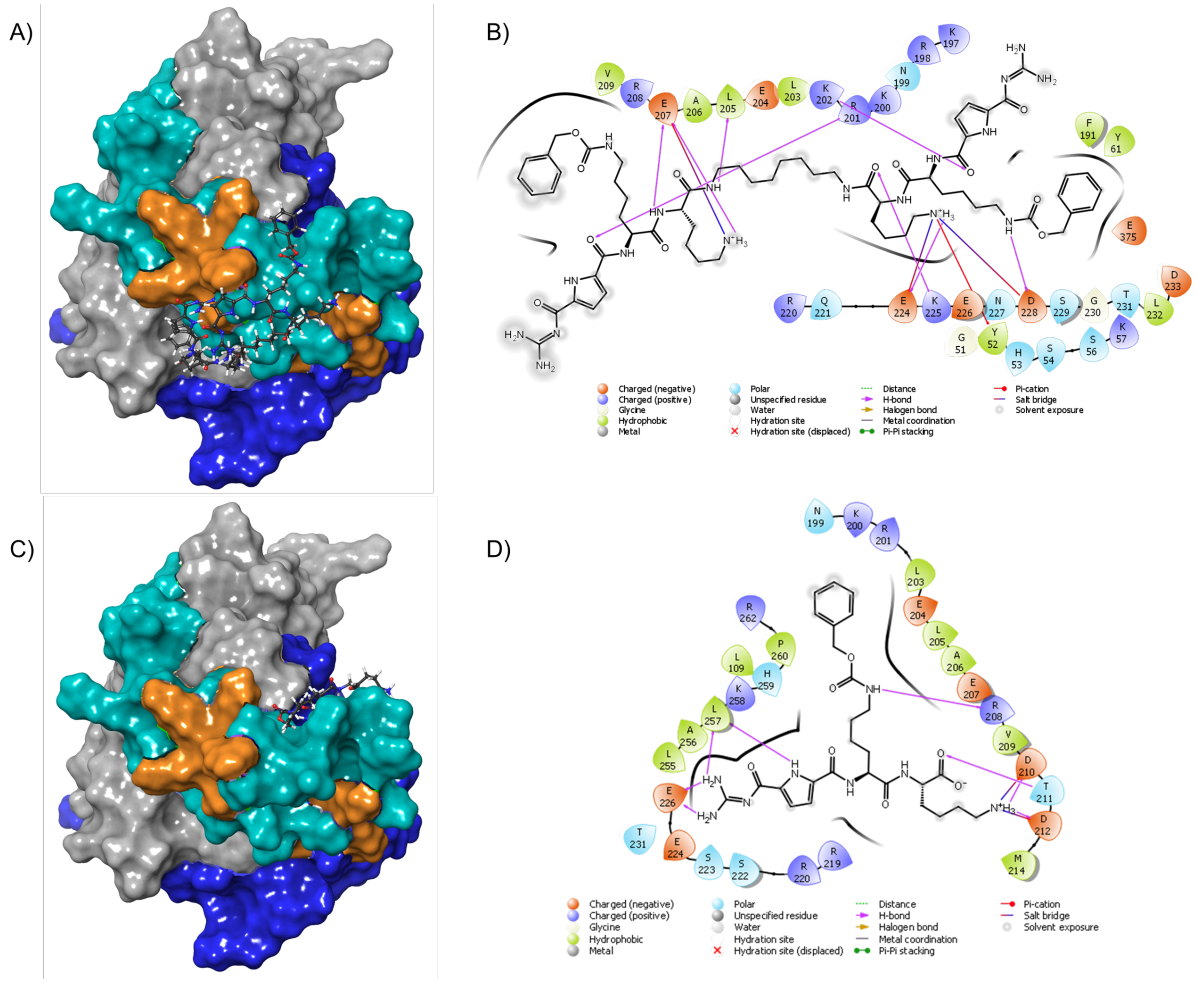


Figure S25: A) In a 3D model, **2GC** covers a large portion of the Taspase1 loop necessary for the Importin α interaction and thus indicates steric hindrance. The proenzyme comprises an α-unit (blue) and a β-unit (gray). The β-unit contains a loop region (turquoise) harboring the basic Taspase1/Importin α interface (^{197}KRNKRK^{202}, ^{217}KKRR^{220} orange), located on two neighboring helices (see Fig. S19). B) 2D ligand-protein interaction diagram reveals a variety of possible interactions of **2GC** and Taspase1 *via* hydrogen bonds (arrowed pink lines). Salt bridges (blue-violet lines) and solvent exposure (blurred grey circles) are also indicated together with polar (cyan), positively (red) and negatively (violet) charged as well as hydrophobic (green) amino acid residues of Taspase1 (aa52-57, 61, 198-209, 224-233). Of note, the hydrophobic cbz protecting group does not interact with the protein. C) In a 3D model of Taspase1 in complex with **1GC**, the monomer only covers a small portion of the loop compared to **2GC** (A). D) 2D ligand-protein interaction diagram revealing no significant interactions between **1GC** and the Importin α/Taspase1 interface. The colour code is the same as in (B), again, relevant Taspase1 amino acid residues are indicated (52-57, 61, 198-209, 224-233). Of note, the cbz group is also not supposed to interact with any hydrophobic groups of Taspase1.

Biological Assays

Cloning

The plasmid for inactive mutant Taspase1$_{D233/T234A}$ was generated as previously described.[3] The plasmid for wildtype Taspase1 was generated as previously described.[4] The plasmid for USF2-GFP was generated as previously described.[5] The gene for Importin α was amplified from a previously described pcDNA3-Importin α-HA plasmid by PCR.[6] During amplification, the ends were modified to introduce ApaI/BamHI restriction sites. Following subcloning into a blunt pJET1.2 vector (Thermo Fisher) according to the CloneJET PCR cloning kit (Thermo Fisher) and plasmid amplification and purification, the Importin α open reading frame was cloned into a modified pET-41b vector containing an N-terminal GST (Glutathione S-transferase) affinity fusion tag and a PreScission protease cleavage site (GeneArt) *via* ApaI/BamHI. The sequence was verified by sequencing (LCG Genomics).

Purification of recombinant proteins

pET22b-Taspase1$_{D233A/T234}$-His and pET22b-WT_Taspase1-His were expressed in *E. coli* BL21 (DE3). Cells were lysed using ultrasonic sheering and enzymatic lysis with lysozyme. The protein was purified using the His tag for affinity chromatography with a HisTrap FF column (GE Healthcare). Following imidazole elution, Taspase1-His-containing fractions were pooled and loaded onto a Superdex 200 HiLoad16/600 column (GE Healthcare) for size exclusion chromatography. pET41-GST-PreScn-Importin α was expressed in *E. coli* BL21 (DE3), cells were lysed using sonication and enzymatic lysis with lysozyme, and the soluble fraction obtained with centrifugation and filtration. The protein was purified using the GST tag for affinity chromatography with a glutathione sepharose GSTrap 4B column (GE Healthcare). Following glutathione elution, the GST-Importin α containing fractions were pooled and loaded onto a Superdex 200 HiLoad16/600 column (GE Healthcare) for size exclusion chromatography. GST-Importin α-containing fractions were pooled, frozen in liquid nitrogen and stored at -20°C.

Pull-down assay

All solutions were prepared with Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) containing 0.1% (v/v) Triton X-100 (Carl Roth) and 1 mM DTT (Carl Roth) (PBST). All incubation steps were carried out at 4°C to preserve the proteins and all centrifugation steps were carried out at 500 xG. Samples taken for later analysis were mixed with 5x sample buffer and heated to 95°C for 5 min. 50 µM Glutathione Sepharose 4B (Merck) were transferred to a Spin Column (IBA Lifescience), equilibrated with 500 µL PBST followed by centrifugation. As already described,[7] 500 µL 2.5 µM GST-Importin α were added to the column, a sample from the "input" fraction was retained, and the column was incubated for 2 h on a rotator. Samples from each "input" fraction enable to control equal loading of the column and are thus prerequisite to later allow comparative quantification of protein amounts. Unbound protein was removed by three washing steps with PBST followed by centrifugation. 500 µL 2.2 µM inactive Taspase1$_{D233/T234A}$-His were pre-incubated with the respective concentration of compound on a rotator for 1 h, and again, a sample for the "input" fraction was retained. The free binding sites on the column were blocked with 1% (w/v) BSA (Carl Roth) in PBST for 30 min on a rotator. The blocking solution was removed from the column by centrifugation for 1 min. Subsequently, Taspase1-His pre-incubated with the compound was added to the column and allowed to bind on a rotator for 1 h. Analogously, a sample from the "unbound" fraction was retained, and unbound protein was removed by three washing steps with PBST followed by centrifugation for 1 min. Samples from the "unbound" fraction allow to compare protein amounts with the "input" fraction to avoid unwanted deviation and thus serve as additional assay control. Generally spoken, the "input", "bound"/"eluted" and "unbound" fractions always have to add up to the same total amount of protein initially used for the assay. Finally, 500 µL 1x sample buffer were added to the column and heated to 95°C for 10 min. Proteins were eluted by centrifugation for 2 min.

SDS-PAGE and Immunoblotting

Here, we used the standard protocols for SDS-PAGE according to Laemmli,[8] and for Immunoblotting according to Towbin.[9] Briefly, for SDS-PAGE, Tris-glycine gels with 7.5% or 10% (v/v) acrylamide in the stacking gel and 4% (v/v) acrylamide in the separating gel were prepared accordingly. For electrophoresis, we used the TetraCell system (BioRad) set to 200 V for 45 min. Proteins were then transferred to a protein-binding membrane using a wet blot tank (Peqlab) set to 360 mA for 90 min at 4°C. To detect the different proteins, the membrane was first reversibly stained with Ponceau S (AppliChem) and then cut between the protein bands according to the Spectra Multicolor Broad Range Protein Ladder (Thermo Fisher). Free binding sites were blocked with 5% (w/v) powdered milk (Carl Roth) or 5% (w/v) Albumin Fraction V (Carl Roth) in Tris buffered saline with Tween-20 (TBST) (Carl Roth) for at least 30 min at room temperature. After that, membranes were incubated with the respective primary antibodies mouse anti-GFP (B-2) 1:1000 (sc-9996, Santa Cruz Biotechnology), mouse anti-α-Tubulin 1:8000 (T6074, Sigma-Aldrich), rabbit anti-Taspase1 1:2000 (sc-85945, Santa Cruz) or mouse anti-Karyopherinα2 1:1000 (sc-55538, Santa Cruz) in 5% (w/v) powdered milk in TBST or mouse anti-Penta-His 1:2000 (34660, Qiagen) in 5% (w/v) Albumin Fraction V in TBST for at least 1 h at room temperature. Unbound antibodies were removed by three washing steps with TBST. Membranes were incubated with the respective secondary antibodies donkey anti-rabbit HPR-coupled 1:10000 (NA934, GE Healthcare) or sheep anti-mouse HPR-coupled 1:10000 (NXA931, GE Healthcare) in 5% (w/v) powdered milk with TBST for 1 h. Unbound antibodies were removed by four washing steps in TBST. For the detection of chemiluminescence, we used the Pierce ECL Plus Western Blotting Substrate (Thermo Fisher) and the Chemidoc Imaging System (BioRad).

Immunoblot quantification

The signals were quantified with Fiji.[10] If necessary, the signal of Taspase1 in the eluted fraction was corrected for Taspase1 bound to the column without Importin α. To correct possible loading differences the signal of Taspase1 in the eluted fraction, values were normalized for the signal of Importin α in the eluted fractions. The data was evaluated using Origin2019 (OriginLab).

Toxicity Assay

1 x 10^4 cells were cultured in Corning 96 Well microplates (Sigma-Aldrich) in 100 µl Dulbecco's modified eagle medium (DMEM) (Thermo Fisher Scientific) supplied with 10% (v/v) fetal calf serum (FCS) (Life Technologies GmbH), Antibiotic-Antimycotic (Life Technologies GmbH) and the respective compound concentration. Since the compounds were dissolved in DMSO (Carl Roth), this solvent was included as a reference. Cells were then incubated at 37°C and 5% CO_2 for 24 h. Toxicity was determined *via* a colorimetric MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) assay. MTS, in the presence of phenazine methosulfate (PMS), produces a formazan product with an absorbance maximum at 490 nm directly proportional to the number of living, metabolically active cells. Here, dehydrogenase enzymes reduce anabolic cofactors such as nicotinamide adenine dinucleotide and its phosphate (NAD/NADP) to NADH/NADHP required for formazan formation. Briefly, the compound-containing medium was removed, and cells were washed with PBS once to remove excess compound. 100 µL fresh DMEM with 10% FCS and Antibiotic-Antimycotic were added to each well. Following addition of 20 µL Cell Titer AQueous One (Promega), absorption at 490 nm was recorded with the plate reader Promega Glow Max (Promega) after 30 min of incubation. Results were normalized to the DMSO references, compared to the untreated cells and are the mean of at least three replicates ± standard deviation.

Semi-in vitro Taspase1 substrate cleavage assay

5 x 10^6 eukaryotic 293T cells were cultured in TC dish 100, standard (Sarstedt) with 10 ml DMEM with 10% (w/v)FCS and 1% (w/v) Antibiotic-Antimycotic and transiently transfected with

a plasmid coding for the confirmed Taspase1 substrate USF2 (Upstream stimulatory factor 2) tagged with GFP.[5] 24 h after transfection, cells were detached, washed and resuspended in 500 µl buffer containing 100 mM HEPES (Applichem), 10% Saccharose (Applichem), 10 mM DTT (Applichem), pH 7.9.[10] Cell lysates were prepared by sonicating the cells for at least four times with a Sonopuls mini 20 ultrasonic homogenizer and the ultrasonic probe MS 1.5 (Bandelin) for 20 s and an amplitude of 90%. After centrifugation at 14000 rpm for 20 min at 4°C, the supernatant was transferred into a new reaction tube. For the *semi-in vitro* Taspase1 substrate cleavage assay, the USF2 cell lysate was incubated with the corresponding compound and 0.4 µg/µl (equals 9 µM) recombinant active Taspase1-His at 37°C and 300 rpm. After 4 and 6 h, samples were collected, mixed with 5x sample buffer and heated to 95°C for 5 min for further analysis by SDS-PAGE and Immunoblotting (see above).

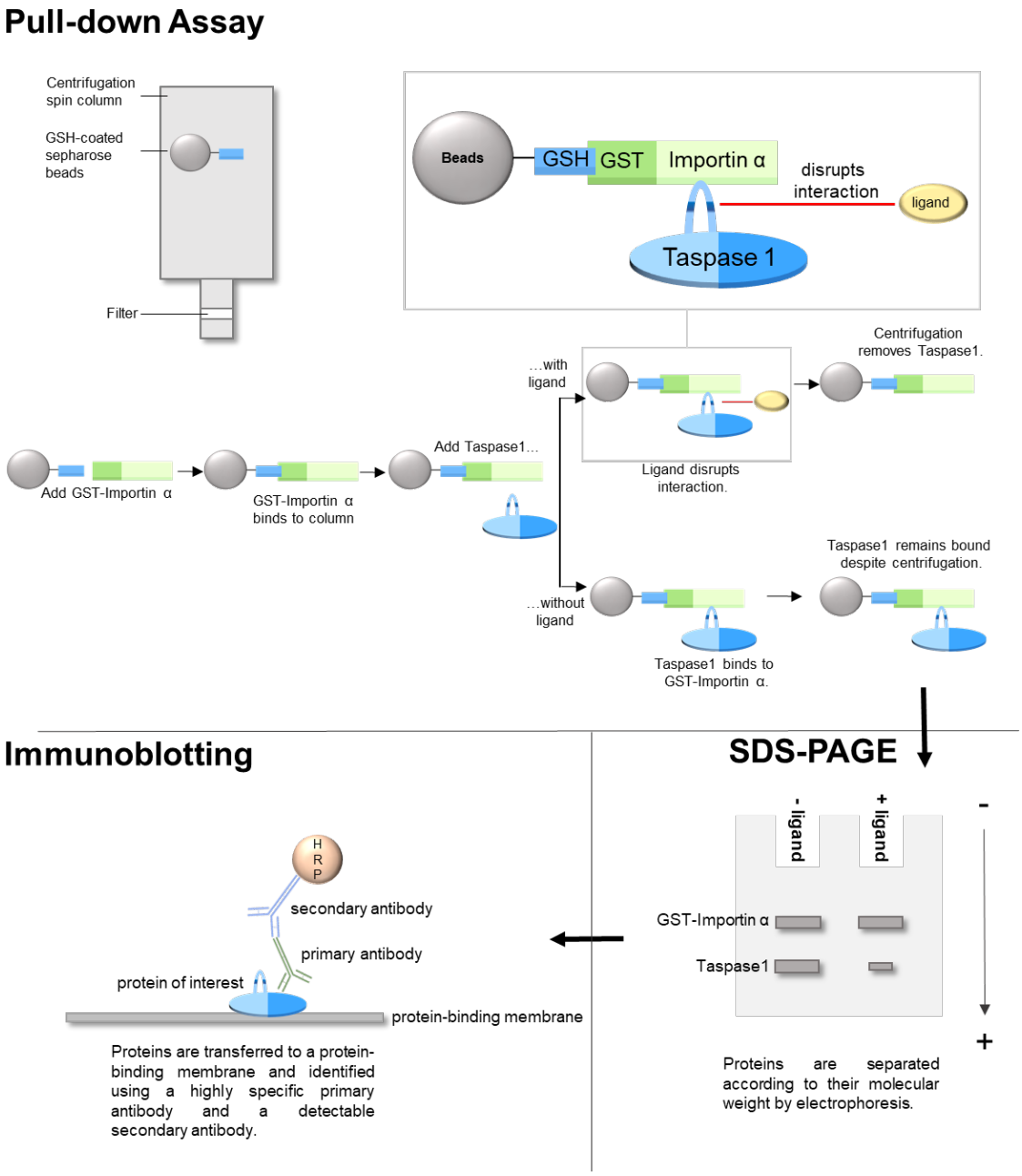


Figure S26: Schematic workflow of the modified pull-down assay.[7] In a spin column, GST-Importin α was fixed on a sepharose matrix coated with glutathione. After GST bound to glutathione with high affinity, unbound protein was removed by centrifugation. Taspase1-His, with or without ligand pre-incubation (as indicated), was added to the column. Unbound protein was again removed by centrifugation. The column was filled with Laemmli buffer that contains ionic detergents as well as reducing agents and heated to 95°C to denature and thus detach all protein from the matrix. SDS-PAGE was applied to separate the proteins according to their molecular weight followed by immunoblot analysis to identify and quantify the respective proteins.

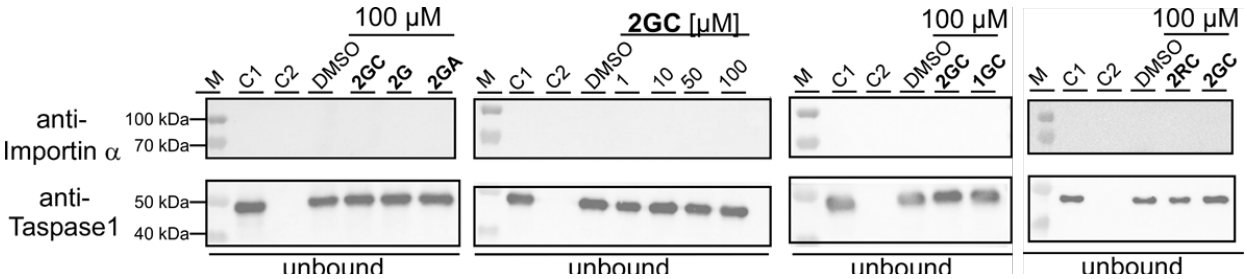
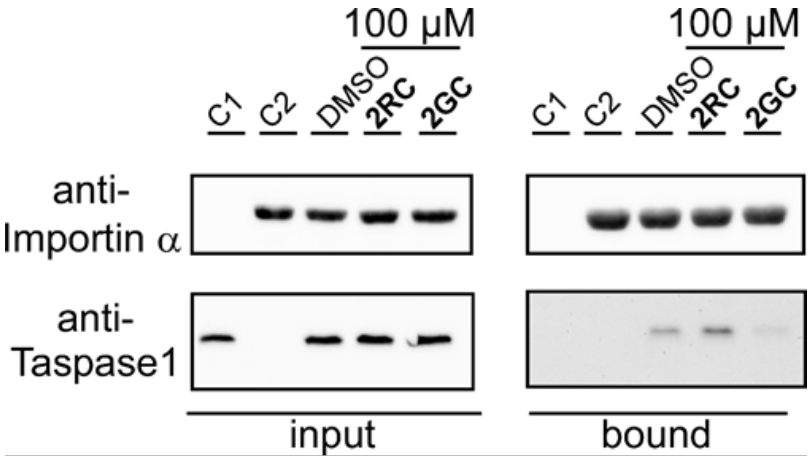


Figure S27: Immunoblot of the "unbound" fraction after incubation with Taspase1 and ligand. Binding of Importin α to the column was demonstrably not affected by the ligands during the assay. Chemiluminescence images were merged with colorimetric images to allow visualization of the marker (M). Controls included only Taspase1 (C1), GST-Importin α (C2) or only the inhibitor solvent (DMSO).



