UNIVERSITÀ DEGLI STUDI DI MILANO
DOCTORAL SCHOOL OF CHEMICAL SCIENCES AND TECHNOLOGIES
DEPARTMENT OF CHEMISTRY
PHD COURSE IN CHEMISTRY, XXXIV CYCLE

Development of Novel Strategies to Enhance the Affinity of Cyclic Peptide Ligands for Integrin Receptors

CHIM/06 Organic Chemistry

Giovanni SACCO
R12379

Tutor: Prof. Dr. Cesare GENNARI

Co-ordinator: Prof. Dr. Daniele PASSARELLA

A.Y. 2021/2022

The presented work was led by:	Prof. Dr. Cesare Gennari	
Doctoral Final Oral Examination:	April, 6th 2022	
Examination Committee:	Chairperson:	Prof. Dr. Sabrina DALLAVALLE Università degli Studi di Milano (I)
	Second Member:	Prof. Dr. Sandrine ONGERI Universitè Paris-Saclay (FR)
	Third Member:	Prof. Dr. Andrea TRABOCCHI Università degli Studi di Firenze (I)

The work herein described was performed at the University of Milan at the Department of Chemistry in the period from October 2018 to December 2021 under the supervision of Prof. Cesare Gennari.

I would like to acknowledge my supervisor, Prof. Cesare Gennari, for giving me the opportunity to work in a multidisciplinary project and in a stimulating environment. I gratefully thank Dr. Alberto Dal Corso, my mentor during my PhD jouney, and all the collaborators: Prof. Luca Pignataro (UNIMI), Prof. Laura Belvisi (UNIMI), Dr. Daniela Arosio (SCITEC-CNR) and Prof. Mayra Paolillo (UNIPV). Without all of you, this project would not have been possible. I would also like to thank all the chemists I've been working and living with every day at UNIMI. I specially thank my friends, "Casa Farruzz" members, for being my second family in Milan and making this journey amazing. I want to thank my girlfriend Emilia, for being special as she is and standing by my side. Finally, I would like to thank and dedicate this research work to my family, in particular to my brother Pio, my father Pino and my mother Maria. They are the anchor to which I will always be able to cling in difficult times and, above all, the first people with whom to enjoy happy moments.

Table of Contents

Abbreviations

2HB	2-hydroxybenzaldehyde
AAZ	Acetazolamide
Ac	Acetyl
ADC	Antibody-drug conjugate
ADCC	Antibody-dependent cellular cytotoxicity
ADCP	Antibody-dependent cellular phagocytosis
APD	Antibody phage display
aq.	Aqueous solution
Boc	*tert*-Butyloxycarbonyl
BSA	Bovine serum albumin
Bu	Butyl
CAIX	Carbonic anhydrase IX
CD3	Cluster of differentiation 3
CDC	Complement dependent cytotoxicity
CLL	Chronic lymphocytic leukemia
DIC	*N*,*N*'-Diisopropylcarbodiimide
DKP	Diketopiperazine
DMAP	4-Dimethylaminopyridine
DMF	*N*,*N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
EDC	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide
EGFR	Endothelial growth factor receptor
ELISA	Enzyme-linked immunosorbent assay
eq	Equivalents
ESI	Electrospray ionisation
Et	Ethyl
Fab	Antigen binding region
FACS	Fluorescence-activated cell scanning
FAK	Focal adhesion kinase
Fc	Fragment crystallizable region
FDA	Food and Drug Administration
Fmoc	9-Fluorenylmethoxycarbonyl
Fv	Variable fragment

HAMA	Human anti-mouse antibody
HAT	Hypoxanthine-aminopterin-thymidine
HATU	O-(7-azabenzotriazol-1-yl)-tetramethyl-uronium hexafluorophosphate
HGPRT	Hypoxanthine-guanine phosphoribosyltransferase
HOAt	1-Hydroxy-7-azabenzotriazole
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
IC	Inhibitory capacity
IgG	Immunoglobulin G
*i*Pr	Isopropyl
J	Scalar coupling constants
mAb	Monoclonal antibody
MAC	Membrane Attack Complex
Me	Methyl
MS	Mass spectroscopy
MW	Molecular weight
NMR	Nuclear Magnetic Resonance
NHS	*N*-Hydroxysuccinimide
NK	Natural killer
PBS	Phosphate-buffered saline
PEG	Polyethylene glycol
Ph	Phenyl
ppm	Part per million
PTX	Paclitaxel
quant.	Quantitative
ROS	Reactive oxygen species
r.t.	Room temperature
SMDC	Small molecue-drug conjugate
SPPS	Solid-phase peptide synthesis
*t*Bu	*tert*-Butyl
tert	Tertiary
TBDPS	Tert-butyldiphenylsilyl
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TMS	Tetramethylsilane
t_R	Retention time
δ	Chemical shift

Amino acid*	One-letter code	Three-letter code
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamic acid	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

* D-amino acids are described by D-Xaa in the three-letter code and with the small letter in the one-letter code.

General Introduction

In the last decades, novel pharmacological strategies for the treatment of various diseases have consisted in the use of monoclonal antibodies (mAbs). This highly attractive class of biotherapeutic agents has emerged as the technology of choice to interact with virtually all protein targets, due to their ability to interact with virtually all types of protein antigens, with exceptionally high affinity and selectivity. However, the structural features of mAbs often limit their efficacy, which explains the high interest of pharmaceutical companies in general methodologies to develop high affinity small molecule drugs. Despite the different advances in screening methodologies, small molecule ligands often show poor efficacy against specific pharmaceutical targets (such as different classes of protein-protein interactions), due to suboptimal binding affinity to the protein of interest. Together with the intensive research of new hits, the development of general strategies to enhance the binding affinity of a given small ligands may pave the way for the development of synthetically-accessible drugs with antibody-like affinity. In this context, this work of Thesis aims at the investigation of two general strategies to enhance the binding affinity of small peptide ligands: the use of multivalency and the development of reversible-covalent ligands (schematically represented in Figure 1).

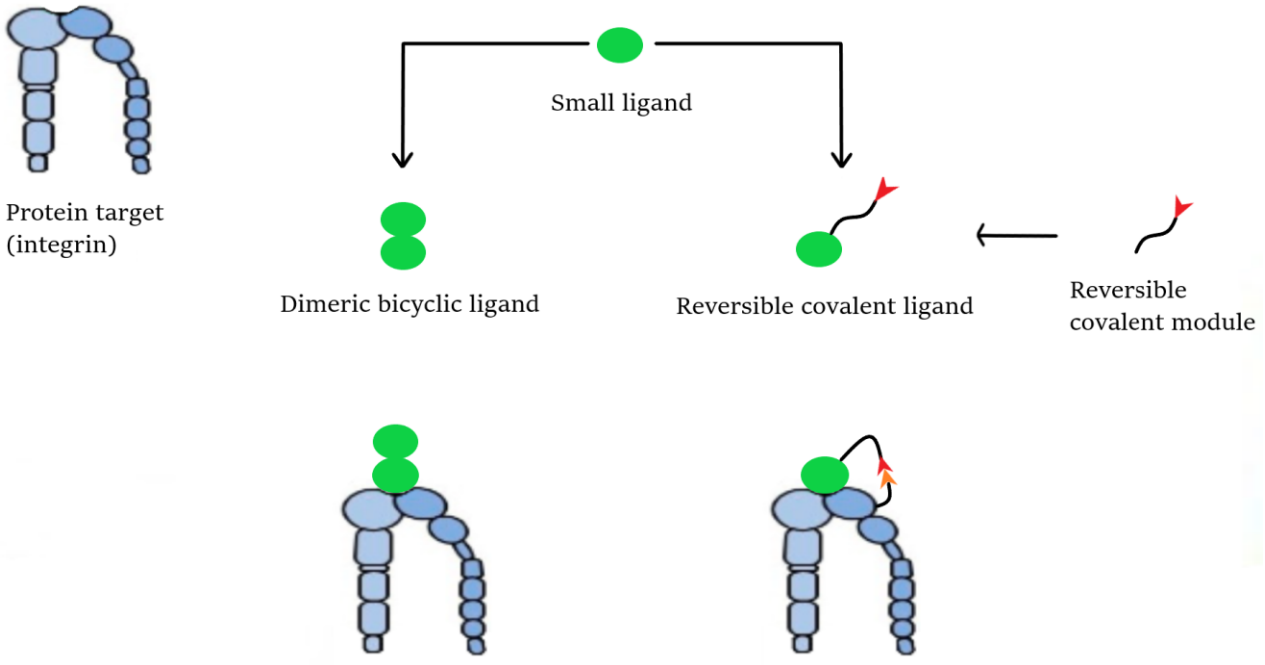


Figure 1 Chemical strategies investigated in this Thesis to promote the binding affinity enhancement of a small ligand. Under each strategy, the schematic interaction of the resulting ligand with the target protein is shown.

In this Thesis, integrins (in particular $\alpha_v\beta_3$ and $\alpha_5\beta_1$) have been selected as case-study protein receptors, whereas the chemical development of novel small molecule ligands concerned cyclic peptides bearing the Arg-Gly-Asp (RGD) integrin binding motif. The multivalency approach is

discussed In Chapter 2. Here, the installation of the RGD pharmacophore into a bicyclic scaffold is described, which led to a condensed and low-molecular-weight dimeric RGD ligand. Our data demonstrated that this design led to higher binding affinity to $\alpha_v\beta_3$ and $\alpha_5\beta_1$ as compared to a monomeric RGD analog, as well as to superior biological activity against cultured U-373-MG glioblastoma cell line. The formation of reversible covalent interactions is discussed in Chapter 3 and 4. In particular, we explored the use of the 2-hydroxybenzaldeyde (2HB) tag electrophilic warhead: 2HB is capable of interacting with the ε-amino group of lysine residues on a given protein, where the resulting Schiff base is stabilized by an internal hydrogen bond, remarkably increasing the imine stability towards hydrolysis. While Chapter 3 describes the development of a new synthetic methodology useful for the preparation of portable 2HB-polyehtylene glycol (PEG) modules, Chapter 4 is based on the design, synthesis and biological evaluation of 2HB-RGD ligands. In this work, we investigated the use of covalent docking as computational model to predict the affinity enhancement of reversible covalent ligands. The final Chapters of this Thesis include an overview of future work and perspectives (Chapter 5), the Experimental Section (Chapter 6) and the final appendix of analytical data (Chapter 7)

From monoclonal antibodies to high-affinity small ligands

1.1 Monoclonal antibodies

Antibodies are glycoprotein found in plasma and extracellular fluids that are secreted by a particular kind of immune system cells known as specialized B lymphocytes or plasma cells. Antibodies, known also as immunoglobulins (Ig) play a key role in adaptive immunity, as they are produced in response to various types of infections, within the process of recognition and neutralization of pathogens. From the structural point of view, antibodies, are composed of four polypeptidic chains. Two identical copies of both heavy chain (~55 kDa) and light chain (~25 kDa) are held together by disulfide bonds among cysteines and noncovalent interactions. The resulting structure is typically represented as a Y-shaped macromolecule (~150 kDa, Figure 2).

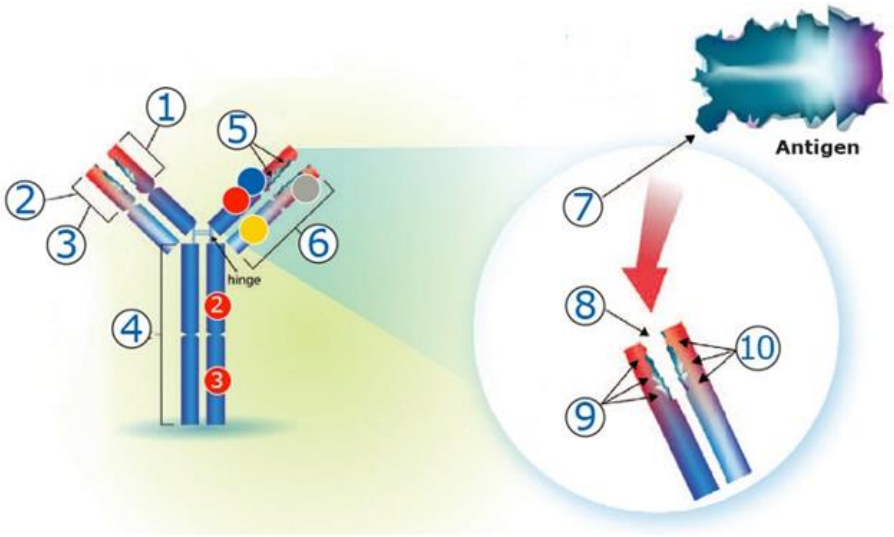


Figure 2 General structure of an antibody and interaction with a generic antigen. 1) Variable region of the heavy chain; 2) Variable fragment (Fv); 3) Variable region of the light chain; 4) Fragment crystallizable region (Fc); 5) Hypervariable region; 6) Antigen binding region (Fab); 7) Epitope; 8) Paratope; 9) Hypervariable region of the light chain; 10) Hypervariable region of the heavy chain. (Adapted from: I. Folz, M. Karow, S. Wasserman. Evolution and Emergence of Therapeutic Monoclonal Antibodies; What Cardiologists Need to Know. *Circulation* **2013**, *127*, 2222–2230)

Light chains consist of a variable N terminal domain of approximately 110 amino acids and a constant domain of equivalent length. In a similar way, the heavy chains is composed by different domains of approximately 110 amino acids: the variable region is located at the N terminal domain, while the constant region is composed by at least three domains, depending on the immunoglobulin subtype. As a result, the N termini of both the light and heavy chains are paired at the extremity of the variable, antigen-binding domains, while the carboxy terminal regions of the two heavy chains join together to form the extremity of the Fc domain. The Fab of the antibody molecule containing the antigen-binding domains and the Fc are connected by a region rich in proline, threonine, and serine, called hinge.

This region gives flexibility to the paratope, providing the antibody the ability to interact with a variety of antigen presentations.

This structure design of Antibodies is functional to carry out two essential tasks:

1. Antigen recognition: antibodies bind to a specific epitope (7, Figure 2) on a given antigen with the arms of the Y structure. Each arm or monovalent antibody fragment (Fab) domain contains a binding site called paratope (8, Figure 2). The high sequence variability of three protein loops of the Fab region (called “complementarity-determining regions”, CDRs, each from five to ten amino acids long) confers the high affinity and specificity of the Fab for the specific antigen;
2. Recruitment of protein or cellular effectors: the fragment crystallizable region (4, Figure 2), shortly called Fc, can interact with different proteins expressed on the surface of immune cells (e.g. to activate cytotoxic cell killing in natural killer cells or phagocytosis in macrophages/dendritic cells), or in the intracellular compartment of endothelial cells (i.e. binding to neonatal Fcγ receptor, which promotes antibody recycling and extend its circulatory half-life) or in the serum (i.e. to initiate complement-dependent cytotoxicity)

Following the recognition of an antigen by an appropriate B cell clone and in the presence of other activating signals (e.g. cytokines/chemokines), B lymphocytes proliferate and evolve into memory B cells as well as into plasma cell clones. The so-obtained clones are able to secrete their B cell receptor (which is transformed into antibody molecules) that diffuses in the body to immobilize the antigen and recruit other cellular/protein effectors through Fc clustering. At the same time, memory B lymphocytes survive at low concentrations in the host until subsequent activation in presence of their specific antigen: these lymphocytes give the host an immunological memory and, in particular, the resulting escalation in antibody response when subjected to a second cycle of infection.[1] Since most antigen structures are highly complex, antibodies present multiple epitopes that are recognized by a large number of different lymphocytes. In this case, the immune response is named “polyclonal” indicating the heterogenous population of lymphocytes (each one encoding for a single antibody) activated to proliferate and differentiate into plasma cells. In contrast, “monoclonal antibodies” (mAbs) are produced by a single and isolated B lymphocyte clone, and are thus able to recognize a single epitope on a specific antigen. mAbs were firstly isolated from sera of patients with multiple myeloma in which the separation and clonal expansion of malignant plasma cells led to the isolation of single clones of identical antibodies. In the1970s, Köhler and Milstein developed a technique that is able to generate monoclonal antibodies of a desired specificity, for which they were awarded the Nobel prize. The technique is known as hybridoma technology (Figure 3)[2] and it is based on the fusion between an immortal myeloma cell line and spleen-derived B lymphocytes from antigen-exposed mammals. In this protocol, mice are first immunized with the antigen of interest, following splenocytes isolation and fusion with immortal myeloma cells, typically performed through

application of an electric field (electrofusion). Alternatively, the fusion among the B-cells and myelomas can be performed using chemical protocols, most often using polyethylene glycol.

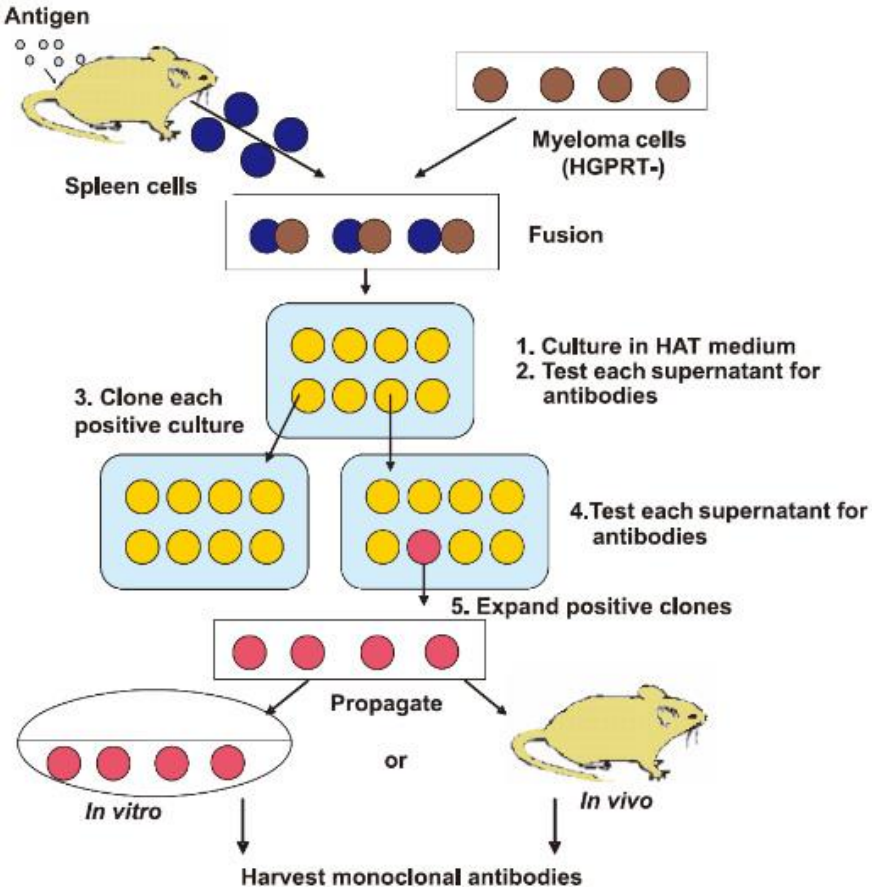


Figure 3 Flow chart for mAbs generation using the hybridoma technology. It involves the production of monoclonal antibodies specific to an antigen of interest. The somatic fusion of B lymphocytes of the spleen with immortal myeloma cells gives rise to a hybridoma cell line that can be perpetually propagated to produce clonally identical antibodies, as these hybridoma cells inherit the indefinite growth properties of myeloma cells and the antibody secretion capabilities of B-lymphocytes. Monoclonal antibodies produced by a unique hybridoma cell line are homogeneous and recognize a single epitope of an antigen. (Adapted from: https://cn.sinobiological.com/resource/antibody-technical/hybridoma-technology)

To warrant the success of the fusion process, the myeloma cells used for this protocol lack the hypoxanthine-guanine phosphoribosyltransferase (HGPRT) gene, making them sensitive to the hypoxanthine-aminopterin-thymidine medium (HAT medium). When fused cells are incubated in the HAT medium, aminopterin blocks the pathway that allows for nucleotide synthesis, a vital pathway for the cell lines in which HGPRT is involved. As a consequence, while unfused myeloma cells undergo apoptosis due to the lack of HGPRT, only the B cell-myeloma hybrids survive, as the HGPRT gene is given by the B cells. Thus, the resulting cell line is able to produce antibodies and is immortal (i.e. it can be expanded at large extent, which is a property of myeloma and cancer cells in general). The resulting hybridoma population is then diluted, in order to isolate cell cultures derived from a single hybridoma clone (i.e. monoclonal expansion). Since the antibodies are produced by the same B cell, they will be directed towards the same epitope, and thus they are monoclonal antibodies. The advent of hybridoma technology was the starting point for the arising success that mAbs had and still have in current medicine. Antibody phage display (APD) is another technology widely used to produce mAbs. APD protocol (Figure 4) is based on genetic modification of

bacteriophages, such as viruses that infect bacteria, in which each gene corresponds to a well-defined mAb. Later on, the antigen is exposed to a phage library and, among the possible mAbs, there is the selection of one mAb. Finally, the selected phage is isolated and propagated through infection in *E. Coli.* This technique allows *in vitro* selection of mAbs of virtually any specificity, facilitating recombinant production of reagents for use in research and clinical diagnostics, as well as for pharmaceuticals for therapeutic use in humans.[3]

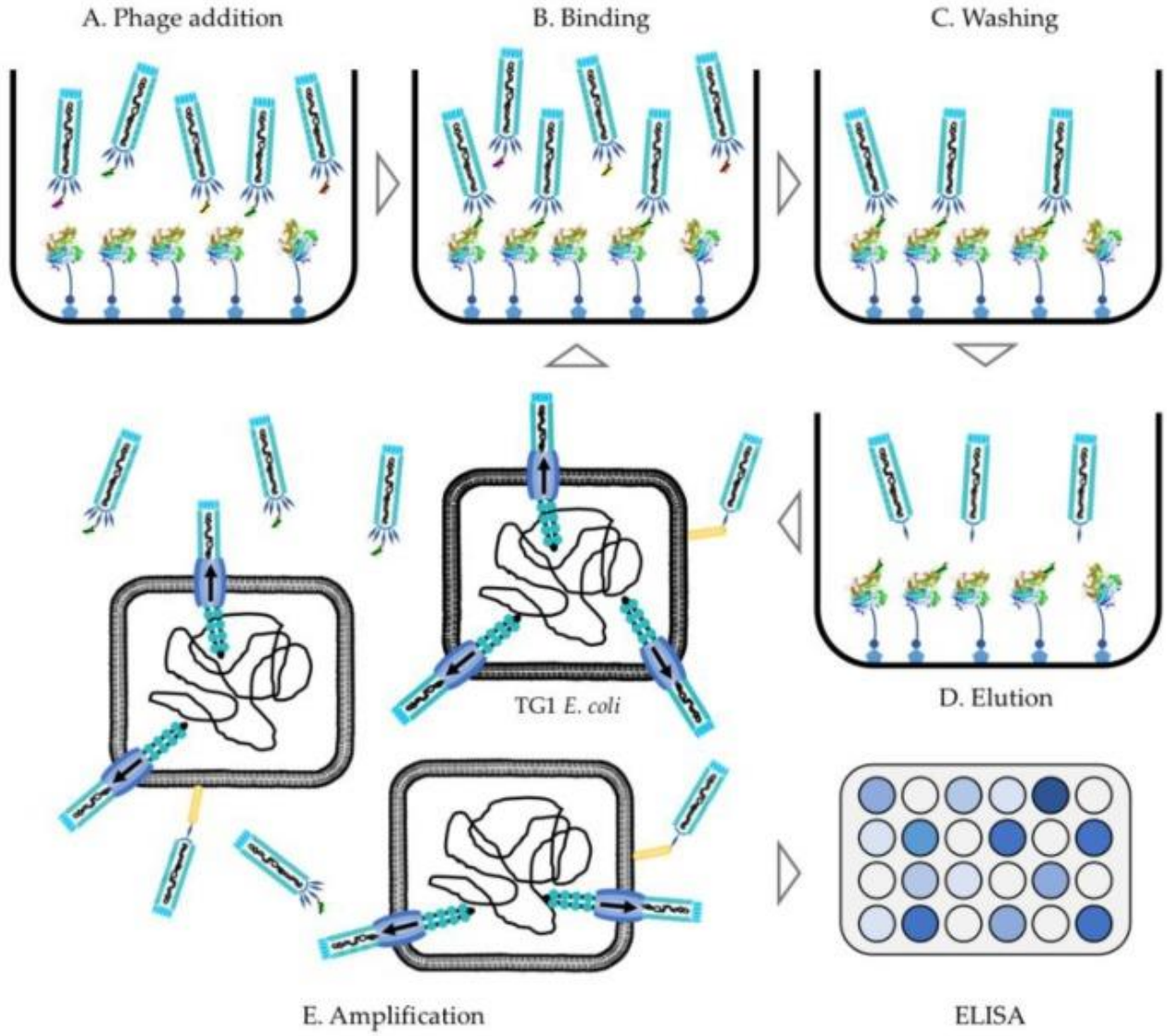


Figure 4 Flow-chart of antibody-phage display (APD) protocol for mAb selection. (Adapted from reference 3)

The applications in which mAbs are involved are various and different among each other. The ability of mAbs to bind a specific epitope of an antigen is used for analytical purposes. The quantitative and qualitative techniques in which mAbs are involved are immunoprecipitation[4], immunoblotting and immunodetection,[5] enzyme-linked immunosorbent assay (ELISA),[6] microarray platforms for high throughput proteomic analysis,[7] X-ray crystallography,[8] fluorescence-activated cell scanning (FACS),[9] immunofluorescence and immunohistochemistry.[10] The ability to bind with a high affinity the related antigen makes mAbs useful also for purification and enrichment protocols for isolated antigen and a cell population expressing a particular antigen, using a fluorescent-activated cell sorter

(FACSort).[9] Notably, the development of mAbs proved particularly relevant in the pharmaceutical industry, where they are used as therapeutic agents for different indications. The implementation of the hybridoma technology to an industrial scale, performed by Ortho Biotech, led to the development of the first murine mAb, Orthoclone OKT3® (Muromonab), that is used as immunosuppressor in patients subjected to organ transplants and the antigen is the cluster of differentiation 3 (CD3), expressed on T cells.[11] Since the approval of Muromonab in 1986, the use of mAbs as drugs grew exponentially. Recently, the 100th mAb-based product has been approved by FDA[12] and 5 out of 10 best-sold drugs in 2021 are mAbs, with Humira® and Keytruda®, ranking respectively in the first and the second position of the chart (Table 1).[13]

Table 1 Top 10 of the most-sold drugs in 2021.

Rank	**Product**	**Company**	**Pharmacological class**	**2021 worldwide sales (US$ millions)**
1	Humira	AbbVie/Eisai	Anti-TNF mAb	19,963
2	Keytruda	Merck & Co.	Anti-PD1 mAb	16,825
3	Revlimid	Bristol-Myers Squibb/BeiGene	Immunomodulator	12,710
4	Eliquis	Bristol-Myers Squibb/Pfizer	Factor Xa inhibitor	10,546
5	Eylea	Regeneron Pharmaceuticals/ Bayer/Santen Pharmaceutical	Anti-VEGF mAb	8,872
6	Opdivo	Bristol-Myers Squibb/Ono Pharmaceutical	Anti-PD1 mAb	8,759
7	Stelara	Johnson & Johnson/Mitsubishi Chemicals	Anti-IL-12/IL-23 mAb	8,445
8	Biktarvy	Gilead Sciences	HIV INSTI/NRTI/ NtRTI	8,418
=9	Imbruvica	AbbVie/Johnson & Johnson	BTK inhibitor	7,607
=9	Xarelto	Bayer/Johnson & Johnson	Factor Xa inhibitor	7,605

1.2 Monoclonal antibodies for cancer therapy

After the development and the consolidation of the hybridoma and phage display technologies, mAbs started to become very attractive as a therapeutic agent for pathologies characterized by the expression of abnormal protein markers, whose structure and pathological activity is not easily blocked by small molecule therapeutics. Among these malignancies, the large variety of cancer types undoubtedly represent the typical applicative field of therapeutic mAbs.

In oncology, common chemotherapeutic agents used for cancer therapy are cytotoxic agents that interfere with the vital mechanism of the cell. Since cancer cells growth is much faster compared to heathy ones, cytotoxic agents should impact preferentially on the development of tumors, sparing healthy cells. Unfortunately, this class of small molecule drugs is characterized by severe side effects, due to the common presence of the vital pathways both in tumor and heathy tissues. This reduces the therapeutic window of the drug, defined as the minimum amount of molecule that generates a therapeutic effect compared to the minimum amount of drug that induces toxicity. To solve this problem, the administration of a drug that is selective for a tumor biological target rather than a healthy tissue would be beneficial for patients. In this context, it became crucial the discovery and development of drugs that act selectively against specific enzymes or receptors that are overexpressed in the tumor environment. On these tumor targets, the activity and selectivity of mAbs would be the solution to the toxicity issue and therefore it would improve the efficacy of the therapy. For this reason, form the 1980s, many mAbs entered the clinical studies. In particular, the first trial concerned patients with relapsed lymphoma, but it turned out to be clinically ineffective.[14] In parallel trials on different patients with different indications showed only a short-lived response to the cure.[15] The fails into the clinical trials of murine mAbs could be ascribed to two main drawbacks. The fact that mAbs were isolated from mice splenocytes was found to generate the so-called human anti-mouse antibody (HAMA) in the patients. The HAMA-mAb complex not only avoids the binding of the mAb to the related antigen, but it also gives rise to allergic reactions up to anaphylactic shock. The presence of a problem related mainly to a non-human portion of the mAb led Winter and coworkers to the substitution of the constant domain of the murine mAb with a human analogue thanks to gene editing.[16] The resulting mAb possess a murine portion and a human portion, thus leading to the so-called chimeric mAb. As a result of the modification of roughly two thirds of murine mAb with the human equivalent, the immunogenicity of the chimeric mAb was lowered and therefore the bioavailability of the molecule increased. In addition to the use of genetically-modified mice, where human lymphocytes are instrumental to produce fully human hybridomas and antibodies (the HuMab Mouse^{TM}), the development of chimeric, humanized and fully human mAbs was also accessible by the advent of the phage display technology, which overcame the need for antibody isolation from a animal host (such as in the hybridoma technology). In 2018, Winter and Smith were awarded the Noble Prize for their work on the phage display technology, which is nowadays a standard for mAb development. The mechanisms by which mAbs are able to kill tumor cells are depicted in Figure 4.

The interaction of the mAb with the related tumor-overexpressed antigen (e.g. a receptor or a natural ligand present in the tumor environment), could lead to the blockade or an agonistic effect on the cancer cell thanks to the direct interaction between the mAb and the antigen or the delivery of a cytotoxic agent covalently bound to the mAb (the so-called antibody-drug conjugates, ADC).

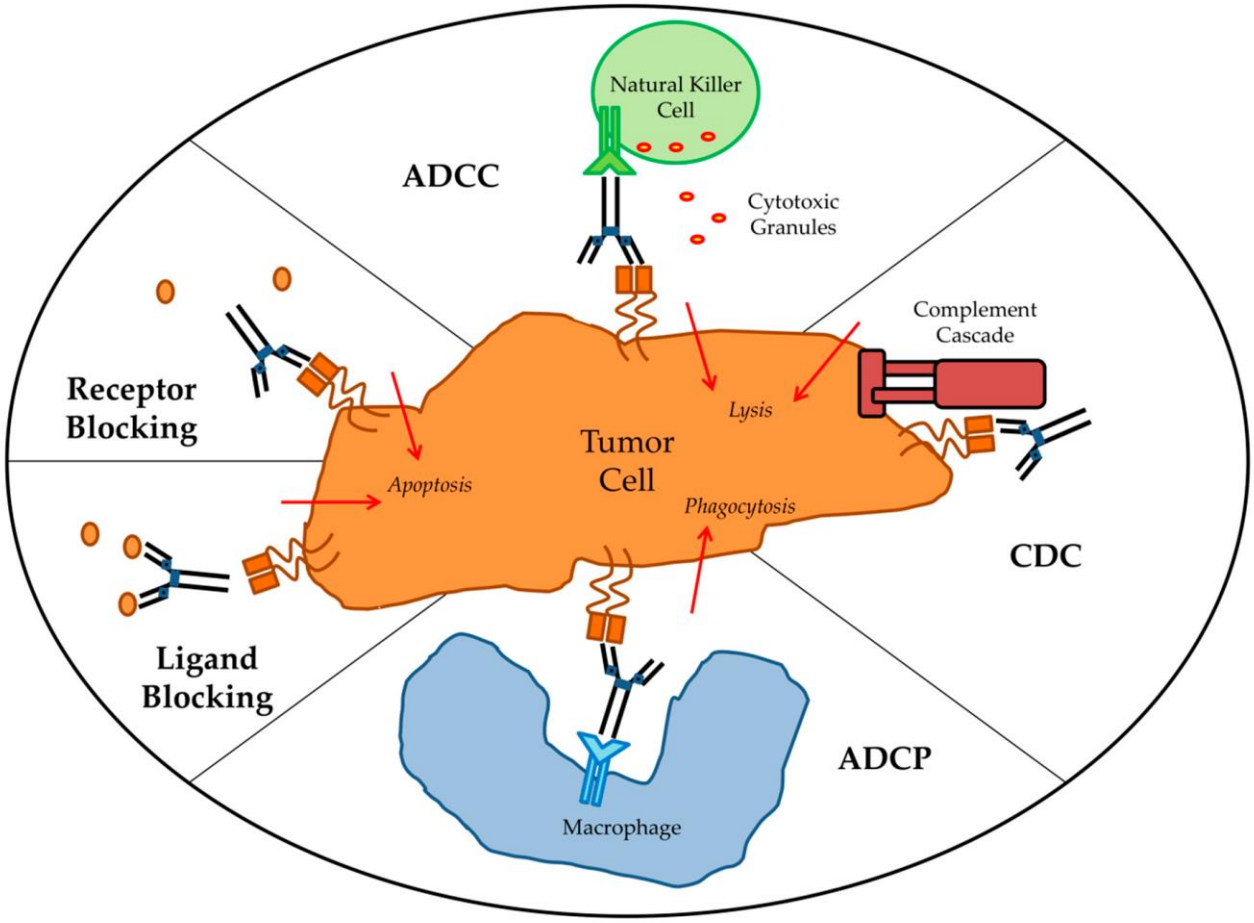


Figure 4 Mechanisms of action of mAb on a tumor cell. Starting from the top and going clockwise, Antibody-dependent cellular cytotoxicity (ADCC), complement dependent cytotoxicity (CDC), antibody-dependent cellular phagocytosis (ADCP), ligand and receptor blockade are highlighted. (Adapted from: D. Zahavi, L. Weiner. Monoclonal Antibodies in Cancer Therapy. *Antibodies* **2020**, *9*, 34)

The other mechanisms rely on an indirect pathway and, in particular, on the activation of the host immune system. These mechanisms are the complement dependent cytotoxicity (CDC), the antibody dependent cellular cytotoxicity (ADCC) and the antibody dependent cellular phagocytosis (ADCP). The CDC is based on the formation of a complex between the protein C1q and the antibody. The mAb-C1q complex trigger the recruitment and deposition on the cell surface of a membrane attack complex (MAC).[17] MAC induces the formation of pores on the cell surface that induce the cell lysis and subsequent death.[18] There are many mAbs able to recruit the complement system. For example, Rituximab, a mAb used for non-Hodgkin's lymphoma, chronic lymphocytic leukemia and other immune diseases, proved to induce tumor cell death by CDC.[19] The CDC recruitment could be enhanced by proper engineering of the mAb structure. For instance, ofatumumab, the anti-CD20 mAb, bears a fully human structure that mediates amplified CDC. It demonstrated greater efficacy

compared to Rituximab in a clinical trial of chronic lymphocytic leukemia (CLL) patients.[20] ADCP is the mechanism that involves the FcγRs on the surface of macrophages. The interaction between the mAb bound to the tumor cell and the FcγRs induce phagocytosis. As a consequence, internalization and degradation of the target cell happens through phagosome acidification. There are not many studies on ADCP but it proved to be a crucial mechanism for mAb therapeutic effect on tumors.[21] ADCC is similar to ADCP but it works with a non-phagocytic mechanism. Discovered by Möller in 1965, ADCP is based on the involvement of effector cells that acts as the actual tumor cell killer.[22] After the interaction between the tumor antigen and the mAb, the antibody recruits the effector cells thanks to the binding between the Fc portion of the antibody and the Fc receptor (FcR) on the effector cell surface. Natural killer (NK) cells are one of the main effectors involved into clinically relevant mAbs[23] but also monocytes, macrophages, neutrophils, eosinophils, and dendritic cells can be recruited.[24] Tumor cell death in the ADCC mechanism is induced by cytotoxic granule release, Fas signaling, and initiation of reactive oxygen species (ROS).[25] The ability of mAbs to induce ADCC is recognized as a crucial factor for mAb therapy success. Research efforts are directed to the design of mAbs with enhanced ADDC induction capability. The enhancement is performed by modification of the Fc portion of the mAb to increase their binding affinity to the activating FcIIIA via site-directed mutagenesis, changing Fc domain glycosylation, and/or removing Fc domain fucosylation.[26] In particular, afucosylated mAbs turned out to be promising in clinical trials.[27] The previously presented mechanisms are the physiological bases for the success of mAbs into the clinical phases. Nowadays, 46 mAbs are approved by FDA for the treatment of various cancers.[28]

Despite the FDA approval of many mAbs for cancer therapy, cases of clinical resistance to the therapies have been reported.[29] The resistance to the therapy can be classified in two ways. The first one is the innate (primary) resistance, where a mutation of the tumor antigen is present before the therapeutic treatment. The second one is the acquired (secondary) resistance, where an inductive mutation of the tumor antigen is the consequence of the exposition of the tumor tissue to the therapeutic agent. In this perspective, the entity of the expression of the proper tumor antigen in the tumor environment is essential for the positive outcome of the mAb therapy. For instance, CD20-expressing tumor cell lines that are continuously exposed to rituximab downgrade the expression of the target at both transcriptional and protein expression level. The beneficial effect of the drug is then lowered by these mutations.[30] Moreover, when mAbs target key receptor for a given signaling pathway, the direct mutation of the downstream effectors may be deleterious for the mAb therapy. For example, it was shown that the main reason of cetuximab limited efficacy against colorectal cancer is the genomic alteration of the EGFR downstream effectors like KRAS, NRAS, BRAF, and PIK3CA. The alteration in this pathway (of both primary and secondary type), makes ineffective the EGFR inhibition mechanism of cetuximab.[31]

Finally, the use of antibodies as a therapeutic agents not only possesses some limitations related to pharmacodynamic issues, highlighted in the previous paragraph, but also to a poor pharmacokinetic

profile, which are mainly related to their large size.[32] For instance, large macromolecules that are present in the blood vessels typically extravasate slowly, retarding the drug accumulation into the tumor site.[33] Moreover, solid tumors are characterized by the so-called “antigen barrier”, which consists in the antigen presentation by cells in perivascular areas (i.e. the surrounding of blood vessels): upon extravasation, mAbs bind tightly to these first layer of cells, and the resulting high density of large antibodies keeps circulating mAbs from diffusing within the tumor tissue and hit all cancer cells.[34] Finally, the long circulatory half-life of mAbs (which results from both their large size and the recycling by neonatal Fcγ receptor in endothelial cells) may be also detrimental in some cases, enhancing the probability of immune reactions or chemical degradation of antibody conjugates.[35]

Reducing the size of the mAb (for example using scFv, Fab or diabody fragments) has been proposed as valuable strategy to improve the rate of extravasation and tissue diffusion, while maintaining the mAb specificity.[36] However, the use of engineered antibodies different from the “canonical” IgG format can higher the immunogenicity risk. Finally, the industrial production of pharmaceutical-grade mAbs and biologics in general is highly expensive, impacting on the costs of treatment and limiting the benefits to the global population.

1.3 High affinity small molecules

1.3.1 Small molecules versus monoclonal antibodies: a case study

The above-described advances in the mAb technology explain the great impact of these classes of pharmaceutical agents in modern medicine. However, the intrinsic limitations of mAbs have prompted current pharmaceutical research to develop small molecule compounds with antibody-like therapeutic effects. In 2018, Neri and coworkers[37] performed an *in vivo* comparative analysis of the tumor-targeting properties of a mAb and a small molecule ligand, both directed against renal cell carcinoma cells expressing the carbonic anhydrase IX (CAIX) antigen.

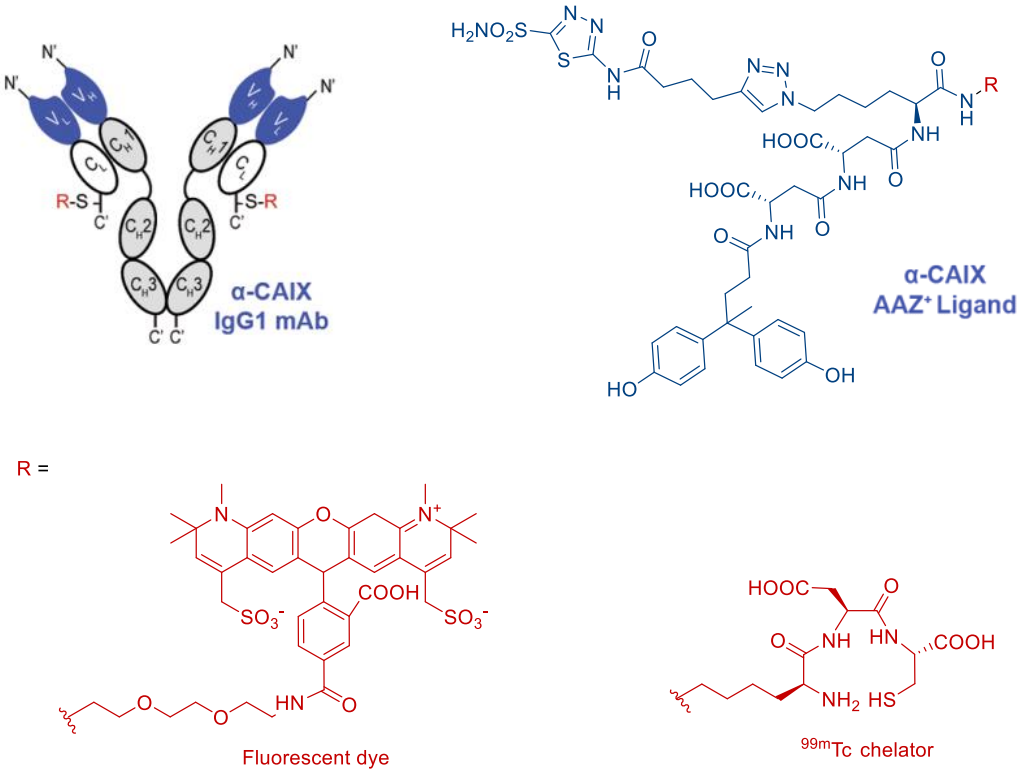


Figure 5 Structure of the anti-CAIX IgG1 mAb and acetazolamide (AAZ) ligand used for the tumor targeting experiment. For the biodistribution experiments, the mAb is radiolabeled with ^{125}I and the AAZ is endowed with a ^{99m}Tc atom enclosed into a proper chelator. For the microscopic imaging experiments, the AAZ is conjugated to a fluorescent dye (Alexa Fluor594). Adapted from reference 37.

The group evaluated the pharmacokinetic performance of the two different carriers (mAb+ and SM+) in parallel with suitable control compounds, such as a generic IgG mAb and a CAIX-inactive small molecule. Both carriers were functionalized with a radioisotope (^{125}I for the antibody and ^{99m}Tc for the small molecule, Figure 5) and injected in mice bearing subcutaneous SKRC-52 tumors (CAIX-expressing renal cell carcinoma cells) in order to monitor and quantify the uptake of these therapeutics in tumors and in blood (Figure 6).

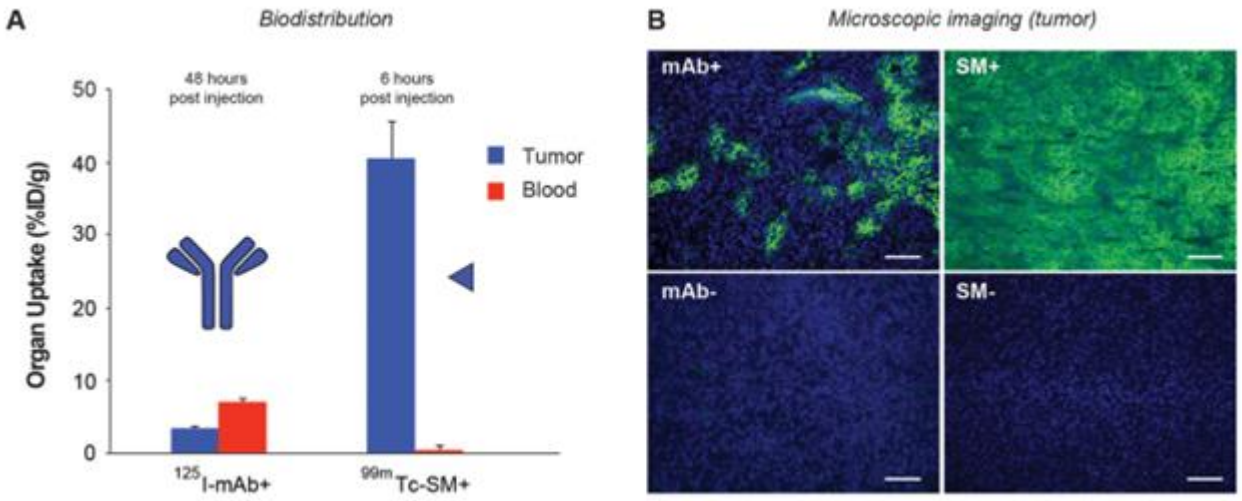


Figure 6 A) Quantitative evaluation of the tumor uptake of radiolabeled anti-CAIX mAb (mAb+) and radiolabeled acetazolamide derivative (SM+); B) Microscopic distribution of mAb+, SM+ and the related radiolabeled negative controls (mAb- and SM-) in SKRC-52 tumor. Adapted from reference 37.

The biodistribution experiment showed that the tumor-to-blood ratio for mAb+ was unfavorable even after 48h, indicating the tendency of mAbs to extravasate slowly. On the other hand, the SM+ showed rapid accumulation in the tumor mass (tumor-to-blood ratio ~ 100:1 6h post injection), highlighting a superior capability to extravasate and accumulate in the tumor tissue compared to mAb+ (Figure 6A). In a similar experiment, the mAb and small molecule products were conjugated to fluorescent dyes (Figure 5), injected in tumor-bearing mice and their distribution in tumor and healthy organs was evaluated by fluorescence microscopy. While the CAIX-targeting SM+ showed a homogenous diffusion in the tumor mass 1h after the injection (Figure 6B), CAIX-targeting mAb+ showed an irregular uptake in the tumor 24 h after the injection, which indicated the preferential accumulation of the IgG within the perivascular areas and the hindered mAb diffusion by the above-mentioned “antigen barrier”. Finally, control compounds mAb- and SM- did not show any accumulation into the tumor site. With this work, Neri and coworkers clearly demonstrated that, due to their superior pharmacokinetic properties, small molecules may be preferred to antibody products, especially for the treatment of solid tumors. However, both the small molecule and the antibody used in this study possessed a very high binding affinity to the cognate protein target. While high-affinity antibodies can be raised against virtually all types of protein target, the identification of high-affinity small ligands is not trivial, stimulating the necessity to develop novel and general strategies to facilitate the discovery and the design of high-affinity small ligands.

1.3.2 High affinity small ligands by fragment-based drug discovery (FBDD)

From the early 2000’s, high-throughput screening (HTS) technique has been established as the gold standard for the discovery of lead compounds in pharmaceutical companies.[38] HTS is based on the screening of large chemical libraries (millions of small molecules) for the activity on a defined biological target with the help of automation, miniaturized assays and large-scale data analysis.[39] This approach led to the discovery of new leads, especially against established classes of targets.

On the other hand, in the case of poorly-established or "undruggable"[40] targets, HTS campaigns often proved negligible results.[41] There is also a problem of the coverage of the chemical space because only a very small part (millions of molecules) of the whole chemical space (10^{63} possible organic molecules)[42] is covered by HTS. Moreover, the structural features of the protein target are often predictive of the ease to develop specific small molecule drugs. The so-called "Druggability" of a protein is defined as the probability to develop a small molecule whose interaction with the protein itself generates a biological effect.[43] "Druggable" proteins typically possess structural features (e.g. hydrophobic pockets) that favors the interaction with a hydrophobic compound, thus increasing the probability to identify potential small molecule ligands. On the contrary, flat and featureless proteins represent preferential targets for mAbs, while being typically considered "undruggable" by small molecules. Among the undruggable targets, protein-protein interactions (PPIs) are one of the most challenging. Many cancer-associated proteins are involved in PPIs[44] and the discovery of new ligands for these targets would be important for the developments of new cures for cancer.

In this scenario, Fragment-based drug discovery (FBDD) consists in an alternative approach to HTS and it proved a promising strategy for the design of new ligands against undruggable targets. FBDD is based on the screening (biophysical or computational) of little molecular fragments that possess a relatively simple chemical structure, a molecular weight lower than 300 Da.[45] While individual fragments show low binding affinity to the target of interest, the whole drug structure results from the assembly of suitable fragments. Even if complementary to HTS campaigns, FBDD allows to decrease the experimental costs and gives an alternative to the design of new drug-like compounds.[46] The impact of this technique on the development of new active principles is significant. In fact, vemurafenib, an inhibitor of oncogenic B-RAF kinase, has been discovered thanks to a FBDD campaign and then it has been approved by FDA for the treatment of late-stage melanoma.[47] Moreover, FBDD has been used to discover new molecules able to interact with targets involved in PPIs.[48] In these contexts, FBDD campaigns against PPIs can be run either to identify allosteric modulators with the aim to change the nature of the PPI,[49] or to develop small molecules capable of disrupting[50] or stabilizing[51] the PPI in order to affect the related biochemical pathway. Once identified the protein target, the usual flow chart of a FBDD campaign in highlighted below (Figure 7).

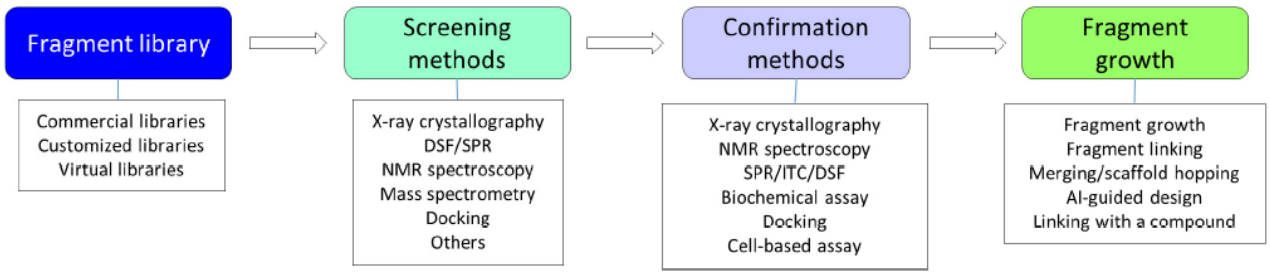


Figure 7 Usual flow chart for a fragment-based drug discovery campaign. Some possible techniques involved in the campaign are listed below each step.

For the fragment library, fragments contained in the library are virtually unlimited and, as anticipated, fragments should be small enough and possess simple structures. The screening methods are

biophysical or computational techniques that are able to quantify the interaction of a fragment to the biological target. Since the binding affinity of the fragment with the protein is usually low, the technique screening method is expected to be very sensitive. In fact, differential scanning fluorescence (DSF), isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), NMR spectroscopy and X-ray crystallography are used for the screening purpose.

The identification of a promising fragment is often followed by hit-to-lead modifications, with the aim to enhance the binding affinity of the resulting molecule. Three methods are mainly used in this step: fragment growing, fragment hopping and fragment linking. The "fragment growing" is the most used technique, and it is based on the addition of different chemical groups on the fragment compound (hit) in order to enhance the potency of the final molecule.[52] Four molecules have been developed with the fragment growing method and then approved by FDA for different indications.[53] Fragment hopping is based on the combination of different chemical features of fragments that are able to bind the target protein. The aim is to transform the non-drug like fragments into a drug-like molecule.[54] In this method, the binding mode of the fragments is crucial for the proper outcome of the optimization. For example, X-ray crystallography or NMR spectroscopy data are vital to gain insights into the binding mode of each fragment. Finally, fragment linking is considered the most powerful fragment optimization technique. Fragment linking is based on the identification of two fragments that are close in space and the linkage of the two fragments. The nature of the linker is important because it should not disrupt the interaction of the two fragments with the protein. Of course, this method can be applied to proteins that have two proximal binding pockets.[55] Whenever it could be applied, the fragment linking protocol allows the dramatic enhancement of the binding affinity.[56] Whenever the lead structure has been defined, the co-crystal between the hit fragment and the protein is a good tool to confirm the binding ability and the features of the ligand-protein interaction.[57] Moreover, structure-activity relationship (SAR) studies of the newly designed molecules is another tool to confirm the binding capability of the lead compound and find the pharmacophoric portions that are fundamental for the formation of the related ligand-protein complex.

1.3.3. High-affinity small ligands by DNA-encoded chemical libraries

The development of DNA-encoded chemical libraries (DELs) is a drug discovery technique alternative to traditional HTS campaigns. In particular, the advantage of DELs over HTS is the capability to screen billions of compounds in a single experiment.[58] The underlying idea of DELs originates from Brenner and Lerner, who postulated a technology reminiscent of phage display, where a small molecule compound (phenotype) is individually associated to a specific DNA sequence (genotype). The latter acts as identification barcode for the specific small ligand, which can be thus identified from a highly heterogeneous mixture through PCR and DNA sequencing.[59] While the early development of DELs consisted in the library assembly on beads, further evolutions of the technique led to the direct coupling of the synthetic compound to the DNA tag, making DEL

an attractive platform for the early-phase drug discovery.[60] A DEL campaign flowchart is represented below (Figure 8).

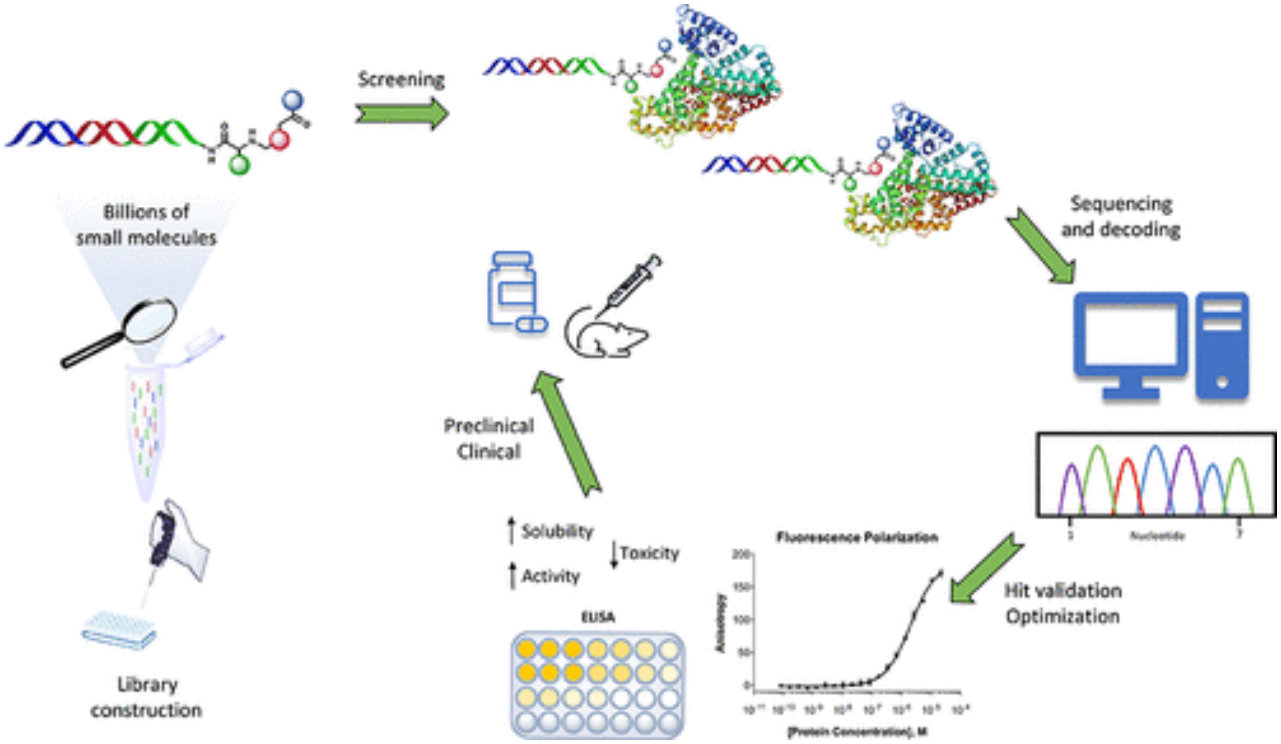


Figure 8 Typical flowchart for a DNA-encoded chemical library (DEL) campaign. DELs are made of synthetic small molecules, individually coupled to unique DNA tags used as amplifiable identification barcodes. The library is exposed to the target protein that will bind a single (or more) ligand-protein complex. The complex is isolated and, thanks to the DNA tag, the small molecule is properly identified and tested in a separate binding assay. If confirmed, the lead compound is tested in preclinical experiments. (Adapted from: A. Gironda-Martínez, E. J. Donckele, F. Samain, D. Neri. DNA-Encoded Chemical Libraries: A Comprehensive Review with Succesful Stories and Future Challenges. *ACS Pharmacol. Transl. Sci.* **2021**, *4*, 1265-1279)

In the DEL technology, the first step consists of the building of a chemical library. Various library design has been proposed. For instance, in “single-pharmacophore” libraries the whole ligand is directly coupled to a single DNA tag, which can be displayed as a single strand or hybridized to a complementary DNA strand. On the contrary, dual-pharmacophore libraries display two different chemical entities on two complementary DNA strands. This approach allows to perform the detection of two binding fragments of non-overlapping binding pockets.

During the DEL synthesis, different sets of building blocks are progressively assembled by split-and-pool protocols. Typically, each chemical modification of the pharmacophore is followed by the addition of a DNA tag (encoding) containing a specific sequence, univocally associated to the inserted building block.

Following the DEL synthesis, the library is screened against the target of interest, typically immobilized on solid support. This incubation allows the isolation of ligand-protein complexes, while unbound fractions of the library are discarded. The protein could be immobilized to magnetic beads, which facilitates the washing steps and the isolation of tagged small molecules that interact with the protein. The latter fraction is isolated through heating elution, and the building blocks that compose the structure of suitable ligands are identified. The use of high-throughput DNA sequencing

technologies gave a huge contribution to the DELs-based campaigns, allowing the screening of libraries that presents millions of compounds.[61] If the DNA part of the library is composed by two main blocks, the data can be depicted as a 3D graph, where the two code are plotted on the x and y axes, whereas the counts (enrichment) of each code is displayed on the z axis.[62] Once identified the hit, structural modifications and ligand optimization can be performed by chemical synthesis in solution, to refine the ligand structure and make it amenable to future development.[63] Up to now, three compounds obtained from a DEL screening campaign are into the clinical trials.[64]

1.3.4. High-affinity small molecules by affinity selection-mass spectrometry

During the last decade, affinity selection-mass spectrometry (AS-MS) has become an important technological platform for drug discovery. In AS-MS, several compounds are screened against the same biomolecular target and binding compounds are detected and identified by their specific molecular weight, which allows the use of native and unmodified ligands and protein targets. The general workflow of a AS-MS screening campaign is depicted below (Figure 9).

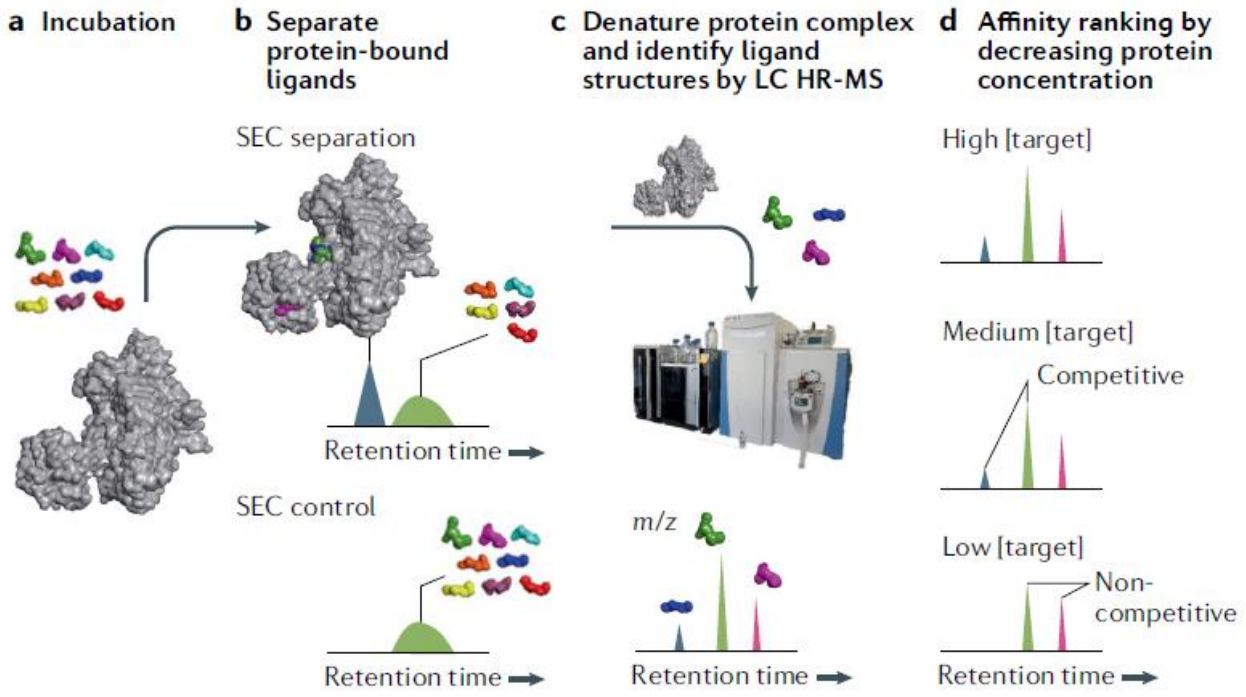


Figure 9 General workflow for an affinity selection-mas spectrometry (AS-MS) screening campaign. It starts with the incubation of a mixture of small molecules with the ligand. It follows the size-exclusion chromatography step in order to remove all the small molecules that did not bind to the protein. The ligand-protein complex is then denaturized, and the small molecules are identified and characterized by LC-HRMS. The confirmation of the hit(s) is performed running a SEC and then LC-HRMS analysis of the small molecule mixture. Each ligand is then ranked by their capability to bind the target in subsequent experiments in which the protein-to-ligand ratio is decreased. (Adapted from: R. Prudent, D. A. Annis, P. J. Dandliker , J. Y. Ortholand, D. Roche. Exploring new targets and chemical space with affinity selection- mass spectrometry. *Nat. Rev. Chem.* **2021**, *5*, 62–71)

Here, the first step consists in the incubation of a mixture of small molecules with the biomolecular target (protein or nucleic acid). At this stage, the target is usually present in higher molar concentrations as compared to the individual small molecules, thus minimizing the ligand competition and enabling the detection of ligands with also low binding affinity for the target. In these conditions, the ligand-target complex is formed, while unbound fractions remain in solution. Following the

incubation step, the complex isolation is performed by either a membrane filtration or washing steps or, more often, by size-exclusion chromatography (SEC). The goal of this step is to separate the unbound small ligands from the ligand-target complex before the denaturing conditions of LC and HRMS analysis. Two SEC-MS methods are typically used: SpeedScreen, developed by Novartis,[65] and Automated Ligand Identification System (ALIS), developed by NeoGenesis pharmaceuticals[66] and lately acquired by Merck. In all these methods, the ligand-target complexes are split by a denaturing LC protocol and then analyzed by high-resolution MS in order to reveal the identity of each hit compound based on their exact mass. Electrospray ion sources (ESIs) are mainly used as MS detector, but some protocols use matrix-assisted laser desorption/ionization (MALDI).[67] The ion separation is performed with a time-of-flight (TOF) or Orbitrap, which both allow a highly accurate analysis and a high resolution (m/z error <5 ppm). Similar to other HTS techniques, the probability to find a binder with AS-MS is mainly related to the quality of the library. AS-MS gives many advantages compared to the other HTS platforms.[68] First of all, thanks to the increasingly high sensitivity of modern MS instruments, very low quantities of both ligands and biomolecular targets are needed. Moreover, no fluorescence, radio-labeling or encoding procedures are needed, which opens for a very general use of this technique. For instance, AS-MS proved a valid method to identify ligands against “undruggable” targets,[69] such as PPIs and protein-nucleic acid interactions. As compared to HTS, AS-MS is considered more cost-effective, since the whole library can be screened in a single step against a target, which is also a property of DEL technology.

