ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

CHIMICA

Ciclo 37

Settore Concorsuale: 03/C1 - CHIMICA ORGANICA

Settore Scientifico Disciplinare: CHIM/06 - CHIMICA ORGANICA

DESIGN AND CHARACTERIZATION OF SHORT PEPTIDIC SEQUENCES MIMICKING THE BEHAVIOR OF LARGER PROTEIC ENTITIES

Presentata da: Tingting He

Coordinatore Dottorato

Luca Prodi

Supervisore

Luca Gentilucci

Co-supervisore

Paola Galletti

Esame finale anno 2025

Abstract

This thesis deals with the design, synthesis, and applications of short peptidic sequences capable of mimicking some behaviors of larger protein entities. Indeed, peptides can represent the minimalist functional, bioactive portion of a protein, while being more stable, less expensive, and less immunogenic. Besides, peptides have the potential to penetrate further into tissues owing to their smaller size.

In the first part of the thesis, peptides have been designed to investigate the interactions between integrins and their protein ligands. Integrins are heterodimers composed of non-covalently bound α and β subunits, and their ligands are glycoproteins expressed on cell surface, or the extracellular matrix proteins. In humans, the combination of 18 α and 8 β subunits generates 24 distinct αβ heterodimers. These receptors represent the most important family of transmembrane receptors that "integrate" extracellular with intracellular environments via transmission of signals. Indeed, integrins mediate bi-directional signaling across the cell membrane (inside-out and outside-in signaling), hence controlling diverse cellular functions, such as spreading, shape, survival, proliferation, apoptosis, etc.

Besides the normal physiological functions, excessive integrin activation can lead to various pathological conditions, encompassing the onset and advancement of cancer, coronary diseases, inflammatory and autoimmune pathologies, etc. For such diseases, anti-integrin antibodies or small-molecule antagonists that prevent receptor activation can be regarded as potential therapeutic options. The RGD-binding integrins, e.g. $\alpha_v\beta_3$, $\alpha_5\beta_1$, $\alpha_v\beta_5$, have garnered significant attention since their discovery in the 1980s, particularly due to their abnormal expression in cancer cells. However, many clinical trials conducted on the RGD-based antagonists have faced repeated failures. Up to now, only a bunch of integrin-based drugs have entered the market as antithrombotic agents, for inflammatory bowel disease, multiple sclerosis, and dry eye disease. This is largely attributed to the fact that ligands proposed as antagonists may also act as partial agonists at low concentrations, which can induce the high-affinity, extended-open conformation of integrins. This triggers outside-in signaling, paradoxically leading to adhesion and unfavorable outcomes in patients. Hence, a deeper understanding of the structural determinants for receptor activation and blockage is currently of the utmost interest.

For $\alpha_4\beta_1$ integrin expressed on leucocytes, the situation is even more challenging, since the exact crystal structure has not yet been reported. In addition, very few potent and selective agonists are currently available for investigating the mechanisms underlying receptor activation. In this context, the first effort of this thesis addresses these challenges. To this aim, we started with a systematic analysis of $\alpha_4\beta_1$ integrin, integrating receptor-ligand interactions of RGD-binding integrins and computational simulations performed with homology or hybrid $\alpha_4\beta_1$ receptors to explain agonism versus antagonism. Subsequently, we constructed a series of LDV-containing cyclopeptides, leading to the identification of several potent ligands against

α_4 integrins. Computational methods were also employed to elucidate possible models for receptor activation or blockage.

In the second part of the thesis, peptide-nanoparticle conjugates are utilized for targeted therapy. As anticipated, despite good initial results in vitro and in animal models, most of clinical trials conducted for integrin antagonists have not met the expectations. Nevertheless, peptidic integrin ligands still maintain noteworthy appeal, when conjugates either covalently or noncovalently to nanoparticles, for targeted drug delivery systems, or to functionalized surfaces, self-assembled monolayers and finally to smart and cell responsive materials. These peptide-material conjugates are expected to allow a range of applications in theranostics, disease monitoring, regenerative medicines, and tissue engineering. In this frame, various nanoplatforms decorated with different integrin-targeting motifs, e.g. LDV- and RGD-based peptides, were prepared, aiming to expand the application of the developed ligands in tumor-targeted therapies.

In the third part of the thesis, peptides have been designed to prevent the assembly of a functional protein. Cancer cells can be selectively targeted also by interfering with specific signaling pathways. The enzyme human lactate dehydrogenase isoform 5 (*h*LDH 5) plays an essential role in cancer metabolism, by catalyzing the final step of glycolysis, that is the conversion of pyruvate to lactate. Since the enzyme is catalytically active only as tetramers, targeting the protein-protein interactions (PPIs) involved in the oligomerization represents a promising strategy to disrupt their function. In this regard, peptides are particularly attractive, as their larger size enables them to effectively target the large flat interfaces of PPIs, which is unlikely to be inhibited by small molecules. Thus, this part focuses on synthesizing peptides capable of inhibiting the tetramerization of *h*LDH 5. As a result, several linear and cyclic peptide inhibitors were identified, demonstrating selective inhibitory efficacy in cancer cells.

In the last part, hybrid peptide-phosphate nanocrystals have been validated as mini-enzymes. Beyond the application as bioactive molecules, peptides can be utilized in organocatalysis. Unlike single amino acids, such as proline, commonly used in 1,2-conjugate addition reactions, peptides present a longer distance between catalytic sites, facilitating also 1,4-addition reactions. However, to be able to rival with the catalytic efficacy of an enzyme, peptide organocatalysts must be easily recycled and reused. Catalysts recycle and reuse is a fundamental issue also from a green and sustainable chemistry perspective. To address this, we explored the in-situ absorption of peptide catalysts using nanocrystalline hydroxyapatite, a fully biocompatible material, which serves as an insoluble solid support for peptides recovery. This strategy was validated in the 1,4-conjugate addition of aldehydes to nitrostyrene, resulted in high enantioselectivity.

Keywords: integrin, receptor-ligand interaction, agonist, peptidomimetics, nanoparticle, lactate dehydrogenase, protein-protein interaction, organocatalysis, hydroxyapatite.

Table of Contents

Chapter IV. Fabrication and Characterization of Versatile Integrin-targeted Nanoplatforms for Diagnostic and Therapeutic Applications 81

Chapter V. Synthesis of Peptides Inhibiting the Assembly of Monomeric Human L-lactate Dehydrogenase into Catalytically-active Homotetramer 109

Chapter I. Introduction

Drug discovery generally focuses on four major categories of biological macromolecules that serve as pharmacological targets: receptors, enzymes, ion channels, and transporters. Among them, the most predominant targets are receptors and enzymes.[1]

Receptors are the proteins within the cytoplasm or on the cell surface membrane that recognize and bind with specific ligands, thereby regulating cellular processes. Molecules upon interaction with their receptor may activate or inactivate a receptor. Agonists activate receptor and produce functional response, while antagonists prevent receptor activation.

Enzymes, on the other hand, are biological catalysts which speed up various biochemical reactions and play a key critical role in numerous cellular activities. Similar to receptor activation or blockage, there are also two possibilities for enzyme-targeted drug design: activation or inhibition of an enzyme. Molecules that increase the activity of an enzyme are called activators, whereas those that decrease the enzymatic activity are termed inhibitors.

The following discussion will primarily focus on these two types of drug targets: receptors, with a specific emphasis on integrin receptors, and enzymes, particularly lactate dehydrogenases (LDHs).

1.1 Integrin Receptors

Integrins are heterodimeric, noncovalently bound proteins expressed on cell surfaces that mediate cell-cell and cell-matrix adhesion.[2,3] Sequencing of the human genome has identified 18 α and 8 β subunits, from which 24 different functional integrins are currently known to be formed. Integrins regulate cellular growth, proliferation, migration, signaling, and cytokine activation and release and thereby play important roles in cell proliferation and migration, apoptosis, tissue repair, as well as in all processes critical to inflammation, infection, and angiogenesis.

Based on ligand specificity, Integrin heterodimers often can be divided into two groups: RGD (Arg-Gly-Asp) binding receptors and non-RGD binding receptors. (Table 1.1)

Table 1.1 Classification of integrin family, based on ligand specificity.

Integrin Receptors		
RGD binding receptors	$\alpha_V\beta_1$, $\alpha_V\beta_3$, $\alpha_V\beta_5$, $\alpha_V\beta_6$, $\alpha_V\beta_8$, $\alpha_5\beta_1$, $\alpha_8\beta_1$, $\alpha_{IIb}\beta_3$	
Non-RGD binding receptors	Collagen receptors	$\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_{10}\beta_1$, $\alpha_{11}\beta_1$
	Laminin receptors	$\alpha_3\beta_1$, $\alpha_6\beta_1$, $\alpha_6\beta_4$, $\alpha_7\beta_1$
	Leukocyte receptors	$\alpha_E\beta_7$, $\alpha_4\beta_1$, $\alpha_4\beta_7$, $\alpha_9\beta_1$, $\alpha_D\beta_2$, $\alpha_L\beta_2$, $\alpha_M\beta_2$, $\alpha_X\beta_2$

The RGD-binding integrins have a common α_V chain or β_1 chain. The minimal recognition motif RGD peptide sequence was first discovered in fibronectin but was later found in several other ECM (extracellular matrix) proteins, such as fibronectin (FN), vitronectin, laminin and osteopontin.[4,5] Among the 24 integrin subtypes known to date, 8 integrin heterodimers recognize the tripeptide RGD motif within ECM proteins, namely: $\alpha_V\beta_1$, $\alpha_V\beta_3$, $\alpha_V\beta_5$, $\alpha_V\beta_6$, $\alpha_V\beta_8$, $\alpha_5\beta_1$, $\alpha_8\beta_1$, and $\alpha_{IIb}\beta_3$. Non-RGD receptors are further subdivided into three categories: leukocyte-specific, collagen-binding and laminin-binding integrins. Leucocyte-specific integrins have a common β_2 chain that is linked to CD18 and bind receptors such as ICAMs (intercellular adhesion molecules) and plasma proteins such as complement components C3b and C4b. Collagen-binding integrins share a common β_1 chain that binds various α chains in integrins $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_{10}\beta_1$, and $\alpha_{11}\beta_1$. Laminin-binding integrins ($\alpha_3\beta_1$, $\alpha_6\beta_1$, $\alpha_7\beta_1$, and $\alpha_6\beta_4$) mediate the adhesion of cells to basement membranes in various tissues.[6]

1.1.1 Integrin structure

As aforementioned, integrins are heterodimeric transmembrane cell adhesion molecules made up of α and β subunits arranged in numerous dimeric pairings. Structurally, in both α and β subunits, a large N-terminal extracellular domain, a single transmembrane region, and a C-terminal intracellular cytoplasmic tail, can be identified (Figure 1.1).[7]

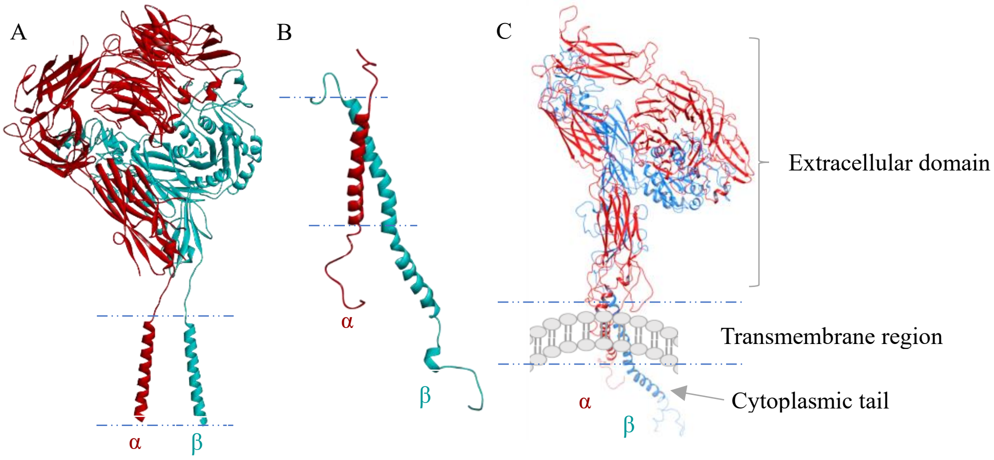


Figure 1.1 Integrin structure. (A) Closed structure of full $\alpha_{IIb}\beta_3$ integrin (PDB code: 8T2V),[8] densities are not visible for the cytoplasmic tails. (B) NMR model of $\alpha_{IIb}\beta_3$ integrin TM/cytoplasmic domain (PDB code: 2KNC).[9] (C) Homology model of full-length integrin $\alpha_{IIb}\beta_3$ (PMDB code: PM0076386).[10]

α-subunit

The α subunit consists of a seven-bladed β-propeller head domain, a thigh domain, two calf domains, a single transmembrane domain, and a short cytoplasmic tail (Figure 1.2 A).[11] Nine of the 18 known Integrin α subunits (α_1, α_2, α_{10}, α_{11}, α_E, α_L, α_M, α_D, and α_X) contain an additional portion called αA or I-domain, which is a domain of approximately 200 amino acids, inserted between blades 2 and 3 in the β-propeller. The I-domain is crucial because it contains the "metal-ion-dependent adhesive site" (MIDAS), which is the main domain responsible for ligand binding (Figures 1.2 B, C).[12] The αI-domain-containing integrins show fairly high homology within their αI domains, which adopts a Rossman fold with five β-sheets surrounded by seven α helices, while the cytoplasmic domains of α chain are highly divergent, only sharing the conserved GFFKR sequence at the membrane-proximal region.[13]

β-subunit

The extracellular region of β subunit (Figure 1.2 A) contains a plexin-sempahorin-integrin (PSI) domain, a hybrid domain, a βI domain, four cysteine-rich epidermal growth factor (EGF) repeats, and a β tail domain (βTD).[14]

The βI domain is structurally homologous to the αI domain of α subunits sharing the same overall fold and, it is inserted into the hybrid domain to form the β head of all integrin β subunits. Additionally, the βI domain features an Mg^{2+} coordinating MIDAS, which is directly involved in the ligand binding in αI-less integrins or indirectly involved in the αI-containing integrins via coordination of Mg^{2+} ion by the Asp/Glu residues; a second site adjacent to MIDAS (ADMIDAS) binding an inhibitory Ca^{2+} ion. and this ADMIDAS site also binds the Mn^{2+} ion leading to a conformational change resulting in an active integrin;[15] and a third site, which is a Ca^{2+}-binding site termed SyMBS (synergistic metal ion-binding site), exerting a positive synergic effect on ligand binding (Figures 1.2 B, D).[12]

In contrast to the α Integrin subunit tails, the β chains share homology in the cytoplasmic tail, with highly-conserved NPX/Y motifs able to bind proteins containing phosphotyrosine binding (PTB) or PTB-like domains.[16]

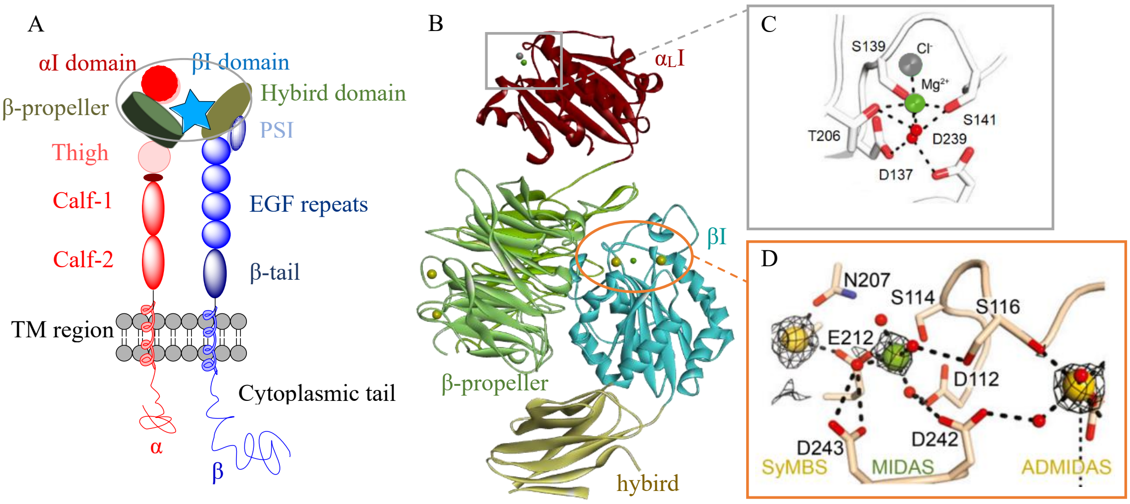


Figure 1.2 Detailed Integrin structure. (A) Schematic structure of Integrin αβ subunits. (B) Representative $\alpha_L\beta_2$ headpiece structure (PDB code: 5E6U). (C) The α_LI-domain MIDAS of $\alpha_L\beta_2$ headpiece. (D) The βI-domain metal ion binding sites of $\alpha_L\beta_2$ headpiece. Mg^{2+}, Ca^{2+}, and Cl^- are shown as green, gold, and silver spheres, respectively.

1.1.2 Integrin activation and signal transduction

Integrin activation is a process of reversible and flexible conformational changes. The solved crystal structures of integrin ectodomains and I domains revealed that integrins exist during activation in the dynamic equilibrium of at least three major conformers: bent-closed (BC, inactive), open-closed (OC, active, low affinity), and open-extended (OE, active, high affinity). Three conformers are schematically shown in Figure 1.3. The delicate balance between different integrin conformations can regulate cell adhesion affinity and signaling strength.

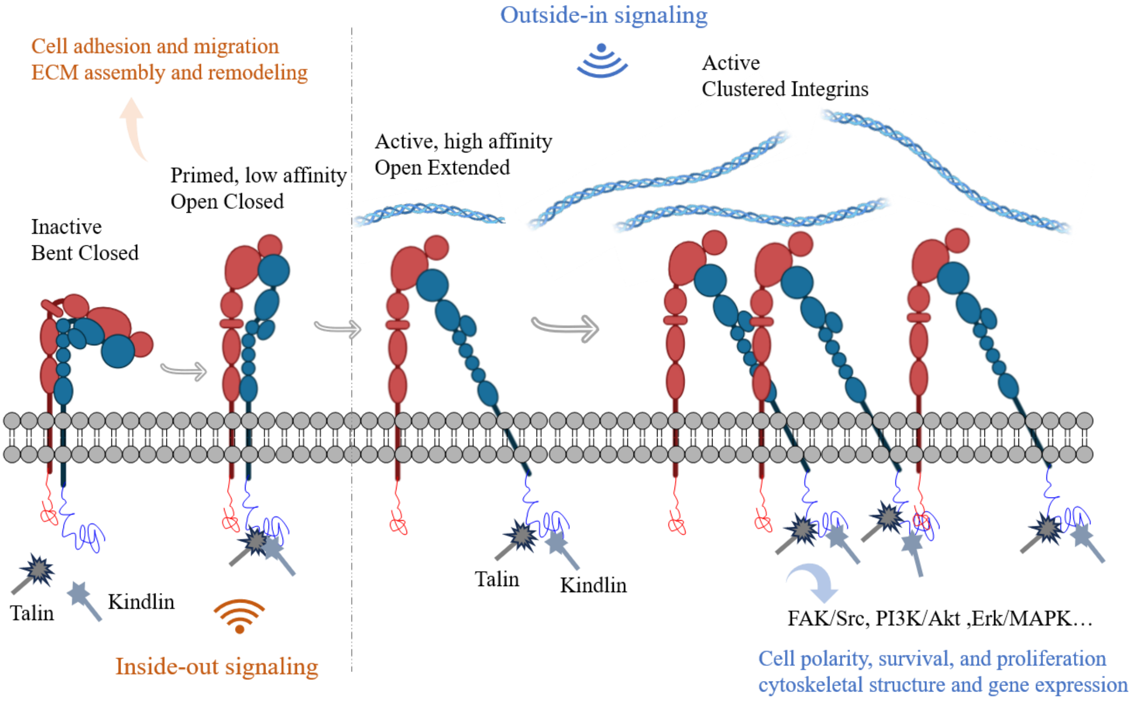


Figure 1.3 Schematic overview of Integrin activation-associated signaling cascades.

As a family of unique transmembrane receptors, integrins can activate bi-directional signaling, referred to as "inside-out" signaling and "outside-in" signaling (Figure 1.3).

Integrins are initially transported from the endoplasmic reticulum to the plasma membrane with the extracellular domains of α- and β-subunits displaying a bent and inactivated or "closed" conformation, while the cytoplasmic domains are linked and stabilized by a salt bridge.

In "inside-out" signaling, intracellular activators like talins or kindlins bind to the cytoplasmic tail of their β-subunits, triggering the conformational changes of integrins and recruiting multivalent protein complexes clustering, which increases their affinity for ligands and thus promotes cell migration and ECM assembly and remodeling (Figure 1.3).[17,18]

Conversely, during "outside-in" signaling, ECM components and other ligands, such as growth factor receptors (GFRs), urokinase plasminogen activator receptor, and transforming growth factor-β (TGF-β) receptor, bind to the external integrin domains, leading to integrin clustering and sending signals into the cells, changing the cell polarity and cytoskeletal structure, inducing gene transcription, and promoting cell survival and proliferation (Figure 1.3).[19] Moreover, recent studies have found that endocytosed integrins also have certain functions as "inside-in" signaling, which can exert a specific role by recruiting focal adhesion kinase (FAK) or enhancing the signaling of the co-trafficking GFRs.[20,21] In summary, integrins enable cells to respond to changes in the extracellular environment and, in turn, affect the extracellular environment itself through bi-directional signaling.

1.1.3 Integrin roles in physiology and pathology

As mentioned earlier, integrins regulate cellular growth, proliferation, migration, signaling, and cytokine activation and release. Consequently, they play critical roles in both physiological processes-such as embryogenesis, organogenesis, the normal growth and tissue development, and pathogenic ones, including tumor metastasis, tumor invasion, angiogenesis, inflammation, ocular diseases, restenosis, and autoimmune diseases, i.e. multiple sclerosis (MS), rheumatoid arthritis (RA), and Crohn's disease.[22,23]

Integrin roles in cancer

Integrins have been found to play multiple key functions in practically every stage of cancer progression, including cancer initiation and proliferation, local invasion and intravasation into the vasculature, survival of circulating tumor cells, priming of the metastatic niche, extravasation into secondary sites and metastatic colonization of new tissues (Figure 1.4).[24] The expression of integrins can differ considerably between normal and tumor tissue and is also associated with cancer types and organotrophic metastasis. For example, integrins $\alpha_v\beta_3$, $\alpha_v\beta_5$, $\alpha_v\beta_6$, $\alpha_4\beta_1$ and $a_5\beta_1$ are usually expressed in most normal epithelia at low or undetectable levels but can be highly upregulated in multiple tumors. In addition, their expressions are thought to correlate with metastasis and poor patient prognosis.[25] The crucial roles of integrins in tumor progression provide valuable clues for cancer diagnosis and targeted therapy. Currently, numerous integrin antagonists/agonists and integrin-targeting nano-therapeutics are under investigation in clinical trials to explore effective regimens and minimize side effects.

It is important to note, however, that in spite of compelling preclinical results demonstrating that the inhibition of integrin $\alpha_V\beta_3$/$\alpha_V\beta_5$ and $\alpha_5\beta_1$ has therapeutic potential, clinical trials with integrin inhibitors targeting those integrins have repeatedly failed to demonstrate therapeutic benefits in cancer patients, and no inhibitors have been registered as anti-cancer drugs.

Nevertheless, it is still widely acknowledged that integrins remain a promising and valid target for cancer therapy, with some key contributions to cancer progression, in particular metastasis, being well-established. Future developments could consider new approaches targeting integrins in other cells of the tumor microenvironment, e.g., cancer-associated fibroblasts and inflammatory/immune cells, rather than focusing solely on tumor cells and vascular cells. Additionally, a deeper understanding of the pharmacological properties of integrin inhibitors, as well as the heterogeneity and redundancy of integrin functions, is essential before advancing novel integrin-targeting agents into clinical trials.

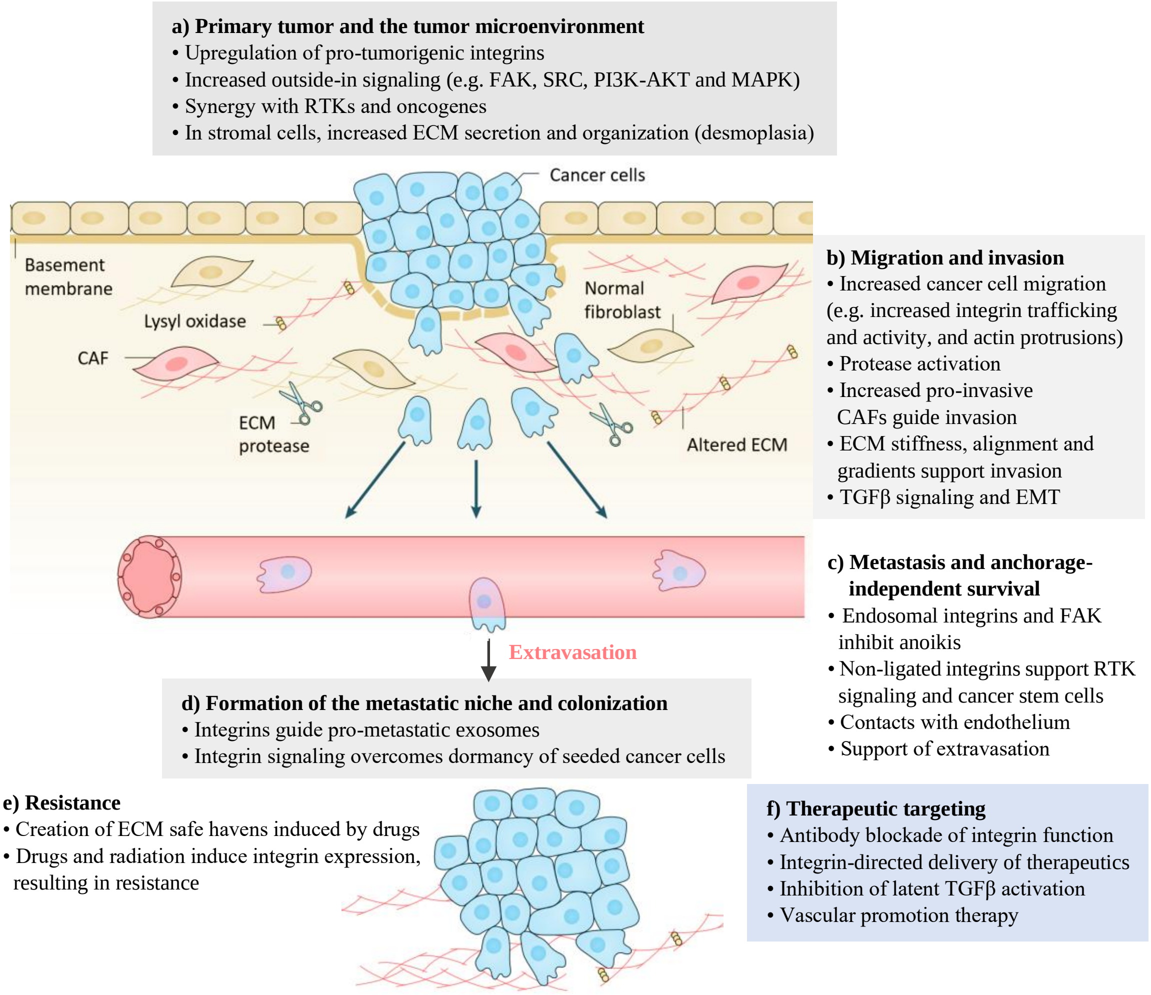


Figure 1.4 Illustration of the multiple roles that integrins play in cancer progression.[24]

Integrin roles in inflammation and infection

Activated integrins are involved in leukocyte extravasation from blood to inflamed tissues. This process consists of a series of sequential molecular interactions known as leukocyte adhesion cascade (Figure 1.5).[26]

In order to home to a site of inflammation, circulating leukocytes roll along the activated endothelial surface via interactions with selectins and chemokines. These interactions trigger inside-out and outside-in signaling cascades[3] that result in the activation of integrins on the leukocytes into high-affinity conformations that facilitate the binding of integrins to their coordinate ligands (e.g., $\alpha_L\beta_2$/ICAM-1; $\alpha_4\beta_1$/VCAM-1 (vascular cell adhesion molecule 1)) and firm adhesion of leukocytes to the endothelial wall. Thereafter, leukocytes move slowly along the surface of the endothelium in a process termed crawling, which ensures the location of an appropriate extravasation site for said leukocytes. Crawling is predominantly mediated by the integrin $\alpha_L\beta_2$ or $\alpha_M\beta_2$ depending on the leukocyte subtype. Subsequently, integrins $\alpha_L\beta_2$, $\alpha_M\beta_2$, and $\alpha_3\beta_1$ interact with various ligands, including ICAM-1/2 and junctional adhesion molecules (JAMs), to facilitate transendothelial migration and detachment of leukocytes (Figure 1.5).[26,27]

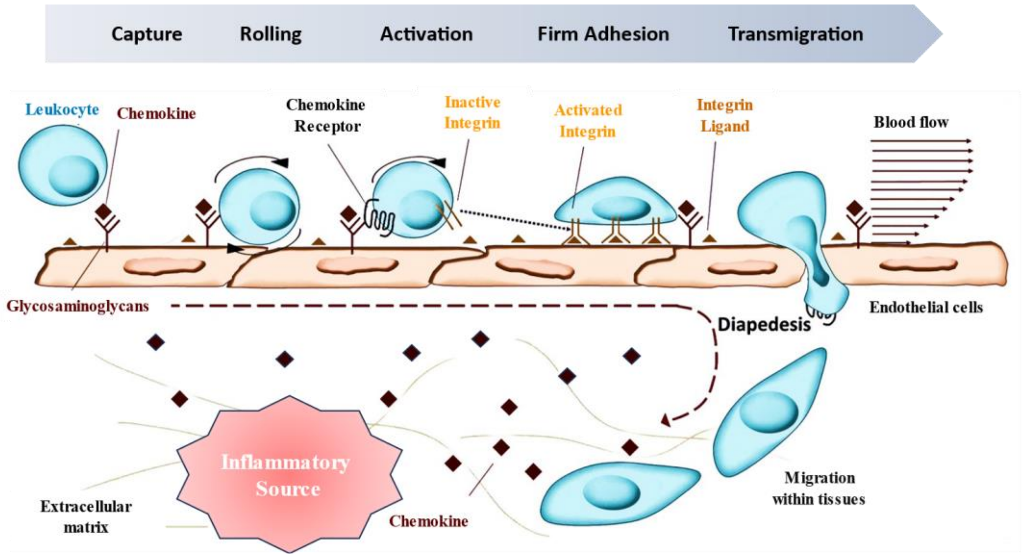


Figure 1.5 Illustration of leukocyte adhesion cascade.[26]

1.1.4 Timeline of milestones in integrin research

The discovery of integrins began in the 1980s, and the term “integrin” was coined in 1986 to highlight their critical roles in maintaining the integrity of the cytoskeletal-ECM linkage.[28]

Among the integrin family, $\alpha_{IIb}\beta_3$ was the first identified family member. Subsequently, the first integrin-targeting drug, Abciximab, was approved in 1994 as an $\alpha_{IIb}\beta_3$ antagonist. After, the first three-dimensional structure of the extracellular domains of $\alpha_v\beta_3$ was reported by Xiong in 2001,[29] providing early evidence for conformational changes in integrins. In 2003, a nano-therapeutic agent with an $\alpha_v\beta_3$-targeting ligand was developed to selectively target angiogenic blood vessels in tumor-bearing mice. The α_L antagonist Efalizumab was approved in 2003 but withdrawn in 2009 due to progressive multifocal leukoencephalopathy (PML).[30] In 2004, the pan-α_4 antagonist Natalizumab was approved for multiple sclerosis. However, a gap in the market followed due to the failure of $\alpha_v\beta_3$ antagonist Cilengitide in clinical phase III trials for the treatment of glioblastoma,[31] impacting α_v-integrin drug discovery. Paradoxically, Cilengitide was found to enhance angiogenesis at low concentrations by inducing conformational changes in $\alpha_v\beta_3$, which is associated with increased tumor growth in vivo. To date, despite α_v-integrin having been explored as a therapeutic target for nearly 40 years, no drugs targeting this integrin have been marketed. In 2014 and 2016, Vedolizumab and Lifitegrast, targeting $\alpha_4\beta_7$ and $\alpha_L\beta_2$ respectively, were approved for inflammatory bowel disease and dry eye disease. CAR T-cell therapy targeting integrins was explored in 2017. A significant breakthrough occurred in 2022 with the approval of Carotegrast by Japan’s PMDA as the first oral anti-integrin drug, the success of the 99mTc-3PRGD2 imaging agent in phase III trials, and positive phase IIa results for the oral $\alpha_v\beta_6/\alpha_v\beta_1$ antagonist PLN-74809.[32] In 2024, FDA has approved ENTYVIO® (Vedolizumab) subcutaneous (SC) administration for maintenance therapy in adults with moderately to severely active Crohn’s disease; Meanwhile, Morf-57, an oral small molecule inhibitor of integrin $\alpha_4\beta_7$, is undergoing phase II clinical trials for Ulcerative

Colitis, with two trials completed and two ongoing; Another promising candidate, 7HP349, works as an allosteric agonist of the integrins VLA-4 (integrin $\alpha_4\beta_1$) and LFA-1 (integrin $\alpha_L\beta_2$). The FDA has granted fast track designation to 7HP349 based on the successful completion of a phase I study that assessed the safety, tolerability and pharmacokinetics in healthy male subjects (Figure 1.6).[33]

To date, the FDA has approved seven drugs targeting various integrins and there are approximately one hundred of integrin-targeting therapies currently in clinical trials, including small molecules, antibodies, peptidomimetics, ADCs, CAR T-cell therapies, and imaging agents. The key lesson from past integrin drug discovery efforts is that success depends on a deep understanding of integrin regulatory mechanisms and addressing unmet clinical needs.[32]

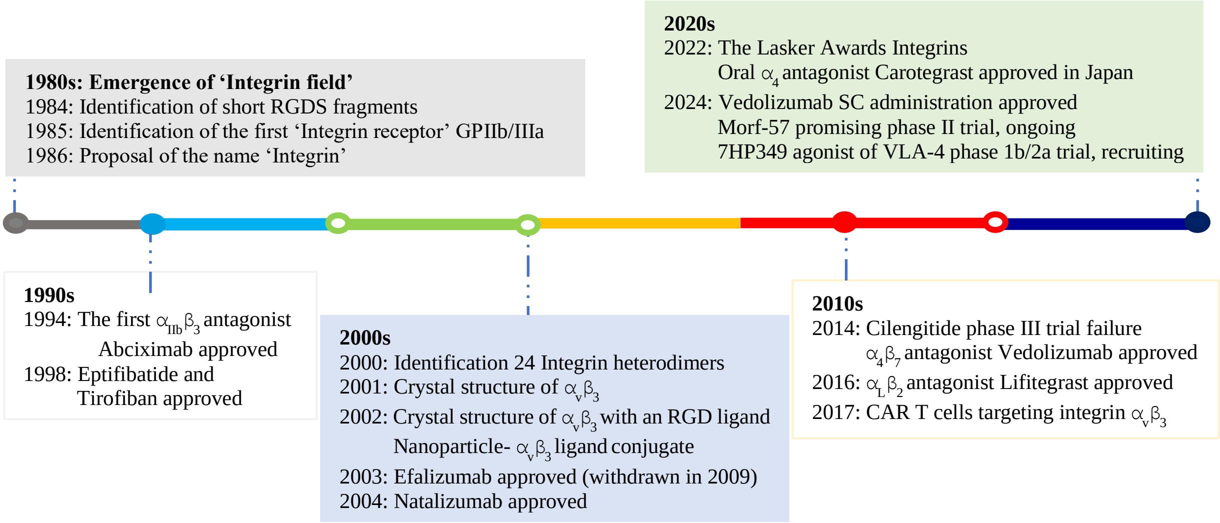


Figure 1.6 Timeline of the historical milestone for the study of integrin.

1.2 Lactate Dehydrogenases (LDHs)

Lactate dehydrogenase (LDH), also known as lactic acid dehydrogenase with an Enzyme Commission Number (EC)[C] of 1. 1. 1. 27, is an intracellular oxidoreductase enzyme that can transfer electrons.[34] LDH plays a central role in several metabolic pathways and is mainly involved in the catalysis of the reversible conversion of pyruvate into lactate, coupled with the oxidation of cofactor NADH to NAD$^+$ (Figure 1.7).

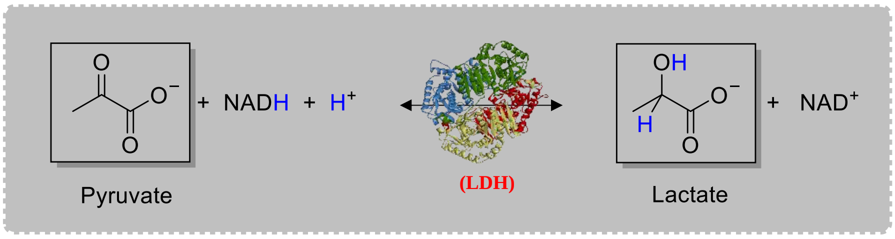


Figure 1.7 The interconversion of pyruvate and lactate, along with the associated conversion of NADH to NAD$^+$, is mediated by the enzyme lactate dehydrogenase (LDH).

1.2.1 LDHs structure and activity

In humans, lactate dehydrogenase (*h*LDH) is a tetrameric molecule (140 kDa) that exists in different isoforms (isozymes) composed of two main types of subunits (35 kDa each): the M-type (LDH-A) and the H-type (LDH-B). These subunits are encoded by different genes: *ldh-a* for the M (muscle) subunit and *ldh-b* for the H (heart) subunit. The two subunits can assemble into homo- or hetero-tetramers, resulting in five major isoforms of functional LDH: H4, H3M1, H2M2, H1M3 and M4, corresponding to five isoforms LDH 1 to LDH 5, respectively. An additional sixth isoform is known as LDH 6 or LDH C. This isoform is mainly found in mature testis and sperm and plays a part in male fertility (Figure 1.8).[35]

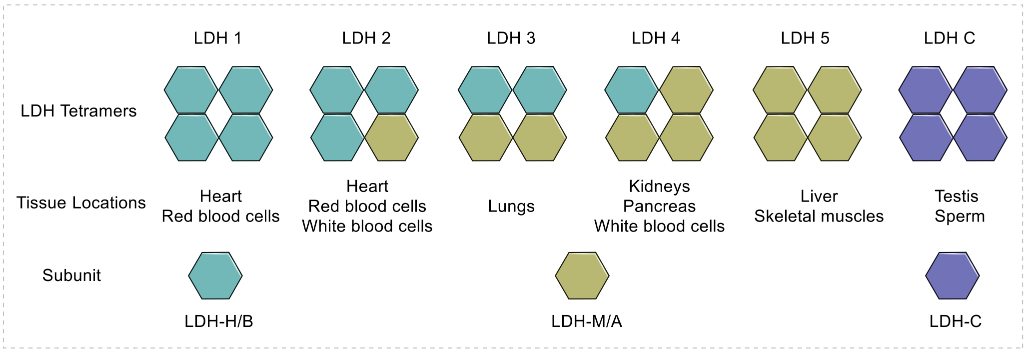


Figure 1.8 The homo- and tetrameric isoforms of lactate dehydrogenase (LDH).

The *h*LDH 5 and *h*LDH 1 subunits are similar in size and share 75% sequence identity, but exhibit different catalytic properties. For example, *h*LDH 5 kinetically favors the conversion of pyruvate to lactate, such as that often found in skeletal muscle and hypoxic tumors. Therefore, *h*LDH 5 is an important isoenzyme under hypoxic conditions and can play a vital part in the development and progression of malignancies. By contrast, *h*LDH 1 has a high capacity in the catalysis of the reverse reaction from lactate to pyruvate, which predominantly occurs in heart and other well-oxygenated tissues.[36,37]

Among all the isoforms, *h*LDH 5 has garnered great interest in recent years, as it was found to be overexpressed in a variety of highly glycolytic human cancers. Although LDH was purified in crystalline form for the first time in 1940 from a heart muscle sample, it was not until 2001 that Brady and co-workers reported the first X-ray structure of *h*LDH A and *h*LDH B in a ternary complex with NADH and oxalate, a structural isostere of pyruvate and a well-known pyruvate competitive inhibitor.[37] Since then, numerous studies using X-ray crystallography, NMR spectroscopy, cryogenic electron microscopy, molecular modeling, and molecular dynamics simulations have been conducted to solve the structures of *h*LDH A and *h*LDH B in their apo, ternary, and inhibitor-bound forms, in order to understand the main binding modes of *h*LDH A.[38-42]

From a structural viewpoint, *h*LDH 5 tetramer shows a D_2 symmetry and comprises 331 amino acids per subunit (Fig. 1.9 A, B). In each monomer, it contains an N-terminal tail that provides a linkage with the C-terminal of the adjacent subunit, which is critical to allow the oligomeric assembly of the tetramer,[43,44] together with two binding sites. The larger one (residues 21-162 and 248-266), characterized by a "Rossmann" domain, is the NADH binding site (also referred to as the cofactor-binding site, Fig. 1.9 D).[38] At this site, NADH cofactors chiefly bind to four residues (Asp^{168}, Arg^{171}, and Thr^{247} and the catalytic His^{195}) located in a groove of the central β-sheet and these residues make key contributions to the catalytic activity of *h*LDH 5. In proximity to the nicotinamide moiety of NADH, a mixed α/β structure forms the substrate-binding site, referred to as the pyruvate-binding site (residues 163-247 and 267-331, Fig. 1.9 C).[38] The substrates, such as pyruvate, interact with three residues (Arg^{171} and Thr^{247} along with Ala^{236}). The two domains form a sort of bilobed structure where a central groove defines the enzyme active site. In addition, residues 95-110 form the so-called "active-site loop", which is involved in the catalytic process and can adopt either an open or a closed conformation, thus regulating the accessibility of substrate and cofactor to the enzyme active site.[37, 45]

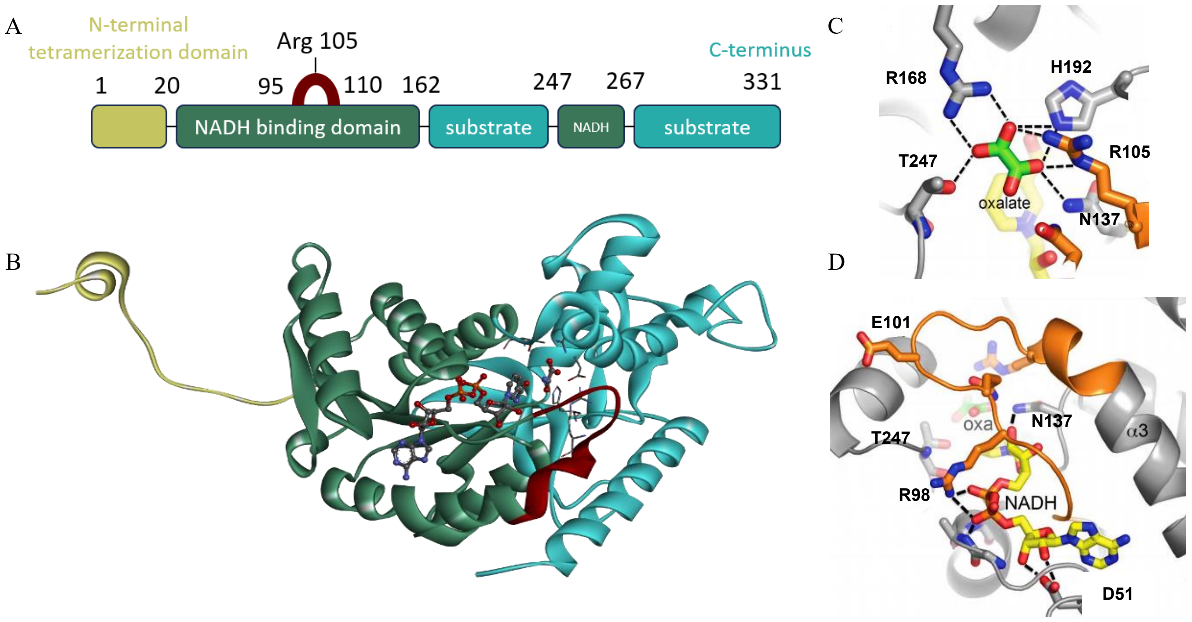


Figure 1.9 Structure and binding modes of *h*LDH 5. (A) Schematic representation of *h*LDH 5 monomer. (B) Crystal structure of an *h*LDH-A subunit. (C) Interactions between oxalate and the LDH-A substrate-binding site. (D) Interactions between NADH and the co-substrate-binding site.[38]

Mechanistically, *h*LDH 5 follows an ordered sequence of events occurring in the millisecond time scale in which NADH binds to the cofactor-binding site with the active-site loop adopting an open conformation, then pyruvate binds to the substrate binding site and, lastly, the active-site loop closes up to provide a largely desolvated ternary complex.[46] Arg^{105} in the active-site loop clamps down on the bound pyruvate. Then a hydride ion is transferred from the nicotinamide ring of NADH to the carbonyl C-atom of the pyruvate.

The observed differences in catalytic activity between *h*LDH A and *h*LDH B mentioned above are thought to be, from a mechanistically viewpoint, the result of altered electrostatic interactions between binding sites and ligands. In fact, the pKa of the active residue His[193] increases from 7.3 in *h*LDH A (likely protonated at physiological pH) to 8.3 in *h*LDH B (non-protonated),[37] representing a potentially exploitable distinction for rational design of *h*LDH A-specific inhibitors.

1.2.2 Targeting LDH for cancer therapeutics

Cancer cells adapt their metabolism to sustain the high energy demand of continuous proliferation. Instead of relying on oxidative phosphorylation as their primary energy source, they shift to glycolysis even in the presence of oxygen phenomenon, known as aerobic glycolysis or the Warburg effect, which is a metabolic hallmark of cancer (Fig. 1.10).[47] By doing this, cancer cells produce energy at a faster rate than with oxidative phosphorylation, obtain intermediates for biosynthesis of macromolecules, and secrete the glycolysis end-product lactate to the extracellular media. Lactate generates an acidic microenvironment that facilitates tumor-associated inflammation, angiogenesis, and immunosuppression, all contributing to tumor survival.[48,49] In addition, reduced oxidative phosphorylation results in lower levels of reactive oxygen species (ROS) and better survival of cancer cells. Therefore, targeting the Warburg effect is an attractive strategy for cancer therapy. As discussed previously, human lactate dehydrogenase (*h*LDH), and more specifically the isoform 5 (LDH 5), plays an essential role in the Warburg effect by catalyzing the final step of glycolysis, i.e., the conversion of pyruvate to lactate, coupled with the regeneration of NAD^+ from NADH.

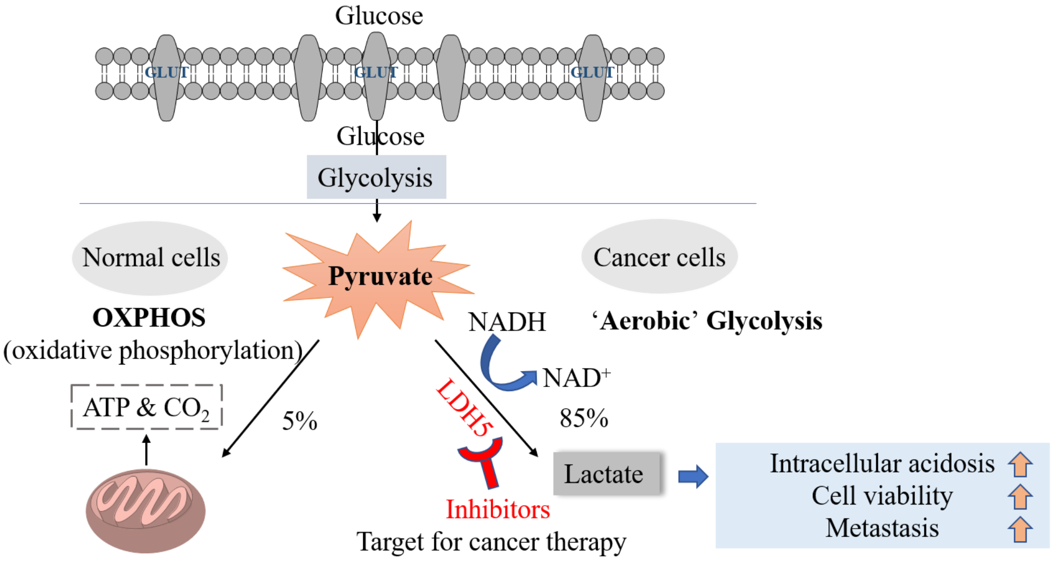


Figure 1.10 Metabolic pathway in normal and cancer cells.

Extensive studies have shown that LDH A is directly or indirectly involved in various aspects of tumor growth, migration, invasion, maintenance, and is associated with poor prognosis in a wide range of tumors. High serum LDH concentrations are also associated with radio-resistance in both primary and metastatic brain tumors. Moreover, multiple studies on various cell lines have shown that attenuation of LDHA increases apoptosis and reduces migration and invasion ability. Notably, human genetic defects in the LDH A gene are nonlethal.[50] Taken together, these studies suggest that LDH A inhibition could be a suitable, safe, and well-tolerated therapy that will impede tumor growth and metastasis.

1.3 Peptide and Peptidomimetics

1.3.1 Peptidomimetics definition and commonly used synthetic strategies

Peptides, typically serving as receptor agonists or antagonists, enzyme substrates or inhibitors, as well as hormones and neurotransmitters, modulate cell-cell signaling pathways in a range of biochemical research and pathological processes.

The use of peptides as therapeutic drugs, however, is complicated by their bioavailability and biostability. Rapid degradation by proteases, poor oral availability, difficult transportation through cell membranes, and nonselective receptor binding are the major limitations of peptides as active pharmaceutical ingredients.[51] Hence suitable chemical modification of native peptides is necessary to bypass the above-mentioned challenges and make them a better drug candidate, which has led to the development of more drug-like molecules known as peptidomimetics.

By definition, peptidomimetics are “compounds whose essential elements (pharmacophore) mimic a natural peptide or protein in 3D space and which retain the ability to interact with the biological target and produce the same biological effect”. Peptidomimetics overcome the inherent obstacles of native peptides by exhibiting higher metabolic stability and improved receptor affinity and selectivity.

The commonly used strategies for designing peptidomimetics include cyclization, replacing peptide bonds with their isostere equivalents, and the incorporation of unnatural amino acids or dipeptide surrogates (Figure 1.11). Among them, macrocyclization can be achieved by lactamization, lactonization, disulfide bridge formation, azide-alkyne cycloadditions or ring-closing metathesis (RCM). Additionally, unnatural amino acids are obtained by several modifications such as D-amino acids, side chain substitution, N-alkylation, isosteric replacement within peptide backbone, and cyclization.[52-55]

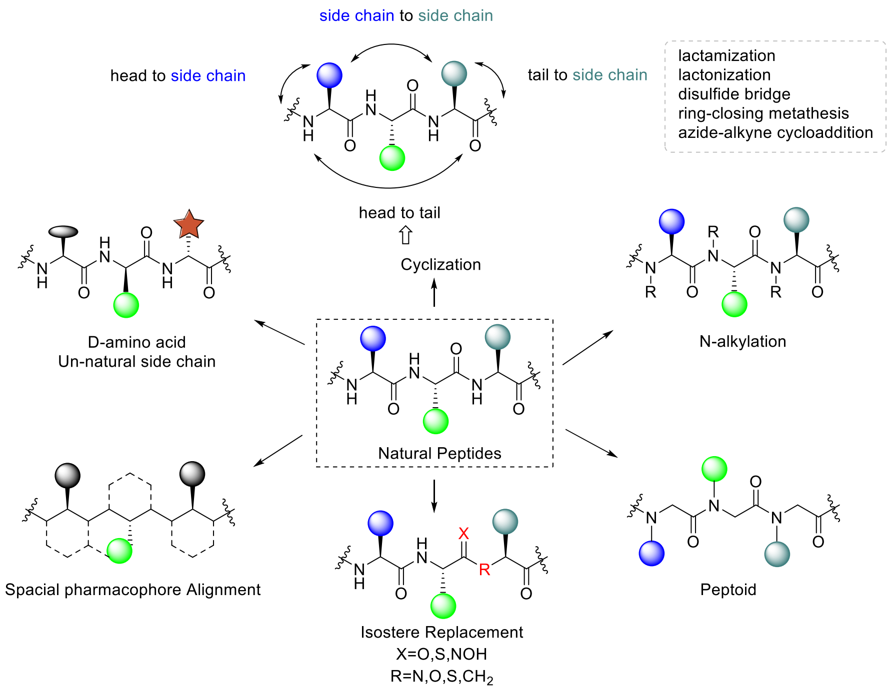


Figure 1.11 Various synthetic strategies of peptidomimetics.

1.3.2 Application in drug discovery and asymmetric catalysis

Peptides/peptidomimetics have emerged as versatile tools in a wide range of fields, including biotechnology, medicinal chemistry, material science, and catalysis. Generally, the design of peptidomimetics begins with an exhaustive structure-activity relationship (SAR) investigation of the parent peptides or proteins to identify the minimal segments or key residues responsible for bioactivity. This is followed by determining the most active 3D arrangement of these key residues using conformationally rigid analogues. Based on the resulting 3D model, the pharmacophores are then recombined using non-peptidic scaffolds. Finally, the synthesized compounds are tested to evaluate their biological activity, enzymatic stability, bioavailability, conformational stability, and ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) profiles.

Here are several selected examples of peptidomimetic strategies applied in the field of Integrins, LDH, and asymmetric catalysis:

Peptidomimetic strategy applied in Integrins

Since RGD was first identified as the minimal recognition sequence in fibronectin by Pierschbacher and Ruoslahti in 1984,[4] numerous peptidic and non-peptidic RGD-based integrin ligands, with various degrees of specificity, have been extensively explored by researchers.

Notably, Kessler and co-workers conducted a systematic study known as spatial screening, which allowed to investigate the 3D disposition of the pharmacophoric groups of the RGD sequence.[56,57] In this context, the potent pentapeptide cyclo(RGDfV) was identified, showing improved receptor affinity and selectivity.[58] Further optimization in their group yielded another two important molecules: c[RGDf-N(Me)V] and c[RGDfK] (Figure 1.12). Notably, c[RGDf-N(Me)V], or cilengitide, was renowned for its outstanding affinity for $\alpha_v\beta_3$ integrins (IC_{50}= 0.65 nM) with improved biostability and oral availability.[59-61] On the other hand, the incorporation of Lys in c[RGDfK][61,62] facilitated its linkage to other molecules, nanoparticles, or fluorescent tags, thus enabling the development of drug delivery or diagnostic systems. Starting from these cyclic structures, the retro (DGR), inverso (rGd), retro-inverso (dGr), and other peptidomimetic strategies were extensively adopted.[52,53,55,63]

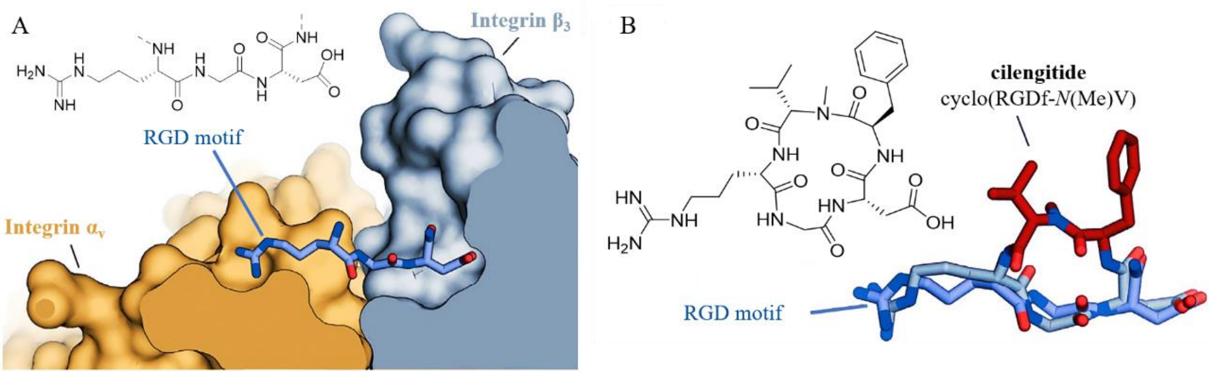


Figure 1.12 (A) Crystal structure of the RGD in fibronectin bound to the $\alpha_v\beta_3$ integrin (PDB code: 4MMX). (B) Chemical structure of the cilengitide and crystal structure (PDB code: 1L5G) superimposed with fibronectin RGD.[55]

Peptidomimetic strategy applied in LDH

Cancer cells undergo metabolic changes compared to normal cells, which favors aerobic glycolysis over oxidative phosphorylation pathway.[47] This altered metabolic pathway supports the survival and proliferation needs for tumor progression. During this metabolic process, lactate dehydrogenase 5 (LDH 5), plays an essential role by catalyzing the conversion of pyruvate to lactate and the regeneration of NAD^+ from NADH. Hence, inhibitors of LDH could be promising candidates for cancer therapy. Currently, there are many small molecules LDH 5 inhibitors that compete with pyruvate and/or NADH for the active site.[64,65] However, to date, no LDH 5 inhibitor has progressed through clinical studies owing to either toxicity caused by cross-reactivity with other NADH-dependent enzymes, or low potency in vivo caused by limited cell penetration and stability.

Alternatively, inhibition of LDH 5 can be realized by disruption of protein-protein interactions (PPIs) between LDH A subunits involved in the oligomerization of LDH 5, as the enzyme is only catalytically active as a tetramer. In this regard, peptides hold great potential because their larger size allows them to better target the PPIs large flat interfaces, which is unlikely to be inhibited by small molecules. Moreover, peptides usually offer higher selectivity and can be further optimized to improve their pharmacokinetic profiles by adopting a peptidomimetic strategy.[52,53]

In 2020, Frédérick and coworkers[66] discovered a short linear peptide, LB8, which can weakly disrupt the tetramerization of lactate dehydrogenase B. They hypothesized that the weak potency of LB8 was due to its poor helical folding and suggested that enhancing its conformational stability could probably improve its potency. In order to increase helicity, they first introduced α-methylalanine (Aib), a well-known helicogenic residue, into the parent linear LB8 peptide, resulting in two linear Aib-containing LB8 analogues, but no improvement was observed. Further NMR 2D ROESY analysis confirmed again the absence of helicity in these peptides. In their second attempt to promote LB8 helicity, the researchers focused on peptide macrocyclization by cysteine stapling with an α-helix-promoting agent. On the basis of LB8 SAR, cysteines were introduced at various i and i + 4 positions, and the peptides were cross-linked using α, α′-dibromo-m-xylene. The resulting cyclopeptides were then evaluated for their binding affinity to LDH-Htr (a truncated version of human LDH H that lacks its 19 N-terminal amino acids). Among these, the most potent analogue, resulting from stapling between positions 1 and 5, showed a 17-fold increase in potency (K_d = 64 μM) compared to LB8 (K_d = 1.05 mM), demonstrating the stapled peptide effectively competed with the N-terminal tetramerization domain of LDHs, leading to the destabilization and disruption of the protein complex and interfering with LDH tetramer formation (Figure 1.13).

To date, only a few examples of using peptides to disrupt the LDH 5 tetramerization have been reported.[44,66-70] This approach remains both challenging and highly promising, given its advantages that small molecules are unlikely to provide, as previously discussed.

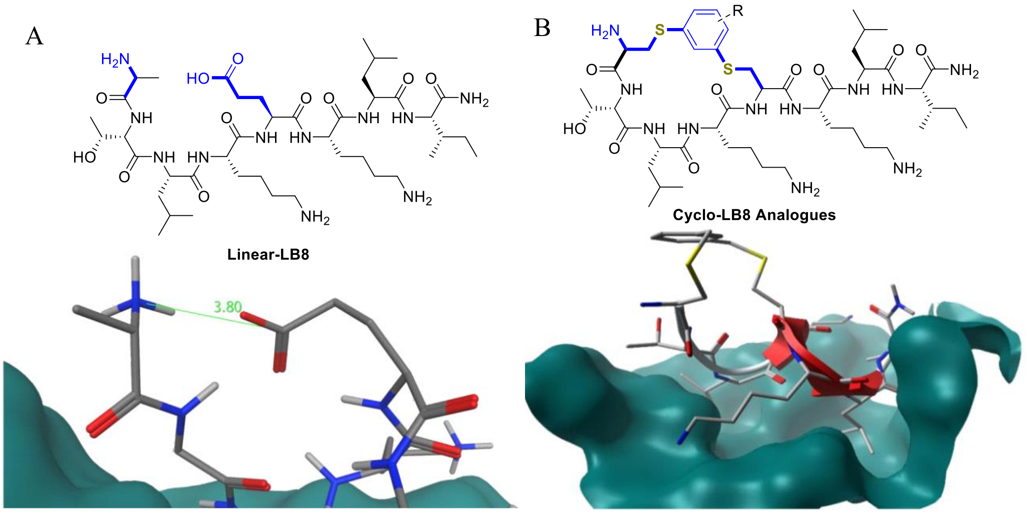


Figure 1.13 Chemical structures and interaction of (A) liner LB8, and (B) cyclo-LB8 analogues with the LDH-H tetramerization site.[66]

Peptidomimetic strategy applied in Asymmetric catalysis

Nature has a vast repertoire of enzyme functions with high catalytic efficiency and specificity. It is hypothesized that enzymes evolved from cooperative interactions among peptides, where functionality emerged from the fusion and diversification of short peptide sequences, identified today as highly conserved regions, responsible for enzyme functionality.

Indeed, peptides have been successfully employed as catalysts in various reactions,[71,72] with the field of organocatalysis showing that even single amino acids such as proline residues are able to catalyze a chemical reaction without the need of being part of an enzyme,[73] demonstrating their potential not only in drug discovery but also in asymmetric catalysis.

However, the efficiency of peptide catalysts in aqueous media remains modest and has met considerable skepticism,[74] likely due to conformational variability and the lack of stabilization from a three-dimensional enzyme structure that allows for a more efficient substrate conversion. Nevertheless, simpler molecules such as small peptides and amino acids can be advantageous because they are easy to synthesize, are modular, tunable, and can act at non-standard pH and temperature conditions. Additionally, the ongoing development of peptidomimetics offers a powerful tool for synthesis of novel peptide-based catalysts through rational design, aiming to precisely position the catalytic site and enhance the catalytic performance of native peptides.

Since the BINOL-derived chiral phosphoric acid (CPA) was first reported in 2004,[75,76] it has been applied to a wide range of transformations, earning its status as a privileged catalyst. However, progress in developing alternative catalyst scaffolds has been relatively slow. The C_2-symmetric scaffold in CPA has been proposed to be crucial for effective catalysis, as this property makes tautomerization of the phosphoric acid proton between the two phosphate oxygens degenerate, which would otherwise provide distinct diastereomeric species with potentially different catalytic activities.

In pursuit of an alternative to the archetypal CPA scaffold, Miller's group postulated that short-sequence peptides could serve as a new family of chiral Brønsted acid catalysts. In this case, catalyst rigidity would perhaps not be necessary, as noncovalent interactions between the flexible peptide backbone and substrates could enable high degrees of selectivity. Moreover, the modularity and tunability of the peptide sequence would enable the development of catalysts with highly diverse structures and functions. Their investigation into phosphothreonine-based peptidic CPA catalysts (pThr-based CPAs) started with studies on the HEH (Hantzsch Ester)-mediated transfer hydrogenation of quinolines (Figure 1.14). Evaluation of a small library of peptides containing an N-terminal pThr residue led to the identification of a Met-containing phosphopeptide, which mediated the model reaction in good yields and up to 93:7 er. These selectivities were comparable to those afforded by the known BINOL-derived CPA (13:87 er). Such high selectivities could be attributed to: (1) the incorporation of critical secondary structural elements into the peptide, in this case a β-turn motif, and (2) secondary interactions between the catalyst and substrate directing group, specifically the 8-urea group.[77]

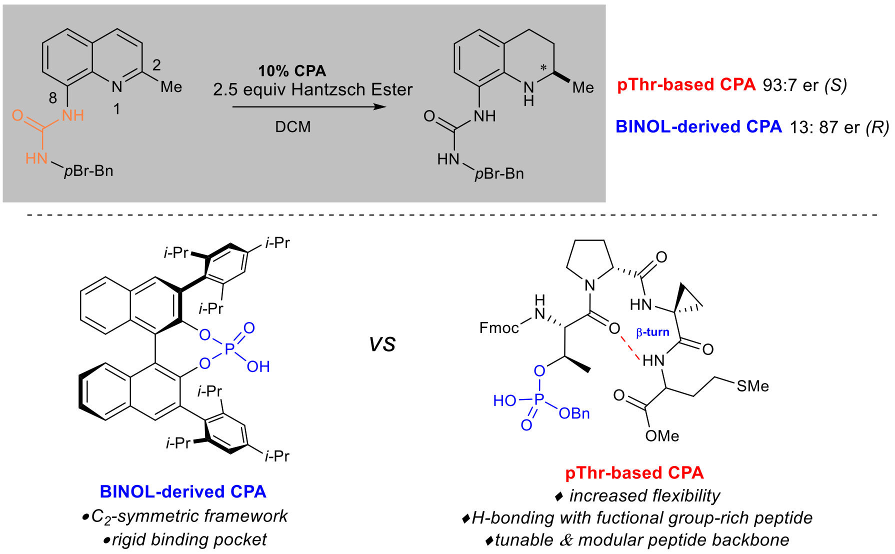


Figure 1.14 The transfer hydrogenation of 8-aminoquinoline catalyzed by BINOL or pThr-based organocatalysts.

1.4 Aim of the Thesis

This thesis aims to design and synthesize bioactive peptides/peptidomimetics capable to mimic, or interfere with protein behavior, with applications as integrin agonists or antagonists, as lactate dehydrogenase inhibitors, and other aspects, such as organocatalysts.

Based on the introduction above, it is evident that integrin receptors and lactate dehydrogenases play essential roles in regulating various physiological and pathological processes. Their broad involvement in a wide range of diseases, including cancer and inflammation, highlights their significant potential as validated drug targets. Furthermore, the development of peptidomimetics strategies offers greater potential for the discovery of highly potent and selective drug molecules.

Firstly, given the crystallographic or cryo-EM structures of $\alpha_4\beta_1$ integrin have not been disclosed, and considering its important role in both physiological and pathological settings, we will thoroughly review the few agonists and antagonists that have been reported to date. We will also speculate on plausible models to explain agonism versus antagonism by comparison with the well-studied RGD-binding integrins and by analysis of computational simulations performed with homology or hybrid receptor structures.

Subsequently, a minilibrary of α/β hybrid cyclopentapeptides containing the LDV recognition motif will be designed and subsequently subjected to biological activity evaluation, in the hope to identify compounds with distinct pharmacological behavior and binding affinity towards α_4 integrin. Moreover, conformational analysis and molecular docking will be performed to elucidate the correlation between secondary structure of the peptides and their ligand-receptor interactions, possibly shedding light on the mechanisms of receptor inhibition or activation.

Building on the results obtained in Chapter 3, several selective and potent $\alpha_4\beta_1$ peptidomimetic ligands will be designed by introducing appropriate modifications for conjugation to gold and PLGA nanoparticles, respectively, thus enabling their potential application in integrin-targeting diagnostics and drug delivery in cancer and inflammation diseases.

Next, recognizing the crucial role of lactate dehydrogenase (LDH) in cancer cell metabolism and that the activity of LDHs is exclusively exerted by their tetrameric forms, we conceive to target LDH tetramerization sites with the ambition of disrupting their oligomeric state. To this end, we will exploit the isolation of recombinant monomeric *h*LDH-A at neutral pH, and its use to identify peptides inhibiting the assembly of the tetrameric enzyme.

Lastly, the potential application of peptides as enantioselective catalysts in a green and sustainable context will be explored. We plan to use tripeptides in amalgam with biocompatible, inert nanocrystalline hydroxyapatite (HAp) as a blended recyclable catalytic system in asymmetric 1,4-conjugate addition reactions and evaluate their catalytic performance.

References

[1] Ashraf, S.; Bashir, N.; Rashid, N.; Chughtai, A. H.; Zia, K. M.; Majeed, S.; Ashiq, M. N.; Murtaza, G.; Najamul-Haq, M. Introduction to Drugs, Drug Targets and Drug Resistance. In *Biochemistry of Drug Resistance*. Springer International Publishing: Cham, **2021**, pp 1-31.

[2] Hynes, R.O. Integrins: A family of cell surface receptors. *Cell* **1987**, *48* (4), 549-554.

[3] Hynes, R.O. Integrins: Bidirectional, Allosteric Signaling Machines. *Cell* **2002**, *110* (6), 673-687.

[4] Pierschbacher, M.D.; Ruoslahti, E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* **1984**, *309* (5963), 30-33.

[5] Ruoslahti, E. RGD AND OTHER RECOGNITION SEQUENCES FOR INTEGRINS. *Annu. Rev. Cell Dev. Biol.* **1996**, *12*, 697-715.

[6] Nieberler, M.; Reuning, U.; Reichart, F.; Notni, J.; Wester, H.J.; Schwaiger, M.; Weinmüller, M.; Räder, A.; Steiger, K.; Kessler, H. Exploring the Role of RGD-Recognizing Integrins in Cancer. *Cancers (Basel)* **2017**, *9*, 116.

[7] Barczyk, M.; Carracedo, S.; Gullberg, D. Integrins. *Cell Tissue Res.* **2010**, *339* (1), 269-280.

[8] Adair, B.D.; Xiong, J.-P.; Yeager, M.; Arnaout, M.A. Cryo-EM structures of full-length integrin $\alpha_{IIb}\beta_3$ in native lipids. *Nat. Commun.* **2023**, *14*, 4168.

[9] Yang, J.; Ma, Y.-Q.; Page, R.C.; Misra, S.; Plow, E.F.; Qin, J. Structure of an integrin $\alpha_{IIb}\beta_3$ transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. *Proc. Natl. Acad. Sci. USA.* **2009**, *106* (42), 17729-17734.