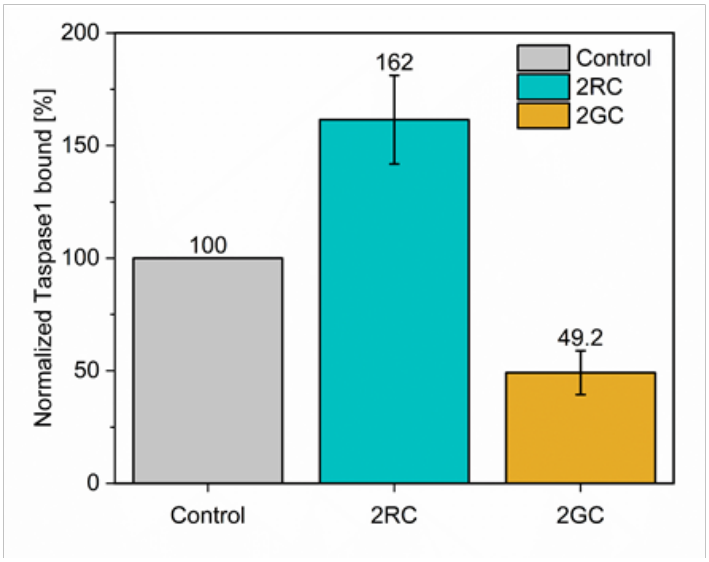


Figure S28: Only the bivalent GCP-containing but not the arginine-based compound allows to efficiently interfere with the Taspase1/Importin α interaction. A) In our pull-down setup, pre-incubation Taspase1 with **2GC** hampers binding to column-bound Importin α in contrast to **2RC**. Controls include either only Taspase1 (C1), GST-Importin α (C2) or DMSO treatment (DMSO). Quantification of results comprises the mean of three replicates ± standard deviation.

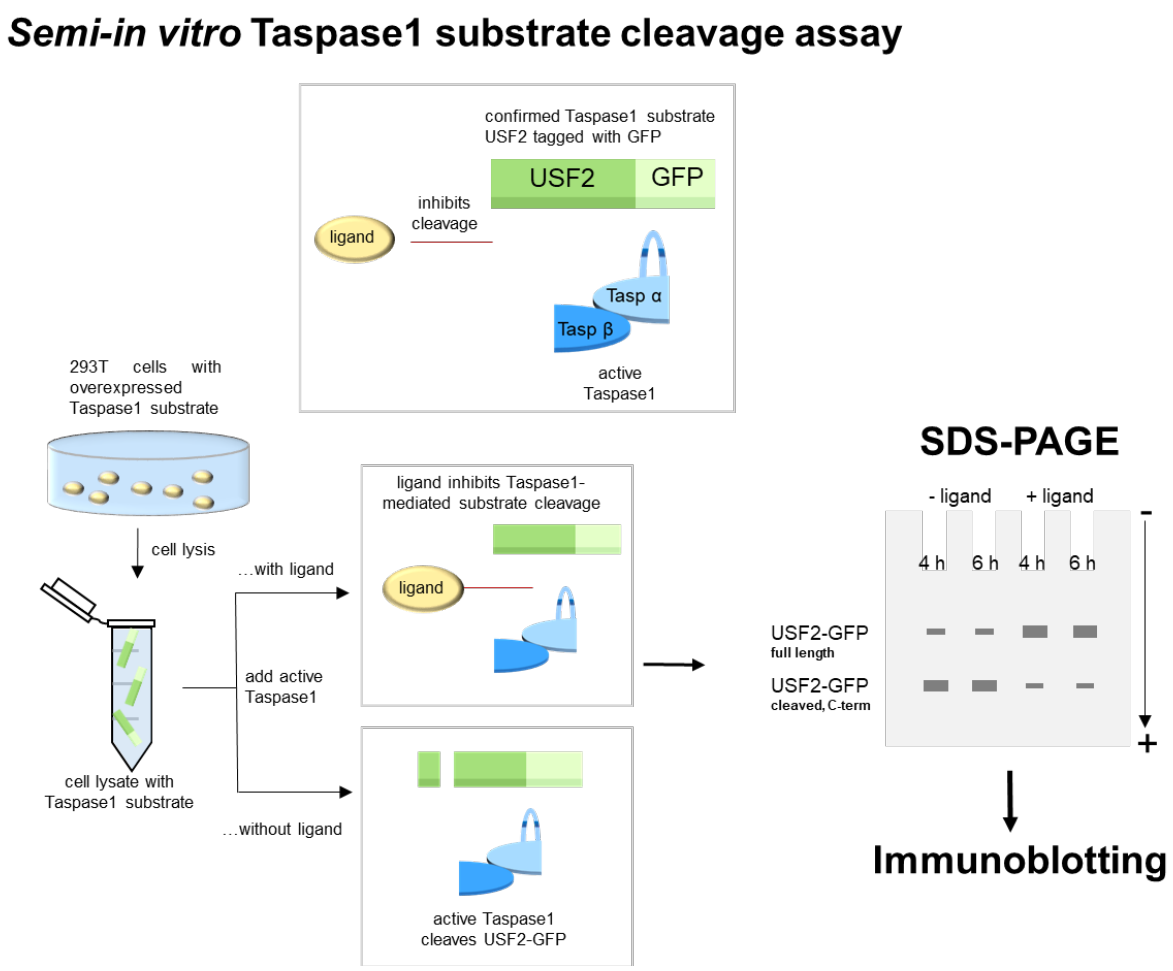


Figure S29: Schematic workflow of the *semi-in vitro* Taspase1 substrate cleavage assay. After transient transfection of eukaryotic 293T cells with a plasmid coding for the Taspase1 substrate USF2 (Upstream stimulatory factor 2) tagged with GFP, cell lysates were prepared by sonicating the cells in 500 µl buffer containing 100 mM HEPES (Applichem), 10% (w/v) Saccharose (Applichem), 10 mM DTT (Applichem), pH 7.9.[5, 11] The USF2 cell lysate was incubated with and without the corresponding compound and 0.4 µg/µl recombinant active Taspase1-His at 37°C and continuous rotation at 300 rpm. Taspase1 cleavage activity was assessed by analyzing substrate cleavage in lysate samples collected after 4 h and 6 h *via* SDS-PAGE and immunoblotting.

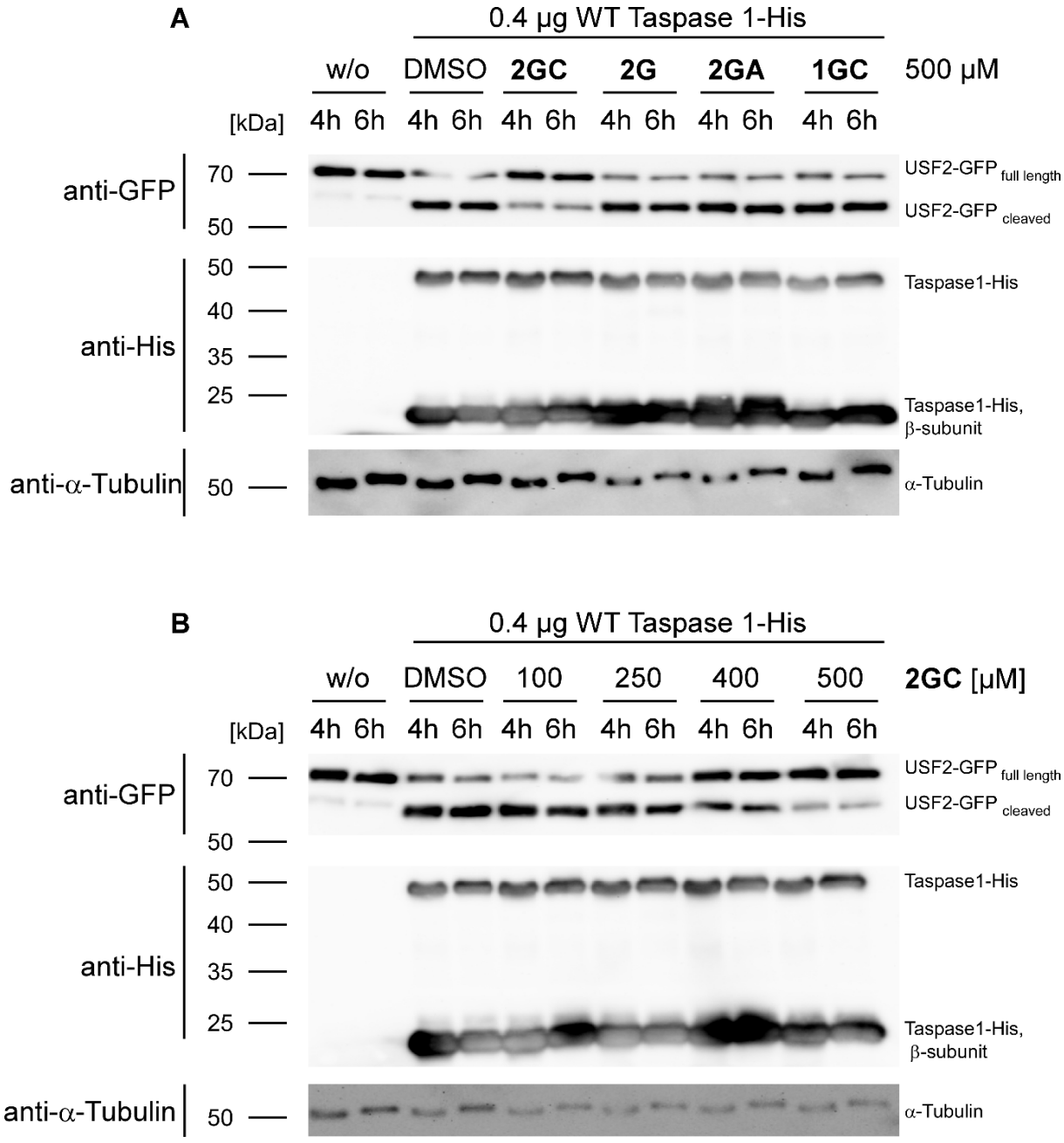


Figure S30: Immunoblot of the *semi-in vitro* Taspase1 substrate cleavage assay indicated an inhibitory effect of **2GC** on Taspase1-mediated USF2 cleavage in contrast to **2G**, **2GA** and **1GC**. 293T cells were transiently transfected with USF2-GFP and 24 h after transfection, whole cell lysates were prepared. The cell lysate was incubated with 0.4 µg recombinant Tapase1-His and 500 µM of the different compounds **2GC**, **2G**, **2GA** and **1GC** (A) or different concentrations of **2GC** (B) at 37°C and continuous rotation at 300 rpm. Controls included cell lysates without active Taspase1 and ligand (w/o) as well cell lysates with active Taspase1 and inhibitor solvent (DMSO). Samples were taken after 4 h and 6 h incubation time and analysed *via* SDS-PAGE and immunoblotting.

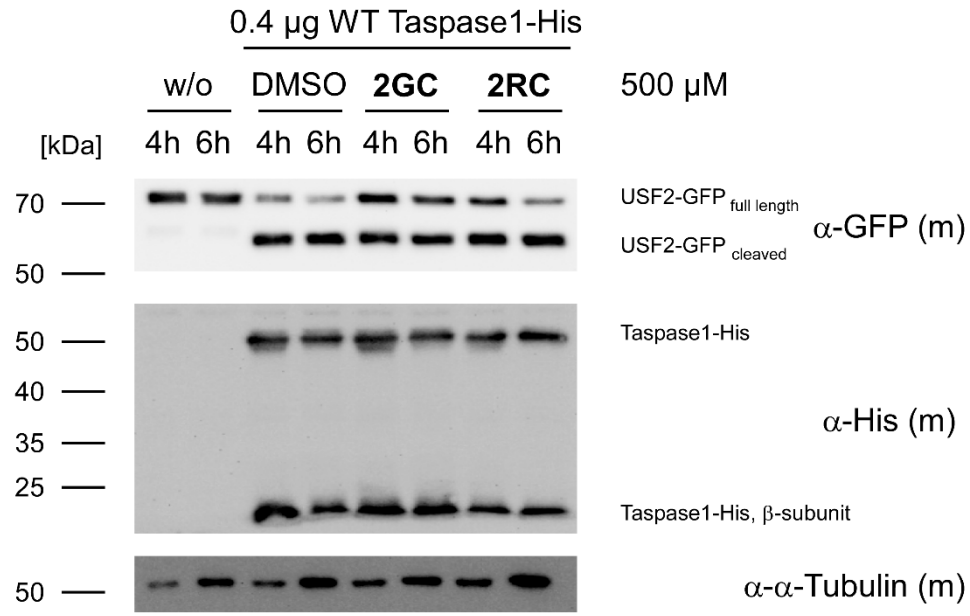


Figure S31: Immunoblot of the *semi-in vitro* Taspase1 substrate cleavage assay indicated an inhibitory effect of **2GC** on Taspase1-mediated USF2 cleavage in contrast to **2RC**. 293T cells were transiently transfected with USF2-GFP and 24 h after transfection, whole cell lysates were prepared. The cell lysate was incubated with 0.4 µg recombinant Tapase1-His and 500 µM of the compound **2GC** or the arginine-containing control **2RC** at 37°C and continuous rotation at 300 rpm. Further controls included cell lysates without active Taspase1 and ligand (w/o) as well cell lysates with active Taspase1 and inhibitor solvent (DMSO). Samples were taken after 4 h and 6 h incubation time and analysed *via* SDS-PAGE and immunoblotting.

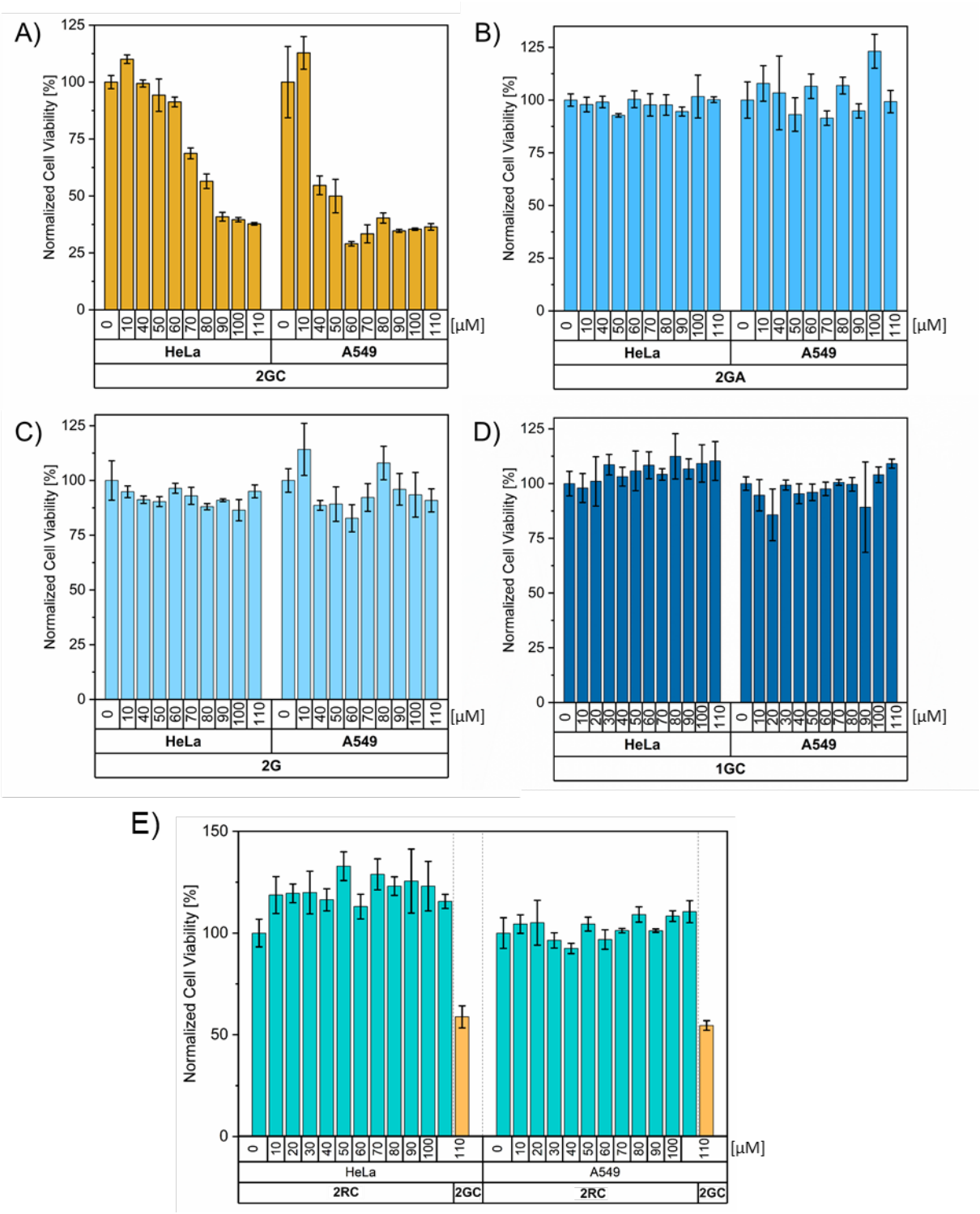


Figure S32: 2GC (orange) affects the cell viability of HeLa and A549 tumor cells in a concentration-dependent manner (A). In contrast, neither the other bivalent compounds **2GA** (blue, B) and **2G** (light blue, C), nor the monovalent building block **1GC** (dark blue, D) decreased tumor cell viability, even when applied at high concentrations of up to 100 µM. This was also evident for the arginine-containing bivalent control compound **2RC** (turquoise, E). Of note, we here again included the highest soluble concentration of **2GC** (110 µM) to allow a direct comparison between the two compounds (orange).

References

[1] C. Schmuck, V. Bickert, M. Merschky, L. Geiger, D. Rupprecht, J. Dudaczek, P. Wich, T. Rehm, U. Machon, *Eur. J. Org. Chem.* **2008**, *2008*, 324-329.

[2] J. van den Boom, F. Trusch, L. Hoppstock, C. Beuck, P. Bayer, *PLoS One* **2016**, *11*, e0151431, 1-13.

[3] J. van den Boom, A. Hensel, F. Trusch, A. Matena, S. Siemer, D. Guel, D. Docter, A. Hoing, P. Bayer, R. H. Stauber, S. K. Knauer, *Nanoscale* **2020**, *12*, 19093-19103.

[4] S. K. Knauer, V. Fetz, J. Rabenstein, S. Friedl, B. Hofmann, S. Sabiani, E. Schroder, L. Kunst, E. Proschak, E. Thines, T. Kindler, G. Schneider, R. Marschalek, R. H. Stauber, C. Bier, *PLoS One* **2011**, *6*, e18253, 1-14.

[5] C. Bier, S. K. Knauer, A. Klapthor, A. Schweitzer, A. Rekik, O. H. Kramer, R. Marschalek, R. H. Stauber, *The Journal of biological chemistry* **2011**, *286*, 3007-3017.

[6] C. Bier, S. K. Knauer, D. Docter, G. Schneider, O. H. Kramer, R. H. Stauber, *Traffic* **2011**, *12*, 703-714.

[7] P. Pasch, A. Höing, S. Ueclue, M. Killa, J. Voskuhl, S. K. Knauer, L. Hartmann, *ChemComm.* **2021**.

[8] U. K. Laemmli, *Nature* **1970**, *227*, 680-685.

[9] H. Towbin, T. Staehelin, J. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 4350-4354.

[10] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, *Nat. Methods* **2012**, *9*, 676-682.

[11] J. van den Boom, M. Mamic, D. Baccelliere, S. Zweerink, F. Kaschani, S. Knauer, P. Bayer, M. Kaiser, *Chembiochem* **2014**, *15*, 2233-2237.

6.3 Publication III: Supplementary information

Supporting information

Dual activity inhibition of threonine aspartase 1 by a single supramolecular bisphosphate ligand

Alexander Höing[a], Robin Struth[b], Christine Beuck[c], Neda Rafieiolhosseini[d], Daniel Hoffmann[d], Roland H. Stauber[e], Peter Bayer[c], Jochen Niemeyer*[b], Shirley K. Knauer*[a]

[a]Molecular Biology II, Center of Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

[b]Organic Chemistry, Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Universitätsstrasse 7, 45141 Essen, Germany

[c]Structural and Medicinal Biochemistry, Center of Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

[d]Bioinformatics and Computational Biophysics, Center of Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany

[e]Molecular and Cellular Oncology/ENT, University Medical Center Mainz, Langenbeckstr.1, 55101 Mainz, Germany

Content

Content 2
List of Figures and Tables 3
Taspase 1 activation 4
Chemical Assays 5
Compound synthesis and purification 5
Computational studies 6
Modelling 6
Energy grid 7
Bead-Spring model of the ligand 8
Simulated Annealing Monte Carlo (SAMC) simulations 8
Biological Assays 9
Cloning 9
Expression and purification of recombinant proteins 9
Pull-down assay 9
SDS-PAGE and immunoblotting 10
Fluorescence anisotropy titration with full-length Taspase 1 12
Fluorescence titration with the Taspase 1 loop 13
Protein NMR spectroscopy 14
Colorimetric cleavage assay 14
Competitive FRET-based cleavage assay 15
Microscopy 16
Confocal laser microscopy for 3D-images 16
Co-localization studies 16
Intracellular biosensor assay 16
Software 18
References 19

List of Figures and Tables

Figures

Figure S1: Activation process of Taspase 1.

Figure S2: Chemical structures and emission spectra of **11d**/**e**/**f.**

Figure S3: Ligand **11d** with deprotonated phosphate groups.

Figure S4: The most probable binding position of **11d** obtained from SAMC simulations.

Figure S5: Immunoblot of the "unbound" fraction after incubation with Taspase 1 and **11d**.

Figure S6: Densitometric quantification of the pull-down results.

Figure S7: Fluorescence anisotropy to determine the binding affinity of **11d**.