Since the AS-MS is still in its infancy, no compounds derived from this platform have entered clinical trials yet, but recent literature reports the use of this technique for the successful identification of many binders against kinases,[70] receptors,[71] GPCRs,[72] RNA[73] and G-quadruplex.[74]

1.4 Integrins as tumor targets

Integrins are a family of receptors that mediates cell adhesion processes. These receptors are transmembrane heterodimeric glycoproteins endowed with an α and a β subunit (Figure 10A), and at least 18 α and eight β subunits are known in humans (Figure 10B).[75] The endogenous ligands of integrin heterodimers are mainly extracellular matrix (ECM) proteins, cell-surface proteins and some soluble ligands.[76]

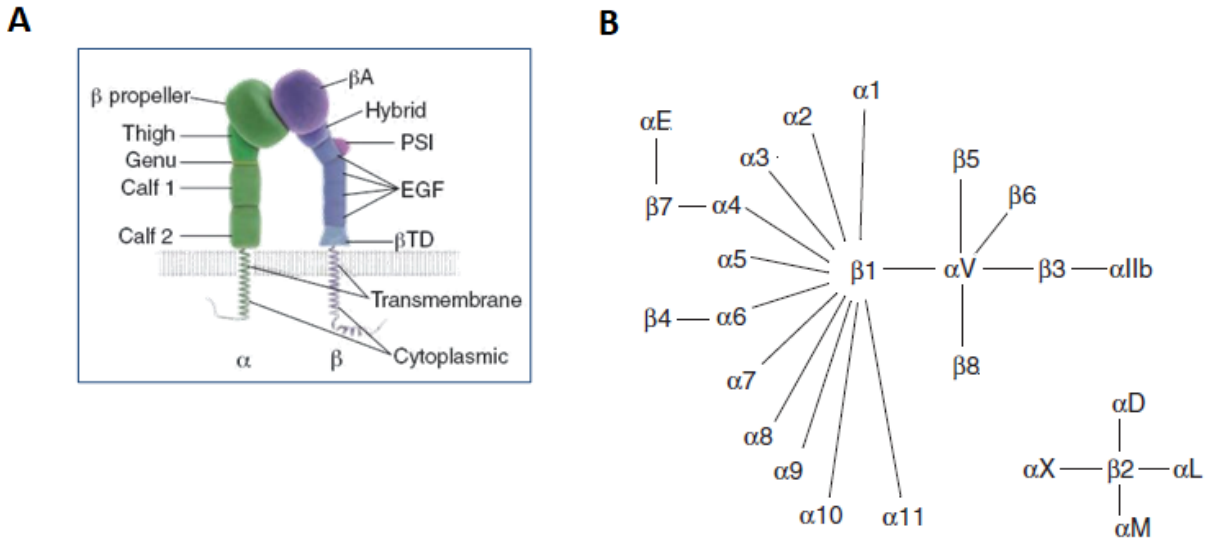


Figure 10 A) Schematic representation of integrin structure; B) Superfamily of human integrin receptors and how they combine to obtain an αβ heterodimer.

From the structural point of view, integrins are characterized by a large extracellular domain (750-1000 amino acids) and small transmembrane and cytoplasmatic domains (25-70 amino acids in most integrin types, except from the β_4 subunit, that is over 1000 amino acids long). The integrin binding pocket is typically displayed between the α and the β subunits.[77]. Recent structural investigations highlighted the presence of three possible conformations (Figure 11):[77] a bent conformation in which the head adopts a ‘‘closed’’, low affinity conformation and the cytoplasmic tails form an inhibitory complex; an extended conformation of the head that is still characterized by a low ligand affinity; and a high affinity form in which the legs and tails separate, and the ‘‘hybrid’’ domain, which is part of the head, moves away from the β-I domain, propeller and α-I domain, promoting conformational changes that create high affinity binding sites on both the head and tail. These integrin conformation plays a key role in mediating “inside-out signaling” mechanisms. In particular, talin and kindlin, two cytoplasmic adaptor proteins, binds to the intracellular domain of the β-integrin, causing a conformational change into the heterodimer, that adopts the active conformation. On the other hand, the binding of high-affinity ligand induces a conformational change of the cytoplasmic tails, promoting a signaling cascade that is required for many integrin-mediated cell mechanisms such as cell adhesion, spreading, migration, proliferation, survival, and differentiation.[78]

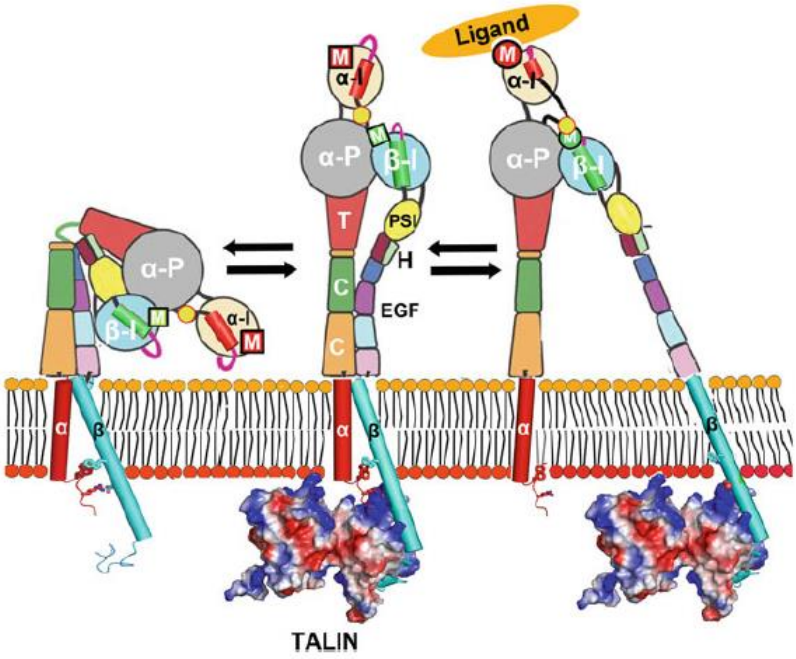


Figure 11 Cartoon of the aXb2 structure derived by crystallography and electro-migration (EM) studies. At left, the bent, low affinity integrin stabilized by bonds between the head, legs and cytoplasmic tails. At center, an unknown trigger causes the integrin to ''stand up'', while maintaining most of its low affinity bonds. At right, binding of activated talin and/or binding of an extracellular ligand, trigger an open, high affinity form of the integrin, with TM helices separated (Adapted from reference 77).

Changes in the ECM and the repertoire of integrins on tumor cells contribute to deregulate integrin signaling in cancer.[79] Integrin signaling alterations can be found in almost all steps of carcinogenesis, ranging from switches in the utilization of α-β heterodimers by the cell, to overexpression of the receptors, activation of downstream effectors of integrin signaling and interference with other signaling pathways.[80] In this perspective, the overexpression of integrins results to be crucial for the proliferation, survival and migration of cancer cells. Moreover, integrin expression on the surface of endothelial cells was found to be a marker of angiogenesis, that is the formation of new blood vessels in non-vascularized and growing tissues. Since the inhibition of angiogenesis was found as a sound therapeutic approach to retard tumor growth and development, specific integrins such as $\alpha_v\beta_3$, $\alpha_v\beta_5$ and $\alpha_5\beta_1$ has been identified as valid biological targets for targeted tumor therapy.[81] In 1984, Ruoslahti and Pierschbacher disclosed the molecular bases of the interaction between extracellular matrix proteins (e.g. fibronectin) and the integrin receptors.[82] In particular, these studies identified the Arg-Gly-Asp tripeptide sequence (RGD, Figure 12) as the key recognition element in fibronectin to engage binding interactions with integrins.

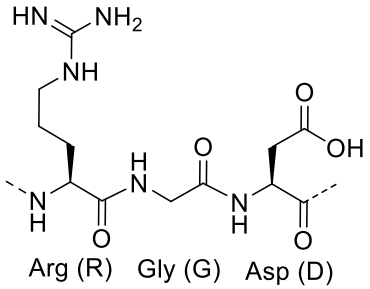


Figure 12 The Arg-Gly-Asp (RGD) recognition motif

After the Ruoslathi work, many research groups have been developing peptides or peptidomimetic ligands capable of binding integrins and interfere with their biological activity in different clinical settings. The first approach was the inclusion of the RGD motif into a cyclic structure, with the aim to give rigidity to the ligand and to stabilize the RGD loop. Among the structures reported in literature, Kessler and coworkers developed the well-known cyclic RGD-bearing integrin ligand called Cilengitide (compound **1**, Figure 13), which has been evaluated in clinical trials as potential antiangiogenic drug.[83]

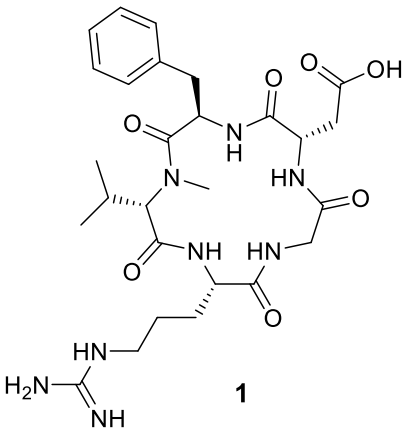


Figure 13 Chemical structure of Cilengitide (**1**)

Xiong and coworkers reported the X-ray analysis of $\alpha_v\beta_3$ integrin crystal structure in complex with Cilengitide, defining the structural basis of the ligand binding interactions (Figure 14).[84] In particular, the C_β of the arginine and the aspartic acid in the RGD motif were found to be presented in a 9 Å distance, which is instrumental for the ligand to engage the so-called “electrostatic clamp”. This interaction is based on two salt bridges: the first one is the interaction between the positively charged guanidine of the arginine and the two carboxylic acid side chains of Asp150 and 218 of the α_v subunit; the second one is the coordination of the Mn^{2+} cation in the metal ion-dependent adhesion site (MIDAS) of the β_3 subunit by the carboxylic acid side chain of the aspartic acid.

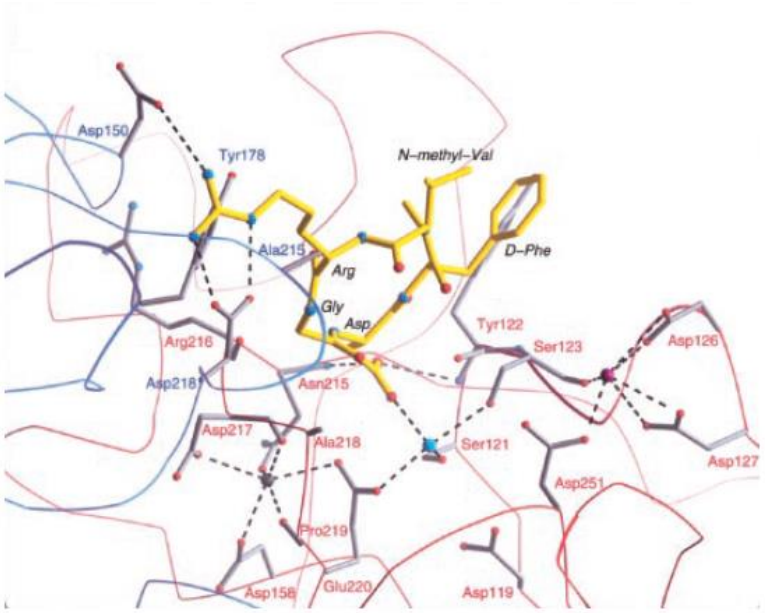


Figure 14 Interactions between the $\alpha_v\beta_3$ integrin binding site and Cilengitide. The ligand is depicted in yellow, the α_v subunit in blue and the β_3 subunit in red. The divalent cation in the β_3 subunit is Mn^{2+}. Hydrogen bonds and salt bridges are represented as dotted lines (Adapted from reference 84)

The disclosure of these data stimulated the development of new peptide and peptidomimetic integrin ligands that resemble the binding pose of cilengitide (Figure 15).[85]

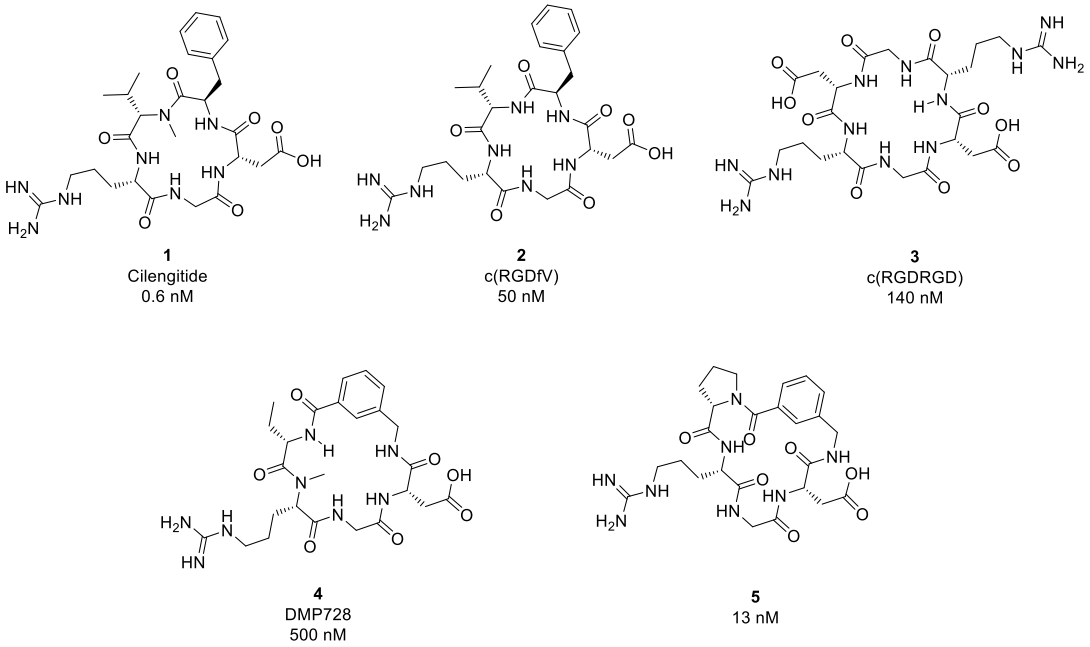


Figure 15 Examples of cyclic RGD peptides. The IC_{50} against $\alpha_v\beta_3$ is reported below each compound. For compound **2**, **4** and **5** the IC_{50} was determined with an ELISA-like assay.[86] For compound **3** the IC_{50} was determined using an assay in which the inhibition of vitronectin- $\alpha_v\beta_3$ complex is evaluated.[87]

Kessler and coworkers performed a conformational NMR analysis of compounds **1** and **2**,[83] showing that the distance between the C_β of the Arg and the Asp in compound **2** is 6.9 Å. Moreover, the N-methylation of the Val residue cause a dramatic enhancement of the binding affinity of **1** compared to **2**. On the other hand, the distance between the C_β of the Arg and the Asp In compound **3** is 5.8 Å,[87] which conceivably represents the cause of the lower binding affinity of **3** as compared to both **1**

and **2**. In parallel, Bach and DeGrado at DuPont Merck showed that the enhancement of the rigidity of a pentacyclic RGD peptide, obtained through the substitution of a L-2-aminobutanoic acid residue in compound **4** with a proline residue in compound **5**, led to an higher affinity of compound **5** for $\alpha_v\beta_3$ integrin receptor when compared to compound **4**.[88] The development and consequent clinical trials of Cilengitide started the era of RGD-containing molecules as antiangiogenic and anticancer compounds.[89] Cyclic scaffolds have also been inserted into the cyclic RGD moiety in order to obtain semipeptidic structures with a rigidity that is comparable to the Cilengitide and an enhanced metabolic stability (Figure 16).[90]

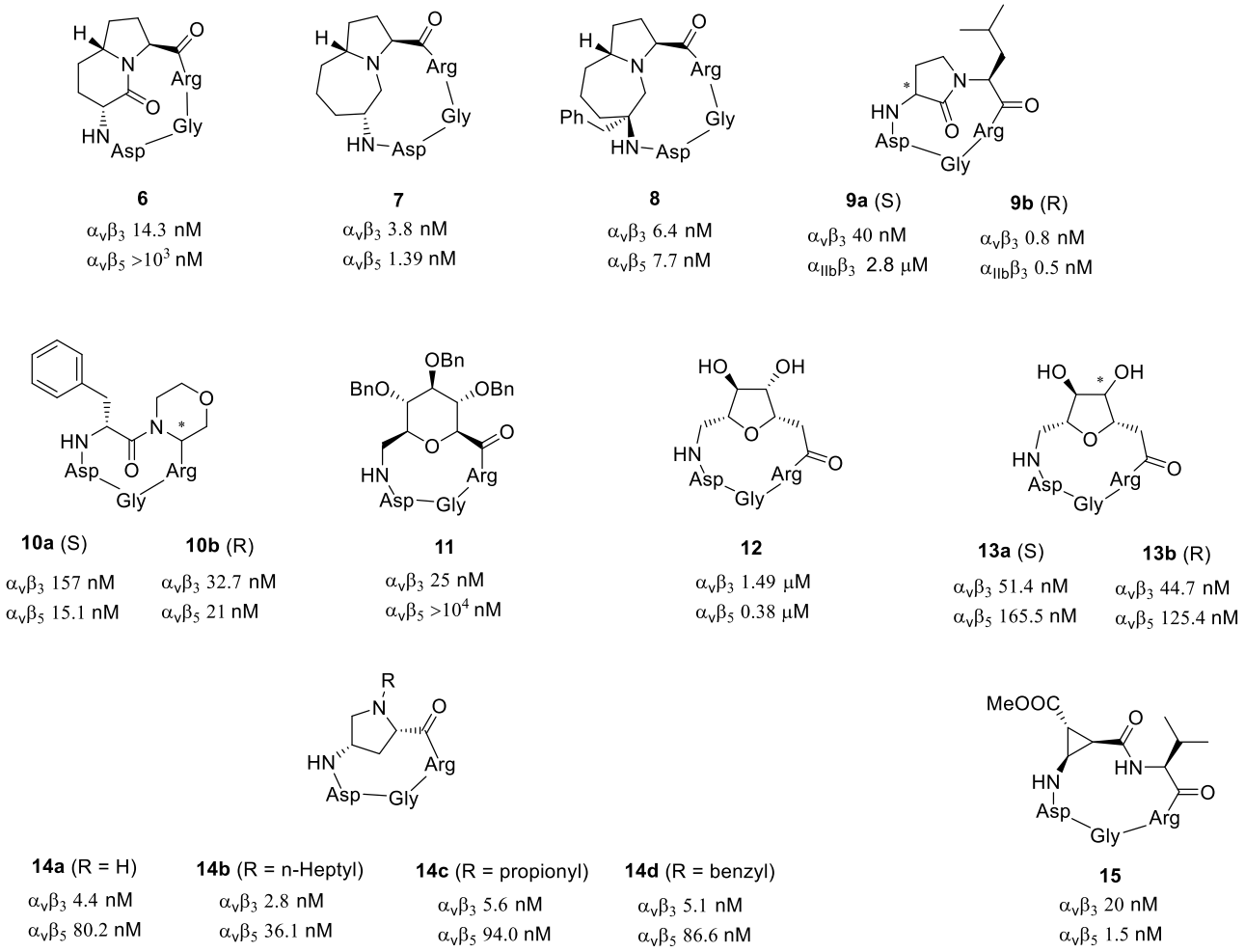


Figure 16 Examples of semipeptidic integrin ligands (**6**-**15**). The IC_{50} on the related integrin is reported below each compound.

In 2009-2012, Gennari and coworkers developed a new class of integrin binders in which a 2,5-diketopiperazine scaffold (DKP1-DKP7, Figure 17A) was installed in the RGD cyclic structure.[91] The use of DKP as peptidomimetic scaffold allowed to enhance the rigidity of the RGD, the metabolic stability of the whole structure as well as the binding affinity, as the DKP scaffold, can participate to the ligand-protein interactions through formation of hydrogen bonds. The chemical diversity could be introduced on N1 and N4 or changing the configuration of C3 and C6.

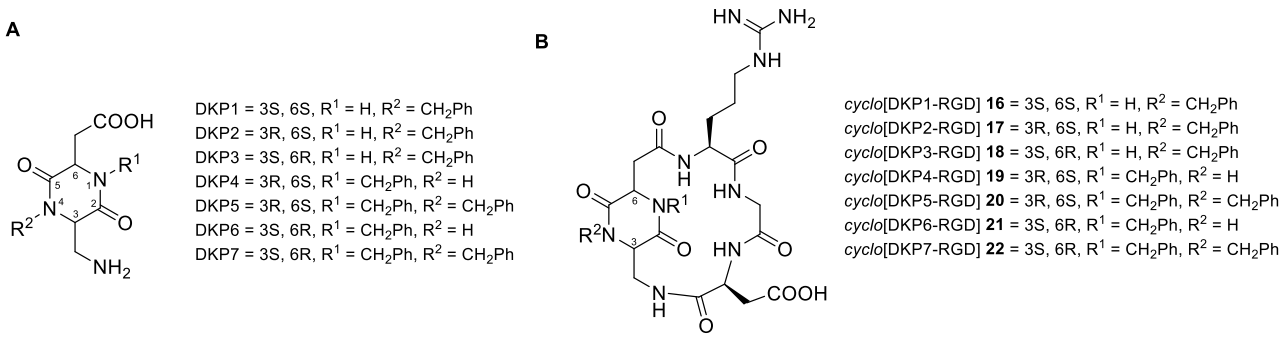


Figure 17 A) Chemical structure of DKP1-DKP7; B) Chemical structure of *cyclo*[DKP-RGD] peptidomimetics.

The group synthesized a library of DKPs and installed them into the RGD cycle, in order to obtain a library of *cyclo*[DKP-RGD] ligands (compounds **16**-**22**, Figure 17B). Competitive binding assays of compounds **16**-**22** on isolated $\alpha_v\beta_3$ and $\alpha_v\beta_5$ receptors have been performed (Table 2).

Table 2 Data collection from competitive binding assays of compounds **16**-**22** on isolated receptors.

Compound	IC_{50} (nM)[a]	
	$\alpha_v\beta_3$	$\alpha_v\beta_5$
1	0.6 ± 0.1[b]	11.7 ± 1.5
16	3898 ± 418	> 10^4
17	3.2 ± 2.7	114 ± 99
18	4.5 ± 1.1	149 ± 25
19	7.6 ± 4.3	216 ± 5
20	12.2 ± 5.0	131 ± 29
21	2.1 ± 0.6	79 ± 3
22	0.2 ± 0.09	109 ± 15

[a] IC_{50} values were calculated as the concentration of compound required for 50% inhibition of biotinylated vitronectin binding. Screening assays were performed by incubating the immobilized integrins $\alpha_v\beta_3$ and $\alpha_v\beta_5$ with increasing concentrations (10^{-12} – 10^{-5} M) of the RGD ligands in the presence of biotinylated vitronectin (1 mg/mL) and measuring the concentration of bound vitronectin in the presence of the competitive ligand. [b] Calculated as the concentration of compound required for 50% inhibition of biotinylated vitronectin binding.[92]

The biological assays catalogued the *cyclo*[DKP-RGD] molecules as low nanomolar integrin binders. Conformational analysis and MC/SD simulations of compounds **17**-**22** have been performed in order to understand better how the ligands arrange in solution. The presence of the DKP not only gives rigidity to the backbone by the formation of a cyclic structure, but also interacts with the RGD backbone with hydrogen bonds that locks the electrostatic clamp into a well-defined conformation. Further *in vitro* studies performed on *cyclo*[DKP3-RGD] **18** showed the high capability of the ligand to inhibit the angiogenesis in human umbilical vein endothelial cells (HUVECs). Compound **18**, however, did not interfere with other vital processes like cell viability and proliferation.[92]

In conclusion, peptide and peptidomimetic RGD compounds proved capable of interacting effectively with integrin receptors and this molecular recognition has been highlighted in a number of

biochemical and biological experiments. Although the potential of this class of compounds for therapeutic use is not well understood, the RGD-integrin interactions at the molecular level are well characterized, offering the opportunity to use RGD ligands as valuable case study to investigate innovative approaches to develop small molecule ligands with antibody-like affinity.

Dimeric bicyclic peptides as high-affinity small ligands

Part of the research work described in this Chapter was published in the following article:

- G. Sacco, A. Dal Corso, D. Arosio, L. Belvisi, M. Paolillo, L. Pignataro, C. Gennari. *Org. Biomol. Chem.* **2019**, *17*, 8913-8917.

2.1 Introduction

Among the possible chemical strategies that have been employed to enhance the binding affinity of small ligands, the exploitation of multivalency is one of the most traditional. In general, the valency of a given ligand is defined as the number of interactions of the same nature that it can engage with the parent receptor during binding.[93] In particular, multivalency can be defined as the interaction between an *n*-valent receptor and an *m*-valent ligand (m, $n > 1$; $m \neq n$),[94] whereas the interaction of multiple and isolated monovalent ligands ($m = 1$) with a multivalent receptor is not defined as multivalent. Multivalency is largely used in nature exploit the cooperativity of single weak interactions. For instance, the internalization of different types of some virus particles into the target cell is based on multivalent interactions between multiple copies of transmembrane proteins expressed by the virus and the receptors on the target cell surface: while the binding affinity of a single viral protein for the receptor may be relatively weak, the cooperativity of many weak interactions allows the tight anchoring of the virus to the cell, which often initiates the infection process. Inspired by nature, the presentation of multiple copies of a pharmacophore portion onto suitable molecular architectures leads to multimeric ligands capable of multivalent interactions with the cognate receptor. In this context, these multimeric compounds are often endowed with “branched” structures, in which a central scaffold is grafted with *m* copies of the monovalent ligand. This multimeric presentation often increases the apparent affinity (or, more appropriately, the “avidity”) of the resulting construct for the targeted receptor. The cooperative effect can result from two binding modalities: the interaction of *n* different ligand units on a multivalent scaffold with *n* different copies of the receptor, defined as “cluster effect” (Figure 18A), or the alternate binding of different ligand copies to the same receptor unit, defined as “rebinding effect” (Figure 18B).[95] In these binding modalities, the ligand-receptor complex formation is promoted by both thermodynamic and kinetic factors.[94,96] From the thermodynamic point of view, a multivalent ligand reduces the loss of entropy typically involved in the formation of monovalent ligand-receptor complexes, thus making the binding event more favorable; from a kinetic point of view, a high ligand “local concentration” (i.e. an increased probability of a new ligand-receptor complex formation upon complex dissociation) is widely accepted as significant component of the cooperative effect.[97]

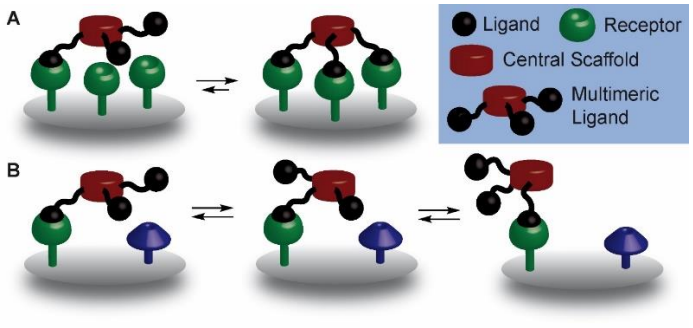


Figure 18 Binging modes of a branched multivalent molecule. A) The "cluster effect" involves multiple copies of the binding portion on different receptor units. B) The "rebinding effect" involves different copies of the binding portion on a single receptor unit.

Owing to the increased complex stability promoted by multivalent interactions, different ligands of clinically-relevant proteins have been developed for therapeutic purposes. For instance, multimeric RGD ligands have been designed and synthesized to effectively bind integrin $\alpha_V\beta_3$ integrin receptor.[98] Gennari and coworkers recently developed a series of multimeric small molecule-drug conjugates (SMDCs) with the aim to investigate their tumor-targeting performances.[99] A fundamental unit of these multivalent structures consisted in the peptidomimetic $\alpha_v\beta_3$ integrin ligand *cyclo*[DKP-RGD] (compound **23** in Figure 19), functionalized with a benzylamino moiety suitable for the conjugation to an azido-modified PEG4 spacer. The cytotoxic payload and linker units of these SMDC products consisted respectively in the microtubue-stabilizing agent Paclitaxel (PTX **24**) and the protease-cleavable Val-Ala dipeptide. The latter was functionalized at its N terminus with different scaffolds, bearing from one to four propargylic moieties: with this design, the integrin ligand and linker-drug modules are joined by triazole ring formation through copper-catalyzed azide-alkyne cycloaddition (CuAAC, the most traditional "click" reaction), leading to different SMDC products with increased valency from 1 to 4. The chemical structures of the resulting conjugates are reported below (compounds **25-28**, Figure 19B). In competitive binding assays, compounds **25-28** showed progressive enhancement of binding affinity for the isolated $\alpha_v\beta_3$ integrin receptor, which increases along the conjugate valency. As a matter of fact, IC_{50} values range from 27.3 ± 9.8 nM of the aromatic monomeric compound **25**, to 1.2 ± 0.5 nM of the trimeric compound **27**. Interestingly, the tetrameric compound **28** showed a binding affinity comparable to compound **27**, which arguably indicates that the functional exploitation of the cooperative effect can be only efficient below a certain valency number (in this case, the trimeric compound). In fact, it has been reported that redundant ligand copies in high-multivalent structures may not contribute to the complex formation (i.e. lowering the so-called "ligand economy") or even interfere with binding.[100]

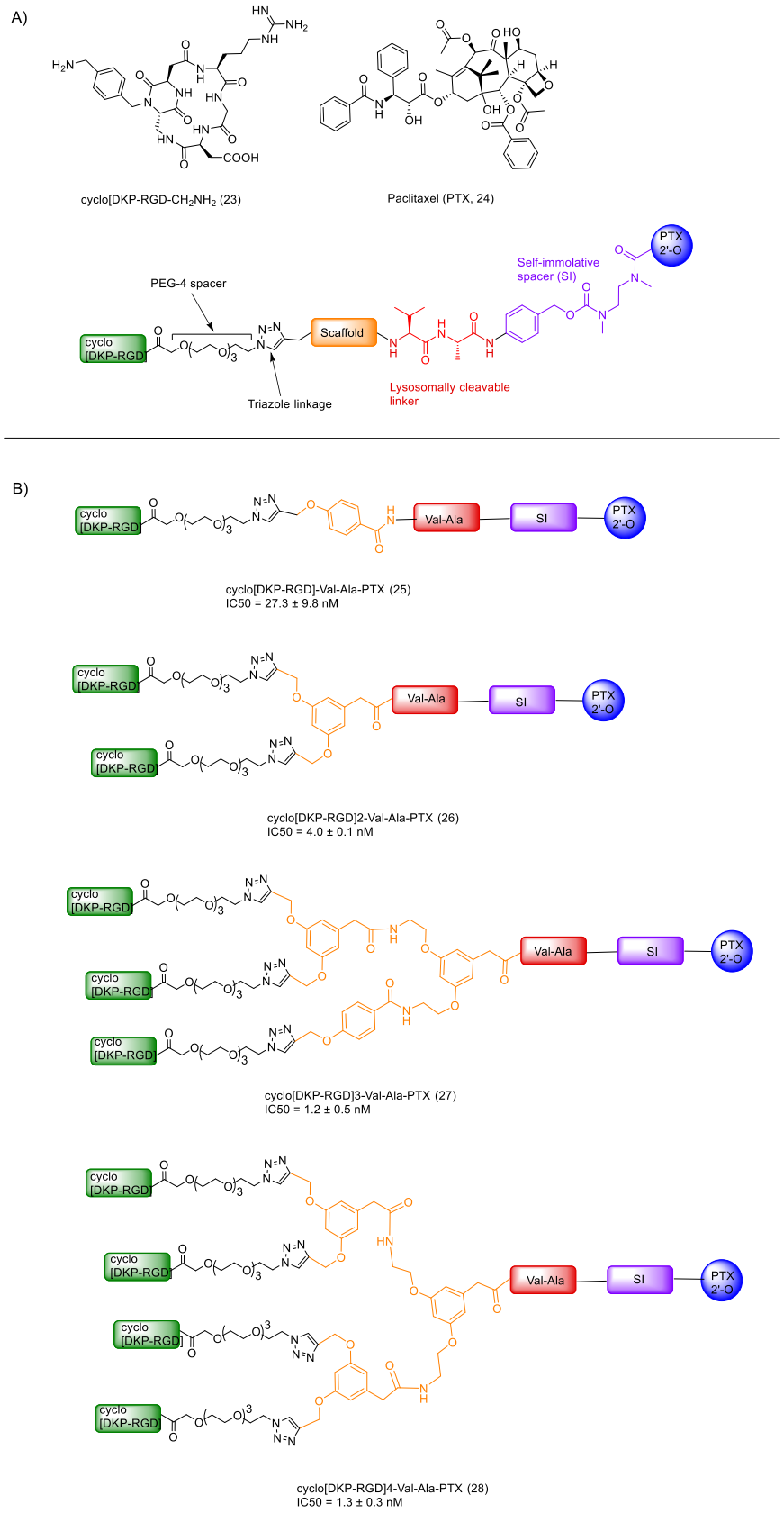
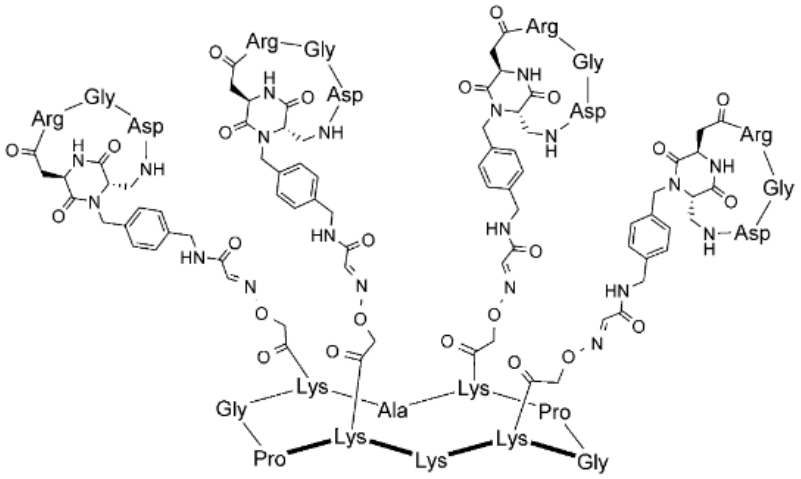


Figure 19 A) Representation of each component of the SMDCs and general structure of multimeric SMDCs. B) Molecular structures of the monomeric (**25**), dimeric (**26**), trimeric (**27**) and tetrameric (**28**) SMDCs, together with the relative IC_{50} values measured in competitive integrin binding assays (Adapted from reference 99).

More recently, the Gennari and Boturyn groups reported the conjugation of the *cyclo*[DKP-RGD] ligand to a cyclic decapeptide called RAFT (Regioselectively Addressable Functionalized Template, compound **29**, Figure 20).[101]



RAFT[cyclo(DKP-RGD)$_4$ **29**
IC_{50} = 0.98 ± 0.27 nM

Figure 20. Molecular structure of the $\alpha_V\beta_3$ integrin tetrameric RAFT[cyclo(DKP-RGD)] ligand **29**. The IC_{50} values of **29** in competitive integrin binding assays is indicated.

Also in this case, the sub-nanomolar IC_{50} of compound **29** indicated a strong cooperative effect of the multimeric presentation, which is beneficial for the resulting stability of the ligand-integrin complex. Considering the two tetrameric RGD constructs **28** and **29**, both endowed with four copies of the same integrin ligand, it is possible to note that RAFT-bearing conjugate **29** showed lower IC_{50} values in competitive integrin binding assays: this data highlights the structural importance of the multimeric scaffold in displaying the correct ligand presentation for a functional avidity for the related receptor.

Moving from recent studies of our group, it has been reported that the design of “branched” multimeric ligands (e.g. similar to compound **28** and **29**) could influence the pharmacokinetic performances *in vivo* of the resulting conjugate,[102] in some cases leading to a lower tumor:organ uptake as compared to monovalent counterparts.[103] To this end, we speculated that a “condensed” multimeric design of small molecule ligands would represent a valid alternative to large “branched” structures. In the case of cyclic peptides, a condensed dimeric structure may be obtained by replicating the same peptide motif responsible for receptor engagement (e.g. the RGD tripeptide in integrin-binding ligands) and displaying it twice in a macrocyclic peptide. However, besides exhibiting a lower stability towards proteolytic degradation, these structures are also characterized by a higher conformational flexibility than small-sized analogues, which typically leads to a decreased binding affinity for the target receptor. In this research context, the formation of bicyclic structures proved beneficial for both the *in vivo* stability and the binding affinity of the resulting ligands, owing to an increased structural rigidity. Due to these conformational properties, bicyclic peptides have been

successfully employed to engage in high-affinity interactions with flat binding sites in “undruggable” protein targets (e.g. interfaces of protein-protein interactions), which increased the worldwide interest in bicyclic peptides as “next-generation therapeutics”.[104] In recent years, the field was significantly pioneered by the seminal work of Winter and Heinis on peptide phage display,[105] that is a highly robust and widely applied technology to develop bicyclic peptide ligands against target proteins. In particular, peptides displayed on filamentous phages are constrained into bicyclic structures via thiol-reactive crosslinking agents (Figure 21), such as 1,3,5-tris(bromomethyl)benzene (TBMB, **30**), 1,3,5-triacryloylhexahydro-1,3,5-triazine (TATA, **31**), 2,4,6-tris(bromomethyl)-1,3,5-triazine (TBMT, **32**), *N,N′,N″*-(benzene-1,3,5-triyl)tris(2-bromoacetamide) (TBAB, **33**) and 1,1',1''-(1,3,5-triazinane-1,3,5-triyl)tris(2-bromoethan-1-one) (TATB, **34**), or simply by disulfide bond formation among cysteines.

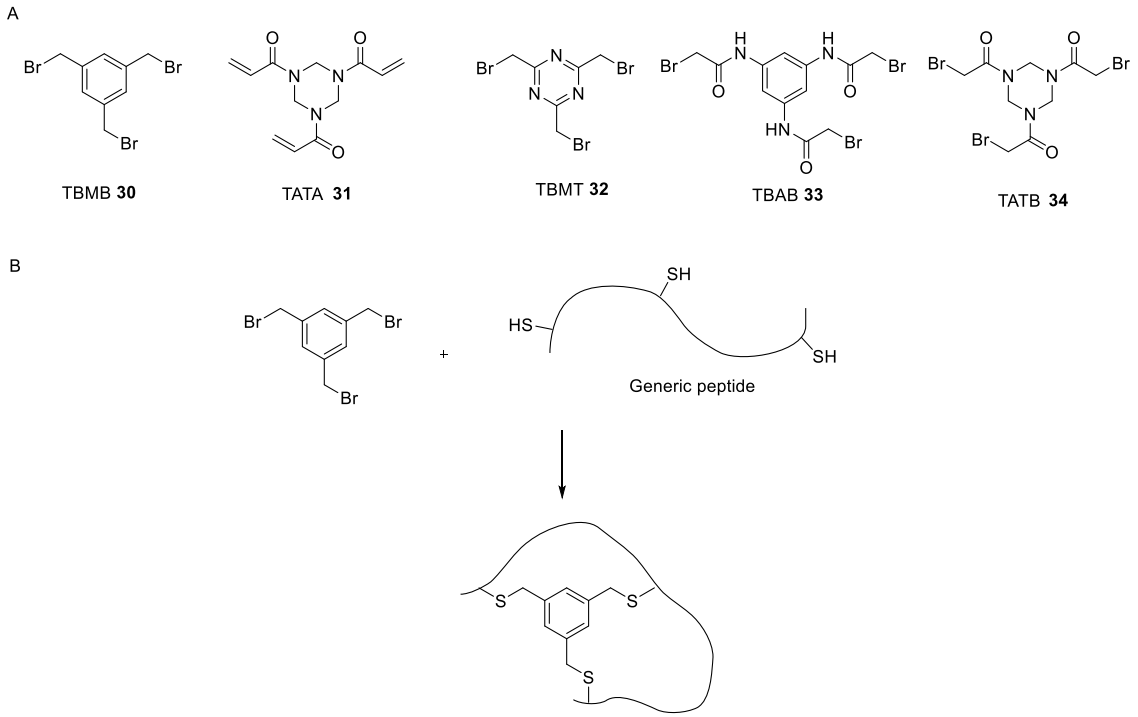
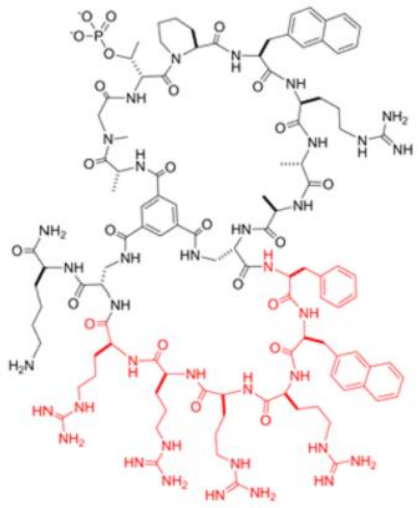


Figure 21 A) Molecular structures of crosslinking agents for the cyclization of linear peptides to obtain bicyclic analogues. B) An example of crosslinking reaction between compound **30** and a generic peptide, endowed with three cysteine residues. In this example, the thiol group of the cysteine side chains to undergo S_N2 nucleophilic substitution on the electrophilic TBMB, leading to a bicyclic structure.

Further advances on bicyclic peptides have been provided by Pei and coworkers, who developed cell-permeable bicyclic peptides to target intracellular proteins,[106] where the target-binding motif and a cell-permeable poly-arginine motifs are individually exhibited on the two rings of a unique bicyclic peptide structure. An example is the bicyclic peptidyl-prolyl cis−trans isomerase (Pin1) inhibitor **35** (Figure 22) reported by Pei and coworkers.[106a]



Pin1 Inhibitor **35**
(K_d = 72 nM)

Figure 22 Molecular structure of the cell-permeable bicyclic peptide **35**, an inhibitor of the Pin1

2.2 Design and synthesis of bicyclic RGD ligands

Inspired by these applications of bicyclic peptides, we envisioned a synergy between multivalent ligands and bicyclic structures by designing a dimeric bicyclic RGD ligand (RGD-2C-RGD, compound **36**, Figure 23A).

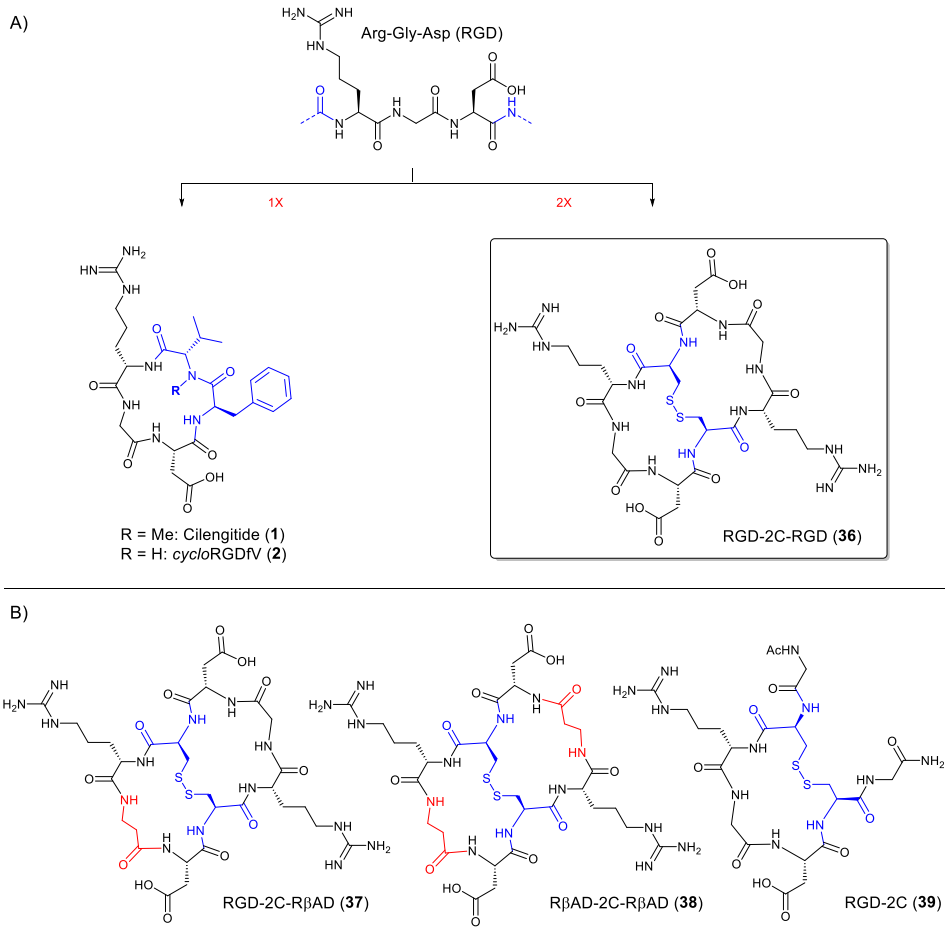
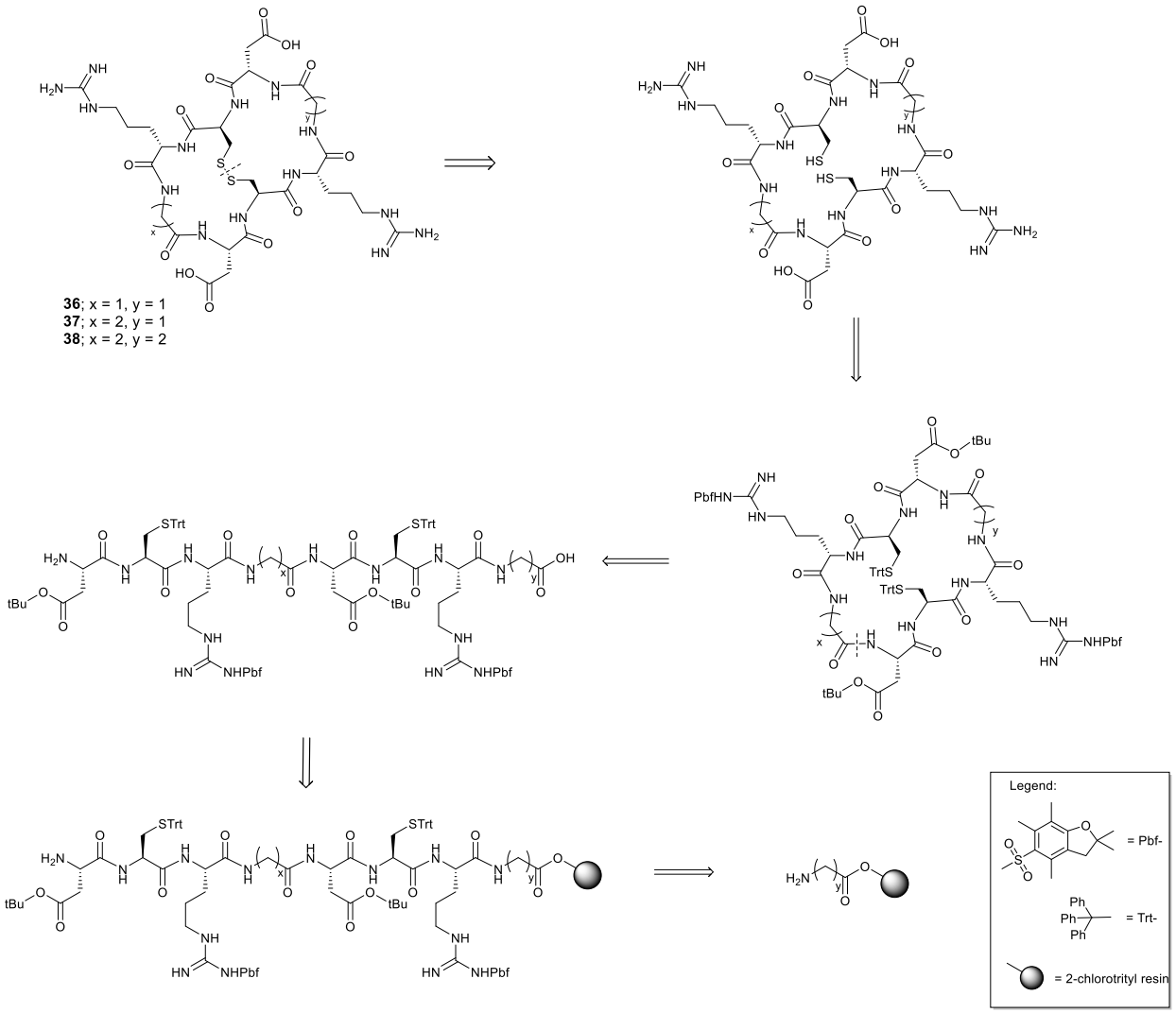


Figure 23 A) Design of a dimeric bicyclic peptide as integrins ligand: two RGD sequences are displayed on a symmetric bicyclic structure (RGD-2C-RGD, compound **36**); B) Molecular structure of the bicyclic peptides RGD-2C-RβAD **37**, RβAD-2C-RβAD **38** as well as the monocyclic RGD-2C peptide **39**.

Unlike traditional monocyclic RGD structures (e.g. Cilengitide **1**), the newly designed compound **36** was characterized by the presence of two RGD portions displayed on a macrocyclic structure. The bicyclic structure is given by a disulfide bond between two Cys side chains. The resulting dimeric structure is C_2-symmetric, where the axis of symmetry crosses the disulfide bond. Since the two RGD motifs in compound **36** are in close proximity on the same molecular structure, we hypothesized that cooperative effects between the two tripeptides for integrin binding may be exhibited through the "statistical rebinding" mode (see Figure 18B). Aiming at investigating evidences of such a multivalent effect and its impact on the ligand binding potency, we designed suitable control compounds bearing a similar chemical structure to the parent compound **36**, while being unable to exhibit a multivalent binding. In particular, as highlighted in Chapter 1, a high-affinity interaction between the RGD tripeptide and the related integrin typically results from the 9-Å distance between the Arg and Asp side chains, which is important for the formation of a functional "electrostatic clamp"

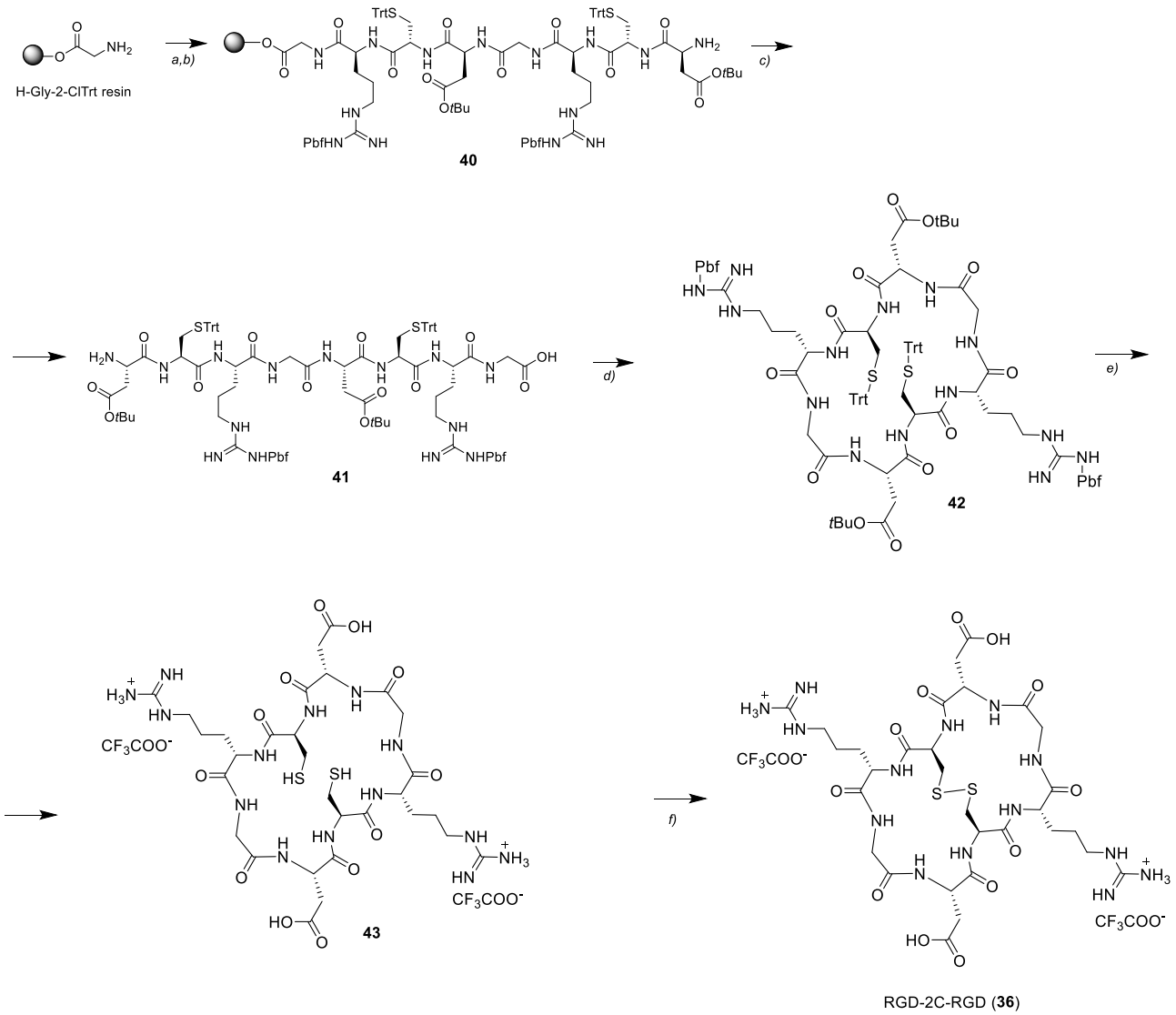
in the integrin binding pocket. Since a longer distance between the Arg and the Asp residues was known to be detrimental for the tripeptide binding affinity, we devised the substitution of the Gly residue with an elongated β-Ala fragment. In particular, bicyclic peptide (RGD-2C-RβAD, compound **37**, Figure 23B) was designed as monomeric analogue of compound **36**. Moreover, a dual substitution of the two Gly residues in **36** with two β-Ala was performed (resulting in the bicyclic peptide RβAD-2C-RβAD, compound **38**, Figure 23B). Finally, a monocyclic RGD peptide (RGD-2C, compound **39**, figure 23B) was designed with the aim to assess the contribution of the bicyclic structure to the integrin binding affinity of the dimeric peptide **36**.

The retrosynthetic analysis of bicyclic peptides **36**-**38** (Scheme 1) is straightforward: the first disconnection performed on the target compounds concerns the disulfide bond, which disrupt the bicyclic structure leading to the corresponding macrocycle. The second disconnection is performed at the amide bond between the Asp amine group and the Gly or βAla carboxylic acids. This specific disconnection was chosen to circumvent the possible epimerization of the C-terminal amino acids during amide bond formation, being a well-known side-reaction involving the stereogenic C_α centers in the carboxylate substrate during the reaction with amide coupling reagents. For this step, protection of the amino acid side chains were devised to drive the selectivity of the amide bond formation and, in particular, well-known acid-labile protecting groups, such as *tert*-butyl for the Asp, trityl for Cys and 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) for Arg, were chosen. The protecting groups allowed both the peptide growth using a Fmoc/*t*Bu solid-phase peptide synthesis (SPPS) protocol on resin, while the use of a mild-acid-labile resin was devised to allow the peptide cleavage from the solid support with intact protecting groups on the side-chains. The synthetic pathway to compound **36** (Scheme 2) started with a SPPS protocol performed on the commercially available H-Gly-2ClTrt resin, in which a Gly residue was coupled to a 2-chlorotrityl (2ClTrt) functionalized resin. The 2ClTrt resin was chosen as solid support because it allows the cleavage of the peptide under mild-acidic conditions, without the simultaneous cleavage of the side chain protecting groups. Moreover, unlike the C-terminal amide isolation in other commercially available resins, 2ClTrt allows the isolation of the C-terminal free carboxylic acid moiety, which can then undergo macrolactamization step. The SPPS protocol followed during the peptide growth was a Fmoc/*tert*-Bu protocol, which consists in the progressive installation of Fmoc-protected amino acids (bearing acid-labile protecting group on the side chain) at the N-terminal amine group on the immobilized peptide. The free carboxylic group of the Fmoc-amino acid was activated using diisopropyl carbodiimide (DIC) and 1-hydroxy-7-azabenzotriazole (HOAt) as additive at 0 °C for 15 min in dry DMF. HOAt was also added as additive, as this is typically important to minimize the probability of amino acid racemization during the coupling step.[107] The activated amino acid was then added to the resin and the coupling step is performed at 75 °C for 10 min under microwave irradiation. After different washing steps, the resin was treated twice with a solution of 20% piperidine in DMF at r.t. in order to perform the Fmoc deprotection.



Scheme 1 Retrosynthetic analysis of bicyclic peptides **36**-**38**.

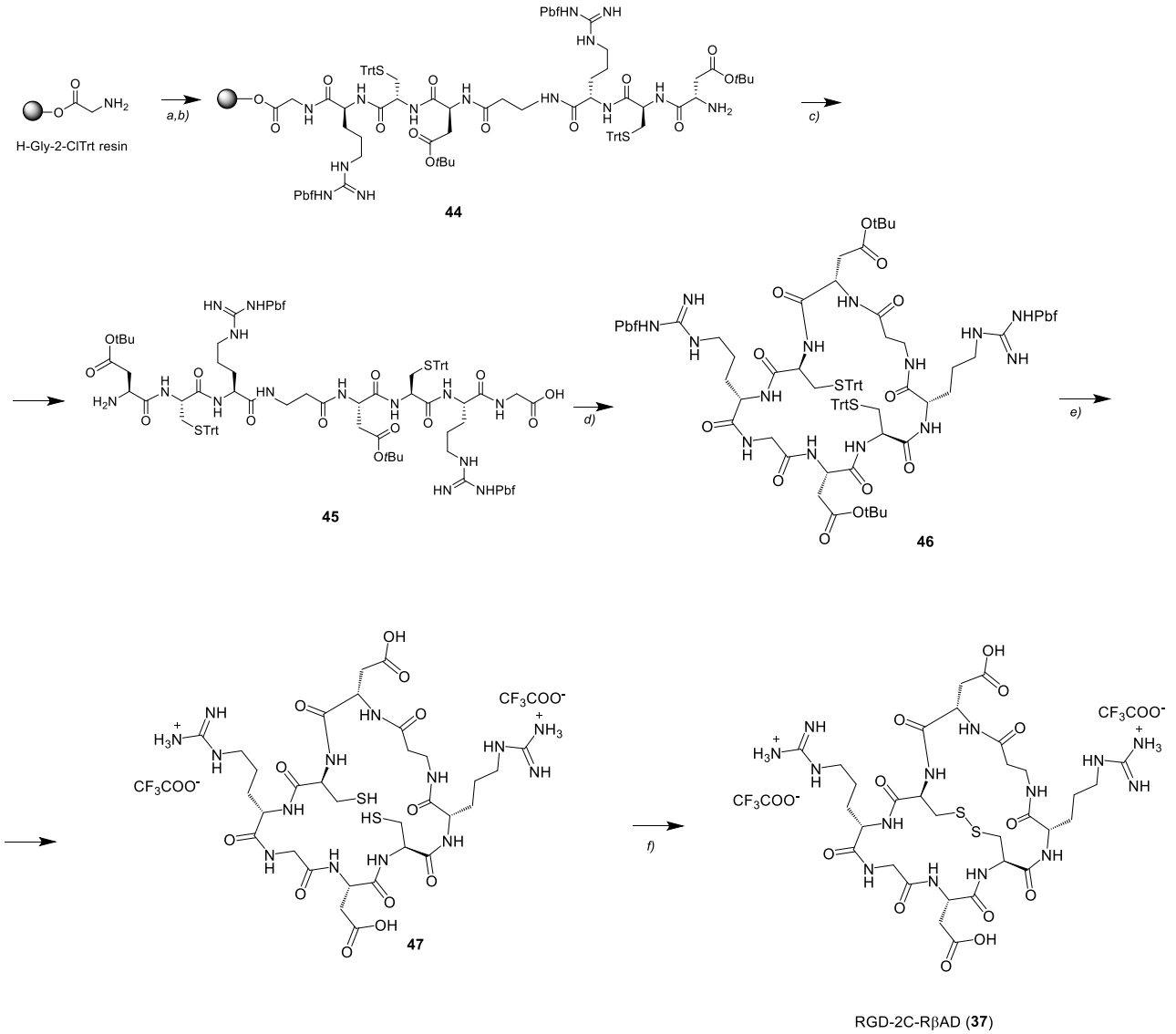
After washing steps, the resin exhibits a free amino group , ready to react in the following coupling step. The Fmoc/*t*Bu protocol was repeated for the full amino acid sequence, obtaining the 2ClTrt-bound fully protected linear peptide **40**. Later on, the linear protected peptide **41** was obtained using a solution of acetic acid, 2,2,2-trifluoroethanol and DCM in a volume ratio of 1:2:7. This cleavage mixture was acidic enough to cleave the peptide from the 2-ClTrt resin, but not acidic enough to initiate the removal of protecting groups on the amino acid side chains. Compound **40** was treated four times with the cleavage mixture and each cleavage step lasted 20 min. Compound **41** was obtained after concentration of the cleavage solution and subsequent precipitation in cold diethyl ether. The crude solid was directly used in the macrolactamization step. Compound **42** was treated with the coupling agent HATU, the additive HOAt and the base *N,N*-diisopropylethylamine (DIPEA) in dry DMF as solvent. The concentration of the macrolactamization step was very low (1,4 mM), in order to achieve the intramolecular cyclization and avoid the formation of intermolecular adducts. The fully protected macrocyclic peptide **42** was filtrated over silica gel and fully deprotected using a mixture of TFA:TIS:H_2O 95:2,5:2,5 v/v.



Scheme 2 REAGENTS AND CONDITIONS: a) Fmoc-AA-OH, DIC, HOAt, DMF, 75 °C (MW), 10 min; b) 20% piperidine in DMF; c) AcOH:2,2,2-trifluoroethanol:DCM 1:2:7; d) HATU, HOAt, *i*Pr_2NEt, DMF; 0 °C to r.t.; g) TFA:TIS:H_2O 95:2.5:2.5 v/v/v, 2 h, 0 °C to r.t.; h) I_2; H_2O:MeCN 1:1, 30 min, r.t.; Y = 2.98% over 5 steps.

The use of triisopropylsilane (TIS) and water is a well-known protocol to quench reactive cationic or radical species that may form during the side chain deprotection. The cleavage cocktail was added to compound **42** at 0 °C and then stirred at r.t. for 2h. Also in this case, compound **43** was precipitated after partial concentration of the reaction mixture and subsequent addition of cold diethyl ether. Finally, to form disulfide bond, the deprotected macrocyclic peptide **43** was treated with an excess of iodine in a 1:1 mixture of water/acetonitrile. Also in this case, the reaction was performed under high dilution conditions (2 mM) in order to avoid intermolecular disulfide oligomerization. The resulting crude was purified with semipreparative reverse-phase (RP) HPLC to obtain the dimeric bicyclic peptide **36** with a 2.98% yield over 5 steps. The synthesis of compound **37** was similar to the one used for compound **36** with the only difference that, during the peptide growth, Fmoc-Gly-OH was substituted with Fmoc-βAla-OH (Scheme 3). A Fmoc/*t*Bu protocol was used also in this case for the peptide growth, leading to the anchored fully protected peptide **44**. The peptide was cleaved from the resin to obtain compound **45** that was used as starting material for the

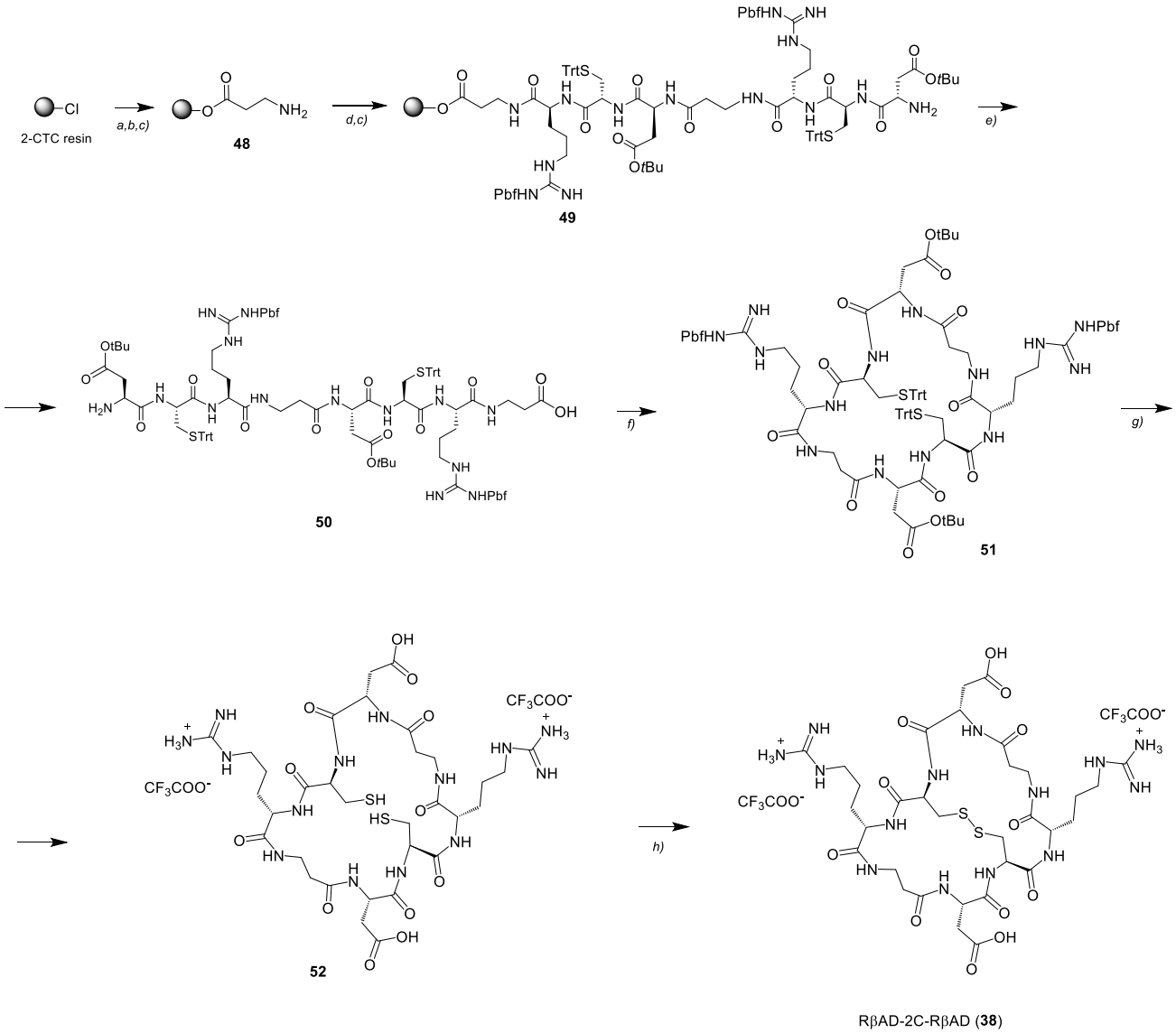
macrolactamization. The macrocyclic protected peptide **46** was obtained and the deprotected using the same cleavage cocktail used for the preparation of compound **43**. The deprotected macrocyclic peptide **47** was then treated with iodine to obtain the monomeric bicyclic ligand **37** with a 4% yield over 5 steps.