[10] Rosano, C.; Rocco, M. Solution properties of full-length integrin $\alpha_{IIb}\beta_3$ refined models suggest environment-dependent induction of alternative bent/extended resting states. *The FEBS Journal* **2010**, *277*, 3190-3202.

[11] Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T. Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily. *J. Cell Biol.* **1989**, *108* (2), 703-12.

[12] Sen, M.; Springer, T. A. Leukocyte integrin $\alpha_L\beta_2$ headpiece structures: The αI domain, the pocket for the internal ligand, and concerted movements of its loops. *Proc. Natl. Acad. Sci. USA.* **2016**, *113* (11), 2940-2945.

[13] Lu, C.F.; Springer, T. A. The alpha subunit cytoplasmic domain regulates the assembly and adhesiveness of integrin lymphocyte function-associated antigen-1. *J. Immunol.* **1997**, *159* (1), 268-78.

[14] Zent, R.; Pozzi, A. Cell-Extracellular Matrix Interaction in Cancer; Springer: New York, NY, USA, **2010**.

[15] Mould, A.P.; Barton, S.J.; Askari, J.A.; Craig, S.E.; Humphries, M.J. Role of ADMIDAS Cation-binding Site in Ligand Recognition by Integrin $\alpha_5\beta_1$. *J. Biol. Chem.* **2003**, *278* (51), 51622-51629.

[16] Smith, M.J.; Hardy, W.R.; Murphy, J.M.; Jones, N.; Pawson, T. Screening for PTB domain binding partners and ligand specificity using proteome-derived NPXY peptide arrays. *Mol. Cell. Biol.* **2006**, *26* (22), 8461-74.

[17] Carman, C.V.; Springer, T.A. Integrin avidity regulation: are changes in affinity and conformation underemphasized? *Curr. Opin. Cell Biol.* **2003**, *15* (5), 547-56.

[18] Wegener, K.L.; Campbell, I.D. Transmembrane and cytoplasmic domains in integrin activation and protein-protein interactions (review). *Mol. Membr. Biol.* **2008**, *25* (5), 376-87.

[19] Giancotti, F.G.; Ruoslahti, E. Integrin signaling. *Science* **1999**, *285* (5430), 1028-32.

[20] Alanko, J.; Mai, A.; Jacquemet, G.; Schauer, K.; Kaukonen, R.; Saari, M.; Goud, B.; Ivaska, J. Integrin endosomal signaling suppresses anoikis. *Nat. Cell Biol.* **2015**, *17* (11), 1412-21.

[21] Gao, Q.; Sun, Z.; Fang, D. Integrins in human hepatocellular carcinoma tumorigenesis and therapy. *Chin. Med. J. (Engl.)* **2023**, *136* (3), 253-268.

[22] Hillis, G.S.; MacLeod, A.M. lntegrins and disease. *Clin. Sci.* **1996**, *91*, 639-650.

[23] Winograd-Katz, S.E.; Fassler, R.; Geiger, B.; Legate, K.R. The integrin adhesome: From genes and proteins to human disease. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 273-288.

[24] Hamidi, H.; Ivaska, J. Every step of the way: integrins in cancer progression and metastasis. *Nat. Rev. Cancer* **2018**, *18* (9), 533-548.

[25] Liu, F.; Wu, Q.; Dong, Z.; Liu, K. Integrins in cancer: Emerging mechanisms and therapeutic opportunities. *Pharmacol. Ther.* **2023**, *247*, 108458.

[26] Bianchi, E.; Molteni, R.; Pardi, R.; Dubini, G. Microfluidics for in vitro biomimetic shear stress-dependent leukocyte adhesion assays. *J. Biomech.* **2013**, *46* (2), 276-283.

[27] Muller, W. A. Mechanisms of Transendothelial Migration of Leukocytes. *Circ. Res.* **2009**, *105* (3), 223-230.

[28] Tamkun, J.W.; DeSimone, D.W.; Fonda, D.; Patel, R.S.; Buck, C.; Horwitz, A.F.; Hynes, R.O. Structure of integrin, a glycoprotein involved in the transmembrane linkage between fibronectin and actin. *Cell* **1986**, *46* (2), 271-82.

[29] Xiong, J.-P.; Stehle, T.; Diefenbach, B.; Zhang, R.; Dunker, R.; Scott, D.L.; Joachimiak, A.; Goodman, S.L.; Arnaout, M.A. Crystal Structure of the Extracellular Segment of Integrin $\alpha_V\beta_3$. *Science* **2001**, *294* (5541), 339-345.

[30] Seminara, N.M.; Gelfand, J.M., Assessing long-term drug safety: lessons (re) learned from raptiva. *Semin. Cutan. Med. Surg.* **2010**, *29* (1), 16-9.

[31] Stupp, R.; Hegi, M.E.; Gorlia, T.; Erridge, S.C.; Perry, J.; Hong, Y.-K; Weller, M. *et al.* Cilengitide combined with standard treatment for patients with newly diagnosed glioblastoma with methylated MGMT promoter (CENTRIC EORTC 26071-22072 study): a multicentre, randomised, open-label, phase 3 trial. *The Lancet Oncology* **2014**, *15* (10), 1100-1108.

[32] Pang, X.; He, X.; Qiu, Z.; Zhang, H.; Xie, R.; Liu, Z.; Gu, Y.; Zhao, N.; Xiang, Q.; Cui, Y. Targeting integrin pathways: mechanisms and advances in therapy. *Sig. Transduct. Target. Ther.* **2023**, *8* (1), 1.

[33] Lokugamage, N.; Chowdhury, I.H.; Biediger, R.J.; Market, R.V.; Khounlo, S.; Warier, N.D.; Hwang, S.A.; Actor, J.K.; Woodside, D.G.; Marathi, U.; Vanderslice, P.; Garg, N.J. Use of a small molecule integrin activator as a systemically administered vaccine adjuvant in controlling Chagas disease. *NPJ Vaccines* **2021**, *6* (1), 114.

[34] Farhana, A.; Lappin, S.L. Biochemistry, Lactate Dehydrogenase. In *StatPearls*, StatPearls. **2024**. PMID: 32491468.

[35] Clausen, J. Lactate dehydrogenase isoenzymes of sperm cells and tests. *Biochem. J.* **1969**, *111*, 207-218.

[36] Kaplan, N.O.; Ciotti, M.M.; Hamolsky, M.; Bieber, R.E. Molecular heterogeneity and evolution of enzymes. *Science* **1960**, *131*, 392-397.

[37] Read, J.A.; Winter, V.J.; Eszes, C.M.; Sessions, R.B.; Brady, R.L. Structural basis for altered activity of M- and H-isozyme forms of human lactate dehydrogenase. *Proteins* **2001**, *43*, 175-185.

[38] Kolappan, S.; Shen, D.L.; Mosi, R.; Sun, J.; McEachern, E.J.; Vocadlo, D.J.; Craig, L. Structures of lactate dehydrogenase A (LDH A) in apo, ternary and inhibitor-bound forms. *Acta Crystallographica Section D* **2015**, *71*, 185-195.

[39] Dempster, S.; Harper, S.; Moses, J.E.; Dreveny, I. Structural characterization of the apo form and NADH binary complex of human lactate dehydrogenase. *Acta Crystallogr. D Biol. Crystallogr.* **2014**, *70*, 1484-1490.

[40] Friberg, A.; Rehwinkel, H.; Nguyen, D.; Pütter, V.; Quanz, M.; Weiske, J.; Eberspächer, U.; Heisler, I.; Langer, G. Structural Evidence for Isoform-Selective Allosteric Inhibition of Lactate Dehydrogenase A. *ACS Omega* **2020**, *5* (22), 13034-13041.

[41] Poli, G.; Granchi, C.; Aissaoui, M.; Minutolo, F.; Tuccinardi, T., Three-Dimensional Analysis of the Interactions between *h*LDH 5 and Its Inhibitors. *Molecules* **2017**, *22* (12), 2217.

[42] Shu, Y.; Yue, J.; Li, Y.; Yin, Y.; Wang, J.; Li, T.; He, X.; Liang, S.; Zhang, G.; Liu, Z.; Wang, Y., Development of human lactate dehydrogenase a inhibitors: high-throughput screening, molecular dynamics simulation and enzyme activity assay. *J. Comput. Aided Mol. Des.* **2024**, *38* (1), 28.

[43] Adams, M.J.; McPherson, A.; Rossmann, M.G.; Schevitz, R.W.; Wonacott, A.J., The structure of the nicotinamide-adenine dinucleotide coenzyme when bound to lactate dehydrogenase. *J. Mol. Biol.* **1970**, *51*, 31-38.

[45] Tang, Y.; Gu, S.; Zhu, L.; Wu, Y.; Zhang, W.; Zhao, C. LDH A: The Obstacle to T cell responses against tumor. *Front. Oncol.* **2022**, *12*.

[46] Qiu, L.; Gulotta, M.; Callender, R. Lactate Dehydrogenase Undergoes a Substantial Structural Change to Bind its Substrate. *Biophys. J.* **2007**, *93*, 1677-1686.

[47] Warburg, O. On the origin of cancer cells. *Science* **1956**, *123*, 309-314.

[48] Ippolito, L.; Morandi, A.; Giannoni, E.; Chiarugi, P. Lactate: A metabolic driver in the tumor landscape. *Trends Biochem. Sci.* **2019**, *44*, 153-166.

[49] de la Cruz-Lopez, K.G.; Castro-Munoz, L.J.; Reyes-Hernandez, D.O.; Garcia-Carranca, A.; Manzo-Merino, J. Lactate in the regulation of tumor microenvironment and therapeutic approaches. *Front. Oncol.* **2019**, *9*, 1143.

[50] Kanno, T.; Sudo, K.; Maekawa, M.; Nishimura, Y.; Ukita, M.; Fukutake, K. Lactate-dehydrogenase M-subunit deficiency - a new type of hereditary exertional myopathy. *Clin. Chim. Acta* **1988**, *173*, 89-98.

[51] Celine, A.; Claudio, S. Converting a Peptide into a Drug: Strategies to Improve Stability and Bioavailability. *Curr. Med. Chem.* **2002**, *9*, 963-978.

[52] Gentilucci, L.; Tolomelli, A.; Squassabia, F. Peptides and Peptidomimetics in Medicine, Surgery and Biotechnology. *Curr. Med. Chem.* **2006**, *13*, 2449-2466.

[53] Gentilucci, L.; De Marco, R.; Cerisoli, L. Chemical modifications designed to improve peptide stability: incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization. *Curr. Pharm. Des.* **2010**, *16*, 3185-3203.

[54] Li Petri, G.; Di Martino, S.; De Rosa, M. Peptidomimetics: An Overview of Recent Medicinal Chemistry Efforts toward the Discovery of Novel Small Molecule Inhibitors. *J. Med. Chem.* **2022**, *65*, 7438-7475.

[55] Mabonga, L.; Kappo, A.P. Peptidomimetics: A Synthetic Tool for Inhibiting Protein-Protein Interactions in

Cancer. *Int. J. Pept. Res. Ther.* **2020**, *26*, 225-241.

[56] Kessler, H.; Gratias, R.; Hessler, G.; Gurrath, M.; Müller, G. Conformation of cyclic peptides. Principle concepts and the design of selectivity and superactivity in bioactive sequences by 'spatial screening'. **1996**, *68* (6), 1201-1205.

[57] Haubner, R.; Finsinger, D.; Kessler, H. Stereoisomeric Peptide Libraries and Peptidomimetics for Designing Selective Inhibitors of the $\alpha_v\beta_3$ Integrin for a New Cancer Therapy. *Angew. Chem., Int. Ed.* **1997**, *36*, 1374-1389.

[58] Haubner, R.; Gratias, R.; Diefenbach, B.; Goodman, S.L.; Jonczyk, A.; Kessler, H. Structural and functional aspects of RGD-containing cyclic pentapeptides as highly potent and selective integrin $\alpha_v\beta_3$ antagonists. *J. Am. Chem. Soc.* **1996**, *118*, 7461-7472.

[59] Reardon, D.A.; Neyns, B.; Weller, M.; Tonn, J.C.; Nabors, L.B.; Stupp, R. Cilengitide: an RGD pentapeptide $\alpha_v\beta_3$ and $\alpha_v\beta_5$ integrin inhibitor in development for glioblastoma and other malignancies. *Future Oncol.* **2011**, *7* (3), 339-54.

[60] Dechantsreiter, M.A.; Planker, E.; Mathä, B.; Lohof, E.; Hölzemann, G.; Jonczyk, A.; Goodman, S.L.; Kessler, H. *N*-Methylated cyclic RGD peptides as highly active and selective $\alpha_v\beta_3$ integrin antagonists. *J. Med. Chem.* **1999**, *42*, 3033-3040.

[61] Kapp, T.G.; Rechenmacher, F.; Neubauer, S.; Maltsev, O.V.; Cavalcanti-Adam, E.A.; Zarka, R.; Reuning, U.; Notni, J.; Wester, H.-J.; Mas-Moruno, C.; Spatz, J.; Geiger, B.; Kessler, H. A Comprehensive Evaluation of the Activity and Selectivity Profile of Ligands for RGD-binding Integrins. *Sci. Rep.* **2017**, *7*, 39805.

[62] Pfaff, M.; Tangemann, K.; Müller, B.; Gurrath, M.; Müller, G.; Kessler, H.; Timpl, R.; Engel, J. Selective recognition of cyclic RGD peptides of NMR defined conformation by $\alpha_{IIb}\beta_3$, $\alpha_v\beta_3$, and $\alpha_5\beta_1$ integrins. *J. Biol. Chem.* **1994**, *269* (32), 20233-8.

[63] Liu, Y.; Mei, L.; Yu, Q.; Zhang, Q.; Gao, H.; Zhang, Z.; He, Q. Integrin $\alpha_v\beta_3$ targeting activity study of different retro-inverso sequences of RGD and their potentiality in the designing of tumor targeting peptides. *Amino Acids* **2015**, *47* (12), 2533-9.

[64] Rani, R.; Kumar, V. Recent Update on Human Lactate Dehydrogenase Enzyme 5 (*h*LDH 5) Inhibitors: A Promising Approach for Cancer Chemotherapy. *J. Med. Chem.* **2016**, *59* (2), 487-496.

[65] Zhang, S.-L.; He, Y.; Tam, K.Y. Targeting cancer metabolism to develop human lactate dehydrogenase (*h*LDH) 5 inhibitors. *Drug Discovery Today* **2018**, *23* (7), 1407-1415.

[66] Thabault, L.; Brisson, L.; Brustenga, C.; Martinez Gache, S.A.; Prévost, J.R.C.; Kozlova, A.; Spillier, Q.; Liberelle, M.; Benyahia, Z.; Messens, J.; Copetti, T.; Sonveaux, P.; Frédérick, R. Interrogating the lactate dehydrogenase tetramerization site using (stapled) peptides. *J. Med. Chem.* **2020**, *63*, 4628-4643.

[67] Nadal-Bufi, F.; Mason, J. M.; Chan, L.Y.; Craik, D.J.; Kaas, Q.; Troeira Henriques, S. Designed β-Hairpins Inhibit LDH 5 Oligomerization and Enzymatic Activity. *J. Med. Chem.* **2021**, *64* (7), 3767-3779.

[68] Nadal-Bufí, F.; Chan, L.Y.; Mohammad, H.H.; Mason, J.M.; Salomon, C.; Lai, A.; Thompson, E.W.; Craik, D.J.; Kaas, Q.; Henriques, S.T. Peptide-based LDH 5 inhibitors enter cancer cells and impair proliferation. *Cell. Mol. Life Sci.* **2022**, *79* (12), 606.

[69] Thabault, L.; Liberelle, M.; Koruza, K.; Yildiz, E.; Joudiou, N.; Messens, J.; Brisson, L.; Wouters, J.; Sonveaux, P.; Frédérick, R. Discovery of a novel lactate dehydrogenase tetramerization domain using epitope mapping and peptides. *J. Biol. Chem.* **2021**, *296*, 10422.

[70] Schoenenberger, G.A.; Wacker, W.E.C. Peptide Inhibitors of Lactic Dehydrogenase (LDH). II. Isolation and Characterization of Peptides I and II*. *Biochemistry* **1966**, *5* (4), 1375-1379.

[71] Metrano, A.J.; Chinn, A.J.; Shugrue, C.R.; Stone, E.A.; Kim, B.; Miller, S.J. Asymmetric Catalysis Mediated by Synthetic Peptides, Version 2.0: Expansion of Scope and Mechanisms. *Chem. Rev.* **2020**, *120*, 11479-11615.

[72] Khettar, I.; Araszczuk, A.M.; Schettini, R. Peptidomimetic-Based Asymmetric Catalysts. *Catalysts **2023**, 13,* 244.

[73] List, B.; Lerner, R.A.; Barbas III, C.F. Proline-catalyzed direct asymmetric Aldol reactions. *J. Am. Chem. Soc.* **2000**, *122*, 2395-2396.

[74] Carvalho, S.; Peralta Reis, D.Q.; Pereira, S.V.; Kalafatovic, D.; Pina, A.S. Catalytic Peptides: the Challenge between Simplicity and Functionality. *Isr. J. Chem.* **2022**, *62*, e202200029.

[75] Akiyama, T.; Itoh, J.; Yokota, K.; Fuchibe, K. Enantioselective Mannich-type reaction catalyzed by a chiral Brønsted acid. *Angew. Chem., Int. Ed.* **2004**, *43*, 1566-1568,

[76] Uraguchi, D.; Terada, M. Chiral Brønsted acid-catalyzed direct Mannich reactions via electrophilic activation. *J. Am. Chem. Soc.* **2004**, *126*, 5356-5357,

[77] Shugrue, C.R.; Miller, S.J. Phosphothreonine as a Catalytic Residue in Peptide-Mediated Asymmetric Transfer Hydrogenations of 8-Aminoquinolines. *Angew. Chem., Int. Ed.* **2015**, *54*, 11173-11176.

Chapter II. Conjecturing about Agonistic and Antagonistic Behavior of $\alpha_4\beta_1$ Integrin Ligands: Lessons from RGD-binding Integrins

Albeit some general information about integrins structures and functions is briefly presented in the Introduction, a systematic study of the $\alpha_4\beta_1$ integrin in this chapter is clearly not redundant, particularly given that most of the research activities conducted in the thesis are based on this receptor and many potential drug molecules in preclinical and clinical research are associated with α_4 integrins. Furthermore, the 3D structure of this receptor has not yet been reported, so structure information inspired from other well-established integrins are of particularly interest for guiding the synthesis of novel, potent, and selective $\alpha_4\beta_1$ integrin ligands.

2.1 Introduction

Integrins are heterodimeric cell-surface receptors that regulate cell-cell adhesion and cellular functions through bidirectional signaling.[1,2] On the other hand, anomalous trafficking of integrins is also implicated in severe pathologies such as cancer, thrombosis, inflammation, allergies, and multiple sclerosis, making them attractive drug targets. However, despite of promising preclinical data, several anti-integrin drugs failed in late-stage clinical trials for chronic indications, with paradoxical side effects.[3] One possible reason is that, at low concentration, ligands proposed as antagonists may also act as partial agonists.[4-9] This can induce the high-affinity, extended-open conformation of integrins, triggering outside-in signaling and paradoxically leading to adhesion and unfavorable outcomes in patients.[10] Hence, the comprehension of the specific structural features for ligands' agonism or antagonism is currently of the utmost interest. For $\alpha_4\beta_1$ integrin, the situation is particularly obscure because nor the crystallographic or the cryo-EM structures are known. In addition, very few potent and selective agonists are available for investigating the mechanism at the basis of receptor activation.

In this account, we discuss the physiological role of $\alpha_4\beta_1$ integrin, the related pathologies, and we review the reported antagonists and the few agonists. Finally, we speculate on plausible models to explain agonism versus antagonism by comparison with RGD-binding integrins and by analysis of computational simulations performed with homology or hybrid receptors.

2.2 Biology and Functions of $\alpha_4\beta_1$ Integrin

In case of tissue injury or infection, the circulating leukocytes are triggered, and migrate to the affected site with the principal aim of eliminating the inflammatory origin and contributing to tissue repair. The movement of leukocytes from the blood across the endothelium, is a well-organized cascade process (cf. Figure 1.5).[11] Initially, selectins, which are expressed on both leukocytes and endothelial cells, interact with their ligands, e.g., P-selectin glycoprotein ligand-1 (PSGL-1), allowing leukocytes to roll onto the endothelial surface. Subsequently, chemokines released locally activate integrins via their G-protein coupled receptors (GPCR). Integrins gain affinity for their CAM ligands, and the firm interaction between them brings to leukocyte arrest. Finally, leukocytes extravasate to the site of infection, and this movement is mediated by specific adhesion molecules (JAM, CD31 (also known as platelet endothelial cell adhesion molecule, PECAM-1)).[12]

The leukocyte subfamily of integrins comprises seven members: four β_2 integrins ($\alpha_D\beta_2$, $\alpha_L\beta_2$, $\alpha_M\beta_2$ and $\alpha_x\beta_2$), two α_4 ($\alpha_4\beta_1$ and $\alpha_4\beta_7$), and $\alpha_E\beta_7$.[13] Their expression varies among the subpopulations of leukocytes. The integrins of the β_2 family, such as the $\alpha_L\beta_2$ (or leukocyte function-associated antigen-1, LFA-1) and $\alpha_M\beta_2$ (Mac-1), bind to their ligands ICAM-1 (CD54) and ICAM-2 (CD102), respectively.[14]

$\alpha_4\beta_1$ Integrin, also known as very late antigen-4 (VLA-4) or CD49d/CD29, is expressed on most leukocytes, and is involved in their homing, trafficking, differentiation, activation, and survival. This integrin recognizes the Leu-Asp-Val (LDV) peptide in FN, the Leu-Asp-Thr (LDT) sequence in the mucosal addressing cell adhesion molecule-1 (MAdCAM-1), and Ile-Asp-Ser (IDS) in VCAM-1 (Vascular cell adhesion molecule 1). The α_4 subunit is also found together with β_7 subunit in the $\alpha_4\beta_7$ dimer. The natural ligand of $\alpha_4\beta_7$ integrin is MAdCAM-1, whose peptide recognition motif is Leu-Asp-Thr (LDT).[15]

Blocking $\alpha_4\beta_1$ integrin could represent an opportunity for the treatment of diverse disorders,[16] because it is involved in the development and sustainment of inflammation, in inflammation-related diseases, and in mobilization or retention of stem cells. This receptor is also implicated in T cell transit through the blood-brain barrier (BBB) in autoimmune encephalitis. In multiple sclerosis (MS), migration of autoreactive T lymphocytes into the central nervous system (CNS) is governed by the interaction between $\alpha_4\beta_1$ integrin and VCAM-1. In allergic conjunctivitis, $\alpha_4\beta_1$ integrin mediates persisting infiltration of neutrophils, eosinophils, and T lymphocytes in the conjunctiva. In asthma and sarcoidosis, this receptor participates in the accumulation of lymphocyte in the lung. Furthermore, $\alpha_4\beta_1$ integrin is expressed in different varieties of cancer, i.e. multiple myeloma, ovarian cancer, and pancreatic cancer, and recent studies have described its potential role in cancer development and in the formation of metastasis.[17]

Not surprisingly, $\alpha_4\beta_1$ integrin has been targeted for the pharmacological treatment of autoimmune diseases such as MS,[18] or inflammatory pathologies. The humanized anti-$\alpha_4\beta_1$ monoclonal antibody (mAb) Natalizumab, blocks in the CNS the interaction of the integrin with adhesion molecules of endothelial cells, and the lymphocyte migration through the BBB, thus effectively preventing formation of new lesions and relapses in MS.[19] However, potential adverse effects of mAb may be severe, including secondary autoimmune diseases or malignancies, and can necessitate treatment discontinuation. Indeed, Natalizumab was withdrawn from the market in 2005 due to its association with the development of progressive multifocal leukoencephalopathy (PML) in patients with MS,[20] leading to decreased enthusiasm for α_4 antagonism. Interestingly, it was readmitted the following year, with a black box warning regarding risk of PML, and with required adherence to a special restricted distribution program.[21] The successful regulatory resubmission of Natalizumab has since revitalized interest in developing α_4 antagonist therapeutics.

2.3 $\alpha_4\beta_1$ Integrin Antagonists

As an alternative to antibodies, small molecule antagonists can be utilized to interfere with integrin-ligand interactions,[16] and they generally offer advantages such as natural non-immunogenicity, lower molecular mass, and a more cost-effective synthesis and development process. The small-molecule antagonists of $\alpha_4\beta_1$ integrin reported in the literature can be grouped into a few classes.[22]

1. *N*-acylphenylalanine derivatives (Figure 2.1 A), such as TBC3486, and Firategrast, a dual antagonist of $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrins, was designed to reduce trafficking of lymphocytes into the CNS for the treatment of multiple sclerosis. While high doses of Firategrast were significantly effective in reducing lesions, low doses proved to be harmful.[6] As a result, it is no longer in pharmaceutical development.

Surprisingly, Carotegrast-methyl (Figure 2.1 A), the first orally available antagonist of α_4 integrin, was launched in May 2022 by EA Pharma Co., Ltd. and Kissei Pharmaceutical Co., Ltd. It exerts an anti-inflammatory effect in vivo by inhibiting both $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrins expressed on lymphocytes. The efficacy and safety of Carotegrast-methyl were confirmed in patients with moderate active ulcerative colitis.[23]

2. Peptide ligands that reproduce the integrin-binding sequences. As mentioned before, FN binds $\alpha_4\beta_1$ integrin through the tripeptide sequence Leu-Asp-Val (LDV), while the integrin-binding sequence in VCAM-1 is the tripeptide Ile-Asp-Ser (IDS).

The LDV-based peptide BIO1211 (Figure 2.1)[24] demonstrated a potent antagonist effect against $\alpha_4\beta_1$ integrin. It effectively suppressed antigen-induced airway hyper-responsiveness in allergic sheep.[25] The high affinity is attributed to the well-known α_4-targeting diphenylurea pharmacophore MPUPA. However, the stability of BIO1211 is rather scarce, as tested in plasma, heparinized blood, and in homogenates of rat liver, lungs, and intestines,[26] and it undergoes rapid clearance in vivo.[27]

3. Peptidomimetics inspired from the IDS or LDV binding sequences. The stability of the peptide sequence can be improved by adopting a peptidomimetic strategy.[28] Selected examples are depicted in Figure 2.1. For instance, constraining the essential binding sequences found in VCAM-1 (IDS) and fibronectin (LDV) have afforded a variety of potent cyclic VLA-4 antagonists, such as the Tyr-cyclopeptide **1** (Figure 2.1).[29]

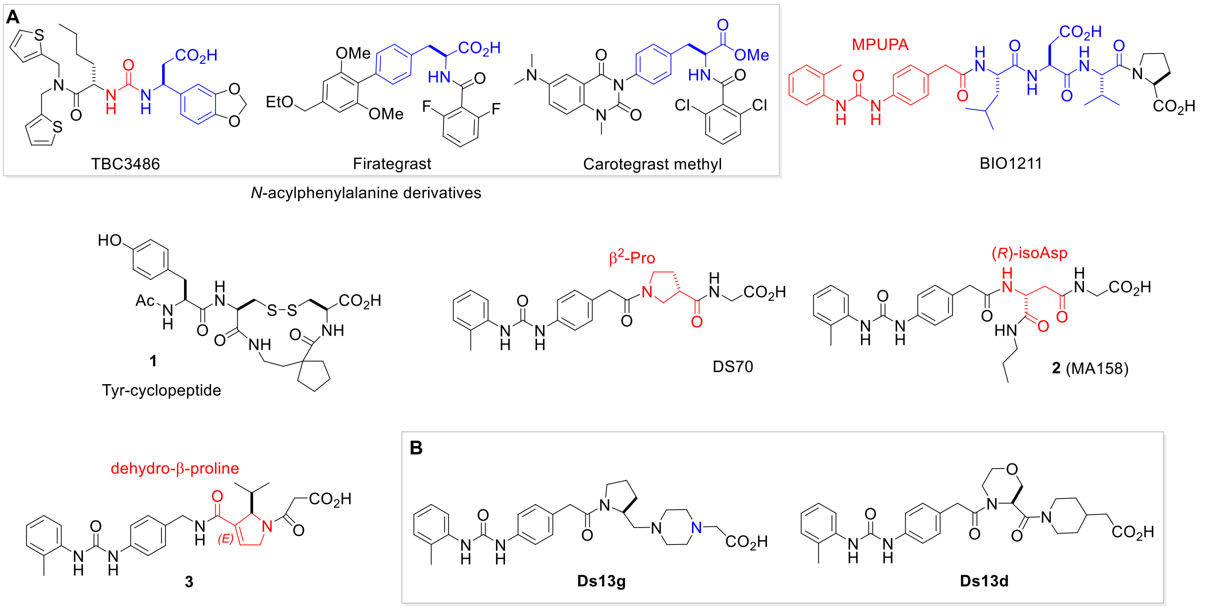


Figure 2.1 Structures of $\alpha_4\beta_1$ integrin antagonists: (A) *N*-acylphenylalanine derivatives TBC3486, Firategrast, Carotegrast-methyl; LDV peptide BIO1211; Tyr-cyclic peptide **1**; peptidomimetics containing a β-residue: DS-70, **2**, and **3**; (B) peptidomimetics Ds13g (pure antagonist) and Ds13d (partial agonist).

To cite another widely utilized approach, the introduction of β-amino acids allowed to increase the enzymatic stability of peptide integrin ligands.[30-32] The minimalist MPUPA-β-Pro-Gly DS-70 (Figure 2.1) was found to be a potent α_4 integrin antagonist with noteworthy stability in mouse serum and significant efficacy in an animal model of allergic conjunctivitis.[33] The sequence **2** containing a *(R)*-isoAsp β^3-core showed low nanomolar affinity (Figure 2.1), being able to prevent the adhesion of FN to Jurkat E6.1 cells, and to prevent FN-induced $\alpha_4\beta_1$ integrin-dependent activation of ERK1/2, AKT, and JNK.[34] Compound **3**, a retro-analogue of BIO1211 containing a dehydro-β-proline ring (Figure 2.1), was an enzymatically stable, potent inhibitor of $\alpha_4\beta_1$/VCAM-1 interaction, with IC_{50} in the nanomolar range.[35] Figure 2.1 also reports the commercially available peptidomimetic Ds13g and the correlated Ds13d (Figure 2.1 B),[36] whose features as integrin ligands are discussed in the next section.

Other orthosteric peptidomimetic $\alpha_4\beta_1$ antagonists have garnered interest as for the treatment of dry eye disease,[37] and dry age-related macular degeneration (AMD).[38] For a comprehensive review of the diverse ligands of $\alpha_4\beta_1$ integrin, together with the affinity values reported by the specific biological assay, see Reference[16].

2.3.1 $\alpha_4\beta_1$ integrin ligands in diagnostics and biomaterials

In addition to the therapeutic applications, the utilization of selective integrin ligands can be also harnessed for diagnostic purposes[39] and for biofunctionalization of diverse materials.[40] For instance, the radionuclide peptidomimetic ^{64}Cu-CB-TE1A1P-PEG4-LLP2A (^{64}Cu-LLP2A) targeting VLA-4 was proposed as a positron emission tomography (PET) imaging biomarker of vaso-occlusive episodes.[41]

Moreover, selective α_4 integrin peptidomimetic ligands have been grafted onto monolayers of fluorescent nanoparticles and yielded cell-adhesive devices for detection and quantification of leucocytes expressing active integrins, which can be utilized for monitoring the severity and evolution of diseases such as asthma.[42] Additionally, a nanostructured surface coated with a α/β hybrid peptide derived from compound **2** (Figure 2.2 A) was able to replicate the high-density multivalency binding between integrin clusters and VCAM-1, displaying notable selectivity for $\alpha_4\beta_1$ integrin-expressing Jurkat cells.[43]

A biomaterial was prepared by direct coating expanded polytetrafluoroethylene (ePTFE) with fibronectin-derived LDV peptide.[44] For its flexibility, biostability, and non-adhesiveness, this polymer is widely utilized in clinical applications, such as in the manufacture of blood-contacting implantable devices.

The ligands LLP2A and LXW7 of integrins $\alpha_4\beta_1$ and $\alpha_v\beta_3$, respectively, have been utilized to modify collagen-based materials. The peptide-modified biomaterial (Figure 2.2 B) improved the adhesion of mesenchymal stem cells (MSCs), osteoblasts, human endothelial progenitor Cells (EPCs), and endothelial cells (ECs). In an adult rat model of calvarial bone defect, the peptide-modified biomaterial increased bone formation and vascularization by synergistically regulating endogenous cells with osteogenic and angiogenic potentials. Similarly, in a fetal sheep model of spinal bone defect, the LLP2A/LXW7-biomaterial promoted bone formation and vascularization, without adverse effects.[45]

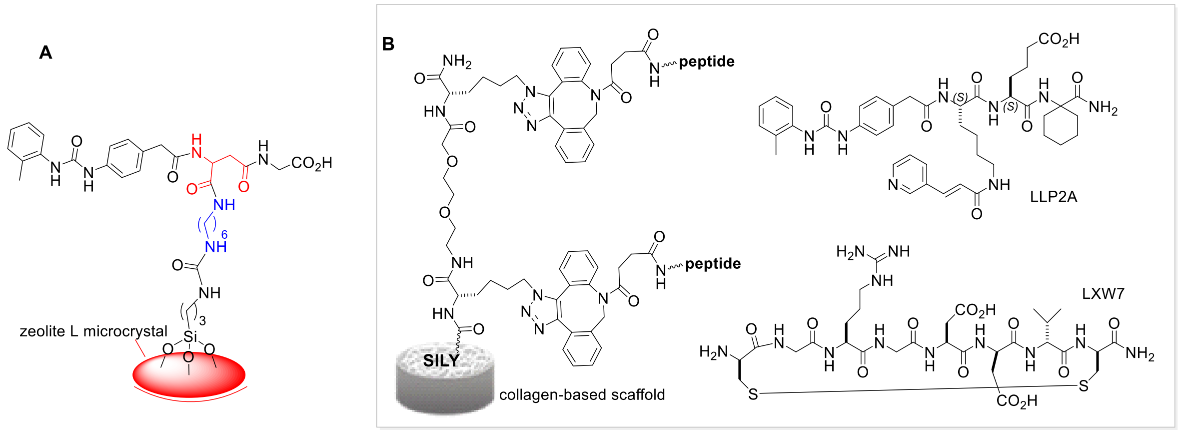


Figure 2.2 (A) Fluorescent zeolite L microcrystal biofunctionalized with an integrin ligand derived from peptidomimetic **2**. (B) Collagen-based scaffold functionalized with integrin ligand peptides LLP2A or LXW7; SILY (RRANAALKAGELYKSILY-NH_2) is a high-affinity collagen-binding peptide.

2.4 Small-molecule Agonists of $\alpha_4\beta_1$ Integrin

Very few potent and selective $\alpha_4\beta_1$ integrin agonists are available currently (Figure 2.3).[46] The urea THI0019 (Figure 2.3) was obtained via two structural modifications of the selective antagonist TBC3486 (Figure 2.1). THI0019 was found to be an agonist, being able to promote cell retention and engraftment.[47] 7HP349, an analog of lead compound THI0019, is a first in class allosteric agonist of integrins $\alpha_4\beta_1$ and $\alpha_L\beta_2$ which promotes both T-cell extravascular trafficking and antigen priming (Figure 2.3).[48] The small ureas **4** and **5**, featuring as β-lactam derivatives, are $\alpha_4\beta_1$ integrin ligands which showed agonistic behavior.[49,50] Recently, the cyclopeptide **6** was designed by connecting the LDV sequence to a 4-amino-L-proline (Amp) scaffold equipped with MPUPA moiety. Unexpectedly, this compound was able to increase the adhesion of $\alpha_4\beta_1$ integrin-expressing cells (Figure 2.3).[51] Their features as an agonistic integrin ligand are discussed in the next sections.

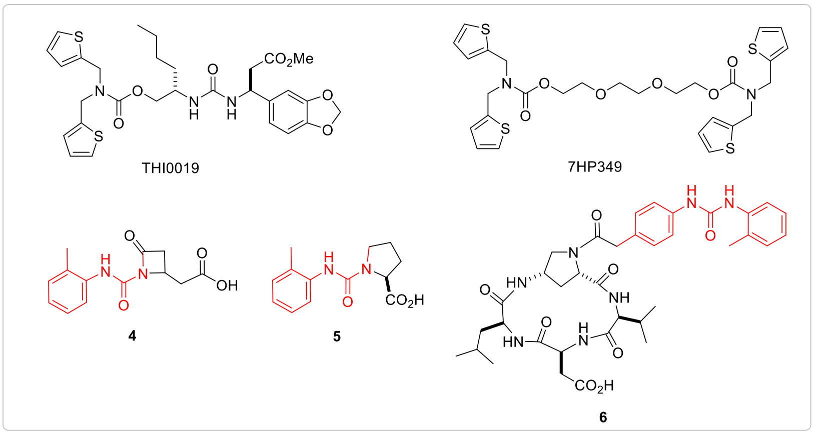


Figure 2.3 Chemical structures of integrin agonists: THI0019; 7HP349; small ureas **4** and **5**; retro-sequence cyclopeptide c[Amp (MPUPA)VDL] (**6**).

Concerning other leukocyte integrins, small-molecule agonists of integrin $\alpha_M\beta_2$ were described by Faridi et al.[52-54] while the first small molecule agonist of integrin $\alpha_L\beta_2$ was discovered by Yang et al.[55]

2.5 Therapeutic Opportunities of $\alpha_4\beta_1$ Integrin Agonists

Agonists of $\alpha_4\beta_1$ integrin could present alternative therapeutic opportunities. The activation of $\alpha_4\beta_1$ integrin might represent an innovative strategy to disturb the migration of leukocytes. After adhesion, leukocyte detachment is required to consent their rolling on the endothelial surface. In this perspective, agonists can preclude the detachment of adherent cells by improving the interaction with integrins.[56]

Activation of the integrins on progenitor cells may be a feasible approach to increase the efficacy of stem cell-based therapies by strengthening cell adhesion and engraftment. Agonists can be useful in the treatment of osteoporosis and for the promotion of bone growth due to their specificity for the $\alpha_4\beta_1$ integrin on mesenchymal stem cells and for bone surface.[57]

The activation of $\alpha_4\beta_1$ integrin could represent a strategy against other diseases. The small molecule $\alpha_4\beta_1$ integrin agonist THI0019 (Figure 2.3) favored cell retention and engraftment in stem cell-based therapies.[47] $\alpha_4\beta_1$ integrin was presumed to play a tumor-protective role in an animal model of colon adenocarcinoma. Indeed, $\alpha_4\beta_1$ depletion led to rapid tumor growth, pointing at small molecule agonists to influence $\alpha_4\beta_1$ expression level in cancer.[58] Furthermore, both $\alpha_4\beta_1$ and $\alpha_L\beta_2$ seem to be involved in the localization of cancer-specific CD8+ effector T cells to the local environment of tumor. Thus, activation of both $\alpha_4\beta_1$ and $\alpha_L\beta_2$ integrin by means of the small-molecule agonist 7HP349 (Figure 2.3) augmented the accumulation of anticancer T cells to the tumor, improving their efficacy. This effect was boosted by concomitant administration of an anti-cytotoxic T-lymphocyte-associated protein 4 agent (anti-CTLA-4).[59] The same small molecule agonist of $\alpha_4\beta_1$ and $\alpha_L\beta_2$ integrins reported above was also tested in combination of a DNA vaccine in a model of Chagas disease. The agonist enhanced both preventive and therapeutic vaccine efficacy, suggesting to adopt integrin agonists to boost immune response mediated by T cells to different types of vaccines.[60] Moreover, 7HP349 has a very safe acute and chronic preclinical toxicity profile.

Pertaining to the other leukocytic integrins, recent evidence supports that α_M (CD11b) integrin is deeply involved in the modulation of proinflammatory signaling. Accordingly, α_M allosteric modulators stimulating the anti-inflammatory activities of α_M integrin, can be exploited in the cure of lupus nephritis, a serious illness of systemic lupus erythematosus, marked by penetration of leukocytes to the kidneys.[61] Lastly, the activation of $\alpha_M\beta_2$ integrin can reduce hypertrophy and mineralization of chondrocytes, therefore agonists of $\alpha_M\beta_2$ integrin have been proposed for the therapy of osteoarthritis since they could lead to reduced inflammatory response.[62]

The β-lactam derived agonist **4** discussed above (Figure 2.3) was utilized for coating biomaterials to stimulate cell adhesion and promoting tissue repair. For instance, Integrin agonist **4**-loaded Strontium-substituted hydroxyapatite (SrHAp) materials significantly enhanced human mesenchymal stem cell (*h*MSC) adhesion and viability, as well as differentiation towards osteoblastic lineage.[63] Merlo et al. further analyzed the effect of the agonist **4** on cell adhesion of equine adipose tissue and mesenchymal stem cells derived from Wharton's jelly, and studied their adhesion to the agonist incorporated in a poly *L*-lactic acid (PLLA) matrix.[64] Moreover, the same agonist was combined with PLLA nanofibers to consent a controlled release of this molecule, resulting in a medication especially appropriate for skin wounds.[65]

2.6 Ligand-integrin Interactions and Conformational Implications

As anticipated, investigation of $\alpha_4\beta_1$ agonists could provide information about the structural requirements for receptor activation. Up to now, there are over 100 solved structures concerning various heterodimers or α- or β-domains, alone or in complex with different ligands. The large majority of structural analyses performed by X-ray, cryogenic electron microscopy (cryo-EM), NMR, and computations, have been conducted for $\alpha_v\beta_3$ and $\alpha_{IIb}\beta_3$ integrins, and contributed to the comprehension of their activation and function. Hence, in the absence of any data for the $\alpha_4\beta_1$ integrin, some deductions can be extrapolated by comparison with RGD-binding integrins, and by analyzing the subunits α_4 and β_1 taken separately from $\alpha_4\beta_7$ and $\alpha_5\beta_1$ integrins.

In general, the α and β subunits consist of a large extracellular ectodomain, a single-span helical transmembrane (TM) domain, and a generally short cytosolic tail. The four domains of the α-chain ectodomain and the seven domains of the β-chain are shown in Figure 1.2 A (Chapter 1, Introduction).

As previously mentioned, in many integrins, including $\alpha_4\beta_1$, the ligand-binding pocket is in a groove at the interface between α and β subunits. A metal-ion-dependent-adhesion-site (MIDAS), is situated in the β-subunit in the absence of an αI domain. For other integrins, including the β_2, MIDAS resides in an extra αI domain in the α subunit. In RGD-binding integrins, the ligand-binding site comprises a pocket to host Arg located in the propeller domain of α subunit, and a pocket for Asp carboxylate containing the MIDAS, normally occupied by pro-adhesive Mg^{2+} ion in the βA domain. There are also two extra regulatory metal binding sites in the β-subunit, a ligand associated metal binding site (LIMBS or SyMBS), and an adjacent to MIDAS (ADMIDAS), each normally occupied by a Ca^{2+} ion (Figure 2.4).

In the following sections, structural features of the well-studied integrins $\alpha_{IIb}\beta_3$ and $\alpha_v\beta_3$ are discussed and compared with integrins $\alpha_5\beta_1$ and $\alpha_4\beta_7$. What stands out is that α_4 and β_1 subunits show marked differences and specificities as compared to other heterodimers.

2.6.1 $\alpha_{IIb}\beta_3$ integrin

For integrin $\alpha_{IIb}\beta_3$, there are almost forty entries in the Protein Data Bank. Despite of being the most investigated, there is still some debate about the precise full structure and orientation of the ectodomain relative to the plasma membrane,[66,67] due to the adoption of diverse experimental procedures, e.g. the use of detergents or averaging over heterogeneous conformational states.

By default, integrins on cell surface are in their resting state, so platelets are free to circulate in blood vessels. In the absence of ligands, this state predominates on cell surfaces (>98%).[68] Acting through their receptors, physiologic agonists stimulate the intracellular binding of talin and kindlin to the cytoplasmic tail of the β-subunit, producing structural rearrangements in the transmembrane (TM) domains that trigger a conformational switch of the ectodomain, generating the active state, competent for extracellular physiologic ligands ("inside-out" activation). Initiation of signals occurs when these ligands bind, and these signals are then transduced through the TM domains to the cytoplasmic tails. This process regulates cell adhesion and various other functions ("outside-in" signaling).

Figure 2.4 shows the detailed structures of βI domain metal coordination sites in unliganded-closed $\alpha_{IIb}\beta_3$ (3FCS, Figure 2.4 A)[69] and liganded-open conformation in the presence of peptide LGGAKQRGDV (2VDR, Figure 2.4 B),[70] respectively. It is clear that in the liganded-open state, the carboxylate of the Asp residue coordinates with Mg^{2+} by expelling a water molecule in the MIDAS. Details of the ligand binding sites for the complexes $\alpha_{IIb}\beta_3$/UR-2922 and $\alpha_{IIb}\beta_3$/tirofiban are also discussed in the next parts.

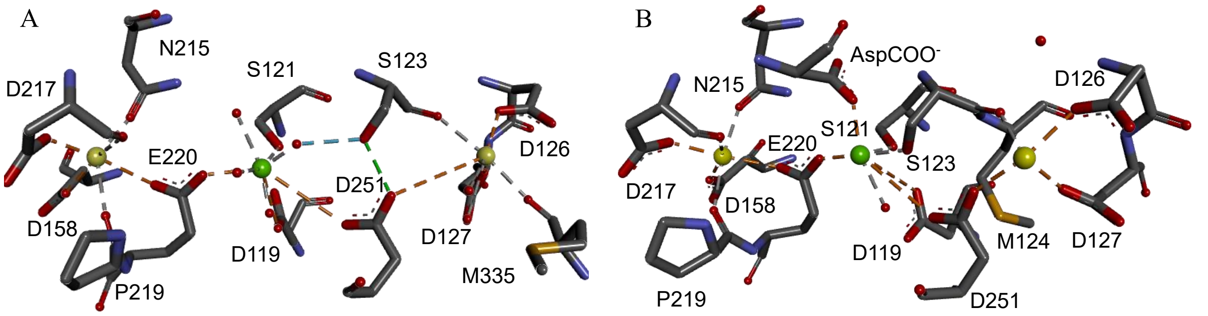


Figure 2.4 (A) Metal ion coordination sites in the βI domain in unliganded-closed $\alpha_{IIb}\beta_3$ (3FCS) and (B) liganded-open $\alpha_{IIb}\beta_3$ (2VDR). Ca^{2+} ions are gold and Mg^{2+} ions are green. Note that only the ligand Asp carboxyl is depicted.

Springer et al. observed eight distinct RGD-bound conformations of the $\alpha_{IIb}\beta_3$ integrin headpiece, by soaking crystals with diverse concentrations of RGD peptides and divalent cations: the closed βI domain conformation, six intermediate βI conformations, and finally the fully open βI with upright hybrid domain. During the extension, diverse regions of the βI domain experienced significant movements, starting with the β1-α1 backbone that hydrogen bonds to the Asp side chain of RGD, followed by ADMIDAS Ca^{2+}, α1 helix, α1' helix, β6-α7 loop, α7 helix, and hybrid domain.[71] In essence, ligand binding at the MIDAS in the βI domain produces a conformational reorganization, transmitted to the opposite end of the βI domain by α7 helix pistoning at the C-terminal connection to the hybrid domain, resulting in its extension.

Using cryo-electron microscopy, Hainein et al. determined 3D structures of full-length human integrin $\alpha_{IIb}\beta_3$ embedded in a lipid bilayer of nanodiscs, while bound to domains of the cytosolic regulator talin and to extracellular ligands. The authors observed four main conformational states in equilibrium, ranging from compact/bent to two partially extended intermediate conformers and finally to a fully extended state. The results indicate that extension of the ectodomain is possible without separating the legs or extending the hybrid domain, and that the ligand-binding pocket is not occluded in any conformations.[72]

Recently, full-length structure obtained by cryo-EM in native cell-membrane nanoparticles (cf. Figure 1.1 A), surprising features were a fully accessible ligand-binding site and the remarkable distance of the α_{IIb} TM domain from that of β_3. Another finding was that the FDA-approved anti-trombotic, $\alpha_{IIb}\beta_3$ inhibitor Eptifibatide, prescribed for myocardial infarction and acute coronary syndromes, induces pseudo-agonist, large conformational rearrangement in the full-length $\alpha_{IIb}\beta_3$, plausibly accounting for impaired hemostasis.[73]

These studies allow to depict a detailed model of the activation process. In ligand-free conditions, Ca^{2+} at ADMIDAS stabilizes the position of α1 helix at the N-terminus and of α7 helix at C-terminus of the βA domain in the closed conformation. Upon inside-out activation, the Ca^{2+}-mediated connection between the α1 and α7 helices is broken, allowing an RGD ligand to access into the binding site. Thereafter, the carboxylate of the RGD peptide makes a salt bridge with Mg^{2+} in MIDAS by displacing a water molecule from the coordination sphere of the cation, and by introducing the Arg side chain into the propeller (Figure 2.4).

Binding the ligand's carboxylate to MIDAS seems to be fundamental to integrin activation, while pulling by the α subunit may be not mandatory.[74] During receptor opening, the α1 helix-ADMIDAS portion undergo a large movement (3.9 Å) toward MIDAS, provoking the reorganization of the adjacent loops and the lowering of the α7 helix, pushing the hybrid domain positioned beneath (Figure 2.4). The swap of Ca^{2+} with Mn^{2+} at ADMIDAS also separates helices α1 and α7, enabling the switch of βA into the ligand-competent state, but the integrin does not attain the extended conformation.

2.6.2 $\alpha_v\beta_3$ integrin

The first crystal structure of the ectodomain of an αI-absent heterodimer in complex with a ligand was reported in 2002 for $\alpha_v\beta_3$/Cilengitide, c[RGDfMeV] (Ca^{2+}).[75] So far, almost twenty structures related to diverse domains of $\alpha_v\beta_3$ have been deposited in Protein Data Bank.

Arnaout et al described the first crystal structure of $\alpha_v\beta_3$ bound to a physiologic ligand, the 10th type III RGD domain of wild-type FN (wtFN10, 4MMX), or to a high-affinity mutant, the pure antagonist (*h*FN10, 4MMZ). Comparison of these structures revealed a central π-π interaction between Trp^{1496} in the RGD-containing loop of the antagonist *h*FN10 and Tyr^{122} (β_3) of the receptor, that blocks conformational changes and traps the integrin in an inactive conformation (Figure 2.5).

In the complex with wild-type FN (4MMX), Pro^{1497} is the closest residue to Tyr^{122}, but in this case the distance is significantly larger. Removing the Trp^{1496} or Tyr^{122} side chains or reorienting Trp^{1496} away from Tyr^{122} converted *h*FN10 into a partial agonist.[76] For *h*FN10, the hydroxyl oxygen of Tyr^{1446} of FN also coordinated the Mn^{2+} at ADMIDAS through a water molecule.

Apparently, the π-π interaction between Trp^{1496} and Tyr^{122} was crucial for blocking conformational changes induced by the binding of ligand, as shown in mutational studies. This hypothesis was substantiated by subsequent work, aimed at converting the partial $\alpha_v\beta_3$ agonist MK-429 into the pure antagonist TDI-4161. The structure of MK-429 was modified by introducing a large bicyclic benzo[d]thiazole, capable to inducing a π-π stacking with Tyr^{122}.[77]

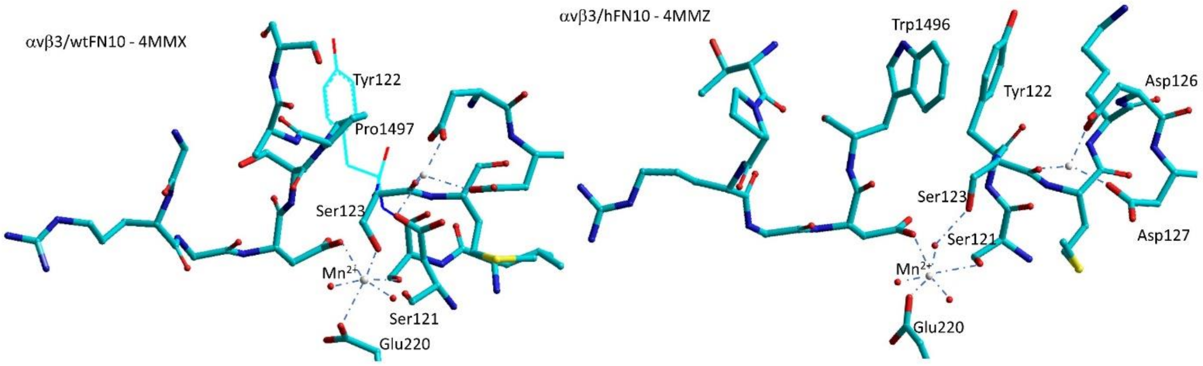


Figure 2.5 Simplified sketch of the binding site of wild-type FN/$\alpha_v\beta_3$ (4MMX) and of the binding site of $\alpha_v\beta_3$ integrin hosting the 10th type III domain of mutant FN (4MMZ), showing the stacking between Trp^{1496} of FN and Tyr^{122} of the β-subunit which seems to prevents the dislocation of β1-α1 loop and subsequent receptor extension.

2.6.3 $\alpha_4\beta_7$ integrin

As said for $\alpha_4\beta_1$, also $\alpha_4\beta_7$ integrin is involved in leukocytes rolling and adhesion, as well as in their firm adhesion on blood vessels. $\alpha_4\beta_7$ integrin mediates rolling adhesion in Ca^{2+}, while in Mg^{2+} or Mn^{2+}, leukocytes adhesion becomes firm.

Springer et al. have reported the structures of the headpiece of $\alpha_4\beta_7$ with the antibody Natalizumab (4IRZ),[78] or with the antibody Fab ACT-1 (3V4P), or with both the small molecule RO0505376 and Fab ACT-1 (3V4V).[79] The inspection of the crystal structure of Natalizumab (4IRZ) with $\alpha_4\beta_7$ revealed that Natalizumab binds to the β-propeller of α_4 subunit in a position which is close but alternative to the orthosteric ligand-binding site (Figure 2.6 A). In contrast, the ACT-1 antibody binds to an extended, disulfide-bonded loop in β_7 subunit, known as the specificity-determining loop (SDL), located in the proximity the ligand-binding site (3V4P, Figure 2.6 B). This difference explains why Natalizumab inhibits both $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrin subfamilies, while ACT-1 only selectively inhibits $\alpha_4\beta_7$. As it turns out in clinical results, Natalizumab is associated with the development of PML when used in patients with multiple sclerosis.[20,21] In contrast, Vedolizumab, the humanized form of the RM-Act-1 mAb, binds exclusively to $\alpha_4\beta_7$ integrin. This selective binding preserves its ability to bind with the ligand VCAM-1, allowing continued immune surveillance within the central nervous system and theoretically eliminating the risk of PML.[80]

RO0505376 is an antagonist of $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrin with IC_{50} values of 32 nM and 42 nM, respectively. The ternary complex formed by $\alpha_4\beta_7$/RO0505376/Fab ACT-1 is contained in 3V4V (Figure 2.6 C).[79] The shape of the binding cleft is clearly dissimilar from that of the RGD-binding integrins. The carboxylate of RO0505376 interacts with the MIDAS cation forming two NH backbone H-bonds (Tyr^{143} and Asn^{235} of β_7). The amide NH of the inhibitor forms an H-bond with backbone carbonyl of Asn^{235} with an orthogonal 2,6-dichlorophenyl moiety in the proximity of the specificity-determining loop. The pyridine ring forms π-π stacking with Phe^{214} and Tyr^{187} of α_4 and a strong 2.6 Å hydrogen bond with Ser^{238} of β_7.

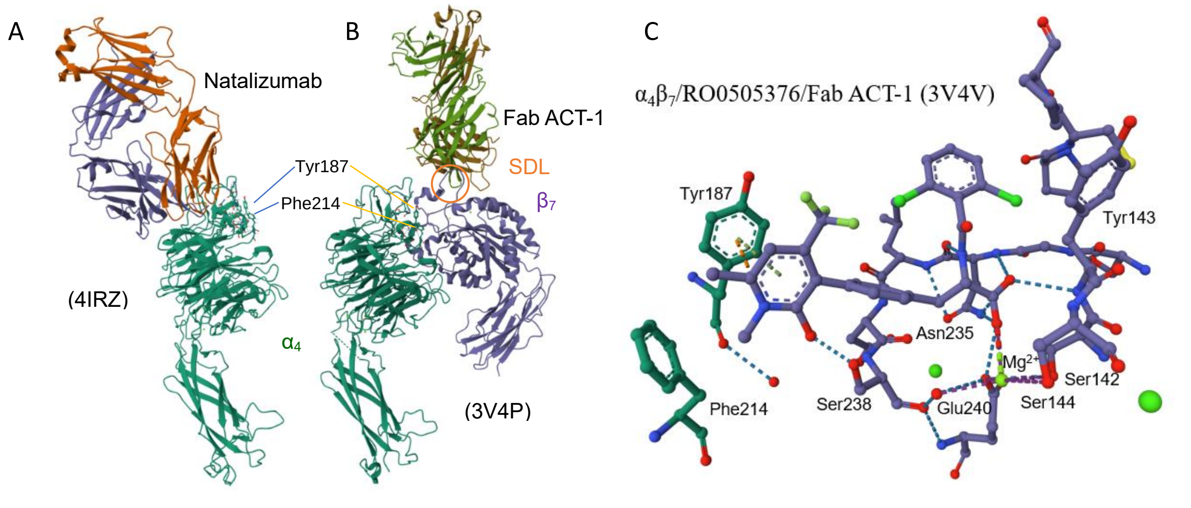


Figure 2.6 (A) Natalizumab Fab bound to α_4 headpiece (4IRZ). (B) Fab ACT-1 in complex with $\alpha_4\beta_7$ headpiece (3V4P). (C) Sketch of the binding site of RO0505376 in $\alpha_4\beta_7$integrin/Fab ACT-1 (3V4V).

2.6.4 $\alpha_5\beta_1$ integrin

Compared to the other RGD-binding integrins, the β_1 transmit quite different types of signals into cells. Integrin $\alpha_4\beta_1$ and RGD-binding $\alpha_V\beta_1$ and $\alpha_5\beta_1$ require markedly different tension thresholds to support cell spreading.[81] Nine X-ray crystal structures related to $\alpha_5\beta_1$ in PDB have been reported so far. The inspection of these structures underlines some structural distinctive features worthy of notice. The comparison between the β_1 and β_3 subunits shows that the former adopts an expand binding pocket, because the residues Arg^{214} and Arg^{216} in β_3 subunit are replaced with Gly^{217} and Leu^{219} in β_1. Unlike the resting structures of β_3 integrins, $\alpha_5\beta_1$ integrin exhibits a half-bent conformation.[82,83] The integrin headpiece of $\alpha_5\beta_1$ is more stable in the closed state than the ectodomain and is thus more difficult to activate.[84]

In 2012, Nagae et al. reported two structures of $\alpha_5\beta_1$ headpiece bound to the allosteric inhibitory antibody SG/19 Fab, i.e., a ligand-free form (3VI3), and a RGD complex (3VI4).[85] The latter was characterized by a conformational state intermediate between closed and open.

Subsequently, Springer at al. performed soaking experiments and obtained structures with linear or cyclic RGD peptides (4WK0, 4WK2, 4WK4, plus 4WJK with no peptide).[86] While the linear RGD peptide induced no movement in the β1-α1 loop (4WK0, Figure 2.7), in the absence of added Ca^{2+} the same ligand produced a measurable movement of the β1-α1 loop and α1-helix, while Ser^{134} displaced a water molecule at the MIDAS (4WK2), On the other hand, binding of a disulfide-cyclized RGD peptide with 20-fold higher affinity in Ca^{2+}, produced a noteworthy conformational change in the β_1-subunit βI domain, as well as the movement of Ser^{134} side chain into direct coordination with the MIDAS, a conformation that is intermediate between low affinity/closed and high affinity/open (4WK4, Figure 2.7).

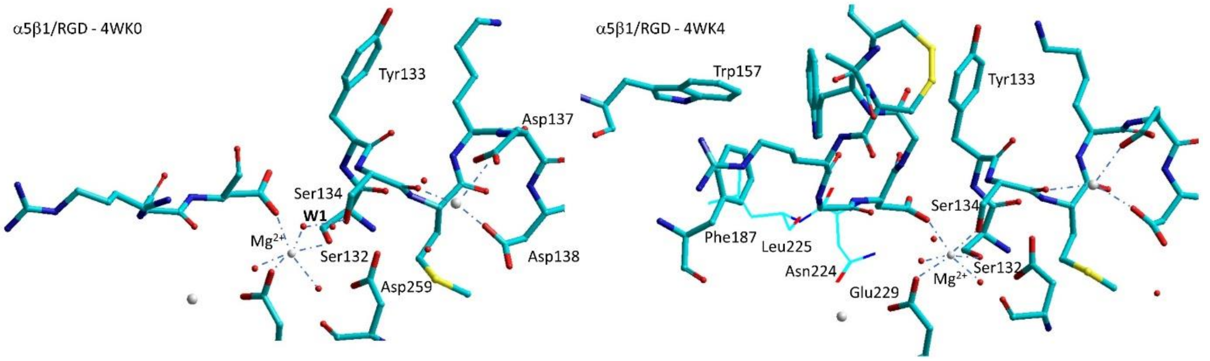


Figure 2.7 Sketches of the metal-ion-dependent adhesion sites for the complexes $\alpha_5\beta_1$ with a linear RGD peptide (4WK0), and with a RGD cyclopeptide (4WK4). The latter differs in particular for the repositioning of Ser^{134} to displace water 1 at MIDAS.

These results support the role of ADMIDAS as a negative regulatory site responsible for integrin inhibition by high concentration of Ca^{2+} and for activation by Mn^{2+}. It has been proposed that during the process of extension/activation of β_1 integrins, Ca^{2+} at the ADMIDAS site becomes highly instable, until it is removed from the site, whereas the cations of LIMBS and MIDAS maintain their positions.[86]

On resuming, ADMIDAS seems to be an important factor in the transmission of conformational changes between the βI and hybrid domains. While the MIDAS Mg^{2+} ion binds the Asp side chain of the RGD motif, the nearby Ca^{2+} ions in the ADMIDAS and SyMBS regulate ligand affinity. Studies with diverse integrins confirmed that low concentrations of Ca^{2+} (~50 μM) enhance binding and higher concentrations of Ca^{2+} (in the range 1-10 mM) inhibit binding. Replacement of the cations with Mn^{2+} has the net effect to substantially enhance integrin binding.

The roles of Mn^{2+} and ADMIDAS in regulating integrin affinity in integrin $\alpha_5\beta_1$ have been studied using antibody Fab fragments specific for the open conformation. As it turned out, Mn^{2+} seems to markedly increases both the population of the extended state and the intrinsic affinity of the $\alpha_5\beta_1$ extended state for the ligand. However, the extended state is populated only partially. That is, Mn^{2+} modifies the conformational ensembles only in part. The study also confirmed the role of ADMIDAS in enlarging the difference in affinity of the low-affinity, of the extended-closed, and the high-affinity extended-open states.[87]

In 2021, Mizuno et al. reported the cryo-EM structures of native human integrin $\alpha_5\beta_1$ in its open conformation in complex with FN and the stabilizing antibody TS2/16 (7NWL), and in its half-bent conformation (7NXD).[88] The complex of integrin $\alpha_5\beta_1$ with FN shows three simultaneous interactions critical for the stabilization of integrin opening, namely, (i) at the synergy site, (ii) at the RGD loop, and (iii) a newly identified binding site proximal to ADMIDAS, specifically centered at R^{1445} and Y^{1446} of FN and D^{137} of ADMIDAS in β_1 βI domain, inducing the translocation of helix α1 to secure integrin opening (Figure 2.8 A). This model explains the conundrum why the binding of RGD peptides was not sufficient for integrin opening in previous studies of $\alpha_5\beta_1$. For the FN receptor $\alpha_5\beta_1$, the extended, open conformation was observed only when both Mn^{2+} and FN were present.[86] That indicates that FN binding does not require prior integrin opening, but instead, the open conformation of $\alpha_5\beta_1$ becomes dominant as a result of the FN binding. Without Mn^{2+}, integrin cannot stably capture FN, as the interaction is transient and short-lived in vitro in the absence of Mn^{2+}. In the resting state, integrin $\alpha_5\beta_1$ uses a unique, incompletely bent conformation. The incomplete bending does not sterically hinder the access of FN to the integrin head.[88]

More recently, David Baker determined the structure of $\alpha_5\beta_1$ in complex with NeoNectin (9CKV, Figure 2.8 B), a de novo $\alpha_5\beta_1$-specific modulating miniprotein binders favoring an extended-open conformation.[89] According to their study, and in comparison to previous research, FN binds to both the α_5-β_1 interface and a synergy site on α_5, while NeoNectin only binds to the interface between α_5 and β_1, suggesting that the synergy site binding is not essential for the conformational switching of integrin $\alpha_5\beta_1$.

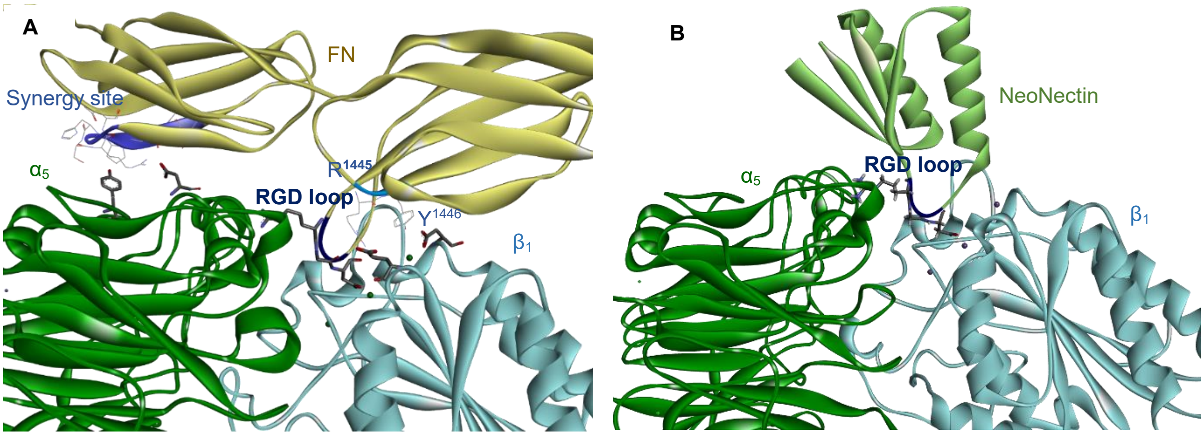


Figure 2.8 (A) Overview of the integrin $\alpha_5\beta_1$-FN interface, including RGD loop, FN synergy site, and a new site near the ADMIDAS centered on R^{1445} and Y^{1446} (7NWL). (B) Cryo-EM map of $\alpha_5\beta_1$ bound to NeoNectin (9CKV).

2.6.5 Agonism or antagonism and dynamics nature of MIDAS

Very recently, Springer et al. reconsidered the role of the water molecules present in MIDAS as discussed above, and proposed a simple but robust theory to explain why small-molecule inhibitors previously regarded as potential therapeutic entities failed in late-stage clinical trials for chronic indications. Plausibly, this may be caused by partial agonism at low concentration, namely the stabilization of the high-affinity, extended-open integrin conformation. The authors utilized diverse techniques, including size-exclusion chromatography in the absence or in the presence of ligands, to measure any increase of the hydrodynamic radius, relatable to extension of the whole structure.[10] In addition, the effects of the ligands on the conformation of the receptor were examined by means of MBC319.4 antibody, specific for the extended states of β_3 integrins.

As a result, failed small-molecule inhibitors in clinic trials of integrins $\alpha_{IIb}\beta_3$ appear to stabilize the open/high-affinity conformation. The reasons of this behavior have been investigated by analyzing the crystal structures of the $\alpha_{IIb}\beta_3$ headpiece in complex with a variety of ligands, such as the “opening” ligands Roxifiban, Lotrafiban, Tirofiban, and many others, in Mg^{2+} or Mn^{2+}, in comparison with “closing” ligands such as UR-2922, Gantofiban, etc. During headpiece extension, the key early step is the movement of the β1-α1 loop toward the α_{IIb} subunit. In the absence of a ligand, the sidechain oxygen of Ser^{123} in the MIDAS motif hydrogen bonds to water 1 (Figure 2.4 A). Whereas in the subsequent open states, movement of the β1-α1 loop repositions the Ser^{123} sidechain oxygen to take the place of water 1 (Figures 2.4 B and 2.9 B). Plausibly, the pure “closing” inhibitors might be longer and more rigid than RGD and block this movement.

An extensive analysis of the crystal structures of the $\alpha_{IIb}\beta_3$ headpiece in complex with inhibitors that either stabilized integrin headpiece opening, closure, or were conformationally neutral (in Mg^{2+} or Mn^{2+}), support a simple structural feature present in diverse $\alpha_{IIb}\beta_3$ antagonists that stabilizes integrins in their bent-closed conformation. Closing inhibitors contain a polar nitrogen that directly stabilizes water 1 (Figure 2.9 A). Other closing antagonists such as BMS4 indirectly stabilizes water 1 through another water 2 by means of the protonated piperazine nitrogen N1, which is closest to the carboxylic acid group of the ligand. Closing inhibitors have no potential for agonism since their binding does not induce integrin conformational change toward the open state. Additionally, and in contrast to both conformationally neutral RUC-4 and opening inhibitors, they stabilize the closed conformation by suppressing conformational change toward the active integrin state.