Figure S8: Fluorescence titration of Taspase 1 loop with **11d**.

Figure S9: Protein NMR spectroscopy.

Figure S10: Optimization of the confocal laser scanning microscope setup.

Figure S11: FRET-based cleavage assay with **11d**.

Tables

Table S1: Names, coordinates, partial charges, and Van der Waals radii of **11d** ligand atoms.

Table S2: Plasmids and primer pairs used for NEB assembly.

Table S3: Raw data generated from the evaluation of the biosensor localization.

Taspase 1 activation

Taspase 1 has to undergo a distinct, multistep activation process to execute its pathobiological cleavage activity.

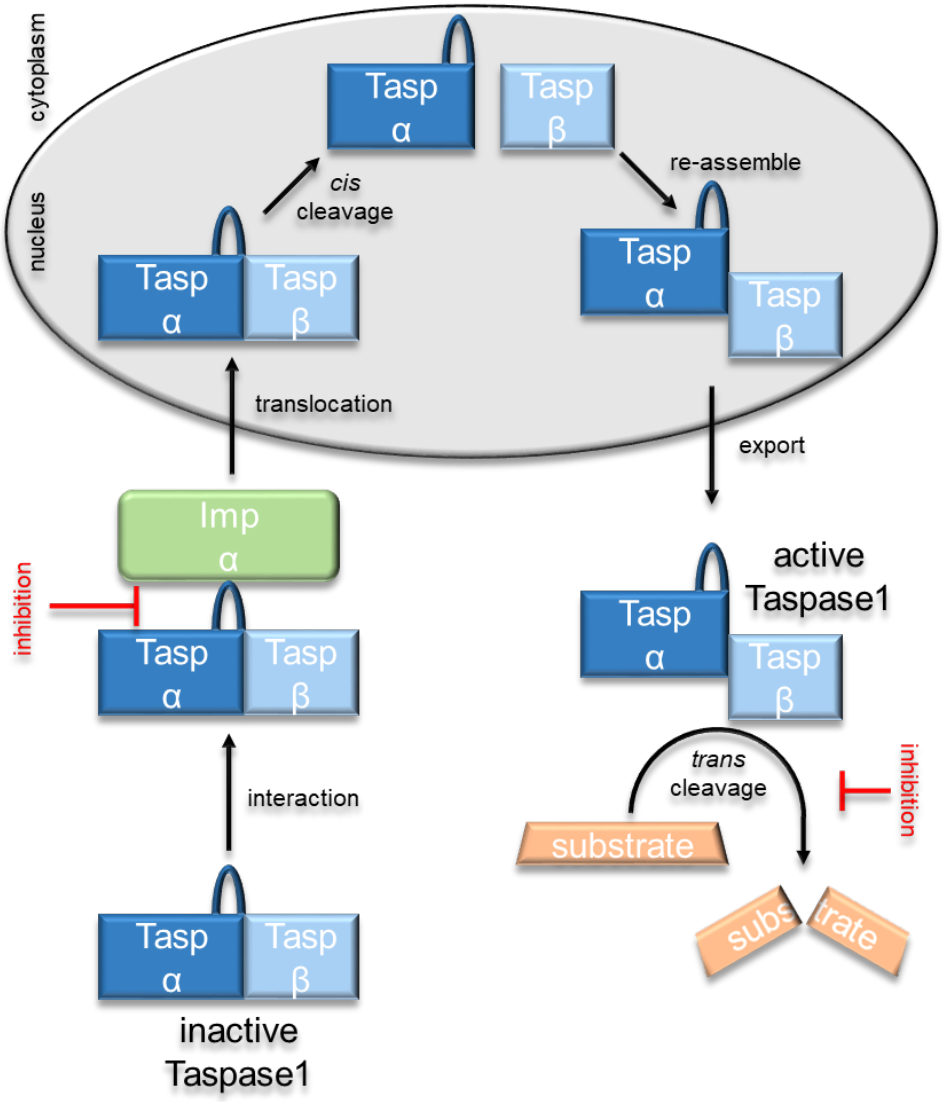


Figure S1: Activation process of Taspase 1. The activation involves the PPI with Importinα and Taspase 1 conversion from an inactive monomer to an active heterodimer through trans cleavage. Inhibition mechanisms are indicated.

Chemical Assays

Compound synthesis and purification

Ligands **11d/e/f** were prepared according to literature procedures.[1-2] For all experiments, ligands **11d/e/f** were used as stock solutions in DMSO (analytical grade), which were prepared by dissolving the solid ligands in an appropriate amount of DMSO to reach a concentration of 100 µM. Aliquots of these stock solutions were used to reach the desired concentrations of **11d/e/f** in further experiments.

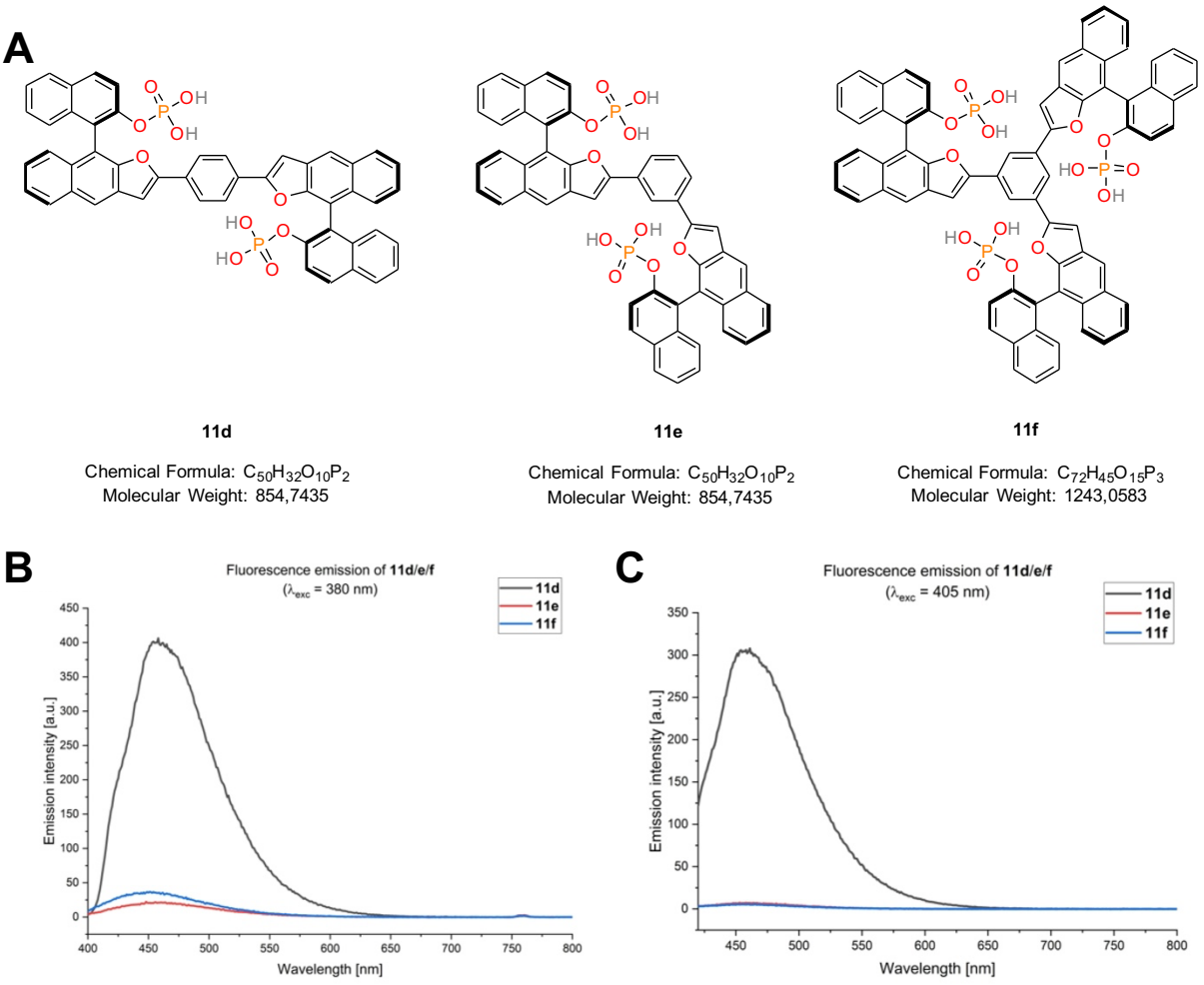


Figure S2: A) Chemical structures of compounds **11d/e/f.** B/C) Emission spectra of **11d/e/f**. **11d** has superior emission intensity at the excitation wavelength used for fluorescence anisotropy (B) and confocal laser scanning microscopy (C) (all: H_2O/DMSO = 95/5, 10 µM).

Computational studies

Modelling

To compute the affinity map of Taspase 1 with **11d**, we proceeded as follows.

Initial structure of **11d** was obtained using Maestro Schrodinger V.3.0.[3] Van der Waals radii and partial charges of ligand atoms (MMFF94_CHARGES) were added automatically by OpenBabel V.3.1.0.[4] Values are provided in Table S1. To specify the protonation state of **11d** which is similar to the protonation state of naphthyl phosphate, there was no experimentally reported value in the literature. Assuming that the protonation state of naphthyl phosphate at pH 7.0 is close to the one of methyl phosphate for which the experimental pK_a values are reported in the literature (pK_{a1} = 1.5, pK_{a2} = 6.3),[5] we used a model of **11d** with doubly deprotonated phosphate groups in our calculations.

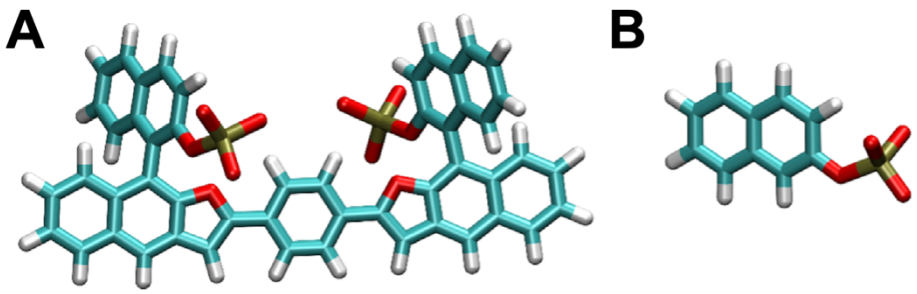


Figure S3: Ligand **11d** with deprotonated phosphate groups (A) and naphthyl phosphate (B).

The model of Taspase 1 that we used in this study was previously generated in the group of Prof. Peter Bayer.[5] This model contained a disordered region comprising aa residues 1 to 39 that was removed for modelling. Charges, van der Waals radii and missing hydrogen atoms were added by PDB2PQR web service v3.1.0 at pH 7.0 with the Amber force field option.[6]

Energy grid

To calculate the electrostatic field of Taspase 1, we solved the nonlinear Poisson-Boltzmann equation with APBS[7] with ionic concentrations of 0.150 mol/L NaCl and relative dielectric permittivities $\epsilon_r^{protein} = 2$ and $\epsilon_r^{water} = 79$. Epitopsy[6] was used to scan the Taspase 1 environment with naphthyl-phosphate using a grid resolution of 0.4 Å, pH 7.0, and a temperature of 298K. Here the reason to use naphthyl phosphate instead of the whole ligand is that Epitopsy assumes the protein and the ligand are rigid structures. Therefore, if the whole ligand with all its degrees of freedom and flexibility is used, the assumption will be far from reality.

Note that in the above-mentioned approach only the interactions of individual naphthyl phosphate with Taspase 1 are obtained. To model the interactions of Taspase 1 with the full **11d** ligand, we combined this approach with further interactions as described.[7]

Table S1: Names, coordinates, partial charges, and Van der Waals radii of **11d** ligand atoms.

	X	Y	Z	Charge	Radius
O	4.513	6.518	-2.081	-1.0333	1.52
O	2.676	4.543	-1.084	-1.0333	1.52
C	5.782	5.463	2.312	-0.15	1.7
C	5.309	5.248	1.014	-0.15	1.7
C	5.906	4.275	0.187	0.0825	1.7
C	6.981	3.525	0.681	-0.15	1.7
C	7.463	3.733	1.982	0	1.7
C	6.858	4.712	2.806	0	1.7
C	8.54	2.981	2.477	-0.15	1.7
C	9.011	3.198	3.775	-0.15	1.7
C	8.413	4.165	4.587	-0.15	1.7
C	7.34	4.921	4.107	-0.15	1.7
O	5.486	4.027	-1.078	-0.3537	1.52
P	4.191	4.824	-1.885	1.3712	1.8
O	4.089	4.202	-3.303	-1.0333	1.52
H	5.307	6.216	2.929	0.15	1.1
H	4.478	5.846	0.669	0.15	1.1
H	9.017	2.227	1.861	0.15	1.1
H	9.842	2.615	4.152	0.15	1.1
H	8.782	4.329	5.591	0.15	1.1
H	6.888	5.666	4.751	0.15	1.1
H	7.44	2.776	0.045	0.15	1.1

Bead-Spring model of the ligand

We have previously developed a coarse-grained bead-spring[8-10] model of a luminescent ligand.[7] We applied the same approach to ligand **11d** such that each of the five chemical groups of the ligand is represented by one bead (see Figure 5B). Each bead is located at the geometric center of the corresponding chemical group. Neighboring beads interact with each other through a harmonic potential. Given the symmetry of ligand **11d**, we consider two sets of spring parameters. We assigned an equilibrium length of 4.5 Å and a spring constant of 25 $k_BT/Å^2$ to the spring connecting the first bead to the second one. The spring connecting the second bead to the third one, has the same spring constant but a different equilibrium length of 7 Å. These parameters are selected such that they keep the beads at reasonable distances. Overlaps between non-bonded beads are avoided by repelling potentials. For more details about the model we refer the reader to our published work.[7] For simplicity, we do not include the angles in our model. Instead, we opt to filter out the final structures obtained from our simulations and only keep those with reasonable angles.

Simulated Annealing Monte Carlo (SAMC) simulations

We did 4000 runs of SAMC to identify the globally optimal configuration of **11d** around Taspase 1 as previously described.[9] At the beginning of each run, we put the central bead at a random grid position within a volume layer around the protein. The thickness of the layer corresponded to the distance between the central bead and the last bead in the chain. Its volume is about 173 nm^3, so that the 4000 initial positions sample the layer at a density of about 23 per nm^3.

Results obtained from our SAMC runs (Figure S4A) suggest that the most likely binding site is in the region around the NLS loop of protein Taspase 1. This binding site is in the positively charged region of the protein surface (Figure S4B).

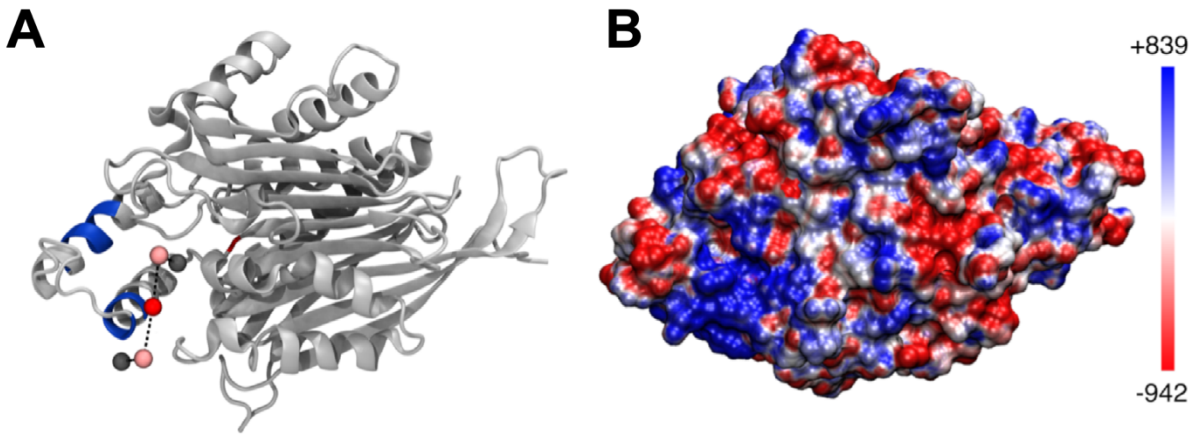


Figure S4: A) The most probable binding position of **11d** obtained from SAMC simulations B) The electrostatic potential on the surface of Taspase 1. The coloring method is based on the surface potential from dark red (most negative) to dark blue (most positive). The values in the color bar are in units of $k_BT/|e|$.

Biological Assays

Cloning

The plasmids for inactive mutant Taspase 1D233/T234A,[11-12] wild type Taspase 1,[13] the Taspase 1 loop$_{178-233}$[14] and GST-PreScn-Importin α[15] were generated as previously described. The plasmid for cleavage probe GST-MLL$_{2700-2850}$-GFP-His was generated by DNA assembly. The backbone was generated by XhoI/NdeI digestion of pET22b-Taspase 1-His.[13] The C-terminal His-tag was kept intact. The fragments for assembly were amplified by PCR with overlaps introduced by the primer (Table S2). 15 fmol backbone was mixed with at least five-fold excess of each insert, mixed with the using the NEBuilder® HiFi DNA Assembly Master Mix (New England Biolabs) and incubated at 50 °C for 30 min. The resulting plasmid was directly transfected into competent *E. coli* NEB-10ß (New England Biolabs). The sequence of the plasmid created was validated by sequencing (LGC Genomics).

Table S2: Plasmids and primer pairs used for NEB assembly.

Fragment	Source	Forward primer	Reverse primer
GST	pC3-GST-USF2-GFP[13]	ctttaagaaggagatatacaATGTCCCCTAT ACTAGGTTATTG	CAGATCCGATTTTGGAGG
MLL2700-2850	JH1117 WT-MLL[16]	atcctccaaaatcggatctgTCTTCAGGTGG AGAGGAAC	agctcctcgcccttgctagcGTCATCACTGTT GTTATTGTC
GFP	pC3-GST-USF2-GFP[13]	GCTAGCAAGGGCGAGGAG	agtggtggtggtggtggtgc*tcgag*CTTGTACAG CTCGTCCATGC

Expression and purification of recombinant proteins

pET22b-Taspase 1$_{D233A/T234}$-His, pET22b-WT_Taspase 1-His and pET41-GST-PreScn-Importinα were expressed and purified as previously described.[15] Expression of the ^{15}N-labelled Taspase 1 loop$_{178-233}$ for NMR titrations was expressed and purified as described previously. pET22b-GST-MLL$_{2700-2850}$-GFP-His was expressed in *E. coli* BL21 (DE3). The cells were lysed using sonication and enzymatic lysis with lysozyme. Cell debris and insoluble fragments were removed using centrifugation and filtration. The protein was purified using a tandem affinity approach utilizing the N-terminal GST-tag and the C-terminal His-Tag. Because of the GFP, the protein could easily be tracked during purification. The soluble fraction was loaded onto a glutathione sepharose GSTrap 4B column (Cytiva). Following glutathione elution, the GST-containing fractions were pooled and loaded onto a HisTrap column (Cytiva). Following imidazole elution, the His-containing fractions were pooled, and the buffer was exchanged for Taspase 1 kinetic buffer (10 % sucrose, 50 µM NaH2PO4, pH 7.4) using a Vivaspin concentrator with a 30 kDa molecular weight cut off. The protein was frozen in liquid nitrogen and stored at -20 °C. Purity was verified using SDS-PAGE and immunoblotting.

Pull-down assay

For the pull-down assay, we used Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) containing 0.1 % (v/v) Triton X-100 (Carl Roth) and 1 mM DTT (Carl Roth) (PBST). Incubation steps were carried out at 4 °C to avoid protein degradation. Centrifugation steps were carried out at 500 xG if not stated otherwise. Samples taken for later analysis were mixed with 5x sample buffer[17] and heated to 95 °C for 5 min. 50 µM slurry of Glutathione Sepharose 4B (Merck) was transferred to a Spin Column (IBA Lifescience) and equilibrated with 500 µL PBST followed by centrifugation. 500 µL 2.5 µM GST-Importinα were added to the column, a sample from the "input" fraction was taken and the column was incubated for 2 h on a rotator. The samples from the "input" fraction were collected to validate that equal amounts of protein and added during the pull-down assay. Unbound protein was removed by three washing steps with PBST followed by centrifugation. 500 µL 2.2 µM inactive Taspase 1$_{D233A/T234A}$-His were pre-incubated with the respective concentration of **11d**/**e**/**f** on a rotator for 1 h. The final DMSO concentration of all samples was adjusted to the highest DMSO concentration used to correct effects on the interaction that were caused by the solvent. Again, a sample for the "input" fraction was retained. The free binding sites on the column were blocked with 1 % (w/v) BSA (Carl Roth) in PBST for 30 min on a rotator. The blocking solution was removed from the column by centrifugation. Subsequently, Taspase 1-$_{D233A/T234A}$-His pre-incubated with **11d** or DMSO was added to the column and incubated on a rotator for 1 h to allow binding to GST-Importinα. Unbound protein was removed by centrifugation. A sample of the "unbound" fraction was taken to validate that the compound did not elute GST-Importinα from the column. We applied three washing steps with PBST followed by centrifugation. Finally, 500 µL 1x sample buffer[19] was added to the column and heated to 95 °C for 10 min to denature all protein on the column. Proteins were eluted by centrifugation for 2 min and analyzed using SDS-PAGE and immunoblotting.