Scheme 3 REAGENTS AND CONDITIONS: a) Fmoc-AA-OH, DIC, HOAt, DMF, 75 °C (MW), 10 min; b) 20% piperidine in DMF; c) AcOH:2,2,2-trifluoroethanol:DCM 1:2:7; d) HATU, HOAt, *i*Pr$_2$NEt, DMF; 0 °C to r.t.; g) TFA:TIS:H_2O 95:2.5:2.5 v/v/v, 2 h, 0°C to r.t.; h) I_2; H_2O:MeCN 1:1, 30 min, r.t.; Y = 4% over 5 steps.

Finally, the synthesis of compound **38** (Scheme 4) was achieved by introducing minor modifications to the previous protocol. Since the C-terminal sequence of the linear peptide features a β-alanine residue, the first step was the functionalization of a 2-chlorotrityl chloride resin (2-CTC) with a Fmoc-βAla-OH. The reaction was an S_N1 on the tertiary trityl carbon and it was performed at r.t for 1 h using DIPEA as base and a 1:1 mixture of DCM and DMF as solvent. After several washing steps, the resin was treated with MeOH in order to cap the unreacted 2-CTCs on the resin. To measure the effective loading of β-alanine residue on resin, a UV protocol was used, consisting in the absorbance measurement at 301 nm of the 9-fluorenylmethyl-piperidine adduct resulting from Fmoc removal.

The measured absorbance was correlated to the concentration of the adduct that is due to the amount of Fmoc protecting groups cleaved from the resin.



Scheme 4 a) Fmoc-ßAla-OH; iPr$_2$NEt, DMF:DCM 1:1, r.t.. 1h; b) MeOH, r.t., 15 min; c) 20% piperidine in DMF; d) Fmoc-AA-OH, DIC, HOAt, DMF, 70 °C (MW), 10 min; e) AcOH:2,2,2-trifluoroethanol:DCM 1:2:7; f) HATU, HOAt, iPr$_2$NEt, DMF; 0 °C to r.t.; g) TFA:TIS:H2O 95:2.5:2.5 v/v/v, 2 h, r.t.; h) I_2; H_2O:MeCN 1:1, 30 min, r.t.; Y = 1% over 6 steps.

The correlation between the loading and the absorbance at 301 nm is reported in the following equation.

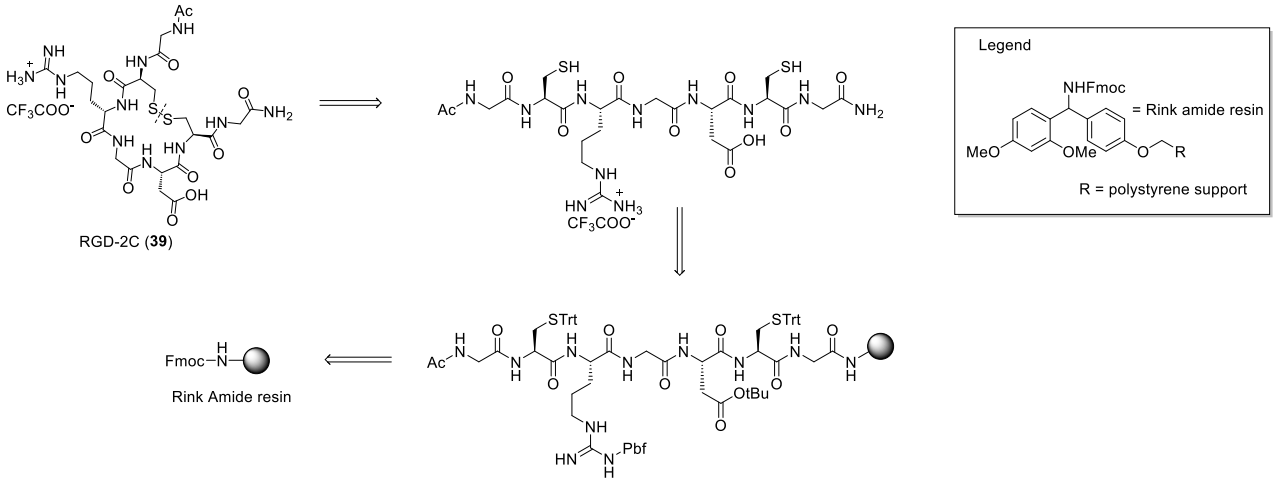
$$X = \frac{A(301\ nm) * V * F_d}{\varepsilon(301\ nm) * m * b}$$

Equation 1 Formula used for the determination of the loading after the Fmoc deprotection. X = loading on the resin (mmol/g); A (301 nm) = absorbance of the solution measured at 301 nm; V = total volume of collected deprotection solution; Fd = dilution factor; ε (301 nm) = 7800 M^{-1} cm^{-1}; m = mass of the resin (g); b = length of the cell (cm).

In this case, the loading was quantitative. The synthetic pathway followed for compound **38** was analogue to the one used for compounds **37**. The peptide growth was, in this case, performed on the 2ClTrt resin functionalized with the β-alanine **48** to obtain compound **49**. The protected peptide anchored to the resin was cleaved using the cleavage mixture used for compounds **41** and **45**,

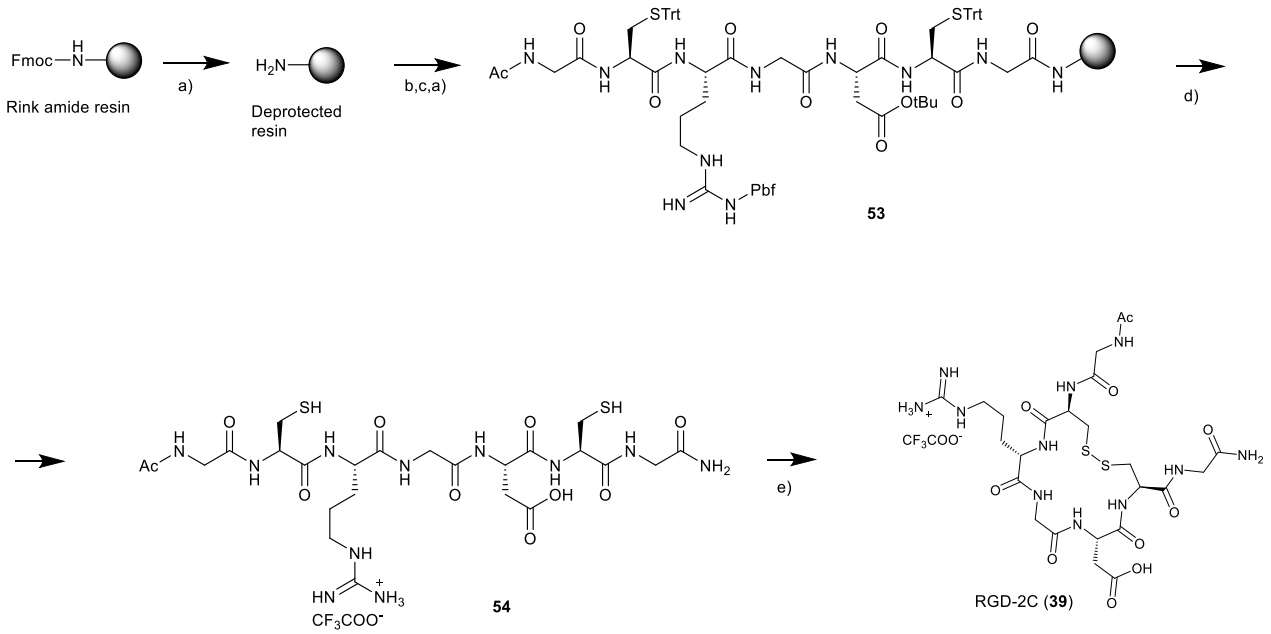
leading to the protected linear peptide **50**. The latter was subjected to a macrolactamization step to obtain the protected macrocyclic peptide **51**, which was then deprotected with the same protocol used for compounds **43** and **47** to obtain the deprotected macrocyclic peptide **52**. The bicycle formation through a disulfide bond was performed on compound **52** to obtain the bicyclic compound **38** with a 1% yield over 6 steps. Giving a look to the final yields obtained for the preparation of compounds **36**-**38**, it should be noted that the efficiency of the synthesis of the bicyclic peptides is not very high. The reasons could be mainly ascribed to the presence of two non-optimized cyclization reactions, such as the macrolactamization and the intramolecular disulfide bond formation, in which there is an intrinsic risk of intermolecular undesired reactions.

In contrast to macrolactamization step for the synthesis of bicyclic peptides, the cyclic structure of compound **39** results directly from the disulfide bond among Cys side chains. As the retrosynthetic analysis of **39** (Scheme 5) did not include the macrolactamization step, we devised SPPS onto a Rink amide resin. During the acidic cleavage step, this resin allows the capping of the C-terminal carboxylic acid as a primary amide, while all side-chain protecting groups are removed. The peptide growth was accomplished using again the Fmoc/*tert*-Bu strategy.



Scheme 5 Retrosynthetic analysis of RGD-2C (**39**).

The synthetic pathway followed to obtain compound **39** is reported below (Scheme 6). The Rink amide resin was initially treated with a 20% piperidine solution in DMF to deprotect the amino group and make it available for the first amino acid coupling. Also in this case, the activation of the Fmoc-protected amino acid was performed using DIC and HOAt in dry DMF and stirring the resulting solution at 0 °C for 15 min. The reaction mixture was then added to the deprotected Rink amide resin and the flask was warmed to 75 °C for 10 min, under microwave irradiation. In contrast to the capping step used for the 2ClTrt resin, the capping of unreacted amino groups on the Rink amide resin was carried out using a 20% acetic anhydride solution in DMF. This coupling/capping/deprotection sequence was repeated for all amino acids, leading the full peptide sequence **53**, anchored to the resin and protected on the amino acid side chains. Compound **53** was cleaved from resin and simultaneously deprotected using a TFA:TIS:H_2O 95:2,5:2,5 cleavage cocktail.



Scheme 6 Reagents and conditions: a) 20% piperidine in DMF; b) Fmoc-AA-OH, DIC, HOAt, DMF, 75 °C (MW), 10 min; c) 20% Ac_2O in DMF; d) TFA:TIS:H_2O 95:2.5:2.5 v/v/v, 2 h, r.t.; e) I_2; H_2O:MeCN 1:1, 30 min, r.t.. Y = 77% over 3 steps.

Following concentration of the cleavage cocktail and subsequent precipitation in cold diethyl ether, the resulting deprotected linear peptide **54** was obtained, featuring a C-terminal primary amide capping as well as a *N*-acetylamide moiety at the N terminus.Compound **54** was used directly without further purifications in the cyclization step, involving the formation of a disulfide bond between the two Cys side chains. Similarly to the disulfide bond formation protocol used for the synthesis of bicyclic peptides, the linear peptide **54** was treated with an excess of iodine in a 1:1 mixture of H_2O/acetonitrile. Also in this case, the cycle formation was performed under high dilution conditions (2 mM concentration). Finally, the crude mixture was purified using semipreparative RP-HPLC and freeze-dried, yielding pure cyclic peptide **39** with a 77% yield over 3 steps.

2.3 *In vitro* biological tests

2.3.1 Integrin Receptors Competitive Binding Assays

In collaboration with Dr. Daniela Arosio (Italian National Research Council), compounds **36**-**39** were tested for their ability to compete with endogenous extracellular matrix proteins for their binding to isolated $\alpha_V\beta_3$ and $\alpha_5\beta_1$ integrin receptors. The assay was based on the simultaneous incubation of synthetic ligands in serial dilutions and a fixed concentration of biotinylated endogenous ligand (i.e. vitronectin for $\alpha_V\beta_3$ and fibronectin for $\alpha_5\beta_1$) into 96-well plates coated with the integrin receptors. Evaluation of the residual bound endogenous ligand was then performed by UV measurements (i.e. addition of HRP-streptavidin and incubation with a substrate reagent solution for colorimetric quantification), followed by the nonlinear regression of the points and calculation of the IC_{50} for the desired compound. The resulting IC_{50} are reported in Table 3.

Table 3 IC_{50} of compounds **36**-**39** determined by competitive binding assays against isolated $\alpha_V\beta_3$ and $\alpha_5\beta_1$ integrin receptors.

	IC_{50}[a] [nM]	
Compounds	$\alpha_V\beta_3$	$\alpha_5\beta_1$
*cyclo*RGDfV (**2**)	1.60 ± 0.90	105 ± 5
RGD-2C-RGD (**36**)	1.02 ± 0.68	263 ± 115
RGD-2C-RβAD (**37**)	5.97 ± 3.77	1000 ± 27
RβAD-2C-RβAD (**38**)	811 ± 59	>100 000
RGD-2C (**39**)	6.39 ± 0.37	728 ± 142

[a] IC_{50} values were determined as the concentration of compound required for 50% inhibition of biotinylated vitronectin binding to integrin $\alpha_V\beta_3$ or biotinylated fibronectin binding to integrin $\alpha_5\beta_1$, as estimated by using GraphPad Prism software. All values are the arithmetic mean ± the standard deviation (SD) of triplicate determinations.

In binding assays against the $\alpha_V\beta_3$ heterodimer, the IC_{50} of dimeric bicyclic RGD ligand **36** lied within the low nanomolar range, comparable to the binding affinity shown by reference peptide *cyclo*RGDfV **2**. On the other hand, **36** showed higher integrin-binding potency than the monomeric bicyclic peptide **37**. As expected, the bicyclic negative control **38** demonstrated a very low integrin binding affinity, with sub-micromolar IC_{50} value. Finally, the comparable IC_{50} values exhibited by the monomeric bicyclic peptide **37** and the monocyclic peptide **39** highlights that the bicyclic scaffold is not a strict requirement for optimal binding profiles.

The trend observed with $\alpha_V\beta_3$ was substantially maintained in the analogue competitive binding assays to purified integrin $\alpha_5\beta_1$. In this case, the reference peptide *cyclo*RGDfV **2** proved the best $\alpha_5\beta_1$ integrin binder among the tested molecules and compound **36** showed a higher binding affinity than the monomeric bicyclic peptide **37**, monocyclic peptide **39** and, by far, bicyclic negative control **38**. Similarly to what observed in the case of $\alpha_V\beta_3$, monocyclic compound **39** and monomeric bicyclic peptide **37** proved comparable integrin-binding affinity.

Overall, these data indicated that the dual presentation of the RGD pharmacophore in dimeric bicyclic peptide **36** leads to an enhanced integrin binding affinity. The effect of multivalency on the binding of compound **36** to the related biological target could be estimated by the calculation of the relative potency Rp and the *Rp/n* values.[108] The *Rp* parameters can be obtained by dividing the measured IC_{50} values of the monomeric bicyclic peptide **37** and the IC_{50} of the dimeric bicyclic peptide **36**. On the other hand, *Rp/n* values are calculated by dividing the previously obtained *Rp* of the dimeric ligand by the valency (*n*) of the ligand (i.e. 2). To verify that the resulting enhanced affinity is not due to a mere *n*-fold increase of the pharmacophore concentration in solution, but to a real multivalent effect, the *Rp/n* value of the multimeric ligand must be higher than 1. Of note, calculated *Rp/n* values of compound **36** for both $\alpha_V\beta_3$ and $\alpha_5\beta_1$ integrin receptors are higher than 1 (*Rp/n* = 2.93 for $\alpha_V\beta_3$ and *Rp/n* = 1.90 for $\alpha_5\beta_1$), accounting for the multivalent effect of this dimeric bicyclic ligand. Conceivably, this multivalent effect is mainly ascribed to a "rebinding effect" (Figure 18B) rather than a "cluster effect" (Figure 18A). This conclusion can be outlined by the fact that integrins bear a single binding site, as well as by considering the size of the condensed bicyclic structure, which is too small to engage two copies of integrin receptors.

2.3.2 *In vitro* biological assays on U373-MG glioblastoma cells

Encouraged by the promising results obtained in competitive binding assays, we wondered if the enhancement of the binding affinity may correlate to an enhancement of the biological activity of the newly designed ligand **36**. In collaboration with Prof. Mayra Paolillo (University of Pavia), we subjected bicyclic compounds **36**–**38** to a series of biological assays using the U-373 MG human glioblastoma cell line (Figure 24), which overexpresses both $\alpha_V\beta_3$ and $\alpha_V\beta_5$ integrin receptors.[109] Since integrins are involved in cell adhesion processes, we initially aimed at investigating the ability of peptides **36**–**38** to induce cell detachment.

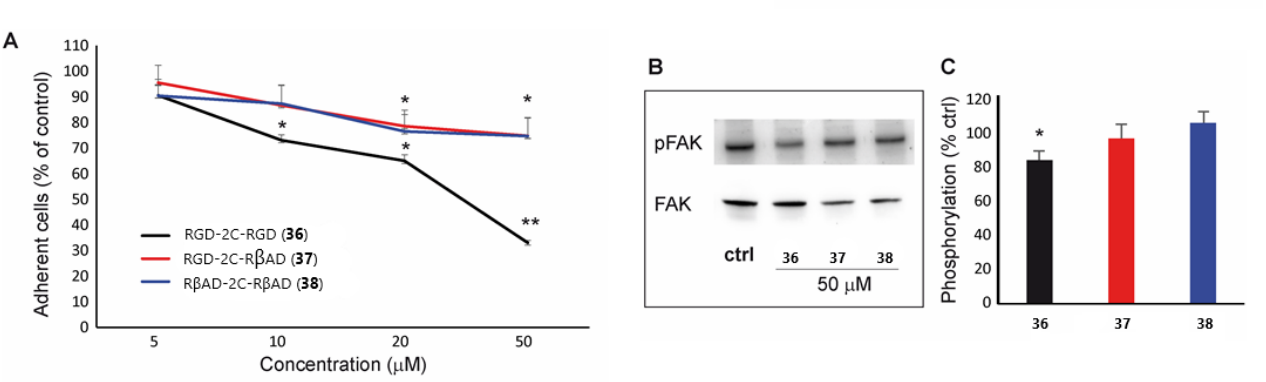


Figure 24 A) Evaluation of cell detachment induced by compounds **36**-**38** on U373-MG glioblastoma cell line; B) Western blot analysis of the effect induced by compounds **36**-**38** on Focal Adhesion Kinase (FAK) phosphorylation; c) Densitometric analysis of Western blot data in figure 22B. pFAK: phosphorylated FAK.

U-373 MG cells were incubated for 48 hours with solutions of bicyclic peptides **36**-**38** at different concentrations (5, 10, 20, and 50 μM). After the incubation, dimeric bicyclic RGD ligand **36** demonstrated to induce cell detachment, with higher efficacy at 50 μM, notably without apparent

toxicity in detached cells. On the other hand, the monomeric analogue **37** was much less effective than **36**, displaying negligible effects on adherent cells similar to negative control **38** (Figure 24A). The high ability of compound **36** to induce cell detachment was then ascribed to its binding affinity to $\alpha_V\beta_3$ integrin with consequent alteration of the integrin signaling pathway and downstream phosphorylation of specific kinases (e.g. Focal Adhesion Kinase, FAK).[110] To confirm the effect of integrin deactivation in the intracellular signaling cascade, we monitored FAK phosphorylation in cells, treated with 50 µM concentration of bicyclic peptides **36**-**38**. FAK phosphorylation was detected by Western blot analysis on U-373 MG cells incubated with compounds **36**-**38** for 48 hours. A significant decrease of FAK phosphorylation was detected only when cells were incubated with dimeric bicyclic ligand **36** (Figure 24B and C). The effect of compound **36** on inhibition of FAK phosphorylation was less pronounced than the effect observed in cell detachment assays. This difference could be ascribed to the involvement of FAK in other signal transduction pathways,[111] not directly connected to $\alpha_V\beta_3$ integrin signaling.

2.4 Conclusions

In this project, we achieved the design and synthesis of a new peptide integrin ligand **36**, characterized by an unprecedented double presentation of the RGD tripeptide into a macrocyclic scaffold, with a bicyclic structure conferred by disulfide bridging. Negative controls **37**-**39** were also prepared and tested in parallel to the lead compound with the aim to rationalize the biochemical and biological effects of the dimeric pharmacophore presentation. Among the prepared compounds, dimeric bicyclic RGD peptide **36** demonstrated a higher binding affinity for both $\alpha_V\beta_3$ and $\alpha_5\beta_1$ proteins, not only with respect to the negative control **38**, but also compared to the monomeric bicyclic analogue **37** and the monocyclic RGD peptide **39**. The superior binding affinity of **36** to both $\alpha_V\beta_3$ and $\alpha_5\beta_1$ integrin receptors led to a strong interference with the integrin signaling pathway of the U373-MG glioblastoma cell line, resulting into a marked cell detachment and a decrease of the FAK phosphorylation. The design of the dimeric bicyclic peptide **36** takes advantage of a multivalent kinetic effect, mainly ascribed to the increase of the local concentration of the pharmacophoric portion in the surroundings of the integrin receptors. Since the dimensions of compound **36** are smaller compared to the usual multivalent scaffolds, the rationale behind the multivalent effect could be found in a “rebinding” effect that cause an increase of the residence time of the RGD tripeptide into the binding pocket. Moreover, the enhanced binding affinity of compound **36** could be also ascribed to weak additional interaction between the second RGD copy of the bicyclic structure and the charged residues on the surface of the integrin receptor. The results presented in this Chapter open for a general use of the dimeric bicyclic design. Indeed, the purposed design could be also applied to other peptide systems, in order to enhance the residence time of small peptides in the binding pocket of clinically-relevant proteins.

2HB-PEG modules as portable tags for the engagement of Lys ε-amino groups

Part of this chapter was published in the following article:

- G. Sacco, S. Stammwitz, L. Belvisi, L. Pignataro, A. Dal Corso, C. Gennari. *Eur. J. Org. Chem.* **2021**, *2021*, 1763-1767.

3.1 Reversible covalent interactions

The identification and optimization of structural interactions that a small molecule ligand can engage with the target protein is a core component of drug discovery processes. In particular, the formation of stable drug-protein complexes is often crucial for the generation of effective therapeutic agents. According to the nature of ligand-protein interactions, it is possible to classify the ligands in two main classes (Figure 25).

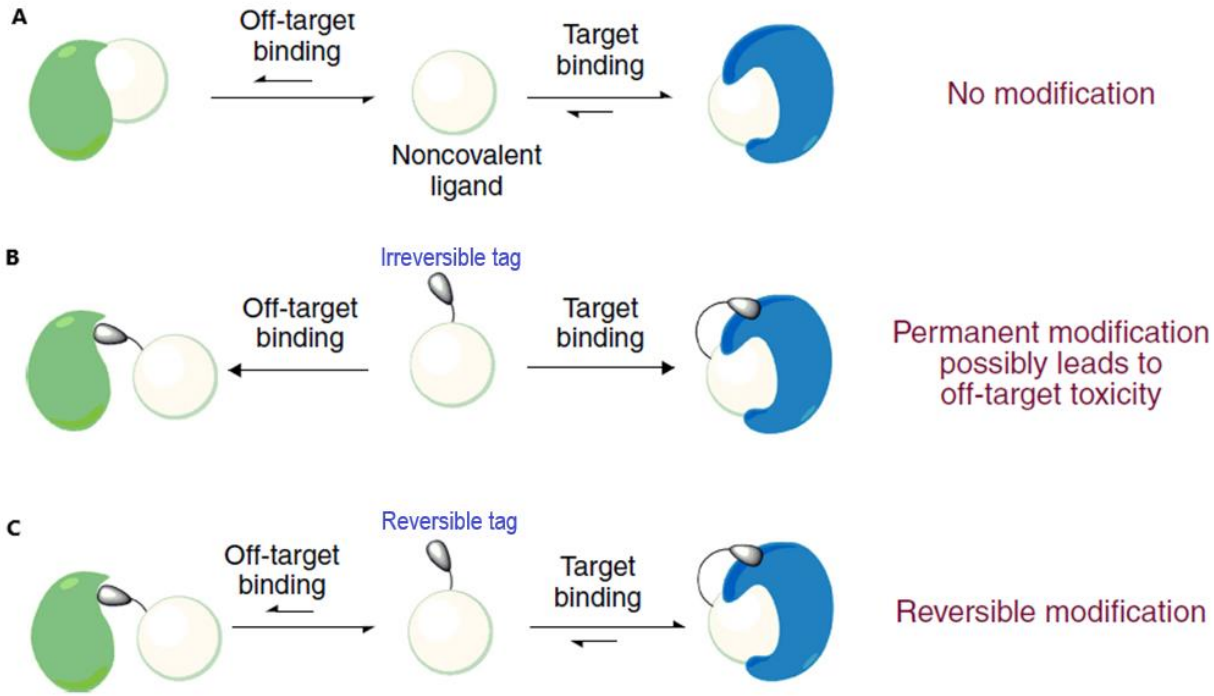


Figure 25 a) Interaction of a noncovalent ligand with both a target protein and an undesired "off-target" protein. In both cases, the ligand is in equilibrium with its complexes with both proteins. Considering the dynamic equilibria involved, the population of unbound and bound ligand is governed by the intrinsic ligand-protein affinities (i.e. the stability of each complexes) as well as by the ligand and protein concentrations; b) Interaction between an irreversible covalent ligand and a target and an off-target protein. In both cases, proteins are modified permanently, leading to both a therapeutic effect and a negligible or, at worst, a toxic effect; c) Interaction between a reversible covalent ligand and both a target and an off-target protein. The reversible interaction between ligand and proteins enhances the stability of the desired ligand-protein complex and decreases off-target binding (Adapted from reference 114).

A first class includes noncovalent ligands (Figure 25A): the interaction between a ligand of this class and a protein is based on an array of weak non-covalent interactions (e.g. hydrophobic interactions,

electrostatic interactions and hydrogen bonding) that, taken together, may provide a high stability to the resulting adduct. An example, already reported in this Thesis, is the interaction between the RGD pharmacophore in cyclopeptides and the $\alpha_V\beta_3$ integrin receptor: the ligand-protein complex is stabilized by *i)* the presence of an electrostatic interaction between the guanidyl group of the Arg (R) and two carboxylic moieties of Asp150 and 215 of the α_V subunit and *ii)* the Asp carboxylic moiety coordination of the MIDAS centre in the β_3 subunit. While the reversible nature of the ligand-protein complex may be the cause of suboptimal therapeutic effects, a second class of drugs represented by covalent ligands is capable of forming a strong covalent bond with the target protein. This class of compounds demonstrated strong therapeutic effects in different cases: one third of the validated enzyme targets are associated to an approved covalent ligand, which has been investigated as a drug.[112] One of the most famous examples of covalent drug is the acetylsalicylic acid, i.e. the well-known Aspirin. Acetylsalicylic acid acts through an irreversible acetylation of cyclooxygenases (COXs), a group of enzymes involved in the inflammatory process and thrombotic events. Due to the reactivity of these covalent inhibitors, it is possible that the non-target proteins are chemically-modified by the drug, leading to either a decrease of therapeutic effects or, at worse, side-toxicity (Figure 25B).[113] A good balance between the "drug-like" noncovalent compounds and the high potency of irreversible covalent ligands is represented by reversible covalent ligands (Figure 25C).[114] In a reversible ligand-protein complex, an intermolecular covalent bond may be formed to enhance the residence time of the ligand into the target binding pocket and to prevent the irreversible modification of off-target proteins. These reversible covalent ligands are composed of two main chemical entities (Figure 26): a noncovalent portion and a reversible reactive tag.[115] Similar to a traditional non-covalent drug, the noncovalent portion is capable of interacting with the binding pocket of the target protein, engaging a number of "canonical" non-covalent interactions. On the other hand, the reversible tag is an electrophilic moiety that, in presence of a nucleophilic side-chain of an amino acid, lead to the formation of a covalent bond that is in equilibrium with its dissociated form.

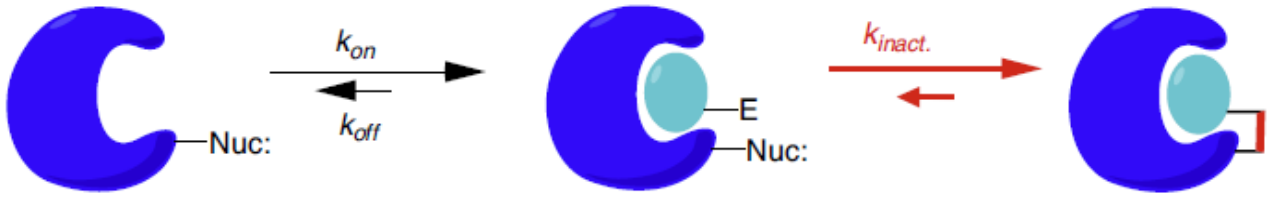


Figure 26 Mechanism of interaction between a reversible covalent ligand and a target protein. "Nuc:" is a nucleophilic side chain of an amino acid such as thiol, alcohol and amine. "E" is an electrophilic tag that is capable to react with the nucleophile and form a reversible covalent bond (Adapted from reference 115).

The nucleophilic side chains usually tagged with reversible covalent ligands are the thiol group of cysteine residue, the alcoholic functionalities of serine, threonine and tyrosine side chains, as well as the ε-amino group of lysine residues. A short list of electrophilic tags used for the engagement of different amino acid side chains is reported below (Figure 27).

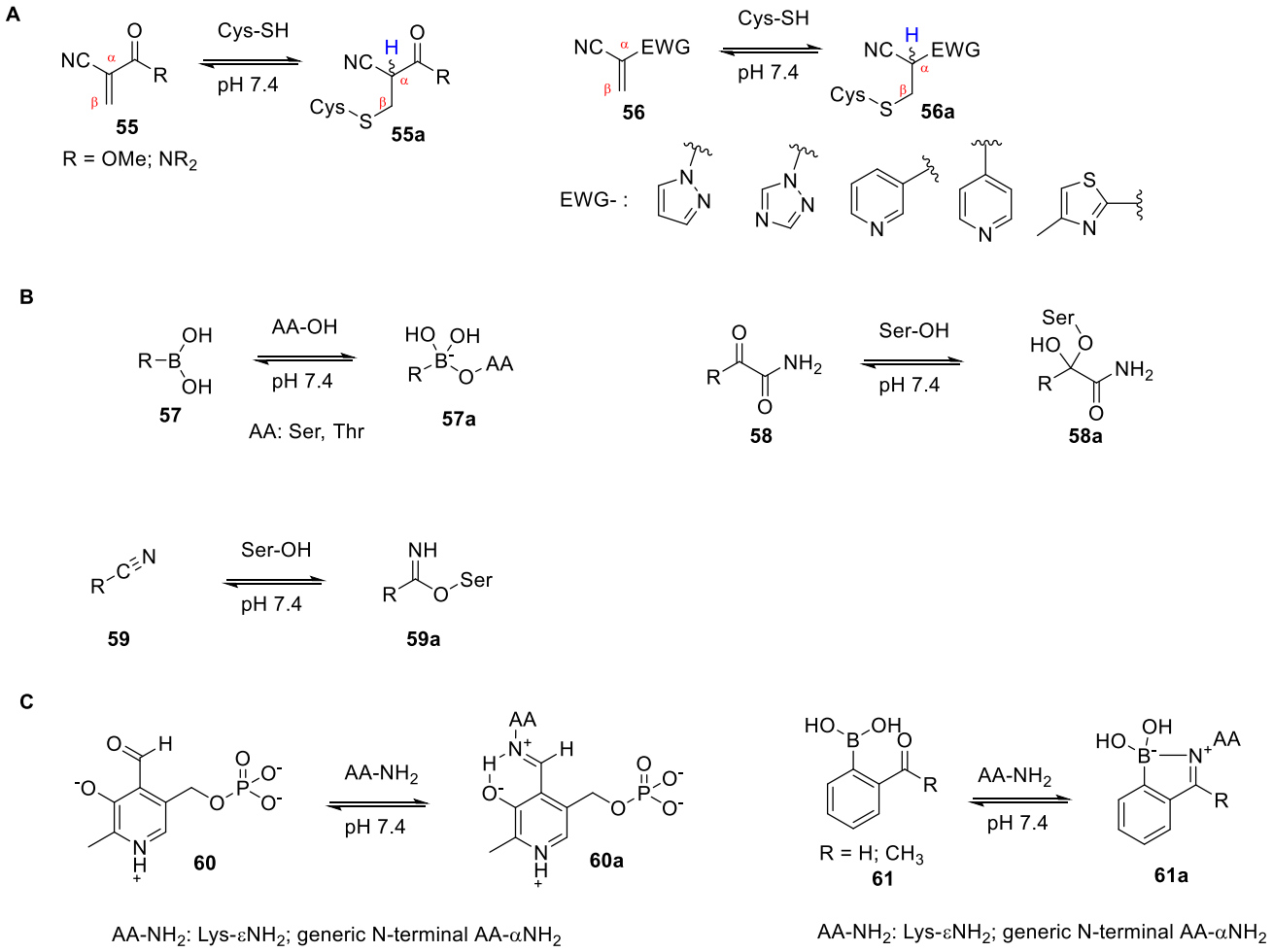


Figure 27 Short selection of electrophilic tag used for reversible covalent interactions with: A) Cysteines; B) Hydroxyl-bearing amino acids (Ser, Thr); C) Amine-bearing amino acids (Lys-εNH$_2$, generic N-terminal AA-αNH$_2$)

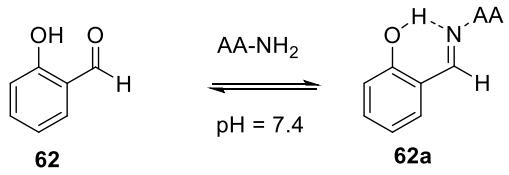
Many electrophilic tags have been developed for cysteine engagement. In particular, these reagents rely on the "soft" nucleophilicity of the thiol group to perform a Michael addition on a proper α,β-unsaturated system. Taunton and coworkers developed a cyanoacrylamide (or cyanoacrylate, compound **55**, Figure 27A) moiety as a cysteine-binding tag.[116] While cyanoacrilamides are excellent Michael acceptors, the so-formed Michael adduct **55a** is prone to a retro-Michael reaction with subsequent elimination of the thiol group, which opens to the use of cyanoacrylamides as reversible-covalent tags. The same group also designed acrylonitriles **56** (Figure 27A) endowed with different electron withdrawing groups.[117] In particular, they found an inverse correlation between the computed proton affinity of the acrylonitrile carbanion and the rate of the β-elimination (retro-Micheal) process, which gives access to the reversibility. Thus, for this system it is possible to tune the intrinsic reversibility of the thiol-Micheal adduct and predict the reactivity of the acrylonitrile with the thiol group of the desired cysteine.

Concerning alcohols, boronic acids (compound **57**, Figure 27B) have been reported as electrophilic tags for the engagement of serine and threonine residues. The boronic acid mechanism of action is based on the formation of an ate complex with hydroxy groups (compound **57a**, Figure 27B), that can be reverted in aqueous solution.[118] It is important to observe that boronic acids studied for

reversible covalent interactions with hydroxyl groups of Ser and Thr are typically present in the active sites of enzymes, rather than on the protein surface. Moreover, the formation of an ate complex corresponds to the formation of a tetrahedral adduct around the B atom, which is a valuable strategy to mimic the transition state of amide hydrolysis generated into a protease binding pocket. In this case, one of the typical examples is bortezomib, an FDA-approved drug for the treatment of multiple myeloma.[119] Bortezomib is a dipeptide-based boronic acid that interacts with the proteasome thanks to the formation of an ate complex between the boronic acid residue and the side chain of a threonine located into the enzyme active site. Another tag developed for the engagement of hydroxyl-bearing amino acids is the α-ketoamide moiety (compound **58**, Figure 27B).[114] The carbonyl group of α-ketoamide interacts with a hydroxyl group of serine located into the binding site to form an hemiketal (compound **58a**, Figure 27B). Similarly to the boronic acid mechanism, the hemiketal **58a** is reversible at physiological conditions and it is able to mimic a tetrahedral transition state. Telaprevir and boceprevir, two reversible covalent ligands approved by the FDA for the treatment of hepatitis C, are characterized by the presence of an α-ketoamide tag. Another functional group capable of engaging serine residues into serine hydrolases is the nitrile group (compound **59**, Figure 27B).[120] In the presence of a hydroxyl group into the enzyme active sites, nitrile-bearing ligands form an imidate species (compound **59a**, Figure 27B) that, in physiological conditions, is in equilibrium with the nitrile. Also this approach led to the development of two FDA-approved drugs, vildagliptin and saxagliptin, used for the treatment of diabetes.[116] Moving on to amine-reactive tags, imine formation is one of the best-known reversible reactions. The equilibrium between the imine and the free amine is regulated by the presence of water: for most of the imines, the physiological conditions move the equilibrium to the free amine, resulting in a very unstable imine bond. To overcome the usual thermodynamic instability of imines, nature uses a pyridoxal phosphate (PLP, compound **60**, Figure 27C) coenzyme that bears a phenol and a phosphate group, both in *orto*- to the reactive aldehyde group. While the former stabilizes the imine thanks to an internal hydrogen bond, the latter contributes to the imine overall stability with an electrostatic stabilization of the iminium ion (compound **60a**, Figure 27C).[121] Taking inspiration from nature, chemists started to develop new tags to engage lysines[122] and N-terminal amino acids. Gois and coworkers designed aromatic 2-formylboronic acid and 2-acetylboronic acid tags (compound **61**, Figure 27C) for reversible protein modification.[123] These functionalized benzaldehydes and acetophenones are able to form an imine that, in presence of the boronic acid in *orto*– position, cyclize to form an iminoboronate (compound **61a**, Figure 27C) with enhanced stability towards hydrolysis compared to traditional imines. The reversibility of the iminoboronate was firstly achieved in the presence of competitive molecules like dopamine, fructose and glutathione, which can either displace the amino group from the imine bond or coordinate the boron atom. Later on, Gao and coworkers demonstrated the reversibility of the iminoboronate in physiological conditions, in the absence of nucleophilic competitors.[124] The 2-formylboronic acid tag found immediate application for the design of new potent protein inhibitors:

researcher at AstraZeneca used the iminoboronate approach to develop a new reversible covalent ligand for induced myeloid leukemia cell differentiation protein Mcl-1.[125] In particular, the group modified the structure of the non-covalent ligand with a 2-formylboronic acid tag to engage a non-catalytic lysine residue. The ligand functionalization with a 2-formylboronic acid tag enhanced the ligand binding affinity by two orders of magnitude compared to the noncovalent parent ligand.

Another interesting tag used for the engagement of amino groups is the 2-hydroxybenzaldehyde (2HB, compound **62**, Figure 28) moiety.



AA-NH_2: Lys-εNH_2; generic N-terminal AA-αNH_2

Figure 28 2-hydroxybenzaldehyde (2HB) tag and engagement with a primary amino group

Similarly to PLP, 2HB is able to form an imine bond (compound **62a**, Figure 28) under aqueous conditions, stabilized by an internal hydrogen bond between the *orto*- phenolic group and the N atom. Also the 2HB found application for the development of reversible covalent ligands. Voxelotor (Oxbryta™) is a small molecule recently approved by FDA for the treatment of sickle cell disease. The small molecule prevents the polymerization of mutant hemoglobin (HbS) thanks to the formation of a stabilized imine between the 2HB tag and the free α-amino group of an N-terminal valine residue.[126] Another important application of 2HB is the engagement of ε-amino groups of non-catalytic lysine residues. Neri and coworkers demonstrated that the incorporation of a 2HB tag into a complementary DNA strand with a DNA-bound non-covalent ligand enhances the stability of the ligand-protein complex thanks to the engagement of a lysine residue close to the ligand binding site.[127] Ligand affinity enhancement was demonstrated not only through annealing of ligand and 2HB conjugates with oligonucleotides, but also by connecting the 2HB and the ligand in the same small molecule structure. In particular, the group functionalized a model benzamidine ligand for urokinase-type plasminogen activator (uPa) with the 2HB tag: the resulting reversible covalent ligand proved a 15-time higher inhibition of uPa than parent benzamidine ligand. These experiments demonstrated that the 2HB moiety can enhance the ligand binding affinity, provided that lysine residues proximal to the ligand binding site are available. In particular, this reversible-covalent approach may be applied in several contexts, due to the fact that Lys are considered “high-frequency” amino acids,[128] with many ε-amino groups exposed to the solvent and in the outer layers of proteins and in the proximity of ligand binding sites. For instance, Figure 29[129] reports crystallographic data for a group of ligand-protein complexes, in which Lys ε-amino groups are highlighted in blue. In all cases, at least two lysine groups are available, which may be exploited for the development of reversible covalent ligands.

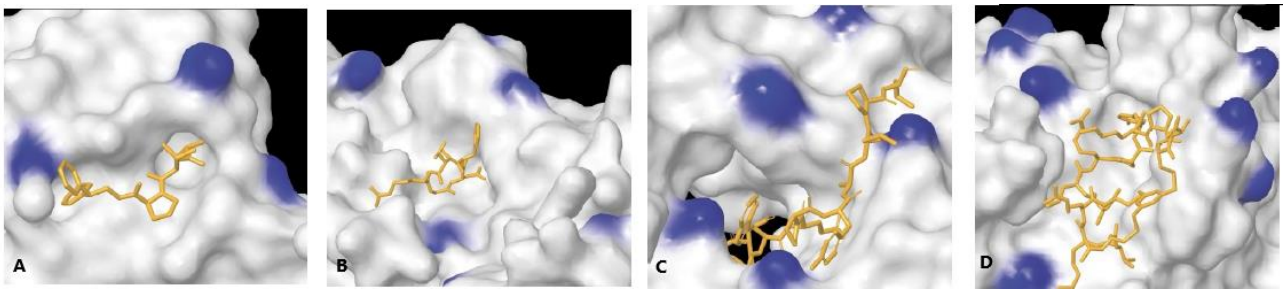


Figure 29 Crystal structures of some literature-reported ligand-protein complexes. In each crystal structure, solvent-exposed ε-amino group of Lys residues are highlighted in blue. Proteins reported are: A) ML–IAP; B) integrin $\alpha_V\beta_3$; C) IL1R1; D) RBBP4.

In general, the design of a reversible covalent ligand endowed with the 2HB moiety must take into account three fundamental aspects:

1) the presence and structural features of taggable lysine residues in the proximity of the ligand binding site;

2) the identification of a spacer, with proper length and flexibility, that allows the 2HB to easily reach the desired Lys residue;

3) the presence, on the ligand structure, of a conjugation site that allows the chemical connection of the 2HB-spacer module to the ligand.

In this context, the design of functionalized 2HB-spacer modules endowed with different reactive handles would give raise to an useful “chemical toolbox” for the 2HB installation into different protein ligands, paving the way to the design of novel reversible covalent ligands.

3.2 Design and retrosynthetic analysis of the 2HB modules

During my PhD work, I explored synthetic strategies for the preparation of 2HB-spacer modules, aimed at facilitating the development of amine-binding reversible-covalent ligands. The model structure of a generic 2HB module is reported in Figure 30A.

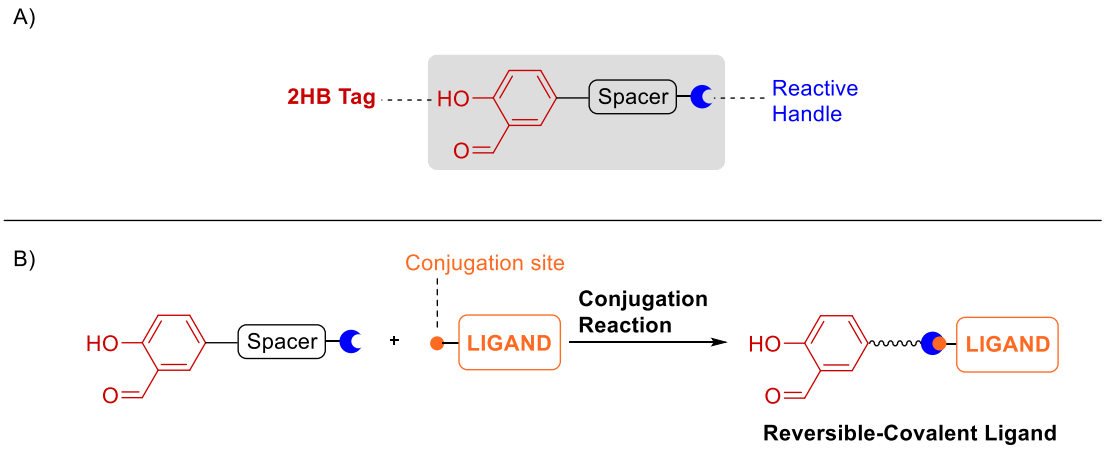


Figure 30 A) General structure of a 2HB module; B) Schematic representation of the reaction between a 2HB module and a generic protein ligand endowed with a proper conjugation site.

The module structure is characterized by three main units: the 2HB tag, a suitable spacer and a reactive handle that, under proper reaction conditions, allows the conjugation of the 2HB tag to a variety of protein ligands (Figure 30B). Polyethylene glycol (PEG) chains represent ideal spacers for the 2HB-ligand connection as they are soluble in water, biocompatible and versatile.[130] On the other hand, the choice of a suitable reactive handle is based on the specific chemical strategy used for the conjugation of the 2HB module to the desired ligand. Among the possible conjugation protocols, the most adopted strategies are cycloaddition reactions[131] involving azide and/or alkyne groups, and amide coupling reactions. Therefore, ideal reactive handles in 2HB-PEG modules are represented by alkyne (as in compound **63**, Figure 31), azide (compound **64**, Figure 31), carboxylic acid (compound **65**, Figure 31) and amine groups (compound **66**, Figure 31). Based on literature reports, the retrosynthetic analysis of 2HB modules relies on the direct alkylation of different hydroxysalicylaldehydes (Scheme 7A).[127,132] Of note, this strategy is characterized by a regioselectivity issue (i.e. the presence of two potentially-reactive hydroxyl groups), which leads to a mixture of products and low isolated yields of the desired mono-functionalized species.

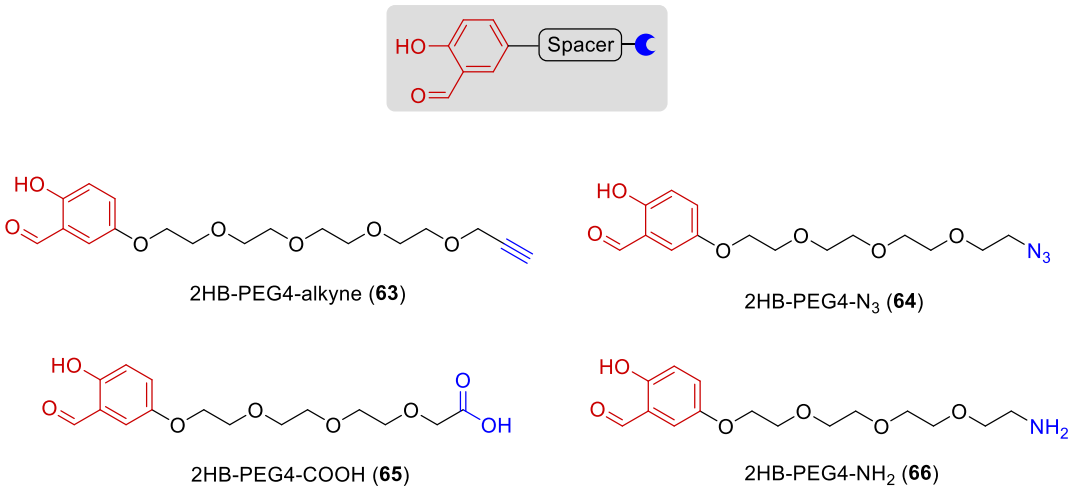
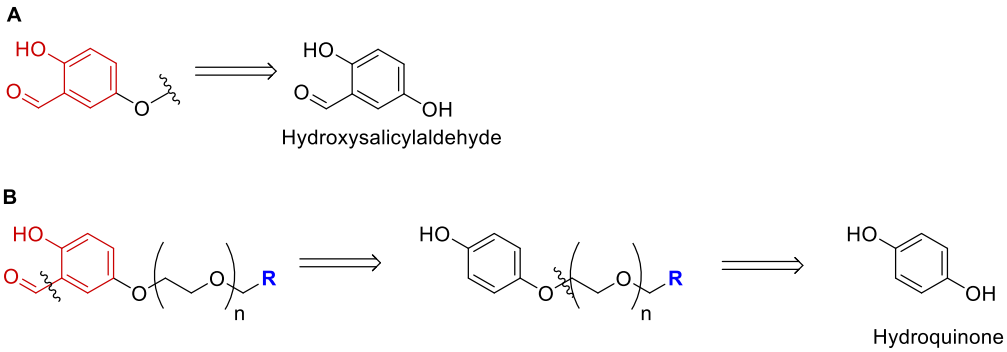
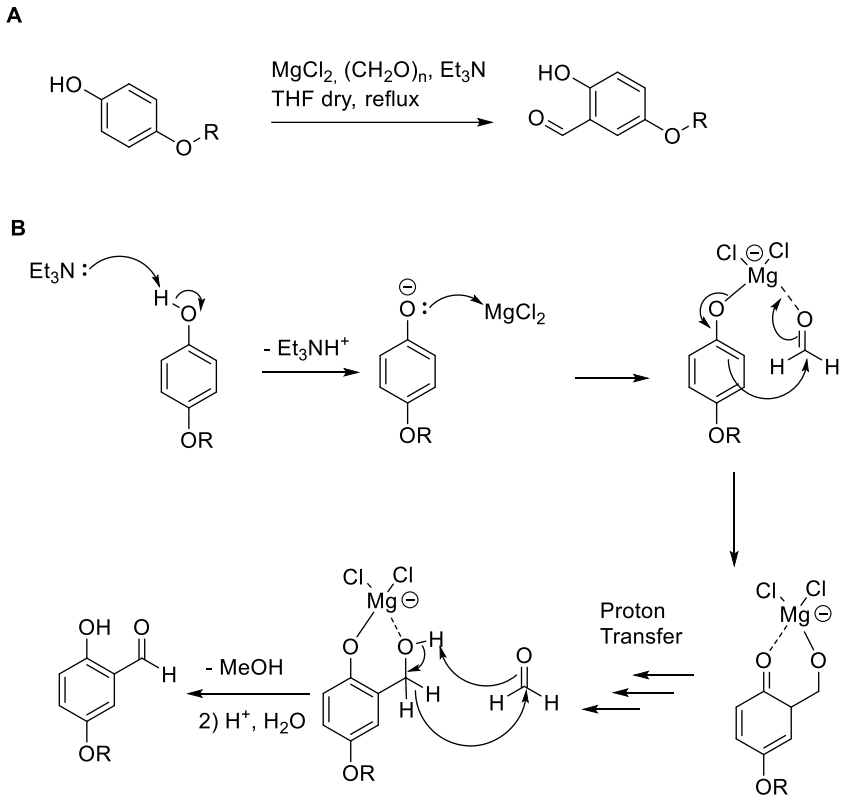


Figure 31 Molecular structures of the four 2HB-PEG modules endowed with an alkyne (compound **63**), an azide (compound **64**), a carboxylic acid (compound **65**) and an amine (compound **66**) as different reactive handles.

A more practical retrosynthetic approach is based on the selective *ortho*-formylation of a phenolic derivative which, in turn, can be generated through non-regioselective mono-alkylation of hydroquinone (Scheme 7B). With this strategy, despite the addition of one synthetic step, the desired 2HB module would be isolated in higher yields. Among the possible formylation protocols, the Skattebøl reaction (i.e. one of the evolution of the original Casiraghi protocol) is selective for phenols, is compatible with many different functional groups and it can be performed with commercially-available and inexpensive chemicals.[133] A generic scheme and the mechanism of the Skattebøl reaction is reported below (Scheme 8).



Scheme 7 A) Retrosynthetic analysis of 2HB modules based on the regioselective alkylation of hydroxysalicylaldehyde; B) New retrosynthetic analysis based on the monoalkylation of the hydroquinone and *o*-formylation of the free phenol.

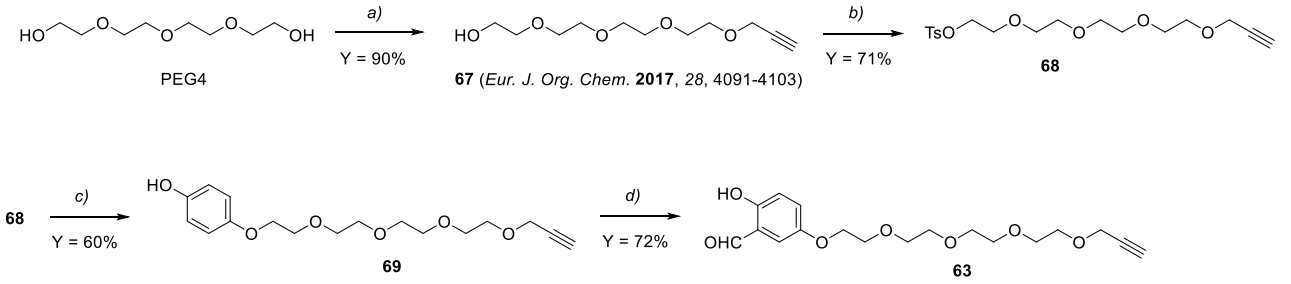


Scheme 8 A) Generic Skattebøl reaction performed on a 4-alkoxyphenol; B) Mechanism of the Skattebøl reaction.

The reaction is performed using a phenol substrate and a stoichiometric excess of paraformaldehyde as source of formaldehyde. Anhydrous magnesium chloride and dry excess triethylamine are used to generate the magnesium phenate. The reaction starts with the deprotonation of the phenol and, in presence of magnesium chloride, formation of the corresponding magnesium phenate. One equivalent of formaldehyde is then coordinated by the Mg counterion of the phenate and an hydroxymethylenation step is performed. After a proton transfer step, the newly formed hydroxymethylene group is oxidized by another equivalent of formaldehyde to the corresponding benzaldehyde, with final formaldehyde reduction to methanol. After acidic work up, the desired 2HB is obtained.

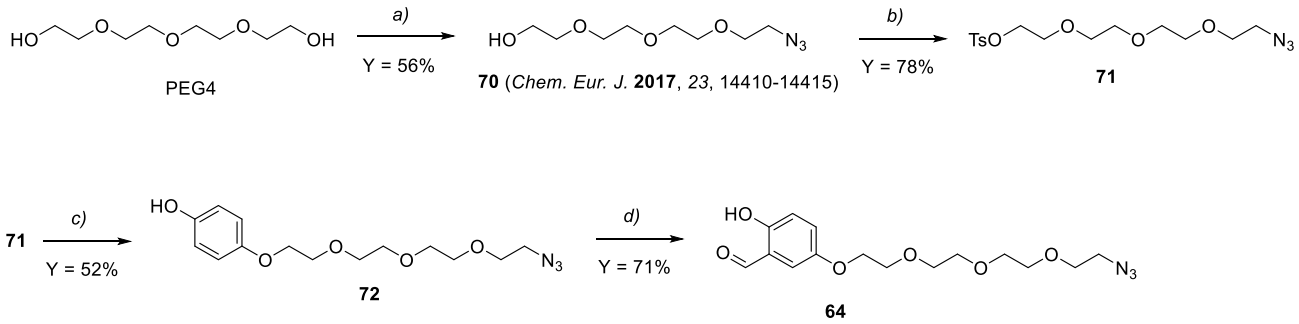
3.3 Synthesis of the modules and coupling to model substrates

The synthesis of the 2HB-PEG4-alkyne module **63** is reported below (Scheme 9).



Scheme 9 REAGENTS AND CONDITIONS: *a)* propargyl bromide, NaH, THF dry, 0°C to r.t.,overnight; *b)* tosyl chloride, Et_3N dry, 4-dimethylaminopyridine, CH_2Cl_2 dry, 0°C to r.t., overnight; *c)* Hydroquinone, K_2CO_3, Bu_4NI, DMF dry, 80°C, overnight; *d)* $MgCl_2$, paraformaldehyde, Et_3N dry, THF dry, reflux, overnight.

The synthesis of module **63** started with a mono-propargylation of commercially available tetraethylene glycol (PEG4), following a literature procedure.[134] PEG4 was initially treated with sodium hydride in dry THF at 0 °C and then propargyl bromide was slowly added to the reaction mixture to obtain compound **67** in high yield. The free alcohol of compound **67** was transformed into a proper leaving group. In particular, treatment of **67** with tosyl chloride in the presence of triethylamine and a catalytic amount of 4-(dimethyamino)pyridine (DMAP) in dry CH_2Cl_2, led to the corresponding tosylate **68**. The latter was used as alkylating agent for the mono-derivatization of hydroquinone: this reaction was performed using compound **68** as limiting reagent, an excess of hydroquinone (3 eq), a small excess of base (K_2CO_3, 1.1 eq) and, to accelerate the substitution reaction, a catalytic amount of tetrabutylammonium iodide (TBAI, 0.1 eq). The reaction was performed in dry DMF and the phenol **69** was obtained in good yield. Compound **69** was used as starting material for the Skattebøl formylation, which led to module **63**. A similar synthetic pathway was followed for the synthesis of the 2HB-PEG4-N_3 module **64** (Scheme 10).

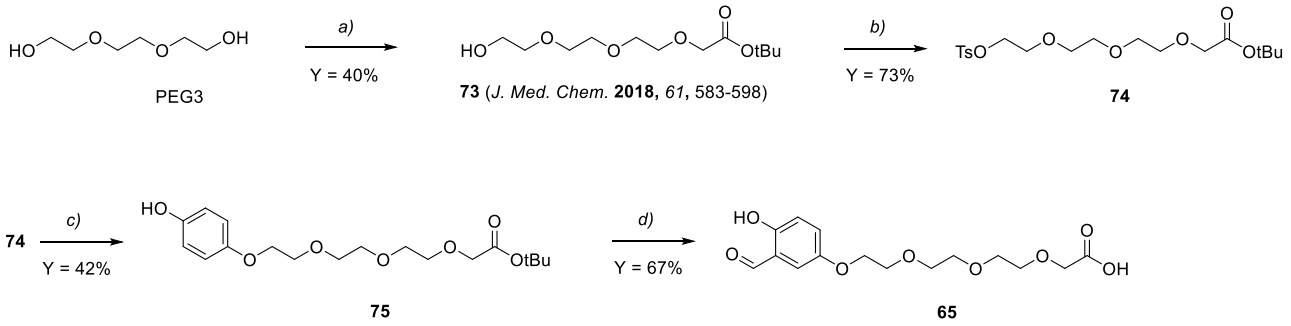


Scheme 10 REAGENTS AND CONDITIONS: *a)* [1] tosyl chloride, Et_3N, 4-dimethylaminopyridine, CH_2Cl_2, 0 °C to r.t., 2h; [2] NaN_3, DMF, 80 °C, overnight; *b)* tosyl chloride, Et_3N, 4-dimethylaminopyridine, CH_2Cl_2, 0 °C to r.t., overnight; *c)* Hydroquinone, K_2CO_3, Bu_4NI, DMF, 80 °C, overnight; *d)* $MgCl_2$, paraformaldehyde, Et_3N, THF, reflux, overnight.

Tetraethylene glycol was converted into azide **70** following a published protocol.[99] In particular, a mono-tosylation reaction was performed using tosyl chloride as limiting reagent, in the presence of triethylamine and a catalytic amount of DMAP in dry dichloromethane. The resulting crude mixture

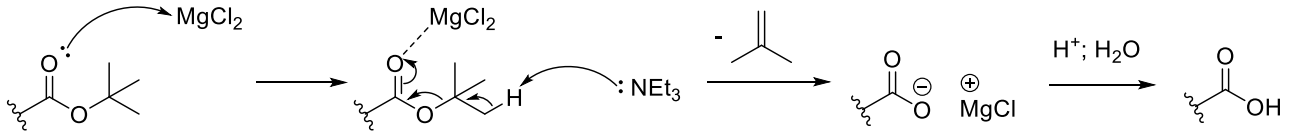
was used as starting material for the nucleophilic substitution, performed using sodium azide in dry DMF. Azide **70** was obtained in good yield and its free hydroxyl group was converted into the corresponding tosylate **71** following the same protocol used for the preparation of compound **68**. Tosylate **71** was used as starting material for hydroquinone mono-alkylation, which led to the ether **72** in good yield. Finally, compound **72** was converted into the desired 2HB-PEG4-N_3 module **64** through Skattebøl formylation.

The 2HB module with the carboxylic acid as reactive handle **65** was synthesized following the synthetic pathway shown in Scheme 11.



Scheme 11 REAGENTS AND CONDITIONS: a) [1] Sodium hydride, DMF, 0 °C to r.t., 1h; [2] t-butylbromoacetate, DMF, 0 °C to r.t, 2h; b) tosyl chloride, Et_3N, 4-dimethylaminopyridine, CH_2Cl_2, 0 °C to r.t., overnight; c) Hydroquinone, K_2CO_3, Bu_4NI, DMF, 80 °C, overnight; d) Anhydrous $MgCl_2$, paraformaldehyde, Et_3N, THF, reflux, overnight;

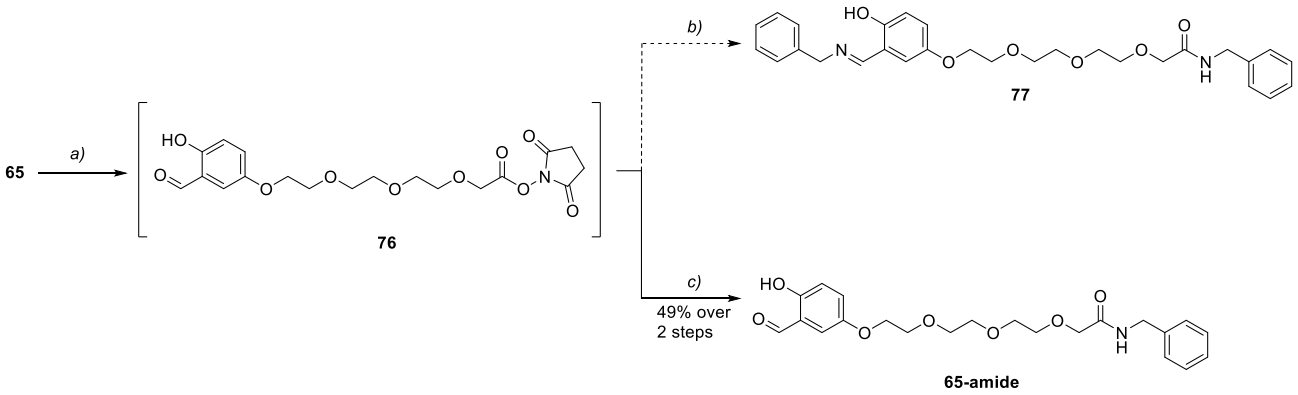
Commercially available triethylene glycol (PEG 3) was functionalized with the carboxylic acid reactive handle protected as *tert*-butyl ester, following a literature-reported procedure:[135] PEG3 was treated with sodium hydride in DMF at 0 °C and *tert*-butyl bromoacetate was added dropwise to the reaction mixture at the same temperature. After reacting for 2 hours at room temperature, work-up and chromatography led to the isolation of *tert*-butyl ester **73**. The free hydroxyl group of **73** was converted into the corresponding tosylate using the same protocol described above for the synthesis of **68** and **71**. As described before, hydroquinone was converted into ether **75**, which was later on subjected to formylation reaction. Interestingly, under the Skattebøl conditions, the phenol *ortho*-formylation occurred together with cleavage of the *tert*-butyl ester group, leading to the direct isolation of 2HB-PEG4-COOH module **65**. The instability of the *tert*-butyl ester in the Skattebøl reaction conditions is conceivably explained by the elimination mechanism shown in Scheme 12.



Scheme 12 Possible mechanism of the *tert*-butyl ester cleavage under Skattebøl reaction conditions.

In particular, the ester functionality is activated by magnesium chloride, which acts as Lewis acid. Moreover, the triethylamine promotes the elimination step according to an E2 mechanism, which is further accelerated by the high temperature. After the isobutylene elimination, the carboxylic acid magnesium salt is formed. While literature reports suggest the compatibility of the 2HB modules **63**

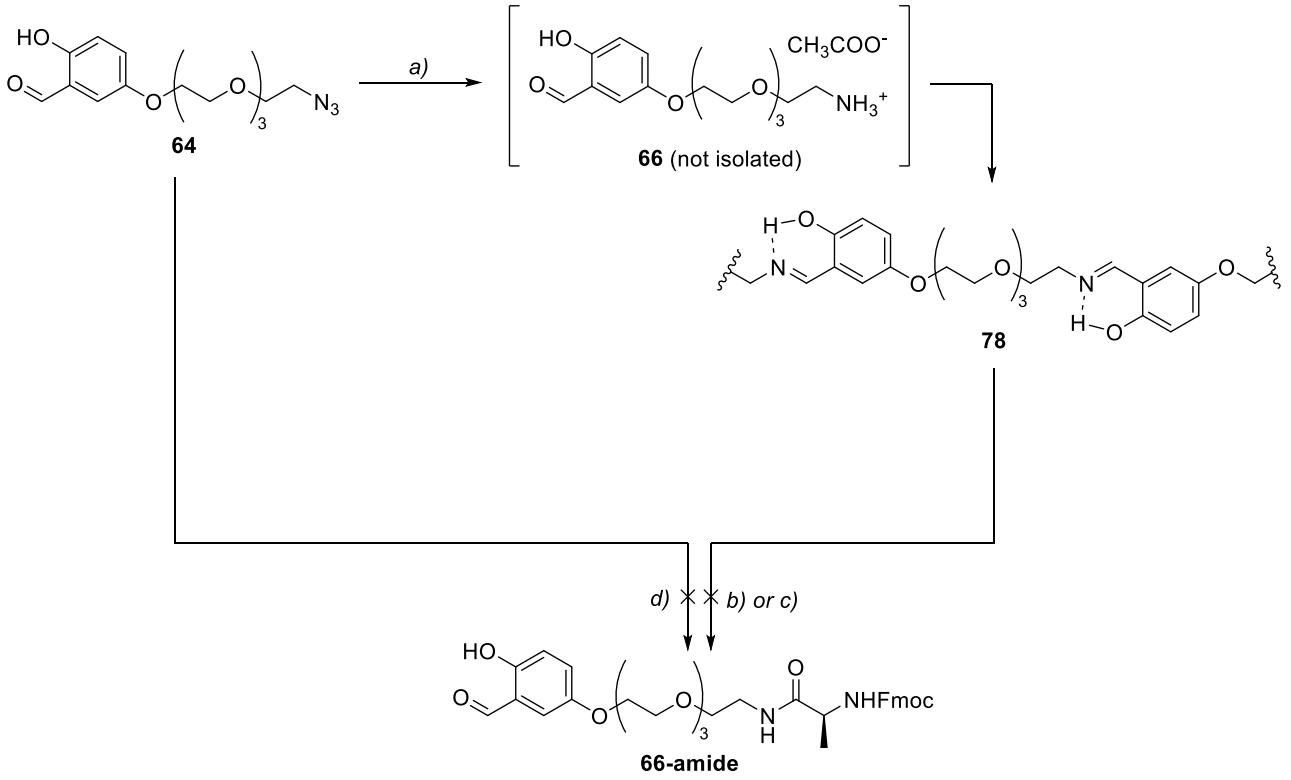
and **64** with the copper-catalyzed azide-alkyne cycloaddition (CuAAC, known as “Click reaction”),[127] we evaluated the reactivity of the 2HB module **65** in a model amide coupling (Scheme 13).