In contrast, with opening ligands, including Tirofiban, Eptifibatide, Lotrafiban, and EF-5154 in Mg^{2+} and Roxifiban in Mn^{2+}, the movement of β1-α1 loop brought the Tyr^{122} backbone into position to hydrogen bond to the drug carboxyl group, while Ser^{123} displaced water 1 to coordinate directly to the MIDAS metal ion (Figure 2.9 B).[10]

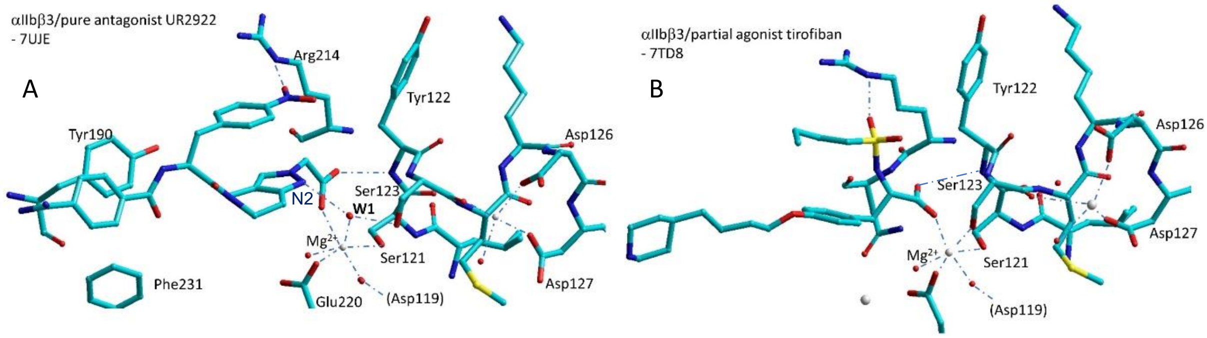


Figure 2.9 (A) The "closing" antagonist UR-2922 stabilizes water 1 between the MIDAS metal ion, the nitrogen N2 of the 1H-pyrazole ring, and Ser^{123} sidechain. (B) With the "opening" ligand tirofiban, β1-α1 loop movement brings Ser^{123} to displaced water 1 to coordinate directly to the MIDAS.

It was proposed that expulsion of this water 1 molecule is a requisite for transition to the open conformation also for α_4 integrins. The experiments were carried out for the dual $\alpha_4\beta_1$/$\alpha_4\beta_7$ integrin antagonist Firategrast, the commercially available peptidomimetics Ds13g and the correlated Ds13d (for the chemical sketches, see Figure 2.1).[36] Respect to the piperidine ring in Ds13d, Ds13g contains an extra basic nitrogen in the piperazine ring. While Firategrast and Ds13d stabilize the open integrin conformation, Ds13g stabilizes the closed conformation.

2.7 Simulations with Homology or Composite Models of $\alpha_4\beta_1$ Integrin

2.7.1 Simulations of $\alpha_4\beta_1$ integrin antagonists

As anticipated in the previous paragraph, the detailed structure of $\alpha_4\beta_1$ integrin is not yet accessible. An alternative to conduct pharmacodynamics investigations is to perform molecular docking simulations. To the purpose, either homology or composite models can be built. Initial attempts by the homology modelling approach required extensive optimization steps.

In 2002, Dixon et al. obtained a homology model of the β_1 subunit based on the I domain of the integrin CD11b (α_M) subunit and analyzed the binding of the LDV cyclopeptide and compound TBC3486 (Figure 2.1).[90] Subsequently, Macchiarulo et al. simulated the binding of the *N*-acylphenylalanine derivative TR-14035 using a homology model based on the crystal structure of $\alpha_v\beta_3$. The so-obtained complex closely reproduced the binding of RO0505376 in $\alpha_4\beta_7$ (Figure 2.6 C).[91] More recently, the deep-learning method AlphaFold was used to generate an optimized homology model of the integrin $\alpha_4\beta_1$.[92]

Frequently, simulations with $\alpha_4\beta_1$ models have been conducted for in-silico screening to select potential inhibitors of the same integrin[93,94] or other β_1 integrins.[95] For instance, molecular docking was used to assist in the design of a blocking polypeptide (antifibrotic 38-amino-acid polypeptide, AF38Pep) for specific inhibition of extra domain A-FN associations with the fibroblast-expressed α_4 integrins.[96]

The availability of the crystallographic structures of $\alpha_4\beta_7$ and $\alpha_5\beta_1$ integrins allowed for obtaining the structure of the heterodimer by merging the α_4 and the β_1 subunits. To create the chimeric receptor, the crystal structure of $\alpha_4\beta_7$ headpiece in complex with Fab ACT-1 and RO0505376 (3V4V)[79] is generally utilized. For $\alpha_5\beta_1$, the situation is more complicated. There are diverse high-resolution crystal structures available of headpiece domains with an RGD peptide (see above). The binding modes of the RGD peptides with respect to the β1-α1 loop are somewhat different, since they refer to closed/inactive or partially open conformations. In two out of four structures (4WK2 and 4WK4), Mg^{2+}-bound water is replaced by Ser^{134}.[86]

Besides, it must be stressed that the assembly of the separate subunits requires a precise optimization of the contacting regions. Finally, simulations of integrins remain particularly challenging because the typical molecular mechanics force fields employed to simulate ligand-receptor interactions fail in the calculation of the metal coordination sphere.

In summary, the receptor models as reported in the literature for performing molecular docking greatly vary, depending on the homology modelling or chimeric approach, of the specific β_1 crystal structure selected, the level of heterodimer optimization, and the computational method, which can range from specialized molecular mechanics, to ab initio quantum mechanics, to hybrid methods.[92] For these reasons, it is not completely unexpected that structurally correlated compounds or even the same molecule, such as BIO1211 (for the chemical structure, see Figure 2.1), yielded rather diverse receptor-bound structures, as reviewed in the next paragraphs.

In a very recent study, Baiula et al. analyzed by molecular docking a minilibrary of minimalist analogues of BIO1211, in which the MPUPA-LDVP sequence was reduced to MPUPA-β-residue-Gly, including the previously mentioned MPUPA-β-Pro-Gly (DS-70) (Figure 2.1).[34] This strategy generated a bunch of pure antagonists with IC_{50} affinity values in a low nanomolar range, such as MPUPA-(*R*)-isoAsp(*N*-propyl)-Gly (**2**, Figure 2.1). These compounds reduced the adhesion of integrin-expressing cells to the natural ligands, without inducing receptor open conformation, nor activation of intracellular signaling pathways.

To obtain the inactive structure of the receptor, the authors paired by protein-protein docking the α_4 and β_1 monomers taken from PDB ID 3V4V and PDB ID 4WK0, respectively.

In the so-obtained model (Figure 2.10 A), the putative ligand binding site resides in a vertical crevice at the α/β interface, as shown for $\alpha_4\beta_7$ integrin. This groove can be partitioned into subpockets of different size, shape, and composition of delimiting residues: the middle subpockets A-E, plus lower (L) and upper (U) subpockets (Figure 2.10 B).[34,97]

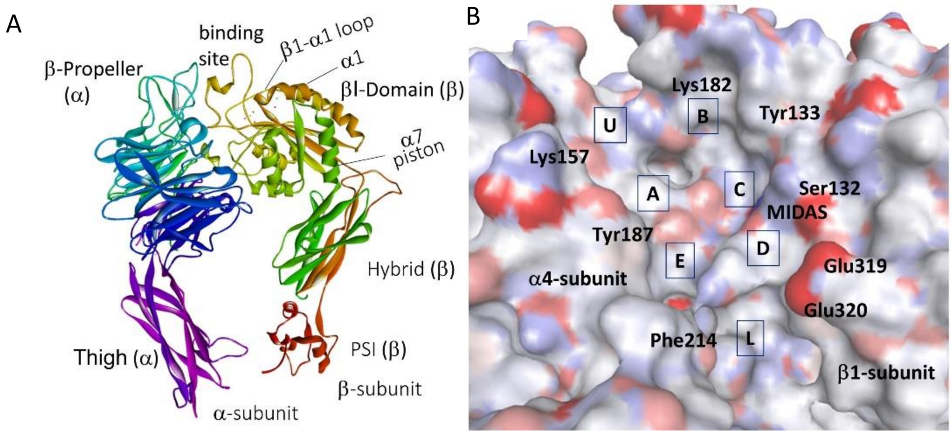


Figure 2.10 (A) Ribbon image of the chimeric $\alpha_4\beta_1$ model. (B) detail of the binding site rendered by the solvent accessible surface, showing subpockets A-E, U, L; the positions of relevant residues are also shown.

In general, in the predicted poses the ligands lie vertically along the $\alpha_4\beta_1$ crevice, with ligand's carboxylate well inserted into subpocket C, so that the main interaction is the salt bridge with Mg^{2+}. As prototypical examples, the sketches of the bioactive conformations of DS-70 and **2** are shown in Figure 2.11. For both, the diphenylurea group finds place in the lower subsite L of the receptor, while the peptide portions attain interactions with the β_1 subunit. For the minimalist peptide **2**, the propyl substituent is allocated in subpocket D. Unfortunately, this structure has the fatal flaw that Asp residue appears to be in the *R* absolute configuration.

The bioactive conformation of **2** and DS70 can be compared to that obtained by da-Silva et al. for MPUPA-LDVP (BIO1211) using a homology model of the $\alpha_4\beta_1$ integrin.[94] In the calculated pose for BIO1211, ligand's backbone folds in a reverse *S*-shape, and AspCOO$^-$ of the peptide makes a salt bridge with Mg^{2+} at MIDAS. The C-terminal Pro occupies a space between subsites B and U on the top of the cavity, while the N-terminal MPUPA finds its position into subpockets E and L (Figure 2.11).

An alternative display of MPUPA-peptides was calculated by Sartori et al.[51] Using the $\alpha_5\beta_1$ integrin complex (3VI4)[85] as a template, the α_4 subunit extrapolated from $\alpha_4\beta_7$ (3V4V)[79] was aligned with α_5, and the model was validated for docking with eight known $\alpha_4\beta_1$ integrin antagonists, including BIO1211 and **3** (for the chemical structures, see Figure 2.1). The analysis of these binding poses revealed that both compounds can interact with Mg^{2+} cation in the β subunit and share a similar disposition within the binding pocket of the receptor, with the common functional groups interacting with the same amino acid residues in the α subunit. Interestingly, for these compounds, the MPUPA group occupied the upper subsite U, which differs from the aforementioned models developed by Baiula and da Silva et al.[34,94] Figure

2.11 reports the binding pose for ligand **3**. This model was further validated with the designed prototype ligand c[Amp(MPUPA)LDVG], which successfully reproduced the ligand-receptor interactions. In particular: (i) the carboxylate group of the Asp residue coordinate with Mg^{2+}; (ii) Key residues Tyr^{187}, Lys^{156}, and Glu^{124} are involved in the ligand-receptor interaction; (iii) Val residue adopts a favorable spatial orientation similar to that observed for BIO1211; and (iv) the MPUPA moiety of both compounds occupies the upper subsite (Figure 2.11). Unfortunately, this compound was found to be inactive in the subsequent $\alpha_4\beta_1$ integrin-mediated cell adhesion assay. Further structural modifications led to the unexpected identification of a retro-sequence agonist c[Amp(MPUPA)VDL], which demonstrated submicromolar activity (for the chemical structures, see Figure 2.3, **6**). To better characterize the agonistic behavior, competition experiments were conducted, revealing that the respective agonist **6** and antagonist BIO1211 behavior were not reciprocally modified. These results seem to exclude the competition between the cyclopeptide **6** and BIO1211 for the same binding site, and the agonist activity could be ascribed to an interaction of this compound in a different region of the receptor.[51] Apparently, further investigations are necessary to elucidate the ligand-receptor interactions for the agonist **6**.

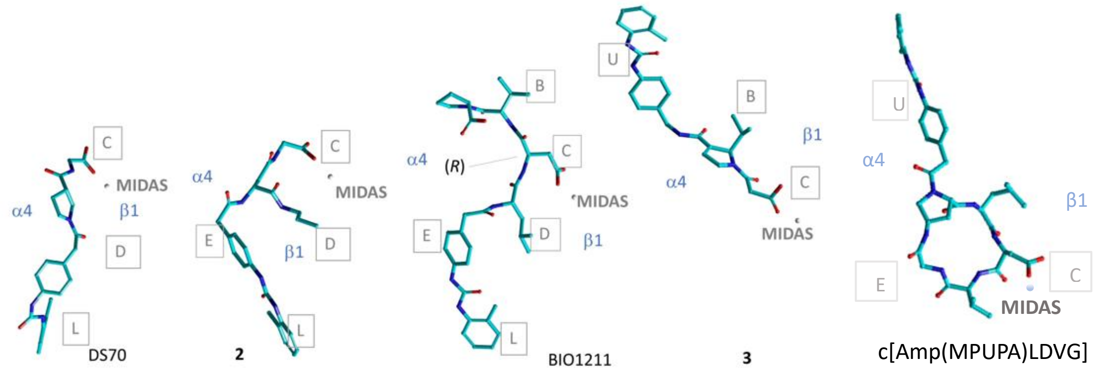


Figure 2.11 Comparison of the predicted binding poses obtained by molecular docking for the MPUPA-antagonists DS-70[34] and **2**[34], **3**[51], BIO1211[94], and c[Amp(MPUPA)LDVG][51].

2.7.2 Simulations of $\alpha_4\beta_1$ integrin agonists

Limited research has been devoted to agonists of leukocyte integrins. The binding poses of small $\alpha_M\beta_2$ agonists composed of a 3-benzylthiazolidine-2,4-dione scaffold connected to a 4-(furan-2-yl)benzoic acid were studied by molecular docking. According to the simulations, the ligands were projected to identify a hydrophobic cleft near the ligand-binding site, indicating the possibility of an allosteric mechanism.[52]

Molecular docking analysis was also performed for the antagonist TBC3486 and related agonists, including urea THI0019, which obtained by the replacement of the carboxylic acid group with a methyl ester, and 7HP349, featuring 2-bis(arylmethyl)carbamate groups in a symmetrical fashion) (Figure 2.12). The $\alpha_4\beta_1$ receptor model was assembled based on the $\alpha_4\beta_7$ and $\alpha_5\beta_1$ crystal structures (3V4V[79] and 3VI4[85], respectively). The predicted binding site partially overlaps the classic binding pocket. Thus, proposed the hypothesis that for the natural ligand to bind, the compounds such as THI0019 would need to be displaced from this site. In addition, neither of the agonists induced the binding of LIMBS mAbs, which bind the β subunit. This finding indicated that the compounds does not engage the MIDAS site, which is consistent with the lack of a carboxylic acid group in their structures. Therefore, it can be concluded that they do not essentially function as ligand mimetics and may induce subtle conformational changes, with their primary effects potentially mediated through the α subunit.[59]

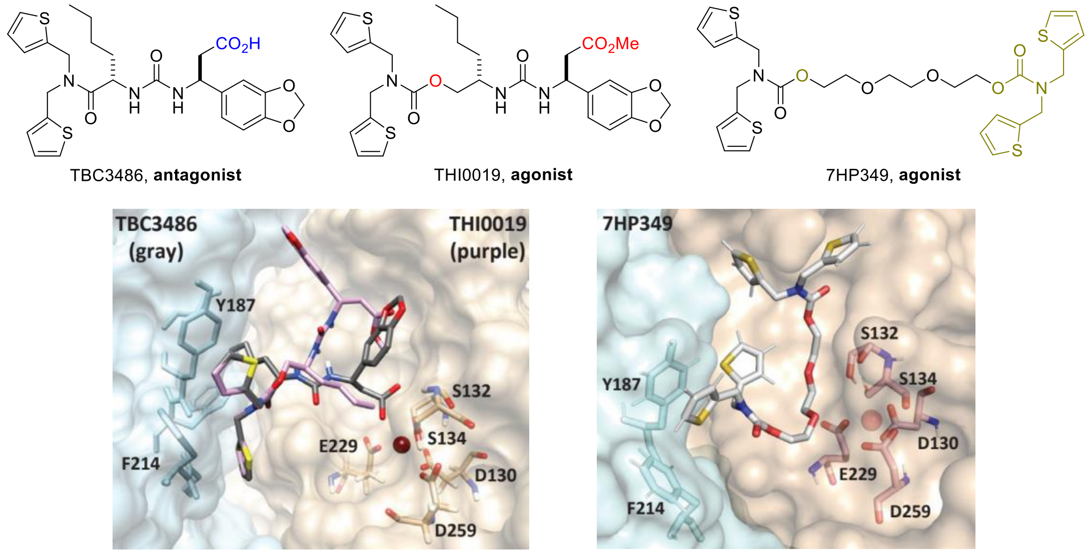


Figure 2.12 The predicted binding modes obtained by molecular docking for the antagonist TBC3486, urea agonists THI0019 and 7HP349 in $\alpha_4\beta_1$.[59] Their chemical sketches are also shown for comparison.

2.8 Conclusion

Inhibitors of intracellular signaling and cell adhesion mediated by $\alpha_4\beta_1$ integrin are regarded as real or promising tools for the treatment of inflammatory diseases, multiple sclerosis, asthma, allergic conjunctivitis, dry eye disease. On the other hand, the agonists of $\alpha_4\beta_1$ Integrins have attracted attention for their potential to impede the recruitment of circulating leukocytes by consistently inhibiting their rolling onto the endothelial surface, thereby hindering their ability to reach sites of inflammation. Also, the activation of $\alpha_4\beta_1$ seems to be a promising strategy to improve cell retention and engraftment in stem cell-based therapies.

At present, very few effective and selective agonists of the α_4 integrins have been reported. As a consequence, very little is known about ligands' characteristics at the basis of receptor blockade or activation. The situation is particularly unclear as the exact structure of the $\alpha_4\beta_1$ integrin has not been disclosed. Extensive crystallographic studies conducted for RGD-binding integrins $\alpha_v\beta_3$ and $\alpha_{IIb}\beta_3$ led to a quite precise depiction of the activation mechanism. However, the proposed models cannot be simply translated to the non-RGD-binding $\alpha_4\beta_1$ integrin. Indeed, unlike RGD-dependent integrins, the binding regions of α_4 integrins have been described as long and wide crevices, open at both ends and capable of the lengthwise accommodation of differently shaped binders.[51,79] Nevertheless, some valuable structural insights can be inferred in light of the crystallographic data above reported for RGD-binding integrins.

In summary, it is generally accepted that small-molecules bind to the integrins without inducing the conformational transition necessary for activation. However, there is evidence that at low concentration several small ligands regarded as antagonists act as partial agonist and promote active conformation of the integrin. Other small-molecules acting as pure antagonists appear to stabilize the inactive conformation by freezing the reorganization of the dynamic interactions about MIDAS, and the dislocation of the β1-α1 loop and ADMIDAS.[76,77,98] The extended conformation of the receptor is better stabilized by ligands with comparatively larger structures, e.g. peptides or cyclopeptides,[86] capable to attain significant interactions with the receptor with high affinity. The discussion presented here emphasizes the urgent need to design and synthesize novel $\alpha_4\beta_1$ integrin ligands to thoroughly explore receptor activation or inhibition. It also can be inferred that, in order to stabilize the extended conformation, agonistic ligands may bind more tightly to the α subunit, as demonstrated by the reported agonist 7HP349.[59]

Note: The work described in this chapter has been published in ***Biomedicines*** with slight modification and has been reproduced here with the permission of the copyright holder.

He, T.; Giacomini, D.; Tolomelli, A.; Baiula, M.; Gentilucci, L.* Conjecturing about Small-Molecule Agonists and Antagonists of $\alpha_4\beta_1$ Integrin: From Mechanistic Insight to Potential Therapeutic Applications. *Biomedicines* **2024**, *12* (2), 316.

References

[1] Hynes, R.O. Integrins: A family of cell surface receptors. *Cell* **1987**, *48* (4), 549-554.

[2] Hynes, R.O. Integrins: Bidirectional, Allosteric Signaling Machines. *Cell* **2002**, *110* (6), 673-687.

[3] Fagerholm, S.C. Integrins in health and disease. *N. Engl. J. Med*. **2022**, *387*, 1519-1521.

[4] Mas-Moruno. C.; Rechenmacher, F.; Kessler, H. Cilengitide: the first anti-angiogenic small molecule drug candidate design, synthesis and clinical evaluation. *Anticancer Agents Med. Chem*. **2010**, *10*, 753-768.

[5] Reynolds, A.; Hart, I.; Watson, A.; Welti, J.C.; Silva, R.G.; Robinson, S.D.; Da Violante, G.; Gourlaouen, M.; Salih, M.; Jones, M.C.; Jones, D.T.; Saunders, G.; Kostourou, V.; Perron-Sierra, F.; Norman, J.C.; Tucker, G.C.; Hodivala-Dilke, K.M. Stimulation of tumor growth and angiogenesis by low concentrations of RGD-mimetic integrin inhibitors. *Nat. Med.* **2009**, *15*, 392-400.

[6] Miller, D.H.; Weber, T.; Grove, R.; Wardell, C.; Horrigan, J.; Graff, O.; Atkinson, G.; Dua, P.; Yousry, T.; Macmanus, D., et al. Firategrast for relapsing remitting multiple sclerosis: a phase 2, randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. **2012**, *11*, 131-139.

[7] Chew, D.P.; Bhatt, D.L.; Sapp, S.; and Topol, E.J. Increased mortality with oral platelet glycoprotein IIb/IIIa antagonists: a meta-analysis of phase III multicenter randomized trials. *Circulation* **2001**, *103*, 201-206.

[8] Topol, E.J.; Easton, D.; Harrington, R.A.; Amarenco, P.; Califf, R.M.; Graffagnino, C.; Davis, S.; Diener, H.C.; Ferguson, J.; Fitzgerald, D., et al. Randomized, double-blind, placebo-controlled, international trial of the oral IIb/IIIa antagonist lotrafiban in coronary and cerebrovascular disease. *Circulation* **2003**, *108*, 399-406.

[9] Théroux, P.; Catella-Lawson, F.; Armstrong, P.; DeCani, J.; Hirsh, J.; Pepine, C.; Ryan, T.J.; Pelletier, G.; Davies, R.; Flather, M.; et al. Inhibition of the platelet glycoprotein IIb/IIIa receptor with tirofiban in unstable angina and non-Q-wave myocardial infarction. *N. Engl. J. Med.* **1998**, *338*, 1488-1497.

[10] Lin, F.-Y.; Li, J.; Xie, Y.; Zhu, J.; Huong Nguyen, T.-T.; Zhang, Y.; Zhu, J.; Springer, T.A. A general chemical principle for creating closure-stabilizing integrin inhibitors. *Cell* **2022**, *185*, 3533-3550.

[11] Muller, W. Leukocyte-endothelial cell interactions in the inflammatory response. *Lab. Invest*. **2002**, *82*, 521-534.

[12] Ferreira, E.F.B.; Silva, L.B.; Cruz, J.V.; Araújo, P.H.F.; Kimani, N.M.; Leite, F.H.A.; Campos, J.M.; Santos, C.B.R. An overview of the $\alpha_4\beta_1$ integrin and the potential therapeutic role of its antagonists. *Curr. Med. Chem*. **2021**, *28*, 5884-5895.

[13] Ou, Z.; Dolmatova, E.; Lassègue, B.; Griendling, K.K. β_1- and β_2-integrins: central players in regulating vascular permeability and leukocyte recruitment during acute inflammation. *Am. J. Physiol. Heart. Circ. Physiol*. **2021**, *320*, H734-H739.

[14] Hyun, Y.M.; Lefort, C.T.; Kim, M. Leukocyte integrins and their ligand interactions. *Immunol. Res*. **2009**, *45*, 195-208.

[15] Newham, P.; Craig, S. E.; Seddon, G. N.; Schofield, N. R.; Rees, A.; Edwards, R. M.; Jones, E. Y.; Humphries, M. J. α_4 Integrin Binding Interfaces on VCAM-1 and MAdCAM-1. *J. Biol. Chem*. **1997**, *272*, 19429-19440.

[16] Baiula, M.; Spampinato, S.; Gentilucci, L.; Tolomelli, A. Novel ligands targeting $\alpha_4\beta_1$ integrin: Therapeutic applications and perspectives. *Front. Chem*. **2019**, *7*, 489.

[17] Lobb, R.R.; Hemler, M.E. The pathophysiologic role of α_4 integrins in vivo. *J. Clin. Invest*. **1994**, *94*, 1722-1728.

[18] Pyka-Fościak, G.; Litwin, J.A.; Lis, G.J. Osteopontin expression and the effect of anti-VLA-4 mAb treatment in experimental autoimmune encephalomyelitis, a mouse model of multiple sclerosis. *Folia Neuropathologica*, **2023**, *61*, 129-137.

[19] Outteryck, O. Natalizumab in relapsing-remitting multiple sclerosis. *Expert. Rev. Neurother*. **2016**, *16*, 471-481.

[20] Shirani, A.; Stüve, O. Neuroimmunology in point for the impact of translational natalizumab for multiple sclerosis: a case. *J. Immunol*. **2017**, *198*, 1381-1386.

[21] Nelson, S.M.L.; Nguyen, T.M.; McDonald, J.W.D.; MacDonald, J.K. Natalizumab for induction of remission in Crohn's disease. *Cochr. Database Syst. Rev.* **2018**, 8, CD006097.

[22] Yang, G.X.; Hagmann, W.K. VLA-4 antagonists: Potent inhibitors of lymphocyte migration. *Med. Res. Rev.* **2003**, *23*, 369-392.

[23] Sugiura, T.; Andou, A.; Hosoi, Y.; Koyama, T. Pharmacological and clinical data of oral alpha 4 integrin antagonist, Carotegrast methyl, CAROGRA. *Nihon Yakurigaku Zasshi/Folia Pharmacol. Jpn.* **2023**, *158*, 203-210.

[24] Lin, Kc; Ateeq, H.S.; Hsiung, S.H.; Chong, L.T.; Zimmerman, C.N.; Castro, A.; Lee, W.C.; Hammond, C.E.; Kalkunte, S.; Chen, L.L.; Pepinsky, R.B.; Leone, D.R.; Sprague, A.G.; Abraham, W.M.; Gill, A.; Lobb, R.R.; Adams, S.P. Selective, tight-binding inhibitors of integrin $\alpha_4\beta_1$ that inhibit allergic airway responses. *J. Med. Chem*. **1999**, *42*, 920-934.

[25] Abraham, W.M.; Gill, A.; Ahmed, A.; Sielczak, M.W.; Lauredo, I.T.; Botinnikova, Y.; Lin, K.C.; Pepinsky, B.; Leone, D.R.; Lobb, R.R.; Adams, S.P. A small-molecule, tight-binding inhibitor of the integrin $\alpha_4\beta_1$ blocks

antigen-induced airway responses and inflammation in experimental asthma in sheep. *Am. J. Respir. Crit. Care. Med*. **2000**, *162*, 603-611.

[26] Karanam, B.V.; Jayra, A.; Rabe, M.; Wang, Z.; Keohane, C.; Strauss, J.; Vincent, S. Effect of enalapril on the in vitro and in vivo peptidyl cleavage of a potent VLA-4 antagonist. *Xenobiotica* **2007**, *37*, 487-502.

[27] Singh, J.; Van Vlijmen, H.; Liao, Y.; Lee, W.C.; Cornebise, M.; Harris, M.; Shu, I.; Gill, A.; Cuervo, J.H.; Abraham, W.M.; Adams, S.P. Identification of potent and novel $\alpha_4\beta_1$ antagonists using in silico screening. *J. Med. Chem*. **2002**, *45*, 2988-2993.

[28] Vagner, J.; Qu, H.; Hruby, V.J. Peptidomimetics, a synthetic tool of drug discovery. *Curr. Opin. Chem. Biol*. **2008**, *12*, 292-296.

[29] Jackson, D. Y. Alpha4 integrin antagonists. *Curr. Pharm. Des*. **2002**, *8*, 1229-1253.

[30] De Marco, R.; Tolomelli, A.; Juaristi, E.; Gentilucci, L. Integrin ligands with α/β-hybrid peptide structure: Design, bioactivity, and conformational aspects. *Med. Res. Rev*. **2016**, *36*, 389-424

[31] De Marco, R.; Mazzotti, G.; Dattoli, S.D.; Baiula, M.; Spampinato, S.; Greco, A.; Gentilucci, L. 5-Aminomethyloxazolidine-2,4-dione hybrid α/β-dipeptide scaffolds as inductors of constrained conformations: Applications to the synthesis of integrin antagonists. *Pept. Sci*. **2015**, *104*, 636-649.

[32] Dattoli, S.D.; De Marco, R.; Baiula, M.; Spampinato, S.; Greco, A.; Tolomelli, A.; Gentilucci, L. Synthesis and assay of retro-$\alpha_4\beta_1$ integrin-targeting motifs. *Eur. J. Med. Chem*. **2014**, *73*, 225-232.

[33] Dattoli, S.D.; Baiula, M.; De Marco, R.; Bedini, A.; Anselmi, M.; Gentilucci, L.; Spampinato, S. DS-70, a novel and potent α_4 integrin antagonist, is an effective treatment for experimental allergic conjunctivitis in guinea pigs. *Br. J. Pharmacol*. **2018**, *175*, 3891-3910.

[34] Baiula, M.; Anselmi, M.; Musiani, F.; Ghidini, A.; Carbone, J.; Caligiana, A.; Maurizio, A.; Spampinato, S.; Gentilucci, L. Design, pharmacological characterization, and molecular docking of minimalist peptidomimetic antagonists of $\alpha_4\beta_1$ integrin. *Int. J. M. Sci*. **2023**, *24*, 9588.

[35] Tolomelli, A.; Baiula, M.; Viola, A.; Ferrazzano, L.; Gentilucci, L.; Dattoli, S.D.; Spampinato, S.; Juaristi, E.; Escudero, M. Dehydro-β-proline containing $\alpha_4\beta_1$ integrin antagonists: Stereochemical recognition in ligand-receptor interplay. *ACS Med. Chem. Lett*. **2015**, *6*, 701-706.

[36] Chiba, J.; Machinaga, N.; Takashi, T.; Ejima, A.; Takayama, G.; Yokoyama, M.; Nakayama, A.; Baldwin, J.J.; McDonald, E.; Moriarty, K.J.; et al. Identified a morpholinyl-4-piperidinylacetic acid derivative as a potent oral active VLA-4 antagonist. *Bioorg. Med. Chem. Lett*. **2005**, *15*, 41-45.

[37] Krauss, A.H.; Corrales, R.M.; Pelegrino, F.S.A.; Tukler-Henriksson, J.; Pflugfelder, S.C.; de Paiva, C.S. Improvement of outcome measures of dry eye by a novel integrin antagonist in the murine desiccating stress model. *Invest. Ophthalmol. Vis. Sci*. **2015**, 56, 5888-5895.

[38] Baiula, M.; Caligiana, A.; Bedini, A.; Zhao, J.; Santino, F.; Cirillo, M.; Gentilucci, L.; Giacomini, D.; Spampinato, S. Leukocyte integrin antagonists as a novel option to treat dry age-related macular degeneration. *Front. Pharmacol*. **2021**, *11*, 617836.

[39] Bellavia, M.C.; Nyiranshuti, L.; Latoche, J.D.; Ho, K.-V.; Fecek, R.J.; Taylor, J.L.; Day, K.E.; Nigam, S.; Pun, M.; Gallazzi, F.; Edinger, R.S.; Storkus, W.J.; Patel, Ravi B.; Anderson, Carolyn J. PET imaging of VLA-4 in a new BRAFV600E mouse model of melanoma. *J. Mol. Biol*. **2022**, *24*, 425-433

[40] Zhao, J.; Santino, F.; Giacomini, D.; Gentilucci, L. Integrin-Targeting Peptides for the Design of Functional Cell-Responsive Biomaterials. *Biomedicines* **2020**, *8*, 307.

[41] Perkins, L.A.; Nyiranshuti, L.; Little-Ihrig, L.; Latoche, J.D.; Day, K.E.; Zhu, Q.; Tavakoli, S.; Sundd, P.; Novelli, E.M.; Anderson, C.J. Integrin VLA-4 as a PET imaging biomarker of hyper-adhesion in transgenic sickle mice. *Blood Adv*. **2020**, 4, 17, 4102-4112.

[42] De Marco, R.; Greco, A.; Calonghi, N.; Dattoli, S. D.; Baiula, M.; Spampinato, S.; Picchetti, P.; De Cola, L.; Anselmi, M.; Cipriani, F.; Gentilucci, L. Selective detection of $\alpha_4\beta_1$ integrin (VLA-4)-expressing cells using peptide-functionalized nanostructured materials mimicking endothelial surfaces adjacent to inflammatory sites. *Pept. Sci*. **2018**, *110*, e23081.

[43] Anselmi, M.; Baiula, M.; Santino, F.; Zhao, J.; Spampinato, S.; Calonghi, N.; Gentilucci, L. Design of α/β-hybrid peptide ligands of $\alpha_4\beta_1$ integrin equipped with a linkable side chain for chemoselective biofunctionalization of microstructured materials. Bi*omedicines* **2021**, *9*, 1737.

[44] Kakinoki, S.; Nishioka, S.; Arichi, Y.; Yamaoka, T. Stable and direct coating of fibronectin-derived Leu-Asp-Val peptide on ePTFE using one-pot tyrosine oxidation for endothelial cell adhesion. *Colloids Surf. B: Biointerfaces* **2022**, *216*, 112576.

[45] Hao, D.; Liu, R.; Fernandez, T.G.; Pivetti, C.; Jackson, J.E.; Kulubya, E.S.; Jiang, H.-J.; Ju, H.-Y.; Liu, W.-L.; Panitch, A.; Lam, K.S.; Kent Leach, J.; Farmer, D.L.; Wang, A. A bioactive material with dual integrin-targeting ligands regulates specific endogenous cell adhesion and promotes vascularized bone regeneration in adult and fetal bone defects. *Bioact. Mater*. **2022**, *20*, 179-193.

[46] Market, R.V.; Biediger, R.J.; Woodside, D.G. Piperazine-Based Agonists of LFA-1 and VLA-4. WO2023018471A1, 9 June **2022**.

[47] Vanderslice, P.; Biediger, R.J.; Woodside, D.G.; Brown, W.S.; Khounlo, S.; Warier, N.D.; Gundlach IV, C.W.; Caivano, A.R.; Bornmann, W.G.; Maxwell, D.S.; et al. Small molecule agonist of very late antigen-4 (VLA-4) integrin induces progenitor cell adhesion. *J. Biol. Chem.* **2013**, *288*, 19414-19428.

[48] Lewis, L.; Schary, W.; De, S.; Vanderslice, P.; Biediger, R.; Hailemichael, Y.; Woodside, D.; Overwijk, W.; Lee, F.; Marathi, U. 603 A randomized, placebo-controlled, first in human phase I single ascending dose (SAD) and multiple ascending dose (MAD) study in healthy male volunteers of the immunostimulant 7HP349. *J. ImmunoTher. Cancer* **2022**, *10*, A632-A632.

[49] Galletti, P.; Soldati, R.; Pori, M.; Durso, M.; Tolomelli, A.; Gentilucci, L.; Dattoli, S. D.; Baiula, M.; Spampinato, S.; Giacomini, D. Targeting integrins $\alpha_v\beta_3$ and $\alpha_5\beta_1$ with new β-lactam derivatives. *Eur. J. Med. Chem*. **2014**, *83*, 284-293.

[50] Martelli, G.; Baiula, M.; Caligiana, A.; Galletti, P.; Gentilucci, L.; Artali, R.; Spampinato, S.; Giacomini, D. Could dissecting the molecular framework of β-lactam integrin ligands enhance selectivity? *J. Med. Chem*. **2019**, *62*, 10156-10166.

[51] Sartori, A.; Bugatti, K.; Portioli, E.; Baiula, M.; Casamassima, I.; Bruno, A.; Bianchini, F.; Curti, C.; Zanardi, F.; Battistini, L. New 4-aminoproline-based small molecule cyclopeptidomimetics as potential modulators of $\alpha_4\beta_1$ integrin. *Molecules* **2021**, *26*, 6066.

[52] Faridi, M. H.; Maiguel, D.; Barth, C. J.; Stoub, D.; Day, R.; Schürer, S.; Gupta, V. Identification of novel agonists of the integrin CD11b/CD18. *Bioorg. Med. Chem. Lett*. **2009**, *19*, 6902-6906.

[53] Maiguel, D.; Faridi, M. H.; Wei, C.; Kuwano, Y.; Balla, K. M.; Hernandez, D.; Barth, C. J.; Lugo, G.; Donnelly, M.; Nayer, A.; Moita, L. F.; Schürer, S.; Traver, D.; Ruiz, P.; Vazquez-Padron, R. I.; Ley, K.; Reiser, J.; Gupta, V. Small molecule-mediated activation of the integrin CD11b/CD18 reduces inflammatory disease. *Sci. Signal*. **2011**, *4*, 1-14.

[54] Faridi, M. H.; Altintas, M. M.; Gomez, C.; Duque, J. C.; Vazquez-Padron, R. I.; Gupta, V. Small molecule agonists of integrin CD11b/CD18 do not induce global conformational changes and are significantly better than activating antibodies in reducing vascular injury. *Biochim. Biophys. Acta* **2013**, *1830*, 3696-3710.

[55] Yang, W.; Carman, C. V.; Kim, M.; Salas, A.; Shimaoka, M.; Springer, T. A. A small molecule agonist of an integrin, $\alpha_L\beta_2$. *J. Biol. Chem*. **2006**, *281*, 37904-37912.

[56] Strozier, R.E. Integrin Agonists or Activating Compounds and Methods for Making and Using Same. US20220118086A1, 10 September **2021**.

[57] Yao, W.; Liu, W.; Lam, K.S.; Xiao, W.; Lane, N. Peptides for Activation of Cell Signaling in Osteoprogenitor Cells. EP3823630A4, 20 April **2022**.

[58] Oh, J.; Magnuson, A.; Benoist, C.; Pittet, M.J.; Weissleder, R. Age related tumor growth in mice is related to integrin α_4 in CD8+ T cells. *JCI Insight* **2018**, *3*, e122961.

[59] Hickman, A.; Koetsier, J.; Kurtanich, T.; Nielsen, M.C.; Winn, G.; Wang, Y.; Bentebibel, S.E.; Shi, L.; Punt, S.; Williams, L.; et al. LFA-1 activation enriches tumor-specific T cells in a cold tumor model and synergizes with CTLA-4 blockade. *J. Clin. Invest.* **2022**, *132* (13), e154152.

[60] Lokugamage, N.; Chowdhury, I.H.; Biediger, R.J.; Market, R.V.; Khounlo, S.; Warier, N.D.; Hwang, S.A.; Actor, J.K.; Woodside, D.G.; Marathi, U.; et al. Use of a small molecule integrin activator as a systemically administered vaccine adjuvant in controlling Chagas disease. *NPJ Vaccines* **2021**, *6*, 114.

[61] Villanueva, V.; Li, X.; Jimenez, V.; Faridi, H.M.; Gupta, V. CD11b agonists offer a novel approach for treating lupus nephritis. *Transl. Res.* **2022**, *245*, 41-54.

[62] Ehirchiou, D.; Bernabei, I.; Chobaz, V.; Castelblanco, M.; Hügle, T.; So, A.; Zhang, L.; Busso, N.; Nasi, S. CD11b Signaling prevents chondrocyte mineralization and attenuates the severity of osteoarthritis. *Front. Cell Dev. Biol.* **2020**, *8*, 611757.

[63] Cirillo, M.; Martelli, G.; Boanini, E.; Rubini, K.; Di Filippo, M.; Torricelli, P.; Pagani, S.; Fini, M.; Bigi, A.; Giacomini, D. Strontium substituted hydroxyapatite with β-lactam integrin agonists to enhance mesenchymal cells adhesion and to promote bone regeneration. *Colloids Surf. B: Biointerfaces* **2021,** *200*, 111580.

[64] Merlo, B.; Baldassarro, V.A.; Flagelli, A.; Marcoccia, R.; Giraldi, V.; Focarete, M.L.; Giacomini, D.; Iacono, E. Peptide mediated adhesion to beta-lactam ring of equine mesenchymal stem cells: A pilot study. *Animals* **2022**, *12*, 734.

[65] Baldassarro, V.A.; Giraldi, V.; Giuliani, A.; Moretti, M.; Pagnotta, G.; Flagelli, A.; Clavenzani, P.; Lorenzini, L.; Giardino, L.; Focarete, M.L.; et al. Poly(l-lactic acid) scaffold releasing an $\alpha_4\beta_1$ integrin agonist promotes nonfibrotic skin wound healing in diabetic mice. *ACS Appl. Bio Mater.* **2023**, *6*, 296-308.

[66] Choi, W.-S.; Rice, W.J.; Stokes, D.L.; Coller, B.S. Three-dimensional reconstruction of intact human integrin $\alpha_{IIb}\beta_3$: New implications for activation-dependent ligand binding. *Blood* **2013**, *122*, 4165-4171.

[67] Huo, T.; Wu, H.; Moussa, Z.; Sen, M.; Dalton, V.; Wang, Z. Full-length $\alpha_{IIb}\beta_3$ cryo-EM structure reveals intact integrin initiate-activation intrinsic architecture. *Structure* **2024**, *32* (7), 899-906, e3.

[68] Li, J.; Springer, T.A. Energy landscape differences among integrins establish the framework for understanding activation. *J. Cell Biol.* **2018**, *217*, 397-412.

[69] Zhu, J.; Luo, B.-H.; Xiao, T.; Zhang, C.; Nishida, N.; Springer, T. A. Structure of a Complete Integrin Ectodomain in a Physiologic Resting State and Activation and Deactivation by Applied Forces. *Mol. Cell* **2008**, *32* (6), 849-861.

[70] Springer, T.A.; Zhu, J.; Xiao, T. Structural basis for distinctive recognition of fibrinogen γC peptide by the platelet integrin $\alpha_{IIb}\beta_3$. *J. Cell Biol.* **2008**, *182*, 791-800.

[71] Zhu, J.; Zhu, J.; Springer, T.A. Complete integrin headpiece opening in eight steps. *J. Cell Biol.* **2013**, *201*, 1053-1068.

[72] Xu, X.P.; Kim, E.; Swift, M.; Smith, J.W.; Volkmann, N.; Hanein, D. Three-dimensional structures of full-length, membrane embedded human $\alpha_{IIb}\beta_3$ integrin complexes. *Biophys. J.* **2016**, *110*, 798-809.

[73] Adair, B.D.; Xiong, J.-P.; Yeager, M.; Arnaout, M.A. Cryo-EM structures of full-length integrin $\alpha_{IIb}\beta_3$ in native lipids. *Nat. Commun.* **2023**, *14*, 4168.

[74] Lin, F.-Y.; Zhu, J.; Eng, E.T.; Hudson, N.E.; Springer, T.A. β-Subunit binding is sufficient for ligands to open the integrin $\alpha_{IIb}\beta_3$ headpiece. *J. Biol. Chem.* **2016**, *291*, 4537-4546.

[75] Xiong, J.-P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S.L.; Arnaout, M.A. Crystal structure of the extracellular segment of integrin $\alpha_v\beta_3$ in complex with an Arg-Gly-Asp ligand. *Science* **2002**, *296*, 151-155.

[76] Van Agthoven, J.F.; Xiong, J.P.; Alonso, J.L.; Rui, X.; Adair, B.D.; Goodman, S.L.; Arnaout, M.A. Structural basis for pure antagonism of integrin $\alpha_v\beta_3$ by a high-affinity form of fibronectin. *Nat. Struct. Mol. Biol.* **2014**, *21*, 383-388.

[77] Li, J.; Fukase, Y.; Shang, Y.; Zou, W.; Muñoz-Félix, J.M.; Buitrago, L.; van Agthoven, J.; Zhang, Y.; Hara, R.; Tanaka, Y.; et al. Novel pure $\alpha_v\beta_3$ integrin antagonists that do not induce receptor extension, prime the receptor, or enhance angiogenesis at low concentrations. *ACS Pharmacol. Transl. Sci.* **2019**, *2*, 387-401.

[78] Yu, Y.; Schürpf, T.; Springer, T. A. How Natalizumab Binds and Antagonizes α_4 Integrins. *J. Biol. Chem.* **2013**, *288* (45), 32314-32325.

[79] Yu, Y.; Zhu, J.; Mi, L.-Z.; Walz, T.; Sun, H.; Chen, J.; Springer, T. A. Structural specializations of $\alpha_4\beta_7$, an integrin that mediates rolling adhesion. *J. Cell Biol.* **2012**, *196* (1), 131-146.

[80] Luzentales-Simpson, M.; Pang, Y.C.F.; Zhang, A.; Sousa, J.A.; Sly, L.M. Vedolizumab: Potential Mechanisms of Action for Reducing Pathological Inflammation in Inflammatory Bowel Diseases. *Front. cell dev. biol.* **2021**, *9*, 612830.

[81] Jo, M.H.; Li, J.; Jaumouillé, V.; Hao, Y.; Coppola, J.; Yan, J.; Waterman, C.M.; Springer, T.A.; Ha, T. Single-molecule characterization of subtype-specific β_1 integrin mechanics. *Nat. Commun.* **2022**, *13*, 7471.

[82] Miyazaki, N.; Iwasaki, K.; Takagi, J. A systematic survey of conformational states in β_1 and β_4 integrins using negative-stain electron microscopy. *J. Cell Sci.* **2018**, *131*, jcs216754.

[83] Su, Y.; Xia, W.; Li, J.; Walz, T.; Humphries, M.J.; Vestweber, D.; Cabañas, C.; Lu, C.; Springer, T.A. Relating conformation to function in integrin $\alpha_5\beta_1$. *Proc. Natl. Acad. Sci. USA.* **2016**, *113*, E3872-E3881.

[84] Li, J.; Springer, T.A. Integrin extension enables ultrasensitive regulation by cytoskeletal force. *Proc. Natl. Acad. Sci. USA.* **2017**, *114*, 4685-4690.

[85] Nagae, M.; Re, S.; Mihara, E.; Nogi, T.; Sugita, Y.; Takagi, J. Crystal structure of $\alpha_5\beta_1$ integrin ectodomain: Atomic details of the fibronectin receptor. *J. Cell Biol.* **2012**, *197*, 131-140.

[86] Xia, W.; Springer, T.A. Metal ion and ligand binding of integrin $\alpha_5\beta_1$. *Proc. Natl. Acad. Sci. USA.* **2014**, *111*, 17863-17868.

[87] Anderson, J.M.; Li, J.; Springer, T. A. Regulation of integrin $\alpha_5\beta_1$ conformational states and intrinsic affinities by metal ions and the ADMIDAS. *Mol. Biol. Cell* **2022**, *33* (6), ar56.

[88] Schumacher, S.; Dedden, D.; Vazquez Nunez, R.; Matoba, K.; Takagi, J.i; Biertümpfel, C.; Mizuno, N. Structural insights into integrin $\alpha_5\beta_1$ opening by fibronectin ligand. *Sci. Adv.* **2021**, *7*, eabe9716.

[89] Wang, X.; Guillem-Marti, J.; Kumar, S.; Lee, D.S.; Cabrerizo-Aguado, D.; Werther, R.; Alamo, K.A.E.; Zhao, Y.T.; Nguyen, A.; Kopyeva, I.; Huang, B.; Li, J.; Hao, Y.; Li, X.; Brizuela-Velasco, A.; Murray, A.; Gerben, S.; Roy, A.; DeForest, C.A.; Springer, T.; Ruohola-Baker, H.; Cooper, J.A.; Campbell, M.G.; Manero, J.M.; Ginebra, M.P.; Baker, D. De Novo Design of Integrin $\alpha_5\beta_1$ Modulating Proteins for Regenerative Medicine. *BioRxiv.* [Preprint] **2024**, Jul 10: 2024. 06. 21. 600123.

[90] You, T.J.; Maxwell, D.S.; Kogan, T.P.; Chen, Q.; Li, J.; Kassir, J.; Holland, G.W.; Dixon, R.A. A 3D structure model of integrin $\alpha_4\beta_1$ complex: I. Construction of a homology model of β_1 and ligand binding analysis. *Biophys. J.* **2002***, 82,* 447-457.

[91] Macchiarulo, A.; Costantino, G.; Meniconi, M.; Pleban, K.; Ecker, G.; Bellocchi, D.; Pellicciari, R. Insights into phenylalanine derivatives recognition of VLA-4 integrin: from a pharmacophoric study to 3D-QSAR and molecular docking analyses. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1829-1839.

[92] Tvaroška, I.; Kozmon, S.; Kóňa, J. Molecular modeling insights into the structure and behavior of integrins: A review. *Cells* **2023**, *12*, 324.

[93] Thangapandian, S.; John, S.; Sakkiah, S.; Lee, K.W. Discovery of potential integrin VLA-4 antagonists using pharmacophore modeling, virtual screening and molecular docking studies. *Chem. Biol. Drug Des*. **2011**, *78*, 289-300.

[94] Vasconcelos, D.; Chaves, B.; Albuquerque, A.; Andrade, L.; Henriques, A.; Sartori, G.; Savino, W.; Caffarena, E.; Martins-Da- Silva, J.H. Development of New Potential Inhibitors of β_1 Integrins through In Silico Methods-Screening and Computational Validation. *Life* **2022**, *12*, 932.

[95] Hatley, R.J.D.; Barrett, T.N.; Slack, R.J.; Watson, M.E.; Baillache, D.J.; Gruszka, A.; Washio, Y.; Rowedder, J.E.; Pogány, P.; Pal, S.; Macdonald, S.J.F. The design of potent, selective and drug-like RGD $\alpha_v\beta_1$ small-molecule inhibitors derived from non-RGD $\alpha_4\beta_1$ antagonists. *ChemMedChem* **2019**, *14*, 1315-1320

[96] Zhang, L.; Yan, H.; Tai, Y.; Xue, Y.; Wei, Y.; Wang, K.; Zhao, Q.; Wang, S.; Kong, D.; Midgley, A.C. Design and evaluation of a polypeptide that mimics the integrin binding site for EDA fibronectin to block profibrotic cell activity. *Int. J. Mol. Sci*. **2021**, *22*, 1575.

[97] Carbone, J.; Ghidini, A.; Romano, A.; Gentilucci, L.; Musiani, F. PacDOCK: A Web server for positional distance-based and interaction-based analysis of docking results. *Molecules* **2022**, *27*, 6884.

[98] Paladino, A.; Civera, M.; Belvisi, L.; Colombo, G. High Affinity vs. native fibronectin in the modulation of $\alpha_v\beta_3$ integrin conformational dynamics: insights from computational analyses and implications for molecular design. *PLoS Comput. Biol*. **2017**, 13, e1005334.

Chapter III. Design and Synthesis of LDV-based $\alpha_4\beta_1$ Integrin Cyclopeptides

3.1 Introduction

In light of the comprehensive discussion above, it is evident that $\alpha_4\beta_1$ integrin plays a significant role in both physiological and pathological conditions, especially in inflammation and cancer diseases. However, very few $\alpha_4\beta_1$ integrin agonists (e.g. β-lactam derivatives **1**, **2**, Figure 3.1)[1-3] have been discovered so far, and the reported antagonists, such as BIO1211 (Figure 3.1), suffer from enzymatic stability problem, thus limiting the investigation of their potential therapeutic efficacy.[4-10] Hence, there is an urgent need to design and synthesize novel potent $\alpha_4\beta_1$ integrin ligands, which can enrich the existing $\alpha_4\beta_1$ small molecule library and enable the computational characterization of the ligands' structural determinants for agonism and antagonism, since the crystal structure is not revealed yet.

In this context, we conceived a mini-library of LDV α/β hybrid cyclopentapeptides (CPPs) **3** and related sequences (Figures 3.1). This approach yielded integrin agonists with diverse affinity for α_4 integrins. Also, the CPPs were utilized as 3D probes for investigating the preferred bioactive conformations, and to analyze $\alpha_4\beta_1$ integrin-ligand binding behavior.

For convenience and clarity, the chemical structures of BIO1211, and **1**, **2** are shown again in the Figure 3.1, as they will be later used as reference compounds in this study. For a more detailed discussion on aspects such as activity and application, please refer to Chapter 2.

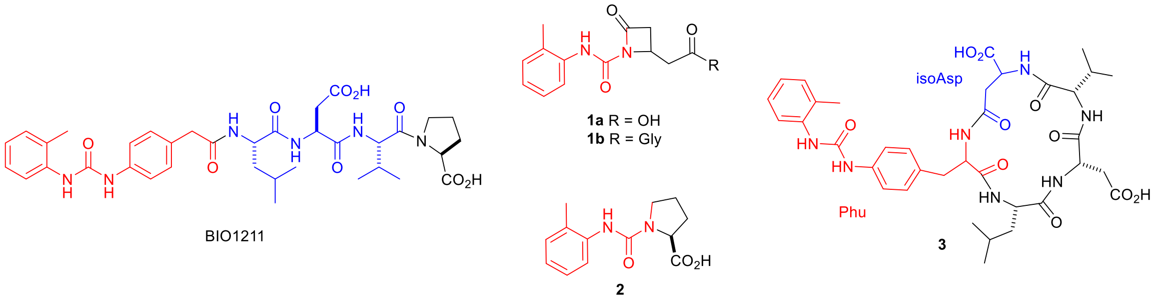
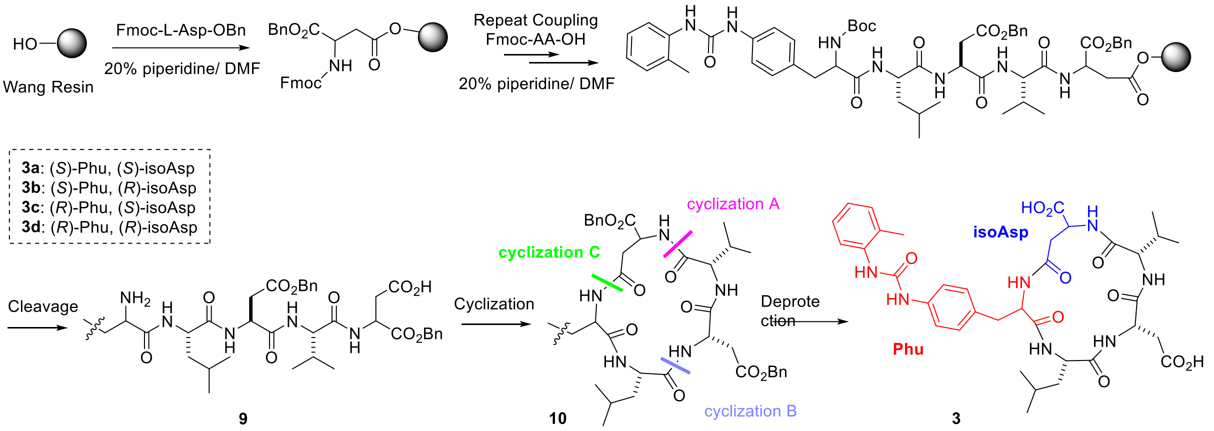


Figure 3.1 Structures of $\alpha_4\beta_1$ integrin ligands. Antagonists: the LDV peptide-urea BIO1211. Agonists: the small urea-derivatives **1** and **2**. Cyclic analogues of BIO1211-**3**, including LDV and the phenylalanine-urea (Phu) residue.

3.2 Results and Discussion

3.2.1 Synthesis of CPPs (3a-3d)

To obtain LDV CPPs equipped with α_4-targeting MPUPA moiety, the diphenylurea moiety was anchored at the 4-position of (*S*)- or (*R*)-Phenylalanine, giving *(S)*- or *(R)*-*p*[3-(*o*-tolyl)urea]phenylalanine (Phu).[11,12] The sequence was complemented with the β-amino acid *(S)*- or *(R)*-isoAsp, to consent macrolactamization while maintaining a free carboxylic group, as at the C-terminus of BIO1211 (Scheme 3.1).



Scheme 3.1 Synthesis of CPPs **7a-7d** via macrolactamization. Alternative routes A and B are also shown. For simplicity, part of Phu can be omitted.

The CPPs of general structure c[(*S/R*)-Phu-Leu-Asp-Val-(*S/R*)-isoAsp] (**3a-3d**) were prepared from linear precursors, obtained in turn by standard solid phase peptide synthesis (SPPS) on Wang resin, with Fmoc-protected amino acids (Scheme 3.1).

The identification of the strategic amide bond for final head-to-tail cyclization was not trivial.[13] Initially, we opted for the convenient cyclization A between the residues isoAsp and Val (Scheme 3.1, cyclization A). However, this approach resulted in poor yields of **10a** (Appendix A: Table S1), which is consistent with the results obtained by Kessler and Marinelli for the cyclization of isoDGR peptides.[14] These authors noted that the presence of isoAsp strongly influenced the conformation of linear peptide precursors and recommended that cyclization could only be achieved if isoAsp was positioned in the middle of the sequence. In contrast to the expectations, the cyclization between Asp and Leu (Scheme 3.1, cyclization B) also gave a negligible yield (Appendix A: Table S1). Much better results were obtained in our case for the ring-forming reaction between Phu and isoAsp (Scheme 3.1, ring closure C).

Hence, the linear sequences **9a-9d** were prepared by standard Fmoc chemistry on a Wang resin (Scheme 3.1, Appendix A: Table S1). The crude precursors **9a-9d** (75-85% pure) were utilized directly for the macrolactamization step under pseudo-high-dilution conditions,[15] giving **10a-10d** (>95% pure after semipreparative RP HLPC). Final deprotection proceeded quantitatively affording the CPPs **3a-3d** (96-98% pure, Table 3.1). The structures were carefully confirmed by ESI-MS, ^{1}H, ^{13}C NMR, and 2D gCOSY spectroscopy.

3.2.2 Cell adhesion assay and competitive solid-phase binding assay (3a-3d)

In vitro experiments were carried out to detect any effects of **3a-3d** on $\alpha_4\beta_1$-mediated cell adhesion, and their selectivity towards $\alpha_4\beta_7$, $\alpha_L\beta_2$, and $\alpha_M\beta_2$ integrins. Although not expressed on leukocytes, $\alpha_5\beta_1$ integrin was also chosen as it shares β_1 subunit with the heterodimer $\alpha_4\beta_1$. Cells were seeded in 96-well plates coated with the specific natural human recombinant ligands (Table 3.1) and allowed to adhere in the presence of increasing concentrations (10^{-10} to 10^{-4} M) of the synthesized CPPs before the determination of the number of adherent cells.

Table 3.1 Effect the LDV CPPs **3a-3d**, and the related **11a**, **11c**, **12a**, **12c**, **13-16**, and reference **BIO1211**, **1a**, **1b**, on integrin-mediated cell adhesion. Data are presented as IC_{50} for antagonists and as EC_{50} for agonists (nM).[a]

Compd. **Sequence**	**Purity (%)** [b]	**FN /Jurkat E6.1** $\alpha_4\beta_1$	**VCAM-1 /Jurkat E6.1** $\alpha_4\beta_1$	**MAdCAM-1 /RPMI8866** $\alpha_4\beta_7$	**Fg/HL60** $\alpha_M\beta_2$	**ICAM-1 /Jurkat E6.1** $\alpha_L\beta_2$	**FN/K562** $\alpha_5\beta_1$
BIO1211 MPUPA-LDVP-OH	-	5.5 ± 4.0 [c] antagonist	4.6 ± 3.0 [c] antagonist	nd	nd	8.4 ± 4.3 [d] antagonist	>5000
1a [e] non peptide, Figure 1	97	15.6 ± 1.5 agonist	13.0 ± 0.8 agonist	>5000	>5000	>5000	>5000
1b [e] non peptide, Figure 1	96	>5000	>5000	>5000	>5000	>5000	9.7 ± 0.5 agonist
3a c[(S)-Phu-LDV-(S)-isoD]	98	50.5 ± 7.8 agonist	35.0 ± 5.9 agonist	31.8 ± 5.5 agonist	>5000	98.2 ± 9.8 agonist	>5000
3b c[(S)-Phu-LDV-(R)-isoD]	97	156 ± 33 agonist	81.8 ± 9.7 agonist	32.1 ± 5.3 agonist	>5000	1110 ± 340 antagonist	>5000
3c c[(R)-Phu-LDV-(S)-isoD]	97	726 ± 28 antagonist	177 ± 57 antagonist	495 ± 89 antagonist	>5000	710 ± 65 antagonist	1950±290 agonist
3d c[(R)-Phu-LDV-(R)-isoD]	99	40.9 ± 4.3 agonist	190 ± 30 agonist	>5000	353 ± 32 antagonist	53.9 ± 5.1 antagonist	>5000
11a *c*[(*S*)-Phu-LDV-(*R*)-β^3Ala]	97	>5000	-	-	-	-	-
11c *c*[(*R*)-Phu-LDV-(*R*)-β^3Ala]	98	>5000	-	-	-	-	-
12a c[(S)-Phu-LAV-(S)-isoD]	95	55.6 ± 2.9 agonist	1.78 ± 0.32 agonist	>5000	53.4 ± 5.4 antagonist	>5000	168 ± 61 agonist
12c *c*[(*R*)-Phu-LAV-(*S*)-isoD]	97	1720 ± 556 agonist	-	-	-	-	-
13 c[(*S*)-Phu-LDV-(*S*)-isoD(NH*Pr*)]	98	>5000	-	-	-	-	-
14 c[(*S*)-Phu-FDV-(S)-isoD]	97	>5000	-	-	-	-	-
15 c[(*S*)-Phu-FAV-(S)-isoD]	98	1061 ± 134 agonist	-	-	-	-	-
16 c[(*R*)-Phu-LD-Phg-(*S*)-isoD]	96	>5000	-	-	-	-	-

[a] $\alpha_4\beta_1$ integrin-mediated cell adhesion was evaluated by assaying Jurkat E6.1 cell adhesion to FN or to VCAM-1; for $\alpha_L\beta_2$ integrin, Jurkat E6.1 cells to ICAM-1; $\alpha_5\beta_1$ integrin, K562 cells to FN; $\alpha_M\beta_2$ integrin, HL60 cells to fibrinogen (Fg); $\alpha_4\beta_7$ integrin, RPMI8866 cells to MAdCAM-1. Values represent the mean ± SD of three independent experiments carried out in quadruplicate. [b] Determined by RP HMLC (General Methods). [c] Reference 10. [d] Reference 9. [e] Compounds previously characterized as integrin agonists; see ref. 2. nd: not determined.

The results of cell adhesion assays are summarized in Table 3.1 and Figure 3.2; the latter reports the heatmaps of adhesion index, a convenient illustration of agonistic or antagonistic behavior of the new synthesized compounds. On the basis of this parameter, an agonist is defined by adhesion index > 1 (displayed in shades of blue), an antagonist by adhesion index < 1 (displayed in shades of orange), and integrin ligands not significantly altering cell adhesion by adhesion index approximately = 1. In addition, concentration-response curves are provided in Appendix (A: Figures S2-S7).

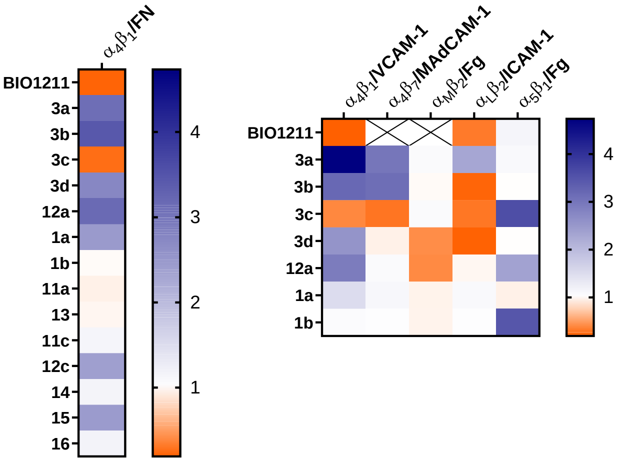


Figure 3.2 Heatmaps of adhesion index: agonist compounds are shown in blue whereas antagonists are displayed in orange. The adhesion index is calculated as the ratio between the number of adhered cells in the presence of the highest CPP concentration (10^{-4} M) and the number of adhered vehicle-treated cells. X: not determined.

To better characterize integrin-ligand interaction, competitive solid-phase ligand binding assays were performed on purified $\alpha_4\beta_1$, $\alpha_4\beta_7$, $\alpha_M\beta_2$, $\alpha_L\beta_2$ and $\alpha_5\beta_1$ integrins, using receptor-specific ligands (Table 3.2) in the presence of increasing concentrations (10^{-10} to 10^{-4} M) of the CPPs.[2]

In the cell adhesion experiments, no significant cell adhesion was observed for bovine serum albumin (BSA)-coated plates (negative control). The reference antagonist BIO1211 inhibited the adhesion of $\alpha_4\beta_1$ integrin expressing Jurkat E6.1 cells to FN and VCAM-1 (IC_{50} 5.5 nM and 4.6 nM, respectively, Table 3.1). In the competitive binding assay on purified $\alpha_4\beta_1$ integrin, BIO1211 confirmed a low nanomolar affinity as previously reported (Table 3.2).[10]

Table 3.2 Binding affinities (IC_{50} Values, nM)[a] of LDV CPPs and BIO1211 on purified integrins.

CPP	$\alpha_4\beta_1$/FN	$\alpha_4\beta_1$/VCAM-1	$\alpha_4\beta_7$/MAdCAM-1	$\alpha_M\beta_2$/Fg	$\alpha_L\beta_2$/ICAM-1	$\alpha_5\beta_1$/FN
BIO1211	8.6 ± 5.1[b]	8.9 ± 3.1	>1000	>1000	5.2 ± 2.1	>1000
3a	43.5 ± 3.5	33.5 ± 4.4	41 ± 7	>1000	83.3 ± 8.7	>1000
3b	133 ± 45	101 ± 35	22.7 ± 6.1	>1000	897 ± 230	>1000
3c	602 ± 32	707 ± 75	183 ± 22	>1000	652 ± 47	193 ± 65
3d	38.2 ± 8.1	28.5 ± 3.1	>1000	244 ± 71	46.7 ± 7.9	>1000
12a	46.1 ± 4.7	41.1 ± 3.1	>1000	47.2 ± 3.1	>1000	203 ± 43
12c	1567 ± 344	>5000	>5000	>5000	>5000	>5000
15	976 ± 168	899 ± 198	>5000	>5000	>5000	>5000
1a [c]	13.3 ± 6.3	10.1 ± 4.9	>5000	>5000	>5000	>5000
1b [c]	>5000	>5000	>5000	>5000	>5000	49 ± 7

[a] IC_{50} values were determined by a competitive solid-phase binding assay to specific ligand (FN for $\alpha_5\beta_1$, VCAM-1 or FN for $\alpha_4\beta_1$, fibrinogen for $\alpha_M\beta_2$, MAdCAM-1 for $\alpha_4\beta_7$ and ICAM-1 for $\alpha_L\beta_2$). [b] Reference 10, Mean ± SD of three independent experiments in triplicate. [c] Compounds previously characterized as integrin agonists, see Ref. 2.

Moreover, previously synthesized and characterized integrin agonists **1a** and **1b** were employed as reference ligands; these compounds were able to increase $\alpha_4\beta_1$- or $\alpha_5\beta_1$-mediated cell adhesion, respectively.[2] As expected, only **1a** increased the adhesion of Jurkat E6.1 cells (Table 3.1), with high affinity towards the isolated integrin (Table 3.2), while **1b** was completely ineffective. Regarding the CPPs, cell adhesion experiments revealed compounds capable to reduce the number of adherent cells promoted by the natural ligands, referred to as antagonists, whereas other ligands increased cell adhesion and therefore were considered to be agonists (Figure 3.2).

The CPPs **3a**, **3b**, and **3d** were able to increase cell adhesion in a concentration-dependent manner (Table 3.1). Remarkably, **3a** showed potency in the nanomolar range (EC_{50}/VCAM-1 35×10^{-9} M, IC_{50}/FN 50.5×10^{-9} M), while **3b** and **3d** displayed a comparatively lower activity (**3b**, EC_{50}/VCAM-1 81.8×10^{-9} M, EC_{50}/FN 156×10^{-9} M; **3d**, EC_{50}/VCAM-1 190×10^{-9} M, EC_{50}/FN 40.9×10^{-9} M), Notably, **3c** was found to be an antagonist with moderate potency (IC_{50}/VCAM-1 177×10^{-9} M, and IC_{50}/FN 726×10^{-9} M).

These results were confirmed by $\alpha_4\beta_1$ affinity evaluation in competitive solid-phase ligand binding: **3a** and **3d** displayed nanomolar IC_{50} values whereas **3b** and **3c** showed a lower affinity for $\alpha_4\beta_1$ (Table 3.2). Binding curves are provided in Appendix (A: Figures S8-S12).

To determine the extent to which experimentally determined binding affinity of CPPs correlates with their potency in modulating integrin-mediated cell adhesion, Pearson (r_P) correlation coefficient was calculated. As regards to $\alpha_4\beta_1$, there was a high positive correlation between binding affinity and FN-mediated cell adhesion potency for all compounds tested ($r_P = 0.9990$, Appendix A: Figure S13), meaning that the highest is the affinity for $\alpha_4\beta_1$ and the highest is the potency in FN-mediated cell adhesion assays.

Notably, regarding the correlation between binding affinity and VCAM-1-mediated cell adhesion potency, a quite low correlation coefficient was determined ($r_P = 0.5920$, Appendix A: Figure S13); for most compounds a correlation was found but some exceptions were identified as those CPPs with the lowest potency for $\alpha_4\beta_1$/VCAM-1 (**3c** and **3d**).

Cell adhesion assays upon different integrin-expressing cell lines were also performed to determine compounds' selectivity (Table 3.1 and Figure 3.2). Nanomolar agonist activity was observed in adhesion experiments with RPMI8866 cells expressing $\alpha_4\beta_7$ integrin to the ligand MAdCAM-1 for the compounds **3a** (EC_{50} 31.8×10^{-9} M) and **3b** (EC_{50} 32.1×10^{-9} M). Ligand binding assays on purified $\alpha_4\beta_7$ integrin confirmed excellent affinity of **3a** and **3b** (Table 3.2). Therefore, they were considered $\alpha_4\beta_1$/$\alpha_4\beta_7$ integrin dual agonists. On the other hand, **3c** was found to be a dual, moderate antagonist of $\alpha_4\beta_1$/$\alpha_4\beta_7$ integrins (for $\alpha_4\beta_7$ integrin, IC_{50} 4.95×10^{-7} M) with a lower affinity (Table 3.2). The reference compounds BIO1211 and **1a** were found to be inactive in the same assays (Tables 3.1, and 3.2), as reported.[2,10]

In the tests for $\alpha_L\beta_2$ integrin, the reference BIO1211 and the CPPs **3b**, **3c** and **3d**, behaved as antagonists with diverse potency in cell adhesion experiments, the most potent among the

CPPS being **3d** (IC_{50} 53.9×10^{-9} M) (Table 3.1). In contrast, **3a** was identified as a potent $\alpha_L\beta_2$ integrin agonist (EC_{50} 98.2×10^{-9} M, Table 3.1). Regarding affinities for isolated $\alpha_L\beta_2$ integrin, BIO1211, **3a** and **3d** showed excellent affinity values, while **3b** and **3c** were able to bind $\alpha_L\beta_2$ integrin with modest affinity (Table 3.2). As described for $\alpha_4\beta_1$, a very high positive correlation between ligand binding affinity and cell adhesion potency was observed also for $\alpha_L\beta_2$, being the correlation index 0.9969 (Appendix A: Figure S13). This means that compounds with a low potency towards $\alpha_L\beta_2$ is able to bind it with a low affinity, and vice versa.