SDS-PAGE and immunoblotting

We used the standard protocols for SDS-PAGE according to Laemmli and for immunoblotting according to Towbin.[17-18] Briefly, for SDS-PAGE, Tris-glycine gels with 12.5 % or 10 % (v/v) acrylamide in the stacking gel and 4 % (v/v) acrylamide in the separating gel were prepared accordingly. For subsequent electrophoresis, we used the TetraCell system (BioRad) set to 200 V for 45 min. Proteins were then transferred to a protein-binding nitrocellulose membrane using a wet blot tank (Peqlab) set to 360 mA for 90 min at 4 °C. The membrane was first reversibly stained with Ponceau S (AppliChem) and the membrane cut between the protein bands according to the Spectra Multicolor Broad Range Protein Ladder (Thermo Fisher) to analyze proteins from the same sample. Free binding sites were blocked with 5 % (w/v) powdered milk (Carl Roth) in Tris buffered saline with Tween-20 (Carl Roth) (TBST) for 60 min at room temperature. After that, membranes were incubated with the respective primary antibodies rabbit anti-Taspase 1 1:2000 (sc-85945, Santa Cruz) or mouse anti-Karyopherinα2 1:1000 (sc-55538, Santa Cruz) in 5% (w/v) powdered milk in TBST over night at 4 °C. Unbound antibodies were removed by three washing steps with TBST. Membranes were incubated with the respective secondary antibodies donkey anti-rabbit HPR-coupled 1:10000 (NA934, GE Healthcare) or sheep anti-mouse HPR-coupled 1:10000 (NXA931, GE Healthcare) in 5% (w/v) powdered milk with TBST for 1 h at room temperature. Unbound antibodies were removed by four washing steps in TBST. For the detection of chemiluminescence, we used the Pierce ECL Plus Western Blotting Substrate (Thermo Fisher) and the Chemidoc Imaging System (BioRad). Automatic exposure time was chosen to avoid overexposure and to ensure equal maximum signal intensity in all blots. Densitometric quantification of the signals was performed using Fiji.[19] If necessary, the signal of Taspase 1 in the eluted fraction was corrected for Taspase 1 bound to the column without Importin α. To correct possible loading differences, the signal of Taspase 1 in the eluted fraction was normalized to the signal of Importin α in the eluted fractions. The data was evaluated using Origin2019 (OriginLab). Homoscedasticity of the samples was tested by Levene-test. Depending on the result, significances were determined by t-test or Welch-t-test to correct differences in the variance of the samples.

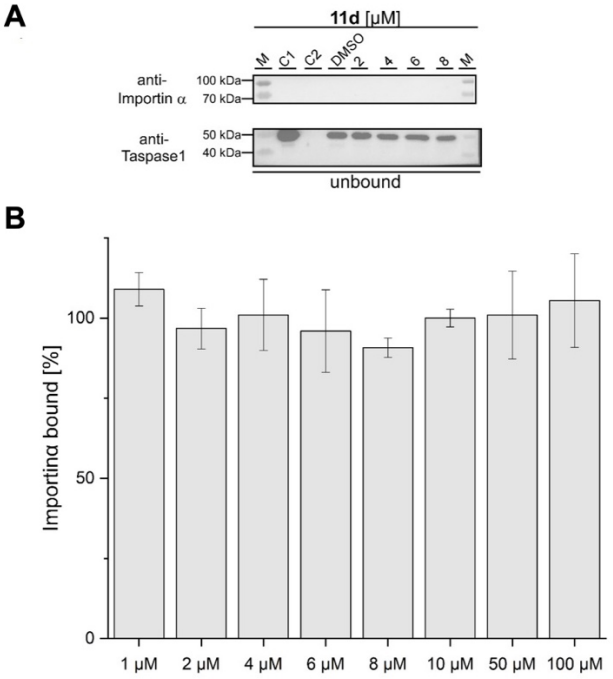


Figure S5: Immunoblot of the "unbound" fraction after incubation with Taspase 1 and **11d**. A) Binding of Importinα to the column was not affected by the ligands during the assay, while unbound Taspase 1 was removed. Chemiluminescence images were merged with colorimetric images to allow visualization of the marker (M). Controls included Taspase 1 (C1), GST-Importin α (C2) alone or an untreated control (DMSO). B) Densitometric quantification of Importinα bound in the "eluted" fraction of the pull-downs shows, that Importinα-binding to the column was not affected by the compound. The results are the mean of three replicates ± standard deviation.

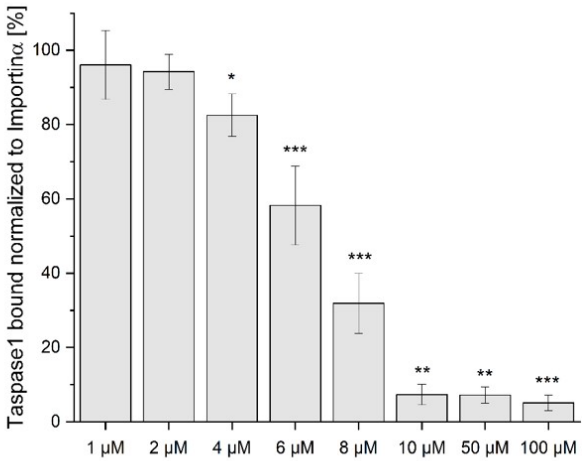


Figure S6: Densitometric quantification of the pull-down results. The results are the mean of three replicates ± standard deviation. *p<0.05; ** p<0.01; ***p<0.001.

Fluorescence anisotropy titration with full-length Taspase 1

Fluorescence anisotropy was performed using a FP-8300 fluorescence spectrometer (Jasco, Pfungstadt, Germany) with high precision cells (Hellma Analytics, Müllheim, Germany). The titration was performed in Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) and all samples were degassed with a MicroCal ThermoVac (Malvern Pananalytical, Kassel, Germany) immediately before the experiment. The concentration of the ligand **11d** was kept constant at 1 µM during the titration with the analyte Taspase $1_{D233A/T234A}$-His. We added increasing volumes (0.5 µL, 0.5 µL, 0.5 µL, 0.5 µL, 1 µL, 2 µL, 4 µL, 8 µL, 16 µL, 32 µL, 64 µL) of a solution consisting of 20 µM protein and 1 µM **11d** to 60 µL 1 µM **11d** at 25 °C. After each titration step, the sample was properly mixed, the change in fluorescence anisotropy measured five times (λ_{exc} 380 nm, λ_{em} 400 nm) and the results averaged. The data was collected with the software Spectra Manager™ with data points representing the mean of three replicates ± standard deviation. Subsequently, the data was fit using the following quadratic binding equation for a one-site specific binding model (Graph Pad Prism 5). The K_D is given as fit ± standard deviation.

$$r = r_0 + \frac{r_{max} * (F + x + K_D) - \sqrt{(F + x + K_D)^2 - 4 * x * F}}{2 * F}$$

r = anisotropy, r_0 = anisotropy in the absence of protein, r_{max} = maximum anisotropy, F = fluorophore concentration, x = protein concentration, K_D = dissociation constant

For the determination of the stoichiometry, we titrated higher concentrations of Taspase 1. The data was fit using the following equation:

$$Y = F_0 + F *$$

A = concentration analyte (**11d**), x = concentration titrant (Taspase 1), F_0 = anisotropy without titrant, F = amplitude, K_D = dissociation constant, n = stoichiometry factor c(analyte)/c(titrant)

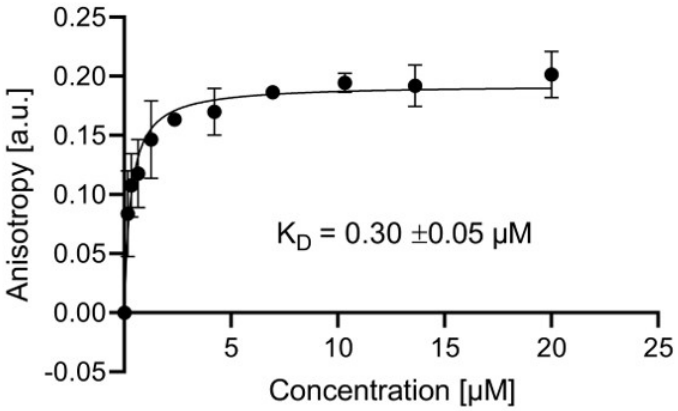


Figure S7: Fluorescence anisotropy to determine the binding affinity of **11d** for Taspase $1_{D233A/T234A}$-His. The results are the mean of three replicates ± standard deviation, K_D is given as fit ± standard error.

Fluorescence titration with the Taspase 1 loop

The Taspase 1 loop with an N-terminal FAM-label was synthesized by GeneCust (Sequence: FAM-SCPPNIMTTRFSLAAF**KRNKRK**LELAERVDTDFMQL**KKRR**QSSEKENDSGTLD, NLS highlighted). In a 96 well-plate Taspase 1 loop was mixed with respective concentrations of **11d**. Untreated controls were treated with respective concentrations of DMSO. While the loop was used at a fixed concentration of 1 µM, we still performed one dilution control per concentration to avoid potential dilution artefacts. Fluorescence was observed using the Promega GloMax (Promega) equipped with the blue filter set (Excitation 490 nm, Emission 510-570 nm). The samples were corrected for potential dilution effects and normalized to untreated controls. To exclude the possibility of the compound binding to the label, we used FAM without peptide as a negative control for the highest compound concentration. The data points are the mean of three replicates ± standard deviation. The data was fit using the following quadratic binding equation for a one-site specific binding model (Graph Pad Prism 5). The K_D is given as fit ± standard deviation.

$$r = r_0 + \frac{r_{max} * (F + x + K_D) - \sqrt{(F + x + K_D)^2 - 4 * x * F}}{2 * F}$$

r = anisotropy, r_0 = anisotropy before titration, r_{max} = maximum anisotropy, F = fluorophore concentration, x = protein concentration, K_D = dissociation constant

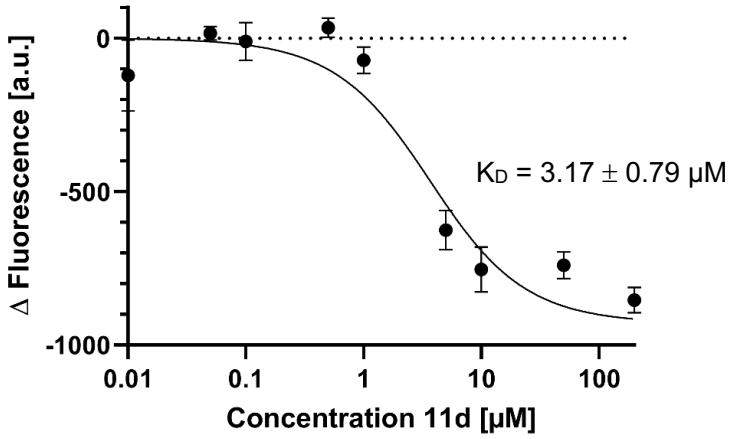


Figure S8: Fluorescence titration of Taspase 1 loop with **11d**. 1 µM FAM-labeled Taspase 1 loop. The results are the mean of triplicates ± standard deviation, the K_D is given as fit ± standard error.

Protein NMR spectroscopy

NMR experiments were recorded at 25 °C on a Bruker 700 MHz Avance Ultrashield NMR spectrometer (Bruker, Germany) equipped with a 5 mm TCI $^{1}H/^{13}C/^{15}N/D$ cryoprobe with z-gradient. Spectra were processed with Topspin 3.5 and analyzed in CARA.[20]

^{15}N-Taspase 1 $loop_{178-233}$ (274 µM) in NMR buffer (1.5 mM KH_2PO_4, 8.9 mM Na_2HPO_4, 136.9 mM NaCl, 2.7 mM KCl, pH 6.5) containing 5 % D_2O was titrated with a 5 mM stock of **11d** in DMSO-d6, yielding a final ligand concentration of 300 µM in the presence of at most 6 % DMSO-d6. $^{1}H,^{15}N$-HSQC spectra were recorded for each titration step. To account for slight shifting of signals due to the presence of DMSO, a control titration with the corresponding volumes of DMSO-d6 without ligand was performed. The relative signal intensities I/I_0 were evaluated,[21] where I represents the signal intensity in the presence of ligand and I_0 the intensity in the DMSO-only reference spectrum. A more than average decrease in intensity also indicates ligand binding due to intermediate exchange kinetics. The amide chemical shift perturbations did not yield useful data because signals already disappeared at small ligand concentrations and thus did not allow tracking of their positions.

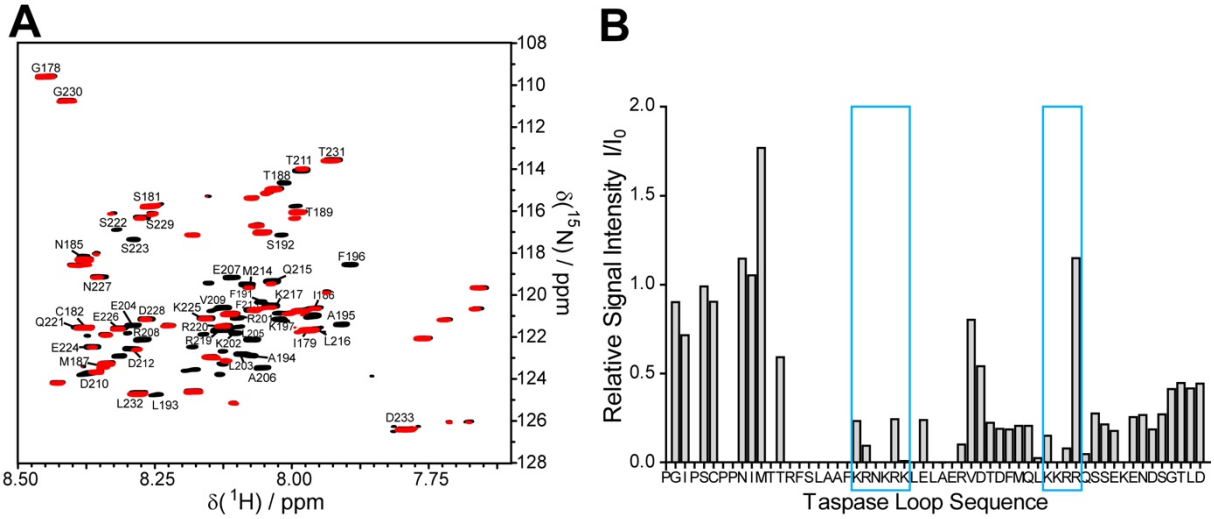


Figure S9: Protein NMR spectroscopy. A) ^{1}H-^{15}N-BEST-TROSY-HSQC spectra of 300 µM ^{15}N-labeled Taspase 1 $loop_{178-233}$ with (red) and without (black) 300 µM **11d**. B) The relative signal intensities show a decrease within Taspase 1's bipartite NLS (blue frame) indicating ligand binding within this region.

Colorimetric cleavage assay

This novel assay utilized a recombinant GST-$MLL_{2700-2850}$-GFP-His with the CS2 cleavage site as a Taspase 1 substrate. If not stated otherwise, all proteins were prepared in Taspase 1 kinetic buffer (10 % sucrose, 50 µM NaH_2PO_4, pH 7.4). Taspase 1 was thawed for 5 min and different aliquots pooled to assure equal protein activity. Taspase 1 and respective concentrations of ligand or DMSO were added into reaction tubes. One negative control contained buffer instead of Taspase 1. In intervals of 15 s the substrate was added to the reaction tubes and the mixture incubated at 37 °C for 90 min. In intervals of 15 s the reaction was stopped by adding 5x sample buffer and heating to 95 °C. The proteins were analyzed using SDS-PAGE and visualized using Coomassie stain solution. The gels were covered with Coomassie staining solution and heated in the microwave at 800 W for 15 s. After 45 min incubation on a rocking shaker the staining solution was removed, and the gel covered in destaining solution. The gel destained overnight on a rocking shaker. For imaging we used a Chemidoc Imaging System (BioRad). Densitometric quantification of the signals was performed using Fiji.[21] The background was subtracted using the rolling ball algorithm with 50 pixels. Of note, the amount of Taspase 1 used was below the detection limit of the later used Coomassie stain, allowing easier quantification of the substrate protein bands. To avoid differences in loading, we only compared the ratio between cleaved and uncleaved substrate per sample and determined the relative cleaved substrate per sample. We normalized the samples to the untreated control to determine the effect of the compound on the proteolytic activity. The data was evaluated using Origin2019 (OriginLab). To correct differences in the variance of the samples, significances were determined using the Welch-t-test.

Competitive FRET-based cleavage assay

This assay is a modified version of an assay described in literature.[22] The peptide used as a substrate was synthesized by Bachem. The N-terminal fluorophore Tide fluor 2 was separated from the C-terminal quencher Tide 2 quencher by the peptide sequence KISQLD/GVDDGC containing a cleavage site for Taspase 1 (cleavage site indicated). HeLa Kyoto cells were lysed using sonication and a standard RIPA buffer. Cell debris was removed by centrifugation at 11000 xG and the supernatant was transferred to a new reaction tube. 8 µM substrate, respective concentrations of lysate and 2 µM **11d** were mixed in a 96 well plate. We included a control without recombinant Taspase 1 to correct for endogenous Taspase 1 in the lysate and possible fluorescence introduced by the compound. Taspase 1-His was thawed for 5 min and added to the mix. Fluorescence was observed using the Promega GloMax (Promega) equipped with the blue filter set (Excitation 490 nm, Emission 510-570 nm). Fluorescence was detected after the intervals given (10x 1 min, 10x 2 min, 10x 10 min, after 180 min total, after 210 min total). The signal in the untreated control was subtracted from the other measurements. The fluorescence in auxiliary units was converted into µM of cleaved substrate using the following equation:

$$[S_C] = \frac{[S_T] * I_{measured} - I_{intact}}{I_{cleaved} - I_{intact}}$$

$[S_C]$ = concentration of cleaved substrate in µM; $[S_T]$ = concentration of total substrate in µM; $I_{measured}$ = observed intensity; I_{intact} = minimum intensity with no substrate cleaved; $I_{cleaved}$ = maximum intensity with all substrate cleaved

The signal in the first 5 min was used to determine the initial velocity with linear regression. The results were normalized to untreated controls to determine the relative activity of Taspase 1 after compound treatment. The data was evaluated using Origin2019 (OriginLab). To correct differences in the variance of the samples, significances were determined using the Welch-t-test.

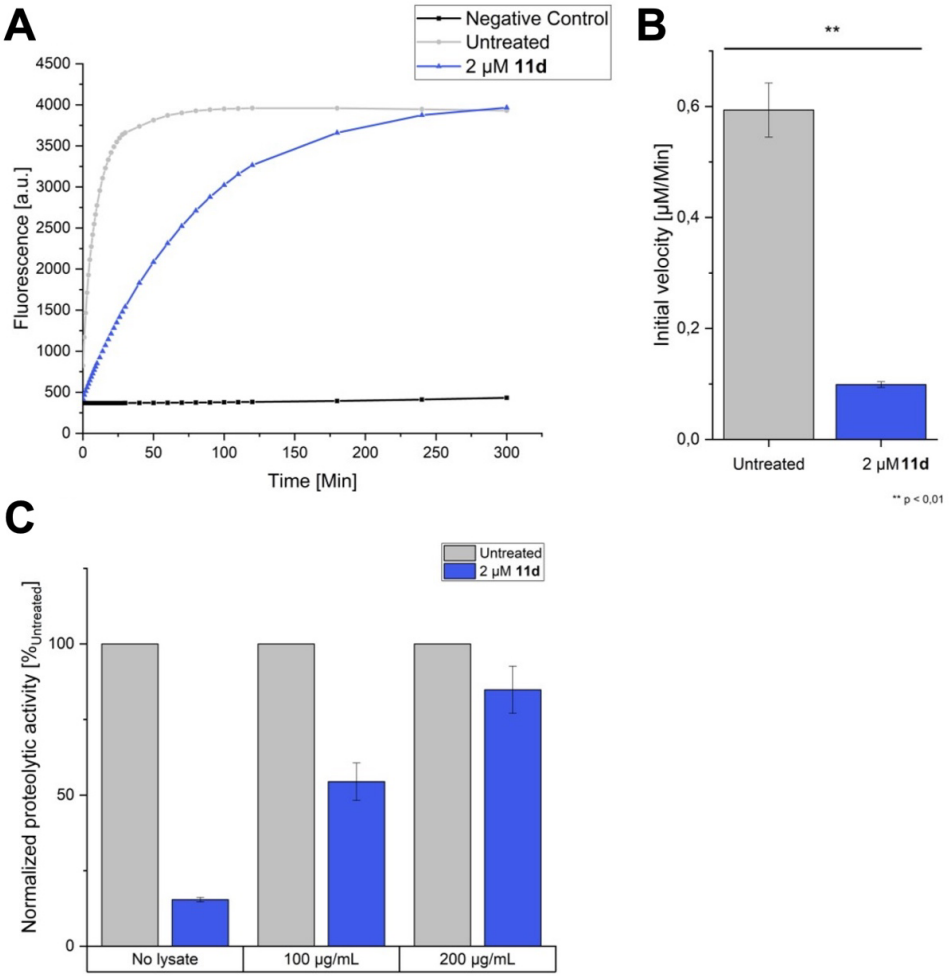


Figure S10: FRET-based cleavage assay with **11d**. A) Exemplary curve showing the effect of 2 µM **11d** on Taspase 1 activity. B) Initial velocity of Taspase 1 enzymatic activity over the first 5 min is significantly decreased by the addition of 2 µM **11d**. The results are the mean of three replicates ± standard deviation. C) A competitive version of the FRET-based cleavage assay shows that **11d** still affects the proteolytic activity of Taspase 1, even if 30- or 60-fold excess of cell lysate is added.