Scheme 13 REAGENTS AND CONDITIONS: *a) N*-hydroxysuccinimide, *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride, CH_2Cl_2, 0 °C to r.t., overnight; *b)* benzylamine(2.4 eq), CH_2Cl_2, r.t, 72 h; *c)* benzylamine (1.1 eq), *i*Pr_2NEt (3.0 eq), CH_2Cl_2, r.t., 1.5 h.

Module **65** was converted into the corresponding *N*-hydroxysuccinimide ester **76** using *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCI, 1.0 eq) and *N*-hydroxysuccinimide (NHS, 1.0 eq) in dichloromethane. After the formation of the active ester species, benzylamine was added to the mixture as model amine. Interestingly, the use of an excess of benzylamine (2.4 eq) led not only to the formation of the desired amide but also to an imine adduct with the 2HB tag (compound **77**). On the other hand, the use of a slight excess of benzylamine (1.1 eq) and a tertiary amine like DIPEA (3.0 eq) led to the formation of the desired product **65-amide** in 49% isolated yield after purification by semi-preparative RP-HPLC.

The synthesis of the 2HB-PEG module **66**, endowed with a primary amino group as reactive handle, was initially attempted through catalytic hydrogenation of azide **64** in the presence of Pd/C in catalytic amounts (10% of supported Pd, 0.1 eq) and 10% of acetic acid in H_2O/THF mixture (Scheme 14). Unfortunately, the desired product was not isolated, as the reaction led to a yellow sticky solid that proved insoluble in common solvents except for water in large volumes. NMR analysis suggested the formation of a mixture of head-to-tail adducts of compound **66** (compound **78** in Scheme 13): the presence of multiple peaks in the chemical shift range of aldehyde and imine protons unveiled the formation of oligomers, with variable numbers of repeating units (Figure 32). A first attempt to de-polymerize compound **78** was performed by oligomer dissolution in water + 0.1% TFA (in order to lower the amine reactivity) and subsequent purification by semi-preparative RP-HPLC but, after the freeze-drying of the desired peak, the undesired oligomer was obtained again. Another depolymerization trial was performed using a model carboxylic acid, such as Fmoc-protected alanine activated as NHS ester, with the aim to obtain the corresponding free module **66-amide**.



Scheme 14 REAGENTS AND CONDITIONS: a) H_2, Pd/C, THF/H_2O 1:1 + 10% AcOH, r.t, 2h; b) Fmoc-Ala-OSu, *i*Pr_2NEt, DMF, r.t, o.n.; c) Fmoc-Ala-OSu, phosphate buffer (pH=7.4), H_2O/DMF, r.t, o.n.; d) H_2, Pd/C, *i*Pr_2NEt, Fmoc-Ala-OSu, THF/H_2O 1:1, r.t, 2h.

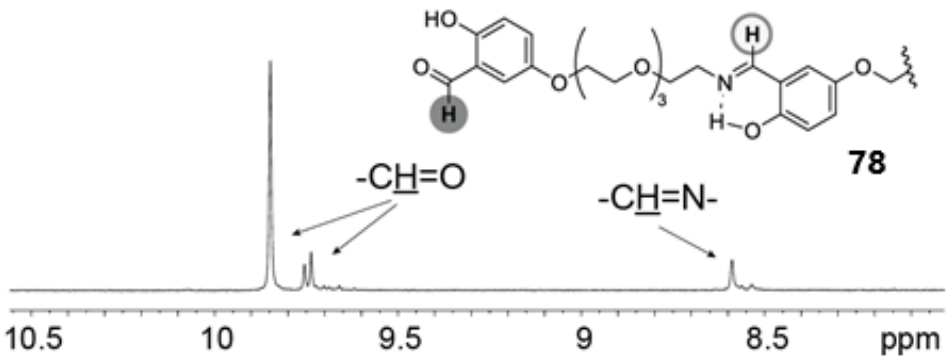
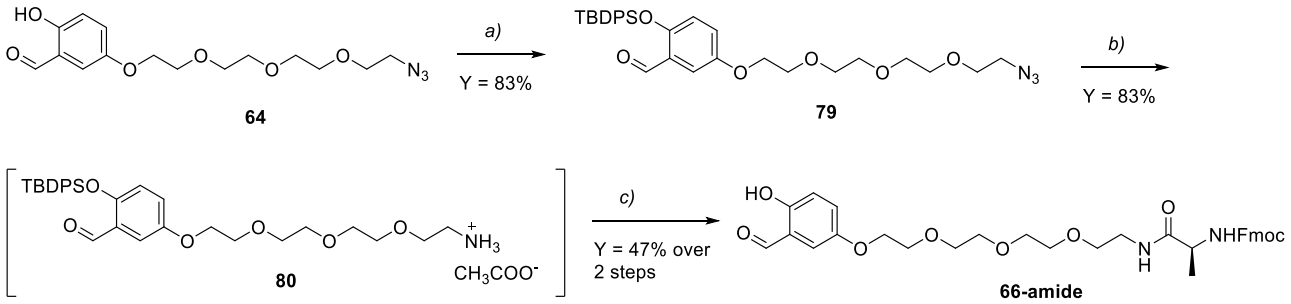


Figure 32 NMR spectrum of the oligomer **78** (solvent: D_2O, ppm range: 8-10.5).

This reaction was run under two different reaction conditions: the first one consisted in the presence of a tertiary amine (DIPEA) in DMF, whereas the second one was carried out in a mixture of DMF and phosphate buffer solution (pH = 7.4). In both cases, the oligomer mixture **78** proved unreactive. Following another strategy, the hydrogenation of azide **64** was also performed in the presence of the Fmoc-Ala-OSu under basic conditions, with the aim to trap the free amine before the imine formation. However, also in this case the oligomer **78** was obtained, no traces of the desired coupling product **66-amide** were detected.

Since the synthesis of compound **66** proved to be troublesome, the use of proper protecting groups was considered to solve the issue. In particular, considering the key role of the *ortho*-phenol in the imine thermodynamic stabilization, we considered the temporary derivatization of this hydroxy group before the synthesis of module **66** and its subsequent coupling to a model substrate. Among the

suitable protecting groups, we selected a silicon-based group such as a *tert*-butyldiphenylsilyl derivative, due to the high stability of this silyl ether under hydrogenation and basic conditions. The synthesis of **66-amide** is reported below (Scheme 15).



Scheme 15 REAGENTS AND CONDITIONS: a) *tert*-butyldiphenylsilyl chloride, Et_3N, 4-dimethylaminopyridine, CH_2Cl_2, 0 °C to r.t., overnight; b) H_2, Pd/C, THF/water, acetic acid, r.t., 2h; c) Fmoc-Ala-OSu, $NaHCO_3$, DMF/water, r.t., overnight.

The 2HB module **64** was reacted with *tert*-butyldihpenylsilyl chloride (TBDPSCl) in the presence of triethylamine and catalytic amounts of DMAP in dry CH_2Cl_2. Under these conditions, the protected module **79** was obtained in 83% isolated yield and used as starting material for the hydrogenation reaction. Following the procedure described above, the acetate salt of the desired amine **80** was obtained: NMR analysis of **80** (Figure 33) clearly showed the absence of oligomers, in strong contrast to what observed with compound **78**, supporting the efficacy of the phenol protection in this synthetic step. Amine **80** was directly subjected to the coupling with Fmoc-Ala-OSu using sodium bicarbonate as the base, in a 1:1 mixture of H_2O/DMF. The reaction was run at 40 °C overnight and, following the purification step performed by semi-preparative RP-HPLC, we realized that the silyl ether protection of the phenol group was cleaved under the slightly acidic purification conditions. This fact led to the isolation of the desired coupled product **66-amide** without the need for a protecting group removal and purification step.

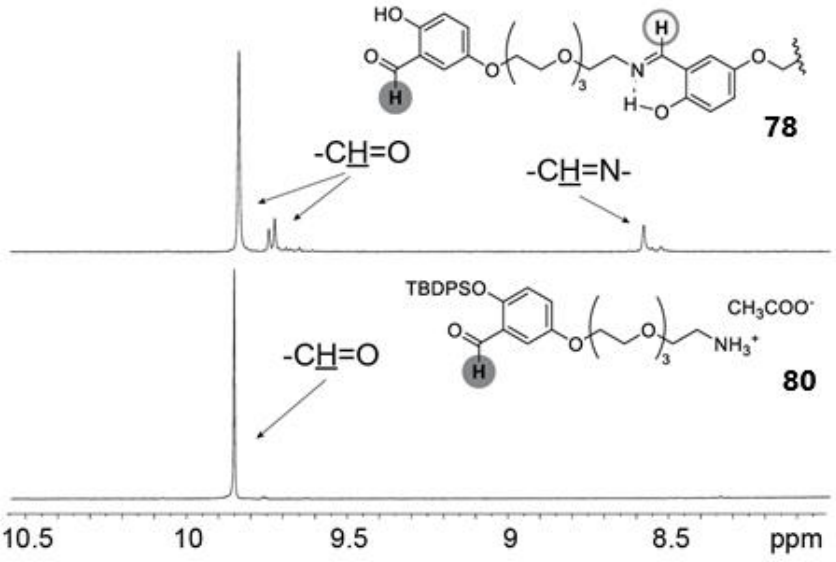


Figure 33 Comparison between the NMR spectrum of the head-to-tail oligomer mixture **78** and the protected amine **80** (solvent: D_2O, ppm range: 8-10.5).

3.4 Conclusions

In this chapter, the synthesis of 2HB modules endowed with (polyethylene)glycol chains as model spacer and different reactive handles (alkyne, azide, carboxylic acid and amine) useful for the bioconjugation to different ligands has been exploited. In particular, we proposed a new synthetic protocol based on the selective *ortho*-formylation (Skattebøl protocol) of the free phenol obtained after a monoalkylation of hydroquinone. The new synthetic pathway revealed to be more efficient compared to the direct alkylation of hydroxysalicylaldehydes, which is characterized by poor regioselectivity and low yields. The protocol led to the straightforward isolation of the 2HB modules with the alkyne, azide and carboxylic acid (compounds **63**-**65**). On the contrary, the synthesis of the 2HB module with the amine as reactive handle (compound **66**) proved more troublesome due to the high reactivity of the 2HB tag towards imine formation with primary amines. In order to solve this issue, the silylation of the free phenolic group of azide **66** (i.e. compound **63**) revealed to be a successful strategy: not only we were able to couple the protected amine **80** to a model carboxylic acid, but also, during the final purification of the coupled product, the removal of the silyl ether protecting group at the phenolic OH took place. Therefore, the desired coupled module **66-amide** was isolated without the need of a dedicated desilylation step. Ideally, the synthetic pathways reported in this chapter may offer various opportunities to conjugate the 2HB modules to different types of ligands (small organic molecules, peptides, peptidomimetics etc.), aiming at the reversible-covalent engagement of Lys ε-amino groups in the target proteins.

2HB-RGD conjugates as reversible covalent $\alpha_V\beta_3$ integrin ligands

4.1 Design of 2HB-RGD ligands

In Chapter 3, the use of 2HB as electrophilic tag to engage ε-amino group of lysine residues was introduced as a general strategy to enhance the binding affinity of small ligands for their cognate receptors. Building on the long-standing expertise of our group in integrin ligands, we decided to develop this strategy in the context of $\alpha_V\beta_3$ integrin ligands. As shown in Figure 34, the structure of a reversible covalent ligand is characterized by the presence of three components. In the context of integrin ligands, a peptide bearing the RGD sequence is responsible for the interaction with the binding pocket of the $\alpha_V\beta_3$ integrin receptor. The resulting interaction contributes to the overall binding affinity through non-covalent interactions between the ligand and the receptor. The 2HB tag is the ligand portion that engages the ε-amino group of a specific lysine residue in a stabilized imine bond, providing a covalent contribution to the total binding affinity. The RGD peptide and the 2-HB tag are joined together by a suitable linker, which should possess a well-defined length and flexibility to display the 2-HB tag in the correct position.

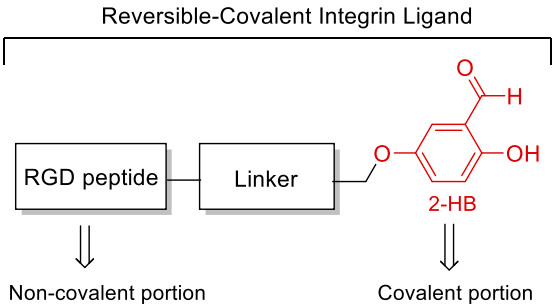
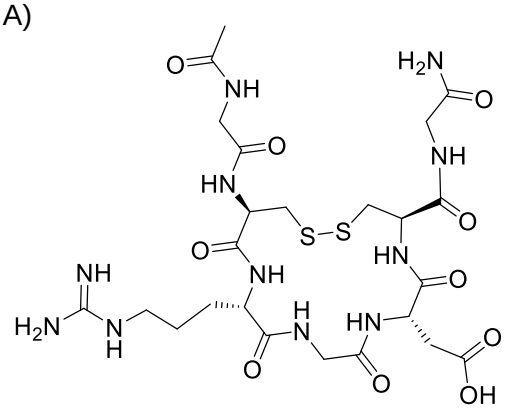


Figure 34. Schematic representation of a generic 2HB-RGD ligand. The RGD peptide is the non-covalent portion of the whole ligand. The linker connects the RGD peptide and the 2-HB tag, which represents the covalent Lys-binding portion.

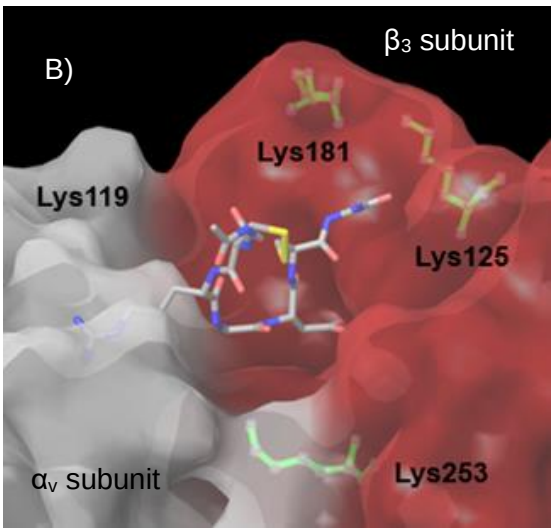
First of all, it is reasonable to assume that the design of effective reversible covalent ligands should start with a careful protein structure analysis and identification of lysine residues proximal to the ligand binding site. To this end, we evaluated the $\alpha_V\beta_3$ structure and the presence of accessible lysine side chains by performing a molecular docking analysis of the cyclic peptide **39** in the integrin binding pocket[84](see Figure 35A and Chapter 2). The outcome of this *in silico* study is reported in Figure 35B: the ligand binding site is surrounded by four lysine residues (Lys119 in the α_V subunit and Lys125, Lys181 and Lys253 in the β_3 subunit), which are potentially available to engage the 2HB tag in reversible covalent imine bonds. In particular, the Lys proximity was assessed by measuring the spatial distance between the ε-amino group of each lysine residue and the capped N and C termini of compound **39**. Since the N-end of **39** is capped with an acetyl group, we measured the distance between each lysine ε-amino group and the methyl group of the acetamido moiety. On

the other hand, the distance between the ligand C-side and each lysine residue was measured starting from the nitrogen atom of the C-terminal primary amide. The measured distances are reported in Figure 35C. Among the Lys residues, the closest ε-amino group to the N-side of peptide **39** belongs to Lys253 in the β_3 subunit, with a distance between 7 and 14.4 Å. As for the **39** C terminus, Lys125 in the β_3 subunit proved the most accessible residue, located at a distance between 7.5 and 8.5 Å from the ligand.

A)



B)



C)

CH_3 lig - NH_3^+ Lys125β_3	15.9-21.8 Å
CH_3 lig - NH_3^+ Lys181β_3	24.5-28.5 Å
CH_3 lig - NH_3^+ Lys253β_3	7-14.4 Å
CH_3 lig - NH_3^+ Lys119α_V	24.9-27 Å

NH_2 lig - NH_3^+ Lys125β_3	7.5-8.5 Å
NH_2 lig - NH_3^+ Lys181β_3	16.7-17.2 Å
NH_2 lig - NH_3^+ Lys253β_3	12.8-13.8 Å
NH_2 lig - NH_3^+ Lys119α_V	20.8-22.2 Å

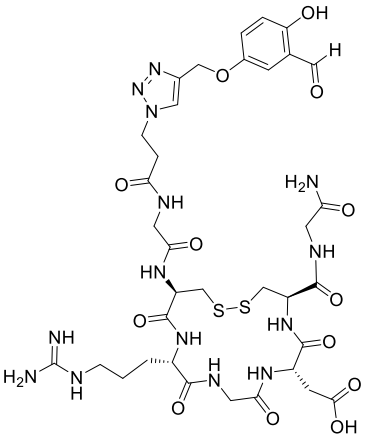
Figure 35 A) Molecular structure of the "parent" ligand RGD-2C (compound **39**); B) molecular docking of **39** into the $\alpha_V\beta_3$ integrin binding site. The lysine residues available for the reversible covalent interaction are highlighted in the picture; C) Measured distances between each lysine residue in $\alpha_V\beta_3$ integrin and, respectively, the N ("CH_3 lig") and the C ("NH_2 lig") termini of compound **39**.

These data highlighted that both the N and C termini of ligand **39** are relatively close to a Lys(ε-NH_2) group, opening to different possibilities of ligand derivatization with the 2HB tag.

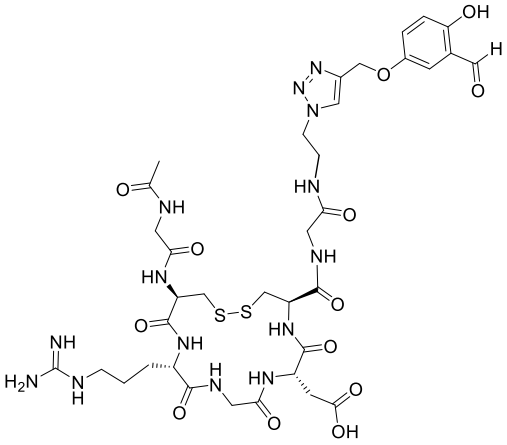
Among the possible linkers that could be used to connect the 2HB tag to **39**, we opted for a triazole connection. In particular, the triazole ring can be easily formed by a copper-catalyzed azide-alkyne cycloaddition, a chemoselective reaction that allows the 2HB installation onto the RGD ligand at late stages of the synthetic pathway. This type of conjugation allows the fine-tuning of the linker's length and flexibility, aimed at the optimal Lys engagement.

Our first approach to the design of reversible-covalent integrin ligands consisted in the 2HB-RGD adducts **81** and **82** shown in Figure 36. The structure of these compounds was based on the parent ligand **39**, with a short linker connecting the 2HB tag to either the N (in **81**) or the C (in **82**) termini of the RGD cyclopeptide. In order to estimate the contribution of the covalent bond to the overall ligand binding affinity for the $\alpha_V\beta_3$ integrin receptor, we also designed two additional compounds, **83** and **84** as "negative controls" of compounds **81** and **82**, respectively. In the control compounds, the 2HB tag was replaced by a non-substituted benzaldehyde, devoid of the *ortho*-phenolic group. This

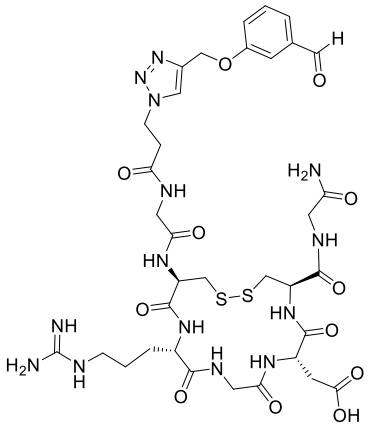
modification was devised according to the 2HB mechanism of action described in Chapter 4, with the phenolic OH group being fundamental for the 2HB imine stabilization in aqueous media. For this reason, the removal of the phenolic hydroxyl group **83** and **84** leads to very similar molecular structures to the 2HB-bearing analogues, albeit unable to form a remarkably stable imine bond with Lys groups in the $\alpha_V\beta_3$ integrin receptor.



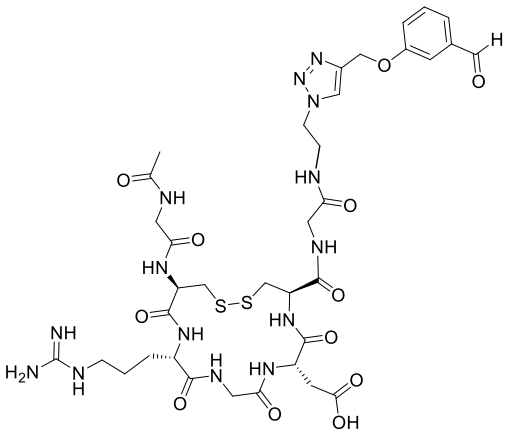
N-side 2HB-RGD (**81**)



C-side 2HB-RGD (**82**)



N-side negative control (**83**)

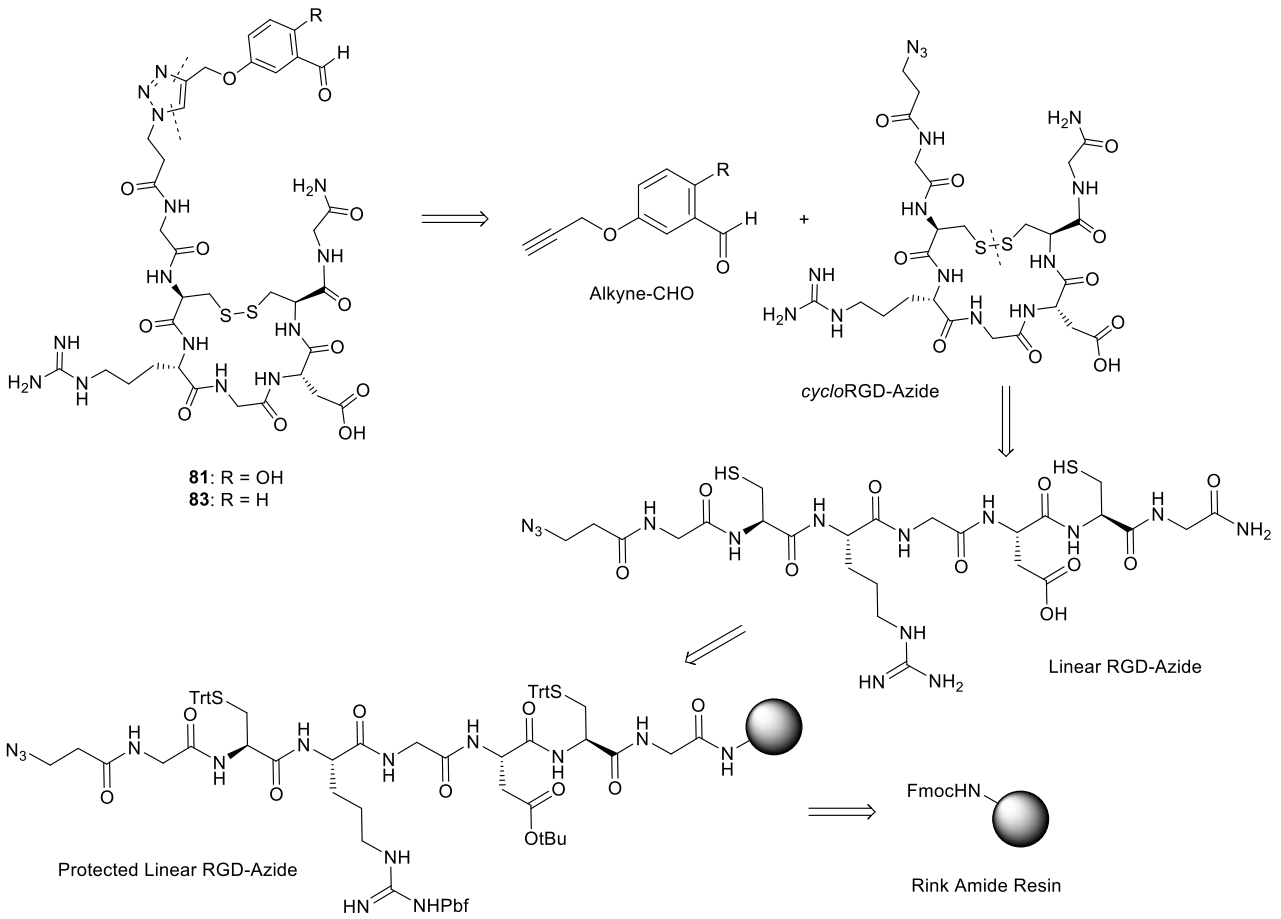


C-side negative control (**84**)

Figure 36 Molecular structures of the 2HB-RGD ligands **81** and **82** and their relative analogues **83** and **84**, devoid of the imine-stabilizing phenolic group.

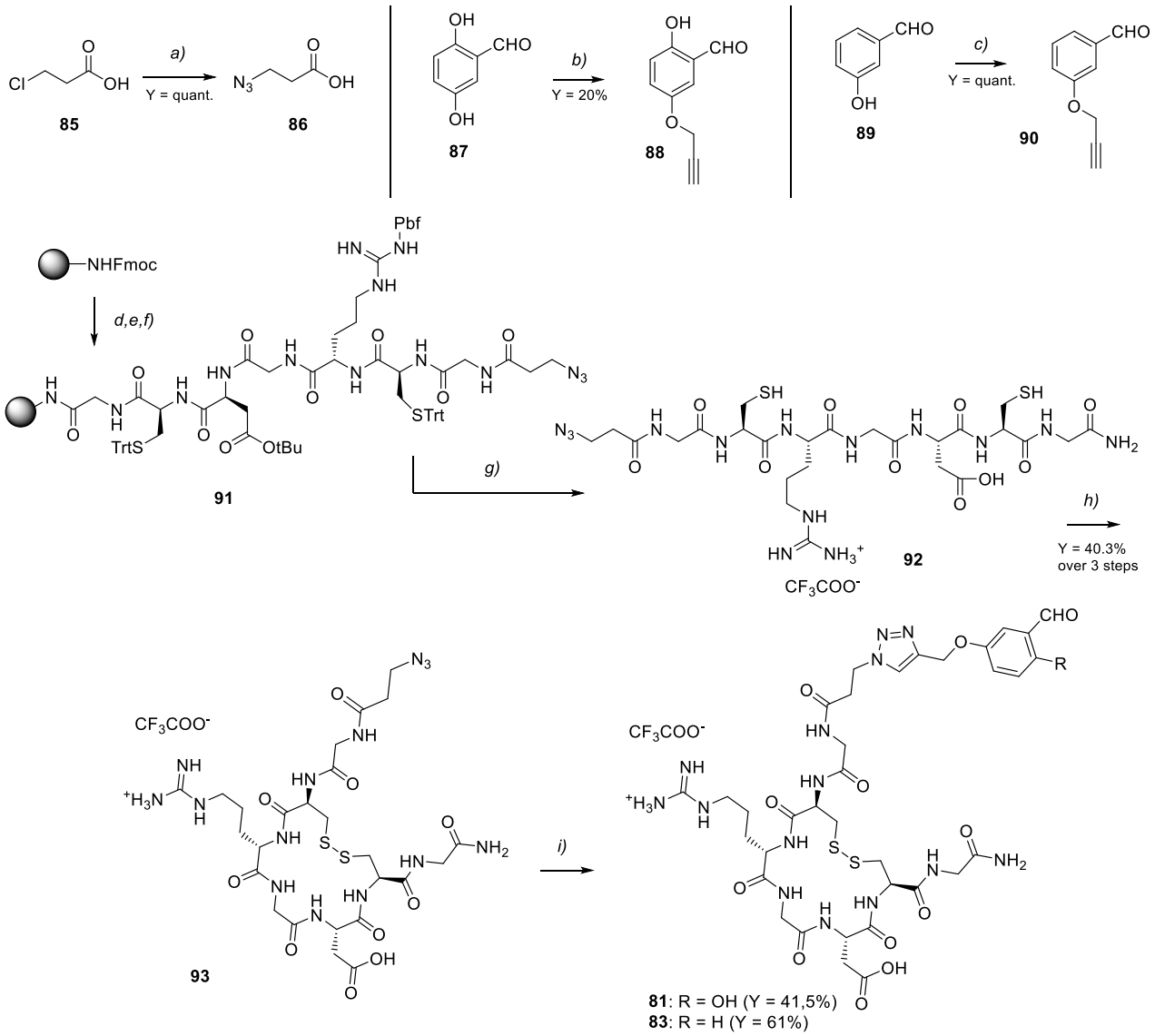
4.2 Synthesis of reversible covalent integrin ligands

As shown in Scheme 16, the retrosynthetic pathways of both 2HB-bearing compound **81** and its benzaldehyde analogue **83** are characterized by the same disconnections, as they both represent N-terminal modifications of the of the naive RGD cyclopeptide. The first disconnection of the target structures consists in the triazole ring formation, which can be easily achieved through a copper-catalyzed azide-alkyne cycloaddition (CuAAC) between the propargylated derivative of aromatic aldehydes and an azide moiety at the RGD N terminus. The latter can be installed onto the peptide structure by using an azide-modified carboxylic acid, connected to the peptide by standard amide coupling. The subsequent retrosynthetic analysis reflects the design of naive peptide **39**, consisting in the disulfide bond opening to obtain the corresponding linear structure, which can be prepared by standard SPPS protocols. We opted for a Rink amide resin for the peptide growth, which allows the capping of the terminal carboxylic acid function as primary amide.



Scheme 16 Retrosynthetic analysis of compounds **81** and **83**.

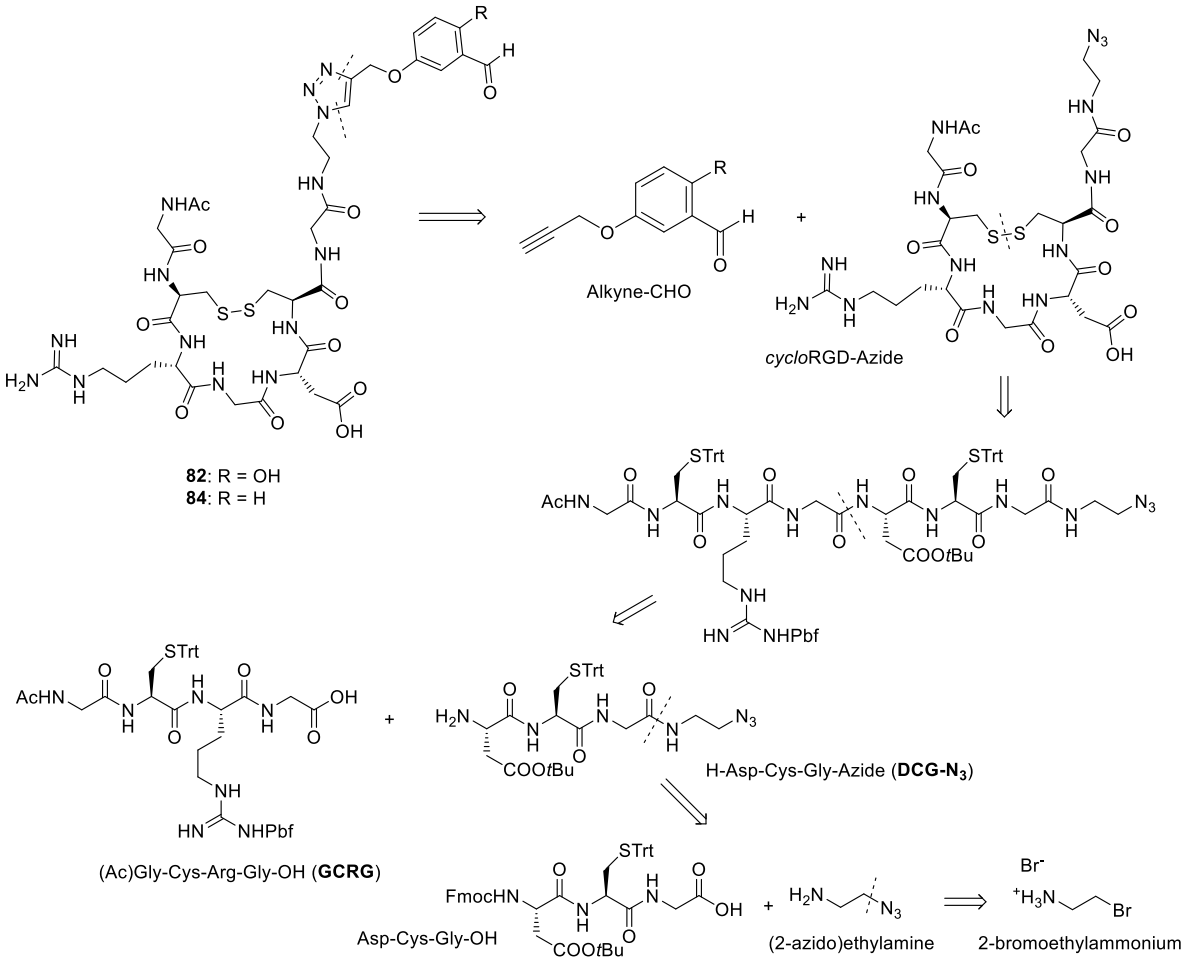
Compounds **81** and **83** were prepared following the synthetic pathway reported in Scheme 17. The preparation of the N-terminal functionalized RGD ligands started with the synthesis of the 3-azidopropanoic acid **86** following a patent procedure,[136] consisting in the nucleophilic substitution of the commercially available 3-chloropropionic acid with sodium azide in boiling water.



Scheme 17 Reagents and conditions: a) NaN_3, H_2O, 100 °C, 22 h; b) Propargyl bromide, K_2CO_3, acetone, 60 °C, 2 h; c) Propargyl bromide, K_2CO_3, DMF dry, r.t.; overnight.; d) 20% piperidine in DMF; e) Fmoc-AA-OH or **86**, DIC, HOAt, DMF, 70 °C (MW), 10 min; f) 20% Ac_2O in DMF; g) TFA:TIS:H_2O 95:2.5:2.5, 2 h, r.t.; h) I_2; H_2O:MeCN 1:1, 30 min., r.t.; i) **88** or **90**, $CuSO_4 \cdot 5H_2O$, sodium ascorbate, degassed H_2O/DMF, 40 °C, overnight.

The synthesis of the 2-HB derivative functionalized in position 5 with a propargyloxy group **88** was performed following a literature procedure[127] in which the commercially available 2,5-dihydroxybenzaldehyde is derivatized in position 5 using propargyl bromide as alkylating agent, potassium carbonate as base and acetone as solvent. Although the reaction conditions of this step were optimized, the desired product **88** was obtained in poor yields (20%), owing to the regioselectivity issues discussed in Chapter 3. Similarly to compound **88**, 3-(propargyloxy)benzaldehyde **90** was obtained using a literature procedure[137] in which the 3-hydroxybenzaldehyde was derivatized at the *meta*-hydroxyl group using propargyl bromide and potassium carbonate in dry DMF. SPPS of the linear heptapeptide was performed using a Rink amide resin and the standard Fmoc/*t*Bu protocol, leading to the protected peptide **91**. A one-pot acidic cleavage and deprotection of **91** was performed using a mixture of trifluoroacetic acid and scavengers (TFA:TIS:H_2O 95:2.5:2.5), which led to the formation of the linear deprotected peptide

92. The latter, after precipitation in cold diethyl ether, was used as starting material for the intramolecular disulfide bond formation without further purifications. The oxidative cyclization of compound **92** was carried out using iodine in a 1:1 mixture of water and acetonitrile. In particular, a low concentration (2 mM) of compound **92** was used to prevent the undesired intermolecular reaction. After purification of the crude mixture with semipreparative RP-HPLC, the cyclic peptide **93** was obtained in 40% yield over three steps. Final compounds **81** and **83** were obtained through CuAAC reaction between the azide-bearing peptide **93**, and alkynes **88** and **90**, respectively. The reactions were performed using the Cu(II) sulphate salt and sodium ascorbate as reducing agent, which enables the *in situ* generation of the active Cu(I) catalytic species in a degassed water/DMF mixture. In both cases, the reaction was run at 40 °C overnight. After purification of the crude mixtures with semipreparative RP-HPLC, compounds **81** and **83** were obtained respectively with 41% and 61% yield.

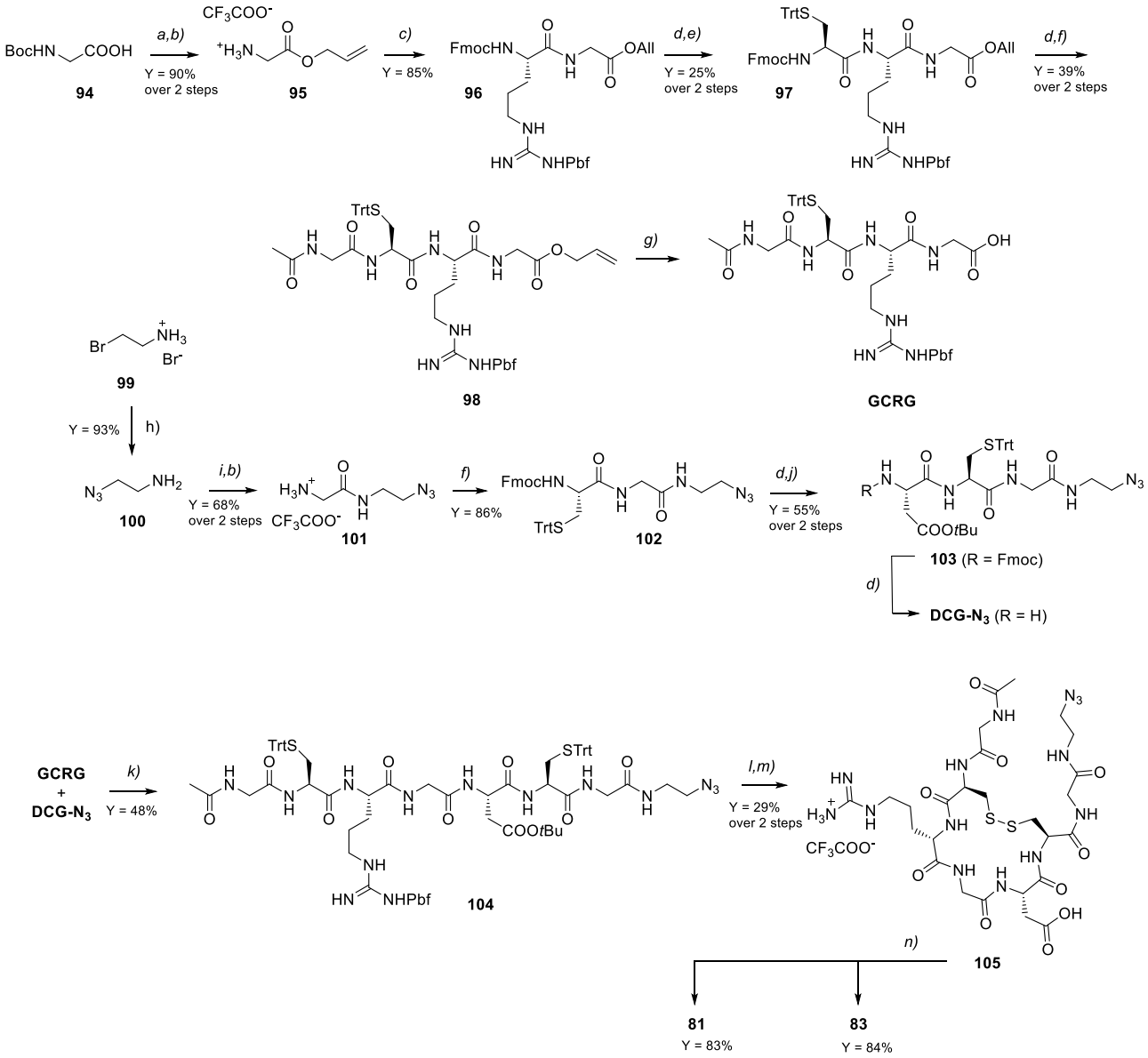


Scheme 18 Retrosynthetic analysis of compounds **82** and **84**.

The retrosynthetic analysis of compounds **82** and **84** is reported in Scheme 18. Also in this case, the first disconnection is performed at the triazole ring, obtaining the propargyl derivative of the aromatic aldehyde and the cyclic RGD peptide bearing an azide moiety at the C terminus. Following disulfide bond opening, the retrosynthetic analysis of the linear peptide is different to that of N-terminal modified ligands **81** and **83**, as the azide installation at the peptide C terminus is not compatible with

standard SPPS protocols. Therefore, we opted for a synthesis of the protected linear peptide in solution, with a first disconnection performed in the middle of the integrin-binding motif, i.e. the amide bond between the glycine and the aspartic acid. This disconnection was devised to avoid epimerization at the C-terminal amino acid during the fragment condensation. The resulting GCRG tetrapeptide, capped at the N-terminal amine as acetamido moiety, can be easily assembled by progressive coupling of the four amino acids, bearing suitable acid-labile protecting groups (e.g. Pbf for the Arg and trityl for Cys side chains). On the other hand, the DCG-N_3 fragment is derivatized at the C-terminus with a (2-azido)ethylamine adaptor, which can be obtained by nucleophilic substitution of the commercially available 2-bromoethylamine hydrobromide.

The synthetic pathway followed for the preparation of compounds **82** and **84** is reported in Scheme 19. The synthesis of the GCRG tetrapeptide started with the allylation of Boc-glycine **94** following a literature procedure,[138] consisting in the carboxylic acid deprotonation with cesium carbonate and its nucleophilic attack to allyl bromide in acetonitrile at room temperature. After aqueous work-up, the intermediate Boc-glycine allyl ester treating with trifluoroacetic acid. After removal of the volatiles in vacuo, the glycine allyl ester **95** was isolated as trifluoroacetate salt and subjected to an amide coupling with arginine. This step was performed using HATU as coupling agent, HOAt as additive and *N,N*-diisopropylethylamine as base. The activated carboxylic acid of Fmoc-arginine (protected at the side chain with an acid-labile Pbf group) was treated with amine **95** and, after purification with column chromatography, the protected dipeptide **96** was obtained in 85% yield. Later on, the Fmoc protecting group in dipeptide **96** was removed with the aim to liberate the amino group and continue the peptide assembly. Fmoc removal was performed treating compound **96** with a 1:1 mixture of acetonitrile and diethylamine: this secondary amine was preferred over the more traditional piperidine due to its lower boiling point, which facilitates its separation from the reaction mixture. After co-evaporation of the diethylamine with acetonitrile and crude filtration over a short plug of silica, the resulting amine was directly used in the next amide coupling. Commercially available Fmoc-cysteine, protected as trityl thioether at the side chain, was activated with the same coupling reagents used in the previous step and then treated with free-base Arg(Pbf)-Gly-OAll. After column chromatography, the corresponding tripeptide **97** was obtained with 25% yield over two steps. Following Fmoc removal from **97** as described above, the corresponding amine was coupled to *N*-acetylglicine, leading to the fully protected tetrapeptide **98** in 39% yield over 2 steps. The following step consisted in the deallylation of compound **98**, which was carried out through a Pd-catalyzed Tsuji-Trost reaction using palladium diacetate and triphenylphosphine to generate the catalytically-active Pd(0) complex. *N*-methylaniline was used as mild nucleophilic scavenger to quench the π-allyl Pd(II) complex and restore the catalyst. After the purification of the crude mixture on a short plug of silica, the GCRG peptide was used in the next step without further purification.



Scheme 19 REAGENTS AND CONDITIONS: a) Allyl bromide, Cs_2CO_3, MeCN dry, r.t., overnight b) TFA:CH_2Cl_2 1:2, 0 °C to r.t., 2h; c) Fmoc-Arg(Pbf)-OH, HATU, HOAt, *i*Pr_2NEt, DMF dry, 0°C to r.t, overnight; d) Et_2NH:MeCN 1:1, r.t; e) Fmoc-Cys(Trt)-OH, HATU, HOAt, *i*Pr_2NEt, DMF dry, 0 °C to r.t, overnight; f) *N*-acetylglycine, HATU, HOAt, *i*Pr_2NEt, DMF dry, 0 °C to r.t, overnight; g) $Pd(OAc)_2$, PPh_3, *N*-methylaniline, dry CH_2Cl_2, 0 °C to r.t, overnight; h) NaN_3, H_2O, 100 °C, overnight; i) Boc-Gly-OH, EDCl, HOBt, Et_3N, DCM dry, r.t., 4h; j) Fmoc-Asp(OtBu)-OH, HATU, HOAt, *i*Pr_2NEt, dry DMF, 0 °C to r.t, overnight; k) [1] EDCl, *N*-hydroxysuccinimide, DMF dry, 0 °C to r.t, overnight; [2] *i*Pr_2NEt, r.t, 2h; l) TFA:TIS:H_2O 95:2.5:2.5, 0 °C to r.t, 2h; m) I_2, H_2O:MeCN 1:1, r.t, 1h; n) Alkyne **88** or **90**, $CuSO_4 \cdot 5H_2O$, sodium ascorbate, degassed H_2O/DMF, 40 °C, overnight.

The synthesis of the DCG-N_3 tripeptide started with a nucleophilic substitution reaction performed with the hydrobromide salt of 2-bromoethylamine (**99**) and sodium azide as nucleophile in water, following a literature procedure.[139] 2-azidoethylamine **100** was isolated and used as starting material for the coupling reaction with Boc-glycine. The acid activation was performed using EDCl as activating agent and, after aqueous work-up of the reaction mixture, the corresponding Boc-protected glycinamide was isolated and immediately treated with trifluoroacetic acid, to yield the trifluoroacetate salt **101** in 68% yield over 2 steps. In this way, the installation of the azide moiety on the C-terminus of the peptide has been successfully performed. The above-mentioned functionalization would not have been possible with a standard SPPS protocol because the

anchoring of the peptide to the resin is always performed on the C-terminus, leading to the impossibility to obtain the desired fonctionalization. Both the amide coupling and the Fmoc removal were performed following the same protocols described for the synthesis of the GCRG tetrapeptide. The first coupling reaction was performed between Fmoc-Cys(Trt)-OH and compound **101**, obtaining the corresponding dipeptide **102** in 86% yield. After Fmoc removal in compound **102** and coupling with Fmoc-aspartic acid (protected at the side chain as tert-butyl ester), the tripeptide **103** was isolated in 55% yield over 2 steps. Later on, Fmoc removal in compound **103** led to the DCG-N_3 tripeptide, which was reacted with the CGRG fragment without further purification. The activation of GCRG was performed using EDCl as activating agent and N-hydroxysuccinimide as additive. The corresponding NHS ester was treated with DCG-N_3 in the presence of N,N-diisopropylethyamine as base. After purification by flash chromatography, the fully protected linear peptide **104** was isolated in 48% yield. The latter was treated with trifluoroacetic acid and scavengers as described above, to perform a complete side chain deprotection and, after precipitation in cold diethyl ether, the corresponding deprotected linear peptide was oxidized to form the macrocyclic ring. This procedure led to the isolation of cyclic peptide **105** in 29% yield over 2 steps. Finally, CuAAC reaction between the azide **105** and either aldehydes **88** or **90** led to the N-functionalized ligands **81** and **83**, which were obtained with 83% and 84% yield respectively, after purification by semipreparative RP-HPLC and lyophilization.

4.3 *In vitro* biological tests and covalent docking studies

2HB-RGD ligands **81** and **82** and the relative negative controls **83** and **84** were tested *in vitro* for their ability to compete with biotinylated vitronectin for the binding to integrin $\alpha_V\beta_3$. As discussed extensively in Chapter 3, this experiment allows the estimation of ligand binding affinity for the integrin receptor, in the presence of an integrin-binding protein competitor. The resulting IC_{50} values are displayed in Table 4.

Table 4 IC_{50} of compounds **81**-**84** determined by competitive binding assays against isolated $\alpha_V\beta_3$ integrin receptor.

	IC_{50}[a] [nM]
Compounds	$\alpha_V\beta_3$
*cyclo*RGDfV (**2**)	1.60 ± 0.90
RGD-2C (**39**)	6.39 ± 0.37
N-side 2HB-RGD peptide (**81**)	129.7 ± 6.4
N-side negative control (**83**)	6.9 ± 1.4
C-side 2HB-RGD peptide (**82**)	3.1 ± 2.2
C-side negative control (**84**)	7.9 ± 2.1

[a]IC_{50} values were determined as the concentration of compound required for 50% inhibition of biotinylated vitronectin binding to integrin $\alpha_v\beta_3$, as estimated by using GraphPad Prism software. All values are the arithmetic mean ± the standard deviation (SD) of triplicate determinations.

The IC_{50} values indicate that the 2HB tag conjugation at the peptide C terminus (ligand **82**) leads to an enhanced affinity for the receptor compared to both the phenol-free benzaldehyde analog **84** and the non-functionalized parent peptide **39**. On the other hand, the 2HB conjugation at the peptide N terminus (ligand **81**) led to a sub-micromolar IC_{50}, i.e. ca 20 times higher than control compound **83**. To rationalize these data, in collaboration with prof. Laura Belvisi (University of Milan), we performed covalent docking experiments. In these *in silico* studies, the 2HB tag of compounds **81** and **82** are forced to form an irreversible covalent bond with each one of the four available lysine residues (Lys119 for the α_v subunit, Lys125,181 and 253 for the β_3 subunit). With this forced intermolecular bond, this docking experiment investigated the fitting of non-covalent interactions (i.e. the ones engaged by the RGD cyclopeptide) in the $\alpha_v\beta_3$ binding pocket. As a result of these studies, the ligand ability to form an array of both covalent and non-covalent interactions with the protein was qualitatively estimated by analyzing 10 different binding poses for each compound, anchored to a specific Lys residue. In particular, a given ligand binding pose was considered optimal whenever the cyclopeptide moiety was superimposable to the X-ray structure of the well-known integrin ligand cilengitide in the $\alpha_v\beta_3$ binding pocket. A summary of the covalent docking studies is reported in Table 5.

Table 5 Number of poses in which the 2HB-RGD ligand, bound to each one of the four available lysines, can reproduce the cilengitide X-ray pose.

Reactive residue	Number of RGD X-ray poses	
	N-side 2HB-RGD peptide (**81**)	C-side 2HB-RGD peptide (**82**)
$Lys^{125}\ \beta_3$	1/10	10/10
$Lys^{181}\ \beta_3$	0/10	5/10
$Lys^{253}\ \beta_3$	5/10	0/10
$Lys^{119}\ \alpha_v$	0/10	0/10

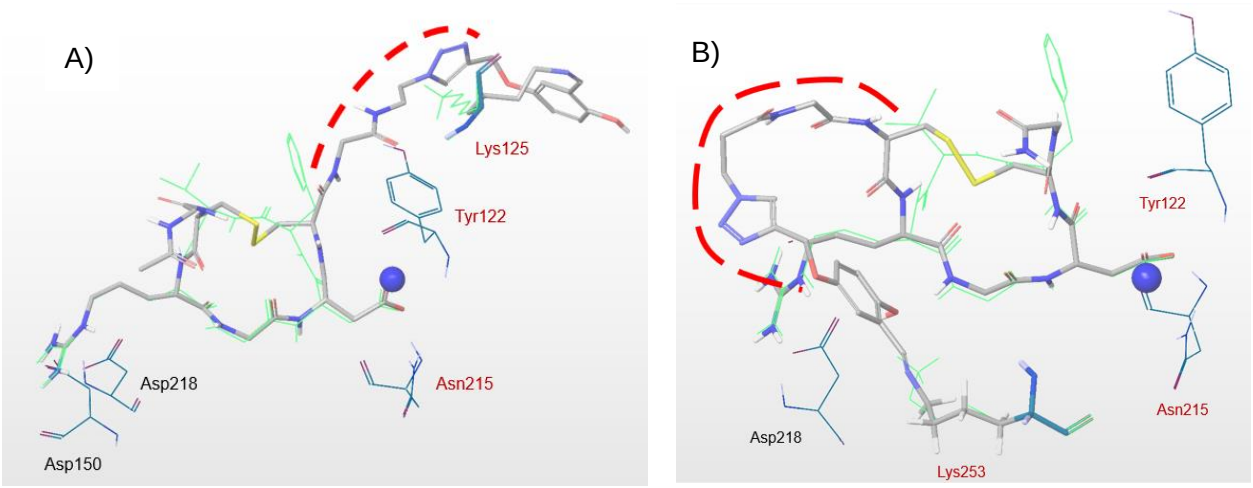


Figure 37 A) Representative covalent docking pose of 2HB-bearing peptide **82** where the cyclic RGD peptide connected to the accessible β_3Lys125 overlaps with the X-ray structure of benchmark Cilengitide ligand (in green) in 10/10 poses. B) representative covalent docking pose of 2HB-bearing peptide **81**, where the forced interaction of N-terminal 2HB with accessible β3Lys253 leads to RGD overlapping with Cilengitide in 5/10 poses.

The data show that compound **81**, when connected to Lys253 in the β_3 subunit, was able to reproduce the cilengitide X-ray pose in only 50% of cases, while migrating from the canonical binding pose in the other 50%. On the other hand, compound **82** bound to Lys125 reproduced the cilengitide binding poses in all computational experiments (10/10). As an example of binding poses undertaken by the 2HB-RGD ligands into the receptor binding pocket, Figure 37A shows one of the ten poses of compound **82** when bound to Lys125 of the β_3 subunit: the covalent bond forced between the 2HB tag and Lys125 leads to an extended conformation of the 2HB-triazole tether (red dotted line) as well as a good overlapping between the X-ray pose of cilengitide (in green) and the one of the cyclopeptide component in **82**. On the other hand, Figure 37B (i.e. one of the five coherent poses) shows that the anchoring of ligand **81** to Lys253 results in a “bent” conformation of the 2HB-triazole tether (red dotted line) which tends to move the RGD peptide away from the canonical binding site during the calculations.

These computational analyses are in good agreement with binding studies, as the good fit of **82** is reflected by a ≈50% lower IC_{50} value compared to control compound **84**. Concerning ligand **81**, these data may indicate that upon non-covalent RGD docking, the imine formation between the 2HB tag of **81** and Lys253 destabilizes the RGD binding in the $\alpha_V\beta_3$ pocket. This is reflected by the sub-micromolar IC_{50} value of **81** for $\alpha_V\beta_3$ integrin, i.e. ≈ 20 times higher than both naive ligand **39** and control compound **83**.

While further studies are still in progress (i.e. ligand structure optimization and binding/covalent docking studies with other integrin heterodimers) the data collected so far demonstrated that covalent docking is a valuable tool to guide the chemical design of reversible-covalent ligands, whenever structural information on the ligand binding pose is available. Moreover, we provided evidences that, in some cases, the ligand docking can promote the 2HB connection to “undesired” Lys residues, which may impair the non-covalent ligand-protein interactions and thus the ligand binding activity.

Conclusions and future perspectives

The inhibition of clinically-relevant proteins with small molecule drugs is a widely developed pharmaceutical strategy, but it is limited in many cases by a low stability of the ligand-protein complex. During my PhD thesis work, our research activity concerned the investigation of general strategies to enhance the stability of ligand-protein adducts, aimed at increasing the therapeutic benefits. Building on the long-standing expertise of our group in integrin ligands, cyclic peptides bearing the RGD binding motif were used as case study for the investigation of two novel strategies, consisting in the design of either multivalent or reversible-covalent ligands (i.e. integrin receptors).

Within the multivalent approach, we designed and synthesized a “condensed” dimeric bicyclic RGD ligand that proved a higher integrin binding affinity than suitable control compounds. This activity gap was observed not only in binding assays with purified $\alpha_V\beta_3$ and $\alpha_5\beta_1$ integrin receptors, but also in cultured U373-MG glioblastoma cells. This work expanded the application of bicyclic peptides, demonstrating that multivalent effects and pharmacophore cooperativity can be also displayed without designing large and dendrimeric structures. In addition to the integrin field, the design of “condensed” dimeric bicyclic peptides may be applied to a variety of cyclic peptides featuring a defined target-binding motif.

Concerning reversible covalent ligands, we focused our efforts on the 2-hydroxybenzaldehyde (2HB) tag, a particular aromatic aldehyde that can engage the ε-amino group of lysine residues in remarkably stable imines. We successfully developed a practical synthetic approach for the synthesis of portable 2HB-PEG modules that can be conjugated to virtually all type of protein ligands (e.g. small organic molecules, peptides and peptidomimetics). Moreover, the reversible covalent approach was applied for the first time to the integrin field, with the design of reversible covalent RGD cyclopeptides. In particular, the 2HB tag installation at the peptide C terminus led to an enhanced binding affinity compared to the parent non-covalent peptide. Interestingly, we demonstrated the importance of a proper structural design, as the 2HB tag installation at the peptide N terminus was found to decrease the ligand binding affinity compared to both the parent ligand and suitable control compounds. To this end, covalent docking calculations were proposed as valuable tools to guide the design of this novel class of reversible-covalent ligands.

These results expand the scope of Lys-targeting reversible covalent ligands, supporting the efficacy of the 2HB tag in binding affinity enhancement. Finally, considering the high frequency of Lys residues on the surface of proteins, the technology discussed herein can be translated to a large number of ligand-protein interactions, paving the way to next-generation small molecule drugs.

Experimental Section

General remarks and procedures

Materials and methods

All manipulations requiring anhydrous conditions were carried out in flame-dried glassware, with magnetic stirring and under a nitrogen atmosphere. All commercially available reagents were used as received. Anhydrous solvents were purchased from commercial sources and withdrawn from the container by syringe, under a slight positive pressure of nitrogen. Compounds **67**,[134] **70**,[99] **73**,[135] **86**,[136] **88**,[127] **90**,[137] **95**[138] and **100**[139] were prepared according to literature procedures, and their analytical data were in agreement with those already published. Reactions were monitored by analytical thin-layer chromatography (TLC) using silica gel 60 F254 pre-coated glass plates (0.25 mm thickness). Visualization was accomplished by irradiation with a UV lamp and/or staining with a potassium permanganate alkaline solution or ninhydrin. Flash column chromatography was performed according to the method of Still and co-workers[140] using Chromagel 60 ACC (40-63 µm) silica gel. Proton chemical shifts are reported in ppm (δ) with the solvent reference relative to tetramethylsilane (TMS) employed as the internal standard ($CDCl_3$ δ = 7.26 ppm; CD_2Cl_2, δ = 5.32 ppm; CD_3OD, δ = 3.31 ppm, D_2O, δ = 4.79 ppm; DMSO-d_6, δ = 2.50 ppm; THF-d_8, δ = 3.58 ppm, 1.72 ppm).[141] The following abbreviations are used to describe spin multiplicity: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, bs = broad signal, dd = doublet of doublet, ddd = doublet of doublet of doublet, ddt = doublet of doublet of triplet. Carbon NMR spectra were recorded on a spectrometer operating at 100.63 MHz, with complete proton decoupling. Carbon chemical shifts are reported in ppm (δ) relative to TMS with the respective solvent resonance as the internal standard ($CDCl_3$, δ = 77.16 ppm; CD_2Cl_2, δ = 54.00 ppm; DMSO-d_6, δ = 39.51 ppm; CD_3OD, δ = 49.05 ppm; THF-d_8 δ = 67.57 ppm, 25.37 ppm). HPLC purifications and HPLC traces of final products were performed on Dionex Ultimate 3000 equipped with Dionex RS Variable Wavelenght Detector (column: Atlantis Prep T3 OBDTM 5 µm 19 x 100 mm; flow 15 mL min−1 unless stated otherwise). High resolution mass spectra (HRMS) were recorded on a Q-TOF Synapt G2-Si instrument available at the MS facility of the Unitech COSPECT at the University of Milan. Low resolution mass spectra (MS) were recorded on Thermo Scientific LCQ Fleet Ion Trap mass spectrometer (ESI source).

General procedures

SEMI-AUTOMATIC SPPS PROTOCOL FOR THE PEPTIDE GROWTH

The semi-automatic SPPS was carried out using a Biotage Initiator™ synthesizer, assisted by microwave (MW) irradiation; Fmoc/tBu strategy was used for the peptide growth with different type of resins, depending on the desired functional group on the C-side of the final peptide chain.

Each coupling step consisted in:

1)activation of the Fmoc-protected amino acid;

2)addition of the activated amino acid to the resin at the synthesizer to start the coupling;

3)steps of washing, deprotection and washing again.

Before these steps, when the first AA of the peptide was not the Gly, it needed a coupling step of the first AA (in our case β-Ala) to the 2-ClTrtCl resin.

To perform the automated SPPS, one solution and four solvents were prepared: 20% piperidine in DMF (*v/v*), DCM, DMF, MeOH and Et_2O. DCM was used both as swelling and washing solvent, DMF was necessary either for washings and as solvent for the reactions of coupling and deprotection, MeOH is used as washing solvent and capping agent, Et_2O is used as drying solvent for resin storage.

A) RESIN PREPARATION

The resin was weighted in a 10 mL Teflon vial (Biotage) and processed with the swelling task. The resin was then ready for peptide synthesis.

B) ACTIVATION OF FMOC-AA-OH

Activation of Fmoc-AA-OH took place as follows: a solution of Fmoc-AA-OH (4.0 eq) in DMF dry (3 mL) was cooled to 0°C in an ice bath. HOAt or Oxyma (4.0 eq), DIPEA (8.0 eq) and DIC (4.0 eq) were added to the solution and the mixture was stirred for 15 min at 0 °C. After that, the reaction mixture was added to the swollen resin (1.0 eq) in order to start the coupling reaction.

C) BIOTAGE INITIATOR™ PROGRAMS

The Biotage Initiator™ programs (“Tasks”) used for the semi-automatic SPPS are reported below. Each task can be modified in every parameter, and it is performed under vortex mixing at 800 rpm.

Swelling task: 3 mL of DCM were added to the vial and the mixture was stirred for 20 min at RT. It followed the removal of DCM and 3 mL of DMF were added to the vial. The mixture was then stirred at RT for 5 min and the DMF was removed in order to obtain swollen beads.

Coupling of Fmoc-β-Ala-OH on resin – capping with MeOH: a solution of Fmoc-β-Ala-OH (398.5 mg, 1.28 mmol, 4.0 eq) and DIPEA (446 µL, 2.56 mmol, 8.0 eq) in 3 mL of a mixture DCM:DMF 1:1 was prepared. The resulting solution was added to the swollen resin (200 mg, 1.0 eq) and the coupling was made at RT for 1 hour. It followed the capping with MeOH (3.0 mL, 15 min at RT). The beads were washed with DMF, DCM, MeOH and DMF (6x3.0 mL, 2x3.0 mL, 2x3.0 mL, 2x3.0 ml, respectively. 20 sec for each wash step).

Coupling MW - Fmoc deprotection: The activated Fmoc-AA-OH residue was added to the resin in the reaction vessel of the synthesizer and the coupling reaction assisted by microwaves was carried out at 75 °C under inert atmosphere for 10 min. At the end of the reaction, the beads were washed twelve times with DMF (3 mL × 20 s for every wash). Two deprotection steps were then performed by adding the deprotection solution (20% *v/v* piperidine in DMF, 3.0 mL for each step) to the beads:

the reaction was performed at r.t. under inert atmosphere for 2 min and 10 min for the first and the second deprotection step, respectively. The beads were washed six times with DMF, twice with DCM and twice with DMF (3.0 mL × 20 s for every wash). At the end of each step, the resin was rinsed and ready for the next coupling reaction.

Coupling MW – Capping - Fmoc deprotection: The activated Fmoc-AA-OH residue was added to the Rink Amide MBHA resin in the reaction vessel of the synthesizer and the coupling reaction assisted by microwaves was carried out at 75 °C under inert atmosphere for 10 min. At the end of the reaction, the beads were washed twelve times with DMF (3 mL × 20 s for every wash). To the washed beads, a capping step was performed adding 3 mL of an acetic anhydride solution (20% Ac_2O in DMF). The beads were stirred for 15 min at r.t. and then they were washed six times with DMF. It followed two deprotection steps that were carried out adding the deprotection solution (20% *v/v* piperidine in DMF, 3.0 mL for each step) to the beads: the reaction was performed at r.t. under inert atmosphere for 2 min and 10 min for the first and the second deprotection step, respectively. The beads were washed six times with DMF, twice with DCM and twice with DMF (3.0 mL × 20 s for every wash). At the end of each step, the resin was rinsed and ready for the next coupling reaction.

Resin drying with Et_2O: 3 mL of Et_2O were added to the vessel and the mixture is stirred for 1 min at RT. Then the beads were washed six times with Et_2O (3.0 mL per each wash, 5x 45 sec and, for the last one, 1x30 sec with a 2 min draining).

D) RESIN STORAGE

The resin attached peptides were stored at -20 °C after the drying task. The next SPPS cycle always starts with a swelling step before continuing the peptide synthesis. The Rink amide MBHA attached peptide were stored swollen at -20°C in DMF dry with the N-terminus equipped with Fmoc protecting group.