Concerning $\alpha_M\beta_2$ integrin, the only modestly active compound was the antagonist **3d**, with an IC_{50} of 3.53×10^{-7} M in the integrin-mediated cell adhesion assay (Table 3.1). In terms of affinity for purified $\alpha_M\beta_2$ integrin, **3d** showed moderate binding with an IC_{50} of 2.44×10^{-7} M (Table 3.2).

Finally, while **3c** showed a scarce but measurable agonistic activity towards $\alpha_5\beta_1$ integrin (EC_{50} 1.95×10^{-6} M), BIO1211, **3a**, **3d**, and **3b** were found to be inactive (Table 3.1) and not able to bind to isolated $\alpha_5\beta_1$ integrin (Table 3.2). Not unexpectedly, **1b** was a potent agonist of this integrin with nanomolar affinity as previously reported (Tables 3.1 and 3.2).[2]

3.2.3 Synthesis of CPPs (11a-c, 12a-c, 13-16)

To better distinguish the pharmacodynamic role of the two carboxylate groups and of some relevant side-chains in receptor binding and in determining agonism or antagonism behavior, the most potent agonist **3a** and the antagonist **3c** were selected for modifications.

The CPP **3a** was modified either by replacing isoAsp^5 with *(R)*-β^3-homoAla, giving c[*(S)*-Phu-LDV-*(R)*-βAla^5] (**11a**), or by replacing Asp^3 with Ala, giving the peptide c[*(S)*-Phu-LAV-*(S)*-isoAsp^5] (**12a**). Topologically, the *(R)* configuration of β^3-homoAla corresponds to the *(S)* configuration of isoAsp (Figure 3.3). Alternatively, the isoAsp^5 carboxylate side chain in **3a** was derivatized to the corresponding propylamide, giving **13**, or the Leu^2 in **3a** was replaced with aromatic Phe, yielding **14**. Peptide **15** was further modified based on **12a** by replacing Leu^2 with Phe.

Similarly, the structure of **3c** was modified by replacing isoAsp^5 with *(R)*-β^3-homoAla, giving c[*(R)*-Phu-LDV-*(R)*-βAla^5] (**11c**), or Asp^3 was replaced with Ala, giving c[*(R)*-Phu-LAV-*(S)*-isoAsp] (**12c**). The *(R)* configuration of β^3-homoAla corresponds to the *(S)* configuration of isoAsp (Figure 3.3). Alternatively, the Val^4 in parent **3c** was substituted with an aromatic Phenylglycine (Phg), yielding **16**.

The CPPs were prepared from the linear precursors **9e-9l** (Appendix A: Table S1) as reported for **3a-3d**. To this purpose, Fmoc-(*R*)-β^3-homoAla-OH was synthesized by adapting a procedure reported in the literature;[16] Fmoc-Asp-propylamide was readily prepared from Fmoc-(*R*)-Asp(O^{t}Bu)-OH and *n*-propylamine. Subsequently, cyclization under pseudo-high-dilution conditions afforded **10d-10l** (Appendix A: Table S1); the CPPs **11a**, **11c**, **12a**, **12c**, **13-16** were obtained after final deprotection. The structures were well-identified by ESI-MS and NMR spectroscopy. For all the detailed chemical sketches of CPPs, see Figure 3.3.

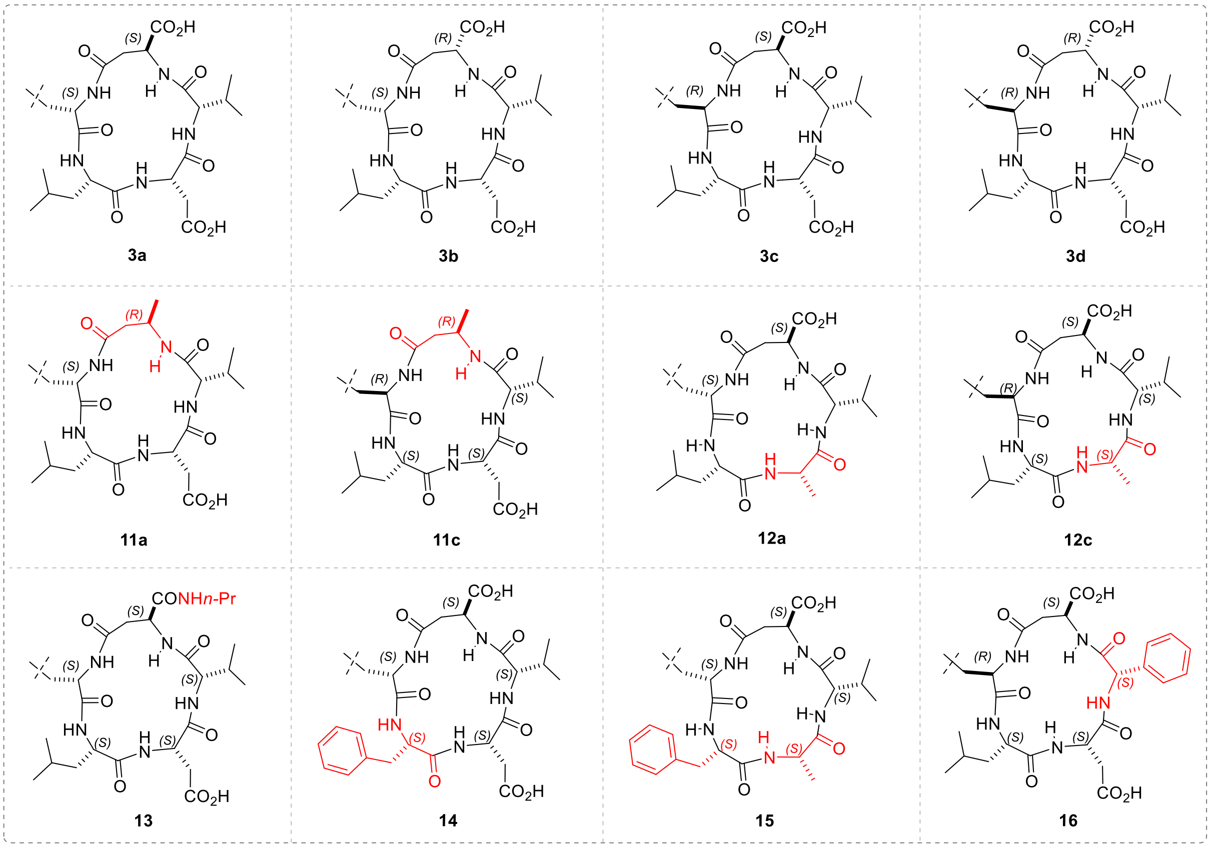


Figure 3.3 Sketches of the CPPs **3a-3d**, and the related **11a**, **11c**, **12a**, **12c**, and **13-16**, the mutated residues are shown in red. For simplicity, part of Phu is omitted.

3.2.4 Cell adhesion assay and competitive solid-phase binding assay (11a-c, 12a-c, 13-16)

The effects of the new CPPs derived from **3a** and **3c** on the adhesion of $\alpha_4\beta_1$ integrin-expressing Jurkat E6.1 cells to the ligand FN were assayed as discussed above (Table 3.1 and Figure 3.2). Apparently, the replacement of the isoAsp^5 with β^3-homoAla in both **3a** and **3c** was not tolerated for the activity towards $\alpha_4\beta_1$ integrins, since **11a** and **11c** became inactive in Jurkat E6.1 cell adhesion assay ($IC_{50} > 5000$ nM, Table 3.1). In a similar way, the derivatization of isoAsp^5 carboxylate into amide in peptide **13** led to a complete loss of activity (Table 3.1).

In contrast, the substitution of Asp^3 by introduction of Ala to give **12a**, **12c** was much better tolerated, albeit **12c** showed a decrease of activity as compared to the parent **3c** (EC_{50} 1.72×10^{-6} M vs 7.26×10^{-7} M). Similar results were confirmed by binding affinity towards purified $\alpha_4\beta_1$ integrin (Table 3.2). Besides, the moderate antagonist behavior of **3c** was converted to agonism in **12c** (see also Computational studies and Appendix A: Figure S20). Intriguingly, CPP **12a** maintained the nanomolar agonist activity of the parent **3a** (EC_{50} 55.6×10^{-9} M, Table 3.1) and excellent binding affinity (IC_{50}/VCAM-1 41.1×10^{-9} M, IC_{50}/FN 46.1×10^{-9} M, Table 3.2).

The CPPS **14-16** showed very modest or null activity in the cell adhesion assay (Table 3.1), with only **15** giving a measurable increase of cell adhesion (EC_{50} 1.06 μM, Table 3.1), and micromolar affinity for the isolated receptor (Table 3.2), confirming the importance of residues Leu and Val.

Further cell adhesion assays using cell lines expressing different integrins and competitive solid-phase binding assays on purified integrins were performed to better characterize the activity of **12a** (Table 3.1), while the other CPPs were neglected, for the scarce to null activity towards $\alpha_4\beta_1$ integrins.

The CPP **12a** showed significantly improved potency as compared to **3a** in the adhesion of Jurkat E6.1 cells to VCAM-1, with an outstanding EC_{50} 1.78×10^{-9} M (Table 3.1). Notably, while **3a** was a dual agonist of $\alpha_4\beta_1/\alpha_4\beta_7$ integrins with similar potency (Table 3.1), **12a** completely lost any activity and binding ability for $\alpha_4\beta_7$ integrins (Tables 3.1 and 3.2, see also the Computational studies and Appendix A: Figure S18). On the other hand, **12a** resulted to be inactive towards $\alpha_L\beta_2$ integrin (Tables 3.1 and 3.2), while becoming a modest agonist for $\alpha_5\beta_1$ integrin (EC_{50} 1.68×10^{-7} M, Table 3.1), with affinity in the submicromolar range for the isolated integrin (Table 3.2). Finally, **12a** was able to bind to and activate $\alpha_M\beta_2$ integrin as antagonist, with an interesting IC_{50} in the nanomolar range (IC_{50} 53.4×10^{-9} M, Table 3.1) and noteworthy nanomolar affinity (IC_{50} 47.2×10^{-9} M, Table 3.2).

3.2.5 Effects of the CPPs on integrin-mediated ERK1/2 phosphorylation

To confirm the agonist or antagonist behavior, the effect of the reference compound BIO1211, **3a-3d** and **12a** on phosphorylation of ERK1/2 (extracellular signal-regulated kinases 1 and 2) in Jurkat E6.1 cells was determined (Figure 3.4). Intracellular signaling generated by the interaction of ECM components with $\alpha_4\beta_1$ integrin produces an increase in the phosphorylation of cytoplasmatic second messengers such as ERK1/2 that contribute to α_4 integrin-mediated cell functions.

The endogenous ligand FN (10 μg/μL), employed as positive control, induced a significant increment of ERK1/2 phosphorylation in comparison to vehicle-treated Jurkat E6.1 cells (Figure 3.4 A). The reference compound BIO1211 (10^{-7}-10^{-9} M), which is defined as a potent α_4 integrin antagonist, significantly prevented ERK1/2 activation induced by FN (Figure 3.4 B).

Similarly to BIO1211, the CPP **3c** (10^{-7}-10^{-9} M) significantly reduced FN-induced intracellular signaling activation, confirming to act as an antagonist (Figure 3.4 A). To further confirm the antagonist behavior, **3c** was administered alone to Jurkat E6.1 cells. In this experimental setting, **3c** did not influence ERK1/2 activation (Figure 3.4 B), thus probably binding to $\alpha_4\beta_1$ without inducing its activation and the resulting downstream signaling.

In contrast, a significant concentration-dependent increase of ERK1/2 phosphorylation was produced by the $\alpha_4\beta_1$ agonists **3a** (Figure 3.4 A), **3b**, **3d**, and **12a** (Figure 3.4 C), confirming their ability to bind the receptor and to induce its activation.

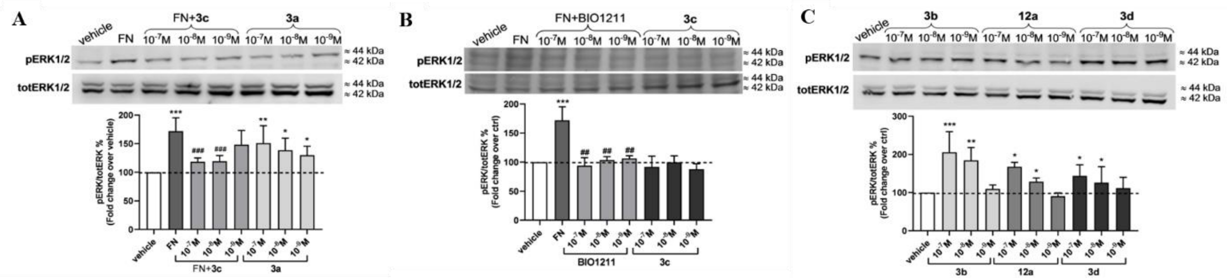


Figure 3.4 Effects of FN (10 µg/mL), the reference compound BIO1211 and **3a-3d**, **12a** on ERK1/2 phosphorylation mediated by $\alpha_4\beta_1$ integrin expressed on Jurkat E6.1 cells. (A, B) The antagonists BIO1211 and **3c** were able to prevent ERK1/2 phosphorylation induced by FN. The antagonist **3c**, administered alone to Jurkat E6.1 cells, did not modify phosphorylation levels of ERK1/2. On the contrary, the agonists **3a** (A), and **3b**, **3d**, **12a** (C), induced ERK1/2 activation in a concentration-dependent manner. Representative western blot shows that Jurkat E6.1 cells plated on FN had a stronger signal for pERK1/2 than vehicle-treated cells (vehicle). The graphs represent densitometric analysis of the bands (mean ± SD; 3 independent experiments); the amount of pERK1/2 is normalized to that of totERK1/2. * $p<0.05$, ** $p<0.01$, *** $p<0.001$ vs vehicle; ## $p<0.01$, ### $p<0.001$ vs FN (Newman-Keuls test after ANOVA).

3.2.6 In vitro enzymatic stability of 3a, 3c

To estimate any increase in enzymatic stability conferred by the α/β hybrid cyclic structure,[17] the representative **3a** and **3c** were incubated in mouse serum in comparison to the reference antagonist BIO1211 (Appendix A: Figure S1). Consistent with other reported studies,[6,7] BIO1211 was found to be poorly stable when added to mouse serum, being almost completely hydrolyzed after 2 h, as determined by RP HPLC analysis. In contrast, **3a** and **3c** appeared significantly more stable, and after 3 h, the remaining amount was estimated at >85%.

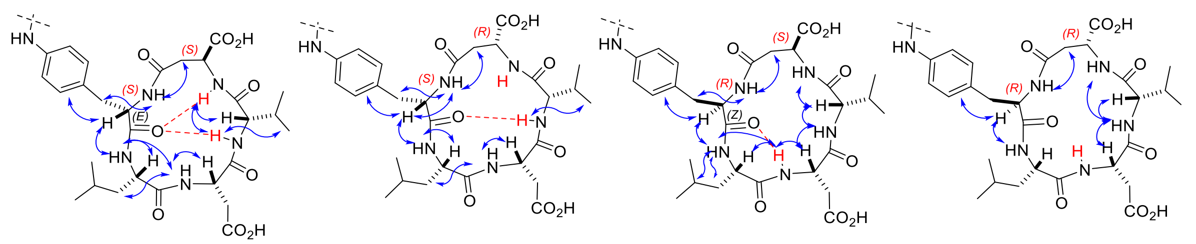
3.2.7 Conformational analysis of the CPPs

Apparently, the LDV CPPs **3a-3d** showed diverse integrin affinities and cell adhesion effects to ligand-coated plates, albeit differing only in the absolute configuration of the residues Phu and/or isoAsp. This suggested that the overall geometry exerts a clear impact on ligand-receptor interactions and binding. Hence, we analysed the 3D conformations of **3a-3d** in solution by NMR spectroscopy and molecular dynamics (MD) simulations.

The NMR analysis was conducted in 8:2 mixtures of DMSO-d_6/H_2O, a highly viscous solvent system recommended as an excellent biomimetic environment.[18,19] For each peptide, ^{1}H-NMR spectra showed a single set of resonances, indicating conformational homogeneity or a rapid interconversion between the conformers. gCOSY analyses allowed the unambiguous assignment of the resonances. Variable temperature (VT) ^{1}H-NMR experiments were used to determine if the amide protons were plausibly involved in intramolecular hydrogen bonding or were solvent exposed (Appendix A: Table S2).[20]

The analyses of the experimental $\Delta\delta/\Delta T$ (ppb K^{-1}) parameters (Appendix A: Table S2) suggest the occurrence of strong hydrogen bonds for Val^4 NH and isoAsp^5 NH in **3a**, while a strong hydrogen bond was supposed for Asp^3 NH in **3c** and **3d** (Figure 3.5).

2D ROESY analyses were performed in the same solvent system (DMSO-d_6:H_2O, 8:2). Cross-peaks intensities were ranked to infer plausible interproton distances (Figure 3.5, and Appendix A: Tables S3-S7).



3a, (*S*)-Phu, (*S*)-isoAsp **3b**, (*S*)-Phu, (*R*)-isoAsp **3c**, (*R*)-Phu, (*S*)-isoAsp **3d**, (*R*)-Phu, (*R*)-isoAsp

Figure 3.5 Sketches of the structures of the cyclic LDV peptides **3a-3d** showing meaningful proton-proton correlations indicated by arrows, as determined by 2D ROESY in DMSO-d_6/H_2O. The amide protons characterized by low |Δδ/ΔT| values (Appendix A: Table S2) are shown in red; predicted hydrogen bonds are shown as red dashed line. For clarity, part of the Phu moiety has been omitted.

The estimated distances were analyzed by simulated annealing and restrained MD simulations, using the AMBER force field[21] in explicit water. In brief, random geometries of each peptide were sampled during a high-temperature unrestrained MD simulation in a box of TIP3P models of equilibrated water molecules.[22] For each random structure, the interproton distances deduced by ROESY were introduced as constraints. As the absence of Hα(i)-Hα(i+1) cross-peaks reasonably excludes the occurrence of *cis*-peptide bonds, the amide bonds angles (ω) were set at 180°.

The structures were subjected to restrained high-temperature simulation with a scaled force field, followed by a period with full restraints, then the system was slowly cooled. The resulting structures were minimized, and the backbones of the structures were clustered by rmsd analysis. For all compounds, this procedure gave one major cluster comprising the large majority of the structures.

The representative structures with the lowest energy and the least number of restraint violations were selected and analyzed. The ROESY-derived structures of **3a** and **3b** (Figure 3.6) show explicit hydrogen-bonds as predicted by VT-NMR analysis. Peptide **3a** is characterized by a clear type II β-turn (βII) centered on Leu^2-Asp^3. In **3b**, Leu^2-Asp^3 appeared to be embedded within an inverse type II β-turn (βII'), plausibly due to the reversal of stereochemistry of the β-residue isoAsp^5. The structures of **3c** and **3d** show similar overall geometries, each showing an inverse γ-turn (γ') centered on Leu^2 (Figure 3.6).

To investigate the dynamic behavior of the LDV CPPs, the structures were analyzed by unrestrained MD simulations at 298 K in a box of explicit TIP3P equilibrated water molecules. During the simulations, the structures of the backbones were maintained, indicating that these conformations plausibly represented stable minima (data not shown).

The secondary structure elements observed for the α/β hybrid **3a-3d** were foreseeable; indeed, β-amino acids are well-known to favor defined secondary structures when introduced in CPPs.[23] These residues exert a significant conformational bias on backbone conformations, and preferably adopt a pseudo-γ-turn at the central position and tend to stabilize γ-turn secondary structures at the opposite side of the macrocycle.[24]

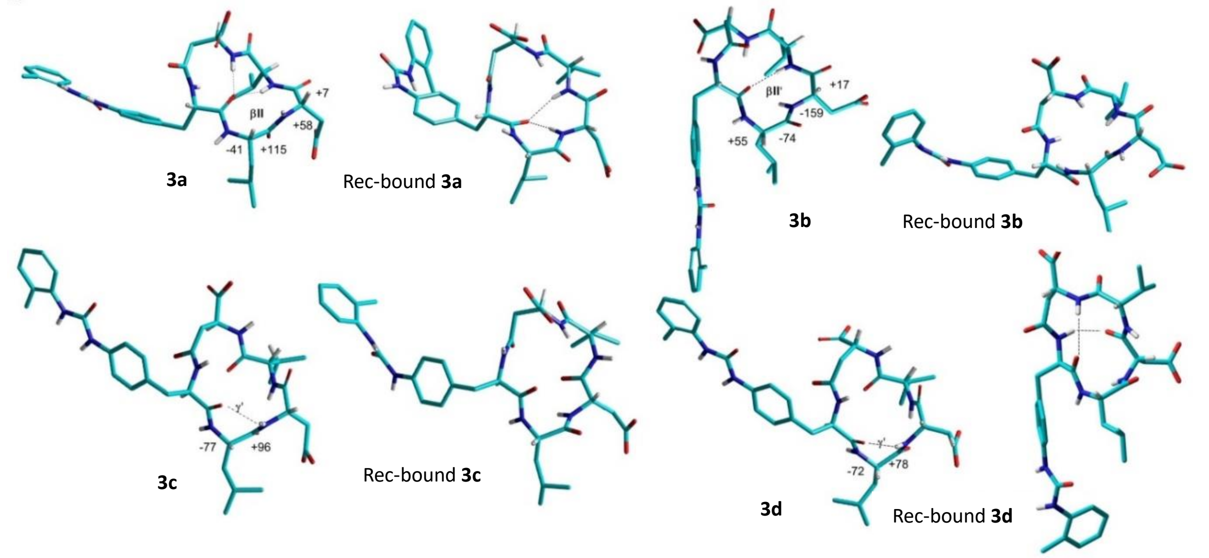


Figure 3.6 Representative lowest energy structures of peptides **3a-3d**, calculated by ROESY restrained MD. The receptor-bound poses as predicted by molecular docking (see Computational studies) are shown for comparison.

As for the other related CPPs, the comparison of the ^{1}H NMR spectra support that **11a** and **13-15** maintain conformations similar to the parent compound **3a**, since the chemical shifts of the resonances for the unaltered residues were practically the same (Appendix A: Figures S22-S33). Similarly, the compounds **11c**, **12c** and **16** showed NMR spectra comparable to the parent compound **3c**. VT-NMR analysis showed for all CPPs the same trends of Δδ/Δt parameters, suggesting that the hydrogen-bonding patterns and secondary structure elements were maintained (Appendix A: Table S2).

Unexpectedly, **12a** displayed differences respect to **3a** in the ^{1}H NMR spectra relative to the resonances of Leu and Phu. In particular, Phu^1 NH and Leu^2 NH in **12a** appeared downfield ($\delta = 9.1$ and 10.6 ppm, respectively), as compared to the parent peptide **3a** ($\delta = 8.2$ and 8.6 ppm, respectively, Figure 3.7 A). Besides, VT-NMR analysis (Appendix A: Table S2) showed for Leu^2 NH an atypical positive Δδ/ΔT (+1.6 ppb K^{-1}). As a consequence of the NMR evidence, the structures **11a**, **11c, 12c**, **13-16**, were not investigated further, while the 3D structure of **12a** in solution was analyzed by 2D ROESY analysis and restrained MD, as reported for **3a-3d**. Eventually, this procedure confirmed that **12a** still maintains a similar conformation of **3a** (Figure 3.6 vs Figure 3.7 B). Possibly, the diverse chemical fields for Phu, Leu resonances might be due to peculiar deshielding effects exerted e.g. by the urea group, rather than to the occurrence of different overall 3D geometries.

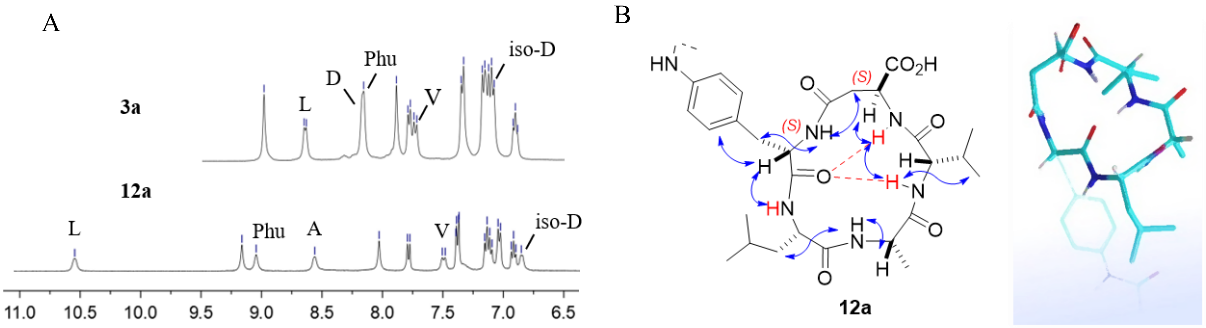


Figure 3.7 Conformational analysis of **12a**. Meaningful proton-proton ROESY correlations are indicated by arrows; the amide protons characterized by low $|\Delta\delta/\Delta T|$ values are shown in red; hydrogen-bonds are shown as red dashed lines. (A) Comparison of the amide-NH regions for **3a** and **12a**. (B) Representative lowest energy structure for **12a**, calculated by ROESY-restrained MD in a 30×30×30 Å box of standard TIP3P water molecules; the MPUPA moiety is rendered in sticks for clarity.

3.2.8 Computational studies

The mechanism by which an agonist such as **3a** is able to increase, while the antagonist **3c** decreases the adhesion of the receptor to the native ligands, appears particularly puzzling. As reviewed in Chapter 2, very few studies have been dedicated to the leukocyte integrin agonists.[25-27] One example was the dual agonist 7HP349, capable to enhance the adhesion of cultured cell lines expressing $\alpha_4\beta_1$ integrin to the natural ligands. Docking of this agonist into the homology model of $\alpha_4\beta_1$ indicated that the ligand binds at a site that partially overlaps the ligand binding pocket. Further experiments showed that 7HP349 did not induce the binding of ligand-induced binding site Abs, which bind the β subunit, supporting a mode of binding in which 7HP349 does not engage the metal ion-dependent adhesion site and does not act like a ligand mimetic.[27] For the CPPs **3**, whose activities are sensitive to the replacement of AspCOO$^-$ that expected to coordinate with Mg^{2+}, another model must be considered.

To investigate the structural elements at the basis of the agonist or antagonist behavior, molecular modeling of the prototypic **3a** and **3c** was performed with Autodock 4.0.[28] In addition, the analysis was extended to the stereoisomers **3b**, **3d**, and to the derivatives **12a**, **12c** and **15**. These CPPs have been selected for their at least measurable affinity for isolated integrin, and clear effects on the adhesion of Jurkat E6.1 cells to the natural ligands (Tables 3.1-3.2).

Simulations of $\alpha_4\beta_1$ integrin are particularly challenging since the precise structure of this integrin is not yet available. In addition, molecular mechanics force fields generally utilized to analyze ligand-receptor interactions lack in descriptions of the highly directional nature of metal coordination. For this reason, the region containing the ligand, MIDAS and the receptor residues in the binding site were treated by hybrid Density-Functional Theory (DFT) combined QM/MM calculation. The $\alpha_4\beta_1$ integrin receptor model (Appendix A: Figure S16) was obtained by combining the crystal structures of the α_4 subunit (PDB ID: 3V4V, Crystal structure of $\alpha_4\beta_7$ headpiece complexed with Fab ACT-1 and RO0505376)[29] and of the β_1 subunit (PDB ID: 4WK4, $\alpha_5\beta_1$, Metal ion and ligand binding of integrin).[30]

The receptor is expected to coordinate a carboxylic group of the ligands through the Mg^{2+} ion of the MIDAS in the αI and βI domains.[29,30] Other metal ion binding sites close to MIDAS are present, i.e. Ca^{2+} ions coordinated by residues in the ADMIDAS, and SyMBS, in the βI domains. SyMBS and ADMIDAS have important roles in regulating ligand binding affinity. In β_1 integrins, the ADMIDAS seems to be a negative regulatory site responsible for integrin inhibition by high concentration of Ca^{2+}, and for activation by Mn^{2+}.[30]

In the resulting model of $\alpha_4\beta_1$ integrin, the plausible ligand-binding pocket appears characterized by a long binding groove at the α/β interface, as reported for $\alpha_4\beta_7$ integrin (Appendix A: Figure S14).[29] This shape is clearly different from that of the Arg-Gly-Asp (RGD)-binding integrins $\alpha_v\beta_3$[31] or $\alpha_{IIb}\beta_3$.[32] In particular, the α_4 subunit completely lacks in the cavity deputed to hosts Arg. Besides, the comparison between the β_1 and β_3 subunits reveals that the former contributes to expand the binding pocket of $\alpha_4\beta_1$ integrin on the β subunit side, since the residues Arg^{214} and Arg^{216} in β_3 subunit are replaced with Gly^{217} and Leu^{219} in β_1.

The best binding conformations of **3a** and **3c** are shown in Figure 3.8, along with **3b** and **3d**, for comparison. The interactions have been analyzed with BIOVIA DSV2021 and with PacVIEW tool in PacDOCK web server.[33] For brevity, herein only the most relevant features of the complexes are discussed; all specific stabilizing interactions and alternative views are discussed in the Appendix A. The calculated poses of **12a**, **12c**, and **15** (Appendix A: Figures S18-S21), and the detail of RO0505376 (Appendix A: Figure S14) into the binding site of $\alpha_4\beta_7$ headpiece (PDB 3V4V), are shown in the Appendix A. The calculated ΔG_{bind} nicely fit the experimental affinities for the isolated integrin: (kcal mol^{-1}) **3a**, -15.81; **3b**, -14.08; **3c**, -15.08; **3d**, -14.72; **12a**, -15.28; **12c**, -12.84; **15**, -13.53.

All CPPs appear to occupy the same location into the crevice between the subunits, in proximity of the MIDAS center. With the only exception of **3d**, within the binding site the Phu^1-LDV-$isoAsp^5$ sequence of all CPPs can be read in a clockwise direction (Figure 3.8). Interestingly enough, for all CPPs but **3d** the coordination to Mg^{2+} in the MIDAS of the β_1 subunit involves the carboxylate side chain of $isoAsp^5$. For the prototypic **3a** and **3c**, this is in line with the experimental observation described above that $isoAsp^5$ carboxylate rather than that of Asp^3 was strictly necessary for receptor binding (**3a** vs **11a**, **12a**, **13**; **3c** vs **11c**, **12c**).

As anticipated, in the docked pose of the agonist **3a**, c[*(S)*-Phu^1-LD^3V-*(S)*-$isoAsp^5$], $isoAsp^5$ COO^- is coordinated to Mg^{2+} in MIDAS, while Asp^3 COO^- interacts with Lys^{157} $NH\zeta^+$ (α_4 subunit) by salt bridge. The large majority of the stabilizing interactions of **3a** involve residues of the α subunit (Figure 3.8, and Appendix A: Figure S17). The aryl rings of Phu^1 lean against the residues Tyr^{187} and Phe^{214} (α_4), and urea C=O is hydrogen-bonded to Lys^{213} NH. The branched isopropylmethyl side chain of Leu^2 finds place in the upper hydrophobic pocket of the α/β-groove, delimited by Leu^{225}, Tyr^{187}, Lys^{157}, all belonging to the α_4 subunit, a cavity which is not utilized by RO0505376 (Appendix A: Figure S14).[29] Val^4 adopts a pseudo axial disposition, perpendicular to the macrocycle plane, making no relevant interactions.

While the antagonist **3c**, c[*(R)*-Phu^1-LD^{3}V-*(S)*-isoAsp^5], diastereoisomer of **3a** for the reversal of configuration at Phu^1, shows fewer interactions with the α subunit, compensated by tight interactions with residues of the β subunit (Figure 3.8, and Appendix A: Figure S17). As for **3a**, isoAsp^{5}COO$^-$ in **3c** is coordinated to Mg^{2+} in MIDAS. Of particular interest is the ionic bond of Asp^{3}COO$^-$ in **3a** with Lys^{182} $NH\zeta^+$ (β_1), an interaction which pulls the CPP scaffold against the β_1 subunit (see for comparison **12c**, c[*(R)*-Phu-LAV-*(S)*-isoAsp], Appendix A: Figure S20). This is in sharp contrast to **3a**, in which Asp^3 interacts with Lys^{157} $NH\zeta^+$ (α_4). The pose of Phu^1 is stabilized by interactions with Phe^{321} (π-π staking), Ser^{227} and Asp^{226} (β_1). Interestingly, Val^4 is in contact with Tyr133 (β_1), as shown in Figure 3.8.

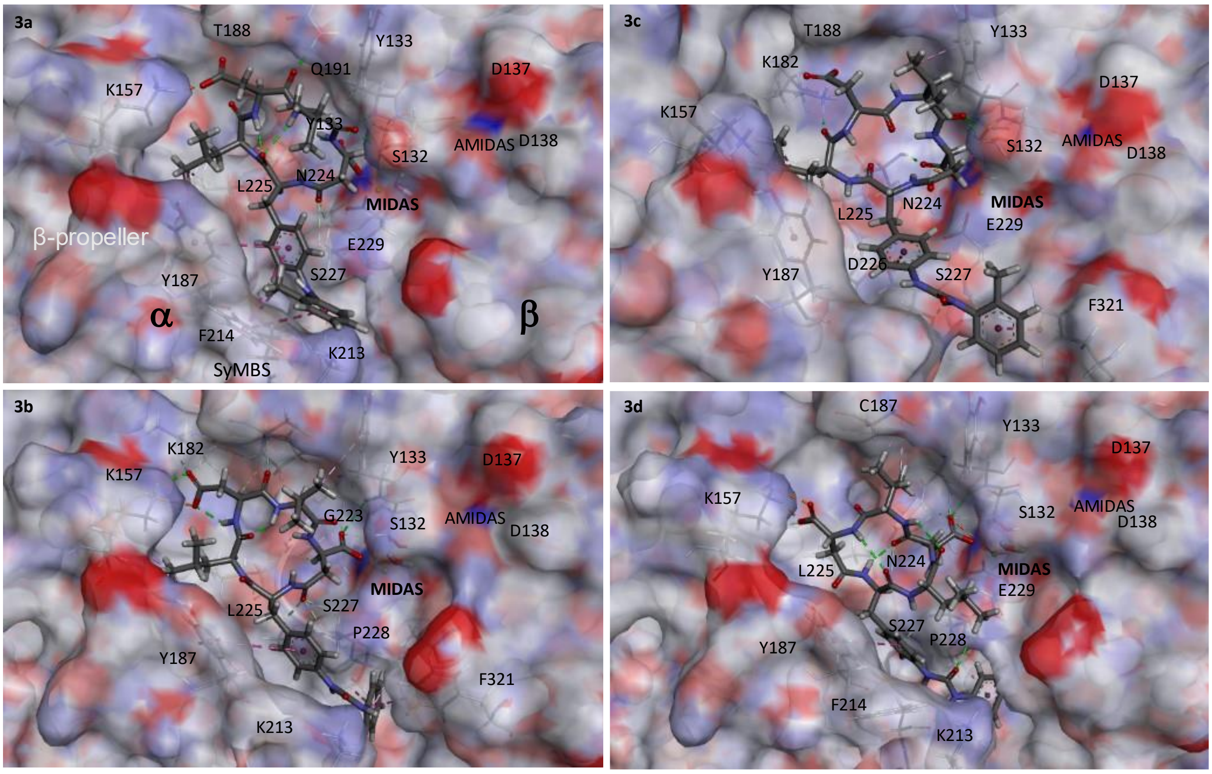


Figure 3.8 Calculated binding conformations of **3a-3d** (right) within the $\alpha_4\beta_1$ integrin binding site. Ligands are rendered in stick and colored by atoms. The integrin binding site is represented by its partially transparent, solid solvent-accessible surface, colored by the atomic interpolated charge. Key receptor residues are represented in tiny stick and non-bonding interactions are indicated as dashed lines. Images were obtained using BIOVIA DSV2021.

The CPP **3b**, c[*(S)*-Phu^1-LDV-*(R)*-isoAsp], differs from **3a** for the inversion of the stereochemistry of isoAsp^5, thus producing a rearrangement of the interactions around MIDAS. Clearly, **3b** shows more balanced interactions with both subunits (Figure 3.8). As for **3a**, isoAsp^{5}COO$^-$ is coordinated to Mg^{2+} (MIDAS), and Asp^{3}COO$^-$ forms a salt bridge with Lys^{157} $NH\zeta^+$ (α_4). Val^4 isopropyl makes some contacts with Ser^{134} (β_1) and Tyr^{133} (β_1). Phu^1 interacts with residues of the α_4 subunit, i.e. Tyr^{187}, Lys^{213}, as well as residues of the β_1 subunit, Ala^{260}, Phe^{321}, and Pro^{228}.

As for **3b**, also **3d** seems to lean against residues of both subunits alike (Figure 3.8). Albeit the docked structure of **3d** occupies the same cleft, the pentapeptide ring appears turned over as compared to the other CPPs. The **3d** c[*(R)*-Phu^1-LD^{3}V-*(R)*-isoAsp^5] sequence can be read in anticlockwise direction within the binding site, upon 180° rotation along an axis passing through Val^4 and Phu^1, so that these residues maintain the same positions. However, because of the rotation, Val^4 adopts a pseudo-equatorial position, in tight contact with Cys^{187} (β_1). The rotation also produces the swap between isoAsp^5 and Asp^3, therefore Mg^{2+} in MIDAS is coordinated to the carboxylate of Asp^3, while isoAsp^5 carboxylate forms a salt bridge with Lys^{157} NHζ^+(α_4). Plausibly, this alternative disposition of the macrolactam ring is dictated by the reversal of configuration at both Phu^1 and isoAsp^5 residues. As for Phu^1, this residue is in contact with Tyr^{187}, Phe^{214}, Lys^{213} of the α_4 subunit, and with Pro^{228}, Ser^{227} of the β_1 subunit.

Concerning the calculated poses of **12a**, **12c** and **15**, these appear similar to those of the parent peptides **3a** and **3c** (Appendix A: Figures S18-S21). Also for these derivatives, the trend of theoretical binding ΔGs is nicely consistent to the experimental binding affinities.

The in-solution and the bioactive conformations of **3a-3d** are presented in Figure 3.6. The inspection of the structures supports the utility of the α/β hybrid CPP scaffolds as conformationally stable probes for investigating integrin binding in the absence of the crystal structure of the receptor.[34] Indeed, the overall geometries are generally maintained at the receptor, with minor differences. For instance, the receptor-bound structure of **3a** shows the intramolecular hydrogen-bond between Phu^1 C=O and Val^4 NH as observed in solution, and a second hydrogen-bond between Phu^1 C=O and Asp^3 NH (Figure 3.6). More pronounced differences can be perceived for **3d**.

For all CPPs, in the bioactive conformation the diphenylurea moiety resides into the lower side of the longitudinal cleft between the α and β subunits, consistent to the specificity of MPUPA for α_4 integrins.[4] This observation aligns with the docking poses reported by da Silva et al.,[35] as discussed in Chapter 2. In the calculated pose for BIO1211, the peptide adopts a reverse *S*-shape, spanning across the interface between the α_4 and β_1 subunits. The N-terminal MPUPA is allocated within the lower side of the α/β groove, as observed for the CPPs (Figure 2.11, BIO1211 and Figure 3.8, CPPs), while the C-terminal Pro is positioned on top of the groove. These authors predicted the interaction of AspCOO$^-$ with the divalent cation in MIDAS. Val occupies the same position as seen for the CPPs, but its position is pseudo-equatorial, so that the branched isopropyl points against the β_1 subunit. Additionally, the side chain of Leu is directed towards the β_1 subunit, and the LDVP sequence presents itself in anticlockwise direction. In this perspective, the geometry of BIO1211 reported by da Silva et al. seems having something in common with the docked pose of **3d**, rather than those of **3a-3c** and the other CPPs, which display the LDV sequence in a clockwise manner.

In a separate docking computation performed by Sartori et al.[25] for BIO1211 and cyclopeptide c[Amp(MPUPA)LDVG] (Figure 2.11), alternative results were obtained, with the

MPUPA group occupying the upper subsites, which differs from our CPPs (Figure 2.11 vs Figure 3.8). However, the LDV sequence was oriented in a clockwise direction, consistent with our observations in **3a**-**3c**, but not for **3d**.

It is important to reiterate that, the reported receptor models for molecular docking may exhibit significant variability. Consequently, it is not entirely surprising that structurally related compounds (c[Amp(MPUPA)LDVG] vs our CPPs **3**), or even the same compound (BIO1211, conducted by Sartori and by da Silva et al., respectively, see Figure 2.11), have resulted in diverse receptor-bound structures, as discussed in Chapter 2 and throughout this section.

With all due caution, the computations with our homology receptor model, which combined hybrid DFT and QM/MM calculations, aroused some structural speculations. Despite a certain similarity, the predicted receptor-bound poses of the most potent agonist **3a** and the antagonist **3c** indeed show some differences, possibly responsible for the alternative behavior of the two compounds in the integrin-mediated cell adhesion to the natural ligands.

In summary, the macrocycle of **3c** appears flattened into the binding site within the propeller and the βI-domain on the integrin head, making many contacts with the β_1 subunit. The computations support the role of the ionic bond Asp^3 COO$^-$-Lys^{182} NHζ^+ (β_1), an interaction which pulls the CPP scaffold against the β_1 subunit, in determining antagonism (Appendix A: Figure S20). Indeed, the substitution of Asp^3 for Ala transformed the antagonist **3c** into the modest agonist **12c** (Tables 3.1-3.2). Besides, the simulations highlight the role of the aryl rings of Phu^1 in the interactions of **3c** with residues adjacent to Glu^{229}, a key residue of the β_1 subunit which belongs to the coordination sphere of both MIDAS and SyMBS.

The characteristics of complexes shown in Figure 3.8 can be interpreted in light of the crystallographic data reported for RGD-binding integrins. In general, it has been established that the mechanism of extension and activation requires a specific reorganization of pre-existing interaction networks around Tyr^{133} in the β1-α1 loop of the β subunit, in the proximity of the ligand recognition site.[36,37] In this perspective, antagonism by **3c** might be the result of the combined compacting effects of Phu^1, that clings to elements of MIDAS and SyMBS, and the bulky isopropyl group of **3c**, that packs against the Tyr^{133}, therefore freezing domain translocation and hinge opening. Consequently, the transmission of the activation signal through α_7-helix downward movement and the hybrid domain swing out in β_1 cannot occur.[38]

On the other hand, the opposite absolute configuration at Phu^1 forces **3a** to log into the binding site lopsided (Appendix A: Figure S15), making fewer contacts with elements of the β_1 subunit. In particular, Asp^3 COO$^-$ makes a salt bridge with Lys^{157} NHζ^+ (α_4), and Phu^1 is in contact only with residues of the α_4 subunit. Nor Tyr^{133} aryl ring or other residues of the β1-α1 loop are tightly packed against the ligand, giving room to the dislocation of β1-α1 loop of the βI domain necessary for receptor activation.

Also **12a**, which shares the same stereochemistry array of **3a**, maintains the lopsided orientation within the receptor, therefore having few contacts with the β_1 subunit (Appendix A:

Figures S18-S19), and indeed proved itself to be a good promoter of cell adhesion (Table 3.1). The other CPPs adopt bioactive conformations which are intermediate between the flat **3c** and the lopsided **3a**, in general making interactions with both subunits, plausibly accounting for their inferior agonist effects (Table 3.1).

Interestingly, in the $\alpha_4\beta_1$-**3a** complex, the distance between the cations at MIDAS and ADMIDAS appears slightly increased by around 0.8 Å as compared to the $\alpha_4\beta_1$-**3c** complex. This seems in contrast to the crystallographic evidence for β_3 integrin. In the inactive conformation, the latter shows an acutely bent conformation. During agonist-induced headpiece opening, movements occur mainly in the β_3 subunit, and the distance between $\beta1$-$\alpha1$ loop elements and the α subunit decreases.[37] The interaction of ligand's carboxylate with MIDAS seems to be necessary for receptor activation, while pulling by the α subunit may be not fundamental.[39] During the conformational transition the ADMIDAS experiences a noteworthy movement towards MIDAS of 3.9 Å, as discussed in Chapter 2.[37]

On the other hand, a moderate increase of the distance between MIDAS and ADMIDAS, as calculated for the agonist **3a**, might make more sense for β_1 integrins. Unlike the resting structures of β_3 integrins, $\alpha_5\beta_1$ integrin exhibited only a half-bent conformation.[40,41] In β_1 integrins, Ca^{2+} in the ADMIDAS seems to be a negative regulatory site responsible for integrin inhibition.[30] Plausibly, the presence in the calculated model of Ca^{2+} at ADMIDAS could contribute to prevent a larger conformational transition. It has been supposed that during receptor activation of β_1 integrins, Ca^{2+} at the ADMIDAS site becomes highly mobile and eventually is expelled from the site, whereas that of LIMBS (SyMBS) and MIDAS remain unchanged. Consistent to this, the inspection of the solid, close water-accessible surface of $\alpha_4\beta_1$-**3a** complex (Figure 3.8) shows that Ca^{2+} of ADMIDAS is more exposed respect to the $\alpha_4\beta_1$-**3c** complex.

Finally, there is evidence that in $\alpha_5\beta_1$ integrin the binding of small peptide ligands is not sufficient for full integrin opening.[30] The extended, open conformation is observed only when both Mn^{2+} and FN are present,[42] while Ca^{2+} binding to the ADMIDAS seems to stabilize the closed conformation.

In this scenario, our data for $\alpha_4\beta_1$ integrins seem to suggest that ligand binding and the overall integrin conformation are less tightly coupled than for other integrins. The small agonist **3a** alone at the binding site seems capable of activating intracellular signaling as an agonist. However, this interaction is not sufficient to induce full receptor opening.[30] Nevertheless, this agonist might act as a promoter of protein-protein interactions (PPIs),[43] being capable to predispose the receptor to adopt a semi-activated conformation and to facilitate Ca^{2+} depletion. The large reorganization of integrin structure would be possible only as a result of subsequent FN binding.

3.3 Conclusion

The CPPs described herein were proposed as potential ligands of α_4 integrin. In particular, the CPP **3c**, c[*(R)*-Phu-LDV-*(S)*-isoAsp], resulted to be an antagonist of α_4 integrins with moderate potency, while **3a**, c[*(S)*-Phu-LDV-*(S)*-isoAsp], appeared to be a potent agonist capable to increase both $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrin-mediated cell adhesion. In addition, **12a**, c[*(R)*-Phu-LAV-*(S)*-isoAsp], was an agonist which selectively promoted the adhesion of $\alpha_4\beta_1$ with low nanomolar potency but not that of $\alpha_4\beta_7$ integrin-expressing cells. Furthermore, these α/β hybrid cyclopeptides demonstrated excellent enzymatic stability, as evidenced by incubation in mouse serum.

In addition, the constrained cyclic LDV peptides could represent suitable probes to explore the structural requirements with respect to the 3D arrangement of the pharmacophoric groups and the interactions with $\alpha_4\beta_1$ integrin. To the purpose, we assembled a homology model of the receptor and optimized it by DFT combined QM/MM calculation to predicted ligands' conformations within the receptor. The binding geometries of **3a** and **3c** showed modest differences, despite of significantly different functions. Albeit speculative, the simulations are suggestive of a possible role of the agonist **3a** as small-molecule PPIs stabilizer, capable of prearranging the receptor in a semi-activated conformation. While the inhibition of PPIs by means of small-molecule drugs that disrupt or prevent a binary protein complex represents a classic approach in pharmacology, the opposite strategy to stabilize PPIs with small molecules is still regarded as an “exotic” approach, scarcely explored in the integrin field.[44]

Further developments might stem from potential applications of the agonist ligands in diagnostics or theranostics. These CPPs might serve as equivalents of the well-known integrin ligand c[RGDfK] which found a wide range of applications for targeting cancer cells, for cell growth and regenerative medicine, etc.[45-47]

3.4 Experimental Section

General Procedures. Unless otherwise stated, standard chemicals and solvents were purchased from commercial sources and used as received without further purification. Target compounds were determined to be ⩾95% pure by analytical HPLC analyses, performed on Agilent 1100 series apparatus, using a reverse-phase column Phenomenex mod. Gemini 3 μm C_{18} 110 Å 100 × 3.0 mm (No 00D-4439-Y0); column description: stationary phase octadecyl-carbon-chain-bonded silica (C_{18}) with trimethylsilyl endcap, fully porous organosilica solid support, particle size 3 μm, pore size 110 Å, length 100 mm, internal diameter 3 mm; mobile phase for neutral compounds: from H_2O/CH_3CN (9:1) to H_2O/CH_3CN (2:8) in 20 min at a flow rate of 1.0 mL·min^{-1}, followed by 10 min at the same composition; DAD (diode-array detection) 210 nm; mobile phase for ionisable peptides: from 9:1 H_2O/CH_3CN/0.1% HCOOH to 2:8 H_2O/CH_3CN/0.1% HCOOH in 20 min, flow rate of 1.0 mL min^{-1}; DAD 254 nm. Semipreparative RP HPLC was carried out with an Agilent 1100 series apparatus, using reverse-phase column ZORBAX mod. Eclipse XDBC18 PrepHT cartridge 21.2 × 150 mm 7 µm (no. 977150-102); column description: stationary phase octadecyl-carbon-chain-bonded silica (C_{18}), double endcapped, particle size 7 µm, pore size 80 Å, length 150 mm, internal diameter 21.2 mm; XSelect Peptide CSH C18 OBD column (Waters), 19 × 150 mm 5 µm (no. 186007021). column description: stationary phase octadecyl-carbon-chain-bonded silica (C_{18}), double endcapped, particle size 5 µm, pore size 130 Å, length 150 mm, internal diameter 19 mm; DAD 210 nm, DAD 254 nm; gradient mobile phase from H_2O/CH_3CN (8:2) to CH_3CN (100 %) in 10 min at a flow rate of 12 $mL{\cdot}min^{-1}$, isocratic mobile phase 1:1 H_2O/CH_3CN /0.1 % TFA in 8 min at a flow rate of 10 $mL{\cdot}min^{-1}$. Routine ESI MS analysis was carried out using an MS single quadrupole HP 1100 MSD detector, with a drying gas flow of 12.5 L min^{-1}, nebulizer pressure 30 psig, drying gas temp. 350 °C, capillary voltage 4500 (+) and 4000 (–), scan 50-2600 amu. High resolution mass spectrometry (HRMS) was performed with a Xevo G2XS QTof apparatus. NMR spectra were recorded on Varian Gemini apparatus (^1H: 400 MHz, ^{13}C: 100 MHz) or Bruker BioSpin GmbH (^1H: 600 MHz, ^{13}C: 150 MHz) at 298 K in 5 mm tubes, using 0.01 M peptide. Solvent suppression was carried out by the solvent presaturation procedure implemented in Varian (PRESAT). Chemical shifts are reported in ppm (δ) and referenced to residual nondeuterated solvent signal as internal standard ($CDCl_3$ ^1H: 7.26 ppm, ^{13}C: 77.16 ppm; $(CD_3)_2SO$: ^1H: 2.50, ^{13}C: 39.52 ppm). The unambiguous assignment of ^1H NMR resonances was based on 2D gCOSY experiments. VT ^1H NMR experiments were carried out over the range 298-348 K; temperature calibration was done with the ethylene glycol OH-CHn chemical-shift separation method. Coupling constants (J) are reported in Hz.

General procedure for SPPS of linear peptides 9. The linear peptides were assembled manually on Wang resin (0.3 g, 1.1 mmol/g loading capacity) using standard procedures. Prior to use, the resin was swollen in DMF (3 mL) for 15 min. In a separate vial, (*S*)- or (*R*)-Fmoc-Asp-OBn (0.3 mmol) and HOBt (0.3 mmol) were dissolved in DMF (4 mL). After 20 min the mixture was added to the resin, followed by DCC (0.3 mmol) and a catalytic amount of DMAP, and the resin was gently shaken for 3 h at RT. Thereafter, a mixture of Ac_2O (10 mmol) and pyridine (10 mmol) was added and shaken for additional 30 min to endcap the unreacted 4-hydroxybenzyl alcohol linkers. The resin was filtered and washed alternatively with DMF, MeOH, and DCM, (3×4 mL each).

Fmoc cleavage was carried out using 20% (*v/v*) piperidine in DMF (5 mL), while gently shaking at RT for 10 min. After washing with DMF and DCM (5 mL), the deprotection was repeated. The resin was then washed sequentially with DMF, MeOH, and DCM, (3×4 mL each). The subsequent coupling reactions were performed by dissolving in a separate vial Fmoc-protected amino acids (0.3 mmol) and HOBt (0.3 mmol) in DMF (4 mL) for 20 min. The last introduced residue was Boc-Phu-OH. The mixture was poured into the reactor followed by DCC (0.3 mmol), and the suspension was shaken for 3 h at RT. Coupling efficacy was monitored by Kaiser test.

Cleavage from the resin and simultaneous removal of Boc protecting group was performed by using a 95:2.5:2.5, *v/v/v* mixture of TFA/TIPS/H_2O (10 mL) while shaking for 2.5 h at RT. The mixture was filtered and the resin washed twice with Et_2O/DCM containing a small portion of TFA. The filtrates were collected and solvents were removed under reduced pressure, and ice-cold Et_2O was added to precipitate the crude peptides as TFA salt, which were recovered by centrifuge and used for the cyclization without further purification. Peptide identity was confirmed by ESI MS (Appendix A: Table S1).

General procedure for Synthesis of the CPPs. The cyclization of the crude peptide was performed under pseudo-high dilution condition. A solution of the linear peptides (0.15 mmol) in DMF (10 mL) was added over 16 h using a syringe pump, to a mixture of of HBTU (0.45 mmol), HOBt (0.45 mmol) and DIPEA (0.9 mmol) in DMF, at RT. Once the addition was complete, the reaction was stirred for additional 2 h. Then, the solvent was distilled at reduced pressure, and the crude peptides were isolated by RP HPLC on a semipreparative C18 column (General methods). Compound's identity was confirmed by ESI MS, in reasonable yield (Appendix A: Table S1).

Removal of benzyl protecting groups was performed by catalytic hydrogenation. A stirred suspension of the protected cyclopentapeptide **10** (0.1 mmol) and a catalytic amount of 10% *w/w* Pd/C in absolute EtOH (10 mL) was stirred under H_2 atmosphere for 12 h at RT. Thereafter, the catalyst was filtered off over Celite® and the solvent was distilled under reduce pressure, to afford the final products **3a-d** in quantitative yield. The purity (Table 3.1) and the identity of the products were determined to be >95% by RP HPLC coupled to ESI MS, and by ^1H and ^{13}C NMR, and 2D gCOSY experiments in 8:2 DMSO-d_6/H_2O.

Cell adhesion assays. For adhesion assays on Jurkat E6.1, RPMI8866, HL60 cells, black 96-well plates were coated overnight at 4 °C with VCAM-1 or ICAM-1 or MAdCAM-1 (5 μg/mL) or Fg (10 μg/mL). The cells were counted, stained with CellTracker green CMFDA (12.5 μM, 30 min at 37 °C, Life Technologies) and after three washes, were preincubated with increasing concentrations of new CPP (10^{-10}-10^{-4} M) or with the vehicle (methanol) for 30 min at 37 °C. Then, cells were plated (500000/well) on coated wells and incubated for 30 min at 37 °C. After three washes, adhered cells were lysed with 0.5% Triton X-100 in PBS (30 min at 4 °C) and fluorescence was measured (Ex 485 nm/Em 535 nm) in an EnSpire Multimode Plate Reader (PerkinElmer, Waltham, MA, USA). For adhesion assays mediated by $\alpha_5\beta_1$ integrin, 96-well plates were coated by passive adsorption with FN (10 μg/mL) overnight at 4 °C. K562 cells were counted and preincubated with various concentrations of the peptides or with the vehicle (methanol) for 30 min at RT. Afterwards, the cells were plated (50,000 cells/well) and incubated at RT for 1 h. The wells were then washed with 1% BSA in PBS (phosphate-buffered saline) to take off nonadherent cells, and 50 μL of hexosaminidase substrate was added; after addition of 100 μL of stopping solution the plates were read at 405 nm. In both types of adhesion assay, the number of adherent cells was determined by comparison with a standard curve made in the same plate. Experiments were carried out in quadruplicate and repeated at least three times. Data analysis and EC_{50} or IC_{50} values were calculated using GraphPad Prism 5.0 (GraphPad Software, San Diego, CA, USA) and concentration-response curves are provided in Appendix A: Figures S2-S7. In addition, to depict agonistic or antagonistic behaviour of the new synthesized compounds, we calculated the adhesion index (Figure 3.2), which is calculated as the ratio between the number of adhered cells in the presence of the highest CPP concentration (10^{-4} M) and the number of adhered vehicle-treated cells. Based on the adhesion index value, it is possible to distinguish between: agonist (adhesion index >1, cell adhesion is increased), antagonist (adhesion index <1, cell adhesion is decreased), compounds not significantly modifying integrin-mediated cell adhesion (adhesion index ≈1, cell adhesion is not altered).

Competitive binding assay on purified integrins. Solid-phase ligand binding assays on purified integrin were conducted as previously described[2] with the following modifications. Regarding $\alpha_5\beta_1$ and $\alpha_L\beta_2$ integrins, black 96-well plates were coated by passive adsorption with FN (0.5 μg/mL) for $\alpha_5\beta_1$ or with ICAM-1 (10 μg/mL, R&D Systems) for $\alpha_L\beta_2$ in carbonate buffer (15 mM Na_2CO_3, 35 mM $NaHCO_3$, pH 9.6) overnight at 4 °C. The following day, wells were blocked with TSB buffer (20 mM Tris-HCl, 150 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM $MnCl_2$, pH 7.5, 1% BSA) for 1 h at room temperature. Purified $\alpha_5\beta_1$ (10 μg/mL) or $\alpha_L\beta_2$ (7 μg/mL) was incubated with the new synthesized compounds, at different concentrations (10^{-4} -10^{-10} M), in coated wells for 1 h at RT. Then, after three washes with PBS buffer, primary antibody (anti-$\alpha_5\beta_1$, BD Bioscience, 1:100 dilution or anti-α_L, Abcam, 1:200 dilution) was added for 1 h at RT. Then, anti-rabbit AlexaFluor488-secondary antibody (ThermoFisher Scientific, 1:400 dilution) was added after three washes with PBS buffer and incubated 1 h at room temperature. After washing three times, fluorescence was measured (Ex 485 nm/Em 535 nm) in Multimode Plate Reader (PerkinElmer, Waltham, MA, USA).

For the evaluation of binding affinity to purified $\alpha_4\beta_1$, $\alpha_M\beta_2$ and $\alpha_4\beta_7$ integrins, competitive solid-phase ligand binding assays were performed as follows. Black 96 well plates were coated overnight at 4 °C with the following endogenous ligands: FN or VCAM-1 (10 μg/mL) for $\alpha_4\beta_1$, MAdCAM-1 (2 μg/mL) for $\alpha_4\beta_7$, and fibrinogen (10 μg/mL) for $\alpha_M\beta_2$, in PBS + 2mM $MgCl_2$ + 0.5% BSA. Afterwards, each well was washed and blocked for 1 h at RT. Purified integrins ($\alpha_4\beta_1$: 0.5 μg/mL; $\alpha_4\beta_7$: 0.5 μg/mL; $\alpha_M\beta_2$: 0.5 μg/mL; R&D Systems) were pre-incubated with serial dilutions of new compounds (10^{-4}-10^{-10} M) for 30 min at RT and then plated into coated wells for 1 h at RT. After 2 washes, primary antibody (for $\alpha_4\beta_1$ and $\alpha_4\beta_7$: rabbit anti-α_4, Abcam, 1:100 dilution; for $\alpha_M\beta_2$: rabbit anti-α_M, Abcam, 1:100 dilution) was added and incubated for 1 h at RT. The plate was washed twice and then was incubated with anti-rabbit AlexaFluor488 secondary antibody (1:400 dilution, ThermoFischer Scientific) for 1 h at RT. After washing three times, fluorescence was measured as described in the previous section.

Experiments were carried out in triplicate and repeated at least three times. Data analysis and IC_{50} affinity values were calculated using GraphPad Prism 9 (GraphPad Software, San Diego, CA, USA) and binding curves are shown in Appendix A: Figures S8-S12.

Western blot analysis. Western blot analysis was performed as previously described, with the following modifications. Jurkat E6.1 cells were cultured for 16/18 h in RPMI medium containing 1% PBS; then, 4×10^6 cells were incubated for 1 h with different concentrations of the most effective cyclic peptides (10^{-7}, 10^{-8}, 10^{-9} M), which were identified as agonists in cell adhesion assays mediated by $\alpha_4\beta_1$ integrin. On the other hand, after 1 h incubation with integrin antagonists, the cells were then seeded on FN (10 μg/mL) coated plates for 1 h. Integrin agonists were not incubated with FN. At the end of the incubation time, Jurkat E6.1 cells were lysed on ice using a mammalian protein extraction reagent (M-PER; Pierce, Rockford, IL, USA) supplemented with a phosphatase inhibitor cocktail. Protein extracts were quantified using a BCA protein assay kit (Pierce, Rockford, IL, USA), separated by 12% SDS-PAGE gel, transferred onto nitrocellulose membranes and immunoblotted with anti-phospho-ERK1/2 (1:1000) (Cell Signalling Technology, Danvers, MA, USA) or anti-total ERK1/2 antibodies (1:2500) (Cell Signalling Technology). Protocols for digital image acquisition and analysis have been previously described.[48] Densitometric analysis of the bands is reported (mean ± SD; n=3); the amount of phosphorylated ERK1/2 (pERK1/2) is normalized to that of total ERK1/2 (totERK1/2). Experiments were replicated independently at least three times. Statistical analyses were performed using one-way ANOVA and post hoc Newman-Keuls test.

In Vitro Enzymatic Stability. Enzymatic stability tests were carried out in triplicate and repeated three times using mouse serum (Sigma-Aldrich). Peptides was dissolved in Tris buffer pH 7.4 to a 10 mM concentration, and 10 mL aliquots were added to 190 mL of serum. Incubations were maintained at 37 °C, and 20 mL aliquots were sampled from the incubation mixtures at the indicated times of 0, 0.15, 0.5, 1.0, 2.0, 3.0 h. Samples were diluted with 90 mL CH_3CN, and enzymatic activity was definitively stopped by adding 90 mL of 0.5% AcOH. After centrifugation (13,000×g for 20 min), the supernatants were separated and the amount of remaining peptide was assessed by RP HPLC.

Conformational Analysis of CPPs. Peptide samples were dissolved in 8:2 DMSO-d_6/H_2O in 5 mm tubes to the final concentration of 0.01 M. At this concentration, the intramolecular aggregation in mixtures of DMSO-d_6 and H_2O is usually unimportant. Besides, self-association of the peptides was excluded based on the reproducibility of the chemical shift of non-exchangeable protons in the concentration range 0.01-0.04 M (not shown). Water suppression was achieved by the PRESAT procedure implemented in Varian. Proton resonance assignment was accomplished through gCOSY. VT ^1H-NMR experiments were recorded over the range of 298-348 K. 2D ROESY experiments were done at RT, phase-sensitive mode, spin-locking field (γb2) = 2000 Hz, mixing time = 250 ms; spectra were processed in the hypercomplex approach; peaks were calibrated on the solvent. Only ROESY-derived constraints were included in the restrained molecular dynamics (MD). Cross-peak intensities were ranked and associated to the distances (Å): very strong = 2.3, strong = 2.6, medium = 3.0, weak = 5.0. The intensities of the cross-peaks arising from protons separated by known distances (e.g., geminal) were found to match with these associations but were discarded. For the absence of Hα(*i*), Hα (*i* + 1) ROESY cross-peaks, all of the ω bonds were set at 180° (f constant: 16 kcal mol^{-1} $Å^{-2}$).

Molecular dynamics simulations. The restrained MD simulations were conducted at 300 K and 1 atm by using the AMBER force field in a 30 × 30 × 30 $Å^3$ box of standard TIP3P models of equilibrated water, periodic boundary conditions dielectric scale factor = 1, and cutoff for the nonbonded interactions = 12 Å; all water molecules closer than 2.3 Å to a solute atom were eliminated, and 50 random structures were generated by a 100 ps simulation at 1200 K; these were subsequently subjected to restrained MD, 50 ps with a 50% scaled force field at 1200 K, then by 50 ps with full distance restraints, force constant = 7 kcal mol^{-1} $Å^{-2}$, after which the system was cooled in 20 ps to 50 K. H-bond interactions were not included, nor were torsion angle restraints. The resulting structures were minimized by 3000 cycles of steepest descent and 3000 cycles of conjugated gradient, and convergence = 0.01 kcal $Å^{-1}$ mol^{-1}. The backbones of the structures were clustered by the rmsd analysis. Unrestrained MD simulations were performed starting with the conformation derived from ROESY in the box of standard TIP3P water for 100 ns at 298 K using periodic boundary conditions, at constant temperature and pressure (Berendsen scheme, bath relaxation constant of 0.2). For 1-4 scale factors, van der Waals and electrostatic interactions are scaled in AMBER to half their nominal value. The integration time step was set to 0.1 fs. The system coordinates were collected every picosecond.

Molecular modelling. The ligand molecules were obtained using a systematic conformational search followed by geometry optimization of the lowest energy structure with MOPAC7 (PM3 Method, RMS gradient 0.01).[49] Since the precise structure of the $\alpha_4\beta_1$ integrin is not yet available, the $\alpha_4\beta_1$ integrin receptor model was obtained by combining the Crystallographic structures of the α_4 subunit (PDB ID: 3V4V) and of the β_1 subunit (PDB ID: 4WK4). This decision was made considering the better homology of the pair β_1/β_7 (52.80%) compared to α_4/α_5 (34.22%). The structural superposition was obtained using the "MatchMaker" procedure implemented in UCSF-Chimera.[50] The pair-wise sequence alignments of the protein fragments were achieved using the blocks substitution matrix 62 (BLOSUM-62) by the Needleman-Wunsch algorithm.[51] The coordinates of the subunits were aligned using residue pairs from the sequence alignments. The superposition/alignment steps were iterated until convergence to perform one or more cycles of refitting of the structures using the sequence alignment and generating a new sequence alignment from the adjusted superposition. The residues at the $\alpha_4\beta_1$ interface were checked and any clashes/overlaps were removed using the Dunbrack 2010 rotamer library, a backbone-dependent rotamer library composed of rotamer frequencies, mean dihedral angles and variances as a function of the backbone dihedral angles.[52] Hydrogen atoms were added with respect of hydrogen bonding network by Reduce software[53] and the PROPKA program[54,55] was employed to estimate the protonation states of the titratable residues. The final model was then validated using the ligands present in the two crystallographic structures used as models (RO0505376 in $\alpha_4\beta_7$ and the cRGD peptide in $\alpha_5\beta_1$). Even considering the obvious differences due to the new combination of the subunits, the complexes resulting from the molecular docking simulations are consistent with the conformations of the original complexes. Molecular docking experiments were performed with Autodock 4.0. We used the Lamarckian Genetic Algorithm which combines global search (Genetic Algorithm alone) to local search (Solis and Wets algorithm). Ligands and receptors were further processed using the Autodock Tools (ADT) software.[56] Gasteiger PEOE[57] charges were loaded on the ligands in ADT and solvation parameters were added to the final structure using the Addsol utility of Autodock. Each docking run consisted of an initial population of 100 randomly placed individuals, a maximum number of 200 energy evaluations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value of 1. For the local search, the so-called pseudo-Solis and Wets algorithm was applied using a maximum of 250 iterations per local search. 250 independent docking runs were carried out for each ligand. The grid maps representing the system in the actual docking process were calculated with Autogrid. The dimensions of the grids were 100×100×100, with a spacing of 0.1 Å between the grid points and the center close to the cavity left by the ligand after its removal. The simpler inter-molecular energy function based on the Weiner force field in Autodock was used to score the docking results. Results differing by less than 1.0 Å in positional root-mean-square deviation (rmsd) were clustered together and were represented by the result with the most favourable free energy of binding. The poses thus obtained were equilibrated by a 5.0 ns of partially restrained MD simulation using the CUDA® version of the GROMACS package[58] with a modified version of the AMBER ff03 force field, a variant of the AMBER ff991 potential in which charges and main-chain torsion potentials have been derived based on QM+continuum solvent calculations and each amino acid is allowed unique main-chain charges. Amber Tools[59,60] was applied to generate the Generalized Amber Force Field (GAFF) files for the unusual residues. The GROMACS molecular topology files (*.gro and *.top) were obtained from the Amber files by Acpype.[61] The MD consisted of 100 ps heating dynamics from 0 to 300 K, followed by equilibration dynamics performed for 5 ns. The MD simulation was performed at constant temperature and volume, with the application of constrained harmonic potentials for the metal ions. After the above-described MD simulations, a combined QM/MM calculation between the ligand and the protein environment was performed using the NWChem 6.1.1 package.[62] The QM region contained the ligand atoms, the MIDAS and the side chains of all major residues of the binding site. The theoretical level used for the QM region was the hybrid DFT of the B3LYP[63] exchange-correlation functional with Grimme's D3 dispersion correction (B3LYP-D3) and the 6-31G(d) basis sets while the MM atoms were subjected to an Amber ff99 force field (B3LYP-D3/6-31G(d) | Amber ff99). Hydrogen link atoms were used for the QM/MM boundary and the non-bonded QM/MM interactions were calculated with a cut-off of 10 Å. Interactions of QM atoms with all MM charges were included in calculations.