Microscopy

Confocal laser microscopy for 3D-images

2 x 10^4 HeLa Kyoto cells were seeded in 200 µl Dulbecco's modified eagle medium (Thermo Fisher Scientific) in µ-Slide 8 Well (Ibidi) supplemented with 10 % (v/v) FCS (Life Technologies GmbH) and Antibiotic-Antimycotic (Life Technologies GmbH). The cells were incubated in DMEM supplied with 50 µM **11d** or respective concentrations of DMSO for 1 h or 24 h at 37 °C and 5 % CO2. Remaining compound was removed by three washing steps with Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) (PBS). The cells were fixed using paraformaldehyde for 20 min at room temperature followed by a washing step with PBS. To stain the outer plasma membrane, we applied Cellbrite green (Biotium) in PBS for 30 min at 37 °C. After that, the cells were washed with PBS three times. For microscopy, we used the "Leica TCS SP8X Falcon" (Leica). 3D-images and maximum projection images of the cells were generated from image stacks with "LasX" (Leica).

Co-localization studies

2 x 10^4 HeLa Kyoto cells were seeded in 200 µl Dulbecco's modified eagle medium (Thermo Fisher Scientific) in µ-Slide 8 Well (Ibidi) supplemented with 10 % (v/v) FCS (Life Technologies GmbH) and Antibiotic-Antimycotic (Life Technologies GmbH). For co-localization studies, the cells were transfected with pC3-Taspase 1-mCherry 4 h after seeding using Lipofectamine 2000 (Invitrogen). The cells were incubated in DMEM supplied with 50 µM **11d** or respective concentrations of DMSO for 1 h or 24 h at 37 °C and 5 % CO2. Remaining compound was removed by three washing steps with Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) (PBS). The cells were fixed using paraformaldehyde for 20 min at room temperature followed by a washing step with PBS. To stain the outer plasma membrane, we applied Cellbrite green (Biotium) in PBS for 30 min ar 37 °C. After that, the cells were washed with PBS three times. For microscopy, we used the "Leica TCS SP8X Falcon" (Leica). 3D-images and maximum projection images of the cells were generated from image stacks with "LasX" (Leica).

Intracellular biosensor assay

The biosensor for Taspase 1 is an established assay for intracellular Taspase 1 activity.[16] HeLa Kyoto cells were seeded to 60 % confluency in 200 µl Dulbecco's modified eagle medium (Thermo Fisher Scientific) (DMEM) in µ-Slide 8 Well (Ibidi) supplemented with 10 % (v/v) FCS (Life Technologies GmbH) and Antibiotic-Antimycotic (Life Technologies GmbH). After 4 h incubation at 37 °C and 5 % CO_2 the cells were transfected with pC3-TS-Cl2+$_R$ 4 h after seeding using Lipofectamine 2000 (Invitrogen). After 4 h medium was exchanged for DMEM supplemented with 50 µM **11d**. The control was treated with respective concentrations of DMSO. 30 h after transfection and after 26 h of compound treatment, the cells were washed three times with Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) (PBS). After that, the cells were fixed with paraformaldehyde for 20 min at room temperature. After that, the cells were washed with PBS three times. For microscopy, we used the "Leica TCS SP8X Falcon" (Leica). Images of the biosensor were randomized using the Fiji macro "Filename randomizer".[21] Based on the intensity of the biosensor signal, its distribution patterns were divided into five categories ("Only nucleus"; "Predominantly nucleus"; "Equally distributed"; "Predominantly cytoplasm"; "Only cytoplasm"; Figure S10B). The last category was later removed since there were no cells displaying this distribution pattern of the biosensor. The number of cells showing each distribution pattern was normalized to the total number of cells evaluated. For each sample the distribution pattern of at least 100 cells was evaluated. The results are the mean of three replicated ± standard deviation. The data was evaluated using Origin2019 (OriginLab). Homoscedasticity of the samples was tested by Levene-test. Depending on the result, significances were determined by t-test or Welch-t-test to correct differences in the variance of the samples. To determine the inhibitory effect of **11d** in the biosensor assay, the ratio of the mean percentage of cells in the category "Predominantly cytoplasm" (P_C) vs. "Predominately nucleus" (P_N) was calculated and set to 1 for the DSMO control, resulting in an inhibitory factor of 10.8.

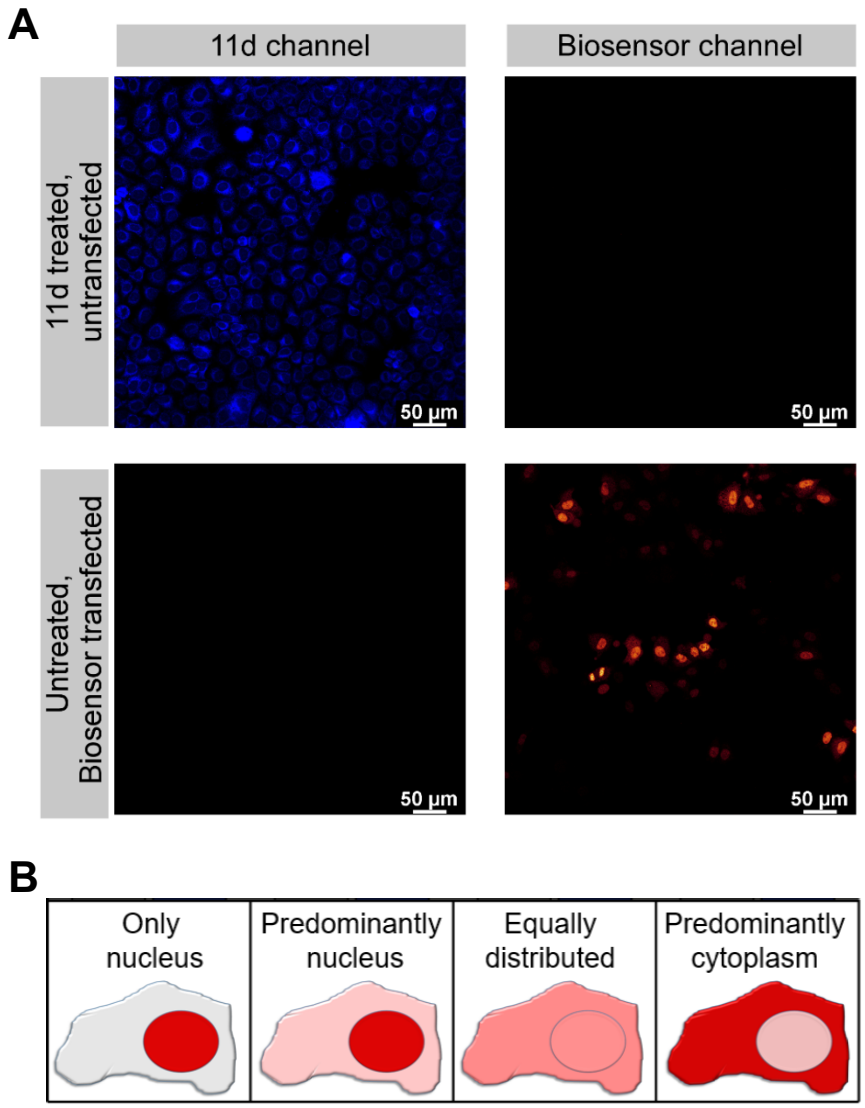


Figure S11: Optimization of the confocal laser scanning microscope setup. A) Neither is the compound **11d** visible in the biosensor channel, nor is the biosensor visible in the **11d** channel in this microscopy setup. B) Assignment of the biosensors' distribution patterns into different categories based on its fluorescence intensity in the different cellular compartments.

Table S3: Raw data generated from the evaluation of the biosensor localization.

Treatment	**Replicate**	**Intracellular location of the biosensor**					**Cells in total**	
		Only nucleus	Predominantly nucleus (P_N)	Equal distribution	Predominantly cytoplasm (P_C)	Only cytoplasm		Ratio P_C/P_N
Control	#1	15	95 (73,1%)	7	13 (10%)	0	130	
Control	#2	16	49 (45%)	26	18 (16,5%)	0	109	
Control	#3	11	50 (41,7%)	35	24 (20%)	0	120	
	mean (%)		53,2%		15,5%			0.29
50µM **11d**	#1	20	35 (29,9%)	16	46 (39,3%)	0	117	
50µM **11d**	#2	6	14 (12,3%)	38	56 (49,1%)	0	114	
50µM **11d**	#3	4	9 (8,4%)	18	76 (71,0%)	0	107	
	mean (%)		16,9%		53,2%			3.15

Software

The atomistic models of the ligands were created in ChemDraw prime 16.0 (PerkinElmer, Waltham, MA, USA). Structures for modelling were obtained using Maestro Schrodinger V.3.0.,[3] OpenBabel V.3.1.0.[4] or PDB2PQR web service v3.1.0 with the Amber force field option.[23] Modelling data was visualized by VMD (version 1.9.4).[24] Images of Western blots and Coomassie gels were acquired by the Chemidoc Imaging System (BioRad), quantified by densitometric analysis with Fiji,[19] and the data was evaluated using Origin2019 (OriginLab). For data acquisition of fluorescence titrations, was the software Spectra Manager™ II (Jasco) or GloMax (Promega) supplied with the instrument were used, and data was plotted with GraphPad Prism 8 (version 8.4.1). Cleavage assay data was evaluated using Origin2019 (OriginLab). NMR data were collected using Topspin 3.5 (Bruker) using the NMRlib 2.0 pulse sequence tools library from IBS (Grenoble, France) available at http://www.ibs.fr/research/scientific-output/software/pulse-sequence-tools/. NMR spectra were processed with Topspin 3.5 (Bruker) and analyzed in CARA (version 1.9.1.7; http://cara.nmr.ch). Relative signal intensities were calculated from the raw chemical shift data and peak intensities using Excel 2016 (Microsoft) and plotted with GraphPad Prism 5.0. Microscopy data was generated by "LasX" (Leica), and images were randomized using the Fiji macro "Filename randomizer".[19] Figure panels were assembled in CanvasDraw 6.0 (ACDsee) or PowerPoint 2016 (Micosoft).

References

(1) Octa-Smolin, F.; van der Vight, F.; Yadav, R.; Bhangu, J.; Soloviova, K.; Wolper, C.; Daniliuc, C. G.; Strassert, C. A.; Somnitz, H.; Jansen, G.; Niemeyer, J. Synthesis of Furan-Annelated BINOL Derivatives: Acid-Catalyzed Cyclization Induces Partial Racemization. *J Org Chem* **2018,** *83* (23), 14568-14587.

(2) Yadav, R.; Kwamen, C.; Niemeyer, J. Development of Fluorescent Chemosensors for Amino-Sugars. *Israel Journal of Chemistry* **2021,** *61*.

(3) Maestro, S., LCC *Maestro*, Schrödinger Release 2021-3; New York, NY, 2021.

(4) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *Journal of cheminformatics* **2011,** *3*, 33.

(5) Kumler, W. D.; Eiler, J. J. The Acid Strength of Mono and Diesters of Phosphoric Acid. The n-Alkyl Esters from Methyl to Butyl, the Esters of Biological Importance, and the Natural Guanidine Phosphoric Acids. *Journal of the American Chemical Society* **1943,** *65* (12), 2355–2361.

(6) Grad, J. N.; Gigante, A.; Wilms, C.; Dybowski, J. N.; Ohl, L.; Ottmann, C.; Schmuck, C.; Hoffmann, D. Locating Large, Flexible Ligands on Proteins. *J Chem Inf Model* **2018,** *58* (2), 315-327.

(7) Rafieiolhosseini, N.; Killa, M.; Tötsch, N.; Grad, J.-N.; Höing, A.; Ottmann, C.; Knauer, S. K.; Voskuhl, J.; Hoffmann, D. Computational model predicts protein binding sites of a luminescent ligand equipped with guanidiniocarbonyl-pyrrole groups. *Beilstein Archives* **2021**.

(8) Milchev, A.; Binder, K. Static and Dynamic Properties of Adsorbed Chains at Surfaces: Monte Carlo Simulation of a Bead-Spring Model. *Macromolecules* **1996,** *29* (1), 343-354.

(9) Underhill, P. T.; Doyle, P. S. On the coarse-graining of polymers into bead-spring chains. *Journal of Non-newtonian Fluid Mechanics* **2004,** *122*, 3-31.

(10) Liu, B.; Wang, J.; Fan, X.; Kong, Y.; Gao, H. An effective bead-spring model for polymer simulation. *J. Comput. Phys.* **2008,** *227* (5), 2794–2807.

(11) Bier, C.; Hecht, R.; Kunst, L.; Scheiding, S.; Wunsch, D.; Goesswein, D.; Schneider, G.; Kramer, O. H.; Knauer, S. K.; Stauber, R. H. Overexpression of the catalytically impaired Taspase1 T234V or Taspase1 D233A variants does not have a dominant negative effect in T(4;11) leukemia cells. *PloS one* **2012,** *7* (5), e34142.

(12) van den Boom, J.; Hensel, A.; Trusch, F.; Matena, A.; Siemer, S.; Guel, D.; Docter, D.; Höing, A.; Bayer, P.; Stauber, R. H.; Knauer, S. K. The other side of the corona: nanoparticles inhibit the protease taspase1 in a size-dependent manner. *Nanoscale* **2020**.

(13) Knauer, S. K.; Fetz, V.; Rabenstein, J.; Friedl, S.; Hofmann, B.; Sabiani, S.; Schröder, E.; Kunst, L.; Proschak, E.; Thines, E.; Kindler, T.; Schneider, G.; Marschalek, R.; Stauber, R. H.; Bier, C. Bioassays to Monitor Taspase1 Function for the Identification of Pharmacogenetic Inhibitors. *PloS one* **2011,** *6* (5).

(14) van den Boom, J.; Trusch, F.; Hoppstock, L.; Beuck, C.; Bayer, P. Structural Characterization of the Loop at the Alpha-Subunit C-Terminus of the Mixed Lineage Leukemia Protein Activating Protease Taspase1. *PloS one* **2016,** *11* (3), e0151431.

(15) Pasch, P.; Höing, A.; Ueclue, S.; Killa, M.; Voskuhl, J.; Knauer, S. K.; Hartmann, L. PEGylated sequence-controlled macromolecules using supramolecular binding to target the Taspase1/Importin α interaction. *Chemical Communications* **2021,** *57* (25), 3091–3094.

(16) Chen, D. Y.; Lee, Y.; van Tine, B. A.; Searleman, A. C.; Westergard, T. D.; Liu, H.; Tu, H.-C.; Takeda, S.; Dong, Y.; Piwnica-Worms, D. R.; Oh, K. J.; Korsmeyer, S. J.; Hermone, A.; Gussio, R.; Shoemaker, R. H.; Cheng, E. H.-Y.; Hsieh, J. J.-D. A Pharmacological Inhibitor of the Protease Taspase1 Effectively Inhibits Breast and Brain Tumor Growth. *Cancer research* **2012,** *72* (3), 736–746.

(17) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970,** *227* (5259), 680-5.

(18) Towbin, H.; Staehelin, T.; Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **1979,** *76* (9), 4350-4.

(19) Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; Tinevez, J.-Y.; White, D. J.; Hartenstein, V.; Eliceiri, K.; Tomancak, P.; Cardona, A. Fiji: an open-source platform for biological-image analysis. *Nature methods* **2012,** *9* (7), 676–682.

(20) Keller, R. L. J. Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment. ETH Zurich, 2005.

(21) Bayer, P.; Matena, A.; Beuck, C. NMR Spectroscopy of supramolecular chemistry on protein surfaces. *Beilstein journal of organic chemistry* **2020,** *16*, 2505–2522.

(22) 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* **2003,** *115* (3), 293–303.

(23) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic acids research* **2004,** *32* (Web Server issue), W665-7.

(24) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **1996,** *14* (1), 33-8, 27-8.

6.4 Publication IV: Supplementary information

Supporting Information

69451 Weinheim, Germany

Recognition of a flexible protein loop in Taspase1 by multivalent supramolecular tweezers

Alexander Höing[a+], Abbna Kirupakaran[b+], Christine Beuck[c], Marius Pörschke[c], Felix Niemeyer[d], Peter Bayer[c], Thomas Schrader[b*], Shirley K. Knauer[a*]

[a] Molecular Biology II, Faculty of Biology, Center of Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany
E-mail: shirley.knauer@uni-due.de
[b] Institute of Organic Chemistry I, Faculty of Chemistry, University of Duisburg-Essen, Essen, Universitätsstrasse 7, 45141 Essen, Germany
Email: thomas.schrader@uni-due.de
[c] Structural and Medicinal Biochemistry, Center of Medical Biotechnology (ZMB), University of Duisburg-Essen, Universitätsstrasse 5, 45141 Essen, Germany
[d] Institute of Organic Chemistry, Faculty of Chemistry, Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Essen, Universitätsstrasse 7, 45141 Essen, Germany

[+] These authors contributed equally.
[*] Corresponding authors.

Supporting information for this article is given via a link at the end of the document.

Abstract: Many protein-protein interactions utilize a multivalent display of epitopes for binding to enhance molecular affinity and specificity. Imitating this natural concept, we here report the sophisticated design of multivalent supramolecular tweezers that allow to address lysine and arginine clusters on a flexible protein surface loop. The unique protease Taspase1 is not only involved in cancer development, but also characterized by a basic bipartite nuclear localization signal (NLS) interacting with Importin α pivotal for proteolytic activation. Newly established synthesis routes enabled us to covalently fuse several tweezer molecules into multivalent ligands. The resulting bi- up to pentavalent constructs were then systematically compared in comprehensive biochemical assays. Indeed, the stepwise increase in valency was robustly reflected by the ligands' gradually enhanced potency to disrupt the interaction of Taspase1 with Importin α, correlated with higher binding affinity and inhibition of proteolytic activity.

DOI: 10.1002/anie.2021XXXXX

SUPPORTING INFORMATION

Table of Contents

Experimental Procedures 3
Synthesis and properties of multivalent tweezers 3
Expression and purification of recombinant proteins 8
Pull-down assay 8
SDS-PAGE and immunoblotting 8
Fluorescence titration with the isolated Taspase1 loop 9
Protein NMR spectroscopy 9
Colorimetric cleavage assay ..... **Fehler! Textmarke nicht definiert.**
Computational studies 10
Software 10
Supporting Figures 11
References 15

Experimental Procedures

Synthesis and properties of multivalent tweezers

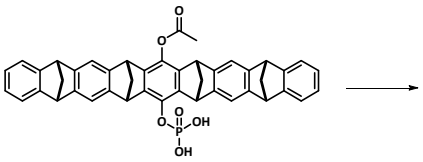
General methods: All commercially available chemicals in this work were purchased from Sigma-Aldrich, Fisher Chemicals, VWR, Fluka, Acros Organics and TCI. Phosphorus oxychloride and Triethyl amine were freshly distilled.

Tweezer compounds were purified by preparative reversed-phase high performance liquid chromatography (RP-HPLC) using a Prominence UFLC system of Shimadzu equipped with the reverse phase column Luna® 5 µm C18 (2), 100 x 21.20 mm from Phenomenex. Separation was achieved by gradient run with acetonitrile/water + 0.1 % TFA at a flow rate of 25 mL per minute with peak detection at 210 nm. Mass spectra were recorded using the Q-TOF mass spectrometer Bruker maXis 4G was by high-resolution electron spray ionization (ESI).

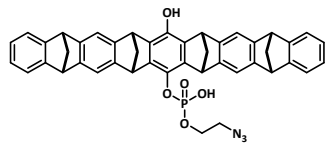
^{1}H, ^{13}C and ^{31}P-NMR spectra were recorded with Bruker AVNeo400 (^{1}H = 400 MHz; ^{13}C = 101 MHz, ^{31}P = 162 MHz) at 25 °C for tweezer compounds synthesized in this work. NMR titrations were carried out on the Bruker DRX 500 (^{1}H = 600 MHz) at 25 °C.

General procedure for the construction of multivalent tweezers: The parent tweezers **MT1**, monoacetoxy-monophosphate tweezer **1**,[1-3] dibut-3-yn-1-yl terephthalate[4] and dibut-3-yn-1-yl 1,3,5-benzenetricarboxylate[4] were synthesized by published methods.

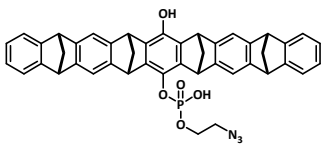
Synthesis of the monoazido diphosphate tweezer 2:



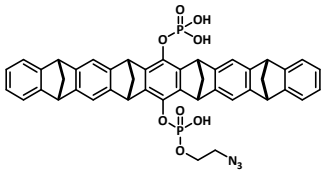
1



1.1



1.2



2

Monoacetoxy monophosphate tweezer 1 (1.50 g, 2.18 mmol, 1.0 eq.) was dissolved in dry pyridine (82 mL) under argon atmosphere. Azidoethanol (5.15 g, 4.49 mL, 59.1 mmol, 27.2 eq.) was added and subsequently trichloroacetonitrile (2.95 g, 2.05 mL, 20.4 mmol, 9.41 eq.). The reaction mixture was heated to 90 °C and stirred for 16 h. After cooling down to room temperature, the solvent was evaporated in vacuo. The solid residue was suspended in HCl (aq., 1 N, 30 mL) and sonicated for 30 min. The suspension was filtered using a D4 frit. After washing the residue with HCl (aq., 1 N, 30 mL), it was washed with H_2O (5 mL). The solid was dissolved in THF (20 mL) and the solvent was removed under vacuum. The residue was passed through a short pad of RP-18 silica (H_2O/MeCN) prior to purification by preparative HPLC (H_2O/MeCN + 0.1 % TFA; gradient). Monoacetoxy monoethylazidophosphate tweezer **1.1** was obtained as a colorless solid.