<u>GENERAL PROCEDURE A FOR PEPTIDE CLEAVAGE FROM 2CTC RESIN WITHOUT LOSS OF SIDE-CHAIN PROTECTING GROUPS</u>

The cleavage reactions from the resin were performed manually, under inert atmosphere and vortex mixing. The resin was swollen as reported before. In the meantime, 12 mL of cleavage cocktail were prepared mixing 1.2 mL of glacial acetic acid, 2.4 mL of 2,2,2-trifluoroethanol (TFE) and bringing the total volume to 12 mL with DCM. The resin was washed 4 times for 20 minutes with 3 mL of cleavage cocktail. Each liquid fraction was collected by flushing nitrogen into the vessel. Cold Et_2O was added to the cleaved peptide solution in the cleavage cocktail, in order to precipitate the peptide. The solid was isolated by centrifugation, dissolved in DCM:MeOH mixture and concentrated in vacuo. The crude product was used as starting material for the next synthetic steps without further purifications.

<u>GENERAL PROCEDURE B FOR SIMULATNEOUS PEPTIDE CLEAVAGE AND DEPROTECTION</u>

The cleavage reactions from the resin were performed manually, under inert atmosphere and vortex mixing. The protected on-beads peptide was swollen first with DMF (3 mL), then with

dichloromethane (3 mL). Under stirring and nitrogen atmosphere, the beads were treated three times with the cleavage cocktail (3.0 mL per 0.1 mmol of resin) 95:2,5:2,5 TFA /TIS/H_2O (*v/v/v*). After 1 h, the liquid phase was filtered off under nitrogen flow and collected in a round bottom flask: the beads were washed with neat TFA (1.0 mL) that was collected. The combined filtered fractions were concentrated and poured in cold diethyl ether, provoking precipitation of the product. Diethyl ether was removed with a centrifuge affording the crude product that was involved in the next step without further purifications.

GENERAL PROCEDURE C FOR MACROLACTAMIZATION OF PROTECTED LINEAR PEPTIDES

In a two-neck round-bottom flask, under inert atmosphere and flame-dried, the protected peptide (1.0 eq) was dissolved in dry DMF (C = 1.4 mM referred to the protected peptide). The solution was cooled to 0 °C and HATU (4.0 eq), HOAt (4.0 eq) and DIPEA (6.0 eq) were added in the order reported before. The reaction mixture changed color form colorless to yellow. The mixture was stirred at 0°C for 1 hour and after it was stirred at RT overnight. The end of the reaction was monitored by TLC (eluent: DCM:MeOH 9:1). The solvent was removed at the high-vacuum pump. The resulting solid was dissolved in AcOEt (40 mL), the organic phase was washed 3 times with $KHSO_4$ 1 M (3x15 mL) and one time with brine (20 mL). The resulting organic phase was dried with sodium sulfate, the solid was filtered and the solvent was removed in vacuo. The resulting crude was purified by flash-chromatography.

GENERAL PROCEDURE D FOR PEPTIDE SIDE-CHAIN DEPROTECION OF ACID LABILE PROTECTING GROUPS

For this synthetic step, a deprotection cocktail made by TFA:TIS:H_2O 95:2.5:2.5 must be prepared in a quantity necessary to obtain a 0.02 M solution referred to the starting material. In a round-bottom flask, the cleavage cocktail was added to the cyclic protected peptide (1.0 eq) at 0 °C. The reaction mixture was stirred at RT for 2 hours, giving different colors from yellow to violet passing through orange. The reaction mixture was concentrated in vacuo using a NaOH trap and cold Et_2O was added to the resulting mixture in order to precipitate the product. The crude was isolated by centrifugation and the resulting solid was used in the next synthetic steps.

GENERAL PROCEDURE E FOR INTRAMOLECULAR DISULFIDE BOND FORMATION

In a round-bottom flask, the deprotected peptide (1.0 eq) was dissolved in a mixture H_2O/ACN 1:1 (C= 2 mM, the solvents must be HPLC grade). I_2 (4 or 20 eq) was added to the solution and the mixture is stirred for 30 minutes at RT. The reaction mixture is concentrated in vacuo and the resulting crude was purified in RP-HPLC and the isolated product is freeze-dried.

GENERAL PROCEDURE F FOR TOSYLATION OF ALCOHOL

To a 0.5 M solution of alcohol in dry DCM, dry Et_3N (1.5 eq) and DMAP (0.2 eq) were added under nitrogen atmosphere and the resulting solution was cooled down to 0°C. Tosyl chloride (1.2 eq) was added and the mixture was stirred at r.t. overnight. The reaction mixture was then diluted with DCM and washed three times with 1N HCl *aqueous* solution (use 1/10 of *aqueous* phase volume

compared to the organic phase per each wash in order to minimize product loss into the *aqueous* phase) and twice with brine. The organic phase was then dried over Na_2SO_4 and concentrated in vacuo. The crude was then purified with flash chromatography.

GENERAL PROCEDURE G FOR HYDROQUINONE MONOALKYLATION

Hydroquinone (2 eq.), K_2CO_3 (1.1 eq) and Bu_4NI (0.1 eq) were dissolved in dry DMF (1/3 of the reaction mixture final volume; C = 0.2 M compared to the limiting reagent) under nitrogen atmosphere. This suspension was stirred at r.t. for 10 min, then tosylate derivative (1 eq.) in dry DMF (the remaining volume was used to dissolve the starting material) was added to the suspension. The reaction mixture was stirred at 80°C overnight. The mixture was concentrated in high-vacuum and the resulting crude was partitioned between AcOEt and a little amount of 1N HCl *aqueous* solution. The organic phase was washed twice with 1N HCl, twice with 5% LiCl *aqueous* solution and twice with brine, dried over Na_2SO_4 and concentrated in vacuo. The crude was then purified with flash chromatography.

GENERAL PROCEDURE H FOR SKATTEBØL FORMYLATION[133]

In a sealed Schlenk tube, anhydrous $MgCl_2$ (3 eq), paraformaldehyde (5 eq) and dry Et_3N (3 eq) were suspended in an aliquote of dry THF (1/3 of the total volume; C = 0.2 M compared to the phenol) under nitrogen atmosphere and stirred at r.t. for 10 min. To this suspension, a solution of phenol (1 eq) in dry THF (the remaining amount) was added to the suspension and stirred at 70°C overnight. The reaction mixture was concentrated in vacuo and the crude was partitioned between AcOEt and a little amount of 1N HCl. The organic phase was washed twice more with 1N HCl, twice with brine, dried over Na_2SO_4 and concentrated in vacuo. The crude was then purified with flash chromatography.

GENERAL PROCEDURE I FOR THE CuAAC (CLICK) REACTION

A stock solution of the azide (0.1 M in degassed water, 1 eq) and the alkyne (0.1 M in degassed DMF, 1.3 eq) are put in a flask under N_2. To this solution, a stock solution of $CuSO_4 * 5\ H_2O$ (0.1 M in degassed water, 1 eq) and sodium ascorbate (0.1 M in degassed water, 1.1 eq) are sequentially added and the reaction mixture was stirred at 40°C overnight. The reaction mixture was dried in vacuo and the resulting crude was purified by semi-preparative RP-HPLC.

GENERAL PROCEDURE J FOR AMIDE BOND FORMATION

A 0.2 M solution of Fmoc-AA-OH (1.1 eq) in dry DMF was cooled down to 0 °C under N_2 atmosphere. To this solution, HATU (1.1 eq), HOAt (1.1 eq) and DIPEA (3 eq) were added sequentially, and the resulting solution was stirred at 0 °C for 15 min. The amine (1 eq) was then added to the previously obtained solution and the reaction mixture was stirred at r.t. overnight. The mixture was concentrated in vacuo and the resulting crude was dissolved in AcOEt. The organic phase was washed once with 1M $KHSO_4$, once with sat. $NaHCO_3$, once with brine. The organic layer was dried over Na_2SO_4 and concentrated to obtain a crude that was purified with flash chromatography to afford the desired protected peptide.

GENERAL PROCEDURE K FOR FMOC DEPROTECTION

N-Fmoc protected compound (1 eq) was dissolved in a 0.1M solution of 1:1 DEA/MeCN mixture. The resulting solution was stirred at r.t until TLC shows the full consumption of the starting material. The reaction mixture was then concentrated, the resulting crude was co-evaporated three times with MeCN to remove the traces of DEA. The crude was filtered over a pad of silica (first elution with 1:1 AcOEt/Hex, followed by amine elution with 10% MeOH in DCM + 1% TEA) and then used as starting material for the subsequent synthetic step.

Biological assays

SOLID PHASE RECEPTOR BINDING ASSAY

Recombinant human integrin $\alpha_v\beta_3$ (R&D Systems, Minneapolis, MN, USA) was diluted to 0.5 µg mL^{-1} in coating buffer containing 20 mmol L^{-1} tris(hydroxymethyl) amino methane-HCl (Tris-HCl; pH 7.4) for tests involving compounds **36**-**39** or phosphate buffer saline (PBS; pH 7.4) for tests involving compounds **81**-**84**, 150 mmol L^{-1} NaCl, 1 mmol L^{-1} $MnCl^2$, 2 mmol L^{-1} $CaCl^2$, and 1 mmol L^{-1} $MgCl^2$. An aliquot of diluted receptor (100 µl $well^{-1}$) was added to 96-well microtiter plates (NUNC MW 96F MAXISORP STRAIGHT) and incubated overnight at 4 °C. The plates were then incubated with blocking solution (coating buffer plus 1% bovine serum albumin) for an additional 2 h at room temperature to block nonspecific binding; this was followed by a 3 h incubation shaking the plate at room temperature with various concentrations (10^{-12}-10^{-5} M) of test compounds in the presence of 1 µg ml^{-1} vitronectin biotinylated by using an EZ-Link Sulfo-NHS-Biotinylation kit (Pierce, Rockford, IL). After being washed, the plates were incubated shaking for 1 h at room temperature with streptavidin biotinylated peroxidase complex (Amersham Biosciences, Uppsala, Sweden). Then the plates were washed again and finally incubated for 30 min in the dark, with Substrate Reagent Solution (100 µl; R&D Systems, Minneapolis, MN), before the reaction was stopped by addition of 2 N H_2SO_4 (50 µl). The absorbance at 415 nm was read in a Synergy HT Multi-Detection Microplate Reader (BioTek Instruments, Inc.). Each data point is the result of the average of triplicate wells and was analyzed by nonlinear regression analysis with the Prism GraphPad Prism software. Each experiment was repeated in triplicate.

Purified recombinant human integrin $\alpha_5\beta_1$ (R&D Systems, Inc., Minneapolis, MN, USA) was diluted to 0.5 mg mL^{-1} in coating buffer containing 20 mm Tris-HCl (pH 7.4), 150 mm NaCl, 1 mM $MnCl_2$, 2 mM $CaCl_2$ and 1 mM $MgCl_2$. Diluted receptor (100 mL/well) was added to 96-well microtiter plates (NUNC MW 96F Maxisorp Straight) and incubated overnight at 4 °C. The plates were then incubated with blocking solution [coating buffer plus 1% bovine serum albumin (Sigma, St. Luis MO, USA)] for additional 2 h at room temperature to block nonspecific binding, followed by 3 h incubation at room temperature with various concentrations (10^{-12}-10^{-4} M) of test compounds in the presence of 1 mg mL^{-1} biotinylated fibronectin (Sigma, St. Luis MO, USA). Biotinylation was performed by using an EZ-Link Sulfo-NHS-Biotinylation kit (Pierce, Rockford, IL, USA) and plates were analysed as described for integrin $\alpha_v\beta_3$.

CELL CULTURE

The U-373 MG human glioblastoma cell lines were purchased from Istituto Zootecnico Regione Lombardia (Brescia, Italy). The cell lines were grown in DMEM supplemented with 5% fetal bovine serum (FBS), 2 mM glutamine, penicillin-streptomycin (10000 u/ml) and cells were grown at 37 °C in controlled atmosphere (5% CO_2/95% air). Confluent cells were split (1:5-1:10 ratio) by trypsinization and used at the third-fourth passage after thawing. For all the experiments the cells

were plated at a density of 10000 cells/cm^2. The reagents used for the cell cultures were from Euroclone, Italy.

REAL TIME QUANTITATIVE RT-PCR (qRT-PCR)

For mRNA expression analysis, RNA was extracted from U-373 MG cells with Quiazol (Qiagen), followed by a DNAse digestion step. RNA quality was assessed by measuring the 260/280 ratio and concentration was estimated at 260 nm. The primers were designed using the Primer3 Input software; and the specificity of each primer was checked by BLAST analysis. Primers used for integrin subunits and for the housekeeping gene RPL6 in quantitative real time RT-PCR reactions have been reported in Table 6.

Table 6. qRT-PCR analysis of gene expression in U-373 MG cells.

Gene	Ct	ACCESSION NUMBER	PRIMER SEQUENCE
α_V	18,32	NM_002210	F: actggcttaagagagggctgtg R: tgccttacaaaaatcgctga
α_5	25,71	NM_000212	R: tcctcaggaaaggtccaatg R: tcctcaggaaaggtccaatg
β_1	29,06	NM_002213	F: agcctatctccacgcacact R: cctcggagaaggaaacatca
β_3	16,05	NM_002205	F: cctgctgtccaccatgtcta R: ttaatggggtgattggtggt
β_5	19,96	NM_133376	F: tccaatggcttaatttgtgg R: cgttgctggcttcacaagta
GADPH	14,78	NM_002046.5	R: cagcaagagcacaagaggaag F: caactgtgaggaggggagatt

At the end of the reaction, a melting curve analysis was carried out to check for the presence of primer dimers. Experiments were performed on three different cell preparations and each run was analyzed in duplicate. Data are expressed as Ct (Table 6), defined as the number of cycles required for the fluorescent signal to cross the threshold. Ct levels are inversely proportional to the amount of target nucleic acid in the sample.

CELL DETACHEMENT ASSAYS

Cells plated in 96 multi-wells in the growth medium (10.000 cells/100 µl per well) were treated with cell culture medium containing 5, 10, 20 50 µM concentrations of compounds 1-3 for 48 hours. Compound **36**, **37** and **38** stock solutions (200 mM in PBS) were diluted in the growth medium and added to the wells. In control wells only the growth medium was added. At the end of treatments, wells were rinsed three times with PBS and floating cells were removed; cell viability was therefore

measured in adherent cells only. 20 µl of MTS reagent (CellTiter 96 AQueous One Solution Cell Proliferation Assay, Promega) were added to each well. After incubation for 3 hours under standard conditions, the absorbance was read in a multiwell plate reader at 450 nm. Six wells were used for each experimental point and each independent experiment was performed three times in quadruplicate.

WESTERN BLOT ANALYSIS

Cells grown in 60 mm dishes were treated for 48 hours with 50 µM compounds **36**-**39**. The cells were then rinsed twice in ice-cold PBS, and 200 ml of cell lysis buffer were added to the dishes (composition: 50mM Tris-HCl pH 7.4, 1% *v/v* NP40, 0.25% w/v sodium deoxycholate, 1 mM phenylmethylsulphonyl-fluoride (PMSF), 1mM Na_3VO4, 1 mM EDTA, 30 mM sodium pyrophosphate, 1 mM NaF, 1 mg/ml leupeptin, 1 mg/ml pepstatin A, 1 mg/ml aprotinin and 1 mg/ml microcystin). After scraping, the cells were sonicated for 10 seconds, centrifuged at 12.000xg for 5 min at 4 °C. The amount of proteins in the supernatant was then measured by the BCA protein Assay Kit (Pierce). For Western blot analysis, 30 µg of proteins were separated by 10% SDS-PAGE at 150 V for 2 hours and blotted onto 0.22 mm nitrocellulose membranes at 50 mA for 16 hours. The membranes were first blocked for 2 hours in Tris buffered saline solution (TBST composition: Tris 10 mM, NaCl 150 mM, 0.1% Tween 20) plus 5% low fat dry milk (TBSTM) and then incubated with the appropriate antibody diluted to 1:1000 in 5% albumin (pFAK) or TBSTM (FAK), for 16 hours at 4 °C under gentle agitation. The membranes were rinsed three times in TBST and then incubated for 2 hours at 21 °C with a goat anti-rabbit IgG horseradish-peroxidase conjugate secondary antibody (Upstate Biotechnology), diluted to 1:10 000 in TBSTM. The membranes were rinsed three times in TTBS and the luminescent signal was captured with the Image Quant LAS4000, General Electric. Experiments were repeated three times in quadruplicate.

Computational experiments

All calculations were performed using the Schrödinger Suite through the Maestro graphical interface [Maestro, version 10.5, Schrödinger, LLC, New York, NY, 2016].

LIGAND PREPARATION

Ionized carboxylate and protonated guanidinium groups have been employed in calculations for the cyclic RGD integrin ligands as the relevant protonation states at pH = 7 for the acid and basic pharmacophoric groups according to Epik module [Epik version 3.5, Schrödinger, LLC, New York, NY, 2016]. The conformation in water solution of a *cyclo*(CRGDC) derivative was previously investigated by NMR spectroscopy and MD simulations.[142] NMR and computational studies indicate flexibility of the macrocycle yet providing support for an extended conformation of the RGD sequence and for the presence of an inverse γ-turn centered on residue Asp of the ligand. In addition, although the inverse γ -turn conformation is not highly populated in the free peptide, the receptor-bound ligand, similarly to Cilengitide in the $\alpha_V\beta_3$-bound X-ray structure, adopts this conformation in docking models to optimize the interactions with the integrin.[142] Accordingly, macrocycle conformations of the new cyclo(CRGDC) ligands (**39**, **81**, **82**) resembling the X-ray $\alpha_V\beta_3$-bound geometry of Cilengitide have been employed in docking calculations (i.e. extended conformation of the RGD sequence and inverse γ -turn centered on Asp residue).

PROTEIN PREPARATION

The crystal structure of the extracellular domain of the integrin $\alpha_V\beta_3$ in complex with the cyclic pentapeptide RGDf(NMe)V Cilengitide (PDB code 1L5G) was used for docking studies. The $\alpha_V\beta_3$ integrin structure was set up for docking as previously reported (residues 1-438 for chain α_V and 107-354 for chain β_3, all bivalent cations modeled as Mn^{2+} ions).[143] Then, the Protein Preparation Wizard using the OPLS2005 force field was run to get the final structures.

MOLECULAR DOCKING

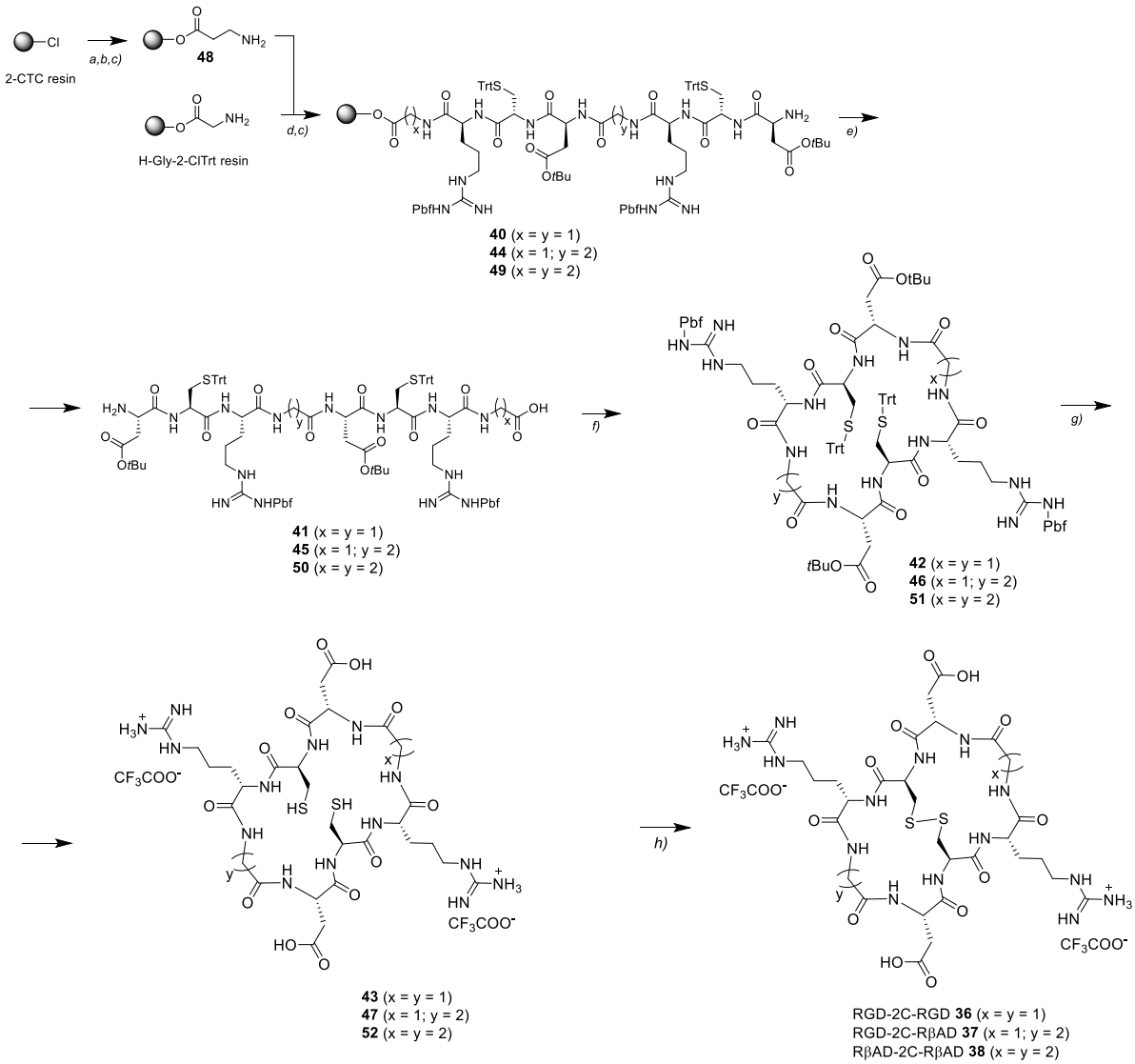
Conventional non-covalent docking calculations were performed using Glide version 7.0 [Glide version 7.0, Schrödinger, LLC, New York, NY, 2016] in the SP (Standard Precision) mode to produce initial docking poses. Receptor grids were generated on the extracellular fragments of $\alpha_V\beta_3$ integrin prepared as described in Protein Preparation. The settings of the flexible-ligand docking protocol were defined as previously reported.[143] The docking protocol was also tested for its ability to reproduce the X-ray binding mode of the cyclic RGD ligand in the receptor crystal structure. Glide was successful in reproducing the experimentally determined binding mode of the cyclic peptide Cilengitide in $\alpha_V\beta_3$ integrin, as it corresponds to the best-scored poses in the docking run. Non-covalent docking poses of the new *cyclo*(CRGDC) ligands show that the RGD peptide can reproduce the X-ray binding mode, and the 2-hydroxybenzaldehyde handle either at the peptide N terminus (compound **81**) or C terminus (compound **82**) can fit unhindered in the $\alpha_V\beta_3$ binding site without forming specific interactions with integrin residues. The covalent docking protocol available in the Schrödinger Suite [Covalent Docking v1.2, Glide, version 7.0, Schrödinger, LLC, New York, NY,

2016] was then applied to generate binding poses of compound **81** and **82** in the X-ray structure of $\alpha_V\beta_3$ while forcing the covalent bond between the 2-hydroxybenzaldehyde moieties and the most accessible Lys(ε-NH_2) groups of both integrin subunits (i.e. α_VLys119, β_3Lys125, β_3Lys181, and β_3Lys253).[144] The docked ligands were confined to an enclosing box with box center in the ligand centroid and automatic box size. A custom imine condensation reaction type was defined to select only aldehyde moieties as reactive sites. Further settings include: docking mode in pose prediction, affinity score calculation using Glide, 10 output poses per ligand reactive site. Analysis of the covalently bound complexes focused on the ability of the cyclic peptides to maintain the canonical X-ray non-covalent interactions of the RGD peptide in the binding pocket while forming the covalent imine bond with a Lys residue.

SYNTHESIS OF BICYCLIC PEPTIDES AND MONOCYCLIC RGD PEPTIDE

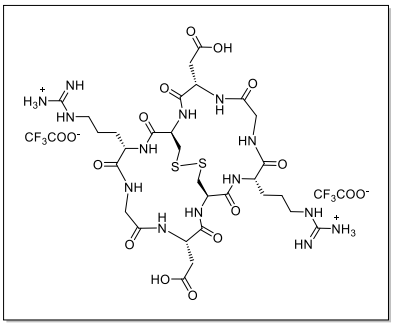
The bicyclic peptides **36**-**38** and the monocyclic peptide **39** were synthesized according to semi-authomatic SPPS protocol and subsequent general procedures.

Synthesis of bicyclic peptides **36**-**38**



REAGENTS AND CONDITIONS: a) Fmoc-ßAla-OH; iPr_2NEt, DMF:DCM 1:1, r.t.. 1h; b) MeOH, r.t., 15 min; c) 20% piperidine in DMF; d) Fmoc-AA-OH, DIC, HOAt, DMF, 70 °C (MW), 10 min; e) AcOH:TFE:DCM 1:2:7; f) HATU, HOAt, iPr_2NEt, DMF; 0 °C to r.t.; g) TFA:TIS:H_2O 95:2.5:2.5 *v/v/v*, 2 h, r.t.; h) I_2; H_2O:MeCN 1:1, 30 min, r.t.;

RGD-2C-RGD (**36**)



$C_{30}H_{48}N_{14}O_{12}S_2$
MW: 1088,97 g/mol

The peptide growth (SPPS protocol) was performed on commercially available H-Gly-2-ClTrt resin (450 mg, 0.234 mmol).

The order and the exact amount of Fmoc-AA-OH used for each coupling step is reported below:

Table 7. Amounts of amino acid derivatives used in the SPPS of compound **40**

Fmoc-AA-OH	Molecular Weight (g/mol)	Amount (mg)
Fmoc-Arg(Pbf)-OH	648.78	607.2
Fmoc-Cys(Trt)-OH	585.71	548.2
Fmoc-Asp(OtBu)-OH	411.15	384.8
Fmoc-Gly-OH	297.31	278.3
Fmoc-Arg(Pbf)-OH	648.78	607.2
Fmoc-Cys(Trt)-OH	585.71	548.2
Fmoc-Asp(OtBu)-OH	411.15	384.8

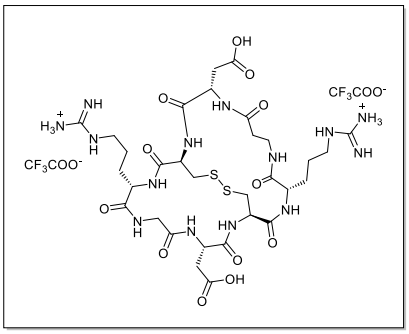
Peptide was cleaved according to General Procedure A. 91.5 mg of solid was obtained and used as starting material for the macrolactamization (General Procedure C). After flash-chromatography (eluent DCM:MeOH 9:1), 67.3 mg of product were obtained, deprotected according to General Procedure D and used in the next step of intramolecular disulfide bond formation (General Procedure E). The crude of this reaction was purified in semi-preparative RP-HPLC using the separation conditions reported below:

Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-15 min: 0-40% B; tr of the product = 9.25 min.

The desired product is freeze-died in water, obtaining its trifluoroacetate salt that appears as a white solid (7.6 mg, 2.98% over 5 steps).

HRMS (ESI) *m/z* calculated for $[C_{30}H_{49}N_{14}O_{12}S_2]^+$: 861.3105; found: 861.3096 $[M+H]^+$

RGD-2C-RβAD (**37**)



$C_{31}H_{50}N_{14}O_{12}S_2$
MW: 1102,99 g/mol

The peptide growth (SPPS protocol) was performed on commercially available H-Gly-2-ClTrt resin (300 mg, 0.156 mmol).

The order and the exact quantity of Fmoc-AA-OH used for each coupling step is reported below:

Table 8. Amounts of amino acid derivatives used in the SPPS of compound **44**

Fmoc-AA-OH	Molecular Weight (g/mol)	Amount (mg)
Fmoc-Arg(Pbf)-OH	648.78	404.8
Fmoc-Cys(Trt)-OH	585.71	365.5
Fmoc-Asp(OtBu)-OH	411.15	256.7
Fmoc-β-Ala-OH	311.33	194.3
Fmoc-Arg(Pbf)-OH	648.78	404.8
Fmoc-Cys(Trt)-OH	585.71	365.5
Fmoc-Asp(OtBu)-OH	411.15	256.7

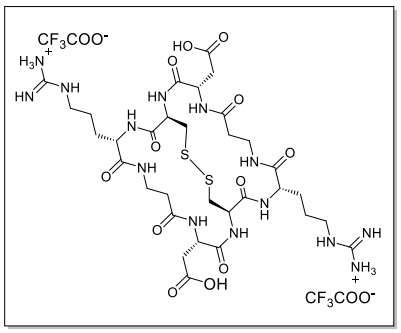
Peptide was cleaved according to General Procedure A. 95.4 mg of solid was obtained and used as starting material for the macrolactamization (General Procedure C). After flash-chromatography (eluent DCM:MeOH 9:1), 59.6 mg of product were obtained, deprotected according to General Procedure D and used in the next step of intramolecular disulfide bond formation (General Procedure E). The crude of this reaction was purified in semi-preparative RP-HPLC using the separation conditions reported below:

Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-15 min: 0 - 40% B; t_r of the product = 9.75 min.

The desired product is freeze-died, obtaining its trifluoroacetate salt that appears as a white solid (6.2 mg, 4% over 5 steps).

HRMS (ESI) *m/z* calculated for $[C_{31}H_{51}N_{14}O_{12}S_2]^+$: 875.3252; found:875.3244 $[M+H]^+$

RβAD-2C-RβAD (**38**)



$C_{32}H_{52}N_{14}O_{12}S_2$
MW: 1117,02 g/mol

The loading of the first AA on the resin and peptide growth (SPPS protocol) started from commercially available 2CTC resin (200 mg).

The loading was determined right after the Fmoc deprotection of the attached AA. Fmoc deprotection of loaded residue was performed manually. The Fmoc group was removed treating the resin twice with 20% (*v/v*) piperidine in DMF (3 mL per each step, first cycle lasted 1 min and the second one lasted 10 min). The liquid phase was collected using a nitrogen flow inside the vessel and the total volume is carefully measured. The beads were washed with DMF, DCM and again with DMF (3 mL per each step, 5 x 30 s per each solvent). A fraction of deprotection solution was diluted 1:1000 with 20% (*v/v*) piperidine in DMF in order to respect the linearity of the Lambert-Beer law at λ=301 nm (if necessary, make further dilutions). 20% (*v/v*) piperidine in DMF solution was the blanc. The loading was calculated using this formula:

$$X = \frac{A(301\ nm) * V * F_d}{\varepsilon(301\ nm) * m * b}$$

Where:

X = loading on the resin (mmol/g);

A (301 nm) = absorbance of the solution measured at 301 nm;

V = total volume of collected deprotection solution (6 mL in this case);

F_d = dilution factor;

ε (301 nm) = 7800 M^{-1} cm^{-1};

m = mass of the resin (g);

b = length of the cell (cm).

In our case, X = 1.74 mmol/g, obtaining a quantitative coupling because the biggest loading reported on the commercially available resin is 1.6 mmol/g.

The order and the exact quantity of Fmoc-AA-OH used for each coupling step is reported in Table 9:

Table 9. Amounts of amino acid derivatives used in the SPPS of compound **49**

Fmoc-AA-OH	Molecular Weight (g/mol)	Amount (mg)
Fmoc-Arg(Pbf)-OH	648.78	892.7
Fmoc-Cys(Trt)-OH	585.71	805.9
Fmoc-Asp(OtBu)-OH	411.15	565.7
Fmoc-β-Ala-OH	311.33	428.3
Fmoc-Arg(Pbf)-OH	648.78	892.7
Fmoc-Cys(Trt)-OH	585.71	805.9
Fmoc-Asp(OtBu)-OH	411.15	565.7

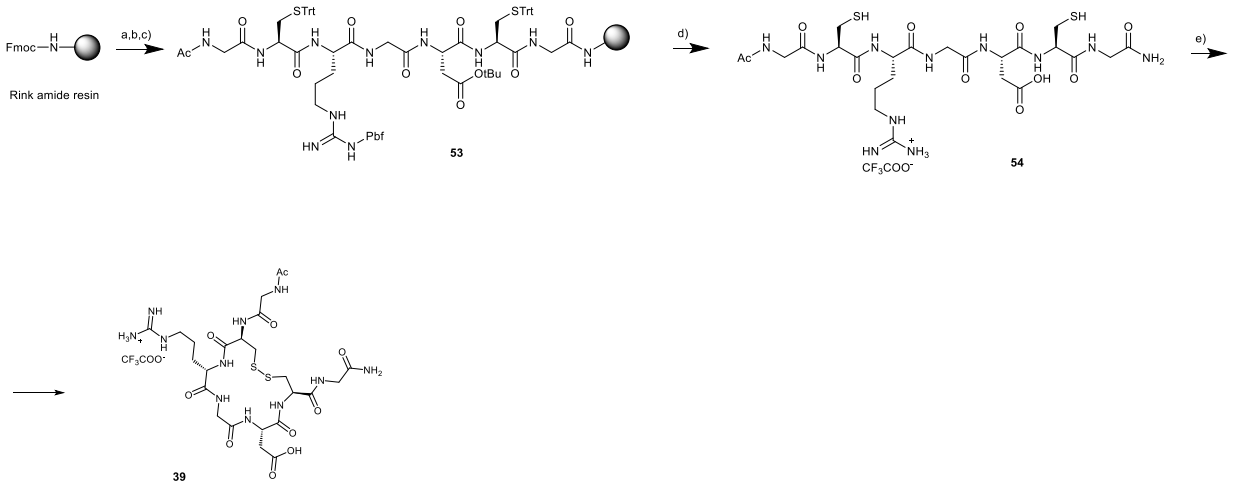
Peptide was cleaved according to General Procedure A. 393.1 mg of solid was obtained and used as starting material for the macrolactamization (General Procedure C). After flash-chromatography (eluent DCM:MeOH 9:1), 134.8 mg of product were obtained, deprotected according to General Procedure D and used in the next step of intramolecular disulfide bond formation (General Procedure E). The crude of this reaction was purified in semi-preparative RP-HPLC using the separation conditions reported below:

Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-15 min: 0-40% B; t_r of the product = 10.4 min.

The desired product is freeze-died, obtaining its trifluoroacetate salt that appears as a white solid (26.1 mg, 1% over 6 steps).

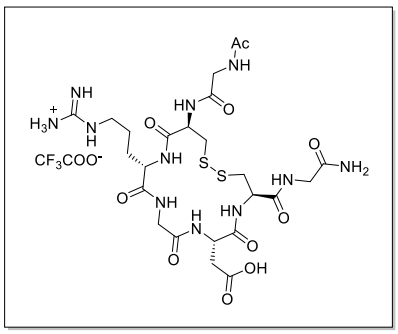
HRMS (ESI) *m/z* calculated for $[C_{32}H_{53}N_{14}O_{12}S_2]^+$: 889.3409; found:889.3401 $[M+H]^+$

Synthesis of monocyclic RGD peptide **39**



Reagents and conditions: a) 20% piperidine in DMF; b) Fmoc-AA-OH, DIC, HOAt, DMF, 70 °C (MW), 10 min; c) 20% Ac_2O in DMF; d) TFA:TIS:H_2O 95:2.5:2.5 *v/v/v*, 2 h, r.t.; e) I_2; H_2O:MeCN 1:1, 30 min, r.t.;

RGD-2C (**39**)



$C_{24}H_{39}N_{11}O_{10}S_2$
MW: 819,79 g/mol

The synthesis was accomplished using a Rink amide MBHA resin (loading = 0.5 mmol/g). The followed synthetic pathway is reported below starting from SPPS step (100 mg, 0.05 mmol). The order and the exact quantity of Fmoc-AA-OH used for each coupling step is reported in Table 10:

Table 10. Amounts of amino acid derivatives used in the SPPS of compound **53**

Fmoc-AA-OH	Molecular Weight (g/mol)	Amount (mg)
Fmoc-Gly-OH	297.31	59.5
Fmoc-Cys(Trt)-OH	585.71	117.1
Fmoc-Asp(OtBu)-OH	411.15	82.3
Fmoc-Gly-OH	297.31	59.5
Fmoc-Arg(Pbf)-OH	648.78	129.7
Fmoc-Cys(Trt)-OH	585.71	117.1
Fmoc-Gly-OH	297.31	59.5

The supported peptide was cleaved and deprotected according to General Procedure B. 80,9 mg of solid was obtained and used as starting material for the oxidation step (General procedure E) without any further purifications. The crude of this reaction was purified in RP-HPLC using the separation conditions reported below:

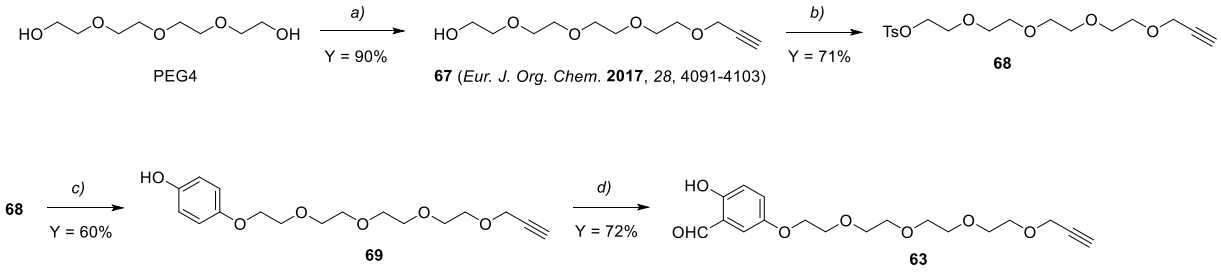
Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-12 min: 0-20% B; t_r of the product = 9.0 min.

The desired product was freeze-died in water, obtaining its trifluoroacetate salt that appears as a white solid (32 mg, 77% over 3 steps).

HRMS (ESI) *m/z* calculated for $[C_{24}H_{40}N_{11}O_{10}S_2]^+$: 706,2401; found: 706,2394 $[M+H]^+$

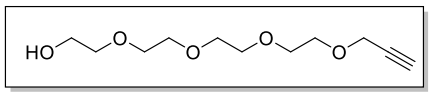
SYNTHESIS OF 2HB-PEG MODULES AND COUPLING TO MODEL SUBSTRATES

Synthesis of 2-HB module **63**



Scheme S1 REAGENTS AND CONDITIONS: *a)* propargyl bromide, NaH, THF dry, 0°C to r.t.,overnight; *b)* tosyl chloride, Et_3N dry, 4-dimethylaminopyridine, CH_2Cl_2 dry, 0°C to r.t., overnight; *c)* Hydroquinone, K_2CO_3, Bu_4NI, DMF dry, 80°C, overnight; *d)* $MgCl_2$, paraformaldehyde, Et_3N dry, THF dry, reflux, overnight.

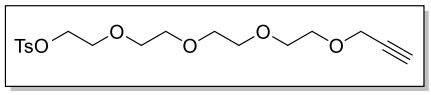
3,6,9,12-Tetraoxapentadec-14-yn-1-ol (**67**)



$C_{11}H_{20}O_5$
MW: 232.13 g/mol

Alkyne **67** was prepared following a published procedure.[134]

3,6,9,12-Tetraoxapentadec-14-yn-1-yl 4-methylbenzenesulfonate (**68**)

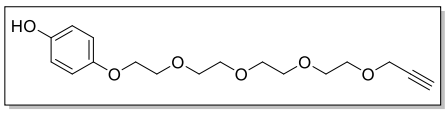


$C_{18}H_{26}O_7S$
MW: 386.14 g/mol

Tosylate **68** was prepared following General Procedure A, starting from 980 mg (4.22 mmol) of alcohol **67**. The crude was purified by flash column chromatography (1:1 Hexane:AcOEt) affording **68** as a colorless oil (1,14 g, 71%). [145]

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.83 – 7.77 (m, 2H), 7.37 – 7.31 (m, 2H), 4.20 (d, J = 2.3 Hz, 2H), 4.18 – 4.14 (m, 2H), 3.73 – 3.57 (m, 14H), 2.45 (s, 3H), 2.42 (t, J = 2.4 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 144.6, 132.8, 129.6, 127.7, 79.5, 74.5, 70.4, 70.3, 70.2, 70.1, 69.1, 68.8, 68.4, 58.1, 21.4; MS (ESI) *m/z* calcd. for $[C_{18}H_{26}O_7SNa]^+$: 409.13 $[M+Na]^+$, found: 409.28.

4-((3,6,9,12-Tetraoxapentadec-14-yn-1-yl)oxy)phenol (**69**)

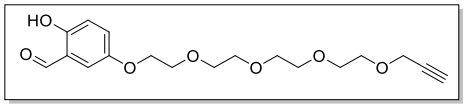


$C_{17}H_{24}O_6$
MW: 324.16 g/mol

Ether **69** was prepared following General Procedure B, starting from 506 mg (1.31 mmol) of tosylate **68**. The crude was purifed by flash chromatography (99:1 CH_2Cl_2:MeOH) affording ether **69** as a pale-brown oil (257 mg, 60%).

^{1}H NMR (400 MHz, $CDCl_3$): δ 6.78-6.71 (m, 4H), 4.19 (d, J = 2.4 Hz, 2H), 4.05-4.03 (m, 2H), 3.82-3.80 (m, 2H), 3.72-3.62 (m, 12H), 2.42 (t, J = 2.4 Hz, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 152.3, 150.5, 116.1, 115.7, 79.6, 74.9, 70.6, 70.6, 70.5, 70.3, 69.9, 69.0, 68.0, 58.3; MS (ESI) *m/z* calcd. for $[C_{17}H_{24}O_6Na]^+$: 347.15 $[M+Na]^+$, found: 347.27.

5-((3,6,9,12-Tetraoxapentadec-14-yn-1-yl)oxy)-2-hydroxybenzaldehyde (**63**)

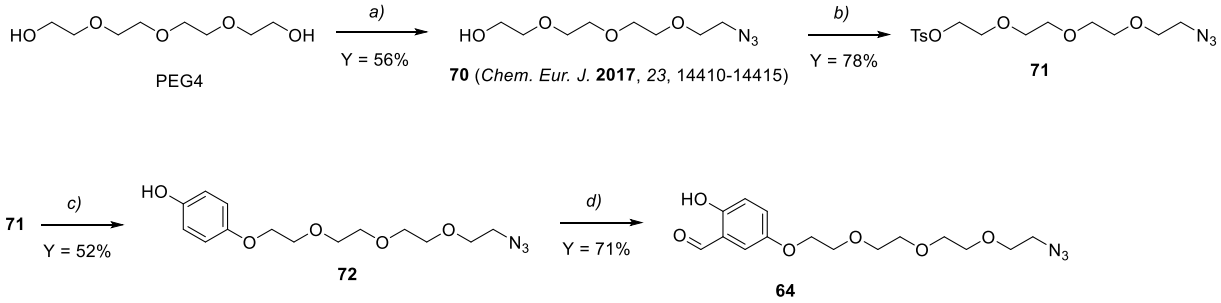


$C_{18}H_{24}O_7$
MW: 352.15 g/mol

Aldehyde **63** was prepared following General Procedure C, starting from 94 mg (0.292 mmol) of phenol **69**. The crude was purified by flash chromatography (99:1 CH_2Cl_2:MeOH) affording aldehyde **63** as a yellow liquid (64 mg, 72%).

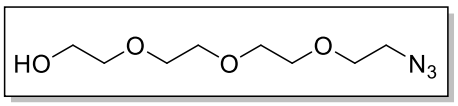
^{1}H NMR (400 MHz, $CDCl_3$): δ 10.65 (s, 1H), 9.84 (s, 1H), 7.18 (dd, J = 9.0, 3.1 Hz, 1H), 7.05 (d, J = 3.1 Hz, 1H), 6.92 (d, J = 9.0 Hz, 1H), 4.20 (d, J = 2.4 Hz, 2H), 4.16-4.09 (m, 2H), 3.89-3.82 (m, 2H), 3.77-3.60 (m, 12H), 2.42 (t, J = 2.4 Hz, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 196.2, 156.2, 151.5, 126.0, 120.1, 118.6, 116.6, 79.7, 74.8, 70.8, 70.6, 70.4, 69.8, 69.1, 68.5, 58.4; HRMS (ESI) *m/z* calcd. for $[C_{18}H_{24}O_7Na]^+$: 375.1414 $[M+Na]^+$, found: 375.1417.

Synthesis of 2-HB module **64**



Scheme S2 REAGENTS AND CONDITIONS: *a)* [1] tosyl chloride, Et_3N, 4-dimethylaminopyridine, CH_2Cl_2, 0 °C to r.t., 2h; [2] NaN_3, DMF, 80 °C, overnight; *b)* tosyl chloride, Et_3N, 4-dimethylaminopyridine, CH_2Cl_2, 0 °C to r.t., overnight; *c)* Hydroquinone, K_2CO_3, Bu_4NI, DMF, 80 °C, overnight; *d)* $MgCl_2$, paraformaldehyde, Et_3N, THF, reflux, overnight.

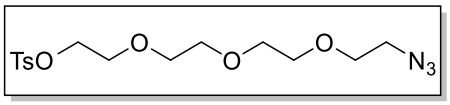
2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethan-1-ol (**70**)



$C_8H_{17}N_3O_4$
MW: 219.12 g/mol

Azide **70** was prepared following a published procedure.[99]

2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (**71**)

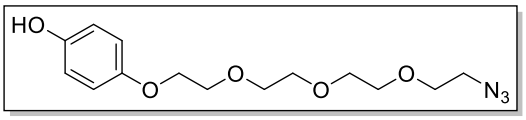


$C_{15}H_{23}N_3O_6S$
MW: 373.13 g/mol

Tosylate **71** was prepared following General Procedure A, starting from 627 mg (2.86 mmol) of alcohol **70**. The crude was purified by flash chromatography (gradient from hexane:AcOEt 6:4 to 100% AcOEt) affording tosylate **71** as a colorless oil (836 mg, 78%).[146]

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.83 – 7.77 (m, 2H), 7.34 (d, J = 8.1 Hz, 2H), 4.19 – 4.13 (m, 2H), 3.74 – 3.55 (m, 12H), 3.38 (t, *J* = 5.1 Hz, 2H), 2.45 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 144.9, 133.0, 129.9, 128.0, 70.8, 70.7, 70.6, 70.1, 69.4, 68.7, 50.7, 21.7; MS (ESI) *m/z* calcd. for $[C_{15}H_{23}N_3O_6SNa]^+$: 396,12 $[M+Na]^+$, found: 396.17.

4-(2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethoxy)phenol (**72**)

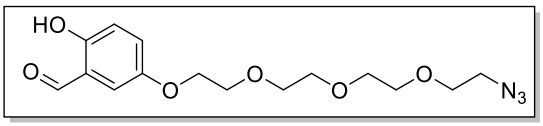


$C_{14}H_{21}N_3O_5$
MW: 311.15 g/mol

Ether **72** was prepared following General Procedure B, starting from 539 mg (1.44 mmol) of tosylate **71**. The crude was purified by flash chromatography (99:1 CH_2Cl_2:MeOH) affording **72** as a pale-brown oil (236 mg, 52%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.80-6.73 (m, 4H), 4.06-4.01 (m, 2H), 3.83-3.80 (m, 2H), 3.74-3.65 (m, 10H), 3.38-3.36 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 152.8, 150.2, 116.2, 115.9, 70.9, 70.8, 70.2, 70.1, 68.2, 50.8; MS (ESI) *m/z* calcd. for $[C_{14}H_{21}N_3O_5Na]^+$: 334.14 $[M+Na]^+$, found: 334.18.

5-(2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethoxy)-2-hydroxybenzaldehyde (**64**)

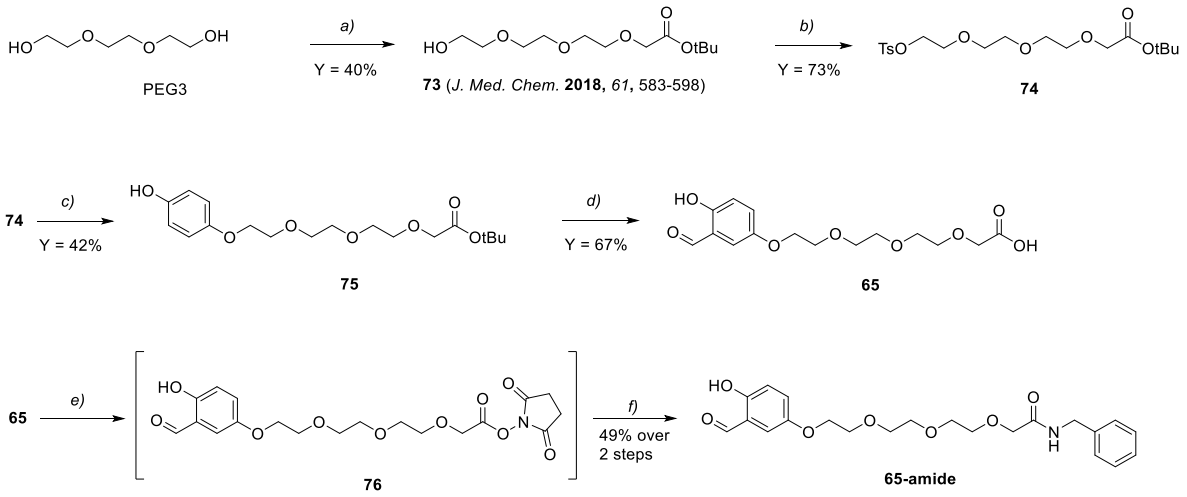


$C_{15}H_{21}N_3O_6$
MW: 339.14 g/mol

Aldehyde **64** was prepared following General Procedure C, starting from 263 mg (0.864 mmol) of phenol **72**. The crude was purified by flash chromatography (99:1 CH_2Cl_2:MeOH) affording **64** as a yellow oil (224 mg, 71%).

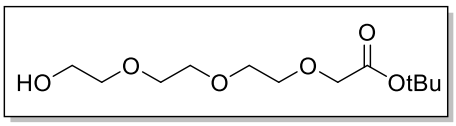
^{1}H NMR (400 MHz, $CDCl_3$) δ 10.64 (s, 1H), 9.83 (s, 1H), 7.16 (dd, J = 9.0, 3.1 Hz, 1H), 7.04 (d, J = 3.1 Hz, 1H), 6.92 (d, J = 9.0 Hz, 1H), 4.13-4.11 (m, 2H), 3.86-3.84 (m, 2H), 3.74-3.64 (m, 10H), 3.37 (t, J = 5.0 Hz, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 196.3, 156.4, 152.2, 126.2, 120.3, 118.9, 116.8, 71.1, 70.9, 70.8, 70.3, 69.9, 68.7, 50.9; HRMS (ESI) *m/z* calcd. for $[C_{15}H_{21}N_3O_6Na]^+$: 362.1323 $[M+Na]^+$, found: 362.1327.

Synthesis of 2-HB module **65** and coupling with benzylamine.



Scheme S3 REAGENTS AND CONDITIONS: *a)* [1] Sodium hydride, DMF, 0 °C to r.t., 1h; [2] *t*-butylbromoacetate, DMF, 0 °C to r.t, 2h; *b)* tosyl chloride, Et_3N, 4-dimethylaminopyridine, CH_2Cl_2, 0 °C to r.t., overnight; *c)* Hydroquinone, K_2CO_3, Bu_4NI, DMF, 80 °C, overnight; *d)* Anhydrous $MgCl_2$, paraformaldehyde, Et_3N, THF, reflux, overnight; *e)* *N*-hydroxysuccinimide, *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride, CH_2Cl_2, 0 °C to r.t., overnight; *f)* benzylamine, *i*Pr_2NEt, CH_2Cl_2, r.t, 1.5 h.

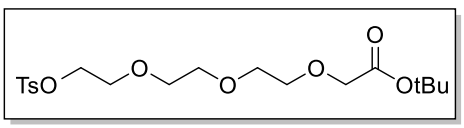
tert-Butyl 2-(2-(2-(2-hydroxyethoxy)ethoxy)ethoxy)acetate (**73**)



$C_{12}H_{24}O_6$
MW: 264.32 g/mol

Ether **73** was prepared following a published procedure.[135]

tert-Butyl 2-(2-(2-(2-tosyloxyethoxy)ethoxy)ethoxy)acetate (**74**)

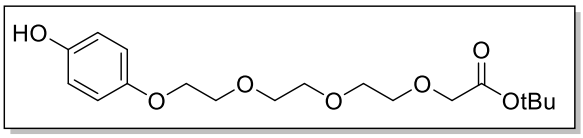


$C_{19}H_{30}O_8S$
MW: 418.50 g/mol

Tosylate **74** was prepared following General Procedure A, starting from 591 mg (2.24 mmol) of alcohol **73** The crude was purified by flash chromatography (gradient from 9:1 to 8:2 Hexane/AcOEt) affording **74** as a pale-yellow oil (687 mg, 73%).[147]

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.80 (d, J = 8.2 Hz, 2H), 7.37 – 7.31 (m, 2H), 4.19 – 4.12 (m, 2H), 4.00 (s, 2H), 3.70 – 3.57 (m, 10H), 2.44 (s, 3H), 1.47 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.9, 145.0, 133.3, 130.1, 128.2, 81.8, 71.8, 70.9, 70.8, 70.7, 70.3, 69.5, 69.2, 68.9, 28.3, 21.9; MS (ESI) *m/z* calcd. for $[C_{19}H_{30}O_8SNa]^+$: 441.15 $[M+Na]^+$, found: 441.12.

tert-Butyl 2-(2-(2-(2-(4-hydroxyphenoxy)ethoxy)ethoxy)ethoxy)acetate (**75**)

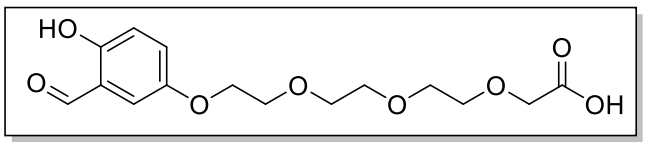


$C_{18}H_{28}O_7$
MW: 356.18 g/mol

Ether **75** was prepared following General Procedure, starting from 687 mg (1.64 mmol) of tosylate **74**. The crude was purified by flash chromatography (98:2 CH_2Cl_2:MeOH) affording **75** as a yellow oil (246 mg, 42%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.77-6.71 (m, 4H), 4.04-4.00 (m, 4H), 3.81-3.79 (m, 2H), 3.72-3.64 (m, 8H), 1.45 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.0, 152.8, 150.4, 116.2, 116.0, 81.9, 70.9, 70.8, 70.8, 70.1, 69.2, 68.3, 28.3; MS (ESI) *m/z* calcd. for $[C_{18}H_{26}O_7Na]^+$: 379.17 $[M+Na]^+$, found: 379.09.

2-(2-(2-(2-(3-Formyl-4-hydroxyphenoxy)ethoxy)ethoxy)ethoxy)acetic acid (**65**)

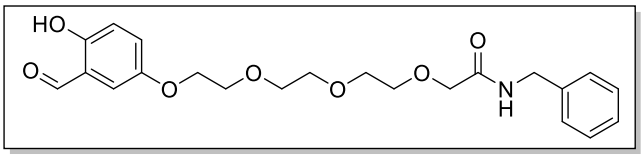


$C_{15}H_{20}O_8$
MW: 328.12 g/mol

Aldehyde **65** was prepared following General Procedure C, starting from 93 mg (0.261 mmol) of phenol **75**. The crude was purified by flash chromatography (CH_2Cl_2:MeOH gradient, from 95:5 to 9:1 + 1% formic acid) affording **65** as a yellow oil (58 mg, 67%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.66 (s, 1H), 9.85 (s, 1H), 7.17 (dd, J = 9.0, 3.1 Hz, 1H), 7.06 (d, J = 3.1 Hz, 1H), 6.92 (d, J = 9.0 Hz, 1H), 4.15-4.12 (m, 4H), 3.86-3.84 (m, 2H), 3.77-3.68 (m, 8H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 196.5, 172.9, 156.4, 152.0, 126.1, 120.3, 118.9, 116.9, 71.4, 70.8, 70.7, 70.4, 69.9, 68.9, 68.6; MS (ESI) *m/z* calcd. for $[C_{15}H_{20}N_3O_6Na]^+$: 351.10 $[M+Na]^+$, found: 351.27.

N-Benzyl-2-(2-(2-(2-(3-formyl-4-hydroxyphenoxy)ethoxy)ethoxy)ethoxy)acetamide (**65-amide**)

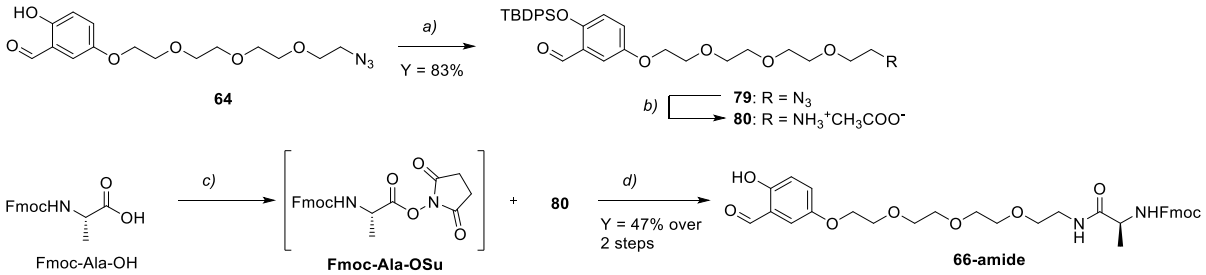


$C_{22}H_{27}NO_7$
MW: 417.46 g/mol

Carboxylic acid **65** (58 mg, 0.136 mmol, 1.0 equiv.) was dissolved in dry CH_2Cl_2 (0.75 ml). *N*-hydroxysuccinimide (16 mg, 0.136 mmol, 1.0 eqiv.) was added and the mixture was cooled to 0 °C. *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCl, 26 mg, 0.136 mmol, 1.0 equiv.) was added and the solution was warmed to r.t. and stirred overnight. Benzylamine (16 µL, 0.150 mmol, 1.1 equiv.) and *i*Pr_2NEt (70 µL, 0.409 mmol, 3.0 equiv.) were dissolved in CH_2Cl_2 (400 µL) and added to the mixture, which was then stirred for 1.5 h. The solution was diluted with CH_2Cl_2 (100 ml) and washed with aq. HCl (1 M, 3 x 10 ml) and brine (10 ml). The organic phase was dried over Na_2SO_4 and concentrated in vacuo. The crude was purified via semipreparative RP-HPLC (flow [ml/min]: 10.00; UV channels:210 nm; 221 nm; A: H_2O + 0.1 % TFA, B: MeCN; 0-2 min: 10 % B; 2-12 min: 10 % B to 100 % B; t_R = 10.5 min) affording **65-amide** as a yellow oil (28 mg, 49%).

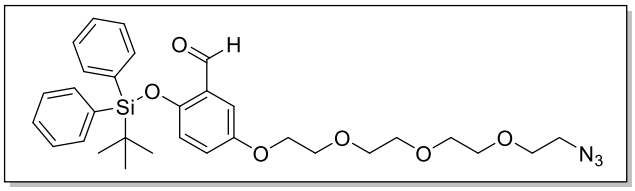
^{1}H NMR (400 MHz, CD_2Cl_2) δ 10.62 (br, 1H), 9.83 (s, 1H), 7.51 (br, 1H), 7.34-7.23 (m, 5H), 7.14 (dd, *J* = 9.0 Hz, 3.1 Hz, 1H), 7.03 (d, *J* = 3.1 Hz, 1H), 6.90 (d, *J* = 9.0 Hz, 1H), 4.45 (d, *J* = 6.1 Hz, 2H), 4.03-4.00 (m, 4H), 3.75-3.71 (m, 2H), 3.70-3.66 (m, 2H), 3.64-3.60 (m, 2H), 3.57 (s, 4H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 196.1, 170.7, 156.3, 151.9, 138.0, 128.7, 127.7, 127.5, 125.9, 120.2, 118.7, 116.6, 71.2, 70.7, 70.5, 70.3, 69.8, 68.4, 43.0; MS (ESI) *m/z* calcd. for $[C_{22}H_{27}NO_7Na]^+$: 440.1680 $[M+Na]^+$, found: 440.1681.

Synthesis of 2-HB derivative **66-amide**



Scheme S4 REAGENTS AND CONDITIONS: *a)* *t*-butyldiphenylsilyl chloride, Et_3N, 4-dimethylaminopyridine, CH_2Cl_2, 0 °C to r.t., overnight; *b)* H_2, Pd/C, THF/water, acetic acid, r.t., 2h; *c)* *N*-hydroxysuccinimide, *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride, DMF, 0 °C to r.t., overnight; *d)* $NaHCO_3$, DMF/water, r.t., overnight.

5-(2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethoxy)-2-((*tert*-butyldiphenylsilyl)oxy)benzaldehyde (**79**)

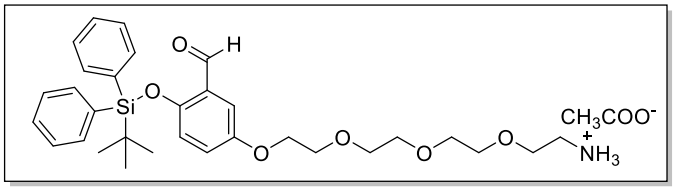


$C_{31}H_{39}N_3O_6Si$
MW: 577.75 g/mol

2HB-azide module **64** (232 mg, 0.685 mmol, 1.0 equiv.), dry Et_3N (158 µL, 1.13 mmol, 1.6 equiv.) and 4-dimethylaminopyridine (9 mg, 0.075 mmol, 0.1 equiv.) were dissolved in dry CH_2Cl_2 (1.5 ml). The solution was cooled to 0 °C and *t*-butyldiphenylsilyl chloride (216 µL, 0.830 mmol, 1.2 equiv.) was added dropwise. The reaction was warmed to r.t. and stirred overnight. A saturated aqueous $NaHCO_3$ solution (1.5 ml) was added to the mixture and the layers were separated. The aq. layer was extracted with EtOAc (3 x 15 ml) and the combined organic layers were dried over Na_2SO_4 and concentrated in vacuo. The product was purified via flash chromatography (gradient from 99.5:0.5 to 98:2 CH_2Cl_2: MeOH), affording silyl eter **79** as a colorless oil (327 mg, 83%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.71 (s, 1H), 7.72-7.70 (m, 4H), 7.47-7.36 (m, 6H), 7.29 (d, J = 3.3 Hz, 1H), 6.76 (dd, J = 9.0 Hz, 3.3 Hz, 1H), 6.44 (d, J = 9.0 Hz, 1H), 4.07-4.05 (m, 2H), 3.81-3.79 (m, 2H), 3.70-3.64 (m, 10H), 3.37-3.35 (m, 2H), 1.11 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 189.9, 153.5, 153.2, 135.6, 132.0, 130.5, 128.2, 126.6, 124.2, 121.6, 110.7, 71.0, 70.9, 70.2, 69.8, 68.1, 50.9, 26.7, 19.8; MS (ESI) *m/z* calcd. for $[C_{31}H_{39}N_3O_6SiNa]^+$: 600.25 $[M+Na]^+$, found: 600.50.

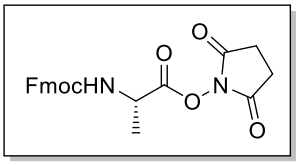
5-(2-(2-(2-(2-Aminoethoxy)ethoxy)ethoxy)ethoxy)-2-((*tert*-butyldiphenylsilyl)oxy)benzaldehyde, acetate salt (**80**)



$C_{33}H_{45}NO_8Si$
MW: 611.81 g/mol

Azide **79** (50 mg, 0.087 mmol, 1.0 equiv.) was dissolved in a 1:1 THF/H_2O mixture + 10 % AcOH (3.15 ml total volume). 10% Palladium on carbon (9.6 mg, 0.009 mmol, 0.1 equiv.) was added and the mixture was stirred under H_2 atmosphere for 2 h. Pd/C catalyst was filtered off and the solvent was removed under reduced pressure, affording **80** as a yellow oil (52 mg) that was directly used in the next reaction without further purification.

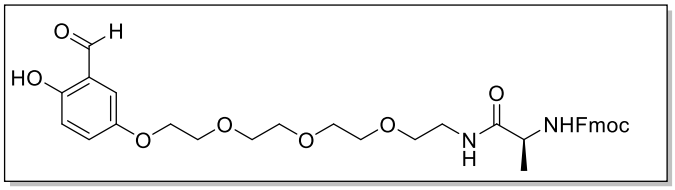
N-(9-Fluorenylmethoxycarbonyl)-L-alanine-*N*-hydroxysuccinate (**Fmoc-Ala-OSu**)



$C_{22}H_{20}N_2O_6$
MW: 408.41 g/mol

A solution of Fmoc-Ala-OH (311 mg, 1 mmol, 1.0 equiv.) and *N*-hydroxysuccinimide (126 mg, 1.1 mmol, 1.1 equiv.) in dry DMF (5 ml) was cooled to 0 °C under a nitrogen atmosphere. To this solution, *N*-Ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDCl, 210 mg, 1.1 mmol, 1.1 equiv.) was added and the mixture was stirred at r.t. overnight. The solvent was removed in vacuo and the resulting crude was partitioned between AcOEt (200 ml) and 1 N HCl (20 ml). The organic phase was washed with 1 N HCl (2 x 20 ml), 5% LiCl (3 x 20 ml) and brine (2 x 20 ml). The organic phase was then dried over Na_2SO_4 and concentrated in vacuo, affording **Fmoc-Ala-OSu** (396 mg) as a white foam, which was used directly in the next step without further purification.