NMR/HRMS-ESI Characterization

c[(S)-Phu-Leu-Asp-Val-(S)-isoAsp] **3a**. **^{1}H-NMR** (400 MHz, 8:2 DMSO-*d*$_6$/H_2O) *δ* 8.98 (s, 1H, PhuNHb), 8.65 (d, *J* = 6.0 Hz, 1H, LeuNH), 8.22-8.14 (m, 2H, AspNH + PhuNH), 7.90 (s, 1H, PhuNHa), 7.80 (d, *J* = 7.6 Hz, 1H, ArH$_6$), 7.75 (d, *J*=9.2 Hz, 1H, ValNH), 7.37 (d, *J*=7.6 Hz, 2H, ArH$_{2'6'}$), 7.22-7.09 (m, 5H, ArH$_{3'5'}$+ArH$_{3,5}$+isoAspNH), 6.94 (dd, *J* = 7.2, 6.8 Hz, 1H, ArH$_4$), 4.60-4.53 (m, 1H, isoAspHα), 4.38-4.31 (m, 1H, PhuHα), 4.29-4.24 (m, 1H, AspHα), 4.16 (dd, *J* = 9.2, 4.0 Hz, 1H, ValHα), 3.79-3.68 (m, 1H, LeuHα), 2.97-2.89 (m, 2H, PhuCHβ+AspHβ), 2.87 (dd, *J* = 14.0, 7.6 Hz, 1H, AspHβ), 2.74 (dd, *J* = 14.0, 2.0 Hz, 1H, PhuCHβ), 2.69 (dd, *J* = 14.4, 2.4 Hz, 1H, isoAspHβ), 2.61 (dd, *J* = 14.4, 4.0 Hz, 1H, isoAspHβ), 2.34-2.25 (m, 1H, ValHβ), 2.23 (s, 3H, ArCH$_3$), 1.73-1.62 (m, 1H, LeuHβ), 1.55-1.43 (m, 1H, LeuHβ), 1.40-1.30 (m, 1H, LeuHγ), 0.91-0.76 (m, 12H, ValCH$_3$+LeuCH$_3$); **^{13}C-NMR** (100 MHz, DMSO-*d*$_6$) *δ* 172.5, 172.1, 171.9, 171.2, 170.5, 170.4, 169.8, 152.7, 138.3, 137.5, 130.7, 130.2, 129.3, 127.5, 126.1, 122.6, 121.0, 117.8, 57.7, 55.5, 52.5, 51.8, 48.4, 37.9, 36.9, 36.0, 35.2, 29.3, 24.2, 23.4, 21.2, 19.8, 17.9, 17.6. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{36}H_{48}N_7O_{10}]^+$ 738.34627, found 738.34654 $[M+H]^+$.

c[(S)-Phu-Leu-Asp-Val-(R)-isoAsp] **3b**. **^{1}H-NMR** (400 MHz, 8:2 DMSO-*d*$_6$/H_2O) *δ* 9.02 (s, 1H, PhuNHb), 8.49 (br.d, 1H, AspNH), 8.38 (d, *J* = 8.4 Hz, 1H, LeuNH), 8.32 (d, *J* = 8.4 Hz, 1H, PhuNH), 8.20 (br.d, 1H, ValNH), 8.01 (br.d, 1H, isoAspNH), 7.94 (s, 1H, PhuNHa), 7.80 (d, *J* = 8.0 Hz, 1H, ArH$_6$), 7.36 (d, *J* = 8.0 Hz, 2H, ArH$_{2'6'}$), 7.18 – 7.11 (m, 4H, ArH$_{3'5'}$+ArH$_{3,5}$), 6.93 (dd, *J* = 7.6, 7.2 Hz, 1H, ArH$_4$), 4.40-4.32 (m, 2H, AspHα+PhuHα), 4.27-4.17 (m, 2H, isoAspHα+LeuHα), 3.96 (dd, *J* = 9.2, 8.8 Hz, 1H, ValHα), 3.03 (dd, *J* = 14.2, 3.8 Hz, 1H, PhuHβ), 2.80-2.68 (m, 4H, PhuHβ+isoAspHβ+AspHβ), 2.38 (dd, *J* = 15.2, 1.6 Hz, 1H, isoAspHβ), 2.23 (s, 3H, ArCH$_3$), 2.13-2.05 (m, 1H, ValHβ), 1.72-1.61 (m, 1H, LeuHβ), 1.50-1.48 (m, 2H, LeuHβ+LeuHγ), 0.87-0.85 (m, 12H, ValCH$_3$+LeuCH$_3$); **^{13}C-NMR** (100 MHz, DMSO-*d*$_6$) *δ* 171.3, 171.0, 170.5, 169.7, 169.2, 152.7, 138.3, 137.5, 130.2, 129.3, 127.5, 126.1, 122.6, 121.1, 118.5, 117.9, 58.1, 52.6, 51.6, 49.1, 48.0, 34.5, 34.3, 31.3, 28.7, 24.2, 22.9, 22.1, 21.6, 19.5, 17.9. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{36}H_{48}N_7O_{10}]^+$ 738.34627, found 738.34599 $[M+H]^+$.

c[(R)-Phu-Leu-Asp-Val-(S)-isoAsp] **3c**. **^{1}H-NMR** (400 MHz, 8:2 DMSO-*d*$_6$/H_2O) *δ* 9.04 (s, 1H, PhuNHb), 8.30 (br.d, 1H, PhuNH), 8.25-8.15 (m, 2H, isoAspNH+LeuNH), 8.02 (br.d, 1H, ValNH), 7.93 (s, 1H, PhuNHa), 7.79 (d, *J* = 8.4 Hz, 1H, ArH$_6$), 7.70 (d, *J* = 6.4 Hz, 1H, Asp-NH), 7.34 (d, *J* = 8.0 Hz, 2H, ArH$_{2'6'}$), 7.17 – 7.11 (m, 2H, ArH$_{3,5}$), 7.06 (d, *J* = 8.0 Hz, 2H, ArH$_{3'5'}$), 6.93 (t, *J* = 7.2 Hz, 1H, ArH$_4$), 4.48-4.42 (m, 1H, AspHα), 4.38-4.35 (m, 1H, PhuHα), 4.32-4.25 (m, 1H, isoAspHα), 4.02-3.92 (m, 1H, LeuHα), 3.58-3.49 (m, 1H, ValHα), 2.79-2.73 (m, 3H, PhuHβ+AspHβ), 2.68-2.55 (m, 3H, isoAspHβ+AspHβ), 2.36-2.26 (m, 1H, ValHβ), 2.22 (s, 3H, ArCH$_3$), 1.38-1.28 (m, 2H, LeuHβ), 1.02-0.92 (m, 1H, LeuCHγ), 0.83 (d, *J* = 6.4 Hz, 3H, LeuCH$_3$), 0.79 (d, *J* = 6.4 Hz, 3H, LeuCH$_3$), 0.72 (d, *J* = 6.0 Hz, 3H, ValCH$_3$), 0.64 (d, *J* = 6.4 Hz, 3H, ValCH$_3$); **^{13}C-NMR** (100 MHz, DMSO-*d*$_6$) *δ* 172.3, 172.0, 171.8, 171.1, 170.1, 169.8, 152.7, 138.4, 137.5, 130.1, 130.0, 129.4, 127.5, 126.1, 122.5, 121.0, 117.6, 109.5, 61.3, 54.9, 51.5, 51.2, 49.6, 36.4, 35.9, 33.6, 31.3, 23.5, 23.0, 22.1, 21.0, 19.1, 17.9. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{36}H_{48}N_7O_{10}]^+$ 738.34627, found 738.34688 $[M+H]^+$.

c[(R)-Phu-Leu-Asp-Val-(R)-isoAsp] **3d.** **^{1}H-NMR** (400 MHz, 8:2 DMSO-*d*$_6$/H_2O) *δ* 8.99 (s, 1H, PhuNHb), 8.55 (d, *J* = 6.4 Hz, 1H, ValNH), 8.51 (d, *J* = 4.8 Hz, 1H, PhuNH), 8.42 (d, *J* = 6.8 Hz, 1H, isoAspNH), 8.24 (d, *J* = 8.4 Hz, 1H, LeuNH), 7.90 (s, 1H, PhuNHa), 7.80 (d, *J* = 7.6 Hz, 1H, ArH6), 7.35 (d, *J* = 7.6 Hz, 3H, AspNH+ArH2’6’), 7.19-7.10 (m, 2H, ArH3,5), 7.09 (d, *J* = 8.0 Hz, 2H, ArH3’5’), 6.94 (dd, *J* = 7.6, 7.2, Hz, 1H, ArH4), 4.54 (dd, *J* = 12.4, 4.4 Hz, 1H, AspHα), 4.42-4.35 (m, 1H, isoAspHα), 4.29-4.21 (m, 1H, PhuHα), 3.97-3.89 (m, 1H, LeuHα), 3.24-3.17 (m, 1H, ValHα), 2.84-2.69 (m, 3H, PhuHβ+AspHβ), 2.682.54 (m, 3H, isoAspHβ+ValHβ), 2.46-2.42 (m, 1H, AspHβ), 2.23 (s, 3H, ArCH$_3$), 1.40-1.31 (m, 1H, LeuHβ), 1.30-1.22 (m, 1H, LeuHβ), 0.84 (d, *J* = 6.4 Hz, 7H, ValCH$_3$+LeuHγ) 0.68 (d, *J* = 6.4 Hz, 3H, LeuCH$_3$), 0.58 (d, *J* = 5.6 Hz, 3H, LeuCH$_3$); **^{13}C-NMR** (100 MHz, DMSO-*d*$_6$) δ 172.0, 171.9, 171.7, 171.3, 170.5, 170.3, 170.1, 152.6, 138.4, 137.4, 130.1, 129.7, 129.4, 127.5, 126.1, 122.6, 121.1, 117.6, 64.5, 55.4, 51.2, 50.3, 48.8, 36.1, 35.8, 34.9, 27.6, 23.3, 23.1, 20.8, 19.5, 19.2, 17.8. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{36}H_{48}N_7O_{10}]^+$ 738.34627, found 738.34701 $[M+H]^+$.

c[(S)-Phu-Leu-Asp-Val-(R)-β^3homoAla] **11a**. **^{1}H NMR** (400 MHz, 8:2 DMSO-*d*$_6$/H_2O) δ 9.01 (s, 1H, PhuNHb), 8.71 (d, *J* = 6.8 Hz, 1H, LeuNH), 8.07 (d, *J* = 7.2 Hz, 1H, AspNH), 7.92 (s, 1H, PhuNHa), 7.81 (d, *J* = 8.4 Hz, 2H, PhuNH+ArH6), 7.59 (d, *J* = 9.6 Hz, 1H, ValNH), 7.37 (d, *J* = 8.8 Hz, 2H, ArH2'6'), 7.17 (d, *J* = 8.8 Hz, 3H, ArH3'5'+ArH3), 7.12 (d, *J* = 8.4 Hz, 1H, ArH5), 7.03 (d, *J* = 7.6 Hz, 1H, β^3AlaNH), 6.93 (dd, *J* = 8.0, 7.2 Hz, 1H, ArH4), 4.41 (dd, *J* = 15.2, 7.2 Hz, 1H, PhuHα), 4.24 (dd, *J* = 12.8, 7.2 Hz, 1H, AspHα), 4.16-4.14 (m, 1H, β^3AlaHβ), 4.11 (dd, *J* = 9.6, 5.6 Hz, 1H, ValHα), 3.66-3.60 (m, 1H, LeuHα), 2.91-2.84 (m, 3H, PuHβ+AspHβ), 2.76 (dd, *J* = 12.8, 8.0 Hz, 1H, PhuHβ), 2.50 (m, 1H, β^3AlaHα), 2.23-2.20 (m, 4H, ArCH$_3$+ValHβ), 2.01 (dd, *J* = 13.2, 6.4 Hz, 1H, β^3AlaHα), 1.76-1.69 (m, 1H, LeuHβ), 1.501.43 (m, 1H, LeuHβ), 1.23-1.18 (m, 1H, LeuHγ), 1.11 (d, *J* = 6.4 Hz, 3H, β^3AlaCH$_3$), 0.86 (d, *J* = 7.2 Hz, 6H, ValCH$_3$), 0.82 (d, *J* = 6.4 Hz, 3H, LeuCH$_3$), 0.78 (d, *J* = 6.0 Hz, 3H, LeuCH$_3$); **^{13}C NMR** (101 MHz, DMSO-*d*$_6$) δ 172.5, 172.2, 171.5, 170.5, 170.4, 169.7, 152.7, 138.4, 137.5, 130.4, 130.2, 129.3, 127.5, 126.1, 122.6, 121.0, 117.8, 57.9, 54.6, 52.7, 52.3, 42.8, 41.3, 37.2, 36.3, 34.8, 24.1, 23.6, 21.1, 20.4, 19.8, 18.0, 17.8. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{36}H_{50}N_7O_8]^+$ 708.37209, found 708.37190 $[M+H]^+$.

c[(S)-Phu-Leu-Ala-Val-(S)-isoAsp] **12a**. **^{1}H-NMR** (400 MHz, 8:2 DMSO-*d*$_6$/H_2O) δ 10.56 (br.d, 1H, LeuNH), 9.18 (s, 1H, PhuNHb), 9.06 (br.d, 1H, PhuNH), 8.58 (br.d, 1H, AlaNH), 8.05 (s, 1H, PhuNHa), 7.80 (d, *J* = 7.6 Hz, 1H, ArH6), 7.51 (d, *J* = 9.6 Hz, 1H, ValNH), 7.40 (d, *J* = 8.4 Hz, 2H, ArH2'6'), 7.18-7.09 (m, 2H, ArH3,5), 7.05 (d, *J* = 8.0 Hz, 2H, ArH3'5'), 6.93 (dd, *J* = 7.6, 7.2, Hz, 1H, ArH4), 6.86 (d, *J* = 6.0 Hz, 1H, isoAspNH), 4.40 (d, *J* = 8.8 Hz, 1H, ValHα), 4.26 (dd, *J* = 7.2, 6.4 Hz, 1H, AlaHα), 4.12-4.05 (m, 1H, isoAspHα), 4.04-3.92 (m, 2H, PhuHα+LeuHα), 2.94-2.80 (m, 2H, PhuHβ+isoAspHβ), 2.69-2.56 (m, 2H, PhuCHβ+isoAspHβ), 2.56-2.49 (m, 1H, ValHβ), 2.24 (s, 3H, ArCH$_3$), 1.87-1.76 (m, 1H, LeuHβ), 1.40 (d, *J* = 7.6 Hz, 3H, AlaCH$_3$), 1.15-1.01 (m, 4H, LeuHβ+ValCH$_3$), 0.92-0.79 (m, 4H, ValCH$_3$+LeuCHγ), 0.73 (s, 3H, LeuCH$_3$), 0.57 (s, 3H, LeuCH$_3$); **^{13}C-NMR** (100 MHz, DMSO-*d*$_6$) δ 176.5, 175.9, 172.2, 171.7, 171.6, 168.6, 152.7, 138.7, 137.5, 130.1, 129.3, 127.6, 126.1, 122.6, 121.1, 117.7, 56.4, 55.9, 51.8, 51.0, 50.3, 36.2, 35.8, 30.8, 28.1, 23.9, 20.0, 19.5, 18.0, 17.1, 16.6. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{35}H_{48}N_7O_8]^+$ 694.35644, found 694.35596 $[M+H]^+$.

c[(S)-PhU-Leu-Asp-Val-(S)-isoAsp(nPr)] **13**. **^{1}H NMR** (400 MHz, 8:2 DMSO-*d*$_6$/H_2O) δ 9.08 (br s, 1H, PhuNHb), 8.73 (d, *J* = 7.2 Hz, 1H, LeuNH), 8.25 (d, *J* = 5.2 Hz, 1H, AspNH), 7.99 (br s, 1H, PhuNHa), 7.96 (br d, 1H, PhuNH), 7.82 (d, *J* = 8.0 Hz, 1H, ArH6), 7.71 (d, *J* = 8.8 Hz, 1H, ValNH), 7.46-7.41 (m, 1H, propyl-NH), 7.38 (d, *J* = 8.4 Hz, 2H, ArH2'6'), 7.22-7.10 (m, 5H, ArH3'5'+ArH3,5+isoAspNH), 6.93 (t, *J* = 7.2 Hz, 1H, ArH4), 4.44 (dd, *J* = 12.4, 7.6 Hz, 1H, isoAspHα), 4.32-4.25 (m, 2H, PhuHα+AspHα), 4.15 (dd, *J* = 8.8, 6.0 Hz, 1H, ValHα), 3.72-3.64 (m, 1H, LeuHα), 3.01-2.94 (m, 2H, propylCH$_2$), 2.91-2.80 (m, 4H, PhuCHβ+AspHβ), 2.65 (dd, *J* = 14.0, 4.0 Hz, 1H, isoAspHα), 2.55-2.50 (m, 1H, isoAspHα), 2.34-2.25 (m, 1H, ValHβ), 2.24 (s, 3H, ArCH$_3$), 1.72-1.63 (m, 1H, LeuHβ), 1.52-1.43 (m, 1H, LeuHβ), 1.42-1.32 (m, 3H, propylCH$_2$+LeuHγ), 0.90-0.74 (m, 15H, ValCH$_3$+LeuCH$_3$ +propylCH$_3$); **^{13}C-NMR** (100 MHz, DMSO-*d*$_6$) δ 172.5, 172.2, 171.4, 170.7, 170.4, 170.2, 170.0, 152.7, 138.5, 137.5, 130.3, 130.2, 129.3, 127.5, 126.1, 122.6, 121.0, 117.8, 58.0, 55.5, 52.5, 51.9, 50.5, 40.4, 37.5, 37.2, 36.0, 34.9, 31.3, 28.7, 24.1, 23.5, 22.2, 21.0, 20.0, 18.0, 17.9, 11.2. **HRMS-ESI/QTOF** *m/z* calcd. for

$[C_{39}H_{55}N_8O_9]^+$ 779.40920, found 779.40883 $[M+H]^+$.

c[(R)-Phu-Leu-Asp-Val-(R)-β^3homoAla] **11c**. **^{1}H NMR** (400 MHz, 8:2 DMSO-d_6/H_2O) *δ* 9.05 (s, 1H, PhuNHb), 8.39 (d, *J* = 4.8 Hz, 1H, PhuNH), 8.15 (d, *J* = 8.4 Hz, 1H, LeuNH), 7.93 (s, 1H, PhuNHa), 7.88 (d, *J* = 7.2 Hz, 1H, β^3AlaNH), 7.85 (d, *J* = 8.4 Hz, 1H, ValNH), 7.82 (d, *J* = 8.0 Hz, 1H, AspNH), 7.79 (d, *J* = 8.0 Hz, 1H, ArH6), 7.36 (d, *J* = 8.4 Hz, 2H, ArH2’6’), 7.15 (t, *J* = 8.2 Hz, 1H, ArH5), 7.09 (d, *J* = 8.4 Hz, 3H, ArH3’5’+ArH3), 6.93 (t, *J* = 7.2 Hz, 1H, ArH4), 4.60 (dd, *J* = 15.2, 8.0 Hz, 1H, AspHα), 4.28 (dd, *J* = 13.6, 6.4 Hz, 1H, PhuHα), 4.01 (dd, *J* = 11.6, 8.0 Hz, 1H, LeuHα), 3.95-3.92 (m, 1H, β^3AlaHβ), 3.63 (t, *J* = 7.2 Hz, 1H, ValHα), 2.91 (dd, *J* = 14.0, 8.4 Hz, 1H, AspHβ), 2.87 (dd, *J* = 12.0, 5.6 Hz, 1H, PhuCHβ), 2.76 (dd, *J* = 14.4, 8.8 Hz, 1H, PhuHβ), 2.55-2.50 (m, 1H, AspHβ), 2.39 (dd, *J* = 13.6 5.6 Hz, 1H, β^3AlaHα), 2.28 (dd, *J* =, 14.0, 6.8 Hz, 1H, ValHβ), 2.23 (s, 3H, $ArCH_3$), 2.16 (dd, *J* = 13.2, 3.6 Hz, 1H, β^3AlaHα), 1.37 (dd, *J* = 19.2, 6.8 Hz, 2H, LeuHβ), 1.26-1.10 (m, 1H, LeuHγ), 1.05 (d, *J* = 6.8 Hz, 3H, β^3AlaCH_3), 0.84 (d, *J* = 6.8 Hz, 3H, ValCH_3), 0.80 (d, *J* = 6.4 Hz, 3H, ValCH_3), 0.70 (d, *J* = 6.0 Hz, 3H, LeuCH_3), 0.62 (d, *J* = 5.6 Hz, 3H, Leu-CH_3); **^{13}C-NMR** (100 MHz, DMSO-d_6) *δ* 171.8, 171.7, 171.3, 171.1, 169.5, 169.2, 152.7, 138.5, 137.5, 130.1, 129.7, 129.5, 127.5, 126.1, 122.6, 121.0, 117.6, 60.4, 55.7, 51.0, 50.7, 42.7, 41.9, 36.3, 36.0, 29.0, 28.5, 23.3, 23.1, 21.0, 19.6, 19.5, 18.7, 17.9. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{36}H_{50}N_7O_8]^+$ 708.37209, found 708.37287 $[M+H]^+$.

c[(R)-Phu-Leu-Ala-Val-(S)-isoAsp] **12c**. **^{1}H NMR** (400 MHz, 8:2 DMSO-d_6/H_2O) *δ* 9.09 (s, 1H, PhuNHb), 8.36 (br s, 1H, PhuNH), 8.12 (d, *J* = 6.8 Hz, 1H, LeuNH), 8.00 (s, 1H, PhuNHa), 7.93-7.88 (m, 2H, ValNH+isoAspNH), 7.80 (d, *J* = 8.0 Hz, 1H, ArH_6), 7.76 (br d, 1H, AlaNH), 7.35 (d, *J* = 7.6 Hz, 2H, $ArH_{2'6'}$), 7.18-7.09 (m, 2H, $ArH_{3,5}$), 7.06 (d, *J* = 7.6 Hz, 2H, $ArH_{3'5'}$), 6.93 (dd, *J* = 7.2, 6.8 Hz, 1H, ArH_4), 4.41-4.30 (m, 2H, PhuHα+isoAspHα), 4.16 (dd, *J* = 7.2, 6.8 Hz, 1H, AlaHα), 4.00-3.92 (m, 1H, LeuHα), 3.71 (dd, *J* = 8.0, 6.4 Hz, 1H, ValHα), 2.78 (d, *J* = 7.2 Hz, 2H, PhuHβ), 2.62-2.52 (m, 2H, isoAspHβ), 2.33-2.24 (m, 1H, ValHβ), 2.23 (s, 3H, $ArCH_3$), 1.39-1.31 (m, 2H, LeuHβ), 1.25 (d, *J* = 6.8 Hz, 3H, AlaCH_3), 1.12-1.03 (m, 1H, LeuHγ), 0.85 (d, *J* = 6.4 Hz, 3H, ValCH_3), 0.80 (d, *J* = 6.4 Hz, 3H, ValCH_3), 0.76 (d, *J* = 6.4 Hz, 3H, LeuCH_3), 0.68 (d, *J* = 6.0 Hz, 3H, LeuCH_3); **^{13}C-NMR** (100 MHz, DMSO-d_6) *δ* 172.6, 171.7, 171.5, 171.1, 169.9, 169.8, 152.7, 138.4, 137.5, 130.2, 130.1, 129.5, 127.6, 126.1, 122.6, 121.1, 117.6, 66.4, 59.9, 54.9, 52.0, 50.1, 36.7, 28.9, 23.7, 23.1, 21.1, 19.2, 18.6, 17.9, 17.7. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{35}H_{48}N_7O_8]^+$ 694.35644, found 694.35665 $[M+H]^+$.

c[(S)-Phu-Phe-Asp-Val-(S)-isoAsp] **14**. **^{1}H-NMR** (600 MHz, 8:2 DMSO-d_6/H_2O) *δ* 8.93 (s, 1H, PhuNHb), 8.72 (d, *J* = 7.2 Hz, 1H, PheNH), 8.21-8.16 (m, 2H, AspNH+PhuNH), 7.89 (s, 1H, PhuNHa), 7.82 (d, *J* = 7.8 Hz, 1H, ArH_6), 7.70 (d, *J* = 9.5 Hz, 1H, ValNH), 7.32 (d, *J* = 8.5 Hz, 2H, $ArH_{2'6'}$), 7.29 (t, *J* = 7.6 Hz, 2H, PheArH), 7.21-7.15 (m, 4H, PheArH+ArH_3), 7.13 (t, *J* = 7.7 Hz, 1H, ArH_5), 7.08 (d, *J* = 8.4 Hz, 2H, $ArH_{3'5'}$), 7.06 (d, *J* = 8.2 Hz, 1H, isoAspNH), 6.93 (dd, *J* = 7.8, 7.2 Hz, 1H, ArH_4), 4.57 (dt, *J* = 8.2, 5.5 Hz, 1H, isoAspHα), 4.31 (td, *J* = 7.6, 5.4 Hz, 1H, PhuHα), 4.26-4.19 (m, 2H, AspHα+ValHα), 4.02 (ddd, *J* = 11.3, 7.1, 4.6 Hz, 1H, PheHα), 3.22 (dd, *J* = 13.8, 4.6 Hz, 1H, PheHβ), 3.06-2.96 (m, 2H, PheHβ+PhuHβ), 2.86 (dd, *J* = 16.6, 7.8 Hz, 1H, PhuHβ), 2.69 (dd, *J* = 14.6, 5.1 Hz, 1H, isoAspHβ), 2.59-2.52 (m, 3H, isoAspHβ+AspHβ), 2.37-2.32 (m, 1H, ValHβ), 2.23 (s, 3H, ArCH3), 0.95-0.87 (m, 6H, ValCH_3); **^{13}C-NMR** (150 MHz, DMSO-d_6) *δ* 172.6, 172.1, 170.6, 170.5, 170.3, 169.7, 152.6, 138.5, 138.2, 137.4, 133.5, 131.1, 130.2, 129.6, 129.3, 129.1, 128.2, 127.4, 126.3, 126.1, 122.6, 121.0, 117.9, 57.6, 55.7, 55.5, 51.8, 48.4, 40.1, 36.9, 35.8, 35.1, 34.7, 30.7, 29.2, 19.8, 17.9, 17.5. **HRMS-ESI/QTOF** *m/z* calcd. for $[C_{39}H_{46}N_7O_{10}]^+$ 772.33062, found 772.33004 $[M+H]^+$.

c[(S)-Phu-Phe-Ala-Val-(S)-isoAsp] **15**. **^{1}H-NMR** (400 MHz, 8:2 DMSO-d_6/H_2O) *δ* 8.94 (s, 1H, PhuNHb), 8.51 (d, *J* = 7.2 Hz, 1H, PheNH), 8.21-8.09 (m, 2H, AlaNH+PhuNH), 7.90 (s, 1H, PhuNHa), 7.86-7.75 (m, 2H, ArH_6+ValNH), 7.35-7.25 (m, 4H, $ArH_{2'6'}$+PheArH), 7.23-7.17 (m, 3H, PheArH), 7.16 (d, *J* = 8.0 Hz, 1H, ArH_3), 7.14-7.08 (m, 2H, ArH_5+isoAspNH), 7.04 (d, *J* = 8.2 Hz, 2H, $ArH_{3'5'}$), 6.93 (t, *J* = 7.4 Hz, 1H, ArH_4), 4.60 (dt, *J* = 9.3, 5.0 Hz, 1H, isoAspHα), 4.22-4.12 (m, 2H, ValHα+PhuHα), 4.12-4.04 (m, 1H, PheHα), 3.98 (t, *J* = 7.1 Hz, 1H, AlaHα), 3.22-3.15 (m, 1H, PheHβ), 3.07-2.98 (m, 1H, PheHβ), 2.64 (t, J = 5.7 Hz, 2H, isoAspHβ), 2.57 (d, J = 7.5 Hz, 2H, PhuHβ), 2.38-2.30 (m, 1H, ValHβ), 2.23 (s, 3H, ArCH3), 1.40 (d, J = 7.1 Hz, 3H, AlaCH$_3$), 0.90 (dd, J = 6.9 Hz, 6H, ValCH$_3$); **^{13}C-NMR** (100 MHz, DMSO-d_6) δ 172.1, 172.02, 172.00, 170.6, 170.5, 169.9, 152.7, 138.4, 138.2, 137.4, 131.1, 130.2, 129.3, 129.1, 128.2, 127.4, 126.3, 126.1, 122.6, 121.0, 117.9, 57.5, 55.8, 55.7, 50.7, 48.1, 36.8, 36.0, 35.2, 29.2, 19.8, 17.9, 17.5, 16.7. **HRMS-ESI/QTOF** *m/z* calcd. for [C$_{38}$H$_{46}$N$_7$O$_8$]$^+$ 728.34079, found 728.34111 [M+H]$^+$.

c[(R)-Phu-Leu-Asp-Phg-(S)-isoAsp] **16**. **^{1}H-NMR** (400 MHz, 8:2 DMSO-d_6/H$_2$O) δ 8.95 (s, 1H, PhuNHb), 8.57 (d, J = 8.3 Hz, 1H, PhuNH), 8.46 (d, J = 6.3 Hz, 1H, PhgNH), 8.14-8.09 (m, 2H, AspNH+LeuNH), 8.04 (d, J = 8.7 Hz, 1H, AspNH), 7.88 (s, 1H, PhuNHa), 7.81 (d, J = 8.2 Hz, 1H, ArH$_6$), 7.35 (d, J = 8.1 Hz, 2H, ArH$_{2'6'}$), 7.33-7.18 (m, 5H, PhgArH), 7.17-7.10 (m, 2H, ArH$_3$+ArH$_5$), 7.05 (d, J = 8.2 Hz, 2H, ArH$_{3'5'}$), 6.94 (td, J = 7.4, 1.3 Hz, 1H, ArH$_4$), 5.15 (d, J = 6.3 Hz, 1H, PhgHα), 4.56-4.47 (m, 2H, PhuHα+isoAspHα), 4.46-4.41 (m, 1H, AspHα), 3.97 (ddd, J = 10.4, 7.2, 4.9 Hz, 1H, LeuHα), 2.82 (d, J = 7.4 Hz, 2H, isoAspHβ), 2.76-2.69 (m, 2H, AspHβ), 2.64-2.52 (m, 2H, PhuHβ), 2.23 (s, 3H, ArCH3), 1.38-1.23 (m, 2H, LeuHβ), 1.21-1.13 (m, 1H, LeuHγ), 0.75 (dd, J = 29.4, 6.5 Hz, 6H, LeuCH3); **^{13}C-NMR** (100 MHz, DMSO-d_6) δ 172.3, 172.1, 171.6, 170.6, 169.7, 169.0, 168.7, 152.6, 138.33, 138.28, 137.5, 130.21, 130.18, 129.6, 129.5, 127.9, 127.5, 126.8, 126.1, 122.6, 121.0, 117.6, 57.7, 54.4, 52.3, 51.5, 49.7, 40.4, 37.3, 37.15, 37.14, 36.1, 23.7, 23.0, 21.3, 17.9. **HRMS-ESI/QTOF** *m/z* calcd. for [C$_{39}$H$_{46}$N$_7$O$_{10}$]$^+$ 772.33062, found 772.33102 [M+H]$^+$.

Note: The work described in this chapter has been published in the ***Journal of medicinal chemistry*** with slight modification and has been reproduced here with the permission of the copyright holder.

Anselmi, M.; Baiula, M.; Spampinato, S.; Artali, R.; He, T.; Gentilucci, L.* Design and Pharmacological Characterization of $\alpha_4\beta_1$ Integrin Cyclopeptide Agonists: Computational Investigation of Ligand Determinants for Agonism versus Antagonism. *J. Med. Chem.* **2023**, *66* (7), 5021-5040.

References

[1] Galletti, P.; Soldati, R.; Pori, M.; Durso, M.; Tolomelli, A.; Gentilucci, L.; Dattoli, S.D.; Baiula, M.; Spampinato, S.; Giacomini, D. Targeting integrins $\alpha_v\beta_3$ and $\alpha_5\beta_1$ with new β-lactam derivatives. *Eur. J. Med. Chem.* **2014**, *83*, 284-293.

[2] Baiula, M.; Galletti, P.; Martelli, G.; Soldati, R.; Belvisi, L.; Civera, M.; Dattoli, S.D.; Spampinato, S.; Giacomini, D. New β-Lactam Derivatives Modulate Cell Adhesion and Signaling Mediated by RGD-Binding and Leukocyte Integrins. *J. Med. Chem*. **2016**, *59*, 9721-9742.

[3] Martelli, G.; Baiula, M.; Caligiana, A.; Galletti, P.; Gentilucci, L.; Artali, R.; Spampinato, S.; Giacomini, D. Could Dissecting the Molecular Framework of β-Lactam Integrin Ligands Enhance Selectivity? *J. Med. Chem.* **2019**, *62*, 10156-10166.

[4] Lin, Kc.; Ateeq, H.S.; Hsiung, S.H.; Chong, L.T.; Zimmerman, C.N.; Castro, A.; Lee, W.C.; Hammond, C.E.; Kalkunte, S.; Chen, L.L.; Pepinsky, R.B.; Leone, D.R.; Sprague, A.G.; Abraham, W.M.; Gill, A.; Lobb, R.R.; Adams. S.P. Selective, tight-binding inhibitors of integrin $\alpha_4\beta_1$ that inhibit allergic airway responses. *J. Med. Chem.* **1999**, *42*, 920-934.

[5] Abraham, W.M.; Gill, A.; Ahmed, A.; Sielczak, M.W.; Lauredo, I.T.; Botinnikova, Y.; Lin, K.C.; Pepinsky, B.; Leone, D.R.; Lobb, R.R.; Adams, S.P. A small-molecule, tight-binding inhibitor of the integrin $\alpha_4\beta_1$ blocks antigen-induced airway responses and inflammation in experimental asthma in sheep. *Am. J. Respir. Crit. Care. Med.* **2000**, *162*, 603-611.

[6] Karanam, B.V.; Jayra, A.; Rabe, M.; Wang, Z.; Keohane, C.; Strauss, J.; Vincent, S. Effect of enalapril on the in vitro and in vivo peptidyl cleavage of a potent VLA-4 antagonist. *Xenobiotica* **2007**, *37*, 487-502.

[7] Fisher, A.L.; DePuy, E.; Jayaraj, A.; Raab, C.; Braun, M.; Ellis-Hutchings, M.; Zhang, J.; Rogers, J.D.; Musson, D.G. LC/MS/MS plasma assay for the peptidomimetic VLA 4 antagonist I and its major active metabolite II: For treatment of asthma by inhalation. *J. Pharm. Biomed. Anal.* **2002**, *27*, 57-71.

[8] Singh, J.; Van Vlijmen, H.; Liao, Y.; Lee, W.C.; Cornebise, M.; Harris, M.; Shu, I.; Gill, A.; Cuervo, J.H.; Abraham, W.M.; Adams, S.P. Identification of Potent and Novel $\alpha_4\beta_1$ Antagonists Using in Silico Screening. *J. Med. Chem.* **2002**, *45*, 2988-2993.

[9] Qasem, A.R.; Bucolo, C.; Baiula, M.; Spartà, A.; Govoni, P.; Bedini, A.; Fascì, D.; Spampinato, S. Contribution of $\alpha_4\beta_1$ integrin to the antiallergic effect of levocabastine. *Biochem. Pharmacol.* **2008**, *76*, 751-762.

[10] Dattoli, S.D.; Baiula, M.; De Marco, R.; Bedini, A.; Anselmi, M. Gentilucci, L.; Spampinato, S. DS-70, a novel and potent α_4 integrin antagonist, is an effective treatment for experimental allergic conjunctivitis in guinea pigs. *Br. J. Pharmacol.* **2018**, *175*, 3891-3910.

[11] Sugg, E.E.; Kimery, M.J.; Ding, J.M.; Kenakin, D.C.; Miller, L.J.; Queen, T.J. Rimele. K.L. CCK-A Receptor Selective Antagonists Derived from the CCK-A Receptor Selective Tetrapeptide Agonist Boc-Trp-Lys(Tac)-Asp-MePhe-NH_2 (A-71623). *J. Med. Chem.* **1995**, *38*, 207-211.

[12] At present, L-Boc-4-amino-phenylalanines is commercially available; at the time of the syntheses of our cyclopeptides, its commercialization was discontinued.

[13] White, C.J.; Yudin, A.K. Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **2011**, *3*, 509-524.

[14] Bochen, A.; Marelli, U.K.; Otto, E.; Pallarola, D.; Mas-Moruno, C.; Di Leva, F.S.; Boehm, H.; Spatz, J.P.; Novellino, E.; Kessler, H.; Marinelli, L. Biselectivity of isoDGR peptides for fibronectin binding integrin subtypes $\alpha_5\beta_1$ and $\alpha_v\beta_6$: Conformational control through flanking amino acids. *J. Med. Chem.* **2013**, *56*, 1509-1519.

[15] Malesevic, M.; Strijowski, U.; Bächle, D.; Sewald, N. An improved method for the solution cyclization of peptides under pseudo-high dilution conditions. *J. Biotechnol.* **2004**, *112*, 73-77.

[16] Caputo, R.; Longobardo, L. Enantiopure β^3-amino acids-2,2-d_2 via homologation of proteinogenic α-amino acids, *Amino Acids* **2007**, *32*, 401-404.

[17] Gentilucci, L.; De Marco, R.; Cerisoli, L. Chemical modifications designed to improve peptide stability: incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization. *Curr. Pharm. Des.* **2010**, *16*, 3185-3203.

[18] Temussi, P.A.; Picone, D.; Saviano, G.; Amodeo, P.; Motta, A.; Tancredi, T.; Salvadori, S.; Tomatis, R. Conformational analysis of an opioid peptide in solvent media that mimic cytoplasm viscosity. *Biopolymers* **1992**, *32*, 367-372.

[19] Borics, A.; Tóth, G. Structural comparison of μ-opioid receptor selective peptides confirmed four parameters of bioactivity. *J. Mol. Graph. Model.* **2010**, *28*, 495-505.

[20] Smith, J.A.; Pease, L.G.; Kopple, K.D. Reverse turns in peptides and Protein. *Crit. Rev. Biochem.* **1980**, *8*, 315-399.

[21] Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

[22] Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. **1983**, *79*, 926-935.

[23] Frank, A.O.; Otto, E.; Mas-Moruno, C.; Schiller, H.B.; Marinelli, L.; Cosconati, S.; Bochen, A.; Vossmeyer, D.; Zahn, G.; Stragies, R.; Novellino, E.; Kessler, H. Conformational control of integrin-subtype selectivity in isoDGR peptide motifs: A biological switch. *Angew. Chem. Int. Ed.* **2010**, *49*, 9278-9281.

[24] Schuman, F.; Müller, A.; Koksch, M.; Müller, G.; Sewald, N. Are β-amino acids γ-turn mimetics? Exploring a new design principle for bioactive cyclopeptides. *J. Am. Chem. Soc*. **2000**, *122*, 12009-12010.

[25] Sartori, A.; Bugatti, K.; Portioli, E.; Baiula, M.; Casamassima, I.; Bruno, A.; Bianchini, F.; Curti, C.; Zanardi, F.; Battistini, L. New 4-aminoproline-based small molecule cyclopeptidomimetics as potential modulators of $\alpha_4\beta_1$ integrin. *Molecules* **2021**, *26*, 6066.

[26] Vanderslice, P.; Biediger, R.J.; Woodside, D.G.; Brown, W.S.; Khounlo, S.; Warier, N.D.; Gundlach IV, C.W.; Caivano, A.R.; Bornmann, W.G.; D.S. Maxwell, McIntyre, B.W.; Willerson, J.T.; Dixon, R.A.F. Small Molecule Agonist of Very Late Antigen-4 (VLA-4) Integrin Induces Progenitor Cell Adhesion. *J. Biol. Chem.* **2013**, *288*, 19414-19428.

[27] Hickman, A.; Koetsier, J.; Kurtanich, T.; Nielsen, M.C.; Winn, G.; Wang, Y.; Bentebibel, S.E.; Shi, L.; Punt, S.; Williams, L.; et al. LFA-1 activation enriches tumor-specific T cells in a cold tumor model and synergizes with CTLA-4 blockade. *J. Clin. Invest.* **2022**, *132* (13), e154152.

[28] Huey, R.; Morris, G.M.; Olson, A.J.; Goodsell, D.S. A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem*. **2007**, *28*, 1145-1152.

[29] Yu, Y.; Zhu, J.; Mi, L.Z.; Walz, T.; Sun, H.; Chen, J.; Springer, T.A. Structural specializations of $\alpha_4\beta_7$, an Integrin that Mediates Rolling Adhesion. *J. Cell Biol.* **2012**, *196*, 131-146.

[30] Xia, W.; Springer, T.A. Metal ion and ligand binding of integrin $\alpha_5\beta_1$. *Proc. Natl. Acad. Sci. USA*. **2014**, *111*, 17863-17868.

[31] Xiong, J.P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S.L; Arnaout, M.A. Crystal structure of the extracellular segment of integrin $\alpha_V\beta_3$ in complex with an Arg-Gly-Asp ligand. *Science*. **2002**, *296*, 151-155.

[32] Springer, T.A.; Zhu, J.; Xiao, T. Structural basis for distinctive recognition of fibrinogen γC peptide by the platelet integrin $\alpha_{IIb}\beta_3$. *J. Cell Biol*. **2008**, *182*, 791-800.

[33] Carbone, J.; Ghidini, A.; Romano, A.; Gentilucci, L.; Musiani, F. PacDOCK: A Web Server for Positional Distance-Based and Interaction-Based Analysis of Docking Results. *Molecules* **2022**, *27*, 6884.

[34] Modlinger, A.; Kessler, H. Spatial screening for the identification of the bioactive conformation of integrin ligands. *Top. Curr. Chem.* **2006**, *272*, 1-50.

[35] Vasconcelos, D.; Chaves, B.; Albuquerque, A.; Andrade, L.; Henriques, A.; Sartori, G.; Savino, W.; Caffarena, E.; Martins-Da-Silva, J.H. Development of New Potential Inhibitors of β_1 Integrins through In Silico Methods-Screening and Computational Validation. *Life* **2022**, *12*, 932.

[36] Van Agthoven, J.F.; Xiong, J.P.; Alonso, J.L.; Rui, X.; Adair, B.D.; Goodman, S.L.; Arnaout, M.A. Structural basis for pure antagonism of integrin $\alpha_V\beta_3$ by a high-affinity form of fibronectin. *Nat. Struct. Mol. Biol*. **2014**, *21*, 383-388.

[37] Zhu, J.; Zhu, J.; Springer, T.A. Complete integrin headpiece opening in eight steps. *J. Cell Biol.* **2013**, *201*, 1053-1068.

[38] Paladino, A.; Civera, M.; Belvisi, L.; Colombo, G. High Affinity vs. Native fibronectin in the modulation of $\alpha_v\beta_3$ integrin conformational dynamics: insights from computational analyses and implications for molecular design. *PLoS Comput. Biol*. **2017**, *13*, e1005334.

[39] Lin, F-Y.; Zhu, J.; Eng, E.T.; Hudson, N.E.; Springer, T.A. β-Subunit Binding Is Sufficient for Ligands to Open the Integrin $\alpha_{IIb}\beta_3$ Headpiece. *J. Biol. Chem*. **2016**, *291*, 4537-4546.

[40] Miyazaki, N.; Iwasaki, K.; Takagi, J. A systematic survey of conformational states in β_1 and β_4 integrins using negative-stain electron microscopy. *J. Cell Sci.* **2018**, *131*, jcs21675.

[41] Su, Y.; Xia, W.; Li, J.; Walz, T.; Humphries, M.J.; Vestweber, D.; Cabañas, C.; Lu, C.; Springer, T.A. Relating conformation to function in integrin $\alpha_5\beta_1$. *Proc. Natl. Acad. Sci. USA*. **2016**, *113*, E3872-E3881.

[42] Schumacher, S.; Dedden, D.; Vazquez Nunez, R.; Matoba, K.; Takagi, J.i; Biertümpfel, C.; Mizuno, N. Structural insights into integrin $\alpha_5\beta_1$ opening by fibronectin ligand. *Sci. Adv.* **2021**, *7*, eabe9716.

[43] Bier, D.; Thiel, P.; Briels, J.; Ottmann, C. Stabilization of Protein-Protein Interactions in chemical biology and drug discovery. *Prog. Biophys. Mol. Biol.* **2015**, *119*, 10-19.

[44] Andrei, S.A.; Sijbesma, E.; Hann, M.; Davis, J.; O'Mahony, G.; Perry, M.W.D.; Karawajczyk, A.; Eickhoff, J.; Brunsveld, L.; Doveston, R.G.; Milroy, L.G.; Ottmann, C. Stabilization of protein-protein interactions in drug discovery. *Expert Opin. Drug Discovery* **2017**, *12*, 925-940.

[45] Zhao, J.; Santino, F.; Giacomini, D.; Gentilucci, L. Integrin-Targeting Peptides for the Design of Functional Cell-Responsive Biomaterials. *Biomedicines* **2020**, *8*, 307.

[46] Kantlehner, M.; Finsinger, D.; Meyer, J.; Schaffner, P.; Jonczyk, A.; Diefenbach, B.; Nies, B.; Kessler, H. Selective RGD mediated adhesion of osteoblasts at surfaces of implants. *Angew. Chem. Int. Ed.* **1999**, *38*, 560-562.

[47] Greco, A.; Maggini, L.; De Cola, L.; De Marco, R.; Gentilucci, L. Diagnostic implementation of fast and selective integrin-mediated adhesion of cancer cells on functionalized zeolite L monolayers. *Bioconj. Chem.* **2015**, *26*, 1873-1878.

[48] Bedini, A.; Baiula, M.; Spampinato, S. Transcriptional activation of human mu-opioid receptor gene by insulin-like growth factor-I in neuronal cells is modulated by the transcription factor REST. *J. Neurochem.* **2008**, *105*, 2166-2178.

[49] Stewart, J.J. MOPAC: a semiempirical molecular orbital program. *J. Comput. Aided. Mol. Des*. **1990**, *4*, 1-105.

[50] Meng, E.C.; Pettersen, E.F.; Couch, G.S. et al. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **2006**, *7*, 339.

[51] Needleman, S.B.; Wunsch, C.D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol*. **1970**, *48*, 443-453.

[52] Shapovalov, M.V.; Dunbrack, R.L. Jr. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* **2011**, *19*, 844-858.

[53] Word, J.M.; Lovell, S.C.; Richardson, J.S.; Richardson, D.C. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **1999**, *285*, 1735-1747.

[54] Søndergaard, C.R.; Olsson, M.H.; Rostkowski, M.; Jensen, J.H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J. Chem. Theory Comput*. **2011**, *7*, 2284-2295.

[55] Olsson, M.H.; Søndergaard, C.R.; Rostkowski, M.; Jensen, J.H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput*. **2011**, *7*, 525-537.

[56] Morris, G.M.; Goodsell, D.S.; Halliday R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem*. **1998**, *19*, 1639-1662.

[57] Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity-a rapid access to atomic charges. *Tetrahedron*. **1980**, *36*, 3219-3228.

[58] Lindahl, E.; Hess, B.; Spoel, D.V.D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *Mol. Model. Annu.* **2001**, *7*, 306-317.

[59] Wang, J., Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157-1174.

[60] Wang, J.; Wang, W.; Kollman, P.A.; Case, D.A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model*. **2006**, *25*, 247-260.

[61] Sousa da Silva, A.W.; Vranken, W.F. ACPYPE-AnteChamber PYthon Parser interfacE. *BMC Res. Notes* **2012**, *5*, 367.

[62] Aprà, E.; Bylaska, E.J.; de Jong, W.A.; et al. NWChem: Past, present, and future. *J. Chem. Phys.* **2020**, *152*, 184102.

[63] Valiev, M.; Bylaska, E.J.; Govind, N.; Kowalski, K.; Straatsma, T.P.; Van Dam, H.J.J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.L.; de Jong, W.A. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun*. **2010**, *181*, 1477-1489.

Chapter IV. Fabrication and Characterization of Versatile Integrin-targeted Nanoplatforms for Diagnostic and Therapeutic Applications

4.1 Introduction

Nanoparticles (NPs) are defined as solid colloidal particles ranging in size from 10 to 1000 nm.[1] However, particles > 200 nm are not practically pursued and nanomedicine often refers to devices < 200 nm, which are comparable to the width of microcapillaries. In particular, for in vivo applications, such as imaging and drug delivery, nanomaterials, which are mostly in the range of 10-200 nm, provide unique properties and advantages, including: 1) enabling micro-scale or surface chemistry that integrates multiple functions on a single platform; 2) delivering large payload of imaging or therapeutic agents; 3) prolonging blood circulation or altering pharmacokinetics and biodistribution for improved delivery; and 4) introducing target specificity to a molecular biomarker or diseased tissue.

NPs accumulate preferentially in tumors or inflammatory sites via passive drug targeting due to the well-known enhanced permeation and retention effect (EPR) and impaired lymphatic drainage in diseased tissues.[2] However, unfunctionalized, native NPs often struggle in the complex environment of the human body.[3] The cellular uptake of naked particles can be very limited, because the cell membrane serves as a semi-permeable barrier, which efficiently protects the cell interior from the environment. In addition, proteins are known to adsorb onto the NP surfaces (so-called “protein corona”) in biological environments due to the high surface free energy, these protein coronas may alter the biological distribution, effects, and toxicity of the NPs.[4]

Interestingly, the accumulation of nanoparticles in tumors can be dramatically increased by adopting active targeting strategies.[5] An actively targeted nanoparticle is usually obtained by functionalizing NP surfaces with ligands for recognition by specific receptors/antigens on target cancer cells. In this regard, integrins may play a pivotal role in tumor targeting because of their overexpression on many tumor cell types and neo-angiogenic vessels. As previously mentioned, integrins are heterodimeric, noncovalently bound transmembrane glycoproteins that mediate cell-cell and cell-matrix adhesion. They regulate a diverse array of cellular functions such as cell proliferation, death, differentiation, and migration. As a result, integrins play an essential role in both physiological and pathological conditions, and notably, disease progression in various tumor types is correlated with the overexpression of the $\alpha_v\beta_3$, $\alpha_v\beta_5$, $\alpha_5\beta_1$, $\alpha_4\beta_1$ and $\alpha_v\beta_6$.[6]

Moreover, nanoparticles can be tailor-made to achieve a wide range of biomedical purposes.[1,5,7,8] For example, different nanomaterials, such as gold, silver, silica, or polymers, can impart unique optical, magnetic, or conductive properties, making nanoparticles highly effective in imaging or therapeutic applications. On the other hand, NPs can enhance drug targeting, increase circulation time, and effectively penetrate biological barriers via surface chemical modifications.

In this chapter, three different types of nanoparticles are discussed: metallic gold nanoparticles (AuNPs), poly(lactic-*co*-glycolic acid) based nanoparticles (PLGA NPs), and silica core/PEG shell nanoparticles (Si-PF127 NPs). For clarity, the chapter is organized into two parts based on the integrin-targeting recognition motifs employed: LDV-based NPs (part 1, AuNPs and PLGA NPs) and RGD-based NPs (part 2, silica core/PEG shell NPs).

Part 1: Preparation and Characterization of Gold and PLGA Nanoparticles, Coated with LDV-based α_4 Integrin Ligands

As discussed in Chapter 2 and 3, integrin $\alpha_4\beta_1$ belongs to the leukocyte-specific receptors and plays a crucial role in the development and sustainment of inflammation, in several inflammation-related diseases. In addition, several types of tumor cells express $\alpha_4\beta_1$ integrin, i.e. melanoma, multiple myeloma and ovarian cancer. Its interaction with VCAM-1 increases transendothelial migration and contributes to metastasis to distant organs.[6] Therefore, the routine blockade of this receptor by drug molecules has been implemented in the pharmaceutical industry, exemplified by the marketed drug, Natalizumab, for the treatment of multiple sclerosis.

Alternatively, the activation of $\alpha_4\beta_1$ integrin could represent a promising therapeutic strategy in specific pathological conditions. In contrast to the blockade of integrin functions by antagonists, the activation of $\alpha_4\beta_1$ integrin might represent an alternative strategy to perturb the progression of cell migration. Hence, agonists can be utilized to prevent the release of adherent cells.[9] However, very few potent agonists are currently available, and their use in vivo may be limited due to issues such as poor enzymatic stability and bioavailability. By adopting a peptidomimetic strategy, the enzymatic stability can be significantly improved, as demonstrated in Chapter 3. Besides, by using appropriate linkers, integrin-targeting peptides can be anchored onto a variety of organic or inorganic nanomaterials, surfaces, or polymers to enrich diverse applications in diagnostics and therapeutics, see selected examples in Chapter 2.

In recent years, metal nanoparticles have generated much interest in biomedicine due to their unique physical and chemical properties and have widely been employed in tumor targeting. Among them, gold nanoparticles (AuNPs) are particularly interesting for cancer diagnosis and therapy. Their relatively easy preparation, tunable size, and facile surface chemistry render AuNPs very attractive as carriers for drug delivery. More importantly, their surface plasmon resonance (SPR) enhanced light scattering and absorption make colloidal AuNPs highly appealing candidates for optical imaging and hyperthermia therapy.[10]

On the other hand, poly(D,L-lactic-co-glycolic acid) (PLGA) is composed of varying ratios of lactic acid and glycolic acid monomer units which plays an important role in drug release. It is the most widely used as a versatile and clinically approved polymer for the preparation of nanoparticles.[11] It is approved for use in drug delivery by Food and Drug Administration (FDA) and European Medicine Agency (EMA). PLGA is metabolized through the kerbs cycle to form H_2O and CO_2 with the nontoxic end products. It is soluble in various organic solvents such as acetone, dichloromethane, tetrahydrofuran, ethyl acetate, and chloroform which is synergistic advantageous for nanoparticle synthesis. Moreover, it has some exceptional properties such as wide range of degradation rates, stability during long-term storage, and high encapsulation efficiency. As a result, PLGA NPs are well-suited drug delivery vehicles, providing sustained and controlled release of therapeutic agents.

Taken together, the conjugation of peptides to NPs represents an effective approach to addressing the intrinsic drawbacks of the peptides, allowing the access to a variety of biomedical uses. And in turn the tumor-targeting precision of NPs can be strongly improved by conjugation with peptide ligands addressing integrin receptors overexpressed by cancer cells.

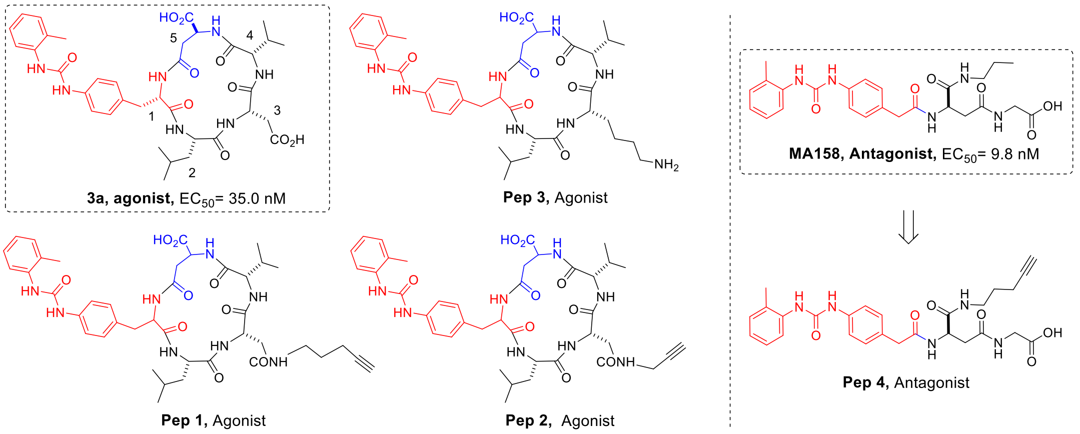
In this part, we address the design and synthesis of linkable $\alpha_4\beta_1$ integrin agonist/antagonist ligands, their conjugation to Au/PLGA nanoparticles, as well as the biological evaluation of the resulting nanostructures for their potential applications in diagnostics and theranostics.

4.2 Results and Discussion

4.2.1 Design and synthesis of linkable $\alpha_4\beta_1$ integrin agonist/antagonist ligands

As discussed in Chapter 3, a minilibrary of LDV-containing cyclopentapeptides decorated with α_4-targeting moiety MPUPA was constructed. Notably, peptide **3a** was identified as a potent agonist with an EC_{50} of 35.0 nM in the cell adhesion assay on $\alpha_4\beta_1$-expressing Jurkat E6.1 cell. Moreover, structure-activity relationship studies revealed that isoAsp^5 carboxylate, rather than Asp^3, is strictly necessary for receptor binding. This allows for the modification and conjugation of Asp^3 to nanomaterials, enabling its potential application in integrin targeted therapy. Accordingly, peptide-alkynes **1** and **2**, varying different carbon lengths of spacers, and amine peptide **3**, obtained by replacing the Asp^3 in parent **3a** with Lys, were designed for conjugation purposes (Figure 4.1).

As for the design of the antagonistic peptide, MA158 (MPUPA-(*R*)-isoAsp(NHPr)-Gly-OH), previously developed by our group and already discussed in Chapter 2, served as the parent modal. MA158 exhibited high affinity in the nanomolar range in $\alpha_4\beta_1$ integrin-mediated cell adhesion assay and demonstrated high enzymatic stability in vitro when incubated in mouse serum.[12] Consequently, the antagonistic peptide ligand **4** was derived from MA158, featuring an alkyne group in order to facilitate anchoring onto the nanoparticles' surface via CuAAC (copper-catalyzed azide-alkyne cycloaddtion) chemistry (Figure 4.1).



3a, agonist, EC_{50}= 35.0 nM

Pep 3, Agonist

MA158, Antagonist, EC_{50}= 9.8 nM

Pep 1, Agonist

Pep 2, Agonist

Pep 4, Antagonist

Figure 4.1 Chemical structures of linkable ligands used in this study.

The linear peptides were prepared by standard solid phase peptide synthesis, with Fmoc-protected amino acids. Asp was modified to corresponding alkyne-substituted amino acids prior to use in the solid-phase synthesis. And the cyclization was carried out under pseudo-high-dilution conditions.[13] For detailed synthesis procedures, see Experimental Section. Finally, the structures were carefully confirmed by ESI-MS and NMR spectroscopy (Appendix B).

4.2.2 Preparation and characterization of peptide-nanoparticle conjugates

Synthesis of peptide-AuNPs conjugates

Firstly, spherical, citrate-capped AuNPs were obtained via the sodium citrate reduction method,[14] with an average diameter of 14 nm. The successful formation of AuNPs was primarily assured by color changes from yellow to clear to black and finally to dark red during the synthesis. The AuNPs were further characterized by UV-vis and DLS analysis. Subsequently, the citrate ions on the AuNPs surface were replaced by HS-PEG_{5000}-COOH through a ligand exchange process. The obtained PEG-AuNPs were characterized by UV-vis, DLS, and TEM analysis. Finally, the peptide conjugation was accomplished either by carbodiimide (CDI) chemistry or by CuAAC reaction, depending on the anchor groups presented on the peptidic ligands. For conjugation via CuAAC, the resulting PEG-AuNPs were further modified with a bifunctional amine-azide spacer via CDI chemistry, introducing a reactive azide group necessary for the final conjugation with alkyne-bearing peptides. Namely, these reactions yielded Pep **3**-AuNPs, Pep **1**-AuNPs, Pep **2**-AuNPs and Pep **4**-AuNPs, respectively. The peptide capped conjugates were well-characterized by using UV-vis spectroscopy, DLS, and TEM. The schematic synthesis route and the representative structures of the conjugates are shown in figure 4.2.

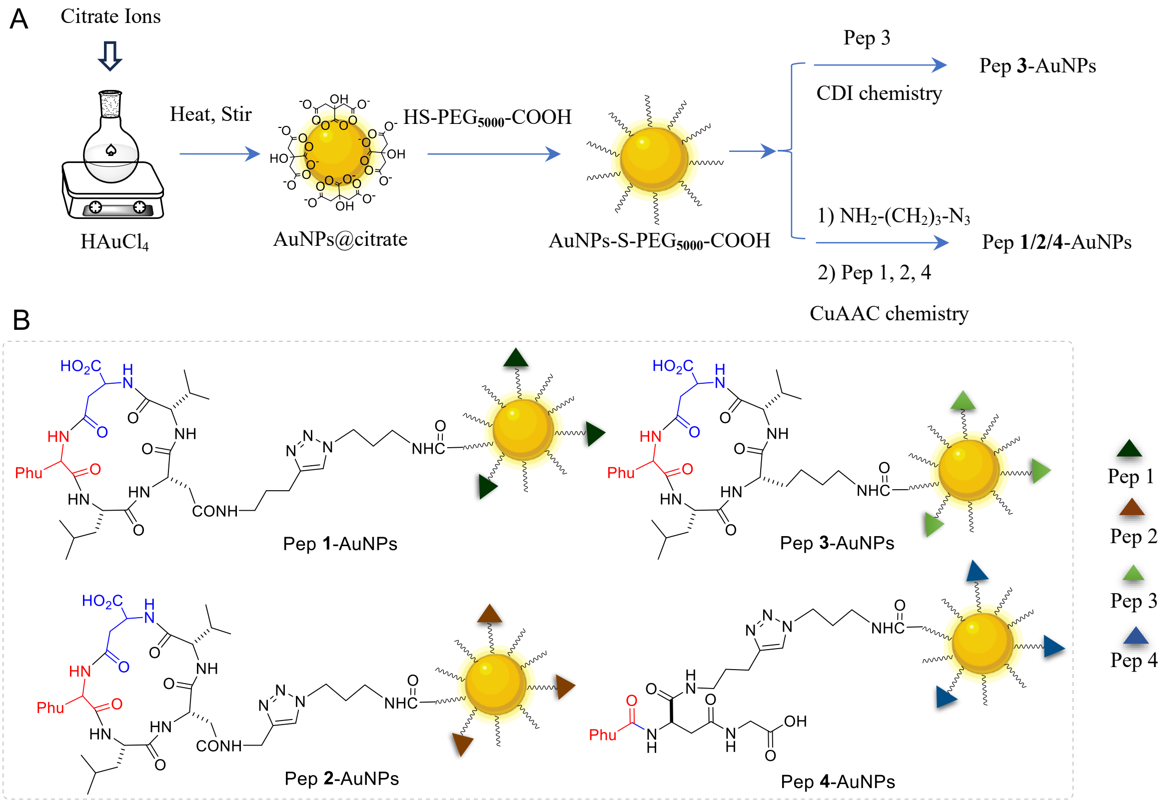
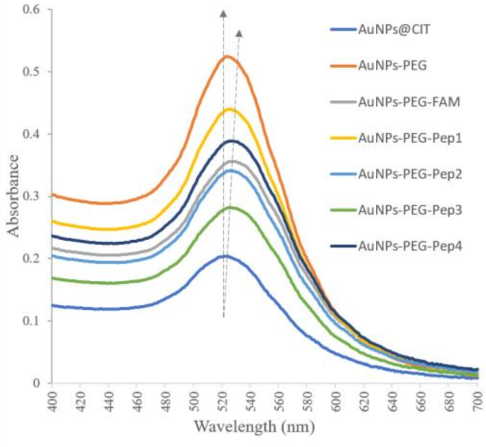


Figure 4.2 (A) Schematic synthesis of Pep-AuNPs. (B) Representative structures of Pep-AuNPs.

Characterization of peptide-AuNP conjugates

As widely reported, small AuNPs (1-20 nm) typically show a strong SPR band at around 520-530 nm in the visible region. The UV-visible spectra of synthesized particles indeed displayed SPR peaks within this range, confirming the formation of the stable AuNPs without aggregations (Figure 4.3). Additionally, a gradual red shift in the absorption maximum was observed for all peptide-AuNP conjugates as compared to the PEGylated AuNPs (λ_{max} = 523 nm) and bare AuNPs@cit (λ_{max} = 522 nm), which is an indication of successful loading of peptides over the surface of the AuNPs (Figure 4.3).



Conjugates	λ_{max} (nm)
AuNPs@cit	522
PEG-AuNPs	523
Pep **1**-AuNPs	525
Pep **2**-AuNPs	526
Pep **3**-AuNPs	525
Pep **4**-AuNPs	526
FAM-AuNPs	527

Figure 4.3 UV-Vis spectra monitoring the SPR band of AuNPs.

Further, the hydrodynamic diameter (d_H) and zeta potential (ζ potential) were measured by DLS before and after functionalization. DLS confirmed the hydrodynamic size of the bare AuNPs and PEG-AuNPs to be 17 nm and 31 nm, respectively. Following functionalization with peptides, the hydrodynamic diameters increased from 31 nm to 45 nm for Pep **3**-AuNPs, and 47 nm for Pep **4**-AuNPs, respectively (Figure 4.4). This increase in size is representative of desired conjugations. Moreover, ζ-potential measurements were conducted to further characterize the functionalzition. Bare AuNPs and PEG-AuNPs exhibited ζ-potential of -30 mV and -45 mV, respectively. Similarly, an increase in zeta potential was observed for all peptide-capped conjugates (-27 mV for Pep **3**-AuNPs, -29 mV for Pep **4**-AuNPs) when compared to naked PEG-AuNPs (-45 mV), as shown in figure 4.4. Overall, DLS measurements both in size and zeta potential confirmed the successful attachment of peptides to gold nanoparticles.

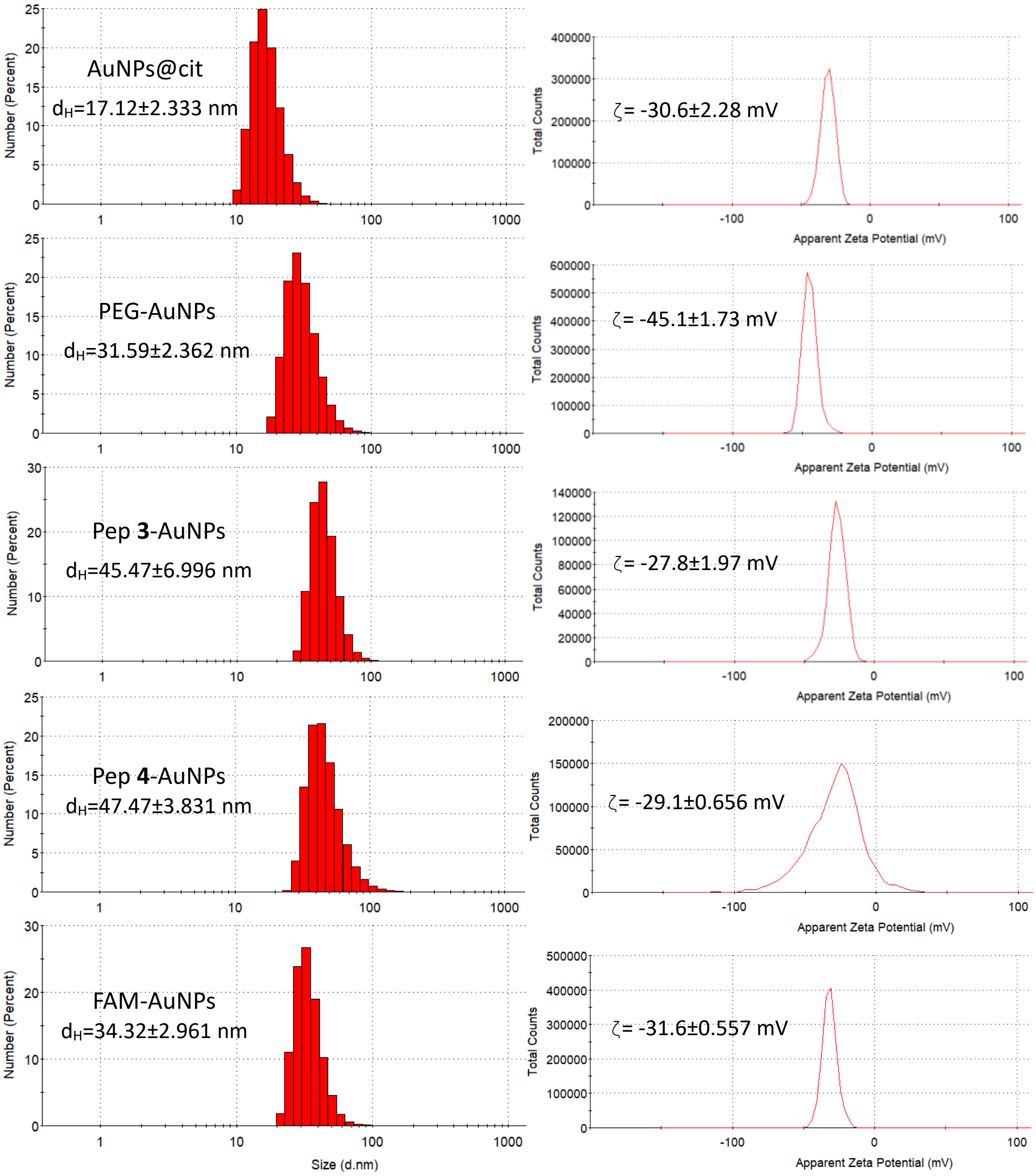


Figure 4.4 Size and zeta potential of AuNPs measured by DLS.

TEM images showed that all the AuNPs indeed exhibited a spherical shape with an average core size of 14 nm (Figure 4.5). It is important to note that the discrepancy in particle size measured by DLS and TEM is likely due to the coatings on the particle surface. Indeed, DLS measures the overall hydrodynamic diameter, which includes the nanoparticle core, the surrounding PEG layer, peptide coating, and any associated solvent molecules, generally leading to larger values. In contrast, TEM measures the core size of the nanoparticles in their dry state, typically showing smaller diameters than DLS. Interstingly, in the TEM images, the peptide-conjugates displayed an organic halo around the AuNPs surface respected to the PEG-AuNPs (Figure 4.5), further confirming the successful conjugation.[15-19]

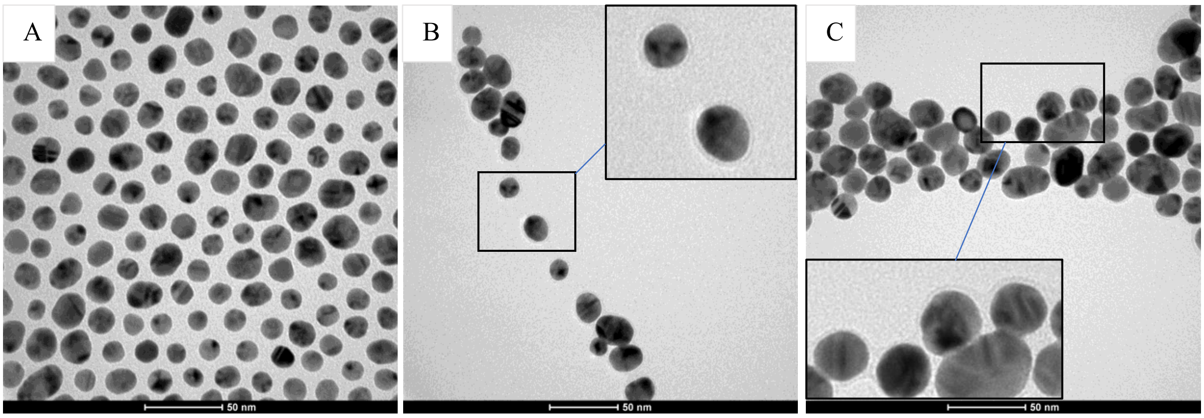


Figure 4.5 TEM images of A) PEG-AuNPs, B) Pep **3**-AuNPs, and C) Pep **4**-AuNPs. Scale bar =50 nm.