Monoacetoxy monoethylazidophosphate tweezer 1.1 (42.3 mg, 0.06 mmol, 1 eq.) was dissolved in 1,4-dioxane (4.40 mL) and treated with 0.88 M sodium hydroxide (0.44 mL, 0.42 mmol, 7 eq.). The reaction mixture was stirred for 3 h at room temperature, prior to addition of HCl (30 mL). The suspension was sonicated twice for 15 min, filtered over a D4 frit. After washing with HCl (aq. 1 M, 2 x 5 mL) and H_2O (5 mL), the residue was dissolved in THF (20 mL) and H_2O (5 mL) was added. The solvent was dried under vacuum. The residue was passed through a short pad of RP-18 silica (H_2O/MeCN) prior to purification by preparative HPLC (H_2O/MeCN + 0.1 % TFA; gradient). The monohydroxy monoethylazidophosphate tweezer **1.2** was obtained as a white solid.

Monohydroxy monoethylazidophosphate tweezer 1.2 (0.15 g, 0.21 mmol, 1.0 eq.) was dissolved in dry THF (30 mL) and cooled down to 0 °C. To this mixture phosphorus oxychloride (0.32 g, 0.19 mL, 2.08 mmol, 9.85 eq.) and triethylamine (0.32 g, 0.44 mL, 3.16 mmol, 15 eq.) were added and stirred for 2 h. The precipitation was filtered off and dried under vacuum. Then, HCl (aq., 1 M, 30 mL) was added, and the suspension was sonicated twice for 15 min. The solid residue was filtered over a D4 frit, washed with (aq., 1 M, 2 x 5 mL) and washed with H_2O (5 mL). The solid residue was dissolved in THF (20 mL) and H_2O (5 mL) was added. The solvent was removed under vacuum. The residue was passed through a short pad of RP-18 silica (H_2O/MeCN) prior to purification by preparative HPLC (H_2O/MeCN + 0.1 % TFA; gradient). The monophosphate monoethylazidophosphate tweezer **2** was obtained as a colorless solid.

Monoacetoxy monoethylazidophosphate tweezer 1.1: Yield: 1.62 g, 2.13 mmol, 98 %; ^{1}H NMR (600 MHz, DMSO-d_6): δ = 7.17 (s, 2H), 7.10 (s, 2H), 7.08 (dd, *J* = 5.3, 3.1 Hz, 4H), 6.83 – 6.75 (m, 4H), 4.35 (q, *J* = 1.7 Hz, 2H), 4.09 (dd, *J* = 8.6, 1.6 Hz, 4H), 4.03 (q, *J* = 1.7 Hz, 2H), 3.64 (dt, *J* = 7.0, 4.7 Hz, 2H), 3.01 (t, *J* = 4.8 Hz, 2H), 2.38 (s, 3H), 2.34 – 2.21 (m, 8H) ppm; ^{13}C NMR (151 MHz, DMSO-d_6): δ = 168.9, 150.2, 150.2, 147.3, 147.2, 146.5, 146.2, 141.7, 141.1, 141.1, 136.7, 135.6, 124.6, 124.5, 121.5, 121.5, 117.2, 116.3, 68.8, 67.8, 65.7, 65.7, 50.3, 50.3, 50.2, 50.1, 47.9, 47.7, 20.5 ppm; ^{31}P NMR (243 MHz, DMSO-d_6): δ = -6.27 ppm; HRMS (ESI pos., MeOH): m/z $[M+H]^+$: 758.2414 (calc.), 758.2466 (found).

Monohydroxy monoethylazidophosphate tweezer 1.2: Yield: 0.26 g, 0.36 mmol, 95 %; ^{1}H NMR (600 MHz, DMSO-d_6): δ = 7.13 (s, 2H), 7.09 (s, 2H), 7.06 (tt, *J* = 4.7, 2.3 Hz, 4H), 6.85 – 6.76 (m, 4H), 4.31 – 4.20 (m, 4H), 4.08 (d, *J* = 1.6 Hz, 4H), 3.24 – 3.17 (m, 2H), 2.70 – 2.62 (m, 2H), 2.33 – 2.17 (m, 8H) ppm; ^{13}C NMR (151 MHz, DMSO-d_6): δ = 150.4, 150.3, 147.3, 147.0, 147.0, 146.9, 142.1, 140.1, 140.1, 136.7, 132.6, 124.6, 124.5, 121.5, 121.2, 116.8, 116.2, 68.0, 67.7, 65.6, 65.6, 50.3, 50.0, 50.0, 47.7, 46.8 ppm; ^{31}P NMR (243 MHz, DMSO-d_6): δ = -6.06 ppm; HRMS (ESI pos., MeOH): m/z $[M+Na]^+$: 738.2128 (calc.), 738.2130 (found).

Monophosphate monoethylazidophosphate tweezer 2: Yield: 0.16 g, 0.20 mmol, 95 %; ^{1}H NMR (600 MHz, DMSO-d_6): δ = 7.15 (s, 2H), 7.11 (s, 2H), 7.08 (p, *J* = 4.1 Hz, 4H), 6.82 – 6.77 (m, 4H), 4.34 (d, *J* = 1.7 Hz, 2H4.32 (t, *J* = 1.6 Hz, 2H), 4.12 – 4.05 (m, 4H), 3.65 (dt, *J* = 6.9, 4.7 Hz, 2H), 3.10 – 3.00 (m, 2H), 2.35 – 2.18 (m, 8H) ppm; ^{13}C NMR (151 MHz, DMSO-d_6): δ = 150.3, 150.3, 147.1, 147.0, 146.7, 146.7, 141.8, 141.0, 141.0, 136.4, 136.4, 135.9, 135.8, 124.5, 124.5, 121.5, 121.3, 116.8, 116.8, 68.9, 67.8, 65.7, 65.7, 50.3, 50.3, 50.2, 48.0, 47.9 ppm; ^{31}P NMR (243 MHz, DMSO-d_6): δ = -5.37, -6.12 ppm; HRMS (ESI pos., MeOH): m/z $[M+H]^+$: 796.1972 (calc.), 796.1979 (found).

General procedure for novel multivalent alkynes: Multivalent alkynes were used as platforms for the multiple click reaction; they were synthesized according to the general procedure detailed below.[4]

Tetrabut-3-yn-1-ylesters (3.1, 3.2) and Pentabut-3-yn-1-ylesters (3.3) of aromatic polycarboxylic acids: The respective aromatic polycarboxylic acid (3.90 mmol, 1.0 eq.) was suspended in dry dichloromethane (50 mL). To this suspension but-3-yn-1-ol (1.11 g, 1.20 mL, 15.7 mmol, 4.0 eq. for tetracarboxylate and 5.0 eq. for pentacarboxylate) and EDC*HCl (3.32 g, 17.3 mmol, 4.4 eq. for tetracarboxylate and 5.5 eq. for pentacarboxylate) were slowly dropwise. DMAP (96.1 mg, 0.78 mmol, 0.2 eq.) dissolved in dry dichloromethane (10 mL) was added to the reaction mixture. After stirring for 16 h at room temperature, AcOH (aq., 0.1 M, 5 mL) was added. The organic was separated from the aqueous phase. The aqueous phase was extracted with dichloromethane (3 x 50 mL). The combined organic layer was washed with AcOH (aq., 0.1 M, 2 x 100 mL), with $NaHCO_3$ (aq., sat., 2x 100 mL) and with NaCl (aq., sat., 100 mL). The organic phase was dried over $MgSO_4$, and the solvent was removed in vacuo. The residue was purified via flash column chromatography (SiO_2, cyclohexane: ethyl acetate 9:1).

Tetra(but-3-yn-1-yl) benzene-1,2,4,5-tetracarboxylate 3.1 was obtained as a colorless oil (32.9 mg, 0.07 mmol, 42 %); ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.10 (s, 2H), 4.45 (t, *J* = 6.8 Hz, 8H), 2.66 (td, J = 6.8, 2.7 Hz, 8H), 2.04 (t, *J* = 2.7 Hz, 4H) ppm; ^{13}C NMR (101 MHz, $CDCl_3$): δ = 165.6, 134.3, 131.8, 131.7, 130.0, 129.0, 79.7, 70.5, 63.9, 19.0, 1.2 ppm; HRMS (ESI pos., MeOH): m/z $[M+Na]^+$: 485.1207 (calc.), 485.1209 (found).

Di(but-3-yn-1-yl) 4',5'-bis(4-((but-3-yn-1-yloxy)carbonyl)phenyl)-[1,1':2',1''-terphenyl]-4,4''-dicarboxylate 3.2 was obtained as a colorless solid (0.37 g, 0.53 mmol, 53 %); ^{1}H NMR (400 MHz, CDCl3): δ 8.06 – 7.81 (m, 8H), 7.51 (s, 2H), 7.38 – 7.07 (m, 8H), 4.39 (t, *J* = 6.8 Hz, 8H), 2.63 (td, *J* = 6.8, 2.7 Hz, 8H), 2.00 (t, *J* = 2.7 Hz, 4H) ppm; ^{13}C NMR (101 MHz, $CDCl_3$): δ 166.2, 145.2, 139.7, 133.0, 130.0, 129.8, 128.8, 80.2, 70.2, 62.9, 19.2 ppm; HRMS (ESI pos., MeOH): m/z $[M+Na]^+$: 789.2459 (calc.), 789.2462 (found).

Penta(but-3-yn-1-yl) benzene-1,2,3,4,5-pentacarboxylate 3.3 was obtained as a colorless oil (87.9 mg, 0.16 mmol, 39 %); ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.52 (s, 1H), 4.39 (t, J = 6.5 Hz, 4H), 4.33 (td, J = 6.5, 2.3 Hz, 6H), 2.93 (t, J = 2.6 Hz, 3H), 2.90 (t, J = 2.6 Hz, 2H), 2.66 (td, J = 6.5, 2.7 Hz, 4H), 2.58 (dtd, J = 11.8, 6.5, 2.7 Hz, 6H) ppm; ^{13}C NMR (101 MHz, DMSO-*d*6): δ = 164.9, 164.0, 163.2, 136.5, 133.28, 130.9, 130.3, 80.5, 80.4, 80.2, 72.9, 72.9, 72.6, 64.8, 64.4, 64.2, 18.2, 17.9, 17.9, 14.1 ppm; HRMS (ESI pos., MeOH): m/z $[M+Na]^+$: 581.1418 (calc.), 581.1424 (found).

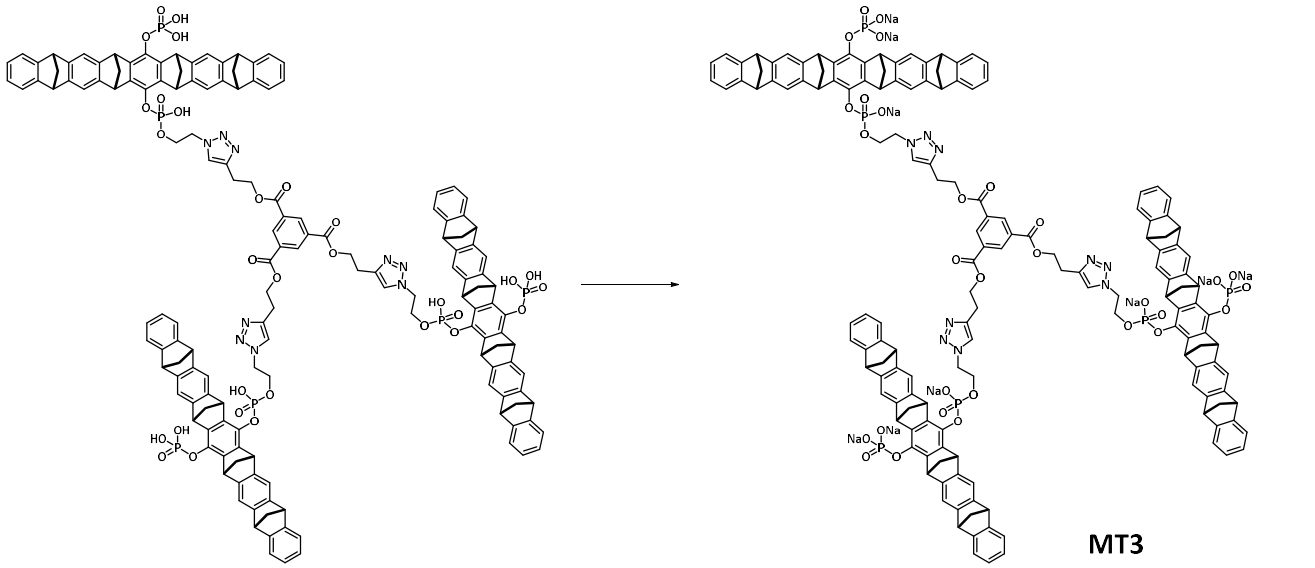
SUPPORTING INFORMATION

Novel multivalent tweezers were synthesized according to general procedures listed below:
Under argon atmosphere monophosphate monoethylazidophosphate tweezer **2** (10.0 mg, 0.01 mmol, 1.0 eq.) and alkyne (1.0 eq.) were dissolved in dry THF (2 mL), H_2O (1 mL) and DIPEA (8.77 mg, 11.9 µL, 0.07 mmol, 5.4 eq.). The transparent solution was stirred for 5 min at room temperature. A solution of copper sulfate pentahydrate (8.30 mg, 0.03 mmol, 3 eq.) and sodium ascorbate (13.0 mg, 0.07 mmol, 5.2 eq.) in H_2O (1 mL) was added to the reaction mixture and stirred for 16 h. Reaction control was performed via HPLC analysis of small samples taken directly from the reaction mixture. After complete consumption of the starting material, another equivalent of azidotweezer **2** was added *in situ* (10.0 mg, 0.01 mmol, 1.0 eq.). Likewise, the whole reagent cocktail, i.e., DIPEA (8.77 mg, 11.9 µL, 0.07 mmol, 5.4 eq.), dry THF (1 mL), copper sulfate pentahydrate (8.30 mg, 0.03 mmol, 3 eq.) and sodium ascorbate (13.0 mg, 0.07 mmol, 5.2 eq.) dissolved in water (1 mL), was again added *in situ*. This sequence was repeated as often as tweezer units could be accommodated in the final product. After full conversion was reached via HPLC control, the solvent was removed under vacuum and HCl (aq., 1 M, 30 mL) was added. The suspension was transferred into a falcon tube and sonicated twice for 15 min. After centrifugation for 15 min at 4000 rpm, the solvent was decanted, and the solid residue was isolated. It was washed with H_2O (30 mL) and with Et_2O (30 mL). After isolation via centrifugation, the multivalent tweezer was dissolved in mixture of THF and H_2O (1:1). After lyophilisation, the desired product was obtained as the free phosphoric acid.

The multivalent diphosphoric acid tweezer derivative was obtained as a colorless solid in moderate to good yield. Conversion to the sodium salt was performed by dissolving the free diphosphoric acid (1.0 eq.) in a 1:1 mixture of MeCN:H_2O (5 ml). Aq. NaOH (0.23 M, 1.0 eq. for each tweezer unit) was added, and the reaction mixture was stirred for 1 h at room temperature. The solvent was removed under vacuum and the compound was lyophilised. The sodium salt product was obtained as a colorless solid in quantitative yield.

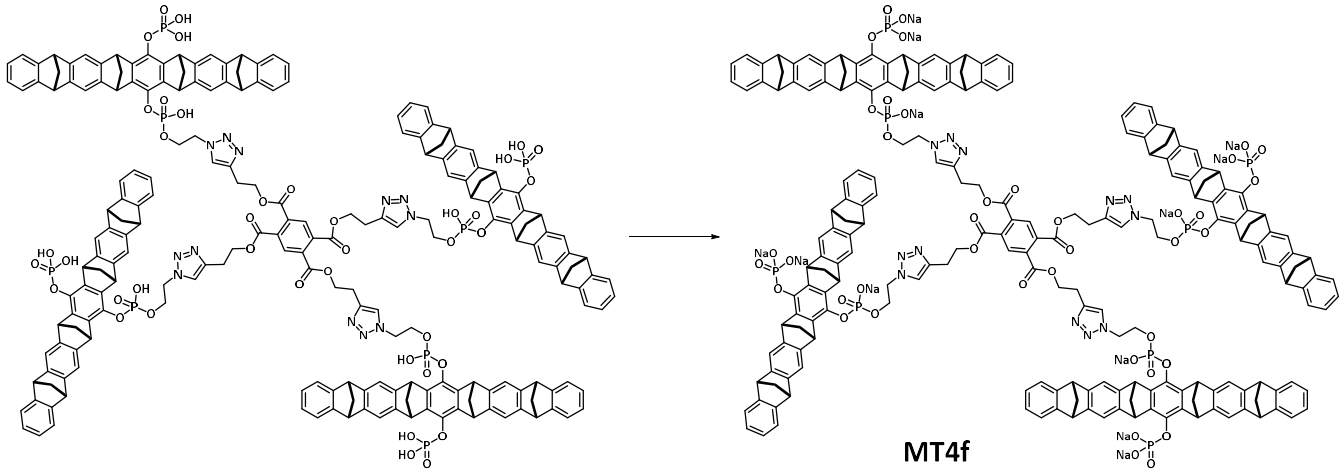
The free phosphoric acid **MT2** was obtained as a colorless solid (9.7 mg, 5.2 µmol, 47 %.); ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.18 (s, 2H), 8.15 (s, 4H), 7.12 – 6.89 (m, 16H), 6.82 – 6.71 (m, 8H), 4.52 (t, *J* = 6.3 Hz, 4H), 4.38 – 4.21 (m, 12H), 4.03 (d, *J* = 24.3 Hz, 12H), 3.12 (t, *J* = 6.3 Hz, 4H), 2.38 – 2.11 (m, 16H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 165.0, 150.3, 146.9, 146.8, 146.7, 143.3, 141.6, 141.0, 133.6, 129.8, 124.5, 123.6, 121.6, 121.3, 116.8, 116.7, 68.9, 67.7, 50.3, 50.24, 48.01, 47.9, 25.0 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -5.43, -6.24 ppm; HRMS (ESI pos., MeOH): m/z $[M+H]^{2+}$: 931.7435 (calc.), 931.7438 (found).

Deprotonation with aq. NaOH afforded the disodium salt **MT2** as a colorless powder in quantitative yield; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.15 (d, J = 14.9 Hz, 2H), 8.10 (s, 2H), 7.37 – 7.27 (m, 2H), 7.11 – 6.86 (m, 16H), 6.75 (dt, J = 7.6, 3.7 Hz, 8H), 4.53 (s, 4H), 4.43 - 4.28 (m, 8H), 4.12 (s, 4H), 4.07 – 3.88 (m, 8H), 3.75 – 3.62 (m, 2H), 2.37 – 1.97 (m, 16H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 150.3, 147.0, 146.6, 141.7, 141.0, 130.3, 125.7, 124.5, 122.1, 121.4, 116.8, 69.1, 67.8, 65.1, 59.5, 59.1, 50.3, 47.9 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -4.86, -4.77 ppm.



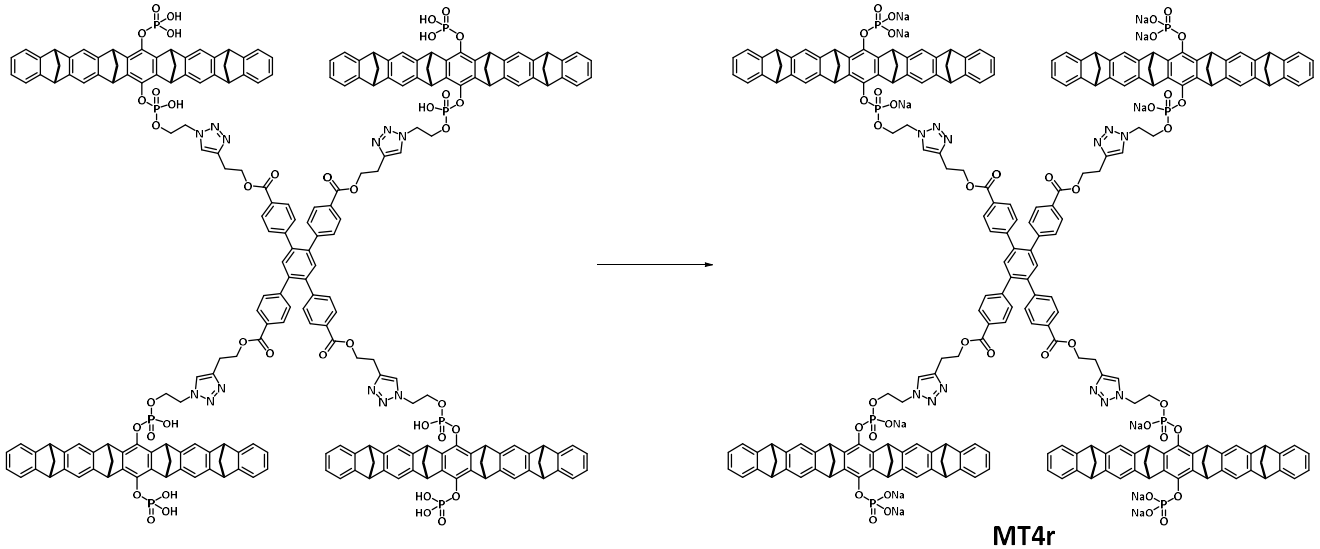
The free phosphoric acid **MT3** was obtained as a colorless solid (40.0 mg, 14.5 µmol, 89 %); ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.66 (s, 3H), 8.16 (s, 3H), 7.25 – 6.92 (m, 24H), 6.75 (p, *J* = 7.4 Hz, 12H), 4.74 – 4.52 (m, 6H), 4.29 (d, *J* = 23.3 Hz, 18H), 4.03 (d, *J* = 17.9 Hz, 18H), 3.30 – 3.08 (m, 6H), 2.38 – 2.09 (m, 24H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 164.1, 150.3, 146.9, 141.6, 141.0, 124.5, 121.5, 121.3, 116.9, 68.9, 67.8, 64.5, 53.5, 50.3, 48.0, 24. 9, 18.1, 16.7 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -5.41, -6.20 ppm; HRMS (ESI pos., MeOH): m/z $[M\text{-}H^+]^{2+}$: 1377.8490 (calc.), 1377.8497 (found).