(9H-fluoren-9-yl)methyl (S)-(1-(3-formyl-4-hydroxyphenoxy)-13-oxo-3,6,9-trioxa-12-azapentadecan-14-yl)carbamate (**66-amide**)



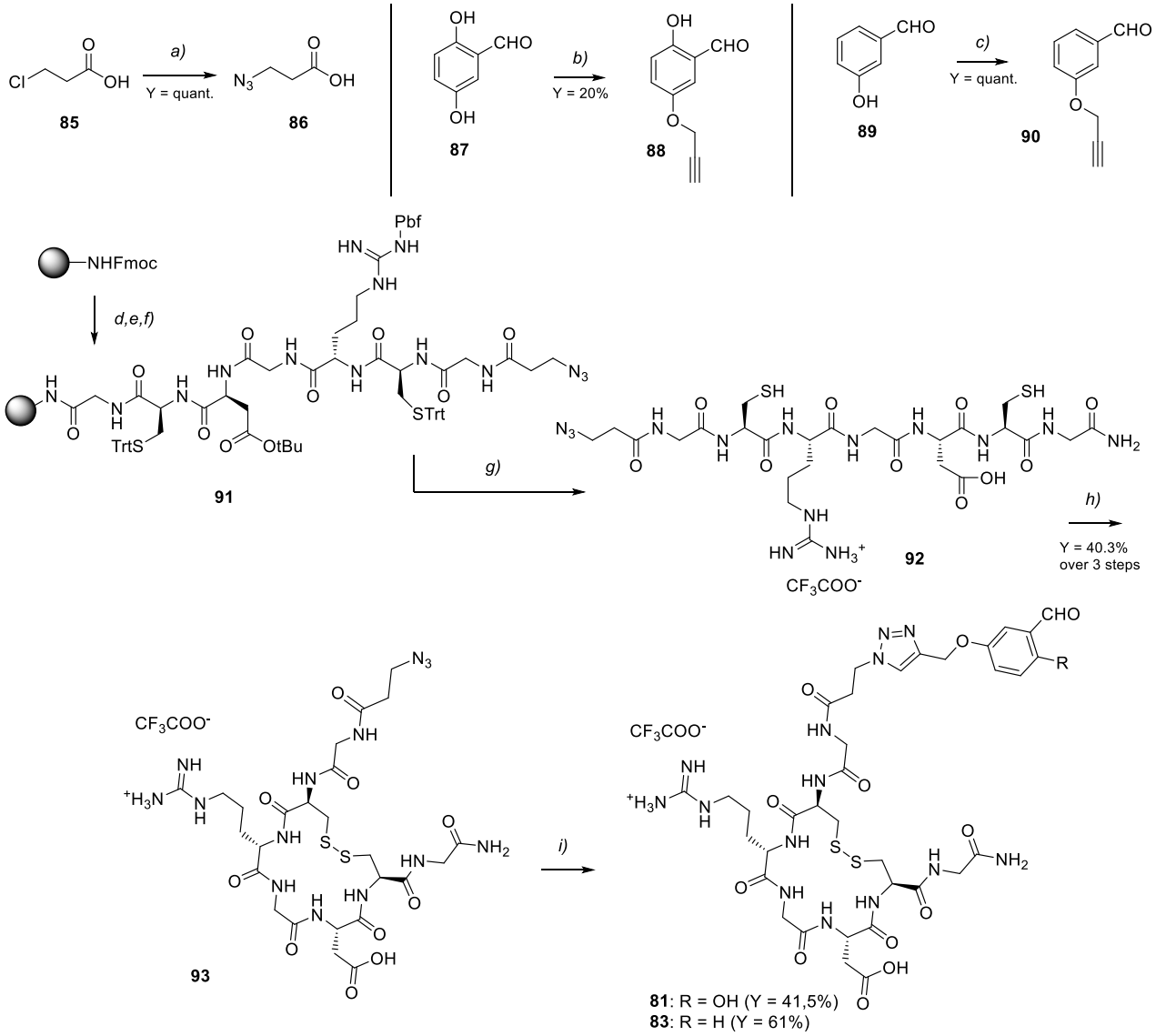
$C_{33}H_{38}N_2O_9$
MW: 606.67 g/mol

Ammonium salt **80** (52 mg, 0.086 mmol, 1.0 equiv.) was dissolved in H_2O (1.2 ml) and $NaHCO_3$ (22 mg, 0.257 mmol, 3.0 equiv.) was added. **Fmoc-Ala-OSu** (68 mg, 0.171 mmol, 2.0 equiv.) was dissolved in DMF (0.8 ml), added to the aqueous mixture and stirred at r.t. overnight. The solvent was removed and the product was purified via semi-preperative RP-HPLC (flow: 10 ml/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1 % TFA, B: MeCN; 0-2 min: 10 % B; 2-12 min: 10 % B to 100 % B; t_R = 11.8 min), affording **66-amide** as a pale-brown oil (24 mg, 47% over 2 steps).

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.63 (br, 1H), 9.80 (s, 1H), 7.75-7.73 (m, 2H), 7.58-7.56 (m, 2H), 7.41-7.37 (m, 2H), 7.32-7.28 (m, 2H), 7.13 (dd, *J* = 9.0, 3.1 Hz, 1H), 7.01 (d, *J* = 3.0 Hz, 1H), 6.90 (d, *J* = 9.0 Hz, 1H), 5.78 (br, 1H), 4.39 (d, *J* = 6.9 Hz, 2H), 4.31-4.23 (m, 1H), 4.19 (t, *J* = 6.9 Hz, 1H), 4.08-4.06 (m, 2H), 3.81-3.79 (m, 2H), 3.72-3.57 (m, 10H), 3.47-3.44 (m, 2H), 1.37 (d, *J* = 6.6 Hz, 3H); ^{13}C-NMR (100 MHz, $CDCl_3$) δ 196.4, 174.0, 156.6, 156.4, 151.9, 143.9, 141.5, 128.0, 127.3, 126.1, 125.3, 120.2, 119.0, 116.9, 70.9, 70.7, 70.6, 70.4, 69.9, 69.6, 68.5, 67.4, 50.9, 47.3, 39.9, 18.9; HRMS (ESI) *m/z* calcd. for $[C_{33}H_{38}N_2O_9Na]^+$: 629.2470 $[M+Na]^+$, found: 629.2470.

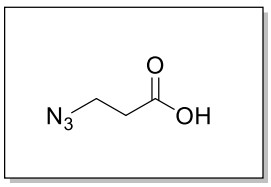
SYNTHESIS OF 2HB-RGD PEPTIDES

Synthesis of N-side 2HB-RGD ligand (**81**) and its negative control (**83**)



Scheme 16 Reagents and conditions: a) NaN_3, H_2O, 100 °C, 22 h; b) Propargyl bromide, K_2CO_3, acetone, 60 °C, 2 h; c) Propargyl bromide, K_2CO_3, DMF dry, r.t.; overnight.; d) 20% piperidine in DMF; e) Fmoc-AA-OH or **86**, DIC, HOAt, DMF, 70 °C (MW), 10 min; f) 20% Ac_2O in DMF; g) TFA:TIS:H_2O 95:2.5:2.5, 2 h, r.t.; h) I_2; H_2O:MeCN 1:1, 30 min., r.t.; i) **88** or **90**, $CuSO_4 \cdot 5H_2O$, sodium ascorbate, degassed H_2O/DMF, 40 °C, overnight.

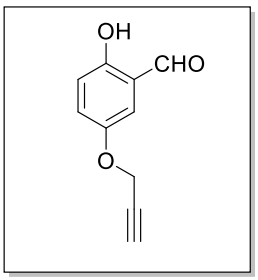
3-azidopropanoic acid (**86**)



$C_3H_5N_3O_2$
MW: 115,09 g/mol

Acid **86** was prepared according to a published procedure.[136]

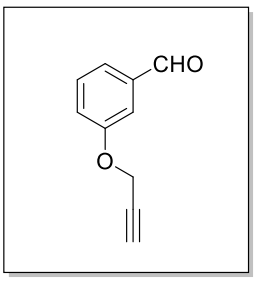
2-hydroxy-5-(prop-2-yn-1-yloxy)benzaldehyde (**88**)



$C_{10}H_8O_3$
MW: 176,17 g/mol

Alkyne **88** was prepared according to a published procedure.[127]

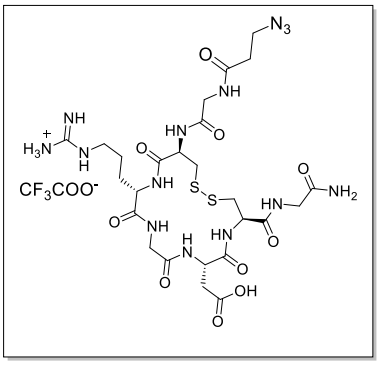
3-(prop-2-yn-1-yloxy)benzaldehyde (**90**)



$C_{10}H_8O_2$
MW: 160,17 g/mol

Alkyne **90** was prepared according to a published procedure.[137]

N-side N_3-RGD-2C (**92**)



$C_{25}H_{41}N_{14}O_{10}S_2$
MW: 874,83 g/mol

The synthesis was accomplished using a Rink amide MBHA resin (loading = 1.1 mmol/g). The followed synthetic pathway is reported below starting from SPPS step (200 mg, 0.22 mmol). The order and the exact quantity of Fmoc-AA-OH used for each coupling step is reported in Table 11:

Table 11. Amounts of amino acid derivatives used in the SPPS of compound **91**

Fmoc-AA-OH	Molecular Weight (g/mol)	Amount (mg)
Fmoc-Gly-OH	297.31	261.6
Fmoc-Cys(Trt)-OH	585.71	515.4
Fmoc-Asp(OtBu)-OH	411.15	361.9
Fmoc-Gly-OH	297.31	261.6
Fmoc-Arg(Pbf)-OH	648.78	570.9
Fmoc-Cys(Trt)-OH	585.71	515.4
Fmoc-Gly-OH	297.31	261.6
3-azidopropanoic acid	115.09	101.3

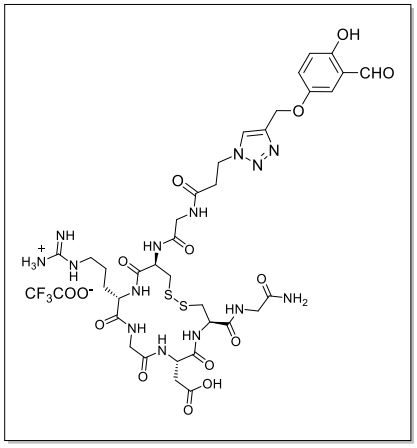
The supported peptide **91** was cleaved and deprotected according to General Procedure B. 130,9 mg of solid was obtained and used as starting material for the oxidation step (General procedure E) without any further purifications. The crude of this reaction was purified in RP-HPLC using the separation conditions reported below:

Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-12 min: 0-30% B; t_r of the product = 10.0 min.

The desired product was freeze-died in water, obtaining its trifluoroacetate salt that appears as a white solid (77.55 mg, 40% over 3 steps).

MS (ESI) *m/z* calculated for $[C_{25}H_{40}N_{14}O_{10}S_2Na]^+$: 783,24; found: 783,42 $[M+Na]^+$

N-side 2HB-RGD-2C (**81**)



$C_{35}H_{48}N_{14}O_{13}S_2$
MW: 1051,00 g/mol

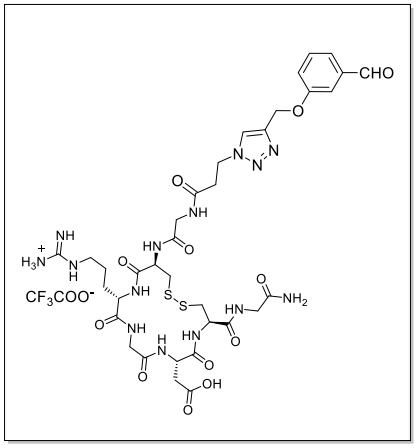
Compound **81** was prepared according to General Procedure I, using compound **92** (28.6 µmol) as the azide and compound **88** (37.2 µmol) as the alkyne. The crude was purified in RP-HPLC using the separation conditions reported below:

Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-12 min: 0-40% B; t_r of the product = 11,5 min.

After freeze-drying, compound **81** was isolated as a fluffy white solid (12.34 mg, 41%)

HRMS (ESI) *m/z* calculated for $[C_{35}H_{48}N_{14}O_{13}S_2]^+$: 937,3039 ; found: 937,3029 $[M+H]^+$

N-side negative control RGD-2C (**83**)



$C_{35}H_{48}N_{14}O_{12}S_2$
MW: 1035,00 g/mol

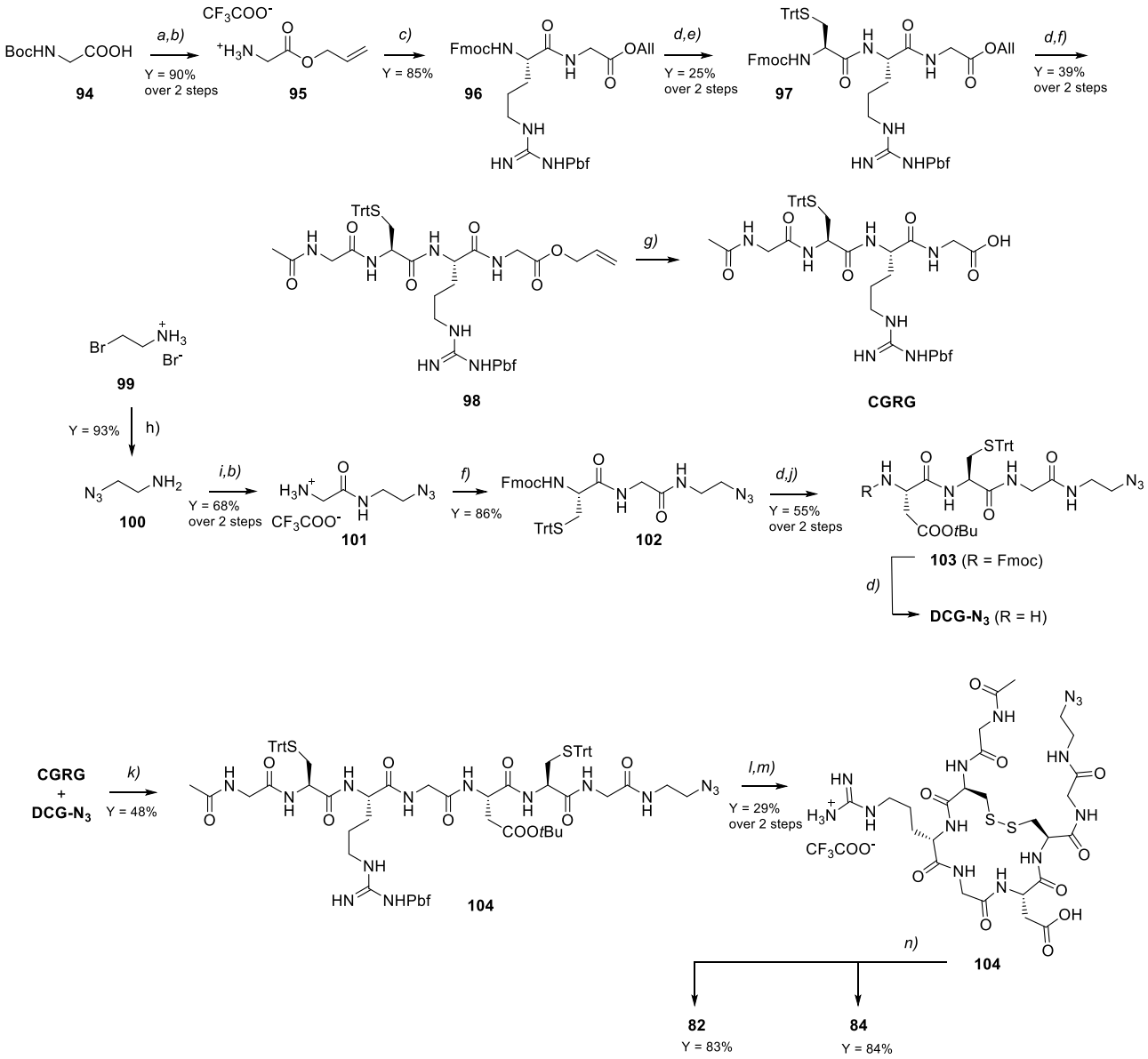
Compound **83** was prepared according to General Procedure I, using compound **92** (28.6 µmol) as the azide and compound **90** (37.2 µmol) as the alkyne. The crude was purified in RP-HPLC using the separation conditions reported below:

Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-12 min: 0-40% B ; t_r of the product = 10,75 min.

After freeze-drying, compound **83** was isolated as a fluffy white solid (18.05 mg, 61%)

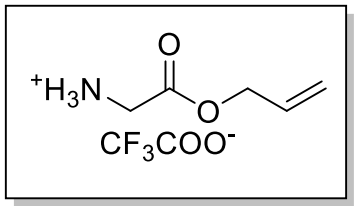
HRMS (ESI) *m/z* calculated for $[C_{35}H_{49}N_{14}O_{12}S_2]^+$: 921,3090 ; found: 921.3081 $[M+H]^+$

Synthesis of C-side 2HB-RGD ligand (**82**) and its negative control (**84**)



Scheme 18 REAGENTS AND CONDITIONS: a) Allyl bromide, Cs_2CO_3, MeCN dry, r.t., overnight b) TFA:CH_2Cl_2 1:2, 0 °C to r.t., 2h; c) Fmoc-Arg(Pbf)-OH, HATU, HOAt, *i*Pr_2NEt, DMF dry, 0°C to r.t, overnight; d) Et_2NH:MeCN 1:1, r.t; e) Fmoc-Cys(Trt)-OH, HATU, HOAt, *i*Pr_2NEt, DMF dry, 0 °C to r.t, overnight; f) *N*-acetylglycine, HATU, HOAt, *i*Pr_2NEt, DMF dry, 0 °C to r.t, overnight; g) $Pd(OAc)_2$, PPh_3, *N*-methylaniline, dry CH_2Cl_2, 0 °C to r.t, overnight; h) NaN_3, H_2O, 100 °C, overnight; i) EDCI, HOBt, Et_3N, DCM dry, r.t., 4h; j) Fmoc-Asp(OtBu)-OH, HATU, HOAt, *i*Pr_2NEt, dry DMF, 0 °C to r.t, overnight; k) [1] EDCI, *N*-hydroxysuccinimide, DMF dry, 0 °C to r.t, overnight; [2] *i*Pr_2NEt, r.t, 2h; l) TFA:TIS:H_2O 95:2.5:2.5, 0 °C to r.t, 2h; m) I_2, H_2O:MeCN 1:1, r.t, 1h; n) Alkyne **88** or **90**, $CuSO_4 \cdot 5H_2O$, sodium ascorbate, degassed H_2O/DMF, 40 °C, overnight.

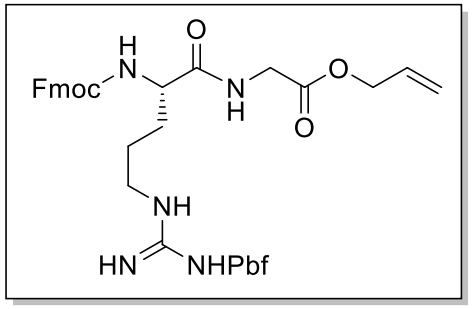
Gly-OAll (**95**)



$C_5H_9NO_2$
MW: 229.15 g/mol

Compound **95** was prepared according to a published procedure.[138]

Fmoc-Arg(Pbf)-Gly-OAll (**96**)

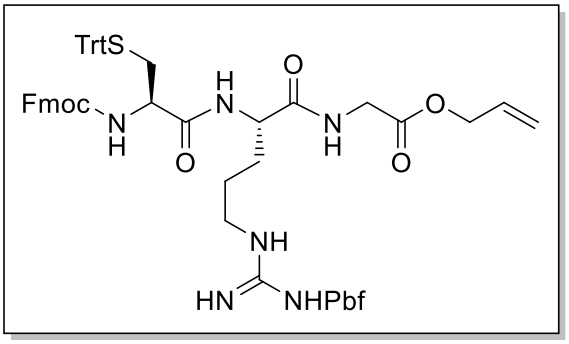


$C_{39}H_{47}N_5O_8S$
MW: 745.89 g/mol

Compound **96** was prepared according to General Procedure J using compound **95** (198.5 mg, 0.87 mmol) as amine and Fmoc-Arg(Pbf)-OH as carboxylic acid. The crude was purified with flash chromatography (DCM:MeOH 95:5) to afford the product as a pale yellow foam (550 mg, 85%).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.77 – 7.71 (m, 2H), 7.59 – 7.53 (m, 2H), 7.37 (t, J = 7.4 Hz, 2H), 7.30 – 7.24 (m, 2H, overlapped with solvent signal), 6.17 (bs, 2H), 5.94 – 5.80 (m, 2H) 5.35 – 5.19 (m, 2H), 4.60 (d, J = 5.8 Hz, 2H), 4.41 – 4.30 (m, 3H), 4.20 - 4.09 (m, 2H), 3.93 (dd, J = 17.7, 5.5 Hz, 1H), 3.36-3.23 (m, 2H), 2.92 (s, 2H), 2.58 (s, 3H), 2.51 (s, 3H), 2.08 (s, 3H), 2.00 – 1.89 (m, 1H), 1.78 – 1.66 (m, 1H), 1.43 (s, 6H) ^{13}C NMR (100 MHz, $CDCl_3$): δ 173.0, 170.1, 159.1, 156.6, 156.5, 144.0, 143.8, 141.4, 138.6, 132.5, 131.6, 127.8, 127.2, 125.3, 124.9, 120.1, 118.9, 117.8, 86.6, 67.2, 66.1, 54.1, 47.2, 43.3, 41.3, 30.3, 28.7, 25.2, 19.4, 18.1, 12.6; MS (ESI) *m/z* calcd. for $[C_{39}H_{47}N_5O_8SNa]^+$: 768.30 $[M+Na]^+$; found 768.50 $[M+Na]^+$.

Fmoc-Cys(Trt)-Arg(Pbf)-Gly-OAll (**97**)

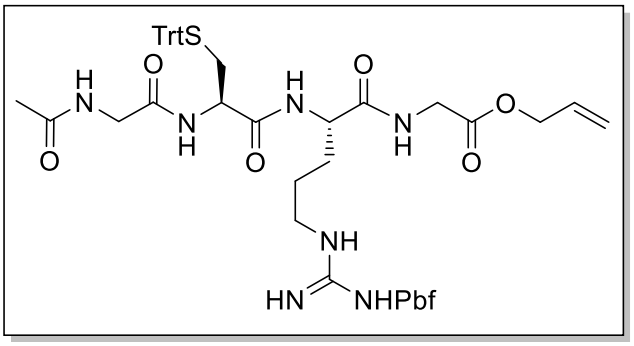


$C_{61}H_{66}N_6O_9S_2$
MW: 1091.35 g/mol

Compound **96** (120 mg, 0.201 mmol) was deprotected according to General Procedure K and the corresponding amine was coupled with Fmoc-Cys(Trt)-OH according to General procedure J. The crude was purified with flash chromatography (DCM:MeOH 95:5) to afford compound **97** as a pale yellow foam (55 mg, 25% over 2 steps).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.64 (t, J = 6.8 Hz, 2H), 7.48-7.41 (m, 3H), 7.32-7.26 (m, 6H, overlapped with solvent signal), 7.21-7.05 (m, 12H), 6.11 (bs, 2H), 5.75 (ddt, J = 16.2, 11.0, 5.8 Hz, 1H), 5.42 (m, 1H), 5.29-5.09 (m, 3H), 4.44 (d, J = 6.0 Hz, 2H), 4.28 – 4.14 (m, 2H), 4.02 (t , J = 7.0 Hz, 1H), 3.88-3.71 (m, 3H), 3.17-2.98 (m, 2H), 2.80 (s, 2H), 2.70-2.61 (m, 1H), 2.56-2.40 (m, 4H), 2.39 (s, 3H), 1.97 (s, 3H), 1.63-1.51 (m, 1H), 1.49-1.38 (m, 2H), 1.33 (s, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 171.9, 170.7, 169.7, 158.9, 156.4, 144.4, 143.9, 143.7, 141.3, 138.5, 132.8, 132.4, 131.7, 129.6, 128.2, 127.8, 127.2, 127.1, 125.2, 124.7, 120.1, 118.9, 117.6, 86.5, 67.3, 66.0, 54.2, 52.7, 47.1, 43.3, 41.2, 40.5, 34.1, 29.8, 29.6, 28.7, 25.2, 19.4, 18.1, 12.6; MS (ESI) *m/z* calcd. for $[C_{61}H_{66}N_6O_9S_2Na]^+$: 1113.42 $[M+Na]^+$; found 1113.70 [M+Na].

Ac-Gly-Cys(Trt)-Arg(Pbf)-Gly-OAll (**98**)



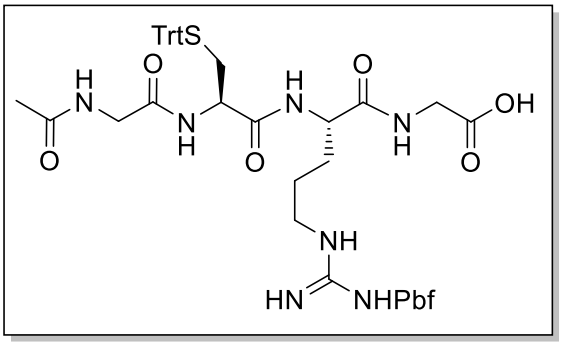
$C_{50}H_{61}N_7O_9S_2$
MW: 968,20 g/mol

Compound **97** (55.2 mg, 0.05 mmol) was deprotected according to General Procedure K and the corresponding amine was coupled with *N*-acetylglycine according to General procedure J. The crude was purified with flash chromatography (from 3% to 10% MeOH in DCM) to afford compound **98** as a pale yellow foam (19 mg, 39% over 2 steps).

^{1}H NMR (400 MHz, DMSO-d^6): δ 8.37-8.20 (m, 3H), 7.90 (d, J = 8.1Hz, 1H), 7.38 – 7,22 (m, 15H), 6.77 (bs, 1H), 6.46 (bs, 1H), 5.96 – 5.86 (m, 1H); 5.33 (dd, J = 17.5, 2.3 Hz, 1H), 5.22 (dd, J = 10.4,

1.2 Hz, 1H), 4.60 – 4.57 (m, 2H), 4.38 – 4.33 (m, 1H), 4.28 – 4.22 (m, 1H), 3.84 – 3.82 (m, 1H), 3.74 – 3.70 (m, 1H), 3.04 – 2.97 (m, 4H), 2.50 (s, 3H), 2.45 – 2.41 (m, 5H), 2.03 (s, 3H), 1.87 (s, 3H), 1.73 – 1.64 (m, 1H), 1.59 – 1.37 (m, 9H), 1.33 – 1.28 (m, 2H);^{13}C NMR (100 MHz, DMSO-d^{6}): *δ* 172.5, 171.0, 170.4, 170.2, 170.1, 158.5, 157.0, 148.8, 145.2, 138.3, 135.0, 133.3, 132.5, 130.1, 129.1, 128.8, 128.5, 127.8, 127.6, 125.3, 118.9, 117.3, 87.3, 65.8, 53.1, 52.6, 43.5, 43.1, 41.6, 34.5, 29.3, 23.4, 20.0, 18.6, 13.3; MS (ESI) *m/z* calcd. for $[C_{50}H_{61}N_7O_9S_2Na]^+$: 990.37 $[M+Na]^+$; found 990.13 $[M+Na]^+$.

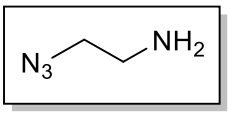
Ac-Gly-Cys(Trt)-Arg(Pbf)-Gly-OH (**GCRG**)



$C_{47}H_{57}N_7O_9S_2$
MW: 928.13 g/mol

$Pd(OAc)_2$ and PPh_3 were dissolved in DCM dry under N_2 and stirred until the solution turns bright yellow. The so-formed solution has been added to a solution of compound **98** in DCM dry under N_2. To the resulting solution, *N*-methylaniline has been added and the reaction mixture was stirred at r.t. and checked by TLC. After 1h, no conversion was detected. Another aliquot of *N*-methylaniline WAS added. After 5h, TLC showed full consumption of compound **98**. The mixture was concentrated in vacuo and then partitioned between AcOEt and 1M $KHSO_4$. The organic phase was dried over Na_2SO_4 and concentrated in vacuo. The crude was filtered over a short pad of silica (1% formic acid in DCM:MeOH 9:1) and used in the next step without further purifications.

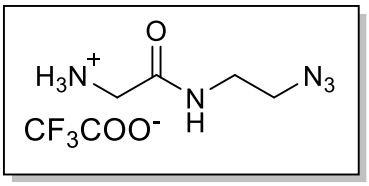
2-azidoethanamine (**100**)



$C_2H_6N_4$
MW: 86,10 g/mol

Compound **100** was prepared according to a published procedure.[139]

2-amino-*N*-(2-azidoethyl)-acetamide (**101**)

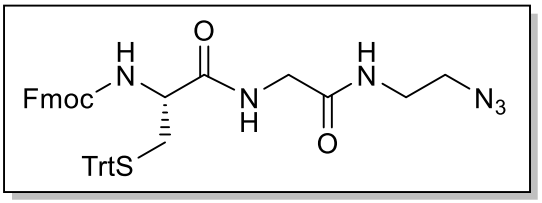


$C_4H_{10}N_5O$
MW: 143,15 g/mol

Boc-Gly-OH (250 mg, 1.43 mmol, 1 eq) was dissolved in DCM dry under N_2 and cooled down at 0°C. To this solution, compound **100** (1M solution in DMF, 2.3 mL, 2.28 mmol, 1.6 eq), triethylamine (218.8 µL, 1.57 mmol, 1.1 eq) and EDCl (300.9 mg, 1.57 mmol, 1.1 eq) were sequentially added. The reaction mixture was stirred at 0°C for 2h and then at r.t. overnight. The mixture was dried in vacuo and the resulting crude was dissolved in AcOEt (50 mL). The organic phase was washed with 1M $KHSO_4$ (3x10 mL), sat. $NaHCO_3$ (1x10 mL) and brine (1x10 mL), dried over Na_2SO_4 and concentrated in vacuo. The crude was dissolved in DCM dry (12.8 mL) under N_2 and then cooled down to 0°C. TFA (6.4 mL) was then added to this solution and the reaction was stirred at r.t. for 1h. The reaction mixture was concentrated in vacuo and the residual TFA was removed with co-evaporation with MeOH (3 x 25 mL) obtaining the trifluoroacetate salt of compound **101** as a brown solid (249 mg, 67.6% over 2 steps).

^{1}H NMR (400 MHz, CD_3OD): *δ* 3.69 (s, 2H), 3.45 (s, 4H). ^{13}C NMR (100 MHz, $CDCl_3$): *δ* 167.7, 51.5, 41.5, 40.0; MS (ESI) *m/z* calcd. for $[C_4H_{10}N_5O]^+$: 144.09 $[M+H]^+$; found 144.00 [M+H].

Fmoc-Cys(Trt)-Gly-N_3 (**102**)

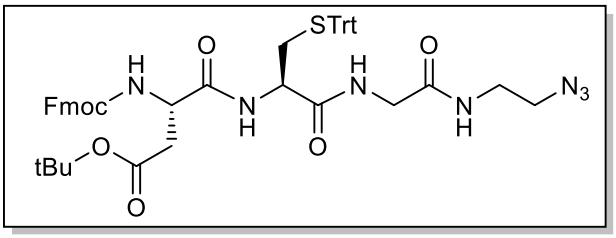


$C_{41}H_{38}N_6O_4S$
MW: 710,85 g/mol

Compound **102** was prepared according to General Procedure *J* using compound **101** (114.4 mg, 0.445 mmol) as amine and Fmoc-Cys(Trt)-OH (236.9 mg, 0.405 mmol) as carboxylic acid. The crude was purified with flash chromatography (AcOEt:Hex 3:2) to afford the product as a white foam (248 mg, 86%).

^{1}H NMR (400 MHz, $CDCl_3$): *δ* 7.82 – 7.75 (m, 2H), 7.57 (d, *J* = 7.5 Hz, 2H), 7.46 – 7.39 (m, 8H), 7.36 – 7.21 (m, 8H, overlapped with solvent signal), 6.81 (bs, 1H), 6.26 (t, *J* = 6.0 Hz, 1H), 4.93 (d, *J* = 6.1 Hz, 1H), 4.47 (d, *J* = 6.4 Hz, 2H), 4.20 (t, *J* = 6.4 Hz, 1H), 3.98 – 3.78 (m, 2H), 3.62 – 3.56 (m, 1H), 3.33 (s, 4H), 2.80 – 2.69 (m, 2H).^{13}C NMR (100 MHz, DMSO-d_6): *δ* 170.1, 169.6, 155.8, 145.2, 144.7, 144.6, 141.6, 130.0, 129.0, 128.6, 128.0, 127.7, 126.2, 121.0, 66.9, 66.8, 54.7, 50.8, 47.5, 43.0, 39.0; MS (ESI) m/z calcd. for $[C_{41}H_{38}N_6O_4SNa]^+$: 733.26 $[M+Na]^+$; found 732.95 [M+Na].

Fmoc-Asp(OtBu)-Cys(Trt)-Gly-N_3 (**103**)

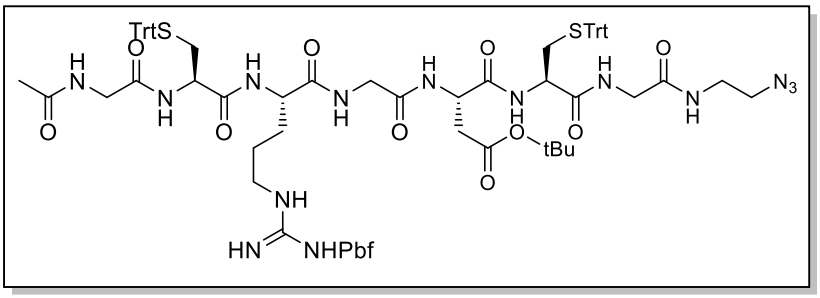


$C_{49}H_{51}N_7O_7S$
MW: 882,05 g/mol

Compound **102** (248.3 mg, 0.349 mmol) was deprotected according to General Procedure K and the corresponding amine was coupled with Fmoc-Asp(OtBu)-OH according to General procedure J. The crude was purified with flash chromatography (gradient AcOEt:Hex from 3:2 to 4:1) to afford compound **103** as a white foam (168 mg, 54.7% over 2 steps).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.77 (d, J = 7.6 Hz, 2H), 7.54 (d, J = 7.5 Hz, 2H), 7.44 – 7.34 (m, 7H), 7.34 – 7.23 (m, overlapped with solvent signal, 9H), 7.23 – 7.15 (m, 3H), 7.09 (bs, 1H), 6.86 (bs, 1H), 6.48 (bs, 1H), 5.66 (d, J = 7.6 Hz, 1H), 4.47 – 4.29 (m, 3H), 4.18 (t, J = 6.9 Hz, 1H), 4.01 – 3.73 (m, 3H), 3.34 (s, 4H), 3.00 – 2.90 (m, 1H), 2.78 – 2.62 (m, 3H), 1.41 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 171.4, 170.9, 169.9, 169.4, 156.2, 144.1, 143.6, 143.5, 141.2, 129.3, 128.1, 127.7, 127.0, 126.9, 124.9, 82.4, 67.3, 53.5, 51.4, 50.1, 47.0, 43.1, 38.8, 37.2, 32.7, 29.6, 27.9; MS (ESI) m/z calcd. for $[C_{49}H_{51}N_7O_7SNa]^+$: 904.35 $[M+Na]^+$; found 904.67 [M+Na].

Ac-Gly-Cys(Trt)-Arg(Pbf)-Gly-Asp(OtBu)-Cys(Trt)-Gly-N_3 (**104**)

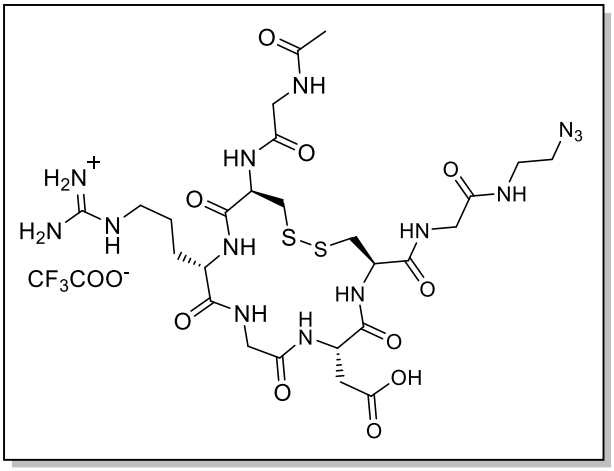


$C_{81}H_{96}N_{14}O_{13}S_3$
MW: 1569,92 g/mol

Carboxylic acid **GCRG** (234 mg, 0.252 mmol, 1 eq) and *N*-hydroxysuccinimide (31.9 mg, 0.277 mmol, 1.1 eq) were dissolved in DMF dry under N_2. To this solution, EDCI (53.1 mg, 0.277 mmol, 1.1 eq) was added at r.t. and the reaction mixture was stirred overnight. Compound **103** (280 mg, 0.3 mmol, 1.2 eq) was deprotected according to General Procedure K and the corresponding amine **DCG-N_3**, without any further purifcations, was added to the reaction mixture containing the NHS derivative of compound **99** together with *N,N*-diisopropylethylamine (131.7 µL, 0.756 mmol, 3 eq). The reaction mixture was stirred at r.t. for 2h. The mixture was concentrated in vacuo and the crude was purified directly with column chromatography (gradient from 5% MeOH to 10% MeOH in DCM), affording compound **104** (190.1 mg, 48%) as a white solid.

MS (ESI) m/z calcd. for $[C_{81}H_{96}N_{14}O_{13}S_3Na]^+$: 1591,63 $[M+Na]^+$, found 1592,20 [M+Na]; m/z calcd. for $[C_{81}H_{95}N_{14}O_{13}S_3]^-$: 1567,64 $[M-H]^-$, found 1568,00 [M-H]; m/z calcd. for $[C_{81}H_{96}N_{14}O_{13}S_3Cl]^-$: 1603,61 $[M+Cl]^-$, found 1603,86 [M+Cl];

Ac-RGD-2C-N_3 (**105**)



$C_{26}H_{42}N_{14}O_{10}S_2$
MW: 888,85 g/mol

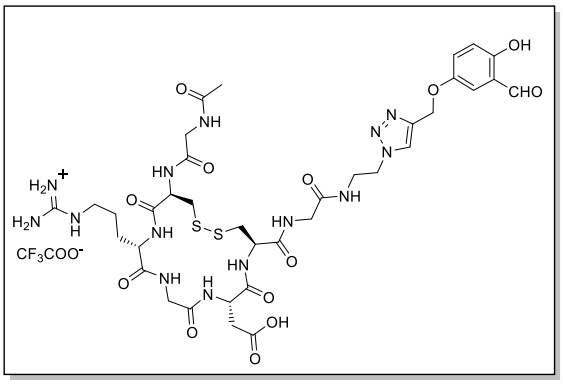
Compound **104** (53.8 mg, 0.0343 mmol) was deprotected according to General Procedure D and used in the next step of intramolecular disulfide bond formation (General Procedure E). The crude of this reaction was purified in semi-preparative RP-HPLC using the separation conditions reported below:

Flow: 10 mL/min; UV channels: 210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-12 min: 0 - 40% B; t_r of the product = 9.5 min.

The desired product is freeze-died, obtaining its trifluoroacetate salt that appears as a white solid (8,96 mg, 29.4% over 2 steps).

MS (ESI) *m/z* calculated for $[C_{26}H_{43}N_{14}O_{10}S_2]^+$: 775.27, found: 775.44 $[M+H]^+$; calculated for $[C_{26}H_{42}N_{14}O_{10}S_2Na]^+$: 797.25, found: 797.44 $[M+Na]^+$

C-Side 2HB-RGD-2C (**82**)



$C_{36}H_{50}N_{14}O_{13}S_2$
MW: 1065,02 g/mol

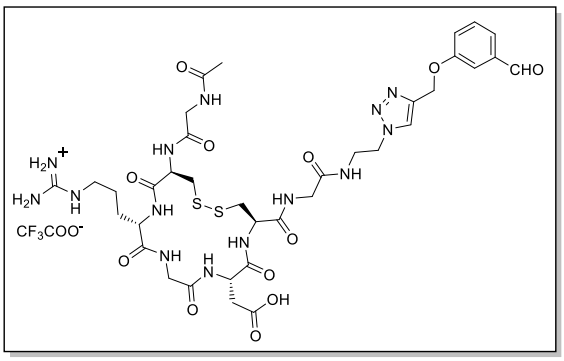
Compound **82** was prepared according to General Procedure I, using compound **107** (77.1 µL; 7.71 µmol) as the azide and compound **88** (100.2 µL; 10.02 µmol) as the alkyne. The crude was purified in RP-HPLC using the separation conditions reported below:

Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-12 min: 0-40% B ; t_r of the product = 11,5 min.

The desired product was freeze-dried in H_2O + 0.1% TFA and obtained as a fluffy white solid (6.86 mg, 83%)

HRMS (ESI) *m/z* calculated for $[C_{36}H_{51}N_{14}O_{13}S_2]^+$: 951,3196 ; found: 951,3195 $[M+H]^+$.

C-Side negative control RGD-2C (**84**)



$C_{36}H_{50}N_{14}O_{12}S_2$
MW: 1049,03 g/mol

Compound **84** was prepared according to General Procedure I, using compound **107** (77.1 µL; 7.71 µmol) as the azide and compound **90** (100.2 µL; 10.02 µmol) as the alkyne. The crude was purified in RP-HPLC using the separation conditions reported below:

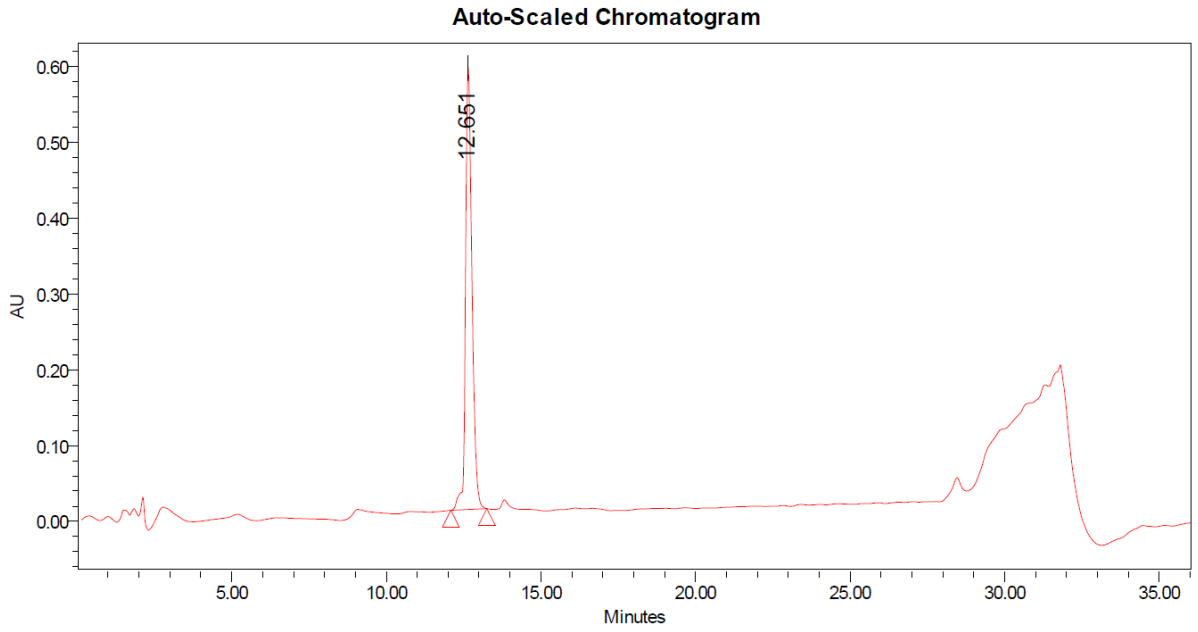
Flow: 10 mL/min; UV channels:210 nm; 221 nm; A: H_2O + 0.1% TFA, B: MeCN without TFA. Gradient: 0-2 min: 0% B, 2-12 min: 0-40% B; t_r of the product = 12 min.

The desired product was freeze-dried in H_2O + 0.1% TFA and obtained as a fluffy white solid (6.83 mg, 84.5%)

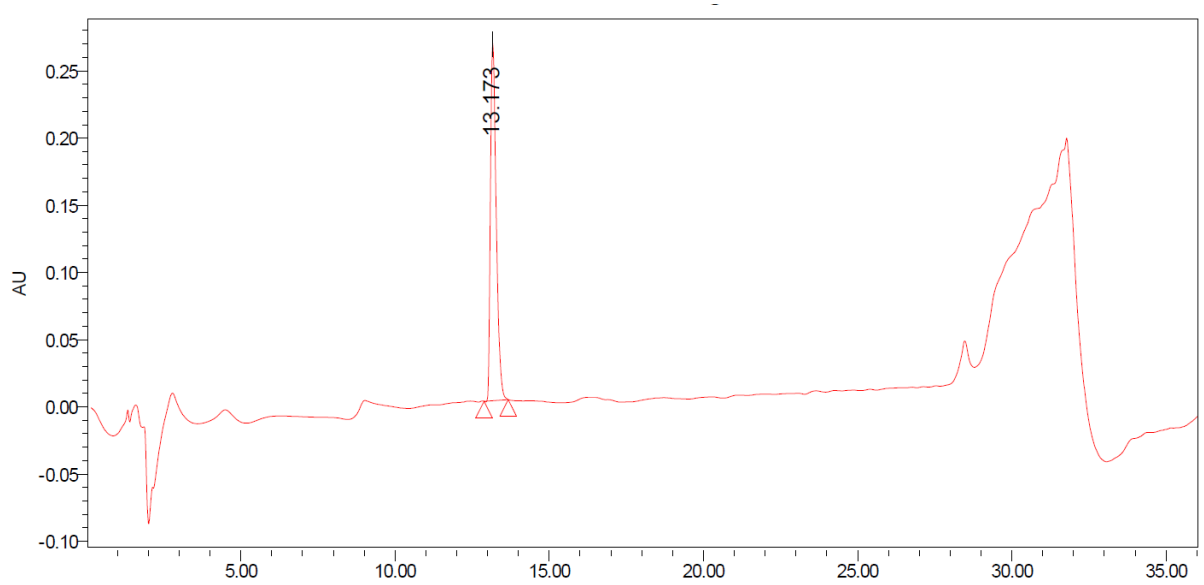
HRMS (ESI) *m/z* calculated for $[C_{36}H_{51}N_{14}O_{12}S_2]^+$: 935,3247 ; found: 951,3235 $[M+H]^+$.

HPLC traces of final products

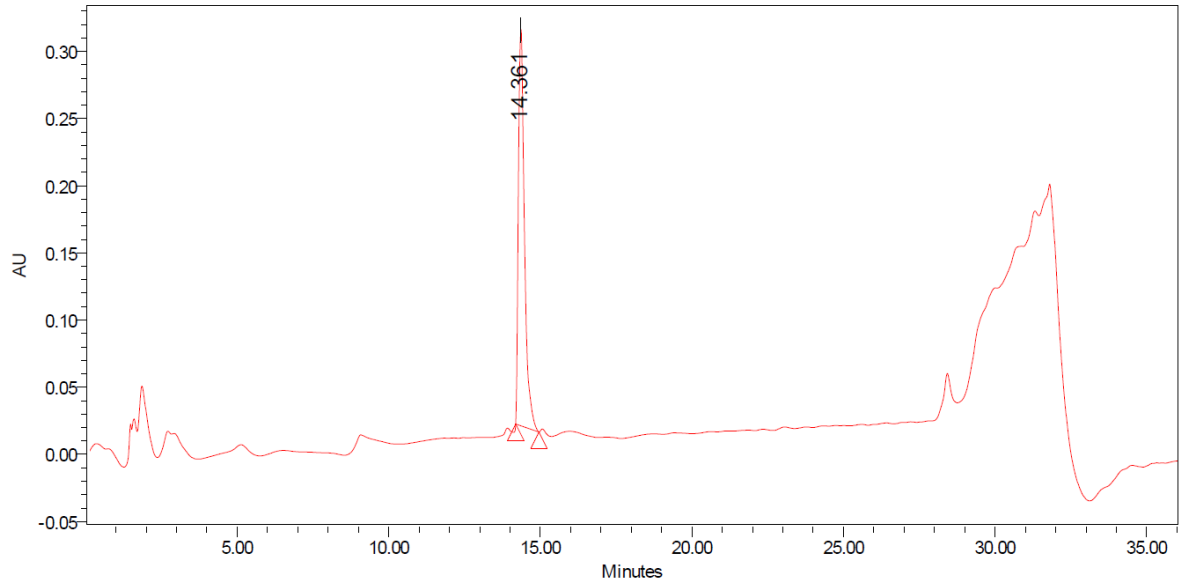
RGD-2C-RGD (**36**)



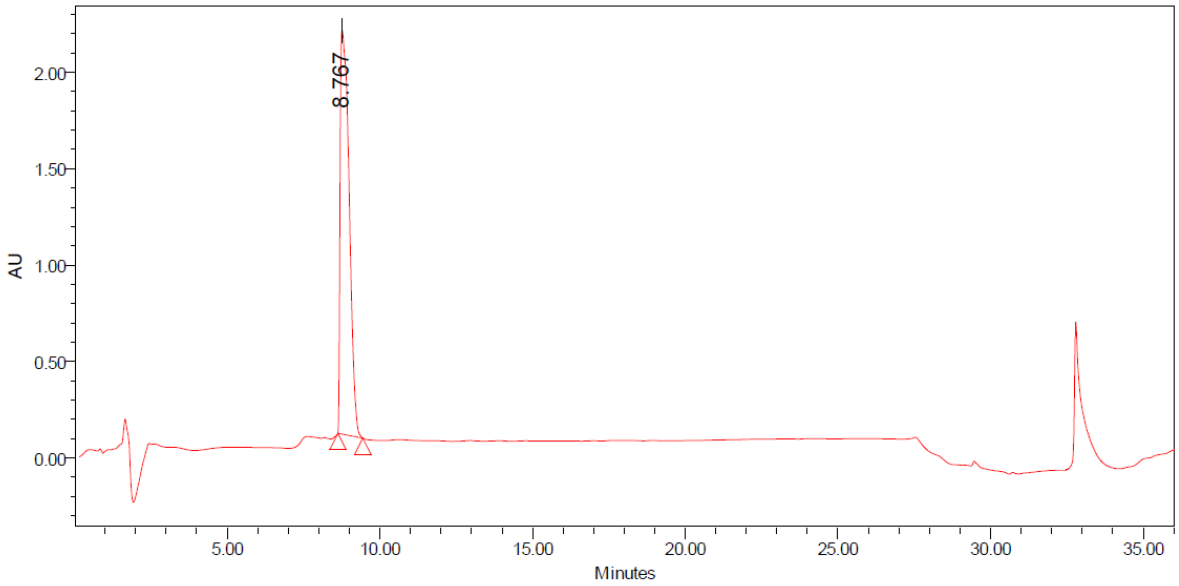
RGD-2C-RβAD (**37**)



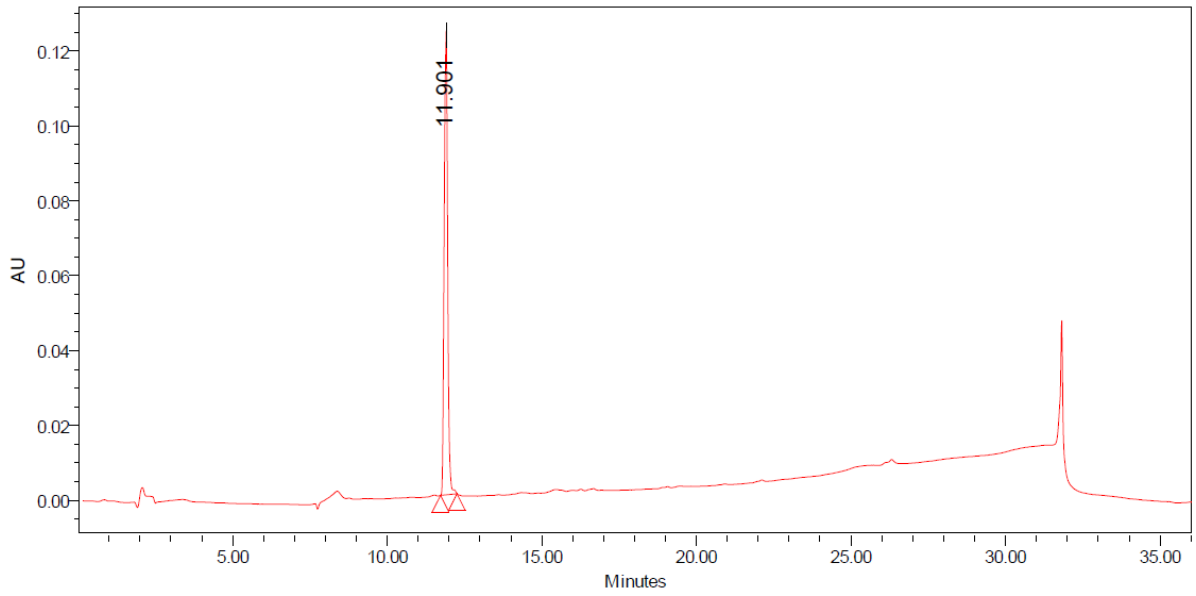
RβAD-2C-RβAD (**38**)



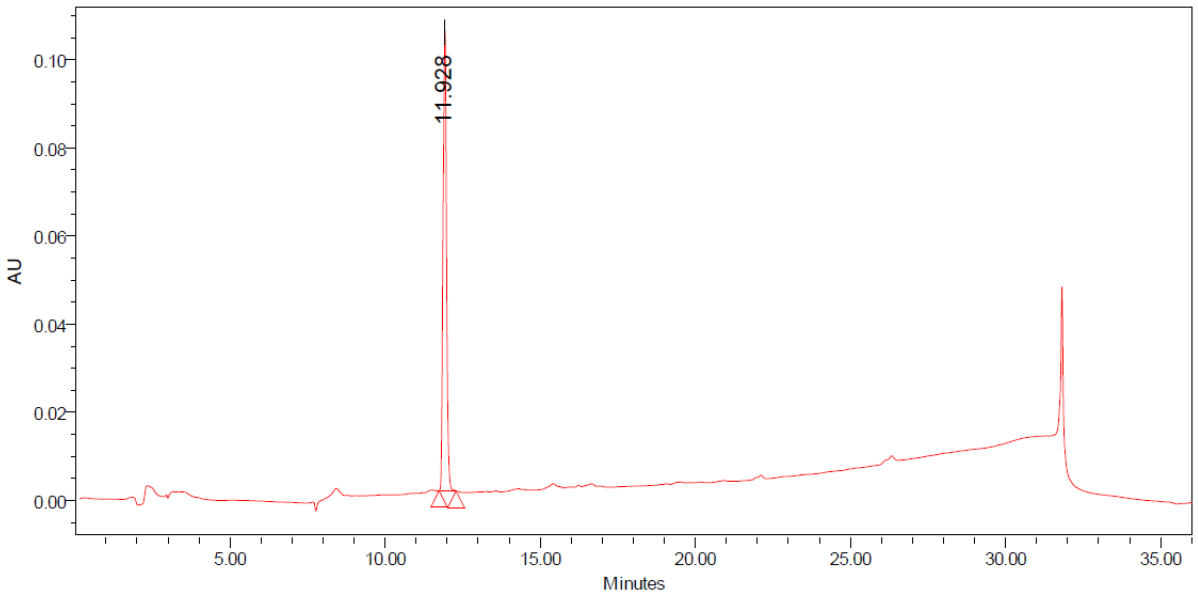
RGD-2C (**39**)



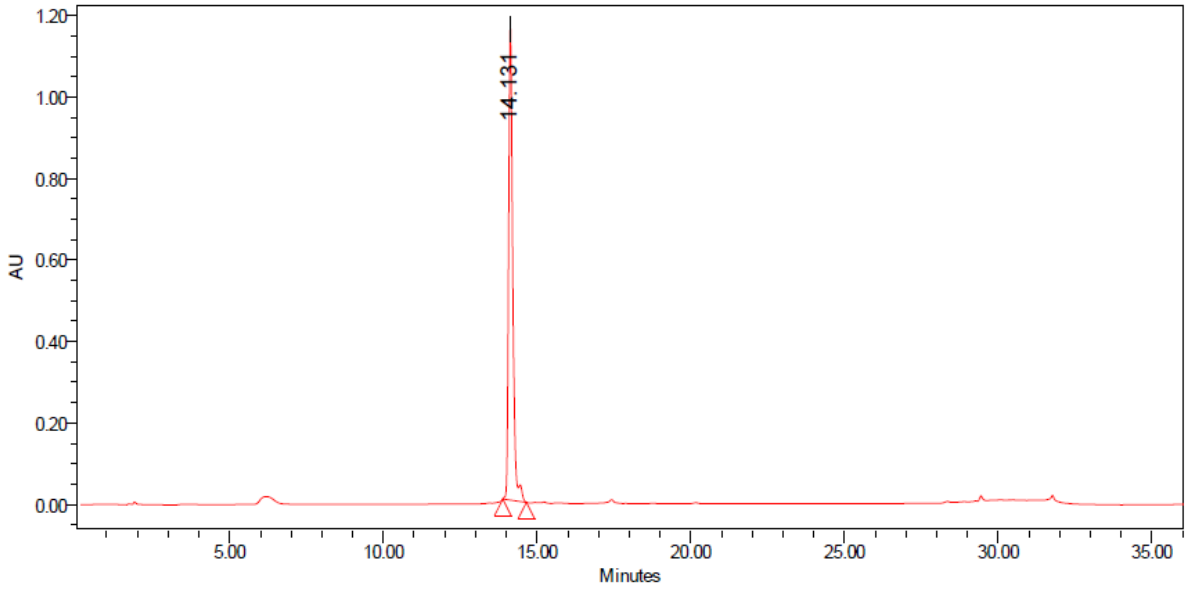
N-Side 2HB-RGD-2C (**81**)



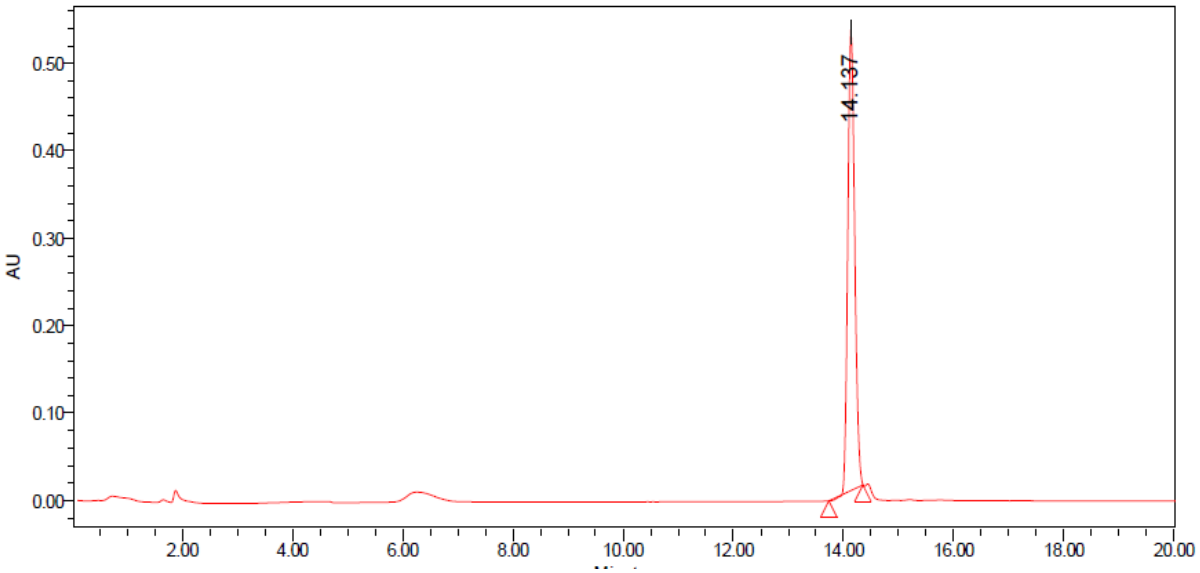
N-Side negative control RGD-2C (**83**)



C-Side 2HB-RGD-2C (**82**)

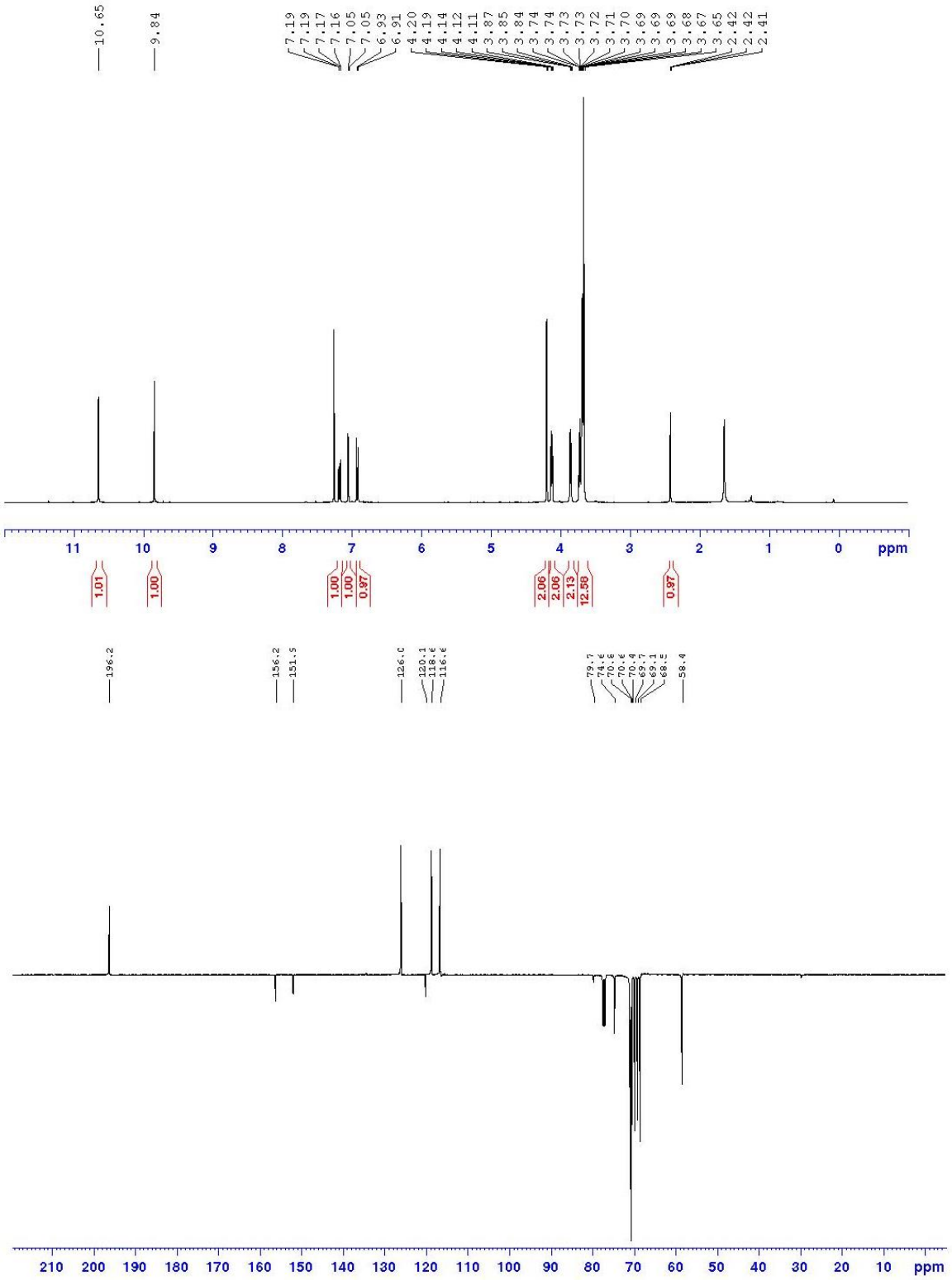


C-Side negative control RGD-2C (**84**)