Next, the number of peptide molecules loaded onto AuNPs was determined by fluorometric quantification method. As for Pep **3**-AuNPs conjugates, obtained by EDC/sulfo-NHS chemistry, fluorescamine assay was conducted for quantification of primary amines on the Lys side chain of peptide **3**. Mechanistically, fluorescamine is not intrinsically fluorescent; however, fluorescence appears immediately after it reacts with primary amino groups at room temperature in a basic aqueous solution; and the excess reagent is subsequently hydrolyzed, yielding a water-soluble, nonfluorescent hydrolyzed product (Figure 4.6). These features and the great sensitivity attainable in the fluorescence measurement have made fluorescamine increasingly popular in peptide analysis.[20]

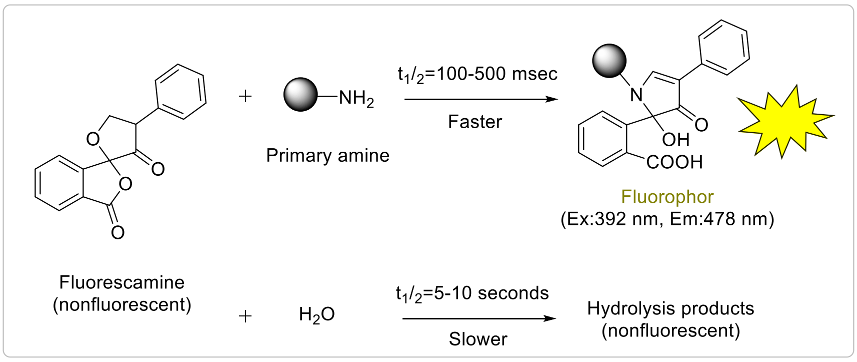


Figure 4.6 Mechanism of fluorescamine assay.

To quantify the amount of nonconjugated peptides, the reaction mixture after conjugation was centrifuged and the supernatants were collected, which contained the unreacted peptide **3**, and later treated with excess of fluorescamine in borate buffer. The resulting in-situ produced fluorescent solution was analyzed by a fluorophotometer. In addition, the calibration curve was obtained by increasing concentrations of standard peptide **3** (0-400 ppm) against fluorescamine under the same condition (Figure 4.7). Finally, the amount of conjugated peptides was evaluated based on the calculation of unloaded peptides in the supernatant and the total peptides added to the reaction. Accordingly, the difference showed that there are ~ 742 Pep **3** molecules per particle, which corresponds to approximately 1.2 Pep **3** molecules per nm² of surface area.

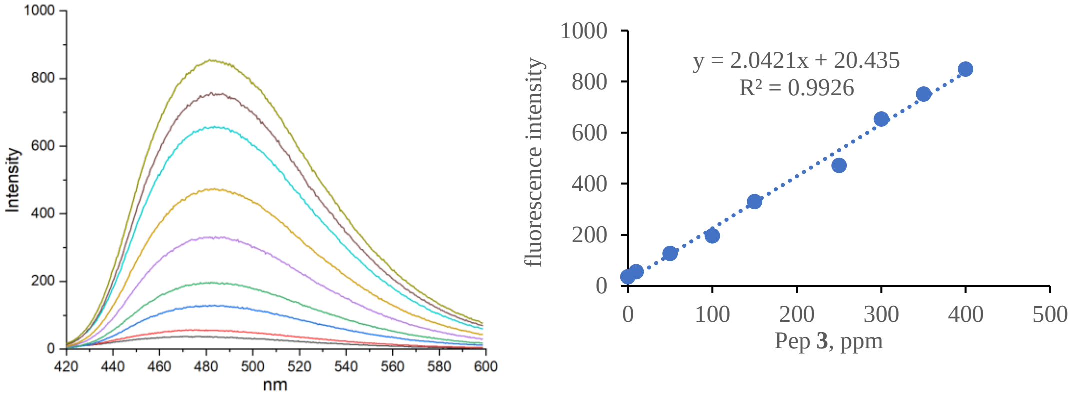


Figure 4.7 Fluorescence emission spectra of fluorescamine interacts with Pep **3** and the relative calibration curve.

On the other hand, for conjugates obtained via CuAAC reactions, the amount of conjugated alkyne-peptide was indirectly determined using a similar fluorescent molecule, FAM-alkyne (Figure 4.8).

FAM-alkyne, a modified fluorescein derivative, was selected as a fluorescent probe to ascertain the effective binding of the alkyne linker to the nanoparticles' surfaces and evaluate the percentage of functionalization (Figure 4.8).[21] Moreover, as the linker bearing the alkyne function are very similar to that of peptide-alkyne derivatives, we thought it could be a suitable study compound to verify the applicability of our conjugation protocol. Thus, the conjugation was conducted between the azide modified AuNPs and fluorescent probe FAM-alkyne under the same conditions used for alkyne peptides **1**, **2**, and **4**.

After conjugation, the resulting FAM-AuNPs were characterized by UV-vis and DLS analysis. Likewise, the SPR band was shifted to a longer wavelength as compared to PEG-AuNPs (λ_{max} = 527 vs λ_{max} = 523) (Figure 4.3). Furthermore, the hydrodynamic diameter was increased to 34 nm, with a zeta potential of -31 mV, as detected by DLS (Figure 4.4). As a result, we believe that the conjugation was successful. However, due to the fluorescence quenching effect of the gold nanoparticles,[22] the resulting FAM-AuNPs conjugates were temporarily non-fluorescent. Therefore, the FAM-AuNPs were digested using KCN, which released the surface-bound ligands thiol-PEG-FAM into solution. The fluorescence intensity of the released FAM derivatives was then measured using a fluorophotometer (Figure 4.8). Known concentrations of standard FAM-alkyne were used to generate the calibration curve (Figure 4.8). Accordingly, it is estimated that there are ~ 1119 FAM-alkyne molecules per particle, which corresponds to approximately 1.8 FAM-alkyne molecules per nm² of surface area. The conjugation is superior to that obtained by carbodiimide chemistry, likely due to the fact that CuAAC reaction is generally more efficient than CDI chemistry for bioconjugation.

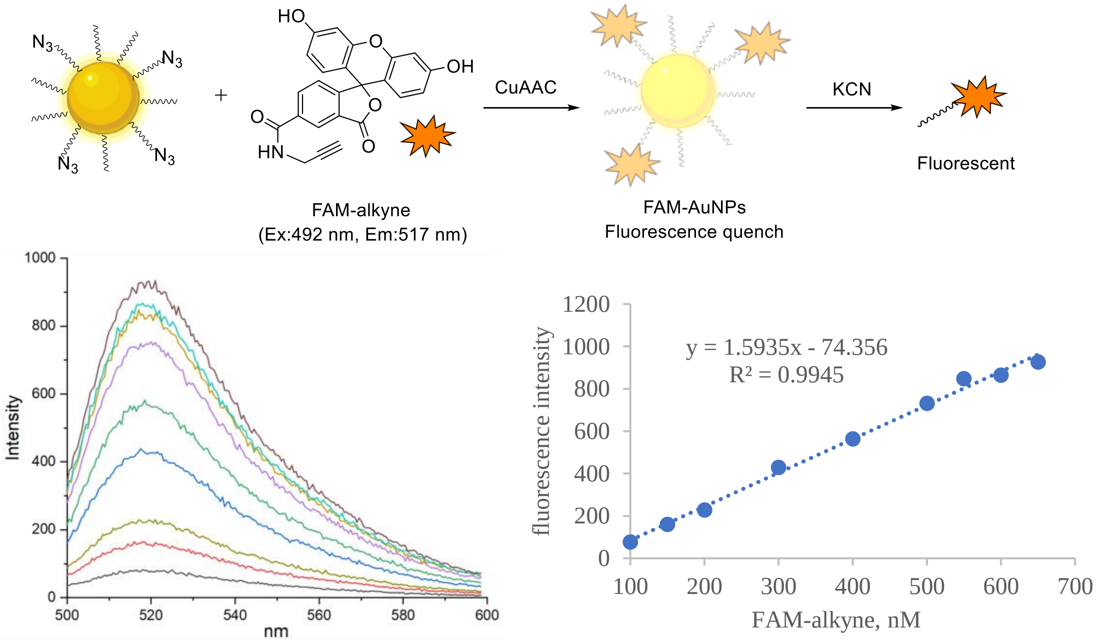


Figure 4.8 Top: mechanism of fluorometric quantification of FAM-AuNPs. Down: Fluorescence emission spectra of released FAM derivative and the relative calibration curve.

Synthesis of peptide-PLGA NPs conjugates

PLGA NPs are advantageous in terms of their sustained release property, protection of the drug from degradation and well-described formulations to deliver various therapeutics with favorable drug release profiles.[11] Hence, the use of biocompatible and biodegradable poly(D,L-lactic-co-glycolic acid) nanoparticles (PLGA NPs) as a drug delivery system has been widely investigated. Herein, we propose the synthesis of LDV-containning integrin ligands functionalized PLGA NPs, encapsulated with an anticancer drug curcumin, aiming to enhance precisely targeted cancer therapy.

To develop LDV-decorated PLGA NPs, we explored both in-synthesis and post-synthesis functionalization strategies.[23] In pre-functionalization method, the conjugation was carried out between carboxylic group of pristine PLGA and amino group of LDV peptide **3** by using EDC/sulfo-NHS chemistry prior to NP formation. In contrast, the post-functionalization approach involved forming the nanoparticles first and then installing the integrin ligands afterward, by using the same EDC/sulfo-NHS chemistry. Both strategies are widely used in the surface modification of NPs. The pre-functionalization approach has advantages such as ease of direct characterization using techniques like NMR and FT-IR.[24] However, a drawback of this method is that the ligand can be buried within the nanoparticle matrix, potentially reducing its accessibility. On the other hand, post-functionalization allows the conjugated ligands to remain more accessible on the NP surface, improving their targeting efficiency.

For nanoparticle fabrication, single emulsification-solvent evaporation method was adopted with some modifications (Figure 4.9).[25] In brief, the drug is co-dissolved with polymer (with or without peptide) in an organic phase, subsequently, the resulting organic solution was added dropwise to the aqueous phase containing the surfactant PVA, then the mixture was emulsified via a sonicator probe. After, the organic solvent was slowly evaporated under stirring at room temperature. Finally, particles were washed with MilliQ water and collected via ultra centrifugation for lyophilization and long-term storage.

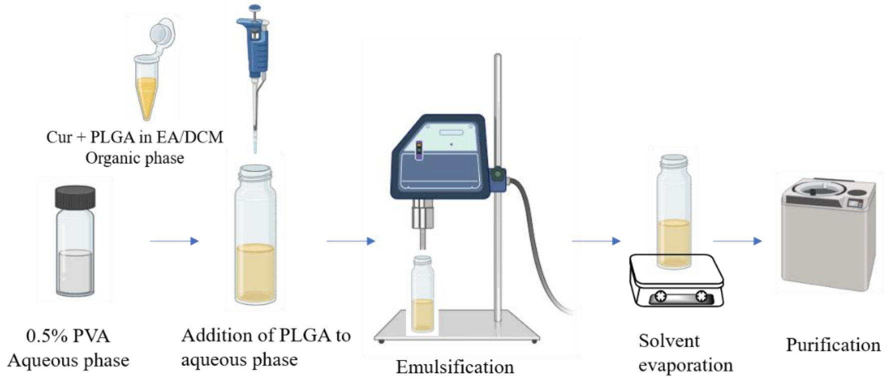


Figure 4.9 Schematic representation of the preparation of PLGA nanoparticles by single emulsification-solvent evaporation method.

The various formulations of PLGA nanoparticles were presented as following (Figure 4.10), which can be functionalized with either targeting peptide, anticancer drug curcumin, or a combination of both.

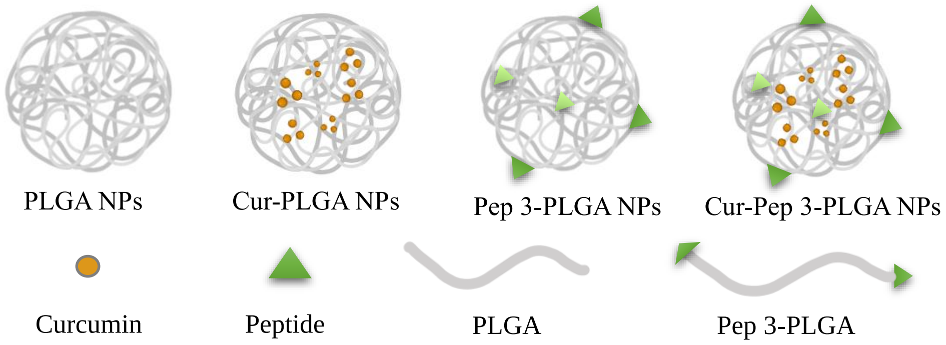


Figure 4.10 Schematic representation of different formulations of PLGA nanoparticles.

Characterization of peptide-PLGA NPs conjugates

As discussed in the last paragraph, we explored both pre- and post-functionalization strategies for surface decoration. In the initial experiment, we used convenient post synthesis functionalization method, as we did for the preparation of peptide grafted gold nanoparticles.

The hydrodynamic diameter and zeta potential of obtained nanomaterials were analysed by DLS measurement (Figure 4.11). For naked PLGA and curcumin encapsulated PLGA nanoparticles, the d_H were determined to be 157 nm and 167 nm, respectively. After conjugation of LDV Pep **3** to carboxylate PLGA nanoparticles via EDC/sulfo-NHS chemistry, a significant increase in size was detected by DLS (468 nm vs 157 nm), while the zeta potential remained almost unchanged, as compared to the bare PLGA NPs (-12.7 mV vs -13.9 mV). In contrast, for Pep **3**-PLGA NPs, obtained by pre-functionalization approach, both size and zeta potential were increased (d_H= 231 nm, zeta potential = -3.9 mV), which is consistent with the results obtained for peptides functionalized AuNPs. Consequently, pre-functionalization method was adopted for further preparations. As for Cur-Pep **3**-PLGA NPs, a hydrodynamic diameter of 239 nm was detected, with a zeta potential of -2.59 mV (Figure 4.11).

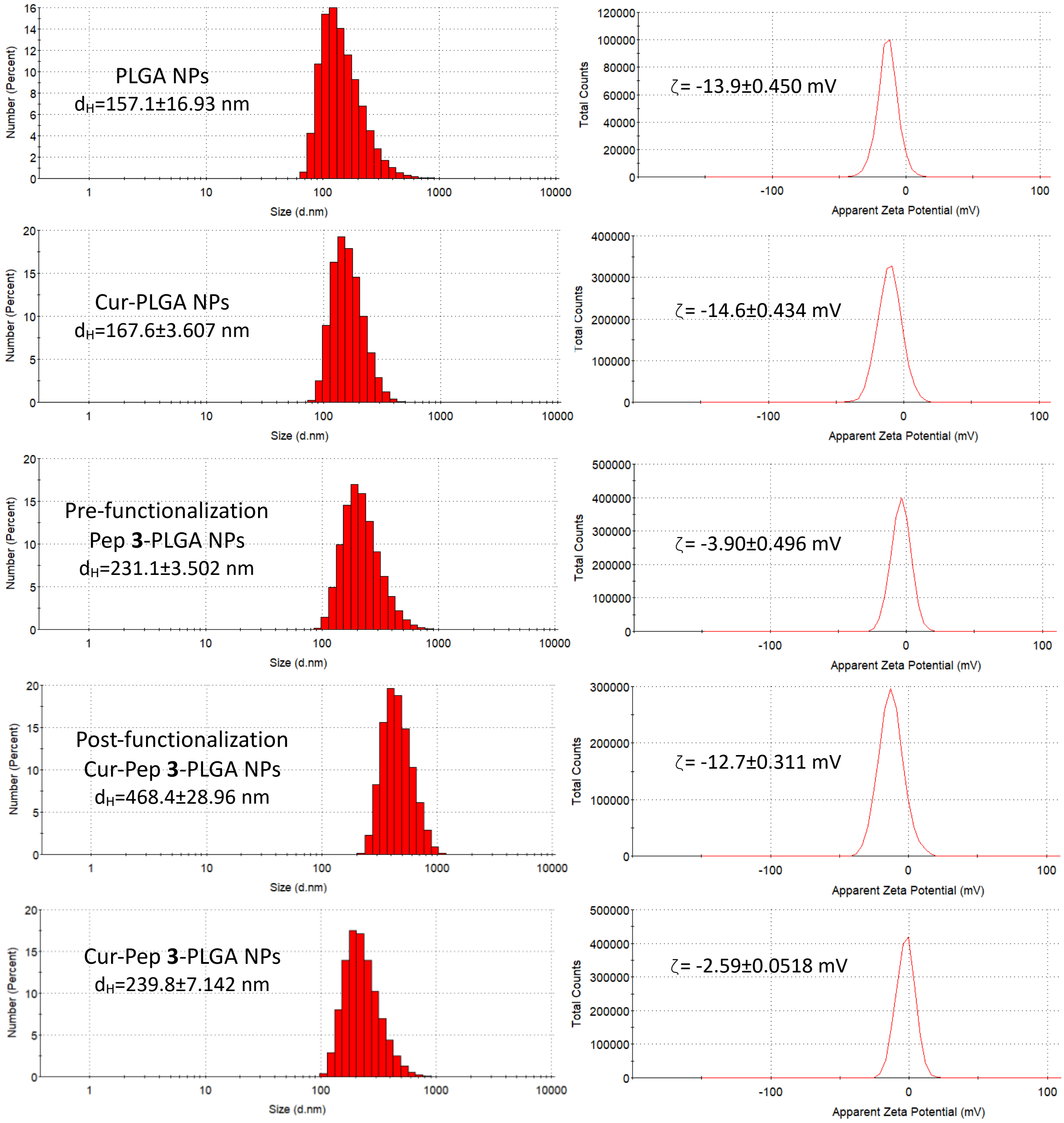


Figure 4.11 Size and zeta potential of PLGA NPs measured by DLS.

4.2.3 Drug loading assay

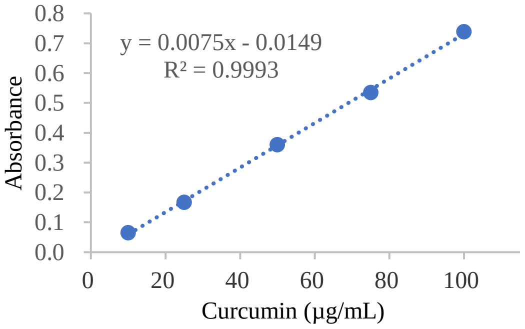
Encapsulation efficiency (EE%) and drug loading capacity (DL%) of curcumin in Cur-PLGA nanoparticles and Cur-Pep **3**-PLGA NPs were determined by UV spectrometry. Briefly, lyophilized NPs were dissolved in acetonitrile (1 mg/mL) using an ultrasonic water bath and a magnetic stirrer. The absorbance at 420 nm was measured in 4-fold diluted samples and the concentration of Curcumin was calculated using a standard curve of Cur constructed in acetonitrile (0-100 µg/mL) (Figure 4.12).

The EE% and DL% of NPs were calculated using following equations:

$$\text{EE\%} = \frac{\text{amount of Cur encapsulated in NPs}}{\text{initially added amount of Cur}} \text{ x } 100$$

$$\text{DL\%} = \frac{\text{amount of Cur encapsulated in NPs}}{\text{weight of NPs after lyophilization}} \text{ x } 100$$

Accordingly, the encapsulation efficiency and drug loading of unconjugated Cur-PLGA NPs were calculated to be 80% and 11%, respectively, while for Cur-Pep **3**-PLGA NPs, which were formed by assembling pre-functionalized Pep **3**-PLGA polymer and curcumin in one-pot, exhibited encapsulation efficiency and drug loading of 74% and 10%, respectively (Figure 4.12). As a result, the encapsulation efficiency and drug loading were not significantly affected before and after peptide conjugation. In contrast, when the peptide conjugation procedure is applied to a drug-loaded nanocarrier (post-functionalization), the encapsulation efficiency of the nanocarrier typically decreases at the end of the process. These reductions are probably the result of drug leakage from the nanocarrier during the additional washing and shaking steps required for the peptide conjugation process.[26]



NPs	EE (%)	DL (%)
Cur-PLGA	80	11
Cur-Pep **3**-PLGA	74	10

Figure 4.12 The calibration curve of standard curcumin in ACN within 0-100 µg/mL range and the canculated encapsulation efficiency and drug loading of PLGA NPs.

To summarize, in this part, we discussed the design and synthesis of linkable $\alpha_4\beta_1$ integrin agonist/antagonist ligands and their conjugation to Au/PLGA NPs. These nanostructures were thoroughly characterized by UV-vis, DLS, and TEM analyses. Currently, a comprehensive biological evaluation of the resulting particles is underway to assess their potential applications in $\alpha_4\beta_1$ integrin-targeted diagnostics and theranostics.

Part 2: Synthesis and Characterization of Silica-Core/PEG-Shell Nanoparticles, Multifunctionalized with RGD Integrin-Targeting and Antitumor Peptides

As the most studied integrin recognition motif, the tripeptide Arg-Gly-Asp (RGD) is recognized by a wide range of integrin superfamily members, including integrins $\alpha_v\beta_3$, $\alpha_v\beta_5$, $\alpha_v\beta_6$, $\alpha_5\beta_1$, and $\alpha_{IIb}\beta_3$, most of which are implicated in tumor progression. Hence, the RGD-based integrin ligands have been extensively investigated as potential antiangiogenic agents for cancer treatment. However, the failure of $\alpha_v\beta_3$ antagonist Cilengitide in clinical phase III trials for the treatment of glioblastoma dampened enthusiasm for this strategy.[27] Alternatively, RGD-conjugates and RGD-based nanoparticles have garnered growing attention as targeting ligands for drug delivery and precision therapy, providing a renewed focus in the field of integrin-targeted therapeutics.[28]

In addition, the target-specific nanoparticles can also be functionalized with a drug peptide ligand, with the aim to improve not only the selectivity of the conjugates, but also their antitumor activity over the free antitumor peptide. Since NPs conjugated to a targeting ligand have a higher capacity to enter cells, carrying the antitumor peptide and thus increasing its local concentration will, in turn, enhance the antitumor activity. In this regard, the retinoblastoma tumor suppressor protein (Rb) plays a vital role in regulating mammalian cell cycle progression and inactivation of this protein is required for entry into the S phase. Rb is inactivated by phosphorylation upon growth factor stimulation of quiescent cells, thereby facilitating the transition from G1 to S phase.[29] Previous studies have shown that the signaling kinase c-Raf (Raf-1) physically and functionally interacts with Rb and contributes to its inactivation, thereby facilitating cell proliferation.[30] Drug peptide ligand Raf analogue which inhibits Rb-Raf-1 binding and therefore inhibits tumor growth and cell proliferation, and vascular endothelial growth factor-mediated capillary tubule formation.

In this context, we designed nanoplatforms with a tumor-homing RGD integrin ligand peptide and/or a cytotoxic peptide ligand Raf analogue, based on inorganic fluorescent NPs coated with a biocompatible organic shell.

We opted for micellar NPs composed of the tri-block surfactant copolymer Pluronic®F127 (PF127, PEG_{100}-PPO_{65}-PEG_{100}) and a dye-doped silica core. The PEG-PPO-PEG block copolymers alone can form micelles in aqueous media with a hydrophobic core, which can be used to non-covalently encapsulate hydrophobic dyes with minimal leakage.[31,32] Micelles are self-organized systems, and the encapsulation of dyes and drugs in these systems is straightforward from the preparation point of view. However, they are dynamic systems, and their stability is strongly influenced by their local concentration, pH, and the presence of other molecules or proteins.[4,32,33] To increase the stability, the hydrophobic core of the micelles was trapped into a silica frame by polymerization of TEOS/TMSCl, giving silica-core/PEG-shell NPs (Si-PF127 NPs).

These particles synergistically combined the features of micelles and silica nanoparticles to address the multiple challenges of targeted delivery. On the other hand, silica NPs as tools to develop targeting probes have several advantages over other nanomaterial and self-organized systems.[34] In particular, silica is photophysically inert, intrinsically non-toxic, and numerous synthetic approaches exist to tailor these nanosystems in terms of size and functionalization. The luminescence emission of these systems depends on the doped dye, allowing for a wide range of emission properties simply by selecting the appropriate doping dyes.

4.3 Results and Discussion

4.3.1 Design and synthesis of linkable cRGD and Raf peptide ligands

As mentioned in the introduction part, the cyclopentapeptide, c[RGDfK], developed by Kessler and co-workers, is a good targeting probe for conjugation to nanoparticles, with an IC_{50} of 2.25 nM towards integrin $\alpha_V\beta_3$ subfamily.[35,36] An alkyne linker was introduced to the side chain of Lys to allow the conjugation via CuAAC reaction. Similarly, for the cytotoxic peptide ligand Raf analogue which inhibits Rb-Raf-1 binding in vivo and therefore inhibits tumor growth and angiogenesis,[30] the N-terminus was exploited for the introduction of alkyne spacer (Figure 4.13).

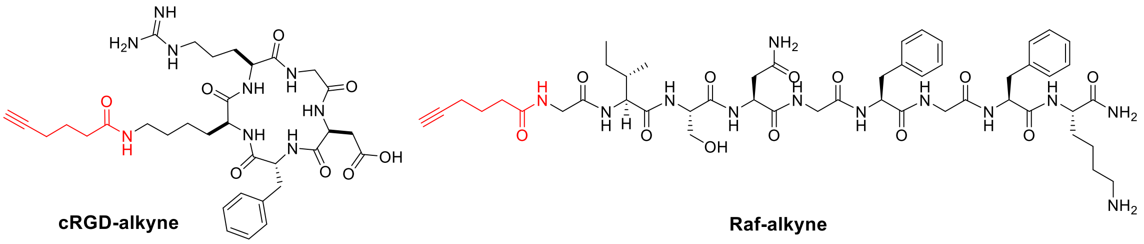


Figure 4.13 Chemical structures of peptide ligands used in this study.

The linear peptides were elongated by using a Fmoc/^{t}Bu solid-phase synthetic strategy. Cyclization was conducted under high-dilution condition[13] via DPPA/$NaHCO_3$ as catalytic system. Alkyne acid linker was introduced just prior to final global deprotection. For detailed synthesis procedure, see Experimental Section. The identities were confirmed by NMR and ESI MS analysis (Appendix B).

4.3.2 Synthesis and characterization of peptide copolymers

The dimesylate derivative of Pluronic®F127, obtained in turn by the treatment of PF127 surfactant with trimethylamine/methanesulfonyl chloride, was subsequently reacted with NaN_3, yielding azide derivitized PF127-N_3 (Figure 4.14).[37] The resulting polymer was confirmed by the changes of proton peaks before and after modifications in NMR spectra, as well as the appearance of characteristic azide peak at 2105 cm^{-1} in FT IR analysis (Figure 4.15).

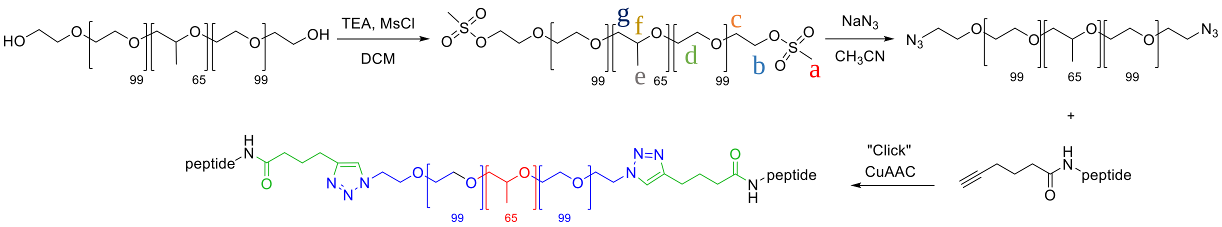


Figure 4.14 Derivatization of Pluronic®F127 copolymer and conjugation to peptides via CuAAC reaction.

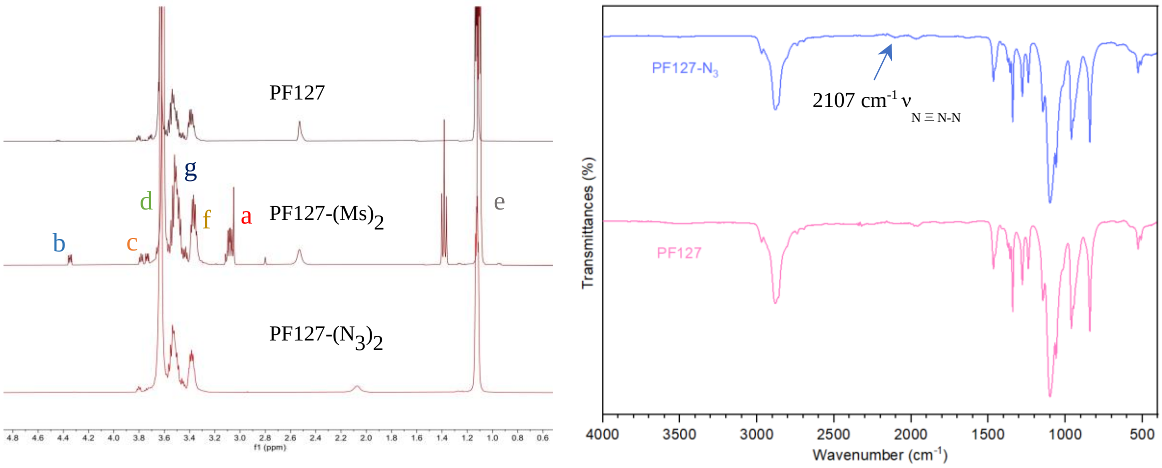


Figure 4.15 Characterization of PF127-N3 by ^{1}H NMR and FT IR.

Subsequently, the resulting azide-modified PF127 polymer was coupled with the Raf-alkyne peptide using copper-catalyzed-azide-alkyne cycloaddition (CuAAC) chemistry in the presence of bipyridine and a catalytic amount of CuI,[38] yielding the Raf peptide-conjugated PF127 polymer (Raf-PF127). The obtained Raf-PF127 conjugates were purified by dialysis against water in 5 kDa cutoff tubes and dried thoroughly before use.

The successful conjugation was confirmed by ^{1}H NMR analysis, where the appearance of a set of tiny peaks corresponding to the aromatic and amide protons (in the 7.0-8.5 ppm region) and α/β protons, as well as other side chain protons (in the aliphatic 0.5-4.5 ppm region) were observed in Raf-PF127, indicating the presence of the Raf peptide, while these peaks were completely absent in the unconjugated PF127-N3 (Appendix B: Figure S7). Moreover, the FT IR spectrum of Raf-PF127 polymer revealed the presence of characteristic peaks of the peptide bonds in the region 1600-1700 cm^{-1} and that the disappearance of the IR signal ascribable to azide at around 2100 cm^{-1}, further confirming the successful attachment of Raf peptide to PF127 polymer (Figure 4.16). A similar pattern in both ^{1}H NMR and FT-IR spectra can be observed in the cRGD-PF127 polymer. In the ^{1}H NMR spectrum, the peaks at 1.14 ppm and 3.40-3.60 ppm were attributed to -CH_3 of PPO and -CH_2CHO of PPO and PEO in PF127, respectively. Moreover, characteristic peaks attributed to the cRGD peptide, appear in the expected regions, confirming the successful conjugation of cRGD to the PF127 polymer (Appendix B: Figure S6). Additionally, in the FT-IR spectrum, peptide bond absorption bands of the cRGD peptide were present (Amide $\nu_{C=O}$ at 1640 cm^{-1}, Amide δ_{N-H} at 1535 cm^{-1}, and ν_{N-H} at 3280 cm^{-1}) (Figure 4.16), consistent with desired conjugations. All these analyses demonstrated that CuAAC chemistry is a robust strategy for the efficient conjugation of peptides to polymeric PF127.

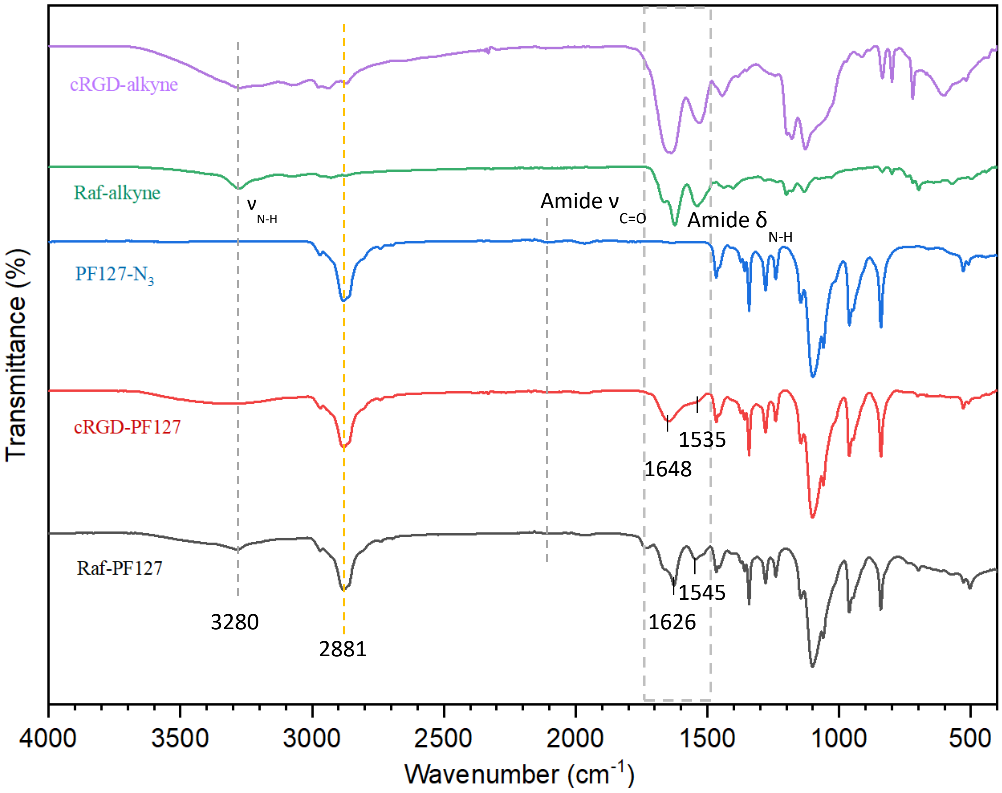


Figure 4.16 FT IR spectra of cRGD-alkyne (purple line), Raf-alkyne (green line), $PF127-N_3$ (blue line), cRGD-PF127 (red line) and Raf-PF127 (black line) copolymer.

4.3.3 Synthesis of silica-core/PEG-shell nanoparticles

Monodispersed fluorescent silica-core/PEG-shell NPs functionalized with peptide moieties and incorporating the dye rhodamine B triethoxysilane (RhB-TES)[39] were expediently obtained using a direct micelle-assisted method.[40] These nanostructures were formed by the condensation of the silica precursor TEOS in an aqueous acid environment in the presence of co-aggregates composed by a mixture of the native Pluronic polymer PF127 and peptide conjugated PF127 (Figure 4.17).[24] These reactions yielded bare Si-PF127 NPs, cRGD-Si-PF127 NPs, Raf-Si-PF127 NPs, and cRGD/Raf-Si-PF127 NPs, respectively. The resulting NPs were purified by dialysis against water at RT using cellulose dialysis tubes (cutoff >12 kDa). The condensation of RhB-TES within the silica core of the NP conferred the desired fluorescent properties to this nanosystem, preventing also the leaking of the fluorophore in the external environment.

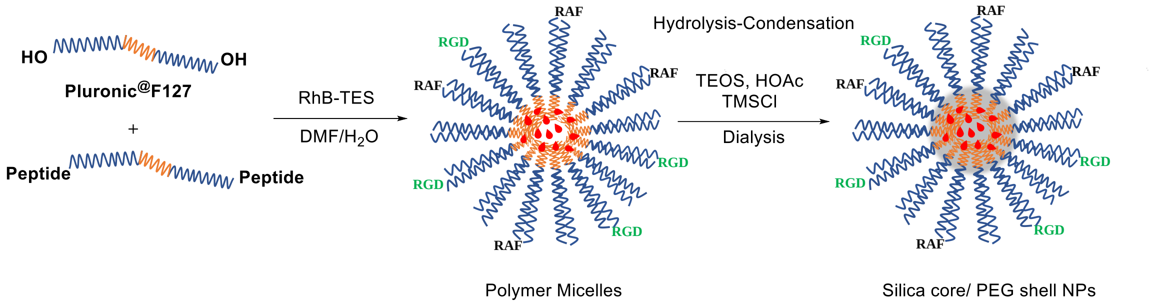


Figure 4.17 Preparation of fluorescent peptides functionalized silica-core/PEG-shell NPs.

4.3.4 Characterization of silica-core/PEG-shell nanoparticles

The hydrodynamic size was analyzed by DLS. Consistent with the reported results,[24,37] the bare NPs showed a d_H of 19.72 ± 1.478 nm, with a PDI value of 0.2 (Figure 4.18).

The effectiveness of the conjugation reaction was also indicated by DLS, since the volume distribution after peptide conjugation increased to 22 nm for cRGD-Si-PF127 NPs, 20 nm for Raf-Si-PF127 NPs and 33 nm for cRGD/Raf-Si-PF127 NPs, compared to that of the bare NPs (Figure 4.18).

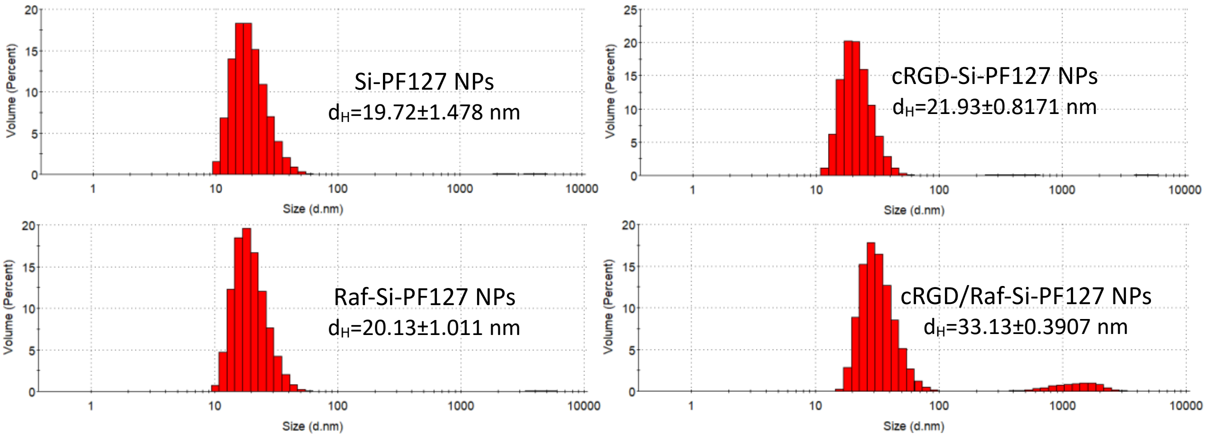


Figure 4.18 DLS hydrodynamic size distribution by volume of silica-core/PEG-shell nanoparticles.

4.3.5 Cell viability assay

The in vitro cell growth inhibitory efficacy was determined for the synthesized NPs by incubating A549, MCF7 and HDFa cells with increasing concentrations of the NPs (0.1, 1.0, 5.0 μM) for 48 h (Figure 4.19).

In general, all samples showed no toxicity at the concentration of 0.1 μM, therefore, these data were not included. On the other hand, all the combinations tested showed negligible toxicity towards healthy HDFa cells at all concentrations. At the concentration of 1 μM, Raf-NPs induced a decrease of viability of about only 5% in MCF7, while the effect was much higher in A549, whose vitality was decreased by 61%. At the concentration of 5 μM, Raf-NPs showed even higher toxicity against A549, reducing their viability by 78%, while the effect was lower in MCF7 cells, whose vitality was decreased by 30%. Similarly, 1μM Raf/cRGD-NPs significantly inhibited the proliferation of $\alpha_v\beta_3$ integrin-overexpressing human lung cancer A549 about 50% while they were almost ineffective towards $\alpha_v\beta_3$ integrin-lowexpressing human lung cancer MCF7. Increasing the concentration of Raf/cRGD-NPs to 5 µM led in general to higher toxicity, which inducing a decrease of viability of about 68% in A549, and of 24% in MCF7 cells (Figure 4.19).

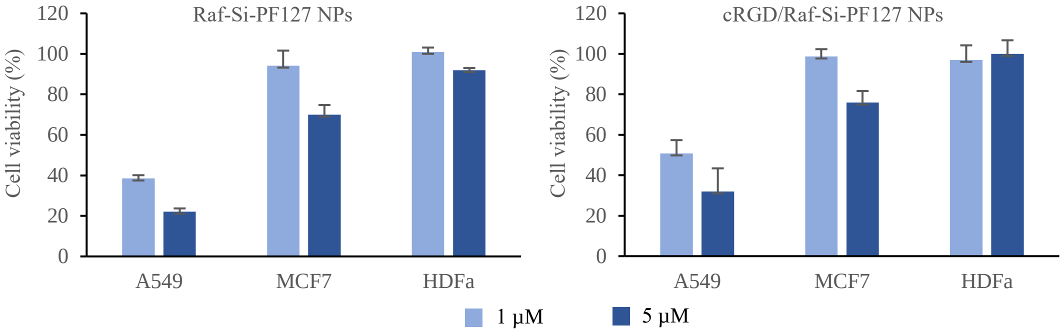


Figure 4.19 Cell viability of peptide-NPs as % of control cells.

It is worth noting that the cytotoxicity of the multifunctional Raf/cRGD-Si-PF127 nanoparticles was slightly inferior to that of Raf-Si-PF127 nanoparticles. This is plausibly due to the reduced amount of cytotoxic Raf peptides in the blend nanosystems, which may not fully compensate for the targeting effect of the cRGD peptide. Indeed, when preparing the Raf/cRGD-Si-PF127 NPs, the amount of Raf-PF127 used in the formulation was only half that of Raf-Si-PF127 NPs. While cRGD enhances the nanoparticle's targeting ability by improving specificity for integrins on tumor cells, the lower concentration of Raf peptide could lead to diminished overall therapeutic efficacy, suggesting a balance between targeting and therapeutic payload is crucial for optimizing the performance of multifunctional nanoparticles.

In summary, we described synthetic nanoplatforms multifunctionalized with a tumor-homing RGD integrin ligand and/or a cytotoxic Raf peptide ligand, constituted by a fluorescent silica core doped with RhB derivative, coated with a PEG shell. Overall, the NPs showed superior toxicity towards cancer cells in the preliminary cell viability assays.

4.4 Conclusion

Currently, there are growing preclinical and clinical studies focusing on integrin-related nanoparticles for diagnostic and therapeutic applications in various diseases. Indeed, numerous integrin targeted nanosystems can be designed with great versatility by optimizing two key elements in this strategy: nanoparticles and integrin ligands. Nanoparticles, such as gold, PLGA, lipid-based, and silica, can be tailored for specific applications like drug delivery, gene therapy, or imaging. Surface modifications like PEGylation can further enhance drug-like properties. On the other hand, integrin ligands can be carefully selected to target specific integrin subfamilies, such as cancer cell overexpressing integrin $\alpha_v\beta_3$ or leukocyte-specific integrin $\alpha_4\beta_1$, ensuring precise targeting of diseased cells.

In this study, several nanoparticles carrier integrin ligands were delicately designed and fabricated, for potential applications in diagnostics and theranostics.

Firstly, building on the results achieved in Chapter 3, several potent LDV-containing integrin ligands decorated metallic gold nanoparticles were prepared. The functionalization was well-characterized by TEM, DLS, and UV-vis analysis. Given the unique optical and electronic properties of AuNPs, they hold great potential for targeted photothermal therapy. Subsequently, biocompatible polymeric PLGA nanoparticles coated with LDV peptides were fabricated. Furthermore, curcumin, an anticancer drug was loaded into the NPs with a drug loading capacity of 10%. Since curcumin is intrinsically fluorescent, this property can be advantageous for visualizing the NPs in cells. Lastly, we also designed nanoplatforms with a tumor-homing RGD-based ligand and/or a cytotoxic peptide Raf analogue, based on inorganic fluorescent silica NPs coated with a biocompatible organic PEG shell. As a result, the nanoparticle conjugates showed moderate to superior toxicity towards A549 and MCF7 cancer cells in the preliminary cell viability assays. These findings highlight the potential of integrin-targeted nanotherapeutics for cancer treatment. To further elucidate the roles that different nanoparticles and integrin ligands could play, in-depth biological evaluations are necessary and currently underway in our laboratory.

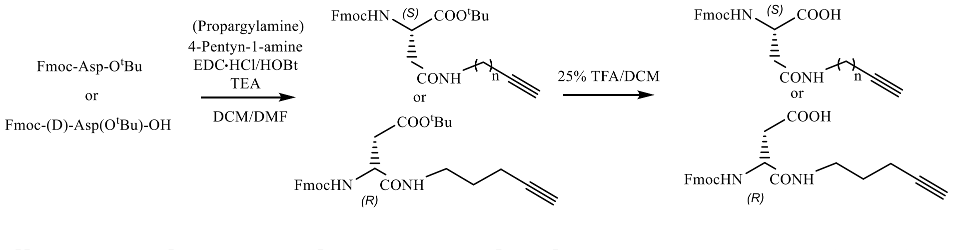
4.5 Experimental Section

Unless otherwise stated, standard chemicals and solvents were purchased from commercial sources and used as received without further purification. Peptide purity was assessed by analytical RP HPLC performed on an 1200 series apparatus Agilent Technologies, Milan, Italy, using an XSelect Peptide CSH C18 column (Waters, Milford, MA, USA), 4.6 mm × 100 mm, 130 Å, 3.5 µm. MS (ESI) analysis was performed using an MS single quadrupole HP 1100 MSD detector (Agilent Technologies, Milan, Italy). Peptides isolation was performed by preparative RP HPLC performed on a Waters 2489 UV/visible Dual Detector equipped with a Waters 1525 Binary HPLC pump, using reverse-phase column XSelect Peptide CSH C18 OBD column (Waters, 19×150 mm, 5 µm, no.186007021). Column description: stationary phase octadecyl-carbon chain-bonded silica (C18), double end-capped, particle size 5 µm, pore size 130 Å, length 150 mm, internal diameter 19 mm; DAD 210 nm, DAD 254 nm. Mobile phase for neutral compounds: isocratic 6:4 H_2O/CH_3CN for 2 min, then gradient from 6:4 H_2O/CH_3CN to 2:8 H_2O/CH_3CN in 10 min, then isocratic 2:8 H_2O/CH_3CN for 2 min; flow rate: 10 mL min^{-1}. Mobile phase for ionizable peptides: isocratic 6:4 H_2O/CH_3CN and 0.1 % HCOOH for 2 min, then gradient from 6:4 H_2O/CH_3CN and 0.1 % HCOOH to 2:8 H_2O/CH_3CN and 0.1 % HCOOH in 10 min, then isocratic 2:8 H_2O/CH_3CN and 0.1 % HCOOH for 2 min. NMR spectra were recorded with a Bruker BioSpin GmbH (^{1}H: 600 MHz, ^{13}C: 151 MHz) at

298 K in 5 mm tubes, using 0.01 M peptide. Chemical shifts are reported in ppm (δ) and internally referenced with $(CD_3)_2SO$: ^1H: 2.50, ^{13}C: 39.52 ppm. Fluorescence measurements were performed with a fluorimeter Perkin Elmer LS50B. DLS measurements were performed with a Zetasizer Nano ZS (Malvern Panalytical, Malvern, UK). The UV-vis absorption spectrum was measured with a microplate spectrophotometer (Multiskan GO, Thermo Scientific, USA). TEM imaging and size analysis were performed using a FEI Tecnai T20 instrument at an accelerating voltage between 80-200 kV (Laboratorio de Microscopias Avanzadas LMA, University of Zaragoza). Samples were prepared by drop casting the AuNP suspension on carbon-coated copper grids (Electron Microscopy Sciences, 200 mesh) and dried under ambient conditions. Particle size distribution was determined with ImageJ software. FT IR analysis was performed on an Alpha FT IR spectrophotometer (Bruker, Billerica, MA, USA).

Synthesis of cyclic LDV peptides 1, 2, 3, and linear peptide 4. cLDV linkable peptides **1**, **2**, and **3** were synthesized following a similar procedure as described in Chapter 3, with minor modifications. In brief, the corresponding linear precursors were assembled by SPPS on a 2-chlorotrityl chloride resin using Fmoc-protected amino acids, and DCC/HOBt as activating agents. For amine peptide **3** (c[*(S)*-Phu-LKV-*(S)*-isoAsp]), (S)-Asp was introduced as Fmoc-Asp-O^{t}Bu; Lys was introduced as Fmoc-Lys(Boc)OH. For peptide-alkyne **1** and **2**, the (*S*)-Asp^5 was introduced as Fmoc-Asp-O^{t}Bu, while (*S*)-Asp^3 was pre-modified to corresponding Fmoc protected alkyne-bearing amino acids before being incorporated into the sequence. Lastly, for linear peptide **4**, the synthesis was performed according to literature, with minor modifications. (*R*)-Asp was introduced as Fmoc protected alkyne-bearing amino acids, similar to **1** and **2**, but was modified from a D-configured β-amino acid.

Synthesis of Fmoc protected alkyne-bearing Asp. A mixture of Fmoc-Asp-O^{t}Bu or Fmoc-(D)-Asp(O^{t}Bu)-OH (0.25 mmol), HOBt (0.4 mmol), EDC.HCl (0.4 mmol), and TEA (0.8 mmol), in 3:1 DCM/DMF (4 mL), was stirred at RT under N_2 atmosphere. After 5 min, the amine-alkyne (0.3 mmol) was added, and the mixture was stirred under N_2 at RT overnight. Then, the solvent was evaporated at reduced pressure, and the residue was diluted with EtOAc (10 mL). The suspension was washed with 0.1 M HCl (5 mL), and a saturated solution of $NaHCO_3$ (5 mL). The organic layer was dried over Na_2SO_4, filtered, and the solvent was evaporated. Purification by flash chromatography over silica gel (hexane/EtOAc 3:1) allowed the isolation of products as white solids.



The fully protected amino acids were treated with 25% TFA in DCM (3 mL) at 0 °C, and the mixture was shaken at RT for 1 h. Then, the solvent was evaporated, and ice-cold Et_2O was added to precipitate the final products as TFA salt in quantitative yield. The products were collected by centrifuge, and used for the next step without further purifications.

General procedure for loading the first amino acid to 2-chlorotrityl chloride (2-CTC) resin. The 2-Chlorotrityl chloride resins ((100-300 mesh, 1.20 mmol/g) were added to a syringe peptide synthesis vessel. The resin was swelled in DCM (15 min). A solution of Fmoc-AA-OH (1 eq.) and DIPEA (1 eq.) in DCM (approximately 10 mL per gram of resin) was added and the resin was shaken for 5 minutes. Then, an extra 1.5 eq. of DIPEA was added and shaking was continued for 1 h. MeOH (0.8 mL/ g of resin) was added to the mixture in order to cap any unreacted functional groups on the resin, and shaken for 15 min. After filtering, the resin was washed with DCM (x3), MeOH (x3), and DMF (x3).

General procedure for Fmoc Deprotection. Fmoc cleavage was carried out using 20% (*v/v*) piperidine in DMF (5 mL), while gently shaking at RT for 10 min. After washing with DMF (5 mL), the deprotection was repeated. The resin was then washed sequentially with DMF, MeOH, and DCM, (3×4 mL each).

General procedure for coupling. The subsequent coupling reactions were performed by dissolving in a separate vial Fmoc-protected amino acids (0.3 mmol) and HOBt (0.3 mmol) in DMF (4 mL) for 20 min. The last introduced residue was Fmoc-Phu-OH. The Fmoc-AA-OH /HOBt mixture was poured into the reactor followed by DCC (0.3 mmol), and the suspension was shaken for 3 h at RT. Coupling efficacy was monitored by Kaiser test.

General procedure for cleavage. Cleavage from the resin was performed by using a mixture of AcOH/TFE/DCM, 1:1:4, *v/v/v* (10 mL) while shaking for 2 h at RT. The mixture was filtered and the resin was washed twice with Et_2O/DCM. The filtrates were collected and solvents were removed under reduced pressure, and ice-cold Et_2O was added to precipitate the crude peptides as TFA salt, which were recovered by centrifuge and used for the cyclization without further purification.

General procedure for cyclization. The cyclization of the crude peptide was performed under pseudo-high dilution condition. A solution of the linear peptides (0.15 mmol) in DMF (10 mL) was added over 16 h using a syringe pump, to a mixture of of HBTU (0.45 mmol), HOBt (0.45 mmol) and DIPEA (0.9 mmol) in DMF at RT. Once the addition was complete, the reaction was stirred for additional 2 h. Then, the solvent was distilled at reduced pressure, and the crude peptides were isolated by RP HPLC on a semipreparative C18 column (General methods).

Removal of acid labile *tert*-butyl and Boc protecting groups was performed by treatment of a 25% TFA in DCM (x2) for 30 min at RT to afford the final products peptides **1-3** in quantitative yield. The purity and the identity of the products were determined to be >95% by RP HPLC coupled to ESI MS, and by NMR in 8:2 DMSO-d_6/H_2O.

Synthesis of 14 nm AuNPs@cit. Citrate stabilized 14 nm AuNPs were synthesized via citrate reduction method. Briefly, prior to use, all glassware was washed with aqua regia and rinsed thoroughly with Milli-Q water. 80.4 mg of $HAuCl_4$ was dissolved in 185 mL MilliQ water in a two-neck round bottom flask. The reaction mixture was stirred, heated to 100 °C on the stirrer at 370 rpm. Once vigorously boiling, 241.8 mg of sodium citrate dihydrate in 5 mL water was quickly added. The reaction mixture was allowed to reflux for another 30 minutes, and the color changed from yellow to clear to black and finally to wine red. The heat and water for the condenser were turned off, and the mixture was allowed to cool to RT. The mixture of synthesized AuNPs was filtered with a 0.2 μm filter to get rid of the larger aggregates of particles. AuNPs stock solution were stored in a glass bottle covered with aluminum foil at 4 °C for further use. The molar concentration of AuNPs was obtained from UV-vis spectroscopy using Beer-Lambert's law ($A = \varepsilon cf$), where A is the absorbance at 450 nm, measured using UV-vis spectrometry by a microDrop plate (Multiskan GO, Thermo Scientific, USA), ε is the molar extinction coefficient ($1.76 \times 10^8\ M^{-1}cm^{-1}$), c is the molar concentration of AuNPs, and f is the path length (0.052). c_{AuNPs}=12.96 nM.

Synthesis of AuNPs@PEG-COOH. The heterobifunctional HS-PEG-COOH (Mercapto-ω-carboxy PEG MW. 5000 Dalton) was grafted to the AuNPs surface by the thiols (SH-) functional groups via ligand exchange.

To 154 mL of resulting citrate-stabilized AuNPs stock solution (12.96 nM), 560 μL of 10% SDS solution, 28 mL of aqueous HS-PEG-COOH (2 mg/mL), and 2.5 mL of 2 M aqueous NaOH were added, bringing the total volume to 200 mL with water. The reaction mixture was shaken for two days to allow complete exchange of citrate ions with HS-PEG-COOH. Afterward, the mixture was centrifuged at 14,000 rpm for 30 minutes at 10 °C. The supernatant was discarded, and the precipitate was resuspended in the same volume of Milli-Q water, followed by two additional washing steps under the same conditions. After the final wash, the concentrated AuNPs@PEG-COOH stock solution was stored in a glass vial, protected with aluminum foil, and kept at 4 °C for future use. c_{AuNPs}= 88.12 nM.

Conjugation of amine LDV peptide 3 to AuNPs@PEG-COOH via carbodiimide chemistry.[41,42] The synthesis of AuNPs functionalized with peptide **3** was based on an EDC/sulfo-NHS coupling reaction between the amine-containing peptide **3** and carboxyl-functionalized AuNPs. In brief, the EDC/sulfo-NHS solution were prepared by dissolving 17 μg of EDC.HCl and 43 μg of sulfo-NHS in 500 μL of 10 mM MES buffer (pH 6) and allowed to premix for 5 minutes. Following this, 500 μL of a 0.5 mg/mL AuNPs@PEG-COOH solution in 10 mM MES buffer (pH 6) was incubated with the prepared EDC/sulfo-NHS solution for 30 min at 37 °C to activate the carboxyl groups on the AuNPs. After activation, the AuNPs were centrifuged at 14,000 rpm for 15 minutes at 4 °C and washed with fresh MES buffer (10 mM, pH 6) to remove any residual EDC/sulfo-NHS. The activated AuNPs were then redispersed in 500 μL of peptide **3** solution (17.64 mM in 10 mM MES buffer, pH 6) and incubated for 3 hours at 37 °C to allow the coupling reaction. After this process, unattached peptides were also eliminated by washing the NPs in MES buffer (15 min, 14,000 rpm, 4 °C), the supernatant containing the unreacted peptides was collected and analyzed by fluorophotometry for quantification. The obtained Pep **3**-AuNPs were resuspended in MilliQ water to a final volume of 500 μL and kept at 4 °C for further use.

Synthesis of AuNPs@PEG-CONH$(CH_2)_3$-N_3. The synthesis procedure followed the same steps as for the conjugation of peptide **3** to AuNPs@PEG-COOH. After EDC/sulfo-NHS activation, instead of the peptide **3** solution, H_2N-$(CH_2)_3$-N_3 in MES buffer was added. This facilitated the coupling of the azide group to the carboxyl-functionalized AuNPs. The resulting azide-functionalized AuNPs (azide-AuNPs) were then resuspended in Milli-Q water to a final volume of 500 μL and stored at 4 °C for future bioconjugation via CuAAC.

Conjugation of alkyne peptides 1, 2, and 4 to azide-AuNPs via CuAAC chemistry.[43,44] Stock solutions: $CuSO_4.5H_2O$: 20 mM (in water), Ligand THPTA: 50 mM (in water), Sodium ascorbate: 100 mM (make fresh before using by adding 1 mL of water to 20 mg), alkyne-peptide: 20 mg/mL, Buffer: 100 mM phosphate pH 7.

In a 2 mL eppendorf tube containing the azide-AuNPs, add the reagents in the following order: a. 442.5 uL of 100 mM phosphate buffer pH 7; b. 25 uL of peptide solution (Pep **1**, Pep **2**, or Pep **4**); c. 7.5 uL of premixed $CuSO_4$ and THPTA (2.5 μl of $CuSO_4$ + 5.0 μL of THPTA); d. 25 ul of sodium ascorbate; e. Mix well, close the tube (to prevent more oxygen from diffusing in), and let the reaction go for 3 hours at 37 °C; f. After, centrifuge at 14,500 rpm, 15 min and 4 °C to remove the supernatants, and then two more washes with 10 mM EDTA solution to remove any copper ions; g. Lastly, wash with Milli-Q water at 14,500 rpm, 15 min and 4 °C and resuspend the AuNPs in Milli-Q water and keep them in the fridge for further use.

Fluoroscamine assay. Stock solutions: 1 mg/mL fluorescamine (in anhydrous ACN), Buffer: 0.1 M borate buffer pH 8.

For fluorometric quantitative analysis of peptides on AuNPs surface, a peptide **3** calibration curve was established. Briefly, to a 2 mL Eppendorf tube, 20 μL of a known concentration of peptide **3** solution was added, followed by 190 μL of borate buffer (pH 8). Before measurements, 90 μL of fluorescamine solution was added to the eppendorf, and the mixture was thoroughly mixed. Then, 150 μL of the resulting reaction mixture was transferred to the cuvette, and the fluorescence emission was recorded rapidly (λ_{Ex}= 390 nm, λ_{Em}= 480 nm, Ex slit=7.5, Em slit=5, under emission mode). For sample measurement, 20 μL of supernatant containing unreacted peptide **3** after conjugation was used.

Determination of FAM-alkyne functionalization. For fluorometric quantitative analysis of FAM on AuNPs surface, a standard FAM-alkyne calibration curve was established. In brief, the fluorescence intensity of known concentrations of FAM-alkyne was recorded (λ_{Ex}= 492 nm, λ_{Em}= 517 nm, Ex slit=5.0, Em slit=5.0, under emission mode), and the maximum absorption at 517 nm was plotted against FAM-alkyne concentrations.

FAM-AuNPs digestion: 135 μL of FAM-AuNPs was mixed with 15 μL of 0.3 M KCN/0.1 M KOH, resulting mixture was incubated at 80 °C for 30 min to allow complete disassociation of gold nanoparticles. After digestion, the solution was subjected to fluorometric analysis, and the maximum absorption at 517 nm was recorded and analyzed.

Synthesis of PLGA-(Pep 3)$_2$ polymer. PLGA (lactide: glycolide = 75:25, Mn = 66,000-107,000, 10 mg) was dissolved in 1 mL of DMSO, EDC·HCl (10 eq.) and sulfo-NHS (10 eq.) were dissolved in 500 μL of DMSO. After the mixture was stirred for 2 h at RT for activation COOH of PLGA, LDV peptide **3** (2.5 eq.) was added to the solution. Then, the solution was stirred overnight at RT. To remove EDC, sulfo-NHS and any non-cross-linked peptide, the reaction mixture was washed with cold EtOH, and the PLGA-(Pep **3**)$_2$ polymer was precipitated, collected by centrifuge. Finally, the functionalized polymer was obtained as a white powder and stored under 4 °C for further use.

Preparation of curcumin loaded PLGA NPs. The nanoparticles were obtained by the single-emulsion solvent-evaporation technique, as previously described. Briefly, curcumin (1 mg) and PLGA (10 mg) were dissolved in a mixture of ethyl acetate (0.3 mL) and methylene chloride (0.1 mL) at RT. This organic phase was rapidly poured into 2 mL of PVA aqueous solution (0.5%, w/v) and emulsified by sonication for 5 min (40% amplitude), resulting in an oil-in-water (O/W) emulsion. Next, the organic solvent was slowly evaporated by stirring under vacuum hood. The particles were then recovered by centrifugation (20,000×g, 30 min, 10 °C) and washed twice with water to remove the surfactant. The nanoparticles were dispersed in 3 mL Milli-Q water, and the resulting nanosuspension was stored under 4 °C for further use.

Synthesis of cyclic RGD peptide and Raf peptide. cRGD peptide was synthesized according to literature.[37] Raf peptide was assembled by standard SPPS on Rink amide resin, Fmoc deprotection and subsequent coupling was done by general procedures, the last introduced residue was 5-hexynoic acid, finally, cleavage from resin and simultaneous removal of acid labile protecting groups, e.g. Boc of Lys, tert-butyl of Ser, was performed by using a 95:2.5:2.5, *v/v/v* mixture of TFA/TIPS/H_2O (10 mL) while shaking for 2.5 h at RT. The mixture was filtered and the resin was washed twice with Et_2O/DCM containing a small portion of TFA. The filtrates were collected and solvents were removed under reduced pressure, and ice-cold Et_2O was added to precipitate the crude peptides as TFA salt, which were recovered by centrifuge and further purified by preparative HPLC.

Synthesis of PF 127-(N_3)$_2$. Pluronic® F127 (6.3 g, 0.5 mmol, 1 eq.) was dried by azeotropic distillation with toluene at reduced pressure. Then, the polymer was solubilized in anhydrous DCM (25 mL) and treated with triethylamine (TEA, 140 μL, 1 mmol, 2 eq.) and methanesulfonyl chloride (MsCl, 77 μL, 1 mmol, 2 eq.) under inert atmosphere. The mixture was stirred at 0 °C for 3 h and then at RT overnight. The solvent was evaporated at reduced pressure, the resulting residue was washed with ice-cold Et_2O, and the solid was collected by precipitation (>95%).

Sodium azide (122 mg, 1.88 mmol, 4 eq.) was added to a suspension of dimesylate-PF127 (6.026 g, 0.47 mmol, 1.0 eq.) in CH_3CN (25 mL), and the mixture was stirred under reflux for

48 h. Subsequently, the solvent was removed under reduced pressure. The solid so obtained was dispersed in a 5% $NaHCO_3$ solution saturated with NaCl (20 mL). The mixture was extracted four times with DCM (4×15 mL). The combined organic phases were dried over Na_2SO_4, filtered, and evaporated at reduced pressure, affording a white solid (95%).

Conjugation of peptides to PF 127-(N_3)$_2$ via CuAAC.[38] PF127-(N_3)$_2$ (0.015 mmol), the oligopeptides (0.06 mmol), and copper iodide (0.07 mmol) were added together into a dry flask. The glassware was capped with a septum and purged with dry argon for several minutes. Then, bipyridine (0.14 mmol) dissolved in 2 mL of degassed N-methyl pyrrolidone (NMP) was added via a degassed syringe. The mixture turned dark brown, indicating complexation of CuI and Bipy. The reaction mixture was stirred for 24 h at RT. The final mixture was diluted in deionized water and subsequently purified by dialysis against pure water (cutoff: 3500). Last, water was removed by rotary evaporation.

Preparation of dye-doped silica core/PEG shell nanoparticles. PF127 (77 mg), PF127-peptides (33 mg), and RhB-TES (1.5 mg) were solubilized in DCM (1 mL) in a 10 mL glass scintillation vial. The organic solvent was evaporated from the homogeneous solution, and the solid residue was subsequently dried under vacuum at rt. The resulting solid mixture was solubilized at 25 °C while stirring in a 1 M AcOH/0.85 M NaCl solution (1.6 mL) for 1 h. Subsequently, TEOS (175 µL) was added after 3 h of stirring, then TMSCl (20 µL) was added. The mixture was kept under stirring for 48 h at 25 °C. The NPs were purified by dialysis against water at RT using cellulose dialysis tubing (Sigma, MW cutoff >12 kDa, avg. diameter 33 mm), and the solution was finally diluted to a total volume of 1.75 mL with water to the final NPs concentration 29 µM.[45]

Cell viability assay. The cytotoxicity of peptide-NPs was evaluated using the cell viability 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) assay, according to the manufacturer’s instructions. Cells were seeded (1.5×10^4 cells/well) and cultured for 48 h. The primary growth medium was replaced by fresh medium, containing NPs at the concentrations of 0.1, 1.0, 5.0 µM. After 48 h, PBS (100 µL) was supplemented with the MTS solution (20 µL/well), incubated for 2 h, and then the absorbance was recorded at 570 nm with a 96-well plate reader. Data were analyzed by Prism software (GraphPad) and expressed as % of controls (untreated cells).

References

[1] Altammar, K.A. A review on nanoparticles: characteristics, synthesis, applications, and challenges. *Front. Microbiol.* **2023**, *14*, 1155622.

[2] Wu, J., The Enhanced Permeability and Retention (EPR) Effect: The Significance of the Concept and Methods to Enhance Its Application. *J Pers Med* **2021**, *11* (8), 771.

[3] Auría-Soro, C.; Nesma, T.; Juanes-Velasco, P.; Landeira-Viñuela, A.; Fidalgo-Gomez, H.; Acebes-Fernandez,

V.; Gongora, R.; Almendral Parra, M.J.; Manzano-Roman, R.; Fuentes, M. Interactions of Nanoparticles and Biosystems: Microenvironment of Nanoparticles and Biomolecules in Nanomedicine. *Nanomaterials (Basel)* **2019**, *9* (10), 1365.

[4] Akhter, M.H.; Khalilullah, H.; Gupta, M.; Alfaleh, M.A.; Alhakamy, N.A.; Riadi, Y.; Md, S. Impact of Protein Corona on the Biological Identity of Nanomedicine: Understanding the Fate of Nanomaterials in the Biological Milieu. *Biomedicines* **2021**, *9* (10), 1496.

[5] Tian, H.; Zhang, T.; Qin, S.; Huang, Z.; Zhou, L.; Shi, J.; Nice, E.C.; Xie, N.; Huang, C.; Shen, Z. Enhancing the therapeutic efficacy of nanoparticles for cancer treatment using versatile targeted strategies. *J. Hematol. Oncol.* **2022**, *15* (1), 132.

[6] Hamidi, H.; Ivaska, J. Every step of the way: integrins in cancer progression and metastasis. *Nat. Rev. Cancer* **2018**, *18* (9), 533-548.

[7] Singh, R.; Lillard, J.W., Jr. Nanoparticle-based targeted drug delivery. *Exp. Mol. Pathol.* **2009**, *86* (3), 215-223.

[8] Mitchell, M.J.; Billingsley, M.M.; Haley, R.M.; Wechsler, M.E.; Peppas, N.A.; Langer, R. Engineering precision nanoparticles for drug delivery. *Nat. Rev. Drug Discov.* **2021**, *20* (2), 101-124.

[9] Strozier, R.E. Integrin Agonists or Activating Compounds and Methods for Making and Using Same. US20220118086A1, 10 September **2021**.

[10] Sharifi, M.; Attar, F.; Saboury, A.A.; Akhtari, K.; Hooshmand, N.; Hasan, A.; El-Sayed, M.A.; Falahati, M. Plasmonic gold nanoparticles: Optical manipulation, imaging, drug delivery and therapy. *J. Control. Release* **2019,** *311-312*, 170-189.

[11] Sharmah, B.; Borthakur, A.; Manna, P., PLGA-Based Drug Delivery Systems: A Promising Carrier for Antidiabetic Drug Delivery. *Advanced Therapeutics* **2024**, *7* (6), 2300424.

[12] Baiula, M.; Anselmi, M.; Musiani, F.; Ghidini, A.; Carbone, J.; Caligiana, A.; Maurizio, A.; Spampinato, S.; Gentilucci, L. Design, pharmacological characterization, and molecular docking of minimalist peptidomimetic antagonists of $\alpha_4\beta_1$ integrin. *Int. J. M. Sci*. **2023**, *24*, 9588.