Deprotonation with aq. NaOH afforded the disodium salt **MT3** as a colorless powder in quantitative yield; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.59 (s, 3H), 8.24 – 7.88 (m, 3H), 7.07 – 6.85 (m, 24H), 6.74 (s, 12H, H-1, H-12, H-13, H-24), 4.55 (s, 6H), 4.41 (d, J = 9.2 Hz, 18H), 4.23 (dd, J = 41.0, 17.2 Hz, 18H), 4.02 (q, J = 13.2, 12.4 Hz, 6H), 2.41 – 1.98 (m, 24H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 164.1, 150.3, 146.9, 141.6, 141.0, 124.5, 121.5, 121.3, 116.9, 68.9, 67.8, 64.5, 53.5, 50.3, 48.0, 24. 9, 18.1, 16.7 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -3.71, -4.74 ppm.



The free phosphoric acid **MT4f** was obtained as a colorless solid (24.0 mg, 6.58 µmol, 51 %); ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.12 (d, *J* = 8.7 Hz, 6H), 7.20 – 6.91 (m, 32H), 6.85 – 6.67 (m, 16H), 4.55 (s, 8H), 4.30 (d, *J* = 17.7 Hz, 24H), 4.04 (d, *J* = 15.1 Hz, 24H), 3.13 (s, 8H), 2.35 – 2.13 (m, 32H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 162.3, 155.1, 150.3, 146.9, 146.7, 141.0, 124.5, 120.9, 120.4, 118.1, 118.1, 116.8, 68.9, 67.8, 50.3, 48.0 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -5.42, -6.24 ppm; HRMS (ESI pos., MeOH): m/z $[M\text{-}H^+]^{3+}$: 1215.8185 (cal), 1215.9496 (found).

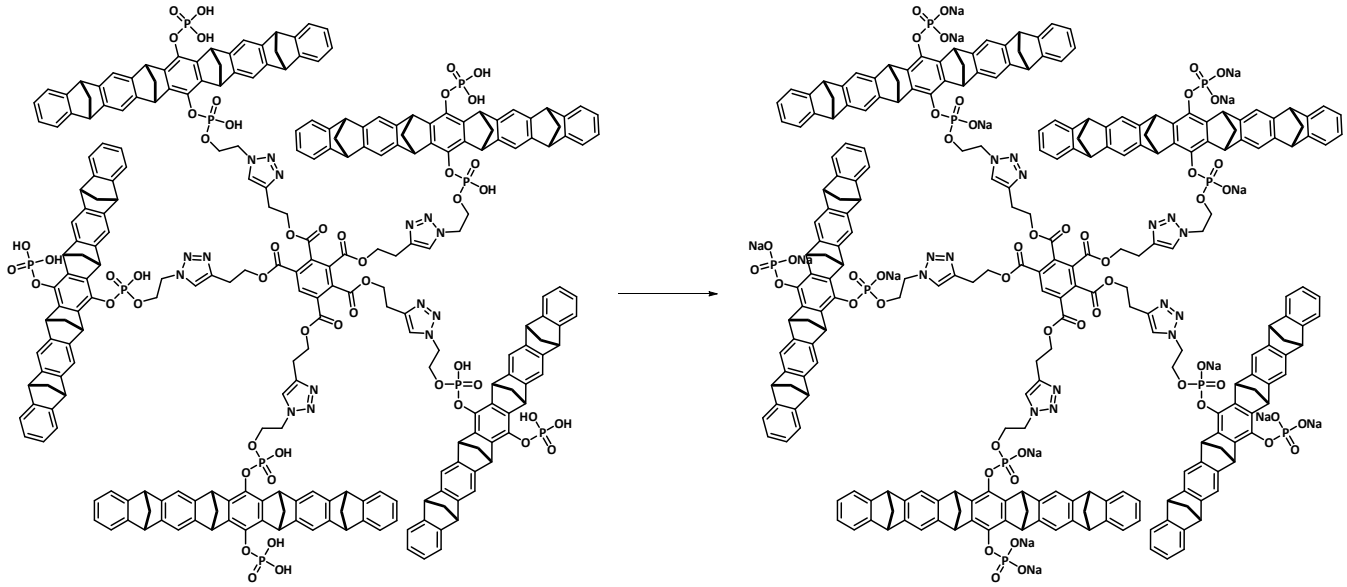
Deprotonation with aq. NaOH afforded the disodium salt **MT4f** as a colorless powder in quantitative yield; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 7.91 (s, 6H), 7.20 – 6.84 (m, 32H), 6.73 (m, 16H), 4.48 (s, 8H), 4.41 -4.13 (m, 24H), 4.05 (m, 24H), 3.63 (s, 8H), 2.35 – 1.99 (m, 32H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 162.3, 155.1, 150.3, 146.9, 146.7, 141.0, 124.5, 120.9, 120.4, 118.1, 118.1, 116.8, 68.9, 67.8, 50.3, 48.0 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -4.72, -3.72 ppm.



MT4r

The free phosphoric acid **MT4r** was obtained as a colorless solid (24.8 mg, 6.23 µmol, 69 %); ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.10 (s, 4H), 7.90 – 7.78 (m, 8H), 7.56 – 7.46 (m, 2H), 7.31 (d, *J* = 8.9 Hz, 8H), 7.13 – 6.95 (m, 32H), 6.75 (dq, *J* = 13.5, 7.3, 5.4 Hz, 16H), 4.52 (s, 8H), 4.42 (s, 8H), 4.37 – 4.28 (m, 8H), 4.21 (s, 8H), 4.17 – 4.08 (m, 8H), 4.04 (d, *J* = 9.3 Hz, 16H), 3.22 – 3.06 (m, 8H), 2.30 – 2.08 (m, 32H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 165.4, 165.4, 159.8, 150.3, 146.9, 146.7, 144.5, 141.7, 141.0, 138.9, 136.5, 130.2, 129.3, 128.4, 124.7, 121.6, 116.9, 68.8, 67.7, 65.1, 50.2, 48.0, 47.9, 25.1 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -5.40, -6.13 ppm; HRMS (ESI neg., MeOH): m/z [M-H]$^{3-}$: 1315.3337 (calc.), 1315.3335 (found).

Deprotonation with aq. NaOH afforded the disodium salt **MT4r** as a colorless powder in quantitative yield; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.09 (s, 4H), 7.90 – 7.78 (m, 8H), 7.56 – 7.46 (m, 2H), 7.31 (m, 8H), 7.13 – 6.95 (m, 32H), 6.75 (m, 16H), 4.52 (s, 8H), 4.42 (s, 8H), 4.37 – 4.28 (m, 8H), 4.21 (s, 8H), 4.17 – 4.08 (m, 8H), 4.04 (d, *J* = 9.3 Hz, 16H), 3.22 – 3.06 (m, 8H), 2.43 – 1.79 (m, 32H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 165.4, 165.4, 159.8, 150.3, 146.9, 146.7, 144.5, 141.7, 141.0, 138.9, 136.5, 130.2, 129.3, 128.4, 124.7, 121.6, 116.9, 68.8, 67.7, 65.1, 50.2, 48.0, 47.9, 25.1 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -5.40, -5.40 ppm.



MT5

The free phosphoric acid **MT5** was obtained as a colorless solid (16.6 mg, 3.66 µmol, 41 %); ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.47 (d, *J* = 21.2 Hz, 1H), 8.10 (d, *J* = 22.7 Hz, 5H), 7.16 – 6.89 (m, 40H), 6.75 (d, *J* = 6.2 Hz, 20H), 4.54 (s, 10H), 4.29 (d, *J* = 17.5 Hz, 30H), 4.04 (d, *J* = 13.6 Hz, 30H), 3.12 (d, *J* = 23.2 Hz, 10H), 2.21 (d, *J* = 36.3 Hz, 40H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 150.3, 147.0, 146.7, 141.6, 141.0, 124.5, 121.5, 121.4, 116.8, 113.9, 69.0, 67.8, 50.3, 49.8, 47.9 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -5.42, -6.25 ppm; HRMS (ESI neg., MeOH): m/z [M-H]$^{4-}$: 1133.0199 (calc.), 1133.0160 (found).

Deprotonation with aq. NaOH afforded the disodium salt **MT5** as a colorless powder in quantitative yield; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.04 (s, 1H), 8.03 (s, 5H), 7.57 – 6.89 (m, 40H), 6.86 – 6.47 (m, 20H), 4.81 -3.57 (m, 80H), 2.31 – 1.86 (d, *J* = 36.3 Hz, 40H) ppm; ^{13}C NMR (101 MHz, DMSO-d_6): δ = 150.3, 147.0, 146.7, 141.6, 141.0, 124.5, 121.5, 121.4, 116.8, 113.9, 69.0, 67.8, 50.3, 49.8, 47.9 ppm; ^{31}P NMR (162 MHz, DMSO-d_6): δ = -5.54, -4.79 ppm.

Expression and purification of recombinant proteins

pET22b-Taspase1$_{D233A/T234}$-His,[1] pET22b-WT_Taspase1-His,[1] pET41-GST-PreScn-Importinα,[2] pET22b-GST-MLL$_{2700-2850}$-GFP-His and the ^{15}N-labelled Taspase 1 loop$_{178-233}$ for NMR titrations were expressed and purified as previously reported.[5-7] Purity was validated using SDS-PAGE and immunoblotting.

Pull-down assay

We used Dulbecco's Phosphate Buffered saline (Sigma-Aldrich) containing 0.1 % (v/v) Triton X-100 (Carl Roth) and 1 mM DTT (Carl Roth) (PBST) for the pull-down assay. If not stated otherwise all incubation steps were carried out at 4 °C to prevent protein degradation and all centrifugation steps were carried out at 500 xG. Samples for immunoblotting were mixed with 5x sample buffer[3] and heated to 95 °C for 5 min. 50 µM slurry of Glutathione Sepharose 4B (Merck) was transferred to a Spin Column (IBA Lifescience) and equilibrated with 500 µL PBST followed by centrifugation. 500 µL 2.5 µM GST-Importinα were added to the column, a sample from the "input" fraction was taken and the column was incubated for 2 h on a rotator. The samples from the "input" fraction were collected to validate, that equal amounts of protein were added to all columns. Unbound protein was removed by three washing steps with PBST followed by centrifugation. 500 µL 2.2 µM inactive Taspase1$_{D233A/T234A}$-His was pre-incubated with the respective compound in PBST with 3 % DMSO (v/v) on a rotator for 1 h. The DMSO concentration of the untreated control was adjusted. Again, a sample for the "input" fraction was retained for later analysis. Free binding sites on the column were blocked with 1 % (w/v) BSA (Carl Roth) in PBST for 30 min on a rotator. The blocking solution was removed from the column by centrifugation. Subsequently, Taspase1-$_{D233A/T234A}$-His pre-incubated with the compounds or DMSO was added to the column and incubated on a rotator for 1 h to allow binding to GST-Importinα. Unbound protein was removed by centrifugation. A sample of the "unbound" fraction was taken to validate that the compound did not elute GST-Importinα from the column. We applied three washing steps with PBST followed by centrifugation. Finally, 500 µL 1x sample buffer[3] was added to the column and heated to 95 °C for 10 min to denature and thereby elute all proteins from the column. Eluted proteins were collected by centrifugation for 2 min and the samples analyzed using SDS-PAGE and immunoblotting.

SDS-PAGE and immunoblotting

We prepared tris-glycine gels with 12.5 % or 10 % (v/v) acrylamide in the stacking gel and 4 % (v/v) acrylamide in the separating gel. A TetraCell system (BioRad) was set to 200 V for 45 min electrophoresis. For the cleavage assay the gels were stained with Coomassie and analyzed. For the pull-down assay the proteins were then transferred to a protein-binding nitrocellulose membrane using a wet blot tank (Peqlab) set to 360 mA for 90 min at 4 °C. To validate the transfer, the membrane was reversibly stained with Ponceau S (AppliChem) and the membrane cut between the protein bands according to the Spectra Multicolor Broad Range Protein Ladder (Thermo Fisher) to analyze proteins from the same sample. 5 % (w/v) powdered milk (Carl Roth) in tris buffered saline with Tween-20 (Carl Roth) (TBST) was used to block free binding sites with for 60 min at room temperature. The membranes were incubated with the respective primary antibodies rabbit anti-Taspase1 1:2000 (sc-85945, Santa Cruz) or mouse anti-Karyopherinα2 1:1000 (sc-55538, Santa Cruz) in 5 % (w/v) powdered milk in TBST over night at 4 °C. After three washing steps with TBST the Membranes were incubated with the respective secondary antibodies donkey anti-rabbit HPR-coupled 1:10000 (NA934, GE Healthcare) or sheep anti-mouse HPR-coupled 1:10000 (NXA931, GE Healthcare) in 5 % (w/v) powdered milk with TBST for 1 h at room temperature. The membrane was washed in TBST four times. We used the Pierce ECL Plus Western Blotting Substrate (Thermo Fisher) and the Chemidoc Imaging System (BioRad) to detect chemiluminescence generated by the HPR-coupled secondary antibodyThe imaging system was set to automatic exposure to avoid overexposure and to ensure equal maximum signal intensities in all blots. Fiji was used for densitometric quantification of the signals.[4] If necessary, the signal of Taspase1 in the eluted fraction was corrected for Taspase1 bound to the column in absence of GST-Importin α. To correct possible loading differences of the gel, the signal of Taspase1 in the eluted fraction was normalized to the signal of Importinα in the eluted fractions. The data was evaluated using Origin2019 (OriginLab).

Fluorescence titration with the isolated Taspase1 loop

The Taspase1 loop with an N-terminal FAM-label was synthesized by GeneCust. (Sequence: FAM-SCPPNIMTTRFSLAAF**KRNKRK**LELAERVDTDFMQL**KKRR**QSSEKENDSGTLD, NLS highlighted. 1 µM Taspase1 loop in PBS was added to each well of a 96 well plate (Corning). As a negative control we used FAM without the loop. The initial fluorescence for each well was determined using the Promega GloMax (Promega) equipped with the blue filter set (Excitation 490 nm, Emission 510-570 nm). After that, the Taspase1-loop was mixed with respective concentrations of compound solved in PBS with 3 % (v/v) DMSO or PBS with 3 % (v/v) DMSO for the untreated controls. The same volume was added to each well to avoid effects caused by dilution of the sample. To the FAM-label without loop we added the highest concentration of compound solved in PBS with 3 % (v/v) DMSO or PBS with 3 % (v/v) to validate that the compound does not bind to the fluorophore instead of the loop. Fluorescence after treatment was determined using the Promega GloMax (Promega) equipped with the blue filter set (Excitation 490 nm, Emission 510-570 nm). The initial fluorescence of each well was subtracted from the fluorescence after treatment to determine the decrease in fluorescence ΔF. The data points are the mean of three replicates ± standard deviation. The data was fit using the following quadratic binding equation for a one-site specific binding model (Graph Pad Prism 5). The K_D is given as fit ± standard deviation.

Protein NMR spectroscopy

NMR experiments were recorded on a Bruker 700 MHz Avance Ultrashield NMR spectrometer (Bruker, Germany) with a 5 mm TCI $^{1}H/^{13}C/^{15}N/D$ cryoprobe with z-gradient at 25 °C. Spectra were processed with Topspin 3.5 and analyzed in CARA.[8]

^{15}N-Taspase 1 $loop_{178-233}$ (250-300 µM) in NMR buffer (1.5 mM KH_2PO_4, 8.9 mM Na_2HPO_4, 136.9 mM NaCl, 2.7 mM KCl, pH 6.5, 10 % D_2O was titrated with a 5 mM stock of molecular tweezer constructs in NMR buffer up to one equivalent of ligand. $^{1}H,^{15}N$-HSQC spectra were recorded for each titration step. The relative signal intensities I/I_0 were evaluated,[9] where I represents the signal intensity in the presence of tweezers and I_0 the intensity in the spectrum without tweezers. A more than average decrease in intensity indicates ligand binding due to intermediate exchange kinetics. The amide chemical shift perturbations did not yield useful data because signals already disappeared at small ligand concentrations and thus did not allow tracking of their positions.

Colorimetric cleavage assay

This modified version of a recently established assay utilizes a recombinant fusionprotein GST-$MLL_{2700-2850}$-GFP-His with the CS2 cleavage site as a Taspase1 substrate. If not stated otherwise, all proteins are prepared in Taspase1 kinetic buffer (10 %sucrose, 50 µM NaH_2PO_4, pH 7.4). Taspase1 was thawed for 5 min and different aliquots pooled to assure equal protein activity among the samples. Taspase1 and respective concentrations of ligand or DMSO were added mixed in PCR strips. One negative control contained buffer instead of Taspase1 to validate, that the substrate is not degraded over time. The samples were cooled down to 4 °C on a TProfessional thermocycler (Biometra). Using a multipipette the substrate was added to each each sample and the program started. All samples were incubated at 37 °C for 90 Min. After that, the temperature was increased to 95 °C for 5 Min to denature all proteins and thereby stop the reaction. 5x sample buffer[3] was added to each sample and the temperature was increased to 95 °C for 5 min again. The proteins were then analyzed using SDS-PAGE and stained with Coomassie for visualization of the protein bands. The gels were covered with Coomassie staining solution and heated in the microwave at 800 W for 15 s to speed up the staining process. The staining solution was removed after 45 Min incubation on a rocking shaker and the gel covered in destaining solution. The destaining solution was repeatedly exchanged until all background stain was removed. For imaging we used a Chemidoc Imaging System (BioRad). Densitometric quantification of the signals was performed using Fiji.[4] To determine the background signal caused by the destained gel itself, we averaged the background of two clear positions in the negative control, where the two cleaved fragements would be and substracted this value from each sample. The background signal was further corrected using a subsequent rolling ball algorithm with 50 pixels. Of note, the amount of Taspase1 used was below the detection limit of Coomassie stain, thereby allowing easier quantification of the substrate protein bands. To avoid differences during the loading of the gels, we only compared the ratio between cleaved and uncleaved substrate per sample and used the relative cleaved substrate per sample for the determination of significance. After that, we normalized the samples to the untreated control to determine the effect of the compound on the proteolytic activity. To correct possible loading differences the signal of Taspase1 in the eluted fraction was normalized to the signal of Importinα in the eluted fractions. The data was evaluated using Origin2019 (OriginLab). To correct differences in the variance of the samples, significances were determined using the Welch-t-test.

SUPPORTING INFORMATION

Computational studies

The protein as well as the isolated loop structure were constructed according to the model from van den Boom.[7] Subsequent preparation was done using the protein preparation wizard (bond order assignment with the help of the CCD database, addition of hydrogens, filling of missing sidechain atoms, deletion of all waters and the generation of het states at pH 7.3 followed by H-bond assignment using PROPKA at pH 7.3 and a restrained minimization to converge the heavy atoms to RMSD below 0.3 Å.

Tweezers were placed manually onto the lysines/arginines in the flexible loop comprising residues 178-233. The NMR results were considered, as they revealed the highest line broadening/intensity drop on the regions from residue 190 to 202 as well as from 217 to 225. From those lysines/arginines the most accessible ones were selected for tweezer placement (R198, R201, R208, K217, K225). The resulting protein structure with five tweezers was relaxed by means of a restrained minimization. Those five tweezers were subsequently connected to the central unit of the MT5 molecule. We chose this according to the higher affinity and therefor relevance.

MD simulations were performed with Desmond from D. E. Shaw. The solvent model was SPC and the box shape orthorhombic with the box size at 10 Å buffer to the protein. Charges were neutralized by the addition of sodium ions and sodium chloride was added to a total concentration of 150 mM. The following molecular dynamics simulation was set up to run for 300 ns in total for the full protein and 1600 ns for the isolated loop. The trajectory was saved resulting in 1000 trajectory frames (energy saved every 1.2 ps). The ensemble class was NPT with a temperature of 300 K and a pressure of 1013 mbar. The model system was relaxed before the productive run by the standard relaxation protocol provided. Calculations were performed on a Nvidia GTX 1070 GPU.

Software

Images of Western blots and Coomassie gels were acquired by the Chemidoc Imaging System (BioRad), quantified by densitometric analysis with Fiji,[10] and the data was evaluated using Origin 2019 (OriginLab). For data acquisition of fluorescence titrations, was the software Spectra Manager™ II (Jasco) or GloMax (Promega) supplied with the instrument were used, and data was plotted with GraphPad Prism 8 (version 8.4.1). Cleavage assay data was evaluated using Origin2019 (OriginLab). NMR data were collected using Topspin 3.5 (Bruker) using the NMRlib 2.0 pulse sequence tools library from IBS (Grenoble, France) available at http://www.ibs.fr/research/scientific-output/software/pulse-sequence-tools/. NMR spectra were processed with Topspin 3.5 (Bruker) and analyzed in CARA (version 1.9.1.7; http://cara. nmr.ch). Relative signal intensities were calculated from the raw chemical shift data and peak intensities using Excel 2016 (Microsoft) and plotted with GraphPad Prism 5.0. All molecular modelling and dynamics calculations were performed in Maestro 13.0, the corresponding images were generated with VMD. Figure panels were assembled in CanvasDraw 6.0 (ACDsee) or PowerPoint 2016 (Microsoft).