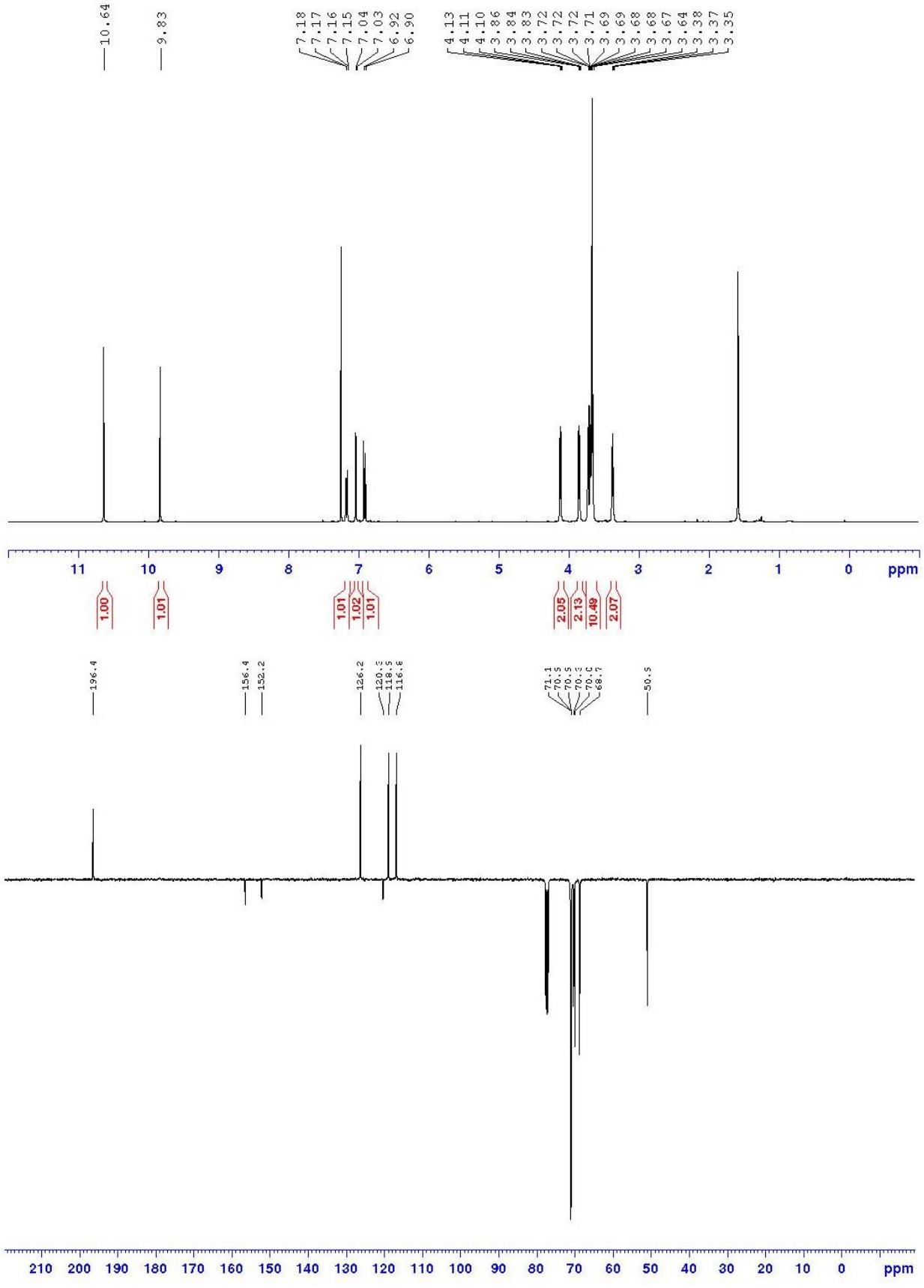


Appendix of NMR data

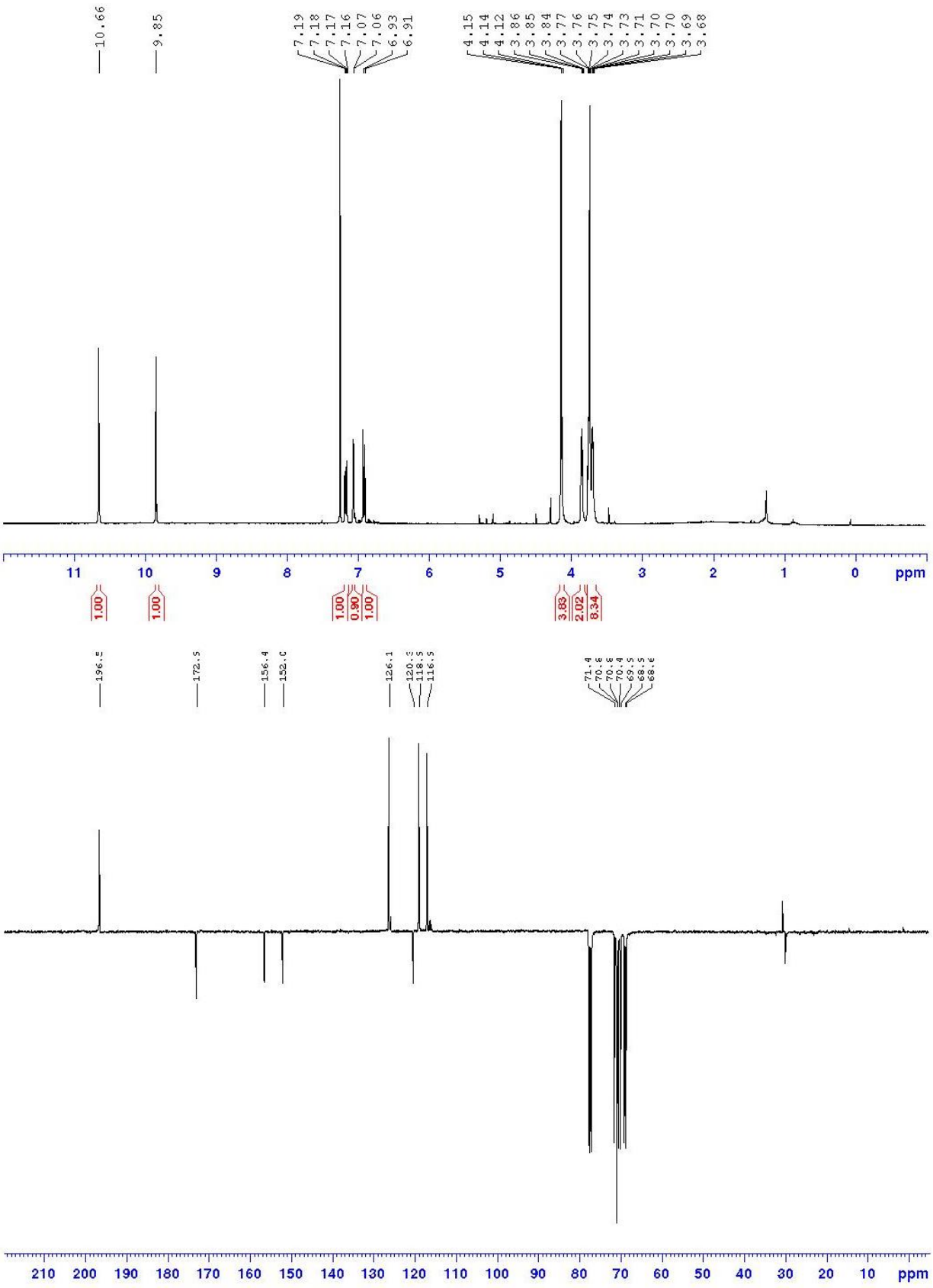
5-((3,6,9,12-tetraoxapentadec-14-yn-1-yl)oxy)-2-hydroxybenzaldehyde (**63**)



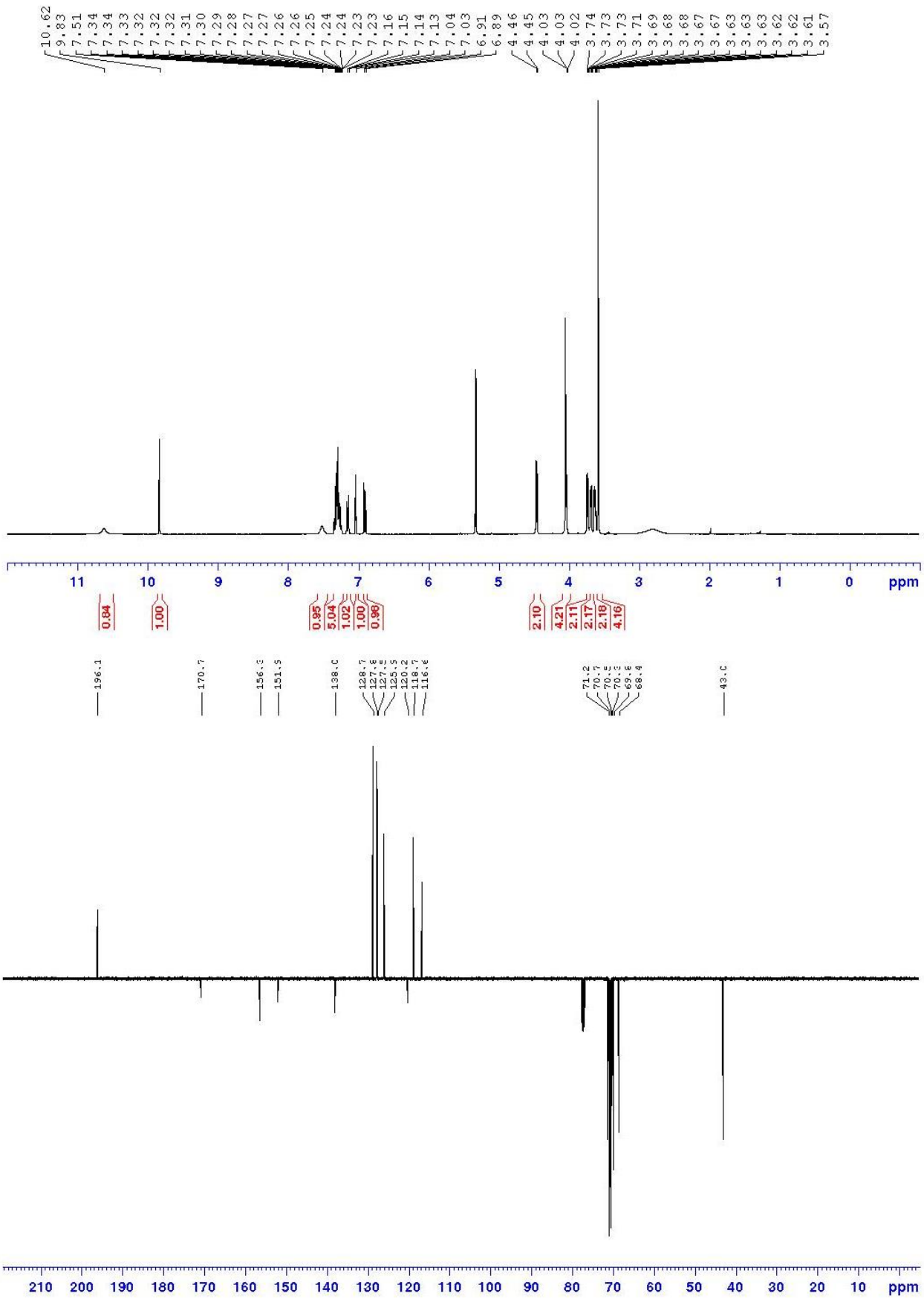
5-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxy)-2-hydroxybenzaldehyde (**64**)



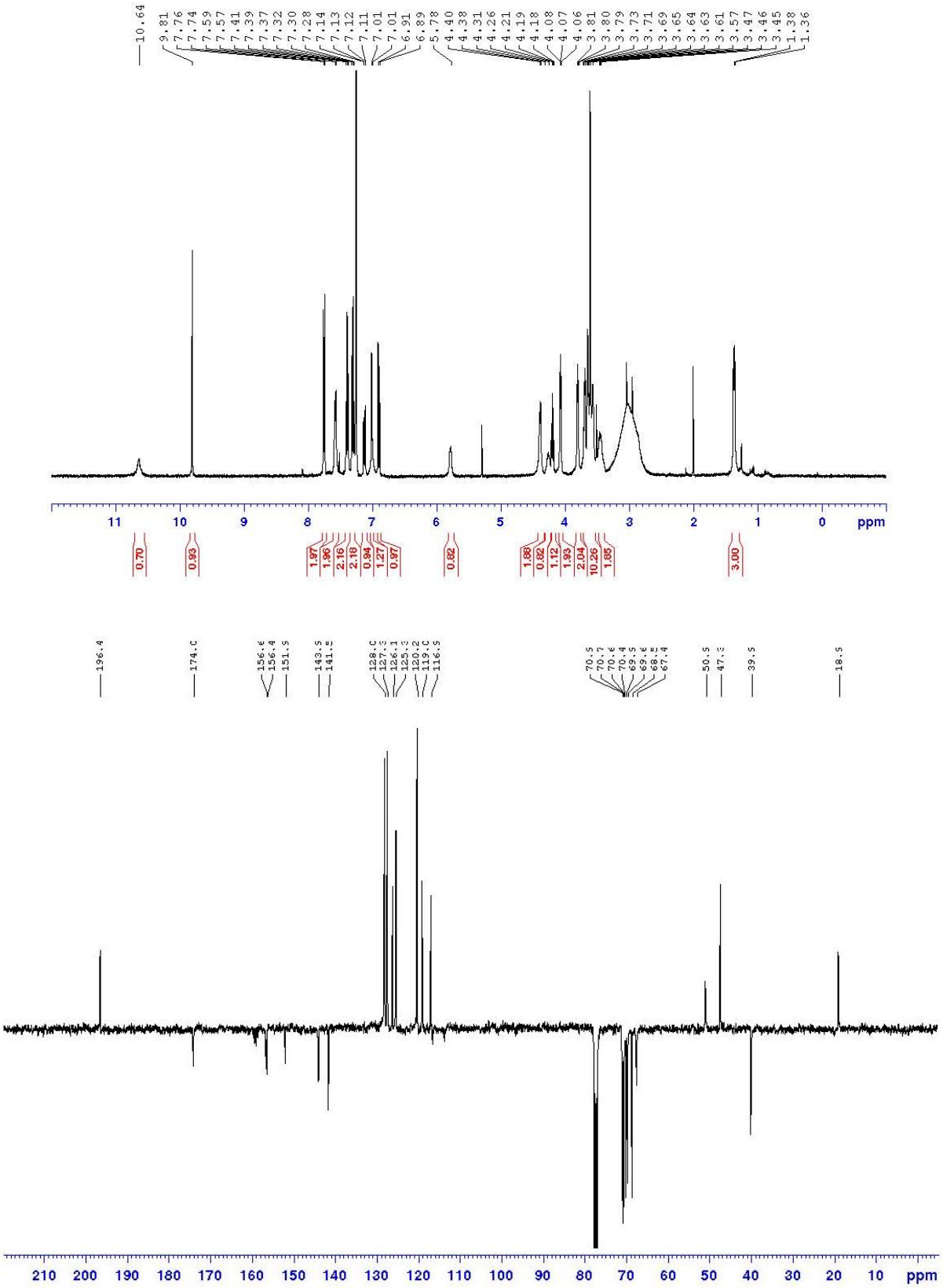
2-(2-(2-(2-(3-formyl-4-hydroxyphenoxy)ethoxy)ethoxy)ethoxy)acetic acid (**65**)



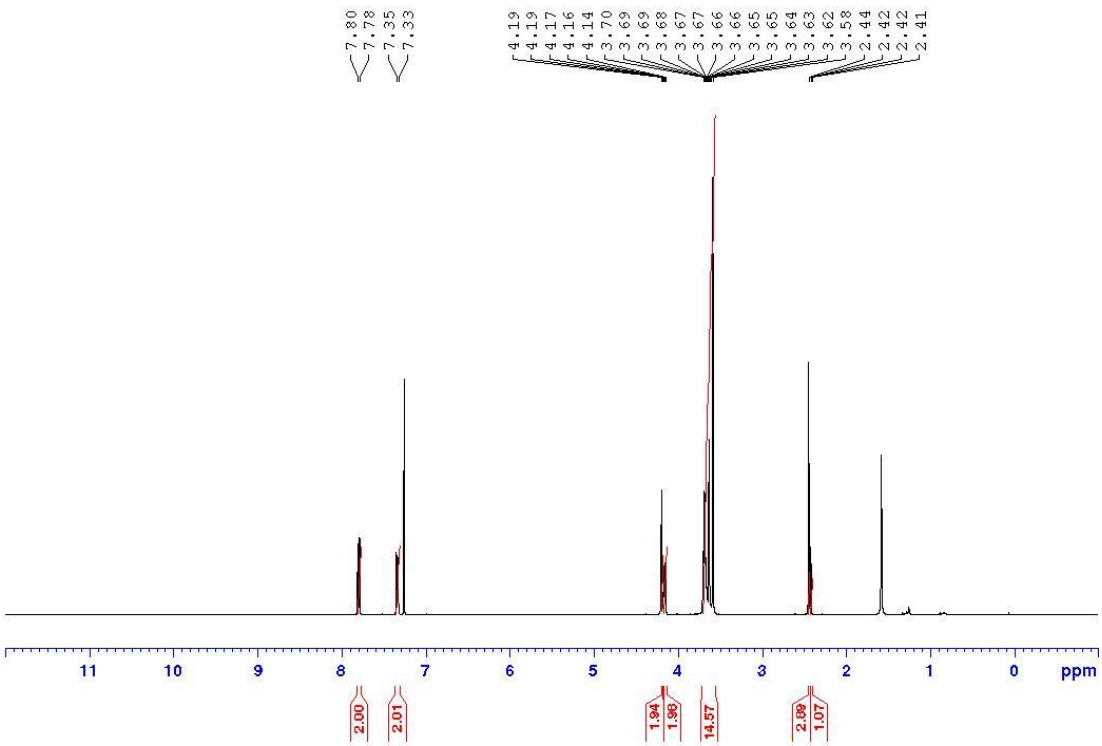
N-benzyl-2-(2-(2-(2-(3-formyl-4-hydroxyphenoxy)ethoxy)ethoxy)ethoxy)acetamide (**65-amide**)

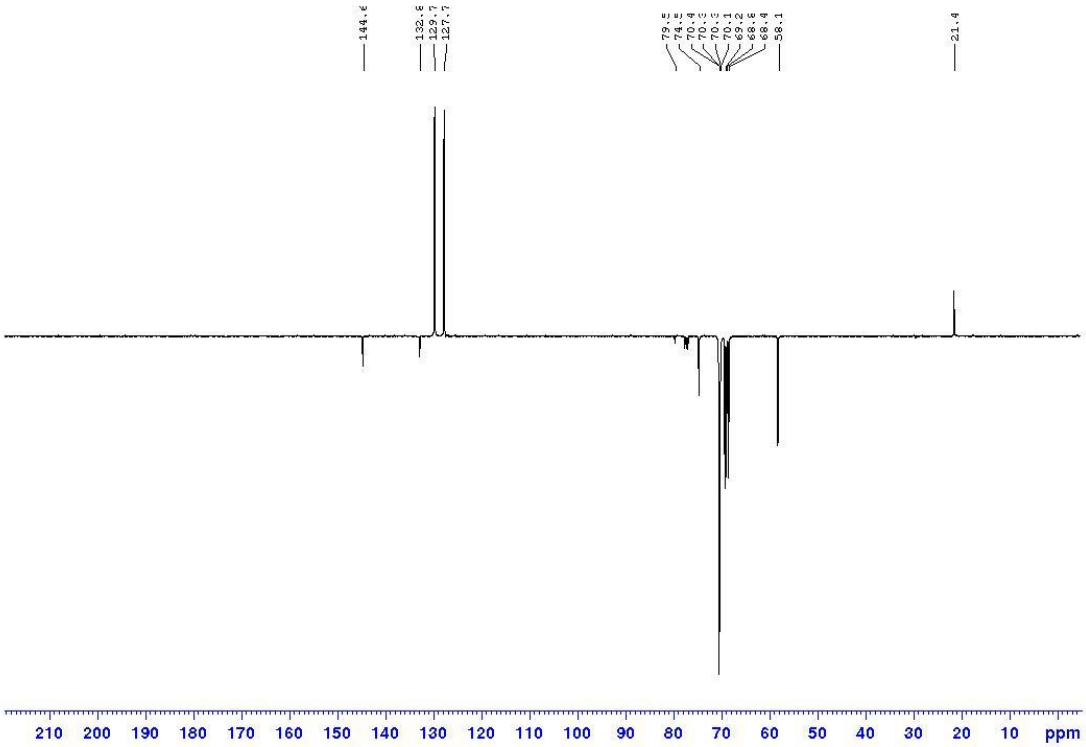


(9H-fluoren-9-yl)methyl (S)-(1-(3-formyl-4-hydroxyphenoxy)-13-oxo-3,6,9-trioxa-12-azapentadecan-14-yl)carbamate (**66-amide**)

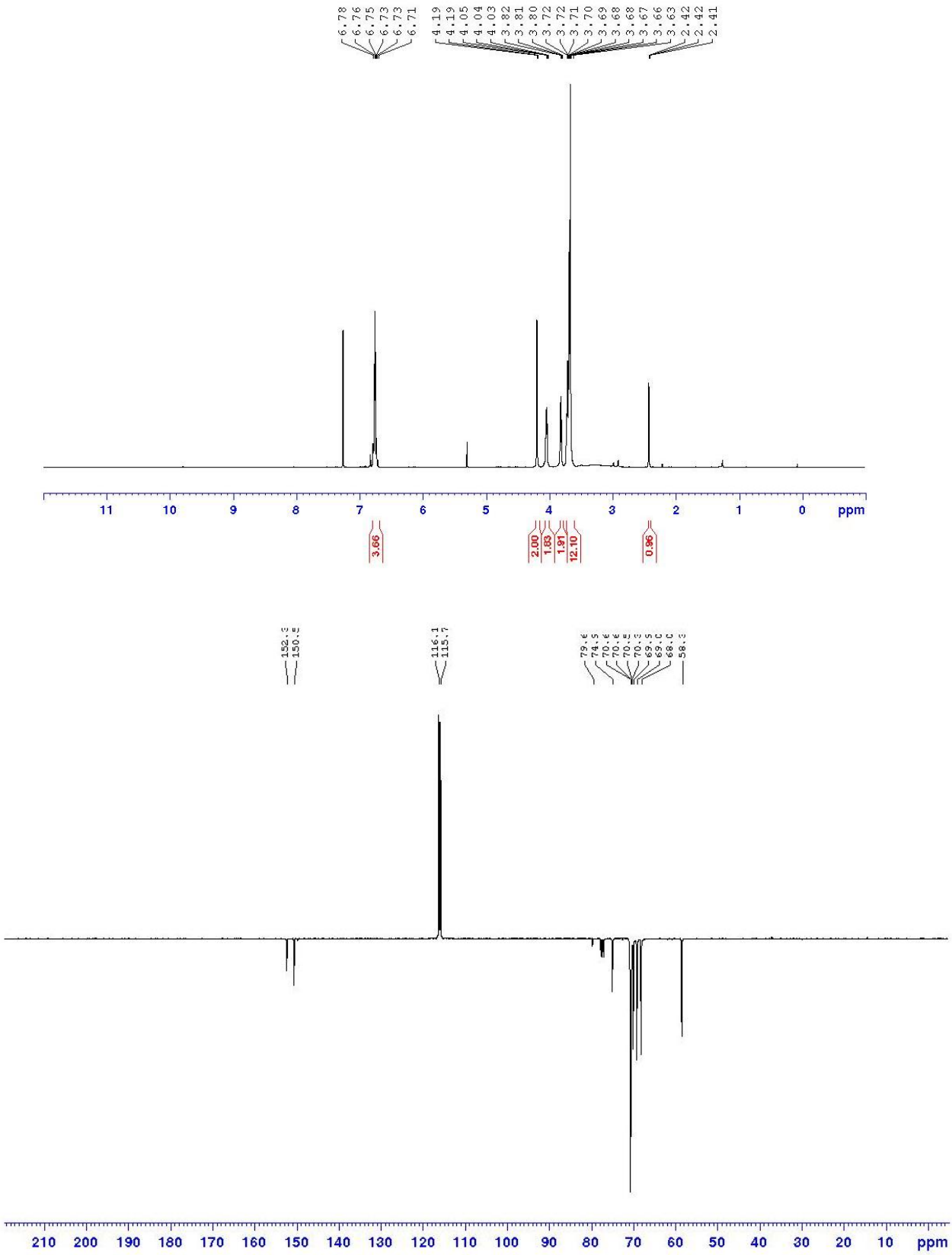


3,6,9,12-Tetraoxapentadec-14-yn-1-yl 4-methylbenzenesulfonate (**68**)

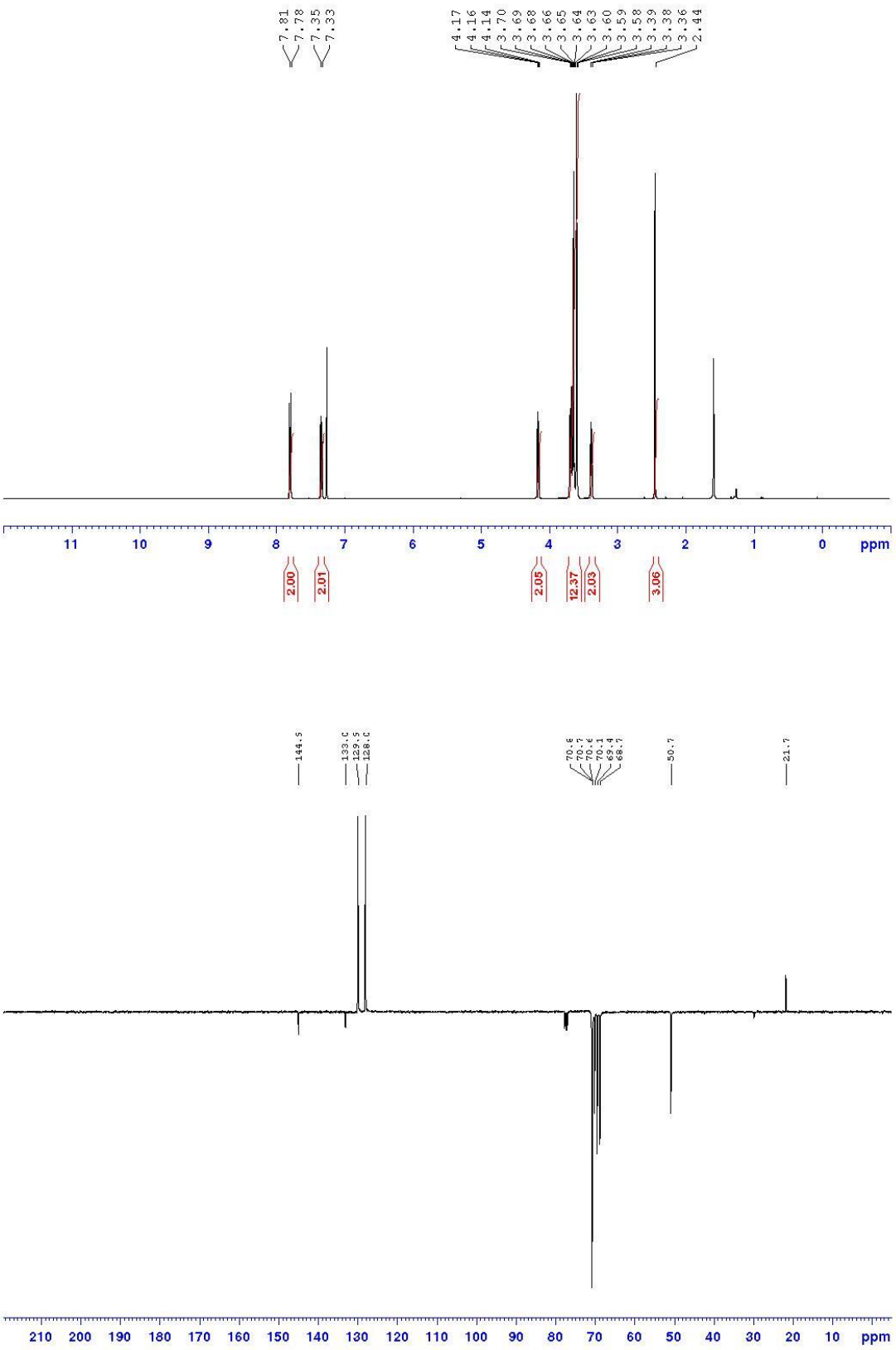




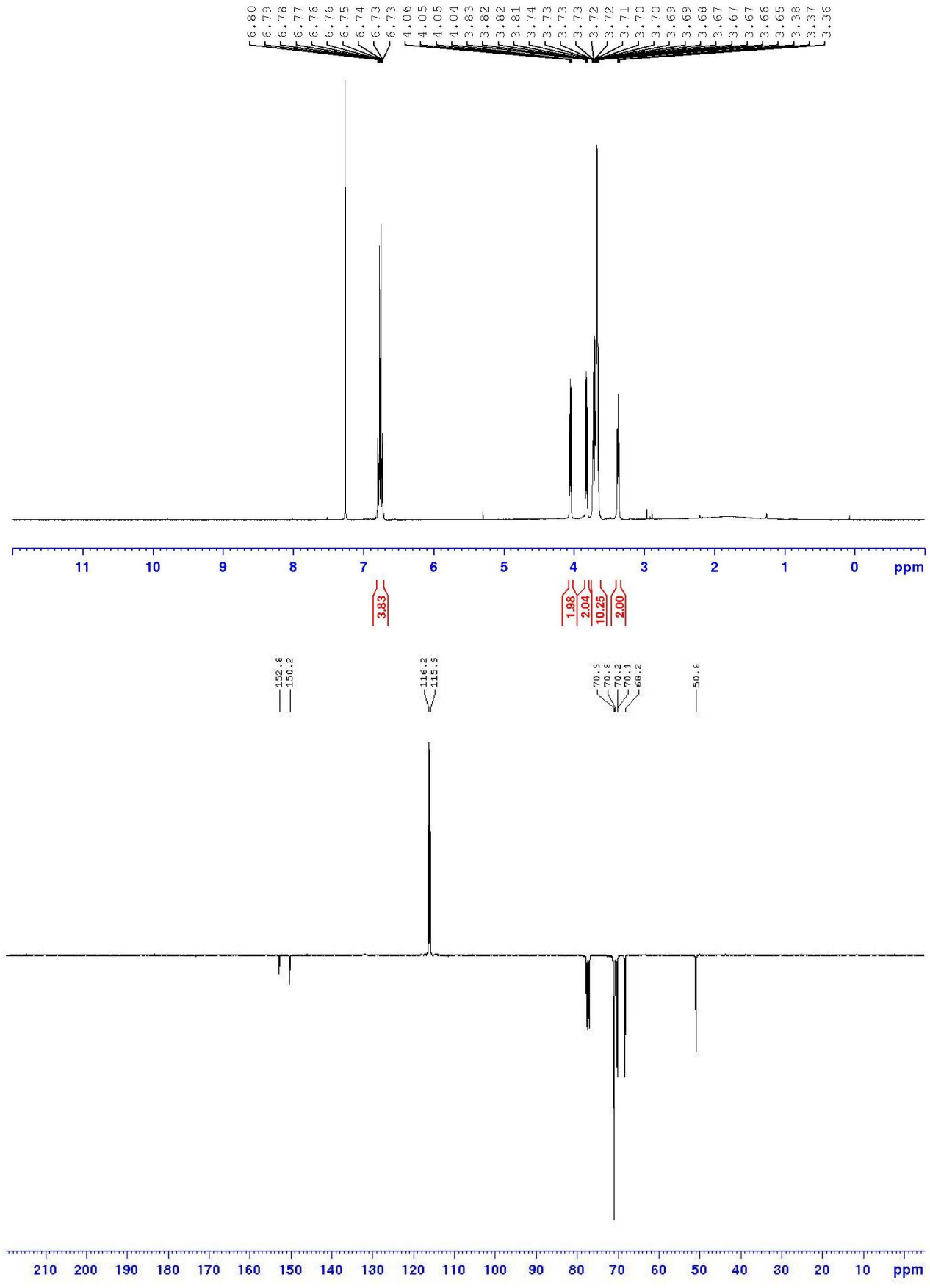
4-((3,6,9,12-tetraoxapentadec-14-yn-1-yl)oxy)phenol (**69**)



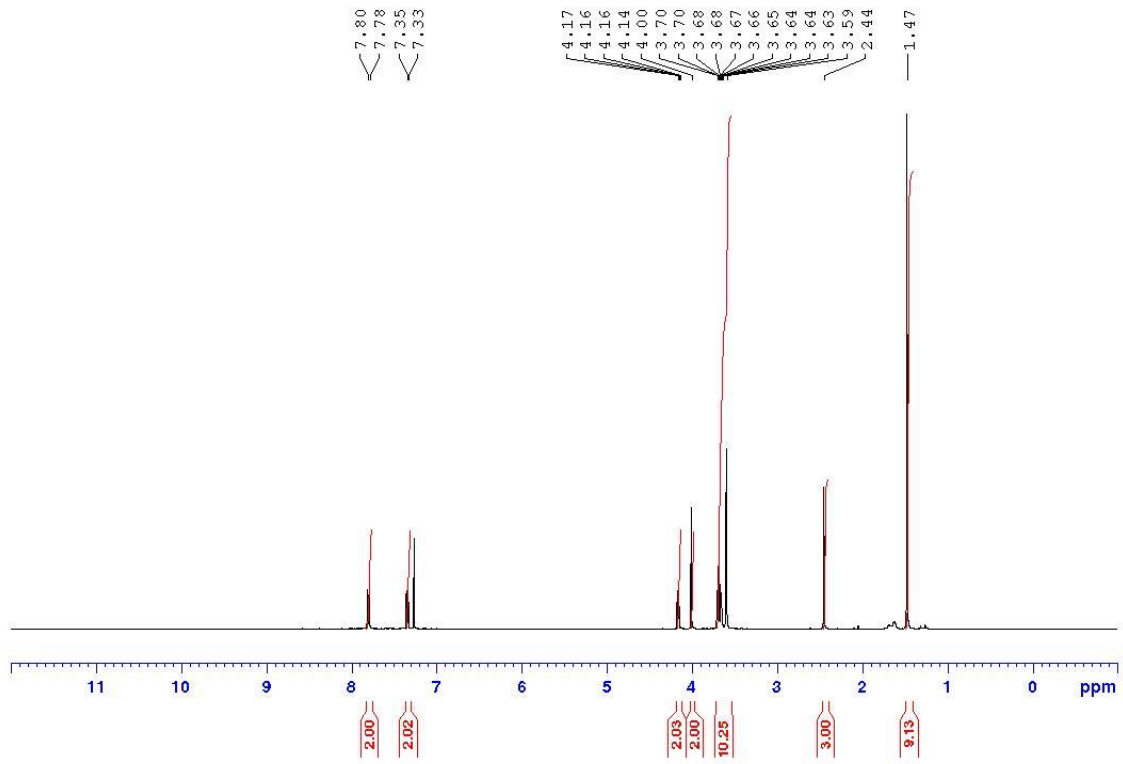
2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (**71**)

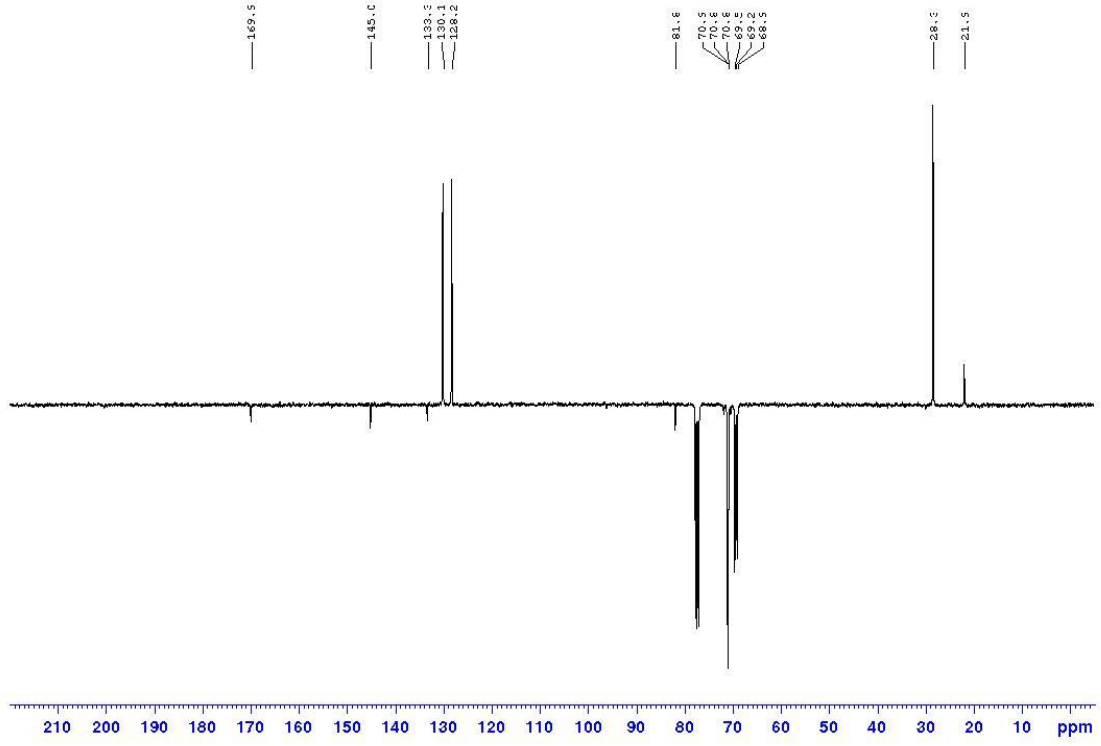


4-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxy)phenol (**72**)

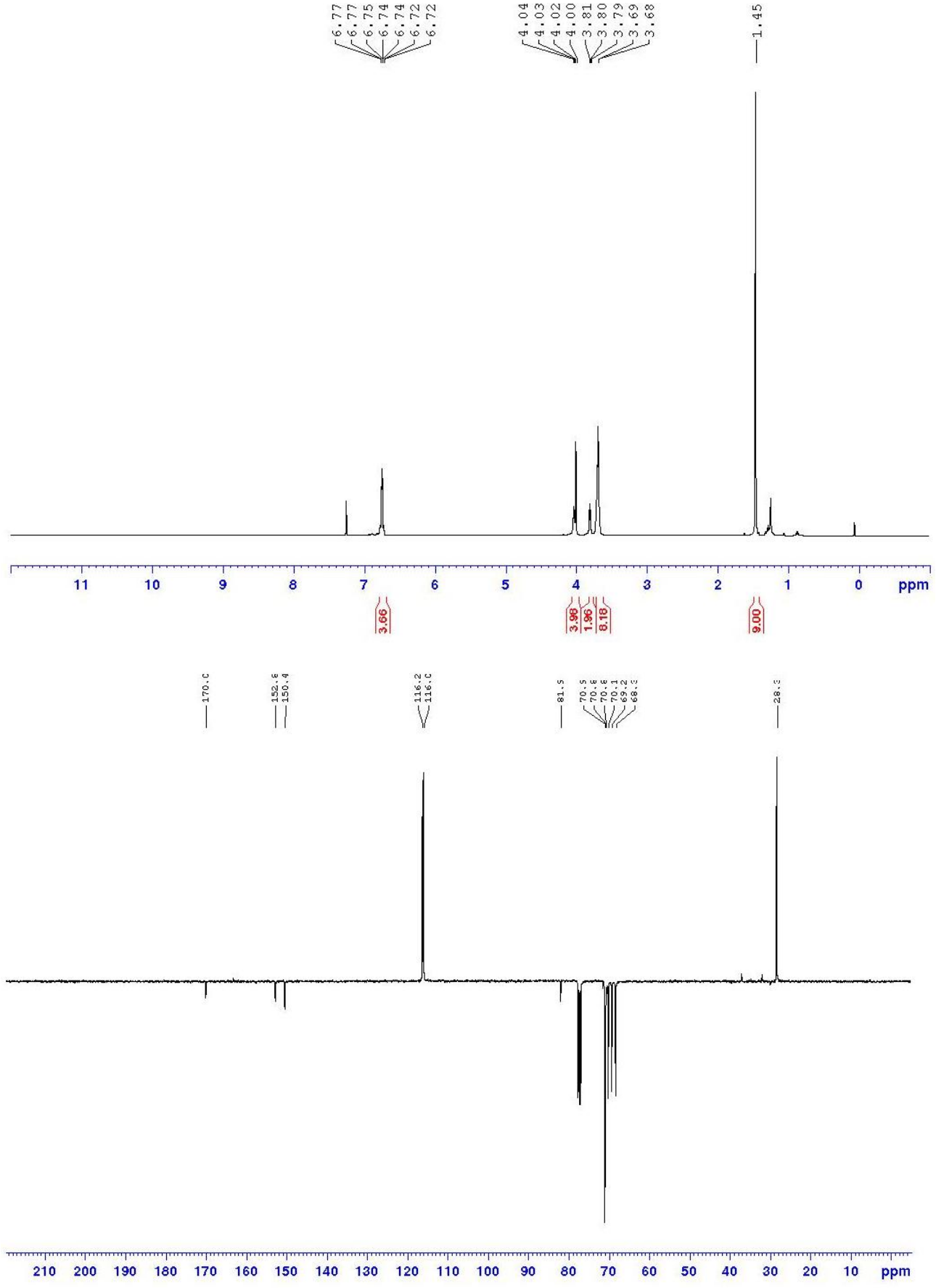


tert-Butyl 2-(2-(2-(2-tosyloxyethoxy)ethoxy)ethoxy)acetate (**74**)

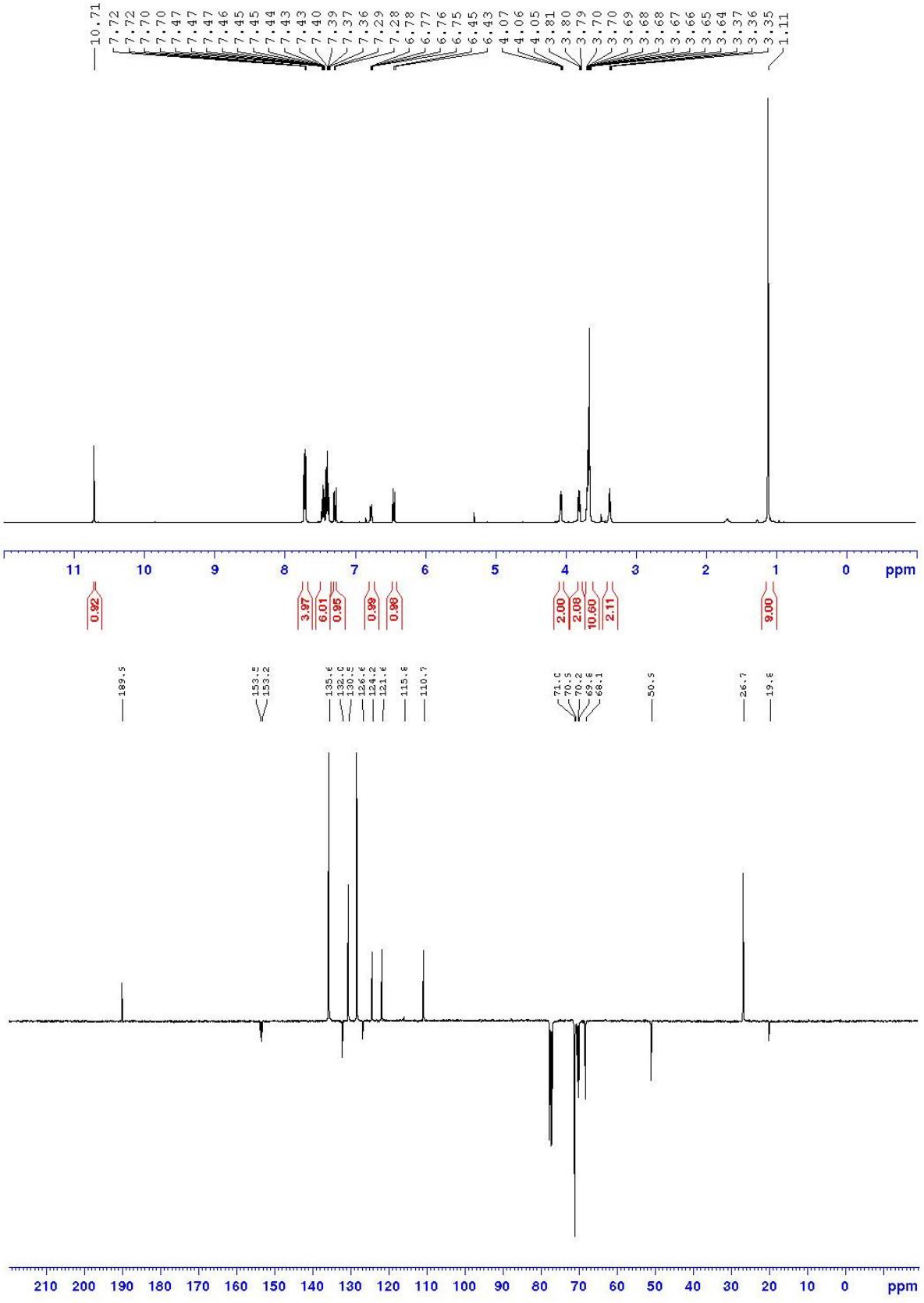




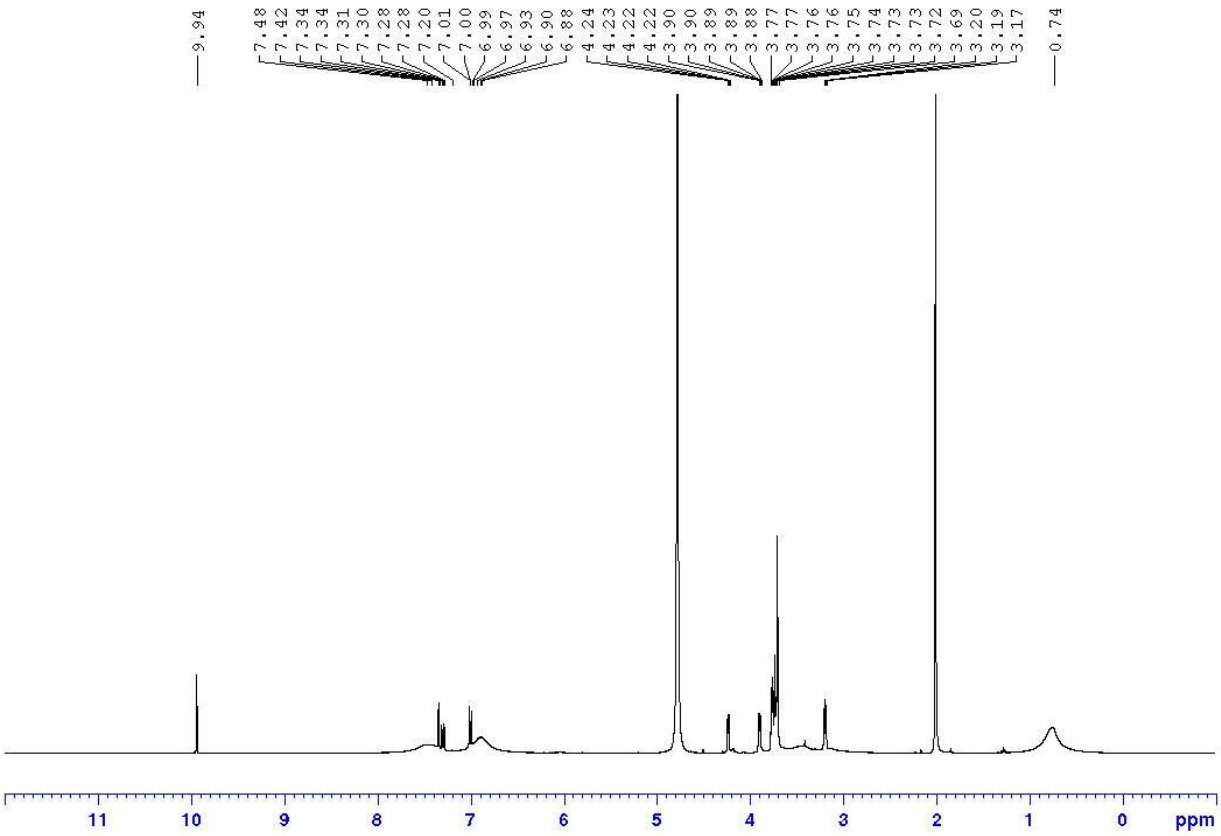
tert-butyl 2-(2-(2-(2-(4-hydroxyphenoxy)ethoxy)ethoxy)ethoxy)acetate (**75**)



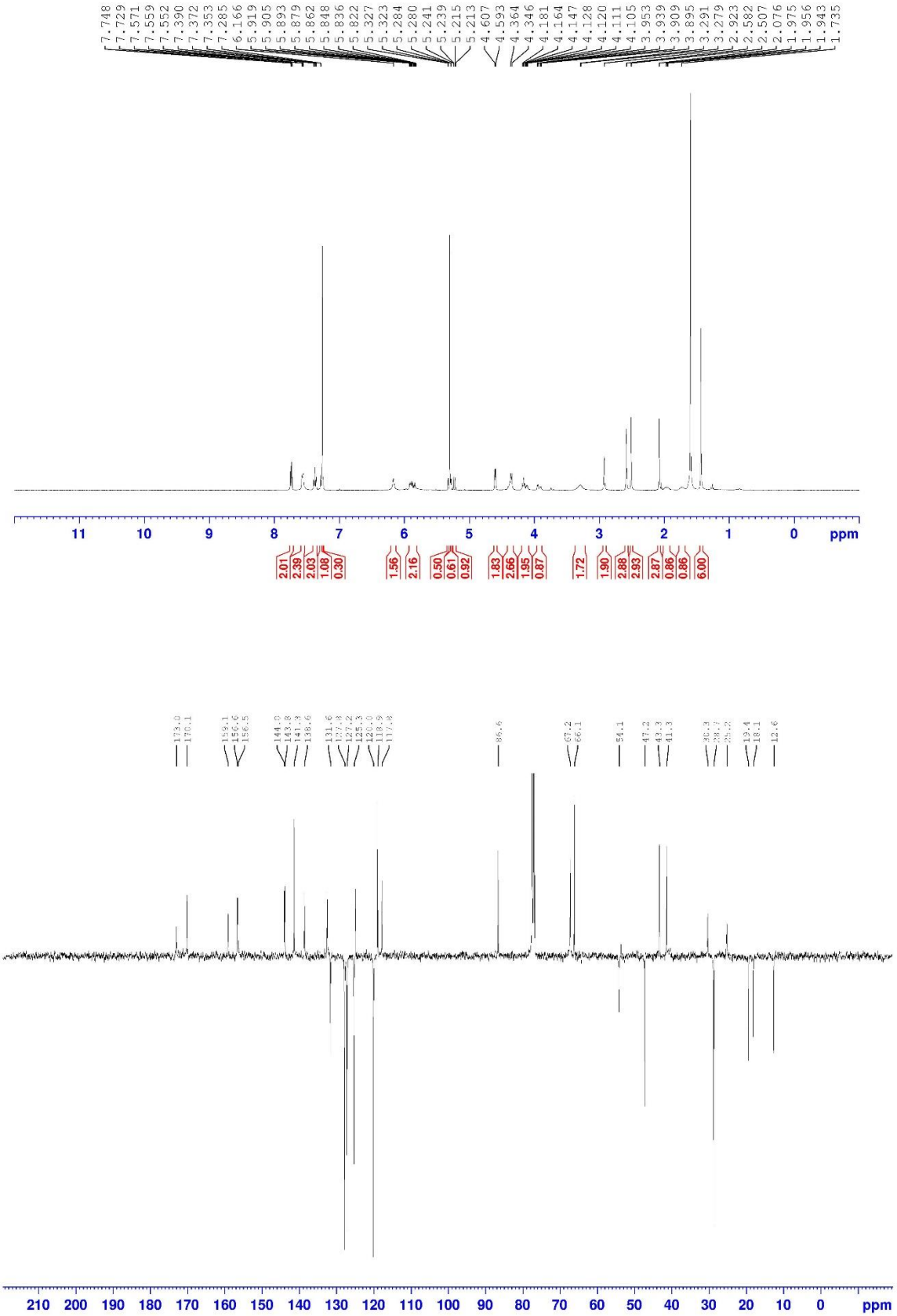
5-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethoxy)-2-((tert-butyldiphenylsilyl)oxy)benzaldehyde (**79**)



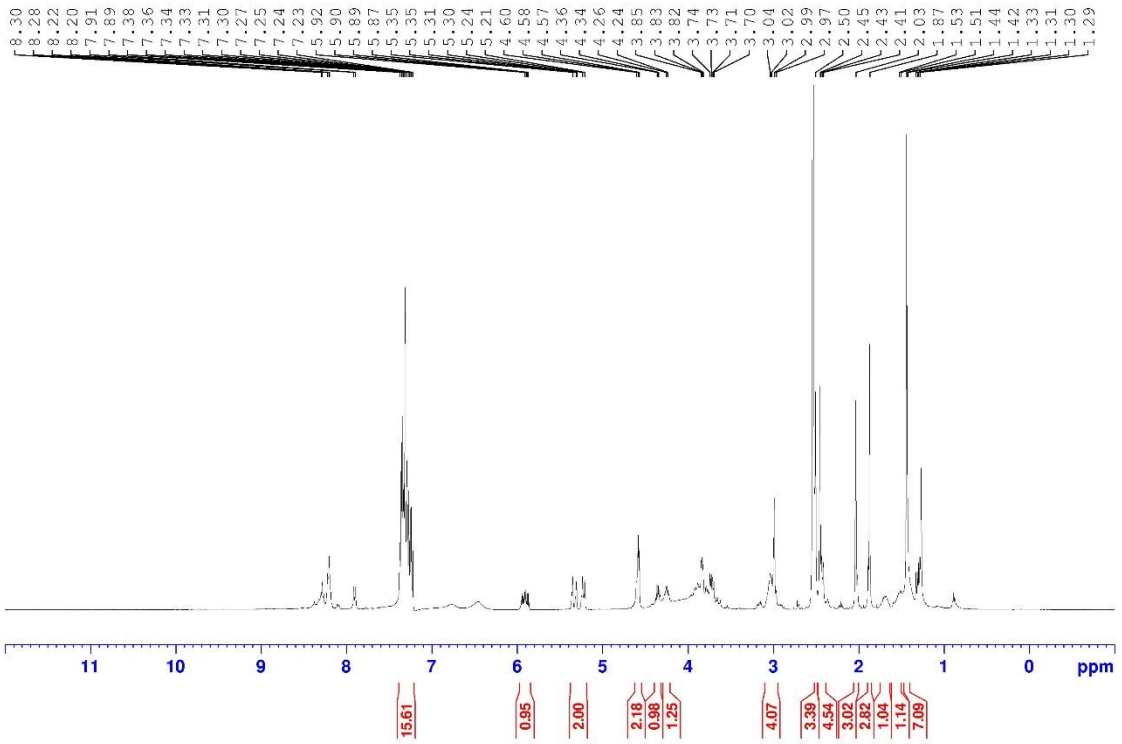
5-(2-(2-(2-(2-aminoethoxy)ethoxy)ethoxy)ethoxy)-2-((tert-butyldiphenylsilyl)oxy)benzaldehyde, acetate salt (**80**)

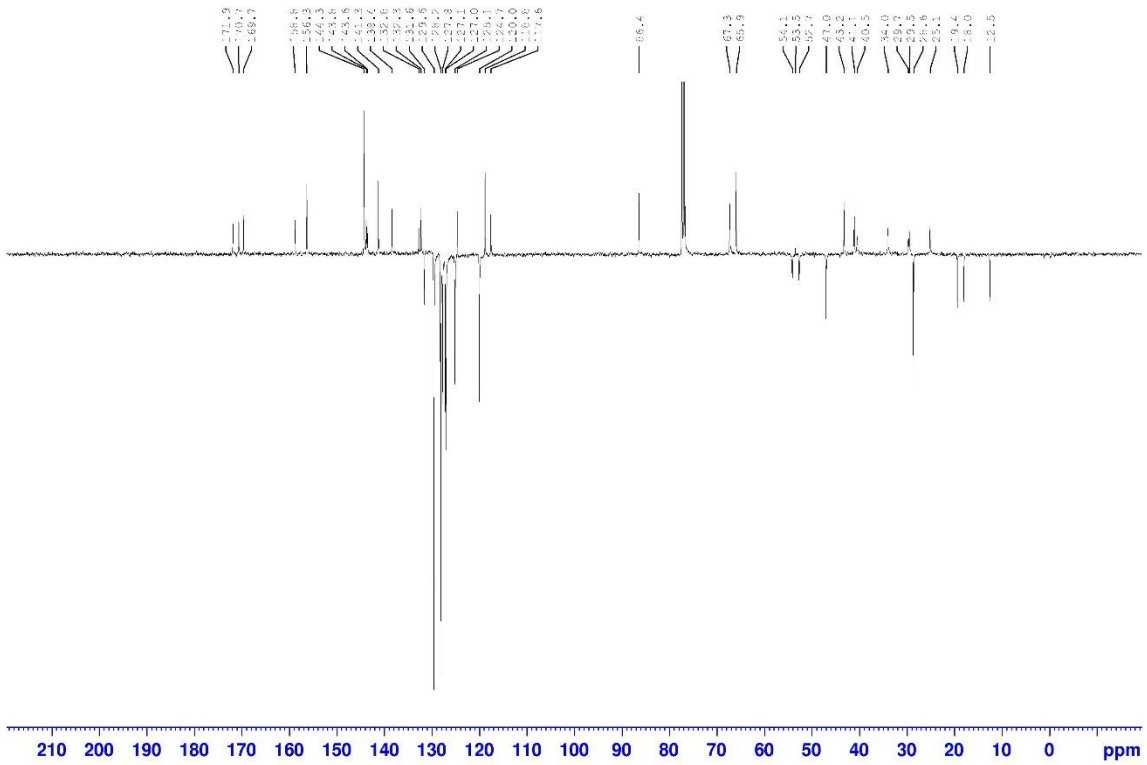


Fmoc-Arg(Pbf)-Gly-OAll (**96**)

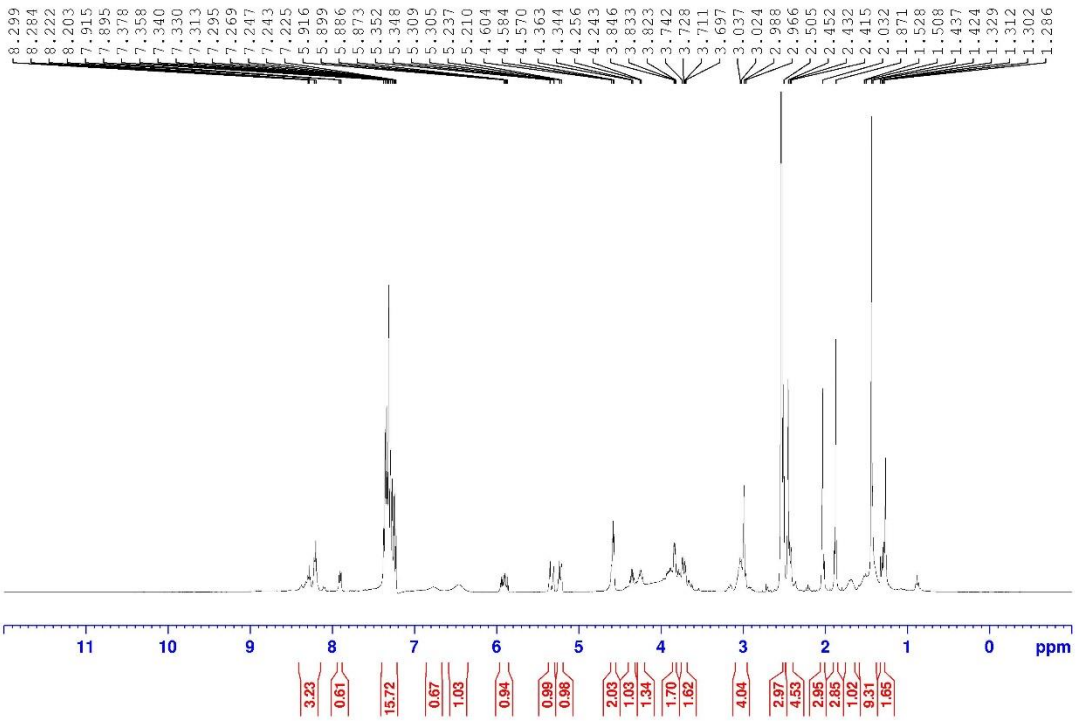


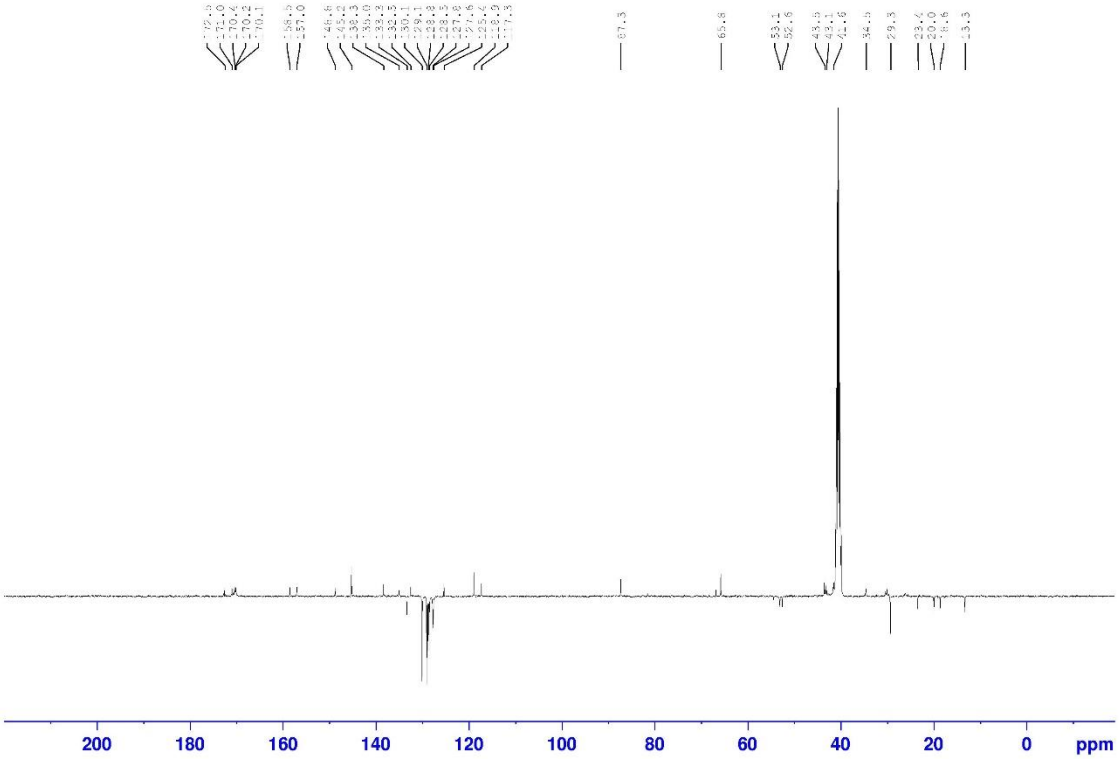
Fmoc-Cys(Trt)-Arg(Pbf)-Gly-OAll (**97**)



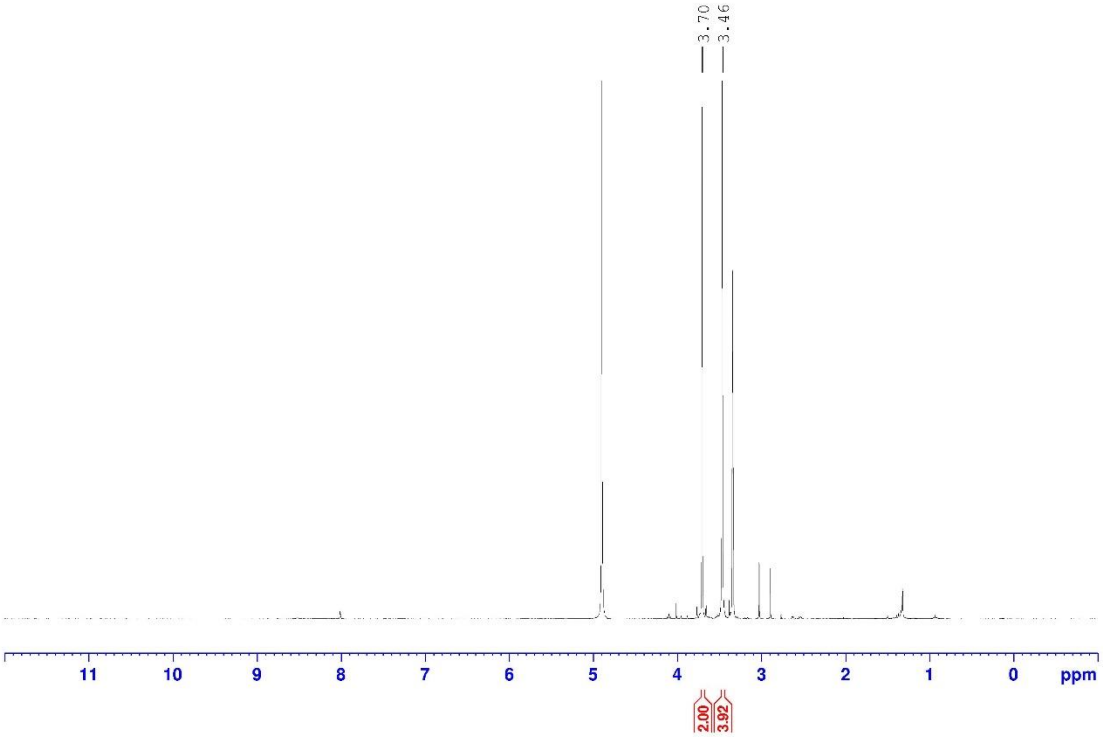


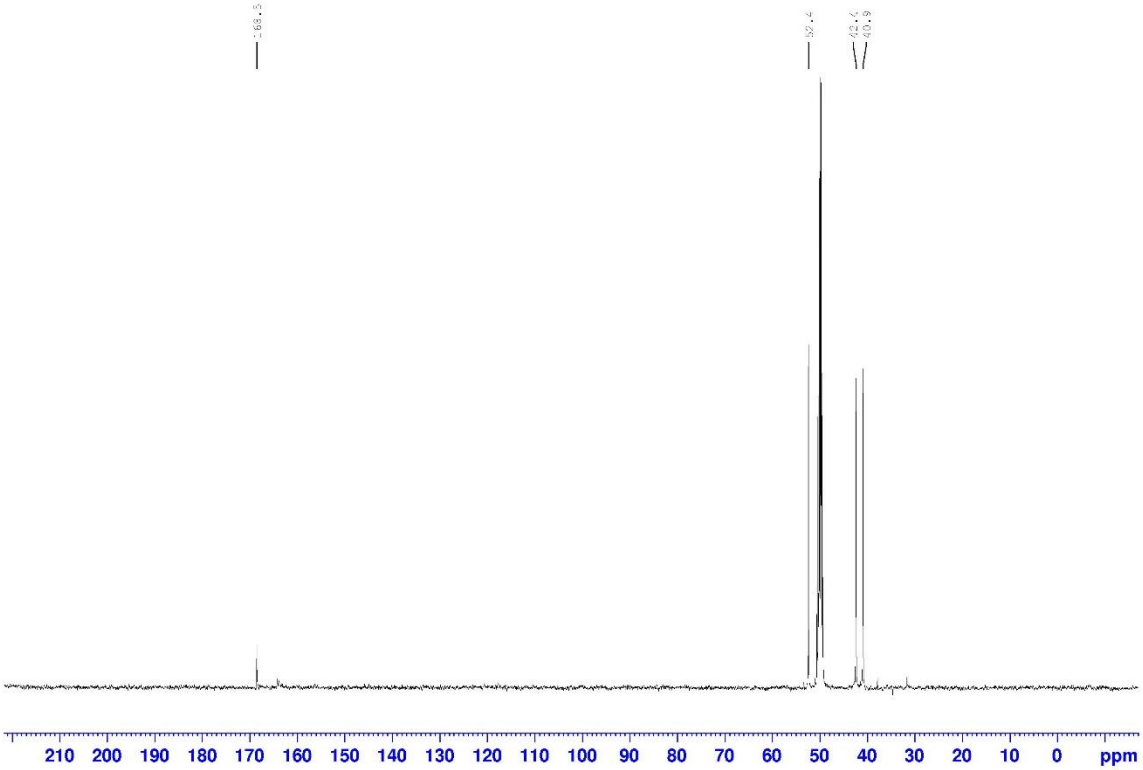
Ac-Gly-Cys(Trt)-Arg(Pbf)-Gly-OAll (**98**)