[13] Malesevic, M.; Strijowski, U.; Bächle, D.; Sewald, N. An improved method for the solution cyclization of peptides under pseudo-high dilution conditions. *J. Biotechnol.* **2004**, *112*, 73-77.

[14] Ji, X.; Song, X.; Li, J.; Bai, Y.; Yang, W.; Peng, X. Size Control of Gold Nanocrystals in Citrate Reduction: The Third Role of Citrate. *J. Am. Chem. Soc.* **2007**, *129* (45), 13939-13948.

[15] Lin, A.Y.; Lunsford, J.; Bear, A.S.; Young, J.K.; Eckels, P.; Luo, L.; Foster, A.E.; Drezek, R.A., High-density sub-100-nm peptide-gold nanoparticle complexes improve vaccine presentation by dendritic cells in vitro. *Nanoscale Res. Lett.* **2013**, *8* (1), 72.

[16] Stewart, M.; Mulenos, M.R.; Steele, L.R.; Sayes, C.M. Differences among Unique Nanoparticle Protein Corona Constructs: A Case Study Using Data Analytics and Multi-Variant Visualization to Describe Physicochemical Characteristics. *Applied Sciences* **2018**, *8* (12), 2669.

[17] Egorova, E.A.; van Rijt, M.M.J.; Sommerdijk, N.; Gooris, G.S.; Bouwstra, J.A.; Boyle, A.L.; Kros, A. One Peptide for Them All: Gold Nanoparticles of Different Sizes Are Stabilized by a Common Peptide Amphiphile. *ACS Nano* **2020**, *14* (5), 5874-5886.

[18] Poonaki, E.; Nickel, A.-C.; Shafiee Ardestani, M.; Rademacher, L.; Kaul, M.; Apartsin, E.; Meuth, S.G.; Gorji, A.; Janiak, C.; Kahlert, U.D. CD133-Functionalized Gold Nanoparticles as a Carrier Platform for Telaglenastat (CB-839) against Tumor Stem Cells *Int. J. Mol. Sci.* **2022**, *23*(10), 5479.

[19] Thobhani, S.; Attree, S.; Boyd, R.; Kumarswami, N.; Noble, J.; Szymanski, M.; Porter, R.A., Bioconjugation and characterisation of gold colloid-labelled proteins. *J. Immunol. Methods* **2010**, *356*, 60-69.

[20] Udenfriend, S.; Stein, S.; Böhlen, P.; Dairman, W.; Leimgruber, W.; Weigele, M. Fluorescamine: a reagent for assay of amino acids, peptides, proteins, and primary amines in the picomole range. *Science* **1972**, *178* (4063), 871-872.

[21] Bari, E.; Serra, M.; Paolillo, M.; Bernardi, E.; Tengattini, S.; Piccinini, F.; Lanni, C.; Sorlini, M.; Bisbano, G.; Calleri, E.; Torre, M.L.; Perteghella, S. Silk Fibroin Nanoparticle Functionalization with Arg-Gly-Asp Cyclopentapeptide Promotes Active Targeting for Tumor Site-Specific Delivery. *Cancers* **2021***, 13,* 1185.

[22] Ghosh, D.; Chattopadhyay, N. Gold and silver nanoparticles based superquenching of fluorescence: A review. *J. Lumin.* **2015**, *160*, 223-232.

[23] Uehara, N., Polymer-functionalized Gold Nanoparticles as Versatile Sensing Materials. *Anal. Sci.* **2010**, *26* (12), 1219-1228.

[24] Santino, F.; Stavole, P.; He, T.; Pieraccini, S.; Paolillo, M.; Prodi, L.; Rampazzo, E.; Gentilucci, L. Preparation of Non-Toxic Fluorescent Peptide-Coated Silica/PEG Nanoparticles from Peptide-Block Copolymer Conjugates. *Micro* **2022***, 2,* 240-256.

[25] Khalil, N.M.; do Nascimento, T.C.F.; Casa, D.M.; Dalmolin, L.F.; de Mattos, A.C.; Hoss, I.; Romano, M.A.; Mainardes, R.M. Pharmacokinetics of curcumin-loaded PLGA and PLGA-PEG blend nanoparticles after oral administration in rats. *Colloids Surf. B: Biointerfaces* **2013**, *101*, 353-360.

[26] Senturk, F.; Cakmak, S., Fabrication of curcumin-loaded magnetic PEGylated-PLGA nanocarriers tagged with GRGDS peptide for improving anticancer activity. *Methods X* **2023**, *10*, 102229.

[27] Stupp, R.; Hegi, M.E.; Gorlia, T.; Erridge, S.C.; Perry, J.; Hong, Y.-K; Weller, M. *et al.* Cilengitide combined with standard treatment for patients with newly diagnosed glioblastoma with methylated MGMT promoter (CENTRIC EORTC 26071-22072 study): a multicentre, randomised, open-label, phase 3 trial. *The Lancet Oncology* **2014**, *15* (10), 1100-1108.

[28] Danhier, F.; Le Breton, A.; Préat, V. RGD-Based Strategies To Target Alpha (v) Beta (3) Integrin in Cancer Therapy and Diagnosis. *Mol. Pharm.* **2012**, *9* (11), 2961-2973.

[29] Wang, S.; Ghosh, R.N.; Chellappan, S.P. Raf-1 physically interacts with Rb and regulates its function: a link between mitogenic signaling and cell cycle regulation. *Mol. Cel. Biol.* **1998**, *18*, 7487-7498.

[30] Dasgupta, P.; Sun, J.Z.; Wang, S.; Fusaro, G.; Betts, V.; Padmanabhan, J.; Sebti, S.M.; Chellappan, S.P. Disruption of the Rb-Raf-1 interaction inhibits tumor growth and angiogenesis. *Mol. Cell. Biol.* **2004**, *24*, 9527-9541.

[31] Swain, J.; Mishra, A.K. Nile red fluorescence for quantitative monitoring of micropolarity and microviscosity of pluronic F127 in aqueous media. *Photochem. Photobiol. Sci.* **2016**, *15*, 1400-1407.

[32] Basak, R.; Bandyopadhyay, R. Encapsulation of hydrophobic drugs in Pluronic F127 micelles: Effects of drug hydrophobicity, solution temperature, and pH. *Langmuir* **2013**, *29*, 4350-4356.

[33] Spicer, C.D.; Jumeaux, C.; Gupta, B.; Stevens, M.M. Peptide and protein nanoparticle conjugates: Versatile platforms for biomedical applications. *Chem. Soc. Rev.* **2018**, *47*, 3574-3620.

[34] Bonacchi, S.; Genovese, D.; Juris, R.; Montalti, M.; Prodi, L.; Rampazzo, E.; Zaccheroni, N. Luminescent silica nanoparticles: Extending the frontiers of brightness. *Angew. Chem. Int. Ed.* **2011**, *50*, 4056-4066.

[35] Kapp, T.G.; Rechenmacher, F.; Neubauer, S.; Maltsev, O.V.; Cavalcanti-Adam, E.A.; Zarka, R.; Reuning, U.; Notni, J.; Wester, H.-J.; Mas-Moruno, C.; Spatz, J.; Geiger, B.; Kessler, H. A Comprehensive Evaluation of the Activity and Selectivity Profile of Ligands for RGD-binding Integrins. *Sci. Rep.* **2017**, *7*, 39805.

[36] Pfaff, M.; Tangemann, K.; Müller, B.; Gurrath, M.; Müller, G.; Kessler, H.; Timpl, R.; Engel, J. Selective recognition of cyclic RGD peptides of NMR defined conformation by $\alpha_{IIb}\beta_3$, $\alpha_v\beta_3$, and $\alpha_5\beta_1$ integrins. *J. Biol. Chem.* **1994**, *269* (32), 20233-20238.

[37] De Marco, R.; Rampazzo, E.; Zhao, J.; Prodi, L.; Paolillo, M.; Picchetti, P.; Gallo, F.; Calonghi, N.; Gentilucci, L. Integrin-Targeting Dye-Doped PEG-Shell/Silica-Core Nanoparticles Mimicking the Proapoptotic Smac/DIABLO Protein. *Nanomaterials* **2020***, 10,* 1211.

[38] Lutz, J.-F.; Börner, H.G.; Weichenhan, K. Combining ATRP and "Click" Chemistry: a Promising Platform toward Functional Biocompatible Polymers and Polymer Bioconjugates. *Macromolecules* **2006**, *39* (19), 6376-6383.

[39] Rampazzo, E.; Bonacchi, S.; Juris, R.; Montalti, M.; Genovese, D.; Zaccheroni, N.; Prodi, L.; Rambaldi, D.C.; Zattoni Andrea, C.; Reschiglian, P. Energy transfer from silica core-surfactant shell nanoparticles to hosted molecular fluorophores. *J. Phys. Chem. B* **2010**, *114*, 14605-14613.

[40] Valenti, G.; Rampazzo, E.; Bonacchi, S.; Khajvand, T.; Juris, R.; Montalti, M.; Marcaccio, M.; Paolucci, F.; Prodi, L. A versatile strategy for tuning the color of electrochemiluminescence using silica nanoparticles. *Chem. Commun.* **2012**, *48*, 4187-4189.

[41] Sanz, V.; Conde, J.; Hernández, Y.; Baptista, P.V.; Ibarra, M.R.; de la Fuente, J.M. Effect of PEG biofunctional spacers and TAT peptide on dsRNA loading on gold nanoparticles. *J. Nanopart. Res.* **2012**, *14* (6), 917.

[42] Aviñó, A.; Cuestas-Ayllón, C.; Gutiérrez-Capitán, M.; Vilaplana, L.; Grazu, V.; Noé, V.; Balada, E.; Baldi, A.; Félix, A.J.; Aubets, E.; Valiuska, S.; Domínguez, A.; Gargallo, R.; Eritja, R.; Marco, M.P.; Fernández-Sánchez, C.; de la Fuente, J.M.; Ciudad, C. J. Detection of SARS-CoV-2 Virus by Triplex Enhanced Nucleic Acid Detection Assay (TENADA) *Int. J. Mol. Sci.* **2022***, 23,* 15258.

[43] Idiago-López, J.; Moreno-Antolín, E.; Eceiza, M.; Aizpurua, J.M.; Grazú, V.; de la Fuente, J.M.; Fratila, R.M. From Bench to Cell: A Roadmap for Assessing the Bioorthogonal "Click" Reactivity of Magnetic Nanoparticles for Cell Surface Engineering. *Bioconjug. Chem.* **2022**, *33* (9), 1620-1633.

[44] Hong, V.; Presolski, S.I.; Ma, C.; Finn, M.G. Analysis and Optimization of Copper-Catalyzed Azide–Alkyne Cycloaddition for Bioconjugation. *Angew. Chem. Int. Ed.* **2009**, *48* (52), 9879-9883.

[45] Rampazzo, E.; Bonacchi, S.; Juris, R.; Genovese, D.; Prodi, L.; Zaccheroni, N.; Montalti, M. Dual-mode, anisotropy-encoded, ratiometric fluorescent nanosensors: Towards multiplexed detection. *Chem. Eur. J.* **2018**, *24*, 16743-16746.

Chapter V. Synthesis of Peptides Inhibiting the Assembly of Monomeric Human L-lactate Dehydrogenase into Catalytically-active Homotetramer

5.1 Introduction

Lactate dehydrogenases are stereospecific enzymes, exclusively generating D- or L-lactate at the expense of pyruvate.[1] In particular, the repertoire of bacterial lactate dehydrogenases (LDHs) contains representatives featuring dimeric or tetrameric quaternary structure, and relying on β-NADH or FAD as the redox cofactor.[2,3] The heterogeneity of prokaryotic LDHs does also concern their catalytic action, which is exerted according to Michaelis-Menten or cooperative kinetics,[2,4] with further complexity conferred to these enzymes by homotropic or heterotropic activation (e.g. by pyruvate and fructose 1,6-bisphosphate, respectively). Remarkably, the structural and functional divergence observed for bacterial LDHs is essentially mirrored when the eukaryotic counterparts are considered. Indeed, LDHs isolated from different eukaryotic sources were reported to be assembled as dimers or tetramers,[5-8] and to be assisted by β-NADH or FAD to exert their catalytic action.[9,10] Quite intriguingly, the occurrence of allosteric transitions in lactate dehydrogenases expressed by eukaryotic vertebrates was early reported,[11] and it was recently interpreted as the outcome of dissociation-association events affecting the tetrameric assembly of rabbit and human LDH.[12,13]

Concerning human lactate dehydrogenases, different homotetrameric forms were identified: (i) the M form (LDH 5, and featuring the LDH-A subunit), prevailing in muscles subjected to transient hypoxic conditions, and accordingly engaged in the reduction of pyruvate to L-lactate;[14,15] (ii) the H form (LDH 1, and containing the LDH-B subunit), mainly expressed in aerobic tissues, where it is supposed to primarily catalyze the oxidation of L-lactate;[16] (iii) the X form (LDH-C), detected in mitochondria isolated from spermatozoa.[17,18] In addition to the enzymes responsible for the interconversion of L-lactate and pyruvate, the occurrence in human mitochondria of a D-lactate dehydrogenase was recently shown.[19] Nevertheless, it is important to note that LDH 5 and LDH 1 represent the forms of human lactate dehydrogenase to which major attention was paid so far, both structurally and functionally speaking. Importantly enough, the quaternary structural analyses of LDH 5 and LDH 1[16] provide the necessary information to design inhibitors targeting the subunit-subunit interactions occurring in these enzymes. As will be later mentioned, this strategy represents an attractive alternative to conventional competitive inhibitors, the selectivity of which is questioned when dealing with dehydrogenases, that is, with enzymes sharing structural homology among their active sites.

Lactate is of primary importance for the energetic metabolism under conditions limiting oxidative phosphorylation, e.g. in cells facing transient hypoxia. However, it was shown that lactate can be produced in fully aerobic tissues,[20,21] and it was therefore proposed that glycolysis does always lead to the generation of lactate.[22] Indeed, the cytosolic lactate produced under fully aerobic conditions is transported to the mitochondrial intermembrane space, where it is oxidized to pyruvate, which is subsequently used by the Krebs cycle. Therefore, any condition limiting the oxidative catabolism of pyruvate (not necessarily oxygen limitation) would translate into an increase of lactate level, a state which is known to take place in cancer cells.[23] Remarkably, a decrease of intracellular pH (pHi) is usually observed when the energetic metabolism is dominated by the glycolytic generation of lactate. It should however be noted that the generation of lactate by LDH does not induce acidification inherently, the reduction of pyruvate at the expense of NADH + H^+ being a proton-consuming reaction. Instead, lactic acidosis represents the outcome of a net release of H^+ from ATP hydrolysis under hypoxic or anoxic environments.[24] A peculiar situation concerns cancer cells, the energetic metabolism of which is mainly driven by glycolysis and lactate release, and whose pHi is nevertheless higher than in normal cells.[25] Remarkably, this phenotype of cancer cells avoids the dissociation of tetrameric LDH-A, thereby preserving its catalytic action of LDH-A.[13]

An interesting link between the quaternary structure of LDH-A and its relevance in the energetic metabolism of cancer cells has recently been reported.[26] In particular, it was shown that tyrosine 10 (Y^{10}) of LDH-A is specifically phosphorylated in quite a number of human cancer cell lines. Moreover, elegant size exclusion chromatography experiments revealed that the phosphorylation of Y^{10} favors the assembly of tetrameric LDH-A at the expense of the enzyme monomeric and dimeric forms, the relative abundance of which is higher in preparations of LDH-A bearing unphosphorylated Y^{10}. Remarkably, these biochemical features do phenotypically translate in the observation that the invasive potential of different forms of malignant cells can be inhibited by suppressing the phosphorylation of LDH-A at Y^{10}.[27] Furthermore, it was reported that pyruvate counteracts the inhibition exerted by metformin on cancer cells proliferation.[28] This effect triggered by pyruvate is linked to the action of LDH, which is necessary to maintain the NAD^+/NADH molar ratio within values appropriate for cells growth and duplication.

Not surprisingly, the importance of LDH-A in the glycolytic pathway of malignant cells prompted a wide search of potent inhibitors of this enzyme. In particular, the active site of LDH-A represents the major target selected so far, and quite a number of competitive inhibitors mimicking pyruvate or β-NADH were synthesized and characterized.[29,30] Albeit highly effective inhibitors belonging to this class were reported (e.g. a glycolate derivative of β-NADH),[31] they face the drawback of lacking selectivity. Structurally speaking, NADH-dependent dehydrogenases do indeed feature a rather high level of similarity, the occurrence of which constrains the landscape of the searchable competitive inhibitors exerting a selective action on a particular member of this enzymatic family. In addition, the structural similarity of the human LDH-A and LDH-B represents an additional concern to the use of competitive inhibitors. Accordingly, to improve the selectivity of LDH antagonists, a couple of allosteric inhibitors (a phthalimide and a dibenzofuran derivative, respectively) were recently synthesized, and it was shown that their IC_{50} values against LDH-A and LDH-B lie in the nM and μM range, respectively.[32]

Notably, a further alternative to competitive inhibitors is represented by compounds targeting protein-protein interactions. In particular, the loss of activity related to the dissociation of tetrameric LDH-A[13,33] prompts the search for peptides able to interfere with the quaternary structure of this enzyme. Peptides can indeed mimic a secondary structure element supposed to be a major determinant of the reciprocal interactions engaged by the subunits of an oligomeric protein.[34] Nevertheless, it should be noted that the effectiveness of peptides in interfering in vivo with protein-protein interactions can be negatively affected by their high degree of conformational freedom and by their facile degradation in host cells. Remarkably, these complications can be overcome using conformationally-restricted cyclic peptides and by including D-amino acids in the primary structure of peptides, respectively. Therefore, peptides represent a promising tool to inhibit the assembly of enzymes that exclusively feature catalytic activity in their oligomeric state. Concerning lactate dehydrogenases, this strategy was first tested using different peptides targeting the C-terminal region of LDH-A.[35] According to dynamic light scattering analyses, a couple of peptides among those tested, i.e. the hexamers IYNLLK and KVVYNV, were found to be competent in inducing a shift of the apparent LDH-A molecular mass from about 240 to ca. 100 kDa.[35] Soon after, very elegant work was performed to identify peptides capable of binding human LDH 1 with high affinity.[36] To this aim, a dimeric form of LDH-B was used for testing the association of candidate peptides to the tetramerization site of LDH-B, i.e. to the C-terminal region of each enzyme subunit. By these means, an interesting observation was obtained: the linear peptide ATLKEKLI was found to feature a K_D for dimeric LDH-B equal to 1 mM, whereas the corresponding constant of cyclic derivatives of this octapeptide was found to decrease down to 25 μM.[36] Moreover, when the cyclic octapeptide best performing against dimeric LDH-B was used to target tetrameric LDH-A and LDH-B, K_D values equal to 117 and 380 μM were determined, respectively. Interestingly, an additional tetramerization site of LDH-B (corresponding to residues located in the E^{62}-F^{72} region of the primary structure) was identified, and a hexadecapeptide peptide directed against this target was observed to bind dimeric LDH-B with a K_D equal to 240 μM.[37] Quite intriguingly, the cyclic octapeptide and the linear hexadecapeptide share the capability of inducing the dissociation of tetrameric LDH-A, according to EC_{50} values equal to 172 and 262 μM, respectively. Remarkably, it was reported that peptides appropriately designed are proficient in inhibiting the activity of human LDH-A.[38] In particular, a heptadecapeptide (denoted cGmC9) consisting of two β-strands connected by two loops was shown to inhibit the reduction of pyruvate catalyzed by LDH-5, and the IC_{50} value for this antagonist action was determined as equal to 2.5 μM. Importantly enough, the cGmC9 peptide inhibits LDH 5 independently of substrate (pyruvate) concentration, therefore ruling out a competitive mechanism for this peptide. It should however be noted that the inhibition of LDH 5 by the cGmC9 peptide was assayed using the tetrameric enzyme exposed to pH 2.5 (in order to dissociate its subunits and let the peptide bind them under these conditions), and thereafter shifting the pH to 7.4. This procedure, besides being efficient, faces the inconvenient of exposing the target enzyme to harsh pH conditions, under which the peptide-enzyme association is also tested.

To provide a useful tool for screening the capability of candidate peptides to inhibit LDH-A activity, we considered of interest to attempt the isolation of homogeneous monomeric LDH-A, the availability of which would greatly improve the robustness of assays designed to test the assembly of LDH 5. In addition, to analyze in detail how the competence of peptides in inhibiting the assembly and the catalytic activity of LDH-A is affected by their structure, we synthesized quite a number of linear and cyclic peptides, containing or not D-amino acids. Accordingly, we report here on a simple procedure to obtain purified human LDH-A monomers, along with the synthesis and characterization of an ensemble of peptides efficiently targeting this enzyme in vitro and in cultured cells. The implications of our observations on the use of peptides to interfere with the assembly of oligomeric proteins are also discussed.

5.2 Results

5.2.1 Overexpression of human LDH-A and purification of its monomeric form

To attempt the overexpression of human LDH-A in *Escherichia coli*, we used a synthetic gene optimized for the codon usage of the bacterial host (Appendix C: Figure S1). Notably, protein extracts obtained from cells overexpressing the synthetic gene were found to contain soluble LDH-A, the purification of which was pursued with standard chromatographic methods (see Methods). In particular, the first purification step was carried out with a Cibacron Blue column, and further separation from contaminants was obtained by means of hydrophobic interaction chromatography. Remarkably, the purification level detected after these two chromatographic steps was rather satisfactory (Appendix C: Figure S2). Nevertheless, we decided to achieve further purification with the aid of gel filtration chromatography. To this aim, we used a Superdex 200 column, conditioned with 10 mM Tris-HCl, pH 7.5. Surprisingly, we were unable to detect any significant amount of tetrameric LDH-A in the eluted fractions, whereas a considerable amount of the enzyme in monomeric form was recovered from the column (Figure 5.1 A). Moreover, it should be noted that the molecular mass calculated from the observed elution volume of LDH-A was determined as equal to 19.8 kDa, a value significantly lower than that expected (36.6 kDa). To test if the isolation of monomeric LDH-A was due to the particular conditions selected for gel filtration, we prepared a second batch of enzyme. This was accomplished according to the procedure previously mentioned, except for the size exclusion chromatography, which was performed using 10 mM HEPES, pH 7.5. Quite intriguingly, monomeric LDH-A was again obtained, indicating that the output of the gel filtration chromatography is not affected by the cationic or anionic nature of the buffer selected for this purification step (Figure 5.1 B). However, when the Superdex 200 column was conditioned with 10 mM Tris-HCl (pH 7.5) supplemented with 125 µM β-NADH and 5 mM oxalate (a competitive inhibitor of lactate dehydrogenases, structurally similar to pyruvate), we recovered tetrameric LDH-A, the molecular mass of which was estimated equal to 157 kDa (Figure 5.1 C). In addition, a very similar output was observed when 10 mM HEPES, 150 mM NaCl (pH 7.5) was used for the size exclusion chromatography, with the peak containing tetrameric LDH-A corresponding to a molecular mass equal to 198 kDa (Figure 5.1 D). Finally, we tested if the addition of NaCl to previously purified monomeric LDH-A triggers its assembly into the corresponding tetramer. Interestingly, when 150 mM NaCl was supplemented to monomeric LDH-A and the enzyme solution accordingly obtained was loaded onto the Superdex 200 column conditioned with 10 mM Tris-HCl (pH 7.5) containing 150 mM NaCl, tetrameric LDH-A was obtained (Figure 5.1 E). Moreover, when the tetramer eluted by means of this size exclusion chromatography was loaded onto the Superdex 200 column conditioned with 10 mM Tris HCl devoid of NaCl, LDH-A was eluted in monomeric form (Figure 5.1 F). Therefore, the purification strategy reported here for human LDH-A is useful to isolate this enzyme in tetrameric or in monomeric form (yielding ca. 10 mg/L of purified enzyme), with the selection between these two forms depending only on the type of buffer used to perform gel filtration chromatography.

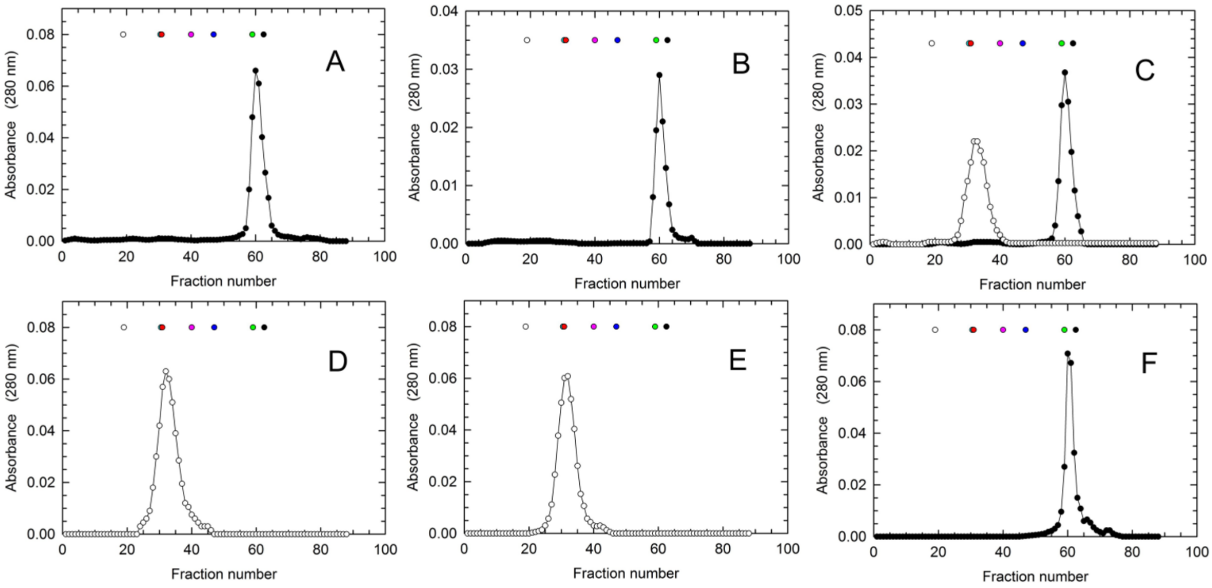
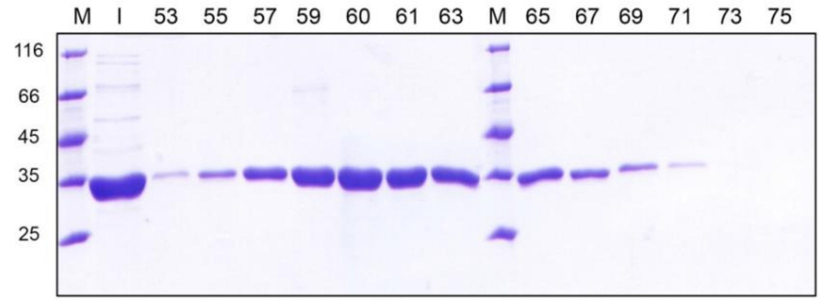


Figure 5.1 Isolation of mono- and tetrameric human LDH-A. The gel filtration chromatography was performed using: (A) 10 mM Tris-HCl (B) 10 mM HEPES (C) 10 mM Tris-HCl (black circles) or 10 mM Tris-HCl supplemented with 125 µM β-NADH and 5 mM oxalate (white circles) (D) 10 mM HEPES, 150 mM NaCl (E) 10 mM Tris-HCl, 150 mM NaCl (F) 10 mM Tris-HCl. The white, cyan, red, magenta, blue, green, and black circles indicate the elution volume of ferritin, catalase, aldolase, albumin, ovalbumin, chymotrypsinogen, and RNase A, respectively. The absorption spectra of these molecular mass markers are reported in Appendix C: Figure S3.

Considering that the deletion of few amino acids (e.g. 20) from the N-terminal of LDH-A is responsible for the dissociation of the tetrameric enzyme into dimers,[39,40] we analyzed by mass spectrometry one of the fractions eluted from the Superdex 200 column that contained monomeric LDH-A (Figure 5.2 A). Importantly enough, by these means we were able to detect peptides that covered most of the enzyme primary structure (from Asp^{6} to Lys^{332}, Figure 5.2 B), indicating that the purification procedure reported here implies the isolation of full-length monomeric LDH-A.

A



B

Peptide	Observed Mr (Da)	Theoretical Mr (Da)	Difference (ppm)	Peptide	Observed Mr (Da)	Theoretical Mr (Da)	Difference (ppm)
6-14	1118.6367	1118.6335	2.85	**172-177**	767.3623	767.5656	-1.70
6-22	2073.0767	2073.0742	1.20	**233-243**	1263.6731	1263.6711	1.59
77-90	1494.7713	1494.7678	2.30	**269-278**	1182.6556	1182.6543	1.05
77-99	2389.3415	2389.3329	3.60	**270-278**	1026.5534	1026.5532	0.19
82-90	1010.4706	1010.4669	3.70	**285-305**	2288.1753	2288.1723	1.33
82-99	1905.0320	1905.0320	0.027	**306-315**	1133.5568	1133.5564	0.33
91-99	912.5781	912.5756	2.74	**306-317**	1374.7363	1374.7354	0.60
91-106	1726.9308	1726.9326	-1.03	**306-318**	1502.8305	1502.8304	0.098
100-112	1555.8105	1555.8066	2.45	**318-328**	1245.6732	1245.6717	1.19
119-126	928.5752	9285.7460	0.63	**318-332**	1762.9252	1762.9254	-0.096
127-132	767.3294	767.3272	2.81	**319-328**	1117.5758	1117.5768	-0.89
156-169	1517.7395	1517.7369	1.77	**319-332**	1634.8303	1634.8304	-0.056
158-171	1550.7652	1550.7624	1.80				

Figure 5.2 Analysis of monomeric human LDH-A by SDS-PAGE and mass spectrometry. (A) Electrophoretic analysis of fractions eluted from the Superdex 200 column equilibrated with 10 mM Tris-HCl (pH 7.5) and used to perform the last purification step of monomeric LDH-A. M: molecular mass markers (their Mr is indicated in kDa at the left); I: input; the fraction numbers are indicated on the top. (B) Identification by mass spectrometry of peptides obtained by in-gel tryptic digestion of monomeric human LDH-A.

5.2.2 Structural properties of monomeric LDH-A

Having observed that β-NADH and oxalate are able to promote the tetramerization of LDH-A monomers, we considered of interest to characterize, by Dynamic Light Scattering (DLS), the association of the enzyme subunits. First, we analyzed the distribution of particles using monomeric LDH-A in 10 mM Tris-HCl, pH 7.5. Notably, when three consecutive DLS assays were carried out with this sample the observed peaks were indicative of a considerable polydispersity of the detected particles (Figure 5.3 A). Moreover, the size distribution of LDH-A particles revealed quite important differences among the three assays, suggesting a considerable spreading over a wide conformational landscape. In contrast, the addition of 125 μM β-NADH to monomeric LDH-A was found to translate into a monodisperse sample (Figure 5.3 B). Furthermore, no major differences were observed among the outputs generated by three successive DLS assays. In particular, under these conditions the molecular mass of LDH-A was determined as equal to 212 ± 35, 212 ± 31, and 190 ± 27 kDa, respectively, yielding a mean value equal to 205 ± 13 kDa. Remarkably, when both 125 μM β-NADH and 10 mM oxalate were added to monomeric LDH-A, the DLS peaks observed with three consecutive analyses featured high reciprocal similarity, and are indicative of a monodisperse sample (Figure 5.3 C). Under these conditions, the values of molecular mass of LDH-A were determined as equal to 153 ± 9, 153 ± 14, and 138 ± 16 kDa, respectively, yielding a mean value of 148 ± 9 kDa, in excellent agreement with an expected value for tetrameric LDH-A equals to 146.3 kDa.

To inspect the secondary structure of monomeric LDH-A, we compared its circular dichroism (CD) spectrum with that of the tetrameric enzyme. It should be noted that both spectra were obtained in PBS buffer, containing 137 mM NaCl. Therefore, the ionic strength of the buffer should induce the association of LDH-A monomers into the tetramer (cf. Figures 5.1 D and E). Interestingly, no major differences were detected between the two samples, with the only exception being the amplitude of molar ellipticity at 208 nm (Figure 5.3 D).

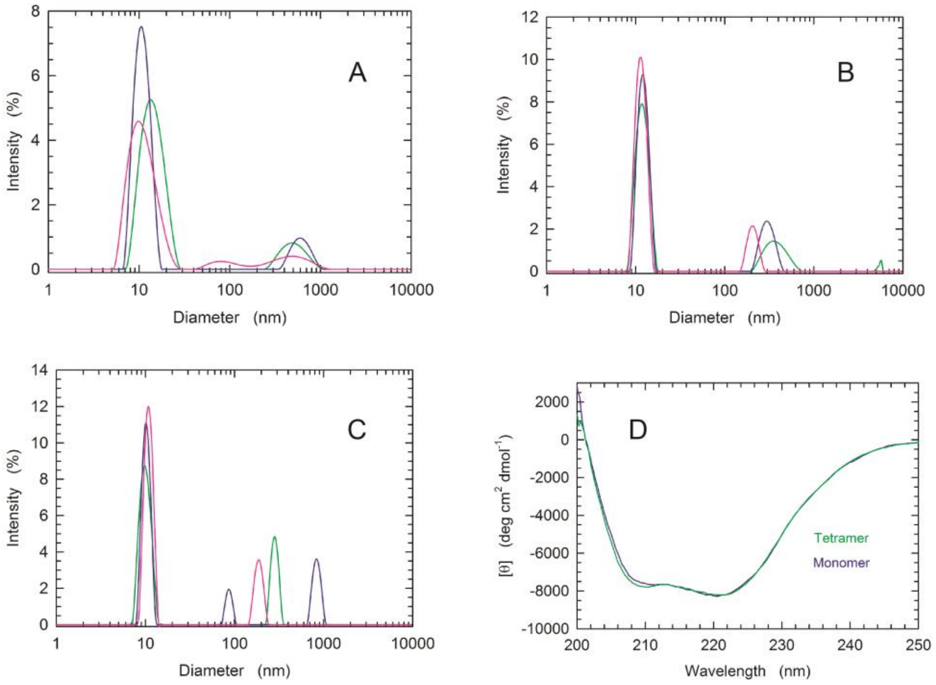


Figure 5.3 Analysis of the quaternary and secondary structure of human LDH-A. (A-C) DLS experiments performed with monomeric LDH-A in: (A) 10 mM Tris-HCl, (B) 10 mM Tris-HCl supplemented with 125 μM β-NADH, (C) 10 mM Tris-HCl containing 125 μM β-NADH and 5 mM oxalate. (D) Far-UV CD spectra of tetrameric and monomeric LDH-A (green and blue line, respectively), in PBS buffer, containing 137 mM NaCl.

5.2.3 Binding of β-NADH by tetrameric and monomeric LDH-A

To further inspect the biochemical features of monomeric LDH-A, we analyzed the binding of β-NADH to LDH-A (at pH 7.5) by Surface Plasmon Resonance (SPR). First, we tested the association between the tetrameric enzyme and the redox cofactor. Unfortunately, both the binding of β-NADH to tetrameric LDH-A and the subsequent dissociation of the binary complex were found to obey a very fast kinetics, whose rate constants could not be determined (Figure 5.4 A). Nevertheless, our observations were useful to determine the K_D of the enzyme-cofactor complex, the value of which was evaluated by means of three independent assays as equal to 28, 22, and 21 µM, yielding a mean value equal to 23 ± 4 µM (Figure 5.4 B). Surprisingly, when the binding of β-NADH to monomeric LDH-A was tested, the association of the redox cofactor to the monomeric enzyme featured a much lower amplitude than that detected in the presence of tetrameric *h*LDH-A (Appendix C: Figure S4, cf. Figure 5.4 A). However, by decreasing to 7.0 the pH of the binding assay, the observed amplitude was more than twice the corresponding response at pH 7.5 (Appendix C: Figure S4). Moreover, a further assay performed at pH 7.0 (i.e. under the more favorable condition) revealed a linear dependence of the SPR response up to very high β-NADH concentrations, e.g. 10 mM (Figures 5.4 C and D). Accordingly, the K_D for the monomer-coenzyme binary complex cannot be determined, with this finding being apparently in contrast with the observations obtained by means of DLS experiments (Figure 5.3). However, it should be considered that the DLS experiments were carried out with free LDH-A monomers, whereas the SPR assays were performed with immobilized enzyme which therefore is incapable to assemble into the tetrameric form. Accordingly, we propose that the binding of β-NADH and the assembly of monomers represent a concerted action, the occurrence of which is hampered by the immobilization of LDH-A.

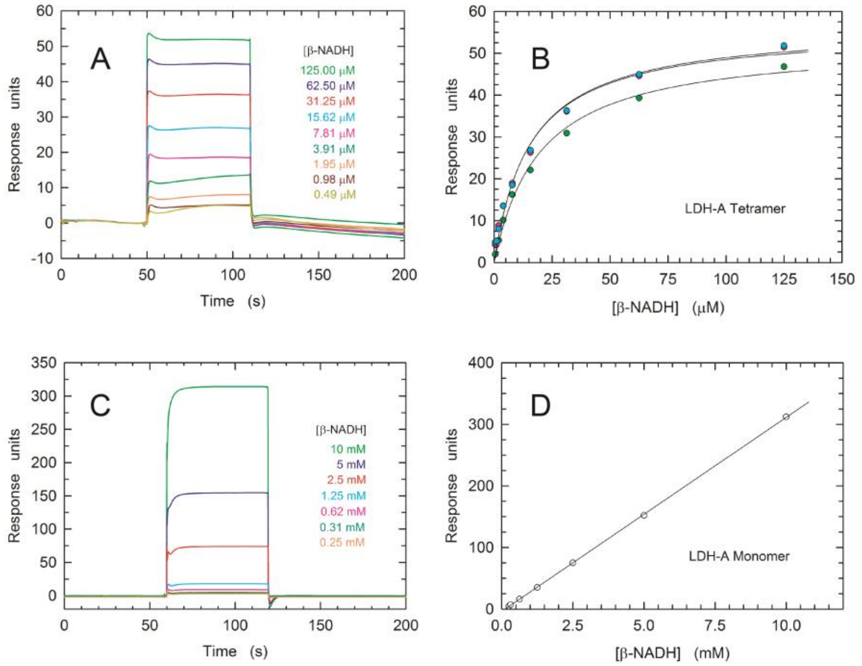


Figure 5.4 Binding of β-NADH to tetrameric and monomeric *h*LDH-A. (A) Kinetics of the association of β-NADH to tetrameric *h*LDH-A as detected by surface plasmon resonance. Sensorgrams were observed loading increasing concentration of β-NADH (0.49-125 μM) on a sensor chip modified with immobilized tetrameric human LDH-A. (B) Determination of the K_D of *h*LDH-A for β-NADH. The response units refer to the data shown in panel A. The continuous lines represent the best fit of a parametric rectangular hyperbola to the experimental observations. (C) Sensorgrams observed loading increasing concentration of β-NADH (0.25-10 mM) on a sensor chip modified with immobilized monomeric human LDH-A. (D) Linear dependence of the response units reported in panel C on the concentration of loaded β-NADH. The continuous line represents the best fit of a linear equation to the experimental observations.

5.2.4 Kinetics of pyruvate reduction by tetrameric and monomeric LDH-A

As previously mentioned, DLS assays revealed that the addition of β-NADH and oxalate to monomeric LDH-A triggers the assembly of LDH-A monomers into the enzyme tetrameric form. Accordingly, the assembled monomers should feature catalytic competence, the extent of which would be, in turn, suggestive of monomers-to-tetramer conversion. Therefore, we performed a set of activity assays using reaction mixtures containing 10 mM Tris-HCl (pH 7.5), 100 μM β-NADH, 500 μM pyruvate, and variable concentrations of LDH-A in monomeric or tetrameric form. To prime this analysis, we compared the kinetics of β-NADH oxidation catalyzed by tetrameric or monomeric LDH-A, at 7.6 nM final concentration of subunits. When the tetrameric enzyme was used, we detected a very fast reaction, reaching equilibrium in about 2.5 minutes (Figure 5.5 A, green line). Conversely, when monomeric LDH-A was used, the reaction rate was definitely slower, taking approximately 20 min to approach completion (Figure 5.5 A, blue line). We propose that this sharp difference represents the outcome of two major factors: i) a partial assembly of monomers into the enzyme homotetramer; ii) during the reaction the limiting substrate (β-NADH) is oxidized, therefore restraining the assembly of monomers. Taking into account that tetrameric LDH-A faces a dilution-induced dissociation, even under conditions of neutral pH,[13] further tests were carried out in the presence of much lower concentrations of enzyme. We indeed reasoned that in the presence of low enzyme concentrations the difference between the activities exerted by tetrameric LDH-A and by assembled monomers should be less pronounced when compared to the difference observed at higher enzyme concentrations. Therefore, LDH-A activity was assayed in the presence of concentrations of LDH-A subunits ranging from 0.0125 to 0.2 nM. As expected, tetrameric LDH-A outperformed the catalytic action of assembled monomers (Figures 5.5 B and C), albeit to a lower extent compared to the difference observed in the presence of 7.6 nM enzyme (Figure 5.5 A). Quantitatively speaking, by fitting a single exponential equation to the experimental observations larger k_{obs} values were determined for the LDH-A tetramer, except under the conditions of maximal dilution (0.0125 and 0.025 nM subunits, Figure 5.5 D).

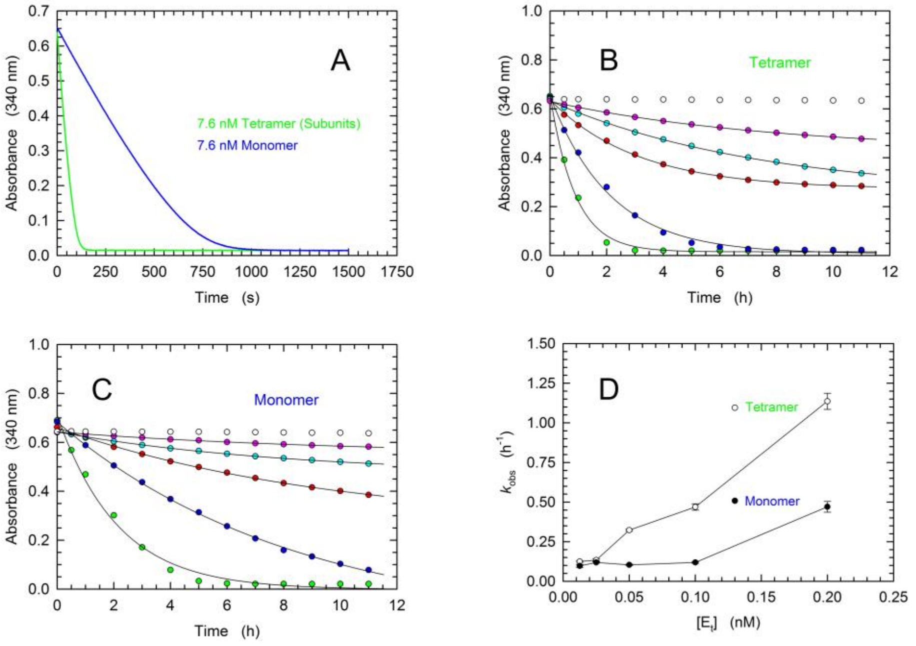
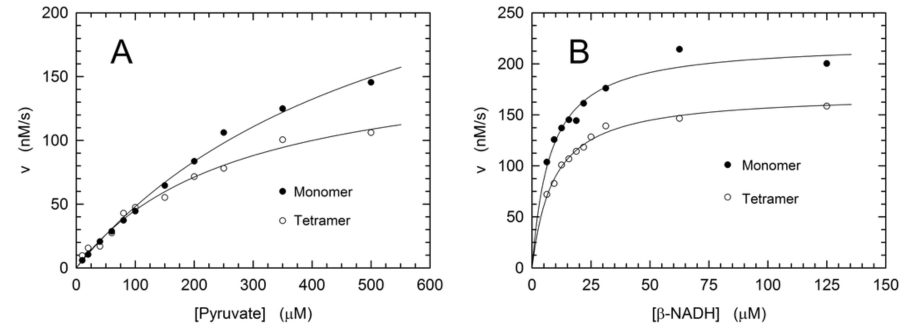


Figure 5.5 Kinetics of β-NADH oxidation catalyzed by tetrameric or monomeric human LDH-A. (A-C) Activity assays were performed using reaction mixtures containing 50 mM Tris-HCl (pH 7.5), 100 μM β-NADH, and 500 μM pyruvate. (A) Kinetics of β-NADH oxidation detected in the presence of 7.6 nM of tetrameric or monomeric LDH-A (green and blue line, respectively). (B, C) Time-course of β-NADH oxidation catalyzed by 12.5, 25, 50, 100, or 200 pM tetrameric (B) or monomeric (C) human LDH-A (magenta, cyan, red, blue, and dark green circles, respectively). The time-course observed in the absence of enzyme is also reported (white circles). (D) Dependence on enzyme concentration of the rate constants determined for the reactions catalyzed by tetrameric (white circles) or monomeric LDH-A (black circles) are reported in panels B and C. To obtain the k_{obs} values, a single-exponential equation was fitted to the experimental observations.

5.2.5 Kinetic parameters of tetrameric and monomeric LDH-A

To further characterize the catalytic action of monomeric LDH-A exposed to β-NADH and pyruvate, and therefore assembled into tetramer (cf. Figure 5.5 C), we determined the enzyme activity as a function of pyruvate or β-NADH concentration. In addition, the kinetics of pyruvate reduction catalyzed by tetrameric LDH-A was assayed under the same conditions used for the monomeric enzyme. Structurally and functionally speaking, the availability of monomeric LDH-A does indeed provide an unprecedented and quite interesting tool to perform quantitative tests of the assembly of monomers into the corresponding catalytically-active tetramer. When reaction velocity was observed as a function of pyruvate concentration in the presence of 125 μM β-NADH, tetrameric LDH-A outperformed its monomeric counterpart, featuring a lower *K*m and a higher k_{cat}, that is, 271 ± 40 vs 561 ± 74 μM and 66 ± 5 vs 32 ± 3 s^{-1} (Figures 5.6 A, C). In addition, the enzyme kinetic parameters were also determined testing reaction velocity as a function of β-NADH concentration, in the presence of 500 μM pyruvate. Under these conditions, tetrameric LDH-A was again observed to outperform the assembled monomers, essentially featuring the same *K*m for the redox cofactor (9.0 ± 0.6 vs 7.7 ± 1.1 μM) and a 3-fold higher k_{cat}, that is, 90 ± 2 and 29 ± 1 s^{-1}, respectively (Figures 5.6 B, C).



C

		Vmax (nM/s)	*K*m (μM)	k_{cat} (s^{-1})	k_{cat}/Km (s^{-1}• μM^{-1})
Pyruvate	Tetramer, 2.56 nM	167.6 ± 12.9	271.0 ± 39.8	65.5 ± 5.0	0.24 ± 0.04
	Monomer, 9.80 nM	317.6 ± 27.3	560.5 ± 73.6	32.4 ± 2.8	0.058 ± 0.009
β-NADH	Tetramer, 1.9 nM	170.9 ± 3.2	9.0 ± 0.6	89.9 ± 1.7	9.9 ± 0.9
	Monomer, 7.6 nM	221.1 ± 8.2	7.7 ± 1.1	29.1 ± 1.1	3.8 ± 0.7

Figure 5.6 Catalytic action of tetrameric and monomeric human LDH-A. (A) Dependence of the initial velocity of β-NADH oxidation observed as a function of pyruvate concentration, in the presence of 125 μM β-NADH and 2.56 nM tetrameric (white circles) or 9.8 nM monomeric (black circles) human LDH-A. The continuous lines represent the best fit of the Michaelis-Menten equation to the experimental observations. (B) Dependence of the initial velocity of β-NADH oxidation observed as a function of β-NADH concentration, in the presence of 500 μM pyruvate and 1.9 nM tetrameric (white circles) or 7.6 nM monomeric (black circles) human LDH-A. The continuous lines represent the best fit of the Michaelis-Menten equation to the experimental observations. (C) Kinetic parameters determined by performing activity assays using monomeric or tetrameric LDH-A, at the indicated final concentrations. Under conditions of variable pyruvate concentration, β-NADH was always used at 125 μM. When reaction mixtures contained variable concentrations of the redox cofactor, pyruvate was invariably added at 500 μM. All the assays were performed at pH 7.5 (Tris-BisTris, 10 mM each).

The divergence between the kinetic parameters determined for monomeric and tetrameric LDH-A should be interpreted in terms of the monomers-to-tetramer assembly. In particular, when the k_{cat} values are considered, the ratio of their values suggests that approximately one half to one third of monomers generates tetramer, in the presence of excess pyruvate and β-NADH, respectively. Moreover, it is interesting to note that the *K*m of assembled monomers for pyruvate is about twice the value determined for tetrameric LDH-A. This observation suggests that pyruvate is capable to stabilize tetrameric LDH-A by binding to a secondary site, featuring low affinity for the monocarboxylic acid.

5.2.6 Design of peptides directed against LDH-A subunit-subunit interactions

The inspection of two quaternary structures available for LDH 5 (PDB codes 1I10[16] and 4OJN[41]) led us to identify the N-terminal region of LDH-A as the target for inhibiting the assembly of the enzyme tetrameric form. In particular, since LDH 5 can be considered as a binary association of dimers, we focused on the interactions between the residues 10-20 of the N-terminal region of one monomer, and the residues 296-303 near the C-terminus of the second monomer (Figure 5.7 A). Accordingly, we designed the linear octapeptide **TH 1** as a potential inhibitor of LDH-A oligomerization, featuring the primary structure GQNGISDL, the sequence of which is identical to the stretch of residues 296-303 of the target protein. To correlate the coordinates of peptides with those of LDH 5 as denoted in the PDB files (shifted to a -1 position due to the lack of the methionine coded by the start codon), the positions of the amino acids in the peptides are hereafter accordingly indicated (e.g. the residues 296-303 translate into 295-302). Interestingly, it can be recognized that the stretch of amino acids 295-302 does clearly feature two turns, centered on the tetrapeptides G^{295}QNG^{298}, and I^{299}SDL^{302}, respectively (Figure 5.7 B). To design cyclic peptides inspired by these turns, we performed molecular modeling simulations. Plausible 3D structures of candidate turn mimetics were estimated by simulated annealing and molecular dynamics (MD) simulations, using the AMBER force field in explicit water.[42] In brief, random geometries of each cyclopeptide were sampled during a high temperature unrestrained MD simulation in a box of TIP3P models of equilibrated water molecules.[43] Each random structure was slowly cooled, the resulting structures were minimized, and the backbones of the structures were clustered by rmsd analysis. Only candidates showing one major cluster comprising the large majority of the geometries were considered. From this cluster the representative structures with the lowest energy were then selected and their conformations compared to those of the G^{295}QNG^{298} (Figure 5.7 C) and I^{299}SDL^{302} (Figure 5.7 F) turns. This procedure resulted in the in-silico selection of two mimetics of G^{295}QNG^{298}, i.e. the cyclopeptides c[GQN-isoD] (**TH 2**, Figure 5.7 D) and [GQN-(R)-isoD] (**TH 3**, Figure 5.7 E), in which either *(S)*- or *(R)*-Asp is framed within a 13-membered ring as a β-amino acid (isoAsp). The structures calculated for the 13-membered cyclotetrapeptides are not unexpected. Indeed, β-amino acids are well known to stabilize defined secondary structures, in particular γ- turns or pseudo-γ-turns.[44]

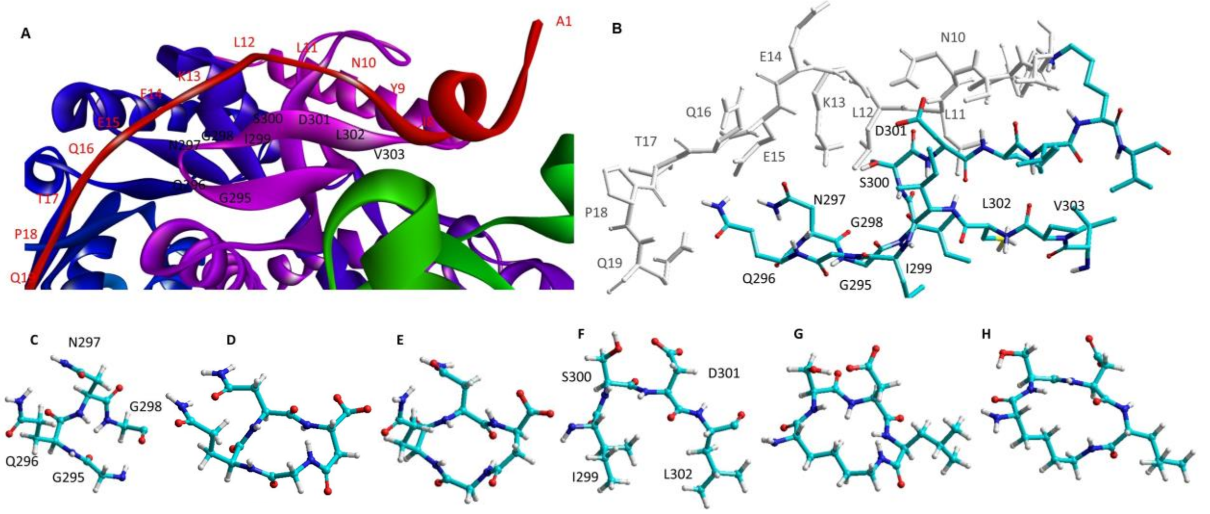


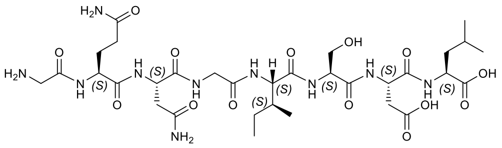
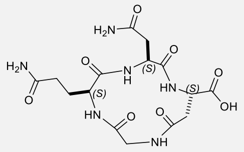
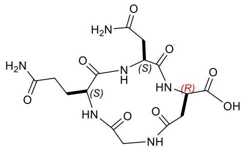
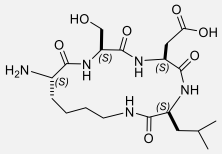
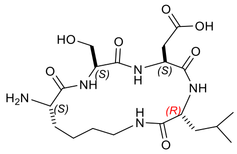
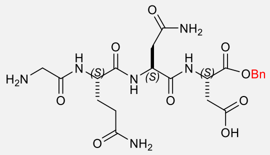
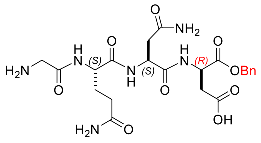
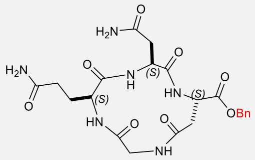
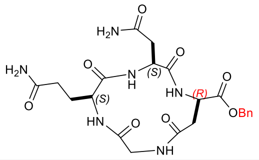
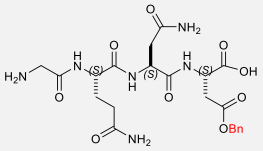
Figure 5.7 Design of the peptides to be tested against monomeric LDH-A. (A) Details of the interactions between the N-terminal region of one monomer with the residues 295-302 located at the C-terminus of a second monomer, and (B) stick representation of the same region. Structural details of: (C) the G^{295}QNG^{298} turn, and (D) the 13-membered cyclopeptides c[GQN-isoD] (**TH 2**), and (E) c[GQN-(*R*)-isoD] (**TH 3**), (F) the I^{299}SDL^{302} turn, and the cyclopeptides (G) c[isoKSDL] (**TH 4**) and (H) c[isoKSD-(*R*)-L] (**TH 5**).

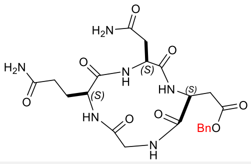
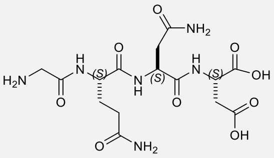
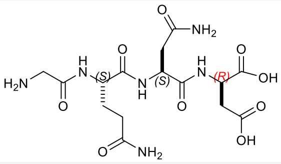
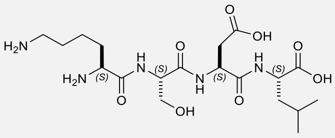
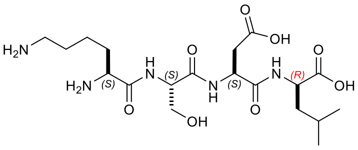
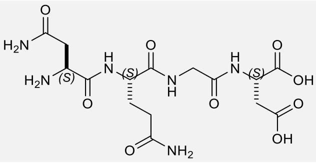
In addition, by means of this strategy two mimetics of I^{299}SDL^{302} were selected (Figure 5.7 F), namely the cyclopeptides c[isoKSDL] (**TH 4**, Figure 5.7 G) and c[isoKSD-*(R)*-L] (**TH 5**, Figure 5.7 H), in which isoLys is involved in the macrolactamization by the side-chain amino group (ε-NH_2). The chemical sketches of the cyclopeptides and of the majority of the intermediates are reported in Table 5.1. These cyclopeptides were prepared by cyclization of linear precursors, obtained in turn by solid phase peptide synthesis. To this aim, the tetrapeptide GQNG was modified into the partially protected GQND-OBn (**TH 6**) or GQN-*(R)*-D-OBn (**TH 7**), while ISDL was modified into Fmoc-KSD(OBn)L (**TH 8**) or Fmoc-KSD(OBn)-*(R)*-L (**TH 9**). These linear sequences were synthesized using standard procedures on a Wang resin using Fmoc-protected amino acids, and DCC/HOBt as activating agents (Appendix C: Table S1). *(S)*- or *(R)*-Asp was introduced as Fmoc-Asp(OBn)-OH, or Fmoc-Asp(OBn); Lys was introduced as Fmoc-Lys(Boc)-OH. Fmoc deprotection was carried out using 20 % (v/v) piperidine in DMF. Cleavage from the resin and simultaneous removal of the acid-labile protecting groups was performed by using TFA in the presence of scavengers.

The cyclization of the partially protected linear peptides was performed under pseudo-high dilution conditions, by slowly adding the peptide to a mixture of $NaHCO_3$ and DPPA in DMF. The crude cyclopeptides c[GQN-isoD(OBn)](**TH 10**), c[GQN-*(R)*-isoD(OBn)] (**TH 11**), c[isoK(Fmoc)-SD(OBn)L] (**TH 12**), c[isoK(Fmoc)-SD(OBn)-*(R)*-L] (**TH 13**), were isolated by RP HPLC on a semipreparative C18 column (Appendix C: Table S1). Removal of benzyl protecting groups was performed in quantitative yield by catalytic hydrogenation, giving c[GQN-isoD] (**TH 2**), c[GQN-*(R)*-isoD] (**TH 3**), while Fmoc deprotection was done as described above, affording c[isoKSDL] (**TH 4**), c[isoKSD-*(R)*-L] (**TH 5**). The purities of the final products were determined to be > 95 % by RP HPLC (Appendix C: Table S1), and their identity was confirmed by ESI-MS, ^{1}H and ^{13}C NMR, and by 2D gCOSY experiments at 600 MHz in DMSO-d_6.

We also prepared and tested the partially protected linear peptides GQND(OBn) (**TH 14**), whose cyclization provided the 12-membered c[GQND(OBn)] (**TH 15**). Furthermore, starting from the partially protected linear peptides described above, we prepared and assayed the fully deprotected linear sequences GQND (**TH 16**), GQN-*(R)*-D (**TH 17**), KSDL (**TH 18**), and KSD-*(R)*-L (**TH 19**) (Table 5.1, and Appendix C: Table S2).

Table 5.1 Structural features of the linear and cyclic peptides assayed as candidate inhibitors of LDH-A activity.

Peptide	Structure	Sequence	Inhibition (%)
TH1		GQNGISDL	36 ± 1
TH2		c[GQN-isoD]	31 ± 3
TH3		c[GQN-(*R*)-isoD]	26 ± 1
TH4		c[isoKSDL]	16 ± 3
TH5		c[isoKSD-(*R*)-L]	7 ± 3
TH6		GQND-OBn	14 ± 7
TH7		GQN-(*R*)-D-OBn	4 ± 2
TH10		c[GQN-isoD-(OBn)]	6 ± 3
TH11		c[GQN-(*R*)-isoD (OBn)]	8 ± 4
TH14		GQND(OBn)	15 ± 4

TH15		c[GQND(OBn)]	9 ± 2
TH16		GQND	22 ± 2
TH17		GQN-(*R*)-D	4 ± 8
TH18		KSDL	5 ± 6
TH19		KSD-(*R*)-L	4 ± 3
DS105		NQGD	0

Note: The SD (n = 3) associated with each estimated value of inhibition is indicated.

5.2.7 Peptides and inhibition of LDH-A activity

As previously mentioned, monomeric LDH-A assembles into the corresponding tetramer when exposed to β-NADH and pyruvate (or β-NADH and oxalate, see Figure 5.1 C), therefore gaining dehydrogenase activity. Accordingly, to test the action of peptides against the assembly of tetrameric LDH-A, we performed activity assays under steady-state conditions, in the presence of monomeric enzyme, 125 μM β-NADH and 0.5 mM pyruvate. In particular, all the peptides were tested at 80 μM (final concentration), and the LDH activity detected in their presence was compared with the catalytic action observed with control samples devoid of any peptides. First, we decided to assay the octameric peptide **TH 1**, the primary structure of which (GQNGISDL) is identical to the LDH-A region spanning residues 295-302. Remarkably, the presence of this octapeptide was found to decrease the observed LDH activity by 36 % (Figure 5.8 A, Table 5.1). Moreover, it is important to note that when the **TH 1** octapeptide was tested against tetrameric LDH-A, no significant inhibition of the enzyme activity was observed. Indeed, the catalytic action exerted by 6 nM tetrameric LDH-A (in 10 mM Tris-HCl, pH 7.5, containing 150 mM NaCl) was determined as equal to 433 ± 11 and 452 ± 19 nM/s in the absence and in the presence of 80 μM **TH 1** octapeptide, respectively. Overall, this insensitivity along with the significant effect elicited by **TH 1** on monomeric LDH-A indicate that the octapeptide inhibits the assembly of LDH-A monomers into the corresponding, catalytically active, tetramer. To dissect the inhibition triggered by **TH 1** on monomeric LDH-A, we decided to evaluate: i) whether or not the two halves of the octapeptide, i.e. GQNG and ISDL, are equally competent in inhibiting LDH-A; ii) the action of linear and cyclic peptides, containing or not D amino acids; iii) the effect, if any, triggered by the presence of protecting groups in the peptides to be assayed. The structures and the distinctive features of all peptides considered are reported in Table 5.1. It should be noted that to allow the cyclization of tetrapeptide GQNG while maintaining a terminal carboxylic group, the sequence was modified to GQND, while for the cyclization of ISDL, while maintaining a terminal amino group, the sequence was modified into KSDL (Table 5.1).

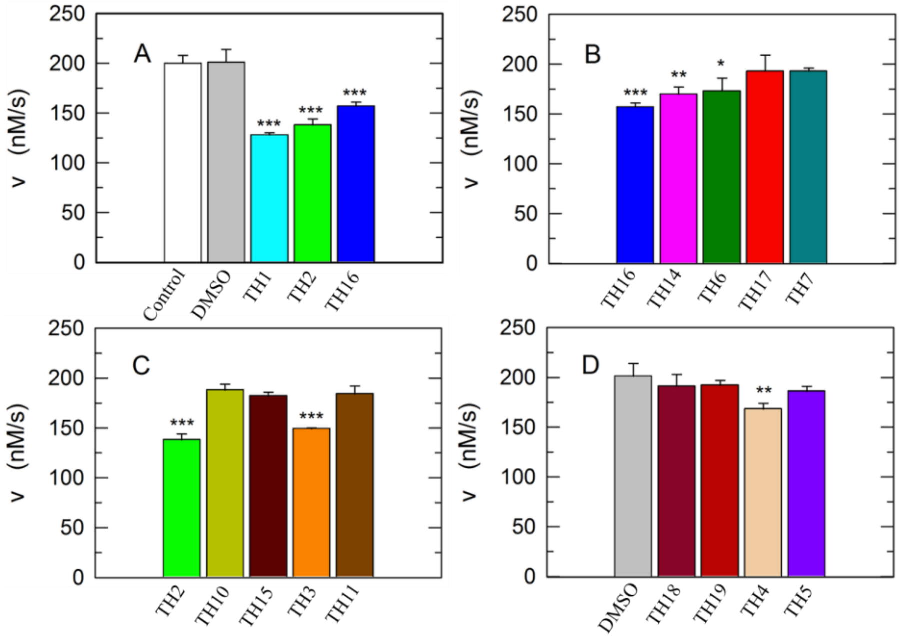


Figure 5.8 Inhibition exerted by linear and cyclic peptides on LDH-A activity detected in vitro. Activity assays were performed with reaction mixtures containing 6.2 nM monomeric LDH-A, 125 µM β-NADH, and 500 µM pyruvate in Tris-BisTris (10 mM each) buffer, pH 7.5 (white bar in panel A). The grey bars (in A and D) represent the activity observed in the presence of the same volume of DMSO (1 %, v/v) carried to the assay mixtures by the peptides to be tested. (A-C) The effect, if any, of tetrameric peptides featuring the GQND primary structure (see also Table 1) on LDH-A activity is shown. The inhibition exerted by the octameric peptide **TH1** is also reported in panel A. The enzyme activity observed in the presence of linear or cyclic peptides is shown in panels B and C, respectively. (D) Extent of LDH-A activity detected in the presence of linear or cyclic tetrapeptides featuring the KSDL primary structure. Error bars represent standard deviation (n = 3). The experimental observations were compared by Student's t-test. The ***, **, and * symbols denote P values lower than 0.001, 0.01, and 0.02, respectively.

Remarkably, the GQND linear tetrapeptide **TH 16** was found to be a quite effective inhibitor, being responsible for a 22 % decrease of LDH-A activity (Figure 5.8 A). In addition, to obtain a scrambled variant of **TH 16** presumably devoid of inhibitory action, we constructed **DS 105**, a linear tetrapeptide featuring the NQGD primary structure. As expected, the catalytic action of LDH-A is insensitive to the presence of **DS 105** in the assay mixture (data not shown). Furthermore, we did not detect a significant inhibition of LDH-A in the presence of **TH 17**, which differs from **TH 16** for the specific substitution of L-aspartate with D-aspartate (Table 5.1, Figure 5.8 B). Therefore, the replacement of a single L-amino acid with its D- counterpart does completely suppress the interference exerted by **TH 16** on the assembly of LDH-A monomers into the corresponding catalytically active tetramer. The effect, if any, of the presence of a protecting benzyl group in the GQND tetrapeptide was then analyzed. In particular, the inhibitory action of different protected peptides was compared with the action of the **TH 16** peptide, and the following observations were accordingly obtained: i) the **TH 14** peptide, bearing a benzyl group bound to the C-carboxyl (Table 5.1), performs less than its progenitor **TH 16** (Figure 5.8 B); ii) the **TH 6** peptide, featuring a benzyl group located at the Cβ-carboxyl (Table 5.1), does weakly perform when compared to **TH 16**, being responsible for 14 % inhibition of LDH-A activity (Figure 5.8 B); iii) the presence in **TH 7** of a benzyl group bound to the C-carboxyl (Table 5.1) does not confer a significant inhibitory action to this tetrapeptide, which therefore behaves as its inactive parental unprotected peptide **TH 17** (Figure 5.8 B). Rather interestingly, when cyclic and unprotected GQND tetrapeptides were assayed, a similar extent of LDH-A inhibition was observed with **TH 2** and **TH 3** (Figure 5.8 C), containing L- and D-aspartate, respectively (Table 5.1). In addition to this, we determined that cyclic GQND **TH 10**, **TH 11**, and **TH 15** tetrapeptides, bearing a benzyl protecting group, were outperformed by their unprotected counterparts (Figure 5.8 C).

Overall, the KSDL tetrapeptides were found less effective than their GQND counterparts (Figure 5.8 C and D, Table 5.1). In particular, the most potent cyclic **TH 4** tetrapeptide in this series was found to inhibit LDH-A activity by 16 % only (Figure 5.8 D), which is significantly lower than the level of inhibition exerted by **TH 2** (Figure 5.8 C).

5.2.8 Molecular modeling

According to our observations, among the cyclopeptides tested c[GQN-isoD] (**TH 2**) and [GQN-*(R)*-isoD] (**TH 3**), feature the highest inhibitory efficiency (Figure 5.8 C). Apparently, the cyclic structure contributes to improve their performances, possibly by imposing a proper 3D structure mimicking the $G^{295}QNG^{298}$ turn. Nevertheless, the presence of a free carboxylic group seems also quite important, as suggested by the very poor efficacy of the corresponding benzyl esters c[GQN-isoD(OBn)] (**TH 10**), and c[GQN-*(R)*-isoD(OBn)] (**TH 11**). To ascertain any role played by the free carboxylic group in the inhibition of LDH-A oligomerization, we predicted by molecular modelling and molecular dynamics the possible interactions between the cyclopeptide c[GQN-isoD] (**TH 2**) and the N-terminal target sequence of the monomer (Figures 5.7 A, B). The structures of the interacting N-terminal residues 9 to 19 and the C-terminal residues 295-302 were obtained from the quaternary structure reported by Read et al.[16] (PDB code 1I10). To prime the analysis, the cyclopeptide **TH 2** was superimposed to the G^{295}QNG^{298} turn, then the C-terminal sequence was removed. The system accordingly obtained was subjected to molecular dynamics simulations at 298 K in a box of explicit TIP3P equilibrated water molecules. During this period, the geometry of the N-terminal 9-19 sequence was restrained by applying a force field to dihedral angles and distances, while the cyclopeptide was maintained unrestrained. This step was followed by a stage with a scaled force field, and finally by a phase of unrestrained dynamics. During the last step, the conformation of the N-terminal sequence 9-19 slightly changed, allowing the formation of a salt bridge between the carboxylate of cyclopeptide **TH 2** and the protonated amino side chain of Lys^{13} of the target N-terminal region (Figure 5.9).

With all due caution, the simulations seem to point out this interaction as a main contributor of the comparatively higher inhibitory efficacy of the cyclopeptides.