Supporting Figures

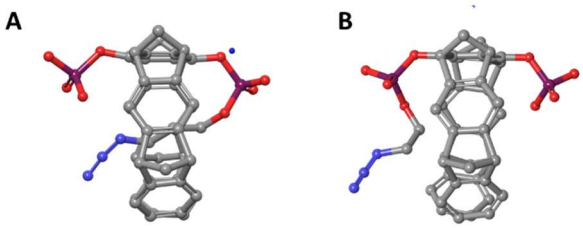


Figure S1. Self-inclusion of one phosphate ester arm inside the tweezer cavity depending on the length of the azidoalkyl chain. A. Monophosphate mono(azidopropyl)phosphate tweezer. **B.** Monophosphate mono(azidoethyl)phosphate tweezer. A Monte-Carlo calculation was performed in explicit water (10 000 steps), followed by a Molecular Dynamics simulation (FF: OPLS-AA, 100 ns, 25°C).

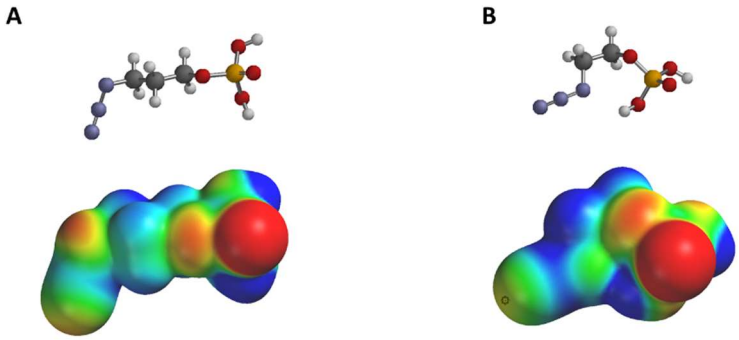


Figure S2. Electrostatic Surface Potential (ESP) of the nitrogen atoms of the azide groups in phosphate-substituted derivatives. A. Azidopropylphosphate $(HO)_2P(O)$-O-CH_2-CH_2-CH_2-N_3. **B.** Aazidoethylphosphate $(HO)_2P(O)$-O-CH_2-CH_2-$N_3$2. Calculated by PM3 implemented in SPARTAN 04 (Wavefunction Inc.) The colour code spans from −30 kcal mol^{-1} (red) to +23 kcal mol^{-1} (blue).

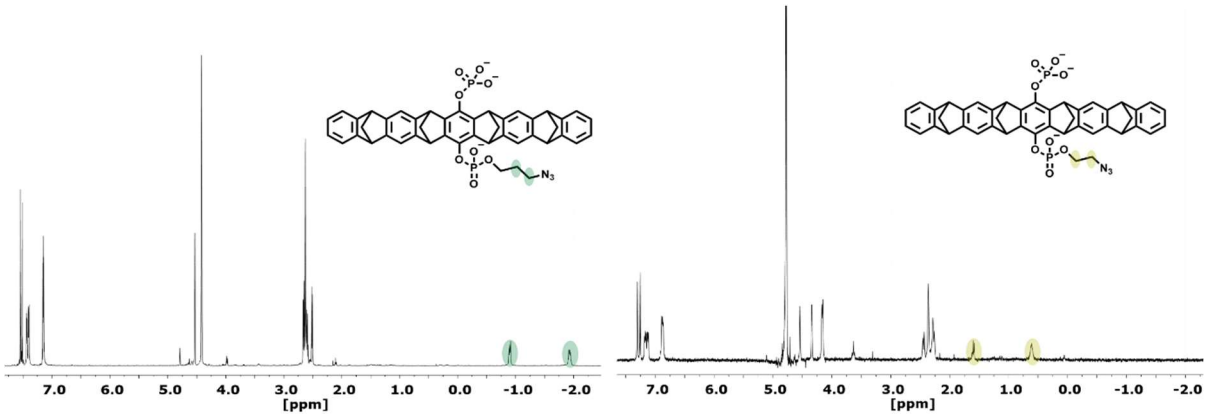


Figure S3. ^{1}H NMR spectra of the two monophosphate monoazidophosphate tweezers with C_2 vs. C_3 alkyl spacer. Note the drastic 2.5 ppm upfield shift of the included methylene protons in the ester arm. **A.** Monophosphate mono(azidopropyl)phosphate tweezer with self-inclusion. **B.** Monophosphate mono(azidoethyl)phosphate tweezer without self-inclusion.

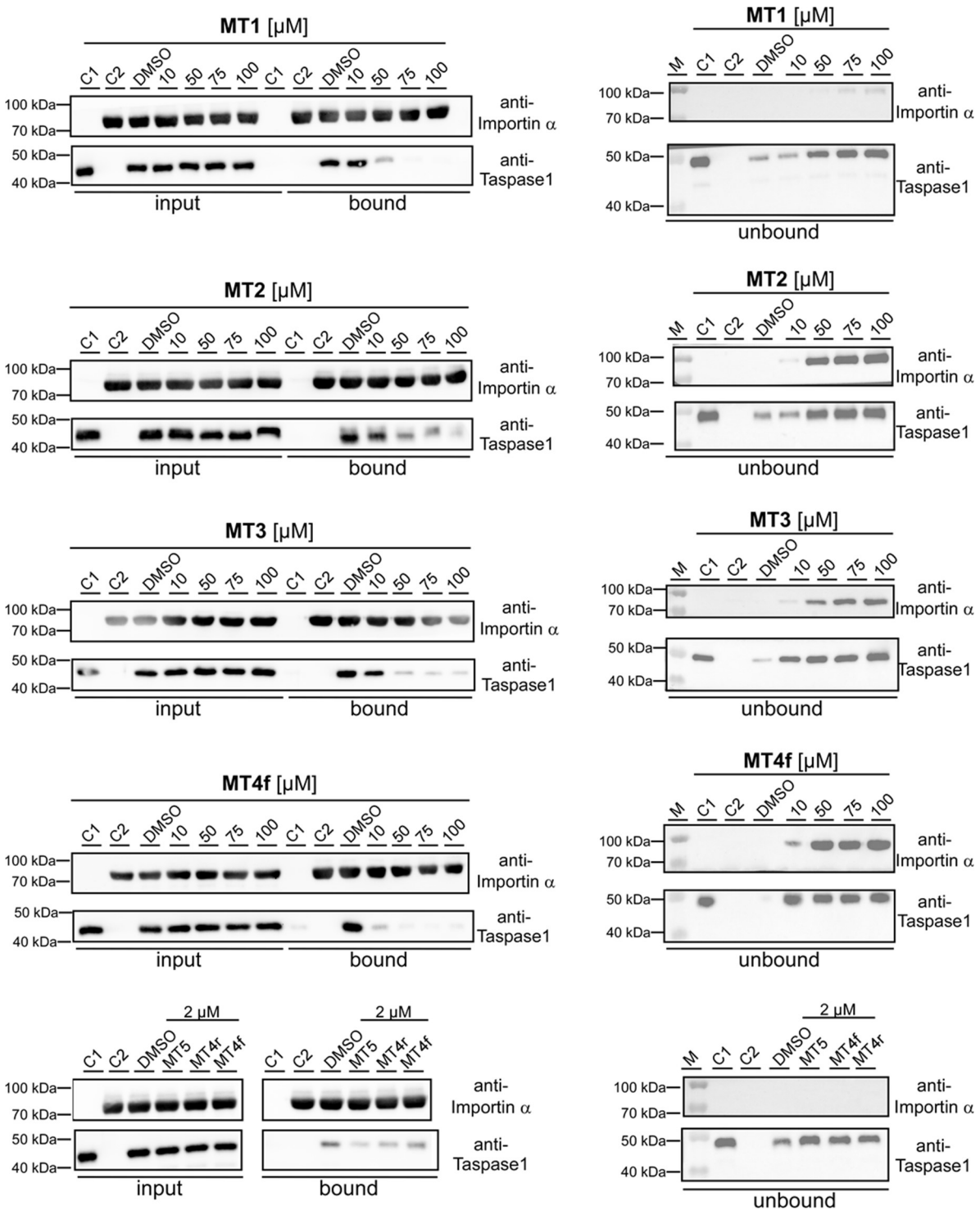


Figure S4. Multivalency increases the tweezers' potential to interfere with the Taspase 1/Importin α interaction. A. Pull-down assays were performed with increasing concentrations of different tweezers as indicated (see also Fig. 4A and Fig. 5A in the main manuscript). Controls: Taspase 1-His (C1) or GST-Importin α (C2) alone were added to the column, and a DMSO-treated sample served as reference. Binding of Importin α to the column was only affected at very high concentrations after Taspase1 was already removed. Controls include Taspase1 (C1), GST-Importinα (C2) or the inhibitor solvent (DMSO). M, molecular weight marker.

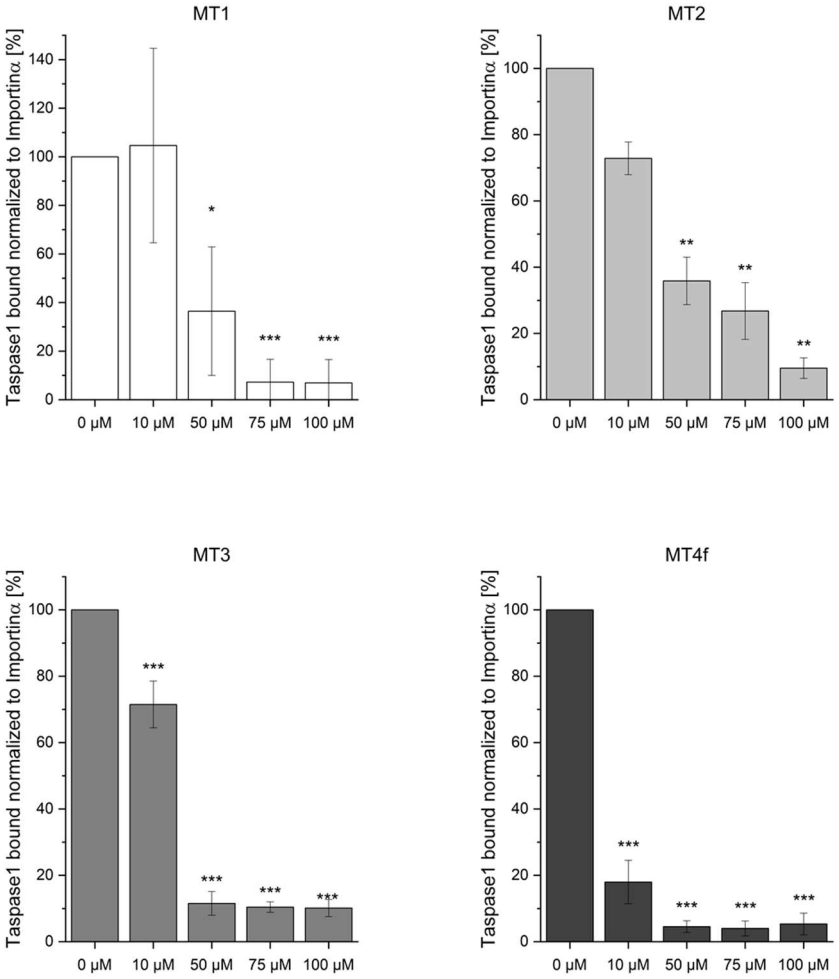


Figure S5. **Multivalency increases the tweezers' potential to interfere with the Taspase 1/Importin α interaction.** Densiometric quantification of pull-down assays performed with increasing concentrations (10-100 μM) of indicated tweezers MT1, MT2, MT3, and MT4f (see Fig. S1, and Fig. 4A in the main manuscript). Controls: Taspase 1-His (C1) or GST-Importin α (C2) alone were added to the column, and a DMSO-treated sample served as reference. Results are given as mean of three replicates including standard deviations (*p < 0.05, **p < 0,01) relative to the DMSO control and were represented as a heatmap in the main manuscript (see Fig. 4B).

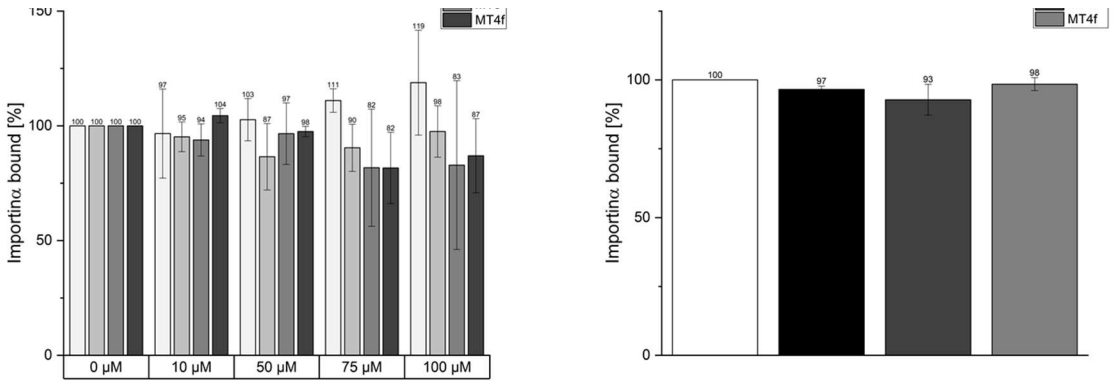


Figure S6. GST-Importinα in the „eluted" fraction of the pull-down assay was quantified to assure that the tweezers do not preferably elute Importin α from the matrix, which would result in a loss of Taspase 1. The results are the mean of three replicates ± standard deviation.

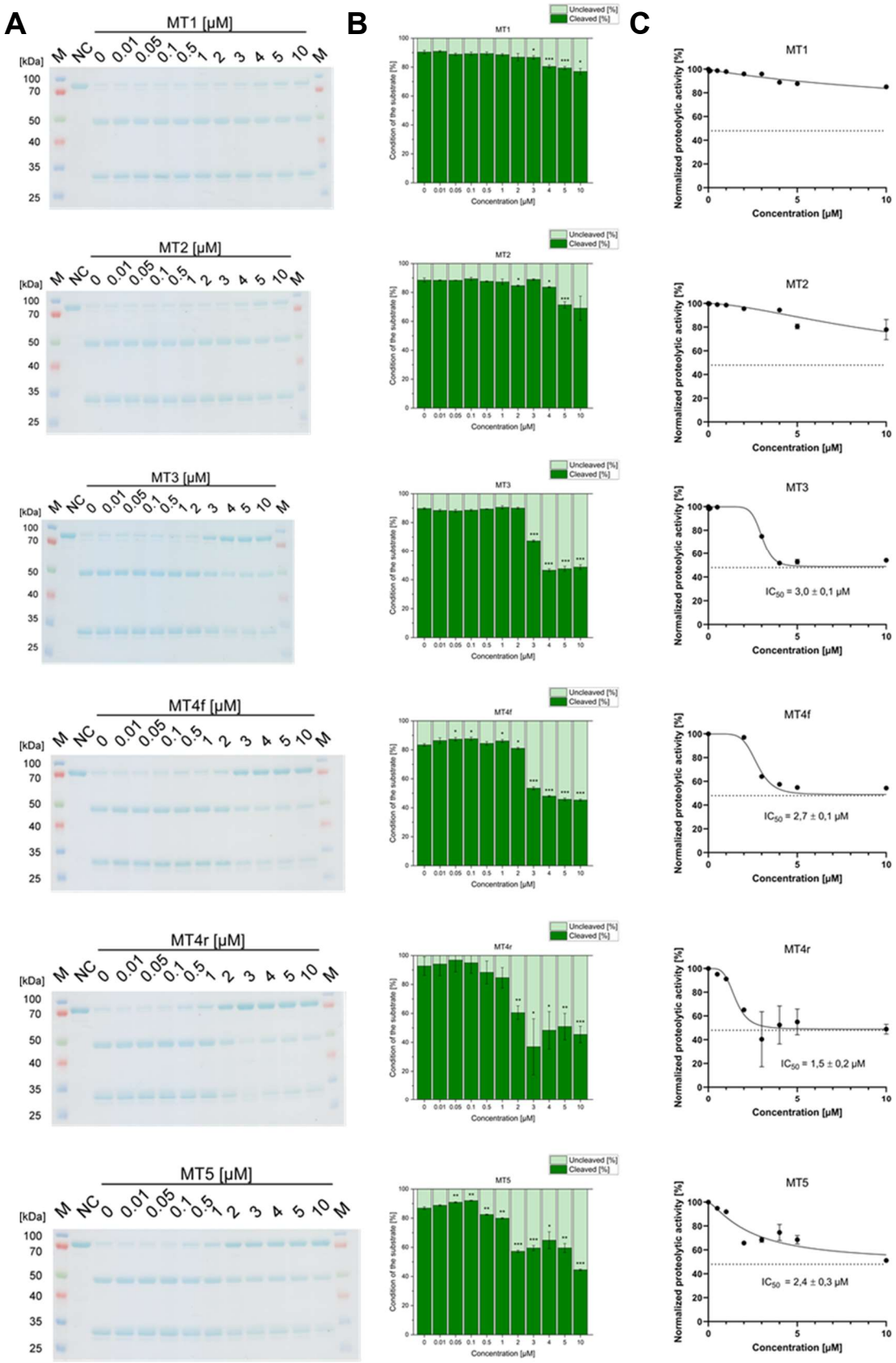


Figure S7. Multivalent supramolecular tweezers interfere with Taspase 1's proteolytic activity. A. Exemplary Coomassie stained gels of the cleavage assay showing the effect of the multivalent tweezer on the digestion of the recombinant Taspase1 substrate GST-MLL-GFP. **B.** Densitometric quantification of the Coomassie stained gels show a significant effect on the digestion of the substrate by Taspase1 after 90 Min incubation. The results are the mean of three replicates ± standard deviation (*p<0.5, **p<0.1, ***p<0.01). **C.** Effect of the multivalent tweezers on the proteolytic activity of Taspase1 normalized to an untreated control. The results are the mean of three replicats ± standard deviation. The IC_{50} are given as fit ± standard error of the mean.

References

[1] C. Heid, A. Sowislok, T. Schaller, F. Niemeyer, F. G. Klarner, T. Schrader, *Chemistry* **2018**, *24*, 11332-11343.
[2] M. Fokkens, T. Schrader, F. G. Klarner, *J Am Chem Soc* **2005**, *127*, 14415-14421.
[3] P. Talbiersky, F. Bastkowski, F. G. Klarner, T. Schrader, *J Am Chem Soc* **2008**, *130*, 9824-9828.
[4] D. Martella, C. Parmeggiani, D. S. Wiersma, M. Piñol, L. Oriol, *Journal of Materials Chemistry C* **2015**, *3*, 9003-9010.
[5] P. Pasch, A. Höing, S. Ueclue, M. Killa, J. Voskuhl, S. K. Knauer, L. Hartmann, *Chemical Communications* **2021**, *57*, 3091–3094.
[6] A. Höing, A. Zimmermann, L. Moews, M. Killa, M. Heimann, A. Hensel, J. Voskuhl, S. K. Knauer, *ChemMedChem* **2021**.
[7] J. van den Boom, F. Trusch, L. Hoppstock, C. Beuck, P. Bayer, *PloS one* **2016**, *11*, e0151431.
[8] R. L. J. Keller, ETH Zurich **2005**.
[9] P. Bayer, A. Matena, C. Beuck, *Beilstein journal of organic chemistry* **2020**, *16*, 2505–2522.
[10] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, *Nature methods* **2012**, *9*, 676–682.

7 Curriculum vitae

For reasons of data protection, the curriculum vitae is not included in the online version of this thesis.

For reasons of data protection, the curriculum vitae is not included in the online version of this thesis.

8 Eidesstattliche Erklärung

Erklärung:

Hiermit erkläre ich, gem. § 7 Abs. (2) d) + f) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich die vorliegende Dissertation selbständig verfasst und mich keiner anderen als der angegebenen Hilfsmittel bedient, bei der Abfassung der Dissertation nur die angegeben Hilfsmittel benutzt und alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet habe.

Essen, den _____ _____

Alexander Höing

Erklärung:

Hiermit erkläre ich, gem. § 7 Abs. (2) e) + g) der Promotionsordnung der Fakultät für Biologie zur Erlangung des Dr. rer. nat., dass ich keine anderen Promotionen bzw. Promotionsversuche in der Vergangenheit durchgeführt habe und dass diese Arbeit von keiner anderen Fakultät/Fachbereich abgelehnt worden ist.

Essen, den _____ _____

Alexander Höing

Erklärung:

Hiermit erkläre ich, gem. § 6 Abs. (2) g) der Promotionsordnung der Fakultät für Biologie zur Erlangung der Dr. rer. nat., dass ich das Arbeitsgebiet, dem das Thema „Rational targeting of Taspase1 nuclear import and biological function by supramolecular ligands“ zuzuordnen ist, in Forschung und Lehre vertrete und den Antrag von Alexander Höing befürworte und die Betreuung auch im Falle eines Weggangs, wenn nicht wichtige Gründe dem entgegenstehen, weiterführen werde.

Essen, den _____ _____

Prof. Dr. Shirley Knauer

Erklärung:

Hiermit erkläre ich, dass durch die Verwendung der Artikel in dieser kumulativen Dissertation keine Urheberrechte verletzt werden.

Essen, den _____ _____

Alexander Höing