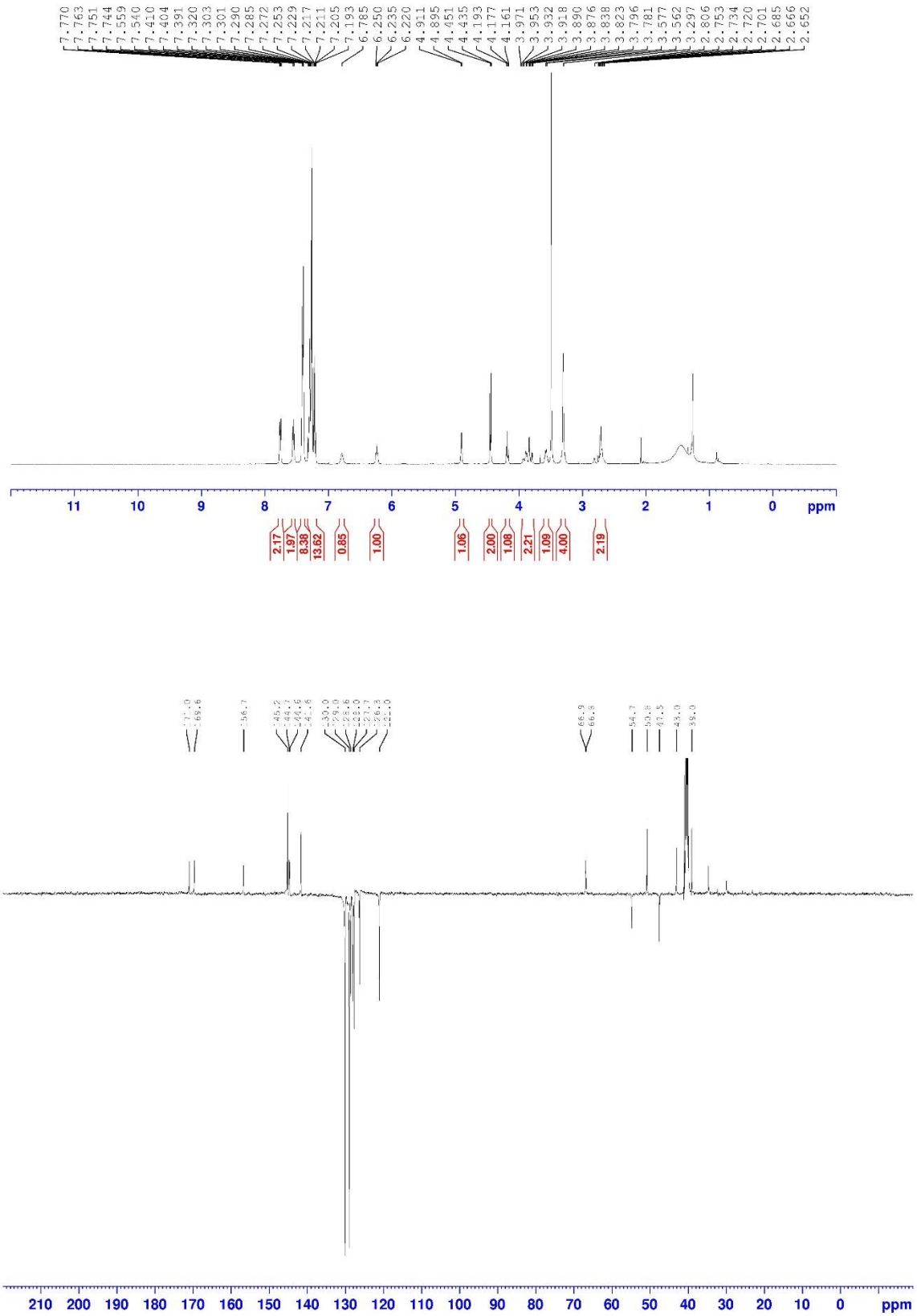


2-amino-N-(2-azidoethyl)-acetamide (**101**)

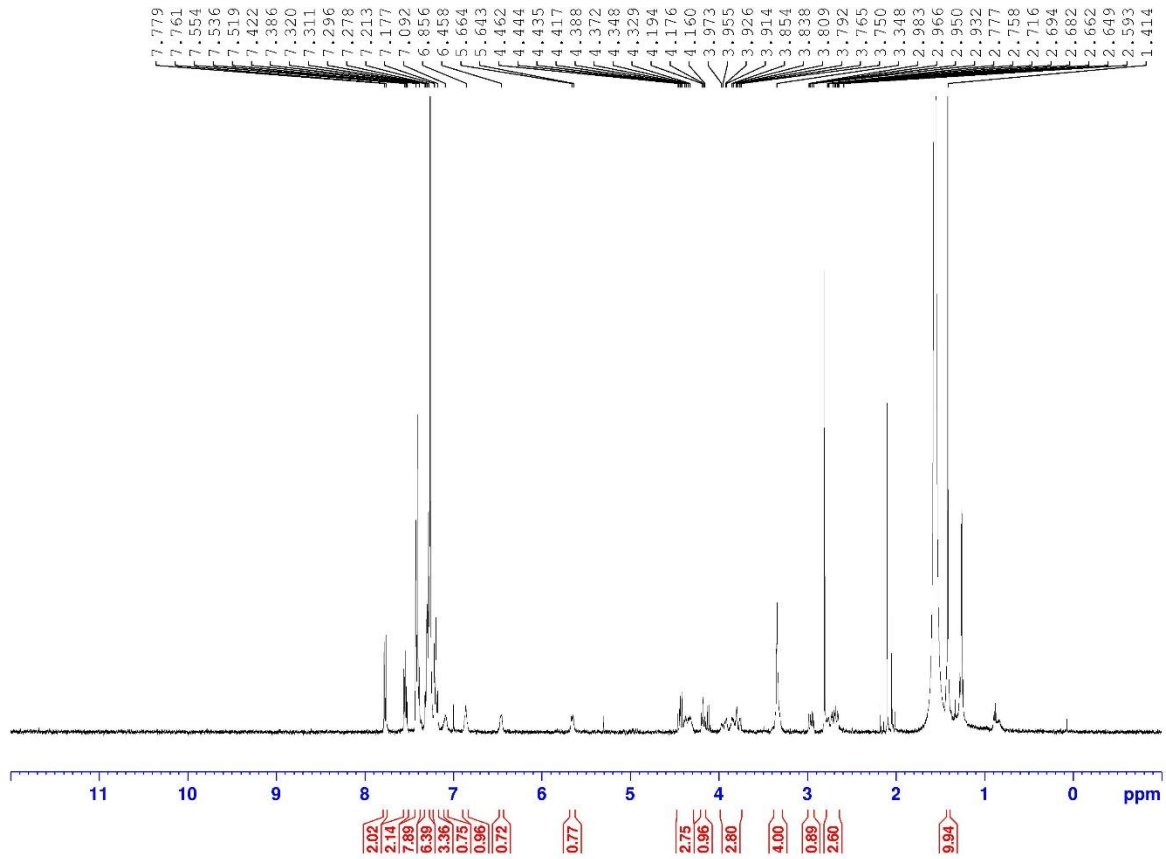


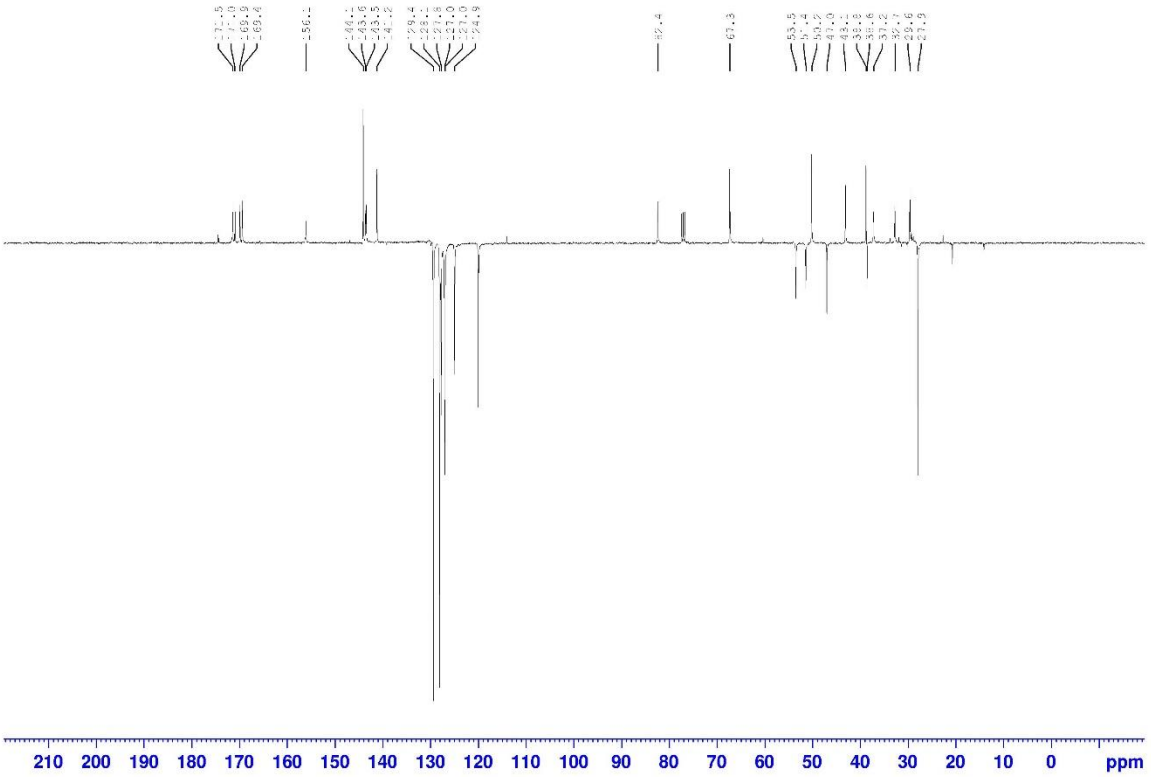


Fmoc-Cys(Trt)-Gly-N_3 (**102**)



Fmoc-Asp(OtBu)-Cys(Trt)-Gly-N_3 (**103**)





References

1 K.C. McCullough, A. Summerfield. Basic concepts of immune response and defense development. *ILAR J* **2005**, *46*, 230-240.

2 G. Köhler, C. Milstein. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **1975**, *256,* 495-497.

3 L. Ledsgaard, M. Kilstrup, A. Karatt-Vellatt, J. McCafferty, A. H. Laustsen. Basics of Antibody Phage Display Technology. *Toxins* **2018**, *10*, 236.

4 J.S. Bonifacino, D.C. Gershlick, E.C. Dell'Angelica. Immunoprecipitation. *Curr. Protoc. Cell Biol.* **2016,** *71*, 7.2.1-7.2.24.

5 S. Gallagher, S.E. Winston, S.A. Fuller, J.G.R. Hurrell. Immunoblotting and Immunodetection. *Curr. Protoc. Cell Biol.* **2011**, *52*, 6.2.1-6.2.28.

6 P. Hornbeck, Enzyme-linked immunosorbent assays. *Curr. Protoc. Immunol.* **2015**, *110*, 2.1.1-2.1.23.

7 a) G.A. Michaud, M. Salcius, F. Zhou, R. Bangham, J. Bonin, H. Guo, M. Salcius, P.F. Predki, B.I. Schweitzer. Analyzing specificity with whole proteome microarrays. *Nature Biotechnol*. **2003**, *21*, 1509-1512; b) U.B. Nielsen, B.H. Geierstanger. Multiplexed sandwich assays in microarray format. *J. Immun. Methods* **2004**, *290*, 107-120.

8 mAbs demonstrated to be particularly useful for the formation of co-crystals between ion channels and the relative mAb, as highlighted in the following examples: a) Y. Jiang, A. Lee, J. Chen, V. Ruta, M. Cadene, B.T. Chait, R. MacKinnon. X-ray structure of a voltage-dependent K 1 channel. *Nature* **2003**, *423*, 33-41; b) Y. Zhou, J.H. Morais-Cabral, A. Kaufman, R. MacKinnon. Chemistry of ion coordination and hydration revealed by a K+ channel-FAb complex at 2.0 Angstrom resolution. *Nature* **2001**, *414*, 43-48.

9 A.L. Givan. Flow Cytometry: First Principles. New York: Wiley-Liss, Inc.

10 a) D.J. Asai, Immunofluorescence microscopy. *Curr. Protoc. Essen. Lab. Tech.* **2015**, *10*, 9.2.1-9.2.23; b) F.M. Hofman, C.R. Taylor, Immunohistochemistry. *Curr. Protoc. Immunol.* **2013**, *103*, 21.4.1-21.4.26.

11 K. Midtvedt, P. Fauchald, B. Lien, A. Hartmann, D. Albrechtsen, B.L. Bjerkely, T. Leivestad, I.B. Brekke. Individualized T cell monitored administration of ATG versus OKT3 in steroid-resistant kidney graft rejection. *Clin. Transplant.* **2003**, *17*, 69-74.

12 A. Mullard. FDA approves 100th monoclonal antibody product. *Nat Rev Drug Discov.* **2021**, *20*, 491-495.

13 L. Urquhart. Top product forecasts for 2021. *Nat. Rev. Drug Disc.* **2021**, *20*, 10.

14 L.M. Nadler, P. Stashenko, R. Hardy, W.D. Kaplan, L.N. Button, D.W. Kufe, K.H. Antman, S.F. Schlossman. Serotherapy of a Patient with a Monoclonal Antibody Directed against a Human Lymphoma-associated Antigen. *Cancer Res* **1980**, *40*, 3147-3154.

15 J. Ritz, S.F. Schlossman. Utilization of monoclonal antibodies in the treatment of leukemia and lymphoma. *Blood* **1982**, *59*, 1-11.

16 L. Riechmann, M. Clark, H. Waldmann, G. Winter. Reshaping human antibodies for therapy. *Nature* **1988**, *332*, 323–327.

17 C.B. Xie, D. Jane-Wit, J.S. Pobe. Complement Membrane Attack Complex: New Roles, Mechanisms of Action, and Therapeutic Targets. *Am J Pathol.* **2020**, *190*, 1138-1150.

18 C.A. Janeway Jr, P. Travers, M. Walport, et al. Immunobiology: The Immune System in Health and Disease. 5th edition. New York: Garland Science; 2001. The complement system and innate immunity.

19 N. Di Gaetano, E. Cittera, R. Nota, A. Vecchi, V. Grieco, E. Scanziani, M. Botto, M. Introna, J. Golay. Complement Activation Determines the Therapeutic Activity of Rituximab in Vivo. *J. Immunol.* **2003**, *171*, 1581–1587.

20 B. Coiffer, S. Lepretre, L.M. Pedersen, O. Gadeberg, H. Fredriksen, M.H.J. Van Oers, J. Wooldridge, J. Kloczko, J. Holowiecki, A. Hellmann, J. Walewski, M. Flensburg, J. Petersen, T. Robak. Safety and Efficacy of Ofatumumab, a Fully Human Monoclonal Anti-CD20 Antibody, in Patients with Relapsed or Refractory B-Cell Chronic Lymphocytic Leukemia: A Phase 1-2 Study. *Blood* **2008**, *111*, 1094–1100.

21 N. Gül. L. Babes, K. Siegmund, R. Korthouwer, M. Bögels, R. Braster, G. Vidarsson, T.L.M. Ten Hagen, P. Kubes, M. Van Egmond. Macrophages Eliminate Circulating Tumor Cells after Monoclonal Antibody Therapy. *J. Clin. Investig.* **2014**, *124*, 812–823.

22 E. Möller. Contact-Induced Cytotoxicity by Lymphoid Cells Containing Foreign Isoantigens. *Science* **1965**,
147, 873–879.

23 P.M. Sondel, K.L. Alderson. Clinical Cancer Therapy by NK Cells via Antibody-Dependent Cell-Mediated Cytotoxicity. *J. Biomed. Biotechnol.* **2011**, *2011*, 379123.

24 F. Nimmerjahn, J.V. Ravetch. Fc Receptors as Regulators of Immune Responses. *Nat. Rev. Immunol.* **2008**, *8*, 34–47.
25 a) G. De Saint Basile, G. Ménasché, A. Fischer. Molecular Mechanisms of Biogenesis and Exocytosis of Cytotoxic Granules. *Nat. Rev. Immunol.* **2010**, *10*, 568–579; b) F. Nimmerjahn, J.V. Ravetch. Analyzing Antibody-Fc-Receptor Interactions. *Methods Mol. Biol.* **2008**, *415*, 151–162; c) C.M. Eischen, P.J. Leibson. Role for NK-Cell-Associated Fas Ligand in Cell-Mediated Cytotoxicity and Apoptosis. *Res. Immunol.* **1997**, *148*, 164–169.
26 a) Z. Liu, K. Gunasekaran, W. Wang, V. Razinkov, L. Sekirov, E. Leng, H. Sweet, I. Foltz, M. Howard, A.M. Rousseau, C. Kozlosky, W. Fanslow, W. Yan. Asymmetrical Fc Engineering Greatly Enhances Antibody-dependent Cellular Cytotoxicity (ADCC) Effector Function and Stability of the Modified Antibodies. *J. Biol. Chem.* **2014**, *289*, 3571–3590; b) P. Umaña, J. Jean-Mairet, R. Moudry, H. Amstutz, J.E. Bailey. Engineered Glycoforms of an Antineuroblastoma IgG1 with Optimized Antibody-Dependent Cellular Cytotoxic Activity. *Nat. Biotechnol.* **1999**, *17*, 176–180; c) J. Davies, L. Jiang, L.Z. Pan, M.J. Labarre, D. Anderson, M. Re. Expression of GnTIII in a Recombinant Anti-CD20 CHO Production Cell Line: Expression of Antibodies with Altered Glycoforms Leads to an Increase in ADCC through Higher Affinity for FcRIII. *Biotechnol. Bioeng.* **2001**, *74*, 288–294; d) R.L. Shields, J. Lai, R. Keck, L.Y. O'Connell, K. Hong, Y. Gloria Meng, S.H.A. Weikert, L.G. Presta. Lack of Fucose on Human IgG1 N-Linked Oligosaccharide Improves Binding to Human FcRIII and Antibody-Dependent Cellular Toxicity. *J. Biol. Chem.* **2002**, *277*, 26733–26740.
27 T. Ishida, T. Joh, N. Uike, K. Yamamoto, A. Utsunomiya, S. Yoshida, Y. Saburi, T. Miyamoto, S. Takemoto, H. Suzushima, K. Tsukasaki, K. Nosaka, H. Fujiwara, K. Ishitsuka, H. Inagaki, M. Ogura, S. Akinaga, M. Tomonaga, K. Tobinai, R. Ueda. Defucosylated Anti-CCR4 Monoclonal Antibody (KW-0761) for Relapsed Adult T-Cell Leukemia-Lymphoma: A Multicenter Phase II Study. *J. Clin. Oncol.* **2012**, *30*, 837–842.
28 The Antibody Society. Therapeutic monoclonal antibodies approved or in review in the EU or US. (date accessed); www.antibodysociety.org/antibody-therapeutics-product-data (checked on 10th november 2021).
29 a) P. McLaughlin, A.J. Grillo-López, B.K. Link, R. Levy, M.S. Czuczman, M.E. Williams, M.R. Heyman, I. Bence-Bruckler, C.A. White, F. Cabanillas, V. Jain, A.D. Ho, J. Lister, K. Wey, D. Shen, B.K. Dallaire. Rituximab Chimeric Anti-CD20 Monoclonal Antibody Therapy for Relapsed Indolent Lymphoma: Half of Patients Respond to a Four-Dose Treatment Program. *J. Clin. Oncol.* **1998**, *16*, 2825–2833; b) S. Benavente, S. Huang, E.A. Armstrong, A. Chi, K.T. Hsu, D.L. Wheeler, P.M. Harari. Establishment and Characterization of a Model of Acquired Resistance to Epidermal Growth Factor Receptor Targeting Agents in Human Cancer Cells. *Clin. Cancer Res.* **2009**, *15*, 1585–1592; c) A. Ahmad. Current Updates on Trastuzumab Resistance in HER2 Overexpressing Breast Cancers. In *Advances in Experimental Medicine and Biology*; Springer: Cham, Switzerland, 2019; Volume 1152, pp. 217–228.
30 M.S. Czuczman, S. Olejniczak, A. Gowda, A. Kotowski, A. Binder, H. Kaur, J. Knight, P. Starostik, J. Deans, F.J. Hernandez-Ilizaliturri. Acquirement of Rituximab Resistance in Lymphoma Cell Lines Is Associated with Both Global CD20 Gene and Protein Down-Regulation Regulated at the Pretranscriptional and Posttranscriptional Levels. *Clin. Cancer Res.* **2008**, *14*, 1561–1570.
31 V. Sforza, E. Martinelli, F. Ciardiello, V. Gambardella, S.Napolitano, G. Martini, C.D. Corte, C. Cardone, M.L. Ferrara, A. Reginelli, G. Liguori, G. Belli, T. Troiani. Mechanisms of Resistance to Anti-Epidermal Growth Factor Receptor Inhibitors in Metastatic Colorectal Cancer. *World J. Gastroenterol.* **2016**, *22*, 6345–6361.
32 N.Krall, J. Scheuermann, D. Neri. Small targeted cytotoxics: current state and promises from DNA-encoded chemical libraries. *Angew. Chem. Int. Ed.* **2013**, *52*, 1384−402.
33 M. S. Dennis, H. K. Jin, D. Dugger, R. H. Yang, L. McFarland, A. Ogasawara, S. Williams, M. J. Cole, S. Ross, R. Schwall, Imaging Tumors with an Albumin-Binding Fab, a Novel Tumor-Targeting Agent. *Cancer Res.* **2007**, *67*, 254 – 261.
34 a) T. Saga, R. D. Neumann, T. Heya, J. Sato, S. Kinuya, N. Le, C. H. Paik, J. N. Weinstein. Targeting cancer micrometastases with monoclonal antibodies: a binding-site barrier. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8999 – 9003; b) G. P. Adams, R. Schier, A. M. McCall, H. H. Simmons, E. M. Horak, R. K. Alpaugh, J. D. Marks, L. M. Weiner. High affinity restricts the localization and tumor penetration of single-chain fv antibody molecules. *Cancer Res.* **2001**, *61*, 4750 – 4755; c) S. I. Rudnick, J. Lou, C. C. Shaller, Y. Tang, A. J. Klein-Szanto, L. M. Weiner, J. D. Marks, G. P. Adams. Influence of Affinity and Antigen Internalization on the Uptake and Penetration of Anti-HER2 Antibodies in Solid Tumors. *Cancer Res.* **2011**, *71*, 2250 – 2259.
35 L. Borsi, E. Balza, M. Bestagno, P. Castellani, B. Carnemolla, A. Biro, A. Leprini, J. Sepulveda,O. Burrone, D. Neri, L. Zardi. Selective targeting of tumoral vasculature: Comparison of different formats of an antibody (L19) to the ED-B domain of fibronectin. *Int. J. Cancer* **2002**, *102*, 75 – 85.
36 T. Olafsen, V. E. Kenanova, G. Sundaresan, A. L. Anderson, D. Crow, P. J. Yazaki, L. Li, M. F. Press, S. S. Gambhir, L. E. Williams, J. Y. Wong, A. A. Raubitschek, J. E. Shively, A. M. Wu. Optimizing

Radiolabeled Engineered Anti-p185HER2 Antibody Fragments for In vivo Imaging. *Cancer Res.* **2005**, *65*, 5907 – 5916.
37 S. Cazzamalli, A. Dal Corso, F. Widmayer, D. Neri. Chemically Defined Antibody− and Small Molecule−Drug Conjugates for in Vivo Tumor Targeting Applications: A Comparative Analysis *J. Am. Chem. Soc.* **2018**, *140*, 1617−1621.
38 R. Macarron, M. N. Banks, D. Bojanic, D. J. Burns, D. A. Cirovic, T. Garyantes, D. V. S. Green, R. P. Hertzberg, W. P. Janzen, J. W. Paslay, U. Schopfer, G.S. Sittampalam. Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.***2011** *10*, 188–195.
39 L. M. Mayr, D. Bojanic. Novel trends in high-throughput screening. *Curr Opin Pharmacol.* **2009**, *9*, 580-588.
40 S. Pathmanathan, I. Grozavu, A. Lyakisheva, I. Stagljar. Drugging the undruggable proteins in cancer: A systems biology approach. *Curr. Op. Chem. Biol.* **2021**, in press. https://doi.org/10.1016/j.cbpa.2021.07.004.
41 S. L. McGovern, E. Caselli, N. Grigorieff, B.K. Shoichet. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* **2002**, *45*, 1712–1722.
42 J. Jesús Naveja J. L. Medina-Franco. Finding Constellations in Chemical Space Through Core Analysis. *Front. Chem.* **2019**, *7*, 510.
43 J. Owens. Determining druggability. *Nat. Rev. Drug Discov.* **2007**, *6*, 187–187.
44 a) M. R. Arkin, Y. Tang, J. A. Wells. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem. Biol.* **2014**, *21*, 1102–1114; b) C. V. Dang, E. P. Reddy, K. M. Shokat, L. Soucek. Drugging the 'undruggable' cancer targets. *Nat. Rev. Cancer* **2017**, *17*, 502–508.
45 a) C. W. Murray, D. C. Rees. The rise of fragment-based drug discovery. *Nat. Chem.* **2009**, *1*, 187–192; b) B. C. Doak, R.S. Norton, M. J. Scanlon. The ways and means of fragment-based drug design. *Pharmacol. Therapeut.* **2016**, *167*, 28–37.
46 D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti. Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **2016**, *15*, 605–619.
47 G. Bollag, P. Hirth, J. Tsai, J. Zhang, P. N. Ibrahim, H. Cho, W. Spevak, C. Zhang, Y. Zhang, G. Habets, E. A. Burton, B. Wong, G. Tsang, B. L. West, B. Powell, R. Shellooe, A. Marimuthu, H. Nguyen, K. Y. J. Zhang, D. R. Artis, J. Schlessinger, F. Su, B. Higgins, R. Iyer, K. D'Andrea, A. Koehler, M. Stumm, P. S. Lin, R. J. Lee, J. Grippo, I. Puzanov, K. B. Kim, A. Ribas, G. A. McArthur, J. A. Sosman, P. B. Chapman, K. T. Flaherty, X. Xu, K. L Nathanson, K. Nolop. Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* **2010**, *467*, 596-599.
48 Y. Si, D. Xu, K. Bum-Erdene, M. K. Ghozayel, B. Yang, P. A. Clemons, S. O. Meroueh. Chemical space overlap with critical protein–protein interface residues in commercial and specialized small-molecule libraries. *ChemMedChem* **2019**, *14*, 119–131.
49 A. P. Turnbull, S. M. Boyd, B. Walse. Fragment-based drug discovery and protein-protein interactions. *Res. Rep. Biochem.* **2014**, *4*, 13–26
50 L. Laraia, G. McKenzie, D. R. Spring, A. R. Venkitaraman, D. J. Huggins. Overcoming chemical, biological, and computational challenges in the development of inhibitors targeting protein-protein interactions. *Chem. Biol.* **2015**, *22*, 689–703.
51 S. A. Andrei, E. Sijbesma, M. Hann, J. Davis, G. O'Mahony, M. W. D. Perry, A. Karawajczyk, J. Eickhoff, L. Brunsveld, R. G. Doveston, L-G. Milroy, C. Ottmann. Stabilization of protein-protein interactions in drug discovery. *Expert Opin. Drug Discovery* **2017**, *12*, 925–940.
52 a) M. Mondal, N. Radeva, H. Fanlo-Virgós, S. Otto, G. Klebe, A. K. H. Hirsch. Fragment Linking and optimization of inhibitors of the aspartic protease endothiapepsin: fragment-based drug design facilitated by dynamic combinatorial chemistry. *Angew. Chem. Int. Ed.* **2016**, *55*, 9422–9426; b) B. Lamoree, R. E. Hubbard. Current perspectives in fragment-based lead discovery (FBLD). *Essays Biochem.* **2017**, *61*, 453–464.
53 a) J. Tsai, J. T. Lee, W. Wang, J. Zhang, H. Cho, S. Mamo, R. Bremer, S. Gillette, J. Kong, N. K. Haass, K. Sproesser, L. Li, K. S. M. Smalley, D. Fong, Y. L. Zhu, A. Marimuthu, H. Nguyen, B. Lam, J. Liu, I. Cheung, J. Rice, Y. Suzuki, C. Luu, C. Settachatgul, R. Shellooe, J. Cantwell, S. H. Kim, J. Schlessinger, K. Y. J. Zhang, B. L. West, B. Powell, G. Habets, C. Zhang, P. N. Ibrahim, P. Hirth, D. R. Artis, M. Herlyn, G. Bollag. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 3041–3046; b) A. J. Souers, J. D. Leverson, E. R. Boghaert, S. L. Ackler, N. D. Catron, J. Chen, B. D. Dayton, H. Ding, S. H. Enschede, W. J. Fairbrother, D. C. S. Huang, S. G. Hymowitz, S. Jin, S. L. Khaw, P. J. Kovar, L. T. Lam, J. Lee, H. L. Maecker, K. C. Marsh, K. D. Mason, M. J. Mitten, P. M. Nimmer, A. Oleksijew, C. H. Park, C. M. Park, D. C. Phillips, A. W. Roberts, D. Sampath, J. F. Seymour, M. L. Smith, G. M. Sullivan, S. K. Tahir, C. Tse, M. D. Wendt, Y. Xiao, J. C. Xue, H. Zhang, R. A. Humerickhouse, S. H. Rosenberg, S. W. Elmore. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat. Med.* **2013**, *19*, 202–208; c) C. Zhang, P. N. Ibrahim, J. Zhang, E. A. Burton, G. Habets, Y. Zhang, B. Powell, B. L. West, B. Matusow, G. Tsang, R. Shellooe, H. Carias, H. Nguyen, A. Marimuthu, K. Y. J. Zhang, A. Oh, R.

Bremer, C. R. Hurt, D. R. Artis, G. Wu, M. Nespi, W. Spevak, P. Lin, K. Nolop, P. Hirth, G. H. Tesch, G. Bollag. Design and pharmacology of a highly specific dual FMS and KIT kinase inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 5689–5694; d) C. W. Murray, D. C. Rees. The rise of fragment-based drug discovery. *Nat. Chem.* **2009**, *1*, 187–192.

54 K. J. Temple, J. L. Engers, M. F. Long, A. R. Gregro, K. J. Watson, S. Chang, M. T. Jenkins, V. B. Luscombe, A. L. Rodriguez, C. M. Niswender, T. M. Bridges, P. J. Conn, D. W. Engers, C. W. Lindsley. Discovery of a novel 3,4-dimethylcinnoline carboxamide M4 positive allosteric modulator (PAM) chemotype via scaffold hopping. *Bioorg. Med. Chem. Lett.* **2019**, *29*, 126678.

55 M. J. Harner, A. O. Frank, S. W. Fesik. Fragment-based drug discovery using NMR spectroscopy. *J. Biomol. NMR* **2013**, *56*, 65–75.

56 M. Mondal, N. Radeva, H. Fanlo-Virgós, S. Otto, G. Klebe, A. K. H. Hirsch. Fragment Linking and optimization of inhibitors of the aspartic protease endothiapepsin: fragment-based drug design facilitated by dynamic combinatorial chemistry. *Angew. Chem. Int. Ed.* **2016**, *55*, 9422–9426.

57 W. Tao, W. Mian-Bin, C. Zheng-Jie, C. Hua, L. Jian-Ping, Y. Li-Rong. Fragment-based drug discovery and molecular docking in drug design. *Curr. Pharmaceut. Biotechnol.* **2015**, *16*, 11–25.

58 A. I. Chan, L. M. McGregor, D. R. Liu. Novel selection methods for DNA-encoded chemical libraries. *Curr. Opin. Chem. Biol.* **2015**, *26*, 55−61.

59 S. Brenner, R. A. Lerner. Encoded Combinatorial Chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 5381−5383.

60 N. Favalli, G. Bassi, J. Scheuermann, D. Neri. DNA-encoded chemical libraries − achievements and remaining challenges. *FEBS Lett.* **2018**, *592*, 2168−2180.

61 L. Mannocci, Y. Zhang, J. Scheuermann, M. Leimbacher, G. De Bellis, E. Rizzi, C. Dumelin, S. Melkko, D. Neri. High-throughput sequencing allows the identification of binding molecules isolated from DNA-encoded chemical libraries. *PNAS* **2008**, *105*, 17670–17675.

62 D. Neri, R. A. Lerner. DNA-Encoded Chemical Libraries: A Selection System Based On Endowing Organic Compounds with Amplifiable Information. *Annu. Rev. Biochem*. **2018**, *87*, 479-502.

63 D. Madsen, C. Azevedo, I. Micco, L. K. Petersen, N. J. V. Hansen. An overview of DNA-encoded libraries: A versatile tool for drug discovery. *Prog. Med. Chem.* **2020**, *59*, 181−249.

64 a) P. A. Harris, S. B. Berger, J. U. Jeong, R. Nagilla, D. Bandyopadhyay, N. Campobasso, C. A. Capriotti, J. A. Cox, L. Dare, X. Dong, P. M. Eidam, J. N. Finger, S. J. Hoffman, J. Kang, V. Kasparcova, B. W. King, R. Lehr, Y. Lan, L. K. Leister, J. D. Lich, T. T. MacDonald, N. A. Miller, M. T. Ouellette, C. S. Pao, A. Rahman, M. A. Reilly, A. R. Rendina, E. J. Rivera, M. C. Schaeffer, C. A. Sehon, R. R. Singhaus, H. H. Sun, B. A. Swift, R. D. Totoritis, A. Vossenkamper, P. Ward, D. D. Wisnoski, D. Zhang, R. W. Marquis, P. J. Gough, J. Bertin. Discovery of a First-in-Class Receptor Interacting Protein 1 (RIP1) Kinase Specific Clinical Candidate (GSK2982772) for the Treatment of Inflammatory Diseases. *J. Med. Chem.* **2017**, *60*, 1247−1261; b) S. L. Belyanskaya, Y. Ding, J. F. Callahan, A. L. Lazaar, D. I. Israel. Discovering Drugs with DNA-Encoded Library Technology: From Concept to Clinic with an Inhibitor of Soluble Epoxide Hydrolase. *ChemBioChem* **2017**, *18*, 837−842; c) J. W. Cuozzo, M. A. Clark, A. D. Keefe, A. Kohlmann, M. Mulvihill, H. Ni, L. M. Renzetti, D. I. Resnicow, F. Ruebsam, E. A. Sigel, H. A. Thomson, C. Wang, Z. Xie, Y. Zhang. Novel Autotaxin Inhibitor for the Treatment of Idiopathic Pulmonary Fibrosis: A Clinical Candidate Discovered Using DNA-Encoded Chemistry. *J. Med. Chem.* **2020**, *63*, 7840–7856.

65 H. Zehender, F. Le Goff, N. Lehmann, I. Filipuzzi, L. M. Mayr. SpeedScreen: the “missing link” between genomics and lead discovery. *J. Biomol. Screen.* **2004**, *9*, 498–505.

66 A. Annis, C.-C. Chuang, N. Nazef in Mass Spectrometry in Medicinal Chemistry Ch. 3, edited by Wanner, K. T. & Höfner, G., Wiley, **2007**.

67 E. C. VanderPorten, M. D. Scholle, J. Sherrill, J. C. Tran, Y. Liu. Identification of small- molecule noncovalent binders utilizing SAMDI technology. *SLAS Discov.* **2017**, *22*, 1211–1217.

68 R. Prudent, D. A. Annis, P. J. Dandliker, J.- Y. Ortholand, D. Roche. Exploring new targets and chemical space with affinity selection- mass spectrometry. *Nat. Rev. Chem.* **2021**, *5*, 62-71.

69 E. Valeur, S. M. Guéret, H. Adihou, R. Gopalakrishnan, M. Lemurell, H. Waldmann, T. N. Grossmann, A. T. Plowright. New modalities for challenging targets in drug discovery. *Angew. Chem. Int. Ed.* **2017**, *56*, 10294–10323.

70 K. M Comess, J. D. Trumbull, C. Park, Z. Chen, R. A. Judge, M. J. Voorbach, M. Coen, L. Gao, H. Tang, P. Kovar, X. Cheng, M. E. Schurdak, H. Zhang, T. Sowin, D. J. Burns. Kinase drug discovery by affinity selection/mass spectrometry (ASMS): application to DNA damage checkpoint kinase Chk1. *J. Biomol. Screen.* **2006**, *11*, 755–764.

71 T. Siu, M. D. Altman, G. A. Baltus, M. Childers, J. M. Ellis, H. Gunaydin, H. Hatch, T. Ho, J. Jewell, B. M. Lacey, C. A. Lesburg, B.-S. Pan, B. Sauvagnat, G. K. Schroeder, S. Xu. Discovery of a novel cGAMP competitive ligand of the inactive form of STING. *ACS Med. Chem. Lett.* **2019**, *10*, 92–97.

72 C. E. Whitehurst, Z. Yao, D. Murphy, M. Zhang, S. Taremi, L. Wojcik, J. M. Strizki, J. D. Bracken, C. C. Cheng, X. Yang, G. W. Shipps Jr, M. Ziebell, E. Nickbarg. Application of affinity selection-mass

spectrometry assays to purification and affinitybased screening of the chemokine receptor CXCR4. *Comb. Chem. High Throughput Screen.* **2012**, *15*, 473–485.
73 N. F. Rizvi, J. A. Howe, A. Nahvi, D. J. Klein, T. O. Fischmann, H.-Y. Kim, M. A. McCoy, S. S. Walker, A. Hruza, M. P. Richards, C. Chamberlin, P. Saradjian, M. T. Butko, G. Mercado, J. Burchard, C. Strickland, P. J. Dandliker, G. F. Smith, E. B. Nickbarg. Discovery of selective RNA- binding small molecules by affinity- selection mass spectrometry. *ACS Chem. Biol.* **2018**, *13*, 820–831.
74 D. A. Flusberg, N. F. Rizvi, V. Kutilek, C. Andrews, P. Saradjian, C. Chamberlin, P. Curran, B. Swalm, S. Kattar, G. F. Smith, P. Dandliker, E. B. Nickbarg, J. O'Neil. Identification of G- quadruplexbinding inhibitors of Myc expression through affinity selection–mass spectrometry. *SLAS Discov.* **2019**, *24*, 142–157.
75 M. Shimaoka, T.A. Springer. Therapeutic antagonists and conformational regulation of integrin function. *Nat. Rev. Drug Discov.* **2003**, *2*, 703-716.
76 R. O. Hynes. Integrins: bidirectional, allosteric signaling machines. *Cell* **2002**, *110*, 673-687.
77 R. C. Liddington. Structural aspects of integrins. *Adv. Exp. Med. Biol.* **2014**, *819*, 111–126.
78 Z. Sun, M. Costell, R. Fassler. Integrin activation by talin, kindlin and mechanical forces. *Nat. Cell Biol.* **2019**, *21*, 25–31.
79 J. Cooper, F.G. Giancotti. Integrin Signaling in Cancer: Mechanotransduction, Stemness, Epithelial Plasticity, and Therapeutic Resistance. *Cancer Cell* **2019**, *35*, 347-367.
80 S. J. Moschos, L. M. Drogowski, S. L. Reppert, J. M. Kirkwood. Integrins and cancer. *Oncology* **2007**, *21*, 13–20.
81 J. S. Desgrosellier, D. A. Cheresh. Integrins in cancer: biological implications and therapeutic opportunities. *Nat. Rev. Cancer* **2010**, *10*, 9-22.
82 M. D. Pierschbacher, E. Ruoslahti. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* **1984**, *309*, 30–33.
83 K.-E. Gottschalk, H. Kessler. The Structures of Integrins and Integrin–Ligand Complexes: Implications for Drug Design and Signal Transduction. *Angew. Chem. Int. Ed.* **2002**, *41*, 3767-3774
84 J.-P. Xiong, T. Stehle, R. Zhang, A. Joachimiak, M. Frech, S. L. Goodman, M. A. Arnaout. Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. *Science* **2002**, *296*, 151-155.
85 L. Auzzas, F. Zanardi, L. Battistini, P. Burreddu, P. Carta, G. Rassu, C. Curti, G. Casiraghi. Targeting αvβ3 Integrin: Design and Applications of Mono- and Multifunctional RGD-Based Peptides and Semipeptides. *Curr. Med. Chem.* **2010**, *17*, 1255-1299.
86 For compound **2**: I. F. Charo, L. Nannizzi, J. W. Smith, D. A. Cheresh. The vitronectin receptor alpha v beta 3 binds fibronectin and acts in concert with alpha 5 beta 1 in promoting cellular attachment and spreading on fibronectin. *J. Cell. Biol.* **1990**, *111*, 2795-2800. For compounds **4** and **5**: S. Mousa, J. Bozarth, M. Forsythe, W. Lorelli, S. Jackson, N. Ramachandran, W. DeGrado, M. Thoolen, T. Reilly. Antiplatelet Efficacy and Specificity of DMP728, a Novel Platelet GPIIb/IIIa Receptor Antagonist. *Cardiology* **1993**, *83*, 374-382.
87 K. Burgess, D. Lim, S. A. Mousa. Synthesis and Solution Conformation of Cyclo[RGDRGD]: A Cyclic Peptide with Selectivity for the αVβ3 Receptor. *J. Med. Chem.* **1996**, *39*, 4520–4526
88 A.C. Bach, R. Espina, S.A Jackson, P.F.W. Stouten, J.L. Duke, S.A. Mousa, W. F. DeGrado. Type II' to type I β-turn swap changes specificity for integrins. *J. Am. Chem. Soc.* **1996**, *118*, 293-294.
89 a) M. Paolillo, M.A. Russo, M. Serra, L. Colombo, S. Schinelli. Small molecule integrin antagonists in cancer therapy. *Mini-Rev. Med. Chem.* **2009**, *9*, 1439-1446; b) Z. Liu, F. Wang, X. Chen. Integrin αVβ3-targeted cancer therapy. *Drug Dev. Res.* **2008**, *69*, 329-339.
90 Compounds **6**-**8**: a) L. Manzoni, L. Belvisi, D. Arosio, M. Civera, M. Pilkington-Miksa, D. Potenza, A. Caprini, E. M. V. Araldi, E. Monferini, M. Mancino, F. Podestà, C. Scolastico. Cyclic RGD-containing functionalized azabicycloalkane peptides as potent integrin antagonists for tumor targeting. *ChemMedChem* **2009**, *4*, 615-632; b) L. Belvisi, A. Bernardi, M. Colombo, L. Manzoni, D. Potenza, C. Scolastico, G. Giannini, M. Marcellini, T. Riccioni, M. Castorina, P. LoGiudice, C. Pisano. Targeting integrins: insights into structure and activity of cyclic RGD pentapeptide mimics containing azabicycloalkane amino acids. *Bioorg. Med. Chem.* **2006**, *14*, 169-180. Compound **9**: R. Haubner, W. Schmitt, G. Hölzemann, S. L. Goodman, A. Jonczyk, H. Kessler. Cyclic RGD Peptides Containing β-Turn Mimetics. *J. Am. Chem. Soc.* **1996**, *118*, 7881–7891. Compound **10**: F. Sladojevich, A. Trabocchi, A. Guarna. Convenient route to enantiopure fmoc-protected morpholine-3-carboxylic acid. *J. Org. Chem*. **2007**, *72*, 4254-4257. Compounds **11**-**13**: a) E. Lohof, E. Planker, C. Mang, F. Burkhart, M. A. Dechantsreiter, R. Haubner, H.-J. Wester, M. Schwaiger, G. Hölzemann, S.L. Goodman, H. Kessler. Carbohydrate Derivatives for Use in Drug Design: Cyclic α_v-Selective RGD Peptides. *Angew. Chem Int. Ed.* **2000**, *39*, 2761 – 2764; b) R. M. van Well, L. Marinelli, C. Altona, K. Erkelens, G. Siegal, M. van Raaij, A. L. Llamas-Saiz, H. Kessler, E. Novellino, A. Lavecchia, J. H. van Boom, M. Overhand. Conformational Analysis of Furanoid ε-Sugar Amino Acid Containing Cyclic Peptides by NMR Spectroscopy, Molecular Dynamics Simulation, and X-ray Crystallography: Evidence for a Novel Turn

Structure. *J. Am. Chem. Soc.* **2003**, *125*, 10822–10829; c) R. M. van Well, H. S. Overkleeft, G. A. van der Marel, D. Bruss, G. Thibault, P. G. de Groot, J. H. van Boom, M. Overhand. Solid-phase synthesis of cyclic RGD-furanoid sugar amino acid peptides as integrin inhibitors. *Bioorg Med Chem Lett.* **2003**, *13*, 331-334; d) G. Casiraghi, G. Rassu, L. Auzzas, P. Burreddu, E. Gaetani, L. Battistini, F. Zanardi, C. Curti, G. Nicastro, L. Belvisi, I. Motto, M. Castorina, G. Giannini, C. Pisano. Grafting aminocyclopentane carboxylic acids onto the RGD tripeptide sequence generates low nanomolar $\alpha_V\beta_3$/ $\alpha_V\beta_5$ integrin dual binders. *J. Med. Chem.* **2005**, *48*, 7675-7687. Compounds **14a-d**: F. Zanardi, P. Burreddu, G. Rassu, L. Auzzas, L. Battistini, C. Curti, A. Sartori, G. Nicastro, G. Menchi, N. Cini, A. Bottoncetti, S. Raspanti, G. Casiraghi. Discovery of subnanomolar arginine-glycine-aspartate-based $\alpha_V\beta_3$/ $\alpha_V\beta_5$ integrin binders embedding 4-aminoproline residues. *J. Med. Chem.* **2008**, *51*, 1771-1782. Compound **15**: S. Urman, K. Gaus, Y. Yang, U. Strijowski, N. Sewald, S. De Pol, O. Reiser. The Constrained Amino Acid β-Acc Confers Potency and Selectivity to Integrin Ligands. *Angew. Chem Int. Ed.* **2007**, *46*, 3796 – 3798.

91 a) M. Marchini, M. Mingozzi, R. Colombo, I. Guzzetti, L. Belvisi, F. Vasile, D. Potenza, U. Piarulli, D. Arosio, C. Gennari. Cyclic RGD peptidomimetics containing bifunctional diketopiperazine scaffolds as new potent integrin ligands. *Chem. Eur. J.* **2012**, *18*, 6195-6207; b) M. Marchini, M. Mingozzi, R. Colombo, C. Gennari, M. Durini, U. Piarulli. Selective O-acylation of unprotected N-benzylserine methyl ester and O,N-acyl transfer in the formation of cyclo[Asp-Ser] diketopiperazines. *Tetrahedron* **2010**, *66*, 9528-9531; c) A. S. M. da Ressurreição, A. Vidu, M. Civera, L. Belvisi, D. Potenza, L. Manzoni, S. Ongeri, C. Gennari, U. Piarulli. Cyclic RGD-peptidomimetics containing bifunctional diketopiperazine scaffolds as new potent integrin ligands. *Chem. Eur. J.* **2009**, *15*, 12184-12188.

92 R. Fanelli, L. Schembri, U. Piarulli, M. Pinoli, E. Rasini, M. Paolillo, M. C. Galiazzo, M. Cosentino, F. Marino. Effects of a novel cyclic RGD peptidomimetic on cell proliferation, migration and angiogenic activity in human endothelial cells. *Vasc. Cell.* **2014**, *6*,11.

93 M. Mammen, S.-K. Choi, G. M. Whitesides. Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chem. Int. Ed.* **1998**, *37*, 2754-2794.

94 C. Fasting, C. A. Schalley, M. Weber, O. Seitz, S. Hecht, B. Koksch, J. Dernedde, C. Graf, E.-W. Knapp, R. Haag. Multivalency as a Chemical Organization and Action Principle *Angew. Chem. Int. Ed.* **2012**, *51*, 10472 – 10498.

95 H. L. Handl, J. Vagner, H. Han, E. Mash, V. J. Hruby, R. J. Gillies. Hitting multiple targets with multimeric ligands *Expert Opin. Ther. Targets* **2004**, *8*, 565-586.

96 M. Weber, A. Bujotzek, R. Haag. Quantifying the rebinding effect in multivalent chemical ligand-receptor systems. *J. Chem. Phys.* **2012**, *137*, 054111.

97 J. Diestler, E. W. Knapp. Statistical mechanics of the stability of multivalent ligand-receptor complexes. *J. Phys. Chem. C* **2010**, *114*, 5287–5304.

98 a) F. Thoreau, L. Vanwonterghem, M. Henry, J.-L. Coll, D. Boturyn. Design of RGD–ATWLPPR peptide conjugates for the dual targeting of αVβ3 integrin and neuropilin-1. *Org. Biomol. Chem.* **2018**, *16*, 4101-4107; b) M. Degardin, D. Thakar, M. Claron, R. P. Richter, L. Coche-Guérente, D. Boturyn. Development of a selective cell capture and release assay: impact of clustered RGD ligands. *J. Mater. Chem. B* **2017**, *5*, 4745-4753; c) L. Sandrin, D. Thakar, C. Goyer, P. Labbé, D. Boturyn, L. Coche-Guérente. Controlled surface density of RGD ligands for cell adhesion: evidence for ligand specificity by using QCM-D. *J. Mater. Chem. B* **2015**, *3*, 5577-5587.

99 A. R. M. Dias, A. Pina, A. Dal Corso, D. Arosio, L. Belvisi, L. Pignataro, M. Caruso, C. Gennari. Multivalency Increases the Binding Strength of RGD Peptidomimetic-Paclitaxel Conjugates to Integrin αVβ3. *Chem. Eur.J.* **2017**, *23*, 14410 –14415.

100 S. J. Kwon, D. H. Na, J. H. Kwak, M. Douaisi, F. Zhang, E. J. Park, J. H. Park, H. Youn, C. S. Song, R. S. Kane, J. S. Dordick, K. B. Lee, R. J. Linhardt. Nanostructured glycan architecture is important in the inhibition of influenza A virus infection. *Nat. Nanotechnol.* **2017**, *12*, 48-54.

101 A. Pina, M. Kadri, D. Arosio, A. Dal Corso, J. Coll, C. Gennari, D. Boturyn. Multimeric Presentation of RGD Peptidomimetics Enhances Integrin Binding and Tumor Cell Uptake. *Chem. Eur. J.* **2020**, *26*, 7492-7496.

102 N. Krall, F. Pretto, D. Neri. A bivalent small molecule-drug conjugate directed against carbonic anhydrase IX can elicit complete tumour regression in mice. *Chem. Sci.* **2014**, *5*, 3640-3644.

103 a) M. Janssen, W. J. G. Oyen, L. F. A. G. Massuger, C. Frielink, I. Dijkgraaf, D. S. Edwards, M. Radjopadhye, F. H. M. Corstens, O. C. Boerman. Comparison of a Monomeric and Dimeric Radiolabeled RGD-Peptide for Tumor Targeting. *Cancer Biother. Radiopharm.* **2002**, *17*, 641-646; b) Z. H. Jin, T. Furukawa, M. Degardin, A. Sugyo, A. B. Tsuji, T. Yamasaki, K. Kawamura, Y. Fujibayashi, M. R. Zhang, D. Boturyn, P. Dumy, T. Saga. $\alpha_V\beta_3$ Integrin-Targeted Radionuclide Therapy with 64Cu-cyclam-RAFT-c(-RGDfK-)4. *Mol. Cancer Ther.* **2016**, *15*, 2076 - 2085; c) Y. Yang, S. Ji, S. Liu. Impact of multiple negative charges on blood clearance and biodistribution characteristics of 99mTc-labeled dimeric cyclic RGD peptides. *Bioconjugate Chem.* **2014**, *17*, 1720-1729.

104 C. A. Rhodes, D. Pei. Bicyclic Peptides as Next-Generation Therapeutics. *Chem. Eur. J.* **2017**, *23*, 12690- 12703.

105 a) K. Deyle, X. D. Kong, C. Heinis. Phage Selection of Cyclic Peptides for Application in Research and Drug Development. *Acc. Chem. Res.* **2017**, *50*, 1866-1874; b) C. Heinis, G. Winter. Encoded libraries of chemically modified peptides. *Curr. Opin. Chem. Biol.* **2015**, *26*, 89-98; c) C. Heinis, T. Rutherford, S. Freund, G. Winter. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* **2009**, *5*, 502-507.

106 a) W. Lian, B. Jiang, Z. Qian, D. Pei. Cell-permeable bicyclic peptide inhibitors against intracellular proteins. *J. Am. Chem. Soc.* **2014**, *136*, 9830-9833; b) B. Jiang, D. Pei. A Selective, Cell-Permeable Nonphosphorylated Bicyclic Peptidyl Inhibitor against Peptidyl-Prolyl Isomerase Pin1. *J. Med. Chem.* **2015**, *58*, 6306-6312; c) T. B. Trinh, P. Upadhyaya, Z. Qian and D. Pei. Discovery of a Direct Ras Inhibitor by Screening a Combinatorial Library of Cell-Permeable Bicyclic Peptides. *ACS Comb. Sci.* **2016**, *18*, 75-85.

107 L. A. Carpino. 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive. *J. Am. Chem. Soc.* **1993**, *115*, 4397–4398.

108 D. Alvarez-Dorta, D. T. King, T. Legigan, D. Ide, I. Adachi, D. Deniaud, J. Désiré, A. Kato, D. Vocadlo, S. G. Gouin, Y. Blériot. Multivalency To Inhibit and Discriminate Hexosaminidases. *Chem. Eur. J.* **2017**, *23*, 9022

109 M. Paolillo, M. C. Galiazzo, A. Daga, E. Ciusani, M. Serra, L. Colombo, S. Schinelli. An RGD small-molecule integrin antagonist induces detachment-mediated anoikis in glioma cancer stem cells. *Int. J. Oncol.* **2018**, *53*, 2683 - 2694;

110 a) F. J. Sulzmaier, C. Jean, D. D. Schlaepfer. FAK in cancer: mechanistic findings and clinical applications. *Nat. Rev. Cancer* **2014**, *14*, 598-610; b) D. S. Harburger, D. A. Calderwood. Integrin signalling at a glance. *J. Cell Sci.* **2009**, *122*, 159-163.

111 S. K. Mitra, D. A. Hanson, D. D. Schlaepfer. Focal adhesion kinase: in command and control of cell motility. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 56-68.

112 J. Singh, R.C. Petter, T. A. Baillie, A. Whitty. The resurgence of covalent drugs. Nat. Rev. Drug Discov. **2011**, *10*, 307-317.

113 J. Uetrecht. Idiosyncratic drug reactions: past, present, and future. *Chem. Res. Toxicol.* **2008**, *21,* 84-92.

114 A. Bandyopadhyay, J. Gao. Targeting biomolecules with reversible covalent chemistry. *Curr. Opin. Chem. Biol.* **2016**, *34*, 110–116.

115 R. Lagoutte, R. Patouret, N. Winssinger. Covalent inhibitors: an opportunity for rational target selectivity. *Curr. Opin. Chem. Biol.* **2017**, *39*, 54–63

116 I.M. Serafimova, M.A. Pufall, S. Krishnan, K. Duda, M.S. Cohen, R.L. Maglathlin, J.M. McFarland, R.M. Miller, M Frodin, J Taunton. Reversible targeting of noncatalytic cysteines with chemically tuned electrophiles. *Nat. Chem. Biol.* **2012**, *8*, 471-476.

117 S. Krishnan, R.M. Miller, B. Tian, R.D. Mullins, M.P. Jacobson, J. Taunton. Design of reversible, cysteine-targeted Michael acceptors guided by kinetic and computational analysis. *J. Am. Chem. Soc*. **2014**, *136*, 12624-12630.

118 G. Springsteen, B.H. Wang. A detailed examination of boronic acid–diol complexation. *Tetrahedron* **2002**, *58*, 5291-5300.

119 R. Bold. "Development of the proteasome inhibitor Velcade (Bortezomib)" by Julian Adams, Ph.D., and Michael Kauffman, M.D., Ph.D. *Cancer Invest.* **2004**, *22*, 328-329.

120 D.A. Bachovchin, B.F. Cravatt. The pharmacological landscape and therapeutic potential of serine hydrolases. *Nat. Rev. Drug. Discov*. **2012**, *11*, 52-68.

121 J. Crugeiras, A. Rios, E. Riveiros, J.P. Richard. Substituent effects on the thermodynamic stability of imines formed from glycine and aromatic aldehydes: implications for the catalytic activity of pyridoxal-5'-phosphate. *J. Am. Chem. Soc.* **2009**, *131*, 15815- 15824.

122 J. Pettinger, K. Jones, M. D. Cheeseman. Lysine-targeting covalent inhibitors. *Angew. Chem. Int. Ed.* **2017**, *56*, 15200–15209.

123 P. M. S. D. Cal, J. B. Vicente, E. Pires, A. V. Coelho, L. F. Veiros, C. Cordeiro, P. M. P. Gois. Iminoboronates: A New Strategy for Reversible Protein Modification. *J. Am. Chem. Soc.* **2012**, *134*, 10299–10305.

124 A. Bandyopadhyay, J. Gao. Iminoboronate formation leads to fast and reversible conjugation chemistry of α-nucleophiles at neutral pH. *Chem. Eur. J.* **2015**, *21*, 14748-14752.

125 G. Akçay, M. A. Belmonte, B. Aquila, C. Chuaqui, A. W. Hird, M. L. Lamb, P. B. Rawlins, N. Su, S. Tentarelli, N. P. Grimster, Q. Su. Inhibition of Mcl-1 through covalent modification of a noncatalytic lysine side chain. *Nat. Chem. Biol.* **2016**, *12*, 931–936.

126 B. Metcalf, C. Chuang, K. Dufu, M. P. Patel, A. Silva-Garcia, C. Johnson, Q. Lu, J. R. Partridge, L. Patskovska, Y. Patskovsky, S. C. Almo, M. P. Jacobson, L. Hua, Q. Xu, S. L. Gwaltney, C. Yee, J. Harris, B. P. Morgan, J. James, D. Xu, A. Hutchaleelaha, K. Paulvannan, D. Oksenberg, Z. Li. Discovery of GBT440, an Orally Bioavailable R-State Stabilizer of Sickle Cell Hemoglobin. *ACS Med. Chem. Lett.* **2017**, *8*, 321–326.

127 A. Dal Corso, M. Catalano, A. Schmid, J. Scheuermann, D. Neri. Affinity enhancement of protein ligands by reversible covalent modification of neighboring lysine residues. *Angew. Chem. Int. Ed.* **2018**, *57*, 17178–17182.
128 S. Gardini, S. Cheli, S. Baroni, G. Di Lascio, G. Mangiavacchi, N. Micheletti, C. L. Monaco, L. Savini, D. Alocci, S. Mangani, N. Niccolai. On Nature's Strategy for Assigning Genetic Code Multiplicity. *PLoS One* **2016**, *11*, e0148174.
129 Figure 27A: Protein: BIR domain of melanoma inhibitor of apoptosis (ML–IAP), Protein Data Bank (PDB): 3F7; literature reference: F. Cohen, B. Alicke, L. O. Elliott, J. A. Flygare, T. Goncharov, S. F. Keteltas, M. C. Franklin, S. Frankovitz, J. P. Stephan, V. Tsui, D. Vucic, H. Wong, W. J. Fairbrother. Orally Bioavailable Antagonists of Inhibitor of Apoptosis Proteins Based on an Azabicyclooctane Scaffold. *J. Med. Chem.* **2009**, *52*, 1723–1730; Figure 27B: Protein: integrin αVβ3, PDB: 1L5G; literature reference: J.-P. Xiong, T. Stehle, R. Zhang, A. Joachimiak, M. Frech, S. L. Goodman, M. A. Arnaout. Crystal Structure of the Extracellular Segment of Integrin αVβ3 in Complex with an Arg-Gly-Asp Ligand. *Science* **2002**, *296*, 151–155; Figure 27C: Protein: type 1 interleukin-1 receptor, PDB: 1G0Y, literature reference: G. P. Vigers, D. J. Dripps, C. K. Edwards III, B. J. Brandhuber. X-ray Crystal Structure of a Small Antagonist Peptide Bound to Interleukin-1 Receptor Type 1. *J. Biol. Chem.* **2000**, *275*, 36927–36933; Figure 27D: Protein: Histone-binding protein RBBP4, PDB: 6ZRD, see: d) P. Hart, P. Hommen, A. Noisier, A. Krzyzanowski, D. Schüler, A. T. Porfetye, M. Akbarzadeh, I. R. Vetter, H. Adihou, H. Waldmann. Structure Based Design of Bicyclic Peptide Inhibitors of RbAp48. *Angew. Chem. Int. Ed.* **2021**, *60*, 1813–1820.
130 a) E. C. Dreaden, S. C. Mwakwari, Q. H. Sodji, A. K. Oyelere, M. A. El-Sayed. Tamoxifen−Poly(ethylene glycol)−Thiol Gold Nanoparticle Conjugates: Enhanced Potency and Selective Delivery for Breast Cancer Treatment. *Bioconjugate Chem.* **2009**, *20*, 2247–2253; b) B. Huang, A. Desai, S. Tang, T. P. Thomas, J. R. Baker. The Synthesis of a c(RGDyK) Targeted SN38 Prodrug with an Indolequinone Structure for Bioreductive Drug Release. *Org. Lett.* **2010**, *12*, 1384–1387.
131 a) E. Lallana, R. Riguera, E. Fernandez. Reliable and Efficient Procedures for the Conjugation of Biomolecules through Huisgen Azide–Alkyne Cycloadditions. *Angew. Chem. Int. Ed.* **2011**, *50*, 8794–8804; b) B. L. Oliveira, Z. Guo, G. J. L. Bernardes. Inverse electron demand Diels–Alder reactions in chemical biology. *Chem. Soc. Rev.* **2017**, *46*, 4895–4950; c) J. Tu, M. Xu, R. M. Franzini. Dissociative Bioorthogonal Reactions. *ChemBioChem* **2019**, *20*, 1615–1627.
132 M. G. Gichinga, S. Striegler. Regioselective alkylation of hydroxysalicylaldehydes. *Tetrahedron* **2009**, *65*, 4917–4922.
133 T. V. Hansen, L. Skattebøl. Ortho-Formylations of Phenols; Preparation of 3-Bromosalicylaldehyde. *Org. Synth.* **2005**, *82*, 64–68; Discussion Addendum: *Org. Synth.* **2012**, *89*, 220-229.
134 Y. Aeschi, S. Drayss-Orth, M. Valášek, F. Raps, D. Häussinger, M. Mayor. Assembly of [2]Rotaxanes in Water. *Eur. J. Org. Chem.* **2017**, *28*, 4091–4103.
135 A. P. Crew, K. Raina, H. Dong, Y. Qian, J. Wang, D. Vigil, Y. V. Serebrenik, B. D. Hamman, A. Morgan, C. Ferraro, K. Siu, T. K. Neklesa, J. D. Winkler, K. G. Coleman, C. M. Crews. Identification and Characterization of Von Hippel-Lindau-Recruiting Proteolysis Targeting Chimeras (PROTACs) of TANK-Binding Kinase 1. *J. Med. Chem.* **2018**, *61*, 583–598.
136 WO2017/95904, 2017, A1, Paragraph 00121; 00139.
137 R.H. Hans, E.M. Guantai, C. Lategan, P.J. Smith, B. Wan, S.G. Franzblau, J. Gut, P.J. Rosenthal, K. Chibale. Synthesis, antimalarial and antitubercular activity of acetylenic chalcones. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 942-944.
138 R.E. Kleiner, Y. Brudno, M.E. Birnbaum, D.R. Liu. DNA-Templated Polymerization of Side-Chain-Functionalized Peptide Nucleic Acid Aldehydes. *J. Am. Chem. Soc.* **2008**, *130*, 4646 – 4659.
139 R. Sanichar, J.C. Vederas. One-Step Transformation of Coenzyme A into Analogues by Transamidation. *Org. Lett.* **2017**, *19*, 1950–1953.
140 W. C. Still, M. Kahn, A. Mitra. Rapid Chromatographic Technique for Preparative Separations with Moderate Resolution. *J. Org. Chem.* **1978**, *43*, 2923 - 2925;
141 G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg. NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **2010**, *29*, 2176 - 2179;
142 A. Spitaleri, S. Mari, F. Curnis, C. Traversari, R. Longhi, C. Bordignon, A. Corti, G.P. Rizzardi, G. Musco. Structural Basis for the Interaction of isoDGR with the RGD-binding Site of $\alpha_V\beta_3$ Integrin. *J. Biol. Chem.* **2008**, *283*, 19757-19768.
143 L. Ferrazzano, D. Corbisiero, E. Potenza, M. Baiula, S. D. Dattoli, S. Spampinato, L. Belvisi, M. Civera, A. Tolomelli. Side chain effect in the modulation of $\alpha_V\beta_3/\alpha_5\beta_1$ integrin activity via clickable isoxazoline-RGD-mimetics: development of molecular delivery systems. *Scientific Reports* **2020**, *10*, 7410.

144 K. Zhu, K. W. Borrelli, J. R. Greenwood, T. Day, R. Abel, R. S. Farid, E. Harder. Docking Covalent Inhibitors: A Parameter Free Approach To Pose Prediction and Scoring. *J. Chem. Inf. Model* **2014**, *54*, 1932-1940.
145 S. Zhu, J. Zhang, G. Vegesna, F. T. Luo, S. A. Green, H. Liu Highly Water-Soluble Neutral BODIPY Dyes with Controllable Fluorescence Quantum Yields. *Org. Lett.* **2011**, *13*, 438–441.
146 P. Gobbo, S. Novoa, M. C. Biesingerb, M. S. Workentin. Interfacial strain-promoted alkyne–azidecycloaddition (I-SPAAC) for the synthesis of nanomaterial hybrids. *Chem. Commun.* **2013**, *49*, 3982-3984.
147 K. He, Z. Zhang, W. Wang, X. Zheng, X. Wang, X. Zhang. Discovery and biological evaluation of proteolysis targeting chimeras (PROTACs) as an EGFR degraders based on osimertinib and lenalidomide. *Bioorg. Med. Chem. Lett.* **2020**, *30*, 127167.