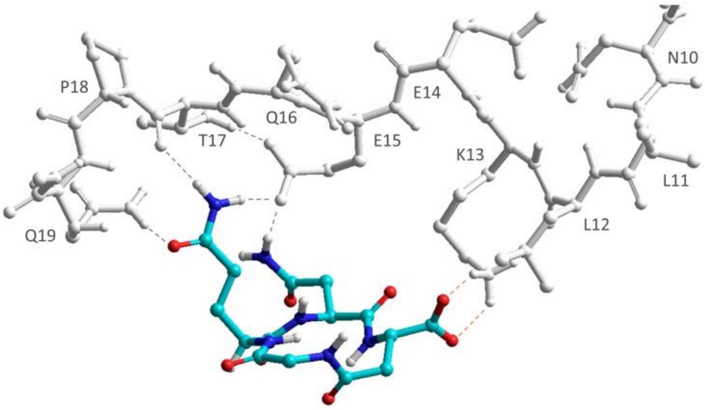


Figure 5.9 Simulation of the association between the **TH 2** peptide and monomeric LDH-A. Predicted interactions between the cyclopeptide c[GQN-isoD] (**TH 2**) and the N-terminal sequence Y^{9}NLLKEEQTPQ^{19}, as simulated by molecular dynamics in a box of explicit TIP3P water molecules; grey and red dotted lines represent hydrogen bonds and salt bridges, respectively.

5.2.9 Inhibitory action of peptides on lactate production by human cells

Having identified peptides able to interfere with the assembly of monomeric LDH-A into catalytically-active tetramer, we considered of interest to test the action of peptides against LDH-A expressed in human cell lines. In particular, we analyzed the lactate secreted by cultured human cells exposed or not to the peptides better performing as inhibitors of LDH-A in vitro, i.e. **TH 1**, **TH 2**, and **TH 16**. Moreover, we used two different cell lines exclusively expressing LDH-A or LDH-B, i.e. MCF7 and BxPC3, respectively. In MCF7 cells, the [mRNA$_{\text{LDH-A}}$]/[mRNA$_{\text{LDH-B}}$] ratio is indeed equal to ca. 150 (Figure 5.10 A), whereas in BxPC3 cells the messenger coding for LDH-A was not detected (Figure 5.10 A). Therefore, by comparing the decrease, if any, of the lactate concentration secreted by these two cell lines the selectivity of peptides acting against LDH-A or LDH-B can be conveniently estimated. When the MCF7 cell line is considered, the linear **TH 16** and the cyclic **TH 2** tetrapeptide performed equally, being both responsible for a concentration of secreted lactate 20 % lower than that detected in their absence (Figure 5.10 B). On the contrary, and rather interestingly, when BxPC3 cells were used, no significant differences were observed between the lactate secreted by cells exposed or not to the tetrapeptides (Figure 5.10 B). Surprisingly enough, the **TH 1** octapeptide did not induce any significant effect on lactate secretion by MCF7 and BxPC3 cells (Figure 5.10 B). We propose that the absence of a detectable phenotype linked to the treatment of cells with **TH 1** is due to a poor import of this octameric peptide, which proved to be rather effective when tested in vitro. It should indeed be noted that the assays with cells were performed in the presence of lipofectamine (see Methods). In addition, it is important to mention that when tested in the absence of lipofectamine both the **TH 16** and the **TH 2** tetrapeptide were found to trigger negligible and not significant effects on lactate production by MCF7 cells. Accordingly, it seems likely that lipofectamine is effective in assisting the penetration of the **TH 16** and **TH 2** peptides into cells, with this action of lipofectamine being much weaker when dealing with a larger molecule, i.e. an octapeptide.

A

	MCF7	BxPC3
LDH-A	1	Not detected
LDH-B	$6.6 \pm 0.9 \cdot 10^{-3}$	1

B

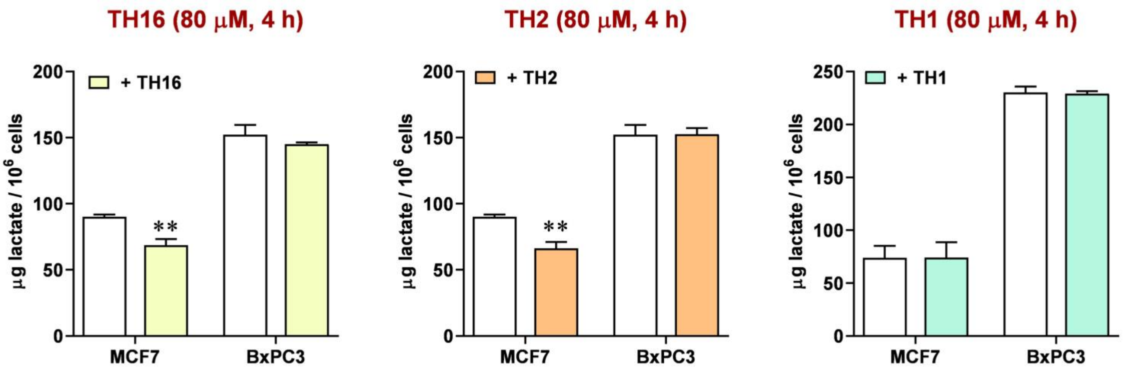


Figure 5.10 Effect of peptides on the secretion of lactate by human cell lines. (A) Relative $[mRNA_{LDH-A}]/[mRNA_{LDH-B}]$ as detected by RT-PCR in MCF7 and BxPC3 cells. (B) The amount of lactate secreted by MCF7 or BxPC3 cells in the absence (white bars) or in the presence (yellow, orange, and cyan bars in A, B, and C, respectively) of different peptides is shown. Error bars represent standard deviation (n = 3). The experimental observations were compared by two-way ANOVA. The ** symbol denote a P value lower than 0.01.

5.3 Discussion

So far, the functional and structural features of human LDH-A were almost exclusively inspected using LDH 5, i.e. the tetrameric form of this enzyme. Accordingly, we considered of interest to isolate homogeneous monomeric LDH-A, the availability of which could be of help to investigate the assembly of tetrameric LDH-A. In addition to this, it should be mentioned that the procedure reported here for the purification of monomeric human LDH-A can be conveniently modified to isolate the corresponding tetramer. In particular, the addition of β-NADH and oxalate or 150 mM NaCl to enzyme samples being subjected to the last purification step, i.e. gel filtration chromatography, translates into the isolation of LDH 5 (Figures 5.1, C-E). Remarkably, these observations are in substantial agreement with those reported by Trausch for rabbit skeletal muscle lactate dehydrogenase.[45] Tetrameric rabbit LDH-A was indeed found to dissociate into dimers when the enzyme was loaded onto a Sephadex G-200 column under conditions of high ionic strength (1 M NaCl), unless 100 μM β-NADH and 10 mM pyruvate were added to the buffer used for the gel filtration chromatography. Therefore, the observations shown here and those obtained by Trausch indicate that both the redox cofactor and the monocarboxylic substrate are important factors promoting the stability of tetrameric LDH-A. Quite intriguingly, by means of gel filtration chromatography we determined the molecular mass of monomeric human LDH-A as equal to 19.8 kDa (Figure 5.1), a value considerably lower than expected. This peculiar observation could represent the output of a particularly slow diffusion of human monomeric LDH-A along the void volume of the chromatographic column, implying an increase of the enzyme elution volume compared to the expected value, which is related to the diffusion of the enzyme into the chromatographic beads. This, in turn, implies that the conformation of monomeric human LDH-A is rather extended, conferring to the enzyme a molecular volume larger than that of a single subunit of tetrameric LDH-A. This feature of the enzyme could be responsible for the unexpected observations obtained when monomeric LDH-A was analyzed by DLS. Indeed, the diameters of monomeric LDH-A molecules would be, according to the DLS experiments, larger than those determined for the same protein in the presence of β-NADH or in the presence of both β-NADH and oxalate (Figure 5.3). It is important to remind that this last condition translates into the assembly of monomers into tetramers, the molecular mass of which was estimated by DLS in excellent agreement with the expected value.

When the competence of monomeric LDH-A in binding β-NADH was tested by surface plasmon resonance (SPR) a weak interaction was detected, the occurrence of which did not translate into an appreciable saturation of the binding site (Figure 5.4). This suggests that the binding of β-NADH is severely inhibited when the mutual association of LDH-A monomers is hampered in the immobilized state. Interestingly, when dissociated rabbit LDH-A was immobilized and exposed to free enzyme subunits the assembly into the tetrameric form occurred, albeit this competence was rapidly lost.[46] Conversely, when free monomeric LDH-A is hosted in solutions containing β-NADH and pyruvate the assembly into the tetrameric forms is triggered, conferring competence in catalytic action. Moreover, the extent of this assembly should be mirrored by the level of catalytic activity, which can be compared with the action exerted by the tetrameric enzyme at an equal concentration of subunits. Accordingly, activity assays were performed as a function of monomeric or tetrameric LDH-A concentration, in the presence of 100 μM β-NADH and 500 μM pyruvate. Remarkably, the observed kinetics of β-NADH oxidation suggest that 20-40 % of LDH-A monomers do assemble into tetrameric enzyme (Figures 5.5, B-D). In addition to these tests, we also determined the kinetic parameters of tetrameric and monomeric LDH-A. Concerning the tetrameric enzyme, k_{cat} values equal to 65.5 and 89.9 s^{-1} were determined under conditions of variable pyruvate or β-NADH concentration, respectively (Figure 5.6 C). The corresponding values for monomeric LDH-A were estimated as equal to 32.4 and 29.1 s^{-1} (Figure 5.6 C), suggesting that 30-50 % of monomers do assemble into the tetrameric form of the enzyme. It should be noted that these values are slightly higher than those (20-40 %) determined detecting the oxidation of β-NADH under conditions exclusively yielding first-order kinetics and implying the depletion of a considerable quantity of the redox cofactor (which with pyruvate is responsible for the assembly of tetrameric LDH-A, Figures 5.5, B-C).

When the *K*m for pyruvate is considered, we determined a value equal to 560 and 271 μM for monomeric and tetrameric LDH-A, respectively (Figure 5.6 C). It should be mentioned that the *K*m value of the tetrameric enzyme is in reasonable agreement with previous observations, yielding values equal to 252 and 398 μM.[13,47] Accordingly, monomeric LDH-A features a rather high *K*m for pyruvate and, as previously mentioned, we interpret this feature as indicative of a secondary site for pyruvate, possibly engaged in the stabilization of the assembled tetramer.

Structurally speaking, the four subunits of LDH 5 are arranged as a binary association of dimers, with their N- and C-terminal regions reciprocally engaged in dimer-dimer interactions.[16] As a target of peptides hopefully able to interfere with the assembly of LDH 5, we selected the N-terminal region of LDH-A. We indeed reasoned that any peptide identified as competent in inhibiting the assembly of monomeric LDH-A into tetrameric LDH-5 in vitro, could also be able to interact with the N-terminal part of the nascent LDH-A displayed by ribosomes during mRNA translation. Moreover, it should be considered that it was previously shown that the N-terminal region of LDH-A represents a better target than its C-terminal part.[38] Accordingly, we designed the **TH 1** octapeptide, featuring the primary structure GQNGISDL, which is identical to the C-terminal part of LDH-A consisting of residues 295-302. Remarkably, the effectiveness of this octapeptide against the assembly of LDH 5 is rather satisfactory, its presence in assay mixtures being responsible for a 36 % decrease of the observed LDH 5 activity. This observation diverges from the output reported for a quite similar peptide tested against LDH-A, i.e. the nonamer Ac-QNGISDLVK-NH_2 (denoted **C1**), which resulted ineffective.[38] The reason for this apparent discrepancy should reside in the different procedures used to evaluate the action of the **TH 1** and **C1** peptides. Indeed, Nadal-Bufi et al. assayed the **C1** nonapeptide using LDH 5 exposed to pH 2.5 to dissociate its subunits, incubating the dissociated enzyme to the peptide for 2 minutes, and finally shifting the pH to 7.4. On the contrary, we used monomeric LDH-A at pH 7.5 as the target of peptides selected as candidates to interfere with the assembly of LDH 5. Therefore, it seems quite likely that exposing both the target and the candidate peptide to pH 2.5 could limit their mutual interactions.

As previously mentioned, we dissected the action of **TH 1** using two different families of tetrapeptides, and we attempted to identify the molecular determinants conferring to the better performers the competence in inhibiting the assembly of LDH 5. Accordingly, we recognized the importance of constructing unprotected peptides, and featuring a cyclic structure. In particular, we observed that the cyclic **TH 2** tetrapeptide outperformed its conjugate linear **TH 16** peptide, with both of them featuring the GQND primary structure. We propose that this difference is likely due to the conformational stability conferred by cyclization to **TH 2**, the binding competence of which to monomeric LDH-A would nevertheless be supported by the stretchy conformation of the target. According to the tertiary structure of LDH-A, the N-terminal region spanning amino acids 9-19 is indeed arranged as a coil, therefore featuring conformational freedom in the monomeric state. Nevertheless, the cyclic nature of **TH 2** implies that a limited number of amino acids, residing on one side of this tetrapeptide, would be able to interact with the target. In particular, we propose that the side containing the couple of residues QN could favorably interact with residues located in the region of the target spanning amino acids 14-19, as previously reported.[35] In addition to the difference between the action of peptides **TH 2** and **TH 16**, it is important to note further unprecedented observations that we have shown here: i) the inhibition exerted on LDH-A by linear peptides is sensitive to their stereochemistry, as indicated by the finding that replacing the L-aspartate in the effective **TH 16** peptide with a D-aspartate yields the ineffective conjugate **TH 17** peptide (Table 5.1); ii) the inhibition of LDH-A triggered by cyclic peptides is insensitive to the presence of L- or D-aspartate, as determined for the **TH 2** and **TH 3** peptides (Table 5.1). Remarkably, these observations suggest the use of conformationally-restricted cyclic peptides, preferably containing D-amino acids to limit their degradation in vivo.

We reported here for the first time on the inhibition of LDH-A exerted in cultured cells by tetrapeptides. Remarkably, by means of assays detecting the lactate secreted by cell lines exclusively expressing LDH-A or LDH-B we were able to show that the tetrapeptides **TH 16** and **TH 2** feature high selectivity against LDH-A. However, the octapeptide **TH 1** that was found to be the best performer in vitro did not elicit any effect in cultured cells, probably because of a limited import of this peptide in the cellular cytosol. Accordingly, it will be of interest to construct chimeric peptides, containing a region targeting LDH-A and an arginine- and/or lysine-rich part, the presence of which could mimic cell-penetrating peptides.[48-50] The heptadecameric cGmC9 peptide constructed and tested by Nadal-Bufi et al. does indeed contain about half of its primary structure designed to contain four arginines,[38] most likely favoring its import into cells' cytosol, the compartment where LDH-A is located and can therefore be targeted. Moreover, and importantly, it was shown that this type of peptides can be successfully used to inhibit the proliferation of cancer cells.[51]

5.4 Conclusion

We have reported here on the isolation of homogeneous human LDH-A in monomeric form by means of a simple purification procedure performed at pH 7.5. Remarkably, this finding provides a convenient and reliable tool to assay candidate inhibitors directed against the assembly of monomeric *h*LDH-A into the corresponding, catalytically active, tetramer. Moreover, we have shown that appropriately designed cyclic peptides can exert this inhibitory action. Considering the relevance of LDH-A in the energetic metabolism of malignant cells, it is our hope that the availability of monomeric *h*LDH-A will be of help to the identification of further peptides repressing glycolysis in cancer cells by means of a strong inhibition of the protein-protein interactions essential to the assembly of tetrameric *h*LDH-A.

5.5 Experimental Section

Purification of LDH-A monomer. To obtain human LDH-A in monomeric form, a gene coding for this enzyme and optimized for *Escherichia coli* codon usage was synthesized (GenScript, Leiden, NL) and cloned into the pET9a expression vector using the *Nde*I and *Bam*HI sites (Appendix C: Figure S1). The construct accordingly obtained was used to transform *E. coli* BL21 (DE3) electrocompetent cells. Transformants were selected using Petri dishes containing LB solid medium supplemented with kanamycin (40 µg/mL), and subsequently purified on LB-kan plates. Single colonies were then picked up to inoculate LB-kan liquid medium, and grown for 15 h at 37 °C under shaking conditions (180 rpm). The pre-cultures accordingly obtained were diluted (1:500) into fresh LB-kan medium and grown at 30 °C for 9 h. Then, 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added, and the induced cultures were incubated at 30 °C for 15 h. Induced cells (300 mL) were harvested by centrifugation (4,500 x g, 20 min, 4 °C), and the pellets were resuspended in 15 mL of 50 mM Na_2HPO_4, 150 mM NaCl, 1 mM EDTA, pH 7.5 (buffer A). Total proteins were extracted by sonication (Misonix-3000 sonifier, output level of 18 W for 15 s, followed by a 15 s cooling interval, for 4 cycles). The crude protein extract accordingly obtained was centrifuged (13,000 x g, 30 min, 4 °C), the pellet was discarded, and the supernatant was immediately loaded onto a HiTrap Blue column (5 mL, GE Healthcare, Piscataway, USA), previously equilibrated with buffer A. After extensive washing of the column with buffer A, LDH-A was eluted with the same buffer supplemented with 0.5 M NaCl. The best fractions (as determined by activity assays and SDS-PAGE analysis) were pooled, and solid NaCl was added to increase its concentration to 3 M in the pooled fractions. The sample accordingly obtained was loaded onto a HiTrap Phenyl FF column, previously equilibrated with Buffer A containing 3.0 M NaCl. After washing the column, a reverse NaCl gradient (3.0-0 M) was applied (10 column volumes), and afterwards LDH-A was eluted with buffer A. The best eluted fractions were pooled, concentrated to 2 mL and loaded onto a Superdex-200 column, equilibrated with 10 mM Tris-HCl, pH 7.5 (or alternatively 10 mM HEPES, pH 7.5). After pooling and concentrating the fractions containing homogeneous LDH-A, aliquots of the concentrated purified enzyme were stored at -20 °C until used. Protein concentration was determined according to Bradford.[52]

Activity assays. The enzyme-catalyzed reduction of pyruvate was assayed by determining the decrease in Absorbance at 340 nm related to the oxidation of β-NADH. The extinction coefficient of β-NADH at 340 nm was considered equal to $6.22 \bullet 10^3$ M^{-1} cm^{-1}.[53] All the assays were performed at 25 °C using a Cary 300 Bio spectrophotometer. To determine the K_m of monomeric or tetrameric LDH-A for pyruvate, reaction mixtures contained 125 µM β-NADH in 50 mM Tris-HCl, pH 7.5. Conversely, to determine the K_m of monomeric or tetrameric LDH-A for β-NADH reaction mixtures contained 500 µM pyruvate in 50 mM Tris-HCl, pH 7.5.

The inhibition exerted by peptides on LDH-A activity was tested by incubating 6.2 nM enzyme and 80 µM peptide in 10 mM Tris-HCl, 10 mM BisTris (pH 7.5) buffer for 5 minutes, at the end of which the reaction was started by the addition of 125 µM β-NADH and 500 µM pyruvate.

Dynamic light scattering. Dynamic light scattering experiments were performed with a Malvern Panalytical (Malvern, UK) Zetasizer Nano ZS system. All the measurements were recorded at 25 °C using solutions previously filtered with 0.2 µm filters, and containing 5 µM enzyme in 10 mM Tris-HCl, pH 7.5. Scattering was evaluated at an angle of 173 degrees. Each individual observed value of enzyme diameter represents the average output of 3 groups of consecutive determinations. Raw data were analyzed with the Zetasizer software (Malvern Panalyticals), release 7.11.

Circular dichroism. CD spectra were recorded over the 200-250 nm wavelength interval at a scan rate equal to 50 nm/min, using a Jasco J-810 spectropolarimeter and a 0.5 cm path-length cuvette. Protein samples were in PBS, and the band width was set to 1 nm. Sixteen scans per sample were acquired and averaged.

Surface plasmon resonance. The binding of β-NADH by monomeric and tetrameric *h*LDH-A was assayed using a Biacore T200 instrument (Cytiva, Marlborough, MA, USA). Both monomeric and tetrameric LDH-A were immobilized on a Xantec CMD500 chip at a concentration of 30 µg/mL in 10 mM HEPES, pH 7, using a standard amine coupling protocol, yielding an observed immobilization level of 6984 and 17424 response units (Rus), respectively. The binding of β-NADH to *h*LDH-A was tested in 50 mM HEPES, pH 7.5, at 25 °C, using a flow rate of 50 µL/min, with 60 and 10 s of association and dissociation time, respectively. Different ranges of β-NADH concentration were used, as indicated in the figures. The elaboration of sensorgrams was achieved using the BIAevaluation software.

Mass spectrometry. To verify the identity of monomeric LDH-A, spots were excised from SDS-PAGE gels and underwent trypsin-in-gel digestion as previously reported.[54] The resulting peptides were analyzed by LC-MS/MS using a Q-Exactive instrument (Thermo-Fisher Scientific, Waltham, MA, USA) equipped with a nano-ESI source coupled with an Ultimate capillary UHPLC as reported elsewhere.[55]

Cell cultures and determination of lactate dehydrogenase expression. To evaluate the competence of peptides in inhibiting the action of LDH-A in cultured cells, two human cell cultures were used: MCF7 and BxPC3, isolated from breast and pancreatic adenocarcinoma, respectively. Cultures were grown in DMEM (MCF7) or RPMI medium (BxPC3), supplemented with 100 U/mL penicillin/streptomycin, 2 mM glutamine and 10 % (v/v) FBS. All the components of the media used for cell culture were obtained from Merck-Millipore. The expression of LDH subunits in both cell lines was evaluated by RT-PCR. RNA extraction, retro-transcription and DNA amplification were performed as previously described.[56] For each culture, the relative expression *ldh-a/ldh-b* was evaluated. The primers used for the RT-PCR reaction were indicated:

a) for LDH-A: 5'-GCAACCCTGCAACGATTTT-3' (forward);
5'-TTCCAGAGGACAAGATCTCAAA-3' (reverse);
b) for LDH-B: 5'-CCAACCCAGTGGACATTCTT-3' (forward);
5'-AAACACCTGCCACATTCACA-3' (reverse).

Colorimetric assays of lactate secretion by cultured human cells. The competence of peptides in inhibiting LDH in cultured cells was tested by assessing the concentration of lactate released by MCF7 and BxPC3 cell cultures. To this aim, cells were seeded in duplicate in 24-wells plates (2×10^5 / well) and let to adhere overnight; the culture medium was then replaced by Krebs-Ringer buffer (300 µL/well). Stock solutions of peptides (8 mM in DMSO) were diluted in Krebs-Ringer buffer to a final concentration of 640 µM and mixed with an equal volume of 4 % (v/v, in Krebs-Ringer buffer) Lipofectamine 2000 (Invitrogen). The final solution was thoroughly mixed, kept for 25 minutes at room temperature, and finally 100 µL were added to each plate well, obtaining a final peptide concentration equal to 80 µM. The concentration of lactate released in the Krebs-Ringer buffer was determined after 4 hours of incubation at 37 °C, using the colorimetric assay previously reported.[56] Each experiment was performed in triplicate and included control samples consisting of untreated cell cultures exposed to the same final concentration of DMSO (1 %, v/v) and Lipofectamine (0.5 %, v/v) to which were exposed the samples treated with peptides.

Synthesis of peptides, general procedures. Unless otherwise stated, standard chemicals and solvents were purchased from commercial sources and used as received without further purification. The purity of target compounds was determined to be ≥ 95 % by HPLC analyses, performed using an Agilent 1200 series apparatus, equipped with a Phenomenex reverse-phase column (Gemini C18, 3 µm, 110 Å, 100 × 3.0 mm, no. 00D4439-Y0). Column description: stationary phase octadecyl-carbon chain-bonded silica (C18) with trimethylsilyl end-cap, fully porous organo-silica solid support, particle size 3 µm, pore size 110 Å, length 100 mm, internal diameter 3 mm; mobile phase for neutral compounds: from 9:1 H_2O/CH_3CN to 2:8 H_2O/CH_3CN in 20 min at a flow rate of 1.0 mL min^{-1}, followed by 10 min at the same composition; DAD (diode-array detection) 210 nm; mobile phase for ionizable peptides: from 9:1 H_2O/CH_3CN and

0.1 % HCOOH to 2:8 H_2O/CH_3CN and 0.1 % HCOOH in 20 min, flow rate of 1.0 mL min^{-1}; DAD 254 nm. Semipreparative RP HPLC was carried out with a Waters 2489 UV/visible Dual Detector equipped with a Waters 1525 Binary HPLC pump, using reverse-phase column XSelect Peptide CSH C18 OBD column (Waters, 19 × 150 mm, 5 μm, no.186007021). Column description: stationary phase octadecyl-carbon chain-bonded silica (C18), double end-capped, particle size 5 μm, pore size 130 Å, length 150 mm, internal diameter 19 mm; DAD 210 nm, DAD 254 nm. Mobile phase for neutral compounds: isocratic 6:4 H_2O/CH_3CN for 2 min, then gradient from 6:4 H_2O/CH_3CN to 2:8 H_2O/CH_3CN in 10 min, then isocratic 2:8 H_2O/CH_3CN for 2 min; flow rate: 10 mL min^{-1}. Mobile phase for ionizable peptides: isocratic 6:4 H_2O/CH_3CN and 0.1 % HCOOH for 2 min, then gradient from 6:4 H_2O/CH_3CN and 0.1 % HCOOH to 2:8 H_2O/CH_3CN and 0.1 % HCOOH in 10 min, then isocratic 2:8 H_2O/CH_3CN and 0.1 % HCOOH for 2 min. Routine ESI MS analysis was carried out using an MS single quadrupole HP 1200 MSD detector, with a drying gas flow of 12.5 L min^{-1}, nebulizer pressure 30 psig, drying gas temperature 350 °C, capillary voltage 4500 (+) and 4000 (−), scan 50-2600 amu. NMR spectra were recorded with a Bruker BioSpin GmbH (^1H: 600 MHz, ^{13}C: 151 MHz) at 298 K in 5 mm tubes, using 0.01 M peptide. Chemical shifts are reported in ppm (δ) and internally referenced with $(CD_3)_2SO$: ^1H: 2.50, ^{13}C: 39.52 ppm. The unambiguous assignment of ^1H NMR resonances was based on 2D gCOSY experiments. Coupling constants (J) are reported in Hz.

The NMR spectra and the HPLC chromatograms of the synthesized peptides are reported in Appendic C: Figures S5 A-Q and S6 A-C, respectively.

General procedure for SPPS of linear peptides. Solid-phase peptide synthesis was performed in polypropylene syringes fitted with a polyethylene porous disc. The linear peptides were assembled manually on Wang resin (0.3 g, 1.1 mmol/g loading capacity) using Fmoc-protected amino acids under standard procedures; (*S*)- or (*R*)-Asp was introduced as Fmoc-Asp(OBn)-OH, or Fmoc-Asp-(OBn); Lys was introduced as Fmoc-Lys(Boc)-OH.

Prior to use, the resin was swollen in DMF (3 mL) for 15 min. In a separate sample tube, (*S*)- or (*R*)-Fmoc-AA-OH (0.6 mmol) and HOBt (81 mg, 0.6 mmol) were dissolved in DMF (4 mL). After 20 min, the mixture was added to the resin, followed by DCC (124 mg, 0.6 mmol) and a catalytic amount of DMAP, and the resin was gently shaken at RT for 3 h. Thereafter, a mixture of Ac_2O (0.92 mL, 10 mmol) and pyridine (0.81 mL, 10 mmol) was added and shaken for additional 30 min to end-cap the unreacted 4-hydroxybenzyl alcohol linkers. The resin was filtered and washed alternatively with DMF, MeOH, and DCM (3 × 4 mL each).

Fmoc cleavage was carried out using 20 % (v/v) piperidine in DMF (5 mL), while gently shaking at RT for 20 min. After washing with DMF and DCM (5 mL), the deprotection was repeated. The resin was then washed sequentially with DMF, MeOH, and DCM (3 × 4 mL).

The subsequent coupling reactions were performed by dissolving in a separate vial Fmoc-protected amino acids (0.6 mmol) and HOBt (81 mg, 0.6 mmol) in DMF (4 mL) for 20 min.

The mixture was poured into the reactor followed by DCC (124 mg, 0.6 mmol), and the suspension was shaken at RT for 3 h. Coupling efficacy was monitored by the Kaiser test. Fmoc cleavage was carried out as reported above.

Cleavage from the resin and simultaneous removal of the acid-labile protecting groups was performed by using a 95:2.5:2.5 v/v/v mixture of TFA/TIPS/H_2O (10 mL) while shaking for 2.5 h at RT. The mixture was filtered and the resin washed twice with 5 % TFA in Et_2O (10 mL). The filtrates were collected and solvents were removed under reduced pressure. Then ice-cold Et_2O was added to precipitate the crude peptides as TFA salts, which were recovered by centrifuge. The linear peptide candidates were isolated by RP HPLC on a semipreparative C18 column. The purity and the identity of the products were determined to be ≥ 95 % by RP HPLC coupled to ESI MS, by ^{1}H and ^{13}C NMR. Linear sequences destined to cyclization were utilized without further purifications.

General procedure for SPPS of cyclopeptides. The cyclization of the crude peptides was performed under pseudo-high dilution conditions, by slowly adding the peptide using a temporized syringe pump (dual-channel KD Scientific model 200). A solution of the linear peptide (0.15 mmol) in DMF (10 mL) was added to a mixture of $NaHCO_3$ (38 mg, 0.45 mmol) and DPPA (65 µL, 0.3 mmol) in DMF (4 mL) at RT over 16 h. Once the addition was complete, the reaction was stirred for additional 6 h. Then the solvent was distilled under reduced pressure, and the residue was suspended in water (5 mL), and extracted 3 times with EtOAc (20 mL). The combined organic solvent was removed at reduced pressure, and the crude peptides were isolated by RP HPLC on a semipreparative C18 column.

Removal of benzyl protecting groups was performed by catalytic hydrogenation. A stirred suspension of the protected peptides (0.1 mmol) and a catalytic amount of 10 % (w/w) Pd/C in absolute EtOH (10 mL) was stirred under H_2 atmosphere at RT for 12 h. Thereafter, the catalyst was filtered off over Celite and the solvent was distilled under reduced pressure, to afford the products in quantitative yield. The purity and the identity of the final products were determined to be ≥ 95 % by RP HPLC coupled to ESI MS, by ^{1}H and ^{13}C NMR, and by 2D gCOSY experiments at 600 MHz in DMSO-d_6.

Molecular modeling. Plausible structures of the cyclopeptides were obtained by simulated annealing using the AMBER force field in a 30 Å × 30 Å × 30 Å box of standard TIP3P models of equilibrated water, periodic boundary conditions dielectric scale factor = 1, and cutoff for the nonbonded interactions = 12 Å; all water molecules closer than 2.3 Å to a solute atom were eliminated. 100 random structures were generated by a 100 ps simulation at 1200 K; then the system was cooled in 20 ps to 50 K. The resulting structures were minimized by 3000 cycles of steepest descent and 3000 cycles of conjugated gradient, convergence = 0.01 kcal $Å^{-1}$ mol^{-1}. The backbones of the structures were clustered by the rmsd analysis.

For the Molecular Dynamics simulations of **TH2**-Y^{9}NLLKEEQTPQ[19], the cyclopeptide **TH2** was initially superimposed to the G^{295}QNG^{298} turn, then the C-terminal sequence was removed. The system was subjected to a 1000 ps molecular dynamics simulations keeping the Y^{9}NLLKEEQTPQ^{19} sequence fully restrained, force constant = 7 kcal mol^{-1} Å^{-2}, for distances, 16 kcal mol^{-1} Å^{-2} for dihedral angles. The simulation was conducted at 300 K and 1 atm by using the AMBER force field in a 60 Å × 60 Å × 60 Å box of standard TIP3P models of equilibrated water, periodic boundary conditions dielectric scale factor = 1, and cutoff for the nonbonded interactions = 12 Å; all water molecules closer than 2.3 Å to a solute atom were eliminated. Then the system was subjected to a 1000 ps simulation with the peptide restrained by a 50% scaled force field, followed by a 1000 ps unrestrained simulation. The resulting structure was minimized by 3000 cycles of steepest descent and 3000 cycles of conjugated gradient, and convergence = 0.01 kcal Å^{-1} mol^{-1}.

Analytical characterization of peptides.

H-Gly-Gln-Asn-Gly-Ile-Ser-Asp-Leu-OH (**TH 1**). **^{1}H NMR** (600 MHz, DMSO-d_6) δ 8.56 (d, J = 7.8 Hz, 1H, GlnNH), 8.37 (d, J = 7.7 Hz, 1H, AsnNH), 8.16 (d, J = 8.2 Hz, 1H, AspNH), 8.03 (d, J = 7.6 Hz, 1H, SerNH), 7.99 (t, J = 6.0 Hz, 1H, GlyNH), 7.94 (br s, 2H, GlyNH$_2$), 7.87 (d, J = 8.0 Hz, 1H, LeuNH), 7.85 (d, J = 8.2 Hz, 1H, IleNH),7.38 (br s, 1H, Asn-CONH$_2$), 7.23 (br s, 1H, Gln-CONH$_2$), 6.91 (br s, 1H, Asn-CONH$_2$), 6.77 (br s, 1H, Gln-CONH$_2$), 5.10 (t, J = 6.0 Hz, 1H, SerOH), 4.62 (td, J = 7.8, 5.4 Hz, 1H, AspHα), 4.57-4.52 (m, 1H, AsnHα), 4.36 (td, J = 7.8, 5.7 Hz, 1H, GlnHα), 4.32 (dt, J = 7.7, 6.1 Hz, 1H, SerHα), 4.23 (dd, J = 8.8, 7.4 Hz, 1H, IleHα), 4.20-4.15 (m, 1H, LeuHα), 3.78 (dd, J = 16.9, 5.8 Hz, 1H, GlyHα), 3.71 (dd, J = 16.9, 5.8 Hz, 1H, GlyHα), 3.64-3.56 (m, GlyHα+SerHβ), 3.56-3.50 (m, 1H, SerHβ), 2.68 (dd, J = 16.7, 5.4 Hz, 1H, AspHβ), 2.60-2.51 (m, 2H, AsnHβ+AspHβ), 2.43 (dd, J = 15.5, 7.7 Hz, 1H, AsnHβ), 2.10 (t, J = 8.0 Hz, 2H, GlnHγ), 1.88 (dt, J = 13.8, 7.5 Hz, 1H, GlnHβ), 1.78 – 1.70 (m, 2H, GlnHβ+IleHβ), 1.65-1.59 (m, 1H, LeuHγ), 1.53-1.46 (m, 2H, LeuHβ), 1.43-1.37 (m, 1H, IleHγ), 1.08-1.03 (m, 1H, IleHγ), 0.88 (d, J = 6.6 Hz, 3H, LeuHδ-CH$_3$), 0.82 (d, J = 6.6 Hz, 6H, LeuHδ-CH$_3$+IleHγ-CH$_3$), 0.80 (t, J = 7.8 Hz, 3H, IleHδ-CH$_3$). **^{13}C NMR** (151 MHz, DMSO-d_6) δ 173.70, 173.64, 171.77, 171.44, 171.05, 170.95, 170.70, 170.45, 169.82, 168.56, 165.85, 61.88, 56.78, 54.79, 52.21, 50.41, 49.81, 49.30, 42.11, 37.06, 36.66, 35.96, 31.15, 28.29, 27.15, 24.27, 24.11, 22.85, 21.37, 15.34, 10.99. **ESI MS** m/z calcd. for $[C_{32}H_{55}N_{10}O_{14}]^+$ 803.4, found 803.3 $[M+H]^+$.

H-Gly-Gln-Asn-Asp-OH (**TH 16**). **^{1}H NMR** (600 MHz, DMSO-d_6) δ 8.56 (d, J = 8.0 Hz, 1H, GlnNH), 8.33 (d, J = 7.8 Hz, 1H, AsnNH), 8.01 (d, J = 8.0 Hz, 1H, AspNH), 7.35 (br s, 1H, Asn-CONH$_2$), 7.24 (br s, 1H, Gln-CONH$_2$), 6.89 (br s, 1H, Asn-CONH$_2$), 6.79 (br s, 1H, Gln-CONH$_2$), 4.58 (qd, J = 7.9, 4.9 Hz, 1H, AsnHα), 4.45 (q, J = 6.9 Hz, 1H, AspHα), 4.37 (td, J = 7.9, 5.6 Hz, 1H, GlnHα), 3.59 (d, J = 9.4 Hz, 2H, GlyHα), 2.62-2.54 (m, 3H, AspHβ+AsnHβ), 2.42-2.35 (m, 1H, AsnHβ), 2.11 (t, J = 8.0 Hz, 2H, GlnHγ), 1.89 (dq, J = 14.1, 7.2 Hz, 1H, GlnHβ), 1.75 (dt, J = 14.0, 7.6 Hz, 1H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-d_6) δ 173.81, 172.48, 171.90, 171.27, 170.65, 170.63, 165.81, 52.11, 49.58, 48.70, 43.72, 37.10, 34.32, 31.21, 28.45. **ESI MS** m/z calcd. for $[C_{15}H_{25}N_6O_9]^+$ 433.2, found 432.9 $[M+H]^+$.

H-Gly-Gln-Asn-D-Asp-OH (**TH 17**). **^{1}H NMR** (600 MHz, DMSO-d_6) δ 8.56 (d, J = 8.2 Hz, 1H, GlnNH), 8.33 (d, J = 7.9 Hz, 1H, AsnNH), 8.06 (d, J = 8.0 Hz, 1H, AspNH), 7.98 (br t, J = 5.7 Hz, 2H, GlyNH$_2$), 7.33 (br s, 1H, Asn-CONH$_2$), 7.18 (br s, 1H, Gln-CONH$_2$), 6.88 (br s, 1H, Asn-CONH$_2$), 6.79 (br s, 1H, Gln-CONH$_2$), 4.61 (td, J = 8.1, 5.4 Hz, 1H, AsnHα), 4.51 (dt, J = 8.1, 6.0 Hz, 1H, AspHα), 4.39 (td, J = 8.1, 5.5 Hz, 1H, GlnHα), 3.59 (br d, J = 5.1 Hz, 2H, GlyHα), 2.68 (dd, J = 16.7, 5.9 Hz, 1H, AspHβ), 2.60 (dd, J = 16.6, 6.1 Hz, 1H, AspHβ), 2.53 – 2.51 (m, 1H, AsnHβ), 2.35 (dd, J = 15.5, 8.2 Hz, 1H, AsnHβ), 2.10 (t, J = 8.0 Hz, 2H, GlnHγ), 1.89 (dtd, J = 13.5, 7.9, 5.6 Hz, 1H, GlnHβ), 1.74 (dq, J = 13.6, 8.0 Hz, 1H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-d_6) δ 173.73, 172.13, 171.73, 171.12, 170.61, 170.58, 165.73, 51.99, 49.47, 48.58, 43.73, 37.22, 36.01, 31.20, 28.51. **ESI MS** m/z calcd. for $[C_{15}H_{25}N_6O_9]^+$ 433.2, found 432.9 $[M+H]^+$.

H-Gly-Gln-Asn-D-Asp-OBn (**TH 7**). **^{1}H NMR** (600 MHz, DMSO-*d*$_6$) *δ* 8.54 (d, *J* = 8.2 Hz, 1H, GlnNH), 8.32 (d, *J* = 7.9 Hz, 1H, AsnNH), 8.29 (d, *J* = 8.0 Hz, 1H, AspNH), 7.40-7.30 (m, 6H, Bn-aromatic + Asn-$CONH_2$), 7.20 (br s, 1H, Gln-$CONH_2$), 6.89 (br s, 1H, Asn-$CONH_2$), 6.80 (br s, 1H, Gln-$CONH_2$), 5.13 (d, *J* = 12.6 Hz, 1H, Bn-CH_2), 5.09 (d, *J* = 12.6 Hz, 1H, Bn-CH_2), 4.67-4.58 (m, 2H, AspHα+AsnHα), 4.37 (td, *J* = 8.2, 5.5 Hz, 1H, GlnHα), 3.59 (s, 2H, GlyHα), 2.73 (dd, *J* = 16.9, 6.1 Hz, 1H, AspHβ), 2.64 (dd, *J* = 16.8, 6.2 Hz, 1H, AspHβ), 2.49-2.46 (m, 1H, AsnHβ), 2.37 (dd, *J* = 15.5, 8.3 Hz, 1H, AsnHβ), 2.10 (t, *J* = 8.0 Hz, 2H, GlnHγ), 1.93-1.85 (m, 1H, GlnHβ), 1.74 (dq, *J* = 13.4, 8.0 Hz, 1H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-*d*$_6$) *δ* 173.69, 171.67, 171.05, 170.86, 170.61, 170.54, 165.81, 135.84, 128.41, 127.94, 127.57, 66.15, 52.07, 49.53, 48.74, 43.72, 37.20, 35.91, 31.24, 28.49. **ESI MS** m/z calcd. for $[C_{22}H_{31}N_6O_9]^+$ 523.2, found 523.2 $[M+H]^+$.

c[Gly-Gln-Asn-isoAsp] (**TH 2**). **^{1}H NMR** (600 MHz, DMSO-*d*$_6$) *δ* 8.40 (t, *J* = 6.0 Hz, 1H, GlyNH), 8.06 (d, *J* = 7.7 Hz, 1H, AspNH), 7.56 (d, *J* = 8.6 Hz, 1H, GlnNH), 7.31 (br s, 1H, Asn-$CONH_2$), 7.22 (br s, 1H, Gln-$CONH_2$), 6.90 (d, *J* = 8.9 Hz, 1H, AsnNH), 6.76 (br s, 1H, Gln-$CONH_2$), 6.74 (br s, 1H, Asn-$CONH_2$), 4.57 (dt, *J* = 8.9, 5.3 Hz, 1H, AsnHα), 4.40 (td, *J* = 7.8, 5.6 Hz, 1H, AspHα), 4.09 (td, *J* = 8.6, 6.7 Hz, 1H, GlnHα), 3.66 (dd, *J* = 14.6, 6.3 Hz, 1H, GlyHα), 3.48 (dd, *J* = 15.0, 6.0 Hz, 1H, GlyHα), 2.81-2.72 (m, 2H, AsnHβ+AspHβ), 2.56 (dd, *J* = 15.5, 5.1 Hz, 1H, AsnHβ), 2.30 (dd, *J* = 15.4, 5.7 Hz, 1H, AspHβ), 2.04 (dt, *J* = 9.0, 6.0 Hz, 2H, GlnHγ), 1.79-1.66 (m, 2H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-*d*$_6$) *δ* 173.30, 172.28, 171.81, 170.93, 170.36, 169.45, 169.32, 54.04, 49.58, 48.47, 44.17, 36.87, 35.45, 31.39, 27.33. **ESI MS** m/z calcd. for $[C_{15}H_{23}N_6O_8]^+$ 415.4, found 415.4 $[M+H]^+$.

c[Gly-Gln-Asn-isoAsp(OBn)] (**TH 10**). **^{1}H NMR** (600 MHz, DMSO-*d*$_6$) *δ* 8.45 (t, *J* = 6.1 Hz, 1H, GlyNH), 8.04 (d, *J* = 7.7 Hz, 1H, AspNH), 7.62 (d, *J* = 8.6 Hz, 1H, GlnNH), 7.38 – 7.30 (m, 6H, Bn-aromatic+Asn-$CONH_2$), 7.23 (br s, 1H, Gln-$CONH_2$), 7.14 (d, *J* = 8.9 Hz, 1H, AsnNH), 6.77 (br s, 1H, Asn-$CONH_2$), 6.74 (br s, 1H, Gln-$CONH_2$), 5.13 (s, 2H, Bn-CH_2), 4.76 (dt, *J* = 8.9, 5.5 Hz, 1H, AsnHα), 4.40 (td, *J* = 7.8, 5.8 Hz, 1H, AspHα), 4.09 (td, *J* = 8.7, 6.7 Hz, 1H, GlnHα), 3.72 (dd, *J* = 14.5, 6.5 Hz, 1H, GlyHα), 3.48-3.46 (m, 1H, GlyHα), 2.80 (dd, *J* = 15.5, 6.1 Hz, 1H, AsnHβ), 2.74 (dd, *J* = 15.5, 7.9 Hz, 1H, AspHβ), 2.61 (dd, *J* = 15.5, 4.9 Hz, 1H, AsnHβ), 2.32 (dd, *J* = 15.4, 5.8 Hz, 1H, AspHβ), 2.09 – 2.00 (m, 2H, GlnHγ), 1.77 (ddt, *J* = 13.6, 9.1, 6.7 Hz, 1H, GlnHβ), 1.70 (ddt, *J* = 15.3, 8.9, 4.3 Hz, 1H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-*d*$_6$) *δ* 173.32, 171.80, 171.00, 170.72, 170.18, 169.56, 169.42, 135.90, 128.41, 127.94, 127.59, 64.91, 54.02, 49.71, 48.83, 44.18, 36.92, 35.40, 31.39, 27.27. **ESI MS** m/z calcd. for $[C_{22}H_{29}N_6O_8]^+$ 505.2, found 505.2 $[M+H]^+$.

c[Gly-Gln-Asn-Asp(OBn)] (**TH 15**). **^{1}H NMR** (600 MHz, DMSO-*d*$_6$) *δ* 8.48 (t, *J* = 6.2 Hz, 1H, GlyNH), 8.06 (d, *J* = 7.7 Hz, 1H, AspNH), 7.65 (d, *J* = 8.6 Hz, 1H, GlnNH), 7.38-7.30 (m, 6H, Bn-aromatic+Asn-$CONH_2$), 7.22 (br s, 1H, Gln-$CONH_2$), 7.13 (d, *J* = 8.8 Hz, 1H, AsnNH), 6.76 (br s, 1H, Asn-$CONH_2$), 6.74 (br s, 1H, Gln-$CONH_2$), 5.13 (s, 2H, Bn-CH_2), 4.75 (dt, *J* = 9.0, 5.5 Hz, 1H, AsnHα), 4.40 (td, *J* = 7.8, 5.7 Hz, 1H, AspHα), 4.09 (td, *J* = 8.6, 6.7 Hz, 1H, GlnHα), 3.71 (dd, *J* = 14.5, 6.5 Hz, 1H, GlyHα), 3.46 (dd, *J* = 14.4, 6.0 Hz, 1H, GlyHα), 2.80 (dd, *J* = 15.5, 6.1 Hz, 1H, AsnHβ), 2.74 (dd, *J* = 15.6, 7.8 Hz, 1H, AspHβ), 2.63-2.60 (m, 1H, AsnHβ), 2.32 (dd, *J* = 15.6, 6.0 Hz, 1H, AspHβ), 2.04 (dt, *J* = 8.9, 6.0 Hz, 2H, GlnHγ), 1.79-1.67 (m, 2H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-*d*$_6$) *δ* 173.32, 171.80, 171.01, 170.71, 170.17, 169.57, 169.43, 135.91, 128.41, 127.94, 127.59, 66.12, 54.03, 49.69, 48.83, 44.18, 36.92, 35.42, 31.39, 27.27. **ESI MS** m/z calcd. for $[C_{22}H_{29}N_6O_8]^+$ 505.2, found 505.2 $[M+H]^+$.

c[Gly-Gln-Asn-D-isoAsp] (**TH 3**). **^{1}H NMR** (600 MHz, DMSO-*d*$_6$) *δ* 8.51 (t, *J* = 6.1 Hz, 1H, GlyNH), 7.70 -7.63 (m, 2H, AspNH+AsnNH), 7.61 (d, *J* = 9.6 Hz, 1H, GlnNH), 7.44 (br s, 1H, Asn-$CONH_2$), 7.25 (br s, 1H, Gln-$CONH_2$), 6.91 (br s, 1H, Asn-$CONH_2$), 6.76 (br s, 1H, Gln-$CONH_2$), 4.51 (td, *J* = 9.6, 4.8 Hz, 1H, AsnHα), 4.28

(ddd, *J* = 11.8, 7.4, 4.2 Hz, 1H, AspHα), 4.18 (td, *J* = 9.4, 7.0 Hz, 1H, GlnHα), 3.67 (dd, *J* = 15.8, 6.0 Hz, 1H, GlyHα), 3.55 (dd, *J* = 15.8, 6.1 Hz, 1H, GlyHα), 2.63 (dd, *J* = 14.2, 4.2 Hz, 1H, AspHβ), 2.55-2.51 (m, 1H, AspHβ), 2.49-2.46 (m, 1H, AsnHβ), 2.33 (dd, *J* = 14.8, 9.8 Hz, 1H, AsnHβ), 2.04 (dt, *J* = 8.6, 6.0 Hz, 2H, GlnHγ), 1.84-1.73 (m, 2H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-d_6) *δ* 173.27, 172.61, 171.60, 170.69, 169.64, 169.18, 169.14, 54.26, 50.82, 50.05, 43.47, 37.25, 36.26, 31.21, 26.73. **ESI MS** m/z calcd. for $[C_{15}H_{23}N_6O_8]^+$ 415.4, found 415.4 $[M+H]^+$.

H-Gly-Gln-Asn-Asp(OBn)-OH (**TH 14**). **^{1}H NMR** (600 MHz, DMSO-d_6) *δ* 8.54 (d, *J* = 8.0 Hz, 1H, GlnNH), 8.29 (d, *J* = 7.9 Hz, 1H, AsnNH), 8.23 (d, *J* = 8.0 Hz, 1H, AspNH), 7.40-7.27 (m, 6H, Bn-aromatic+Asn-CONH$_2$), 7.23 (br s, 1H, Gln-CONH$_2$), 6.89 (br s, 1H, Asn-CONH$_2$), 6.79 (br s, 1H, Gln-CONH$_2$), 5.10 (br d, *J* = 2.3 Hz, 2H, Bn-CH$_2$), 4.64 (q, *J* = 6.5 Hz, 1H, AspHα), 4.59 (td, *J* = 8.1, 5.1 Hz, 1H, AsnHα), 4.33 (q, *J* = 7.4 Hz, 1H, GlnHα), 3.57 (s, 2H, GlyHα), 2.65 (d, *J* = 6.0 Hz, 2H, AspHβ), 2.55-2.52 (m, 1H, AsnHβ), 2.38 (dd, *J* = 15.4, 8.6 Hz, 1H, AsnHβ), 2.10 (t, *J* = 8.0 Hz, 2H, GlnHγ), 1.88 (dq, *J* = 14.3, 7.6 Hz, 1H, GlnHβ), 1.74 (dt, *J* = 14.1, 7.8 Hz, 1H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-d_6) *δ* 173.74, 171.81, 171.18, 170.96, 170.75, 170.65, 165.94, 135.89, 128.42, 127.95, 127.60, 66.11, 54.92, 52.16, 49.47, 48.82, 37.07, 36.09, 31.22, 28.39. **ESI MS** m/z calcd. for $[C_{22}H_{31}N_6O_9]^+$ 523.2, found 523.2 $[M+H]^+$.

c[Gly-Gln-Asn-D-isoAsp(OBn)] (**TH 11**). **^{1}H NMR** (600 MHz, DMSO-d_6) *δ* 8.53 (t, *J* = 6.1 Hz, 1H, GlyNH), 7.83 (d, *J* = 7.2 Hz, 1H, AspNH), 7.72 – 7.60 (m, 2H, AsnNH+GlnNH), 7.45 (br s, 1H, Asn-CONH$_2$), 7.39-7.24 (m, 6H, Bn-aromatic+Gln-CONH$_2$), 6.92 (br s, 1H, Asn-CONH$_2$), 6.76 (br s, 1H, Gln-CONH$_2$), 5.15 (d, *J* = 12.6 Hz, 1H, Bn-CH$_2$), 5.12 (d, *J* = 12.6 Hz, 1H, Bn-CH$_2$), 4.52 (td, *J* = 9.4, 4.9 Hz, 1H, AsnHα), 4.41 (ddd, *J* = 11.7, 7.2, 4.0 Hz, 1H, AspHα), 4.18 (q, *J* = 8.6 Hz, 1H, GlnHα), 3.68 (dd, *J* = 15.8, 6.1 Hz, 1H, GlyHα), 3.53 (dd, *J* = 15.9, 6.2 Hz, 1H, GlyHα), 2.67 (dd, *J* = 14.2, 4.1 Hz, 1H, AspHβ), 2.59-2.54 (m, 1H, AspHβ), 2.48-2.45 (m, 1H, AsnHβ), 2.31 (dd, *J* = 14.8, 9.5 Hz, 1H, AsnHβ), 2.03 (q, *J* = 6.9 Hz, 2H, GlnHγ), 1.77 (q, *J* = 7.9 Hz, 2H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-d_6) *δ* 173.25, 171.76, 170.91, 170.72, 169.71, 169.13, 168.83, 135.83, 128.45, 128.06, 127.77, 66.17, 54.23, 50.78, 50.27, 43.52, 37.13, 36.06, 31.20, 26.71. **ESI MS** m/z calcd. for $[C_{22}H_{29}N_6O_8]^+$ 505.2, found 505.2 $[M+H]^+$.

H-Gly-Gln-Asn-Asp-OBn (**TH 6**). **^{1}H NMR** (600 MHz, DMSO-d_6) *δ* 8.54 (d, *J* = 8.1 Hz, 1H, GlnNH), 8.33-8.25 (m, 2H, AsnNH+AspNH), 7.42-7.27 (m, 6H, Bn-aromatic+Asn-CONH$_2$), 7.22 (br s, 1H, Gln-CONH$_2$), 6.90 (br s, 1H, Asn-CONH$_2$), 6.80 (br s, 1H, Gln-CONH$_2$), 5.11 (d, *J* = 12.6 Hz, 1H, Bn-CH$_2$), 5.09 (d, *J* = 12.6 Hz, 1H, Bn-CH$_2$),, 4.65 (dt, *J* = 8.1, 6.2 Hz, 1H, AspHα), 4.60 (td, *J* = 8.1, 4.9 Hz, 1H, AsnHα), 4.36 (td, *J* = 8.0, 5.6 Hz, 1H, GlnHα), 3.58 (s, 2H, GlyHα), 2.67 (qd, *J* = 16.8, 6.1 Hz, 2H, AspHβ), 2.54-2.51 (m, 1H, AsnHβ), 2.38 (dd, *J* = 15.7, 8.5 Hz, 1H, AsnHβ), 2.10 (t, *J* = 8.0 Hz, 2H, GlnHγ), 1.88 (dtd, *J* = 16.1, 7.6, 5.7 Hz, 1H, GlnHβ), 1.73 (dq, *J* = 13.4, 8.0 Hz, 1H, GlnHβ). **^{13}C NMR** (151 MHz, DMSO-d_6) *δ* 173.73, 171.63, 171.08, 171.02, 170.66, 170.59, 165.80, 135.86, 128.41, 127.96, 127.61, 66.14, 52.02, 49.43, 48.75, 40.06, 37.11, 35.89, 31.21, 28.50. **ESI MS** m/z calcd. for $[C_{22}H_{31}N_6O_9]^+$ 523.2, found 523.2 $[M+H]^+$.

H-Lys-Ser-Asp-D-Leu-OH (**TH 19**). **^{1}H NMR** (600 MHz, DMSO-d_6) *δ* 8.66 (d, *J* = 7.9 Hz, 1H, SerNH), 8.45 (d, *J* = 7.7 Hz, 1H, AspNH), 8.19 (d, *J* = 5.0 Hz, 2H, Lys-αNH$_2$), 7.96 (d, *J* = 8.2 Hz, 1H, LeuNH), 7.82 (t, *J* = 5.8 Hz, 2H, Lys-εNH$_2$), 4.60 (td, *J* = 8.0, 4.9 Hz, 1H, AspHα), 4.44 (q, *J* = 6.6 Hz, 1H, SerHα), 4.21 (ddd, *J* = 10.3, 8.1, 4.6 Hz, 1H, LeuHα), 3.85 (q, *J* = 5.9 Hz, 1H, LysHα), 3.69 (dd, *J* = 10.5, 6.0 Hz, 1H, SerHβ), 3.50 (dd, *J* = 10.5, 6.8 Hz, 1H, SerHβ), 2.80 – 2.75 (m, 2H, LysHε), 2.72-2.67 (m, 1H, AspHβ), 2.55 (dd, *J* = 16.7, 8.3 Hz, 1H, AspHβ), 1.71 (dt, *J* = 10.9, 6.6 Hz, 2H, LysHβ), 1.64-1.51 (m, 4H, LysHδ+LeuHγ+ LeuHβ), 1.48-1.43 (m, 1H, LeuHβ), 1.38-1.32 (m, 2H, LysHγ), 0.85 (d, *J* = 6.3 Hz, 3H, LeuHδ-CH$_3$), 0.80 (d, *J* = 6.4 Hz, 3H, LeuHδ-CH$_3$).

^{13}C NMR (151 MHz, DMSO-*d*$_6$) *δ* 173.87, 171.73, 170.42, 169.81, 168.64, 62.04, 54.68, 52.00, 50.43, 49.82, 38.61, 36.22, 30.59, 26.57, 24.22, 22.96, 21.22, 21.20. **ESI MS** m/z calcd. for $[C_{19}H_{36}N_5O_8]^+$ 462.2, found 462.0 $[M+H]^+$.

H-Lys-Ser-Asp-Leu-OH (**TH 18**). **^{1}H NMR** (600 MHz, DMSO-*d*$_6$) *δ* 8.65 (d, *J* = 7.8 Hz, 1H, SerNH), 8.38 (d, *J* = 8.1 Hz, 1H, AspNH), 8.16 (d, *J* = 5.4 Hz, 2H, Lys-αNH$_2$), 7.95 (d, *J* = 8.0 Hz, 1H, LeuNH), 7.78 (t, *J* = 5.9 Hz, 2H, Lys-εNH$_2$), 4.64 (td, *J* = 8.1, 5.0 Hz, 1H, AspHα), 4.42 (dt, *J* = 8.0, 6.2 Hz, 1H, SerHα), 4.18 (ddd, *J* = 9.9, 8.0, 5.1 Hz, 1H, LeuHα), 3.86-3.83 (m, 1H, LysHα), 3.66 (dd, *J* = 10.7, 6.1 Hz, 1H, SerHβ), 3.57 (dd, *J* = 10.7, 6.3 Hz, 1H, SerHβ), 2.76 (dt, *J* = 10.2, 5.9 Hz, 2H, LysHε), 2.70 (dd, *J* = 16.8, 5.0 Hz, 1H, AspHβ), 2.57-2.51 (m, 1H, AspHβ), 1.73-1.62 (m, 2H, LysHβ), 1.60-1.44 (m, 5H, LysHδ+LeuHγ+ LeuHβ), 1.41-1.30 (m, 2H, LysHγ), 0.88 (d, *J* = 6.6 Hz, 3H, LeuHδ-CH$_3$), 0.83 (d, *J* = 6.5 Hz, 3H, LeuHδ-CH$_3$). **^{13}C NMR** (151 MHz, DMSO-*d*$_6$) *δ* 173.72, 171.69, 170.52, 169.53, 168.49, 64.94, 54.75, 51.89, 50.45, 49.38, 38.52, 35.98, 30.53, 26.49, 24.18, 22.88, 21.35, 21.10. **ESI MS** m/z calcd. for $[C_{19}H_{36}N_5O_8]^+$ 462.2, found 462.0 $[M+H]^+$.

c[isoLys-Ser-Asp-D-Leu] (**TH 5**). **^{1}H NMR** (600 MHz, DMSO-*d*$_6$) *δ* 8.75 (br s, 1H, SerNH), 8.20 (d, *J* = 7.9 Hz, 1H, AspNH), 7.62 (d, *J* = 9.3 Hz, 1H, LeuNH), 7.56 (br d, *J* = 6.9 Hz, 1H, Lys-εNH), 4.36–4.17 (m, 3H, SerHα+AspHα+LeuHα), 3.72 (dd, *J* = 11.3, 5.5 Hz, 1H, SerHβ), 3.64 (dd, *J* = 8.7, 4.0 Hz, 1H, LysHα), 3.58 (dd, *J* = 11.4, 3.9 Hz, 1H, SerHβ), 3.35 – 3.30 (m, 1H, LysHε), 2.80-2.72 (m, 1H, LysHε), 2.65 (dd, *J* = 16.2, 5.3 Hz, 1H, AspHβ), 2.40 (dd, *J* = 16.2, 5.1 Hz, 1H, AspHβ), 1.68-1.64 (m, 2H, LeuHβ+LysHβ), 1.57-1.45 (m, 4H, LysHβ+LeuHγ+LysHδ +LeuHβ), 1.40-1.25 (m, 3H, LysHδ+LysHγ), 0.86 (d, *J* = 6.5 Hz, 3H, LeuHδ-CH$_3$), 0.80 (d, *J* = 6.4 Hz, 3H, LeuHδ-CH3). **^{13}C NMR** (151 MHz, DMSO-*d*$_6$) *δ* 173.72, 172.89, 171.45, 170.83, 170.16, 64.91, 56.71, 52.94, 50.62, 50.33, 37.85, 37.71, 32.13, 26.83, 24.23, 23.32, 21.07, 20.78. **ESI MS** m/z calcd. for $[C_{19}H_{34}N_5O_7]^+$ 444.2, found 444.1 $[M+H]^+$.

c[isoLys-Ser-Asp-Leu] (**TH 4**). **^{1}H NMR** (600 MHz, DMSO-*d*$_6$) *δ* 8.51 (d, *J* = 7.6 Hz, 1H, LeuNH), 8.36 (br s, 1H, SerNH), 8.32 (br s, 1H, AspNH), 8.04 (t, *J* = 5.6 Hz, 1H, Lys-εNH), 4.26-4.16 (m, 2H, SerHα+AspHα), 3.99 (ddd, *J* = 11.2, 7.5, 4.6 Hz, 1H, LeuHα), 3.68 (dd, *J* = 11.3, 5.6 Hz, 1H, SerHβ), 3.59 (dd, *J* = 11.5, 4.8 Hz, 1H, SerHβ), 3.25 (dd, *J* = 8.8, 3.6 Hz, 1H, LysHα), 3.20-3.12 (m, 1H, LysHε), 2.92-2.83 (m, 1H, LysHε), 2.59 (dd, *J* = 15.6, 4.9 Hz, 1H, AspHβ), 2.21 (dd, *J* = 15.5, 5.3 Hz, 1H, AspHβ), 1.66-1.61 (m, 1H, LeuHγ), 1.56-1.48 (m, 3H, LeuHβ+LysHβ), 1.41-1.31 (m, 3H, LysHβ+LysHδ), 1.26-1.13 (m, 2H, LysHγ), 0.85 (d, *J* = 6.5 Hz, 3H, LeuHδ-CH3), 0.78 (d, *J* = 6.4 Hz, 3H, LeuHδ-CH3). **^{13}C NMR** (151 MHz, DMSO-*d*$_6$) *δ* 173.35, 172.82, 171.82, 171.58, 170.13, 64.90, 56.13, 54.22, 52.50, 51.51, 38.67, 36.13, 28.44, 24.29, 23.13, 21.93, 21.07, 17.26. **ESI MS** m/z calcd. for $[C_{19}H_{34}N_5O_7]^+$ 444.2, found 444.1 $[M+H]^+$.

Note: The work described in this chapter has been published in ***Protein Science*** with slight modification and has been reproduced here with the permission of the copyright holder.

Stefan, A.; Gentilucci, L.; Ruffolo, F.; Rossi, V.; Sordi, S.; He, T.; di Stefano, G.; Santino, F.; Brigotti, M.; Scotti, C.; Iamele, L.; de Jonge, H.; Piaz, F.D.; Santarcangelo, D.R.; Hochkoeppler, A.* Peptides inhibiting the assembly of monomeric human l-lactate dehydrogenase into catalytically active homotetramer decrease the synthesis of lactate in cultured cells. *Protein Sci.* **2024**, *33* (10), e5161.

References

[1] Dennis, D.; Kaplan, N.O. D- and L-lactic acid dehydrogenases in *Lactobacillus plantarum*. *J. Biol. Chem.* **1960**, *235*, 810-818.

[2] Garvie, E.I. Bacterial lactate dehydrogenases. *Microbiol. Rev.* **1980**, *44*, 106-139.

[3] Steinbüchel, A.; Schlegel, H.G. Nad-linked L(+)-lactate dehydrogenase from the strict aerobe *Alcaligenes eutrophus*. *Eur. J. Biochem.* **1983**, *130*, 321-328.

[4] Arai, K.; Ishimitsu, T.; Fushinobu, S.; Uchikoba, H.; Matsuzawa, H.; Taguchi, H. Active and inactive state structures of unliganded *Lactobacillus casei* allosteric L-lactate dehydrogenase. *Proteins* **2009**, *78*, 681-694.

[5] Kaplan, N.O.; Ciotti, M.M.; Stolzenbach, F.E. Reaction of pyridine nucleotide analogues with dehydrogenases. *J. Biol. Chem.* **1956**, *221*, 833-844.

[6] Kaplan, N.O.; Ciotti, M.M.; Hamolsky, M.; Bieber, R.E. Molecular heterogeneity and evolution of enzymes. *Science* **1960**, *131*, 392-397.

[7] Selander, R.K.; Yang, S.Y. Horseshoe crab lactate dehydrogenases: evidence for dimeric structure. *Science* **1970**, *169*, 179-181.

[8] Takenaka, Y.; Schwert, G.W. Lactic dehydrogenase: III. Mechanism of the reaction. *J. Biol. Chem.* **1956**, *223*, 157-170.

[9] Cristescu, M.E.; Innes, D.J.; Stillman, J.H.; Crease, T.J. D- and L-lactate dehydrogenases during invertebrate evolution. *BMC Evol. Biol.* **2008**, *8*, 268.

[10] Eventoff, W.; Rossmann, M.G.; Taylor, S.S.; Torff, H.-J.; Meyer, H.; Keil, W.; Kiltz, H.-H. Structural adaptations of lactate dehydrogenase isozymes. *Proc. Natl. Acad. Sci. USA*, **1977**, *74*, 2677-2681.

[11] Fritz, P.J. Rabbit muscle lactate dehydrogenase 5; a regulatory enzyme. *Science* **1965**,*150*, 364-366.

[12] Iacovino, L.G.; Rossi, M.; Di Stefano, G.; Rossi, V.; Binda, C.; Brigotti, M.; Tomaselli, F.; Pasti, A.P.; Dal Piaz, F.; Cerini, S.; Hochkoeppler, A. Allosteric transitions of rabbit skeletal muscle lactate dehydrogenase induced by pH-dependent dissociation of the tetrameric enzyme. *Biochimie.* **2022**, *199*, 23-35.

[13] Pasti, A.P.; Rossi, V.; Di Stefano, G.; Brigotti, M.; Hochkoeppler, A. Human lactate dehydrogenase A undergoes allosteric transitions under pH conditions inducing the dissociation of the tetrameric enzyme. *Biosci. Rep.* **2022**, *42*, BSR20212654.

[14] Valvona, C.J.; Fillmore, H.L.; Nunn, P.B.; Pilkington, G.J. The regulation and function of lactate dehydrogenase A: therapeutic potential in brain tumor. *Brain Pathol.* **2016**, *26*, 3-17.

[15] Woodford, M.R.; Chen, V.Z.; Backe, S.J.; Bratslavsky, G.; Mollapour, M. Structural and functional regulation of lactate dehydrogenase-A in cancer. *Future Med. Chem.* **2020**, *12*, 439-455.

[16] Read, J.A.; Winter, V.J.; Eszes, C.M.; Sessions, R.B.; Brady, R.L. Structural basis for altered activity of M- and H-isozyme forms of human lactate dehydrogenase. *Proteins* **2001**, *43*, 175-185.

[17] Evrev, T.; Zhivcov, S.; Russev, L. LDH isozymes in testicular cultures and human testes. *Hum. Hered.* **1970**, *20*, 70-73.

[18] LeVan, K.M.; Goldberg, E. Properties of human testis-specific lactate dehydrogenase expressed from *Escherichia coli*. *Biochem. J.* **1991**, *273*, 587-592.

[19] De Bari, L.; Moro, L.; Passarella, S. Prostate cancer cells metabolize D-lactate inside mitochondria via a D-lactate dehydrogenase which is more active and highly expressed than in normal cells. *FEBS Lett.* **2013**, *587*, 467-473.

[20] Brooks, G.A. Lactate: glycolytic end product and oxidative substrate during sustained exercise in mammals-the "lactate shuttle". In: Gilles R Ed. Circulation, respiration, and Metabolism: Current Comparative Approaches. Springer Verlag, Berlin, **1985**, pp 208-218.

[21] Schurr, A.; Payne, R.S. Lactate, not pyruvate, is neuronal aerobic glycolysis end product: an *in vitro* electrophysiological study. *Neuroscience* **2007**, *147*, 613-619.

[22] Rogatzki, M.J.; Ferguson, B.S.; Goodwin, M.L.; Gladden, L.B. Lactate is always the end product of glycolysis. *Front. Neurosci.* **2015**, *9*, 22.

[23] Warburg, O. On the origin of cancer cells. *Science* **1956**, *123*, 309-314.

[24] Zilva, J.F. The origin of the acidosis in hyperlactatemia. *Ann. Clin. Chem.* **1978**, *15*, 40-43.

[25] White, K.A.; Grillo-Hill, B.K.; Barber, D.L. Cancer cell behaviors mediated by dysregulated pH dynamics at a glance. *J. Cell Sci.* **2017**, *130*, 663-669.

[26] Fan, J.; Hitosugi, T.; Chung, T.-W.; Xie, J.; Ge, Q.; Gu, T.-L.; Polakiewicz, R.D.; Chen, G.Z.; Boggon, T.J.; Lonial, S.; Khuri, F.R.; Kang, S.; Chen, J. Tyrosine phosphorylation of lactate dehydrogenase A is important for NADH/NAD^+ redox homeostasis in cancer cells. *Mol. Cell Biol.* **2011**, *31*, 4938-4950.

[27] Jin, L.; Chun, J.; Pan, C.; Alesi, G.N.; Li, D.; Magliocca, K.R.; Kang, Y.; Chen, Z.G.; Shin, D.M.; Khuri, F.R.; Fan, J. Phosphorylation-mediated activation of LDHA promotes cancer cell invasion and tumour metastasis. *Oncogene* **2017**, *36*, 3797-3806.

[28] Gui, D.Y.; Sullivan, L.B.; Luengo, A.; Hosiso, A.M.; Bush, L.N.; Gitego, N.; Davidson, S.M.; Freinkman, E.; Thomas, C.J.; Vander Heiden, M.G. Environment dictates dependence on mitochondrial complex I for NAD^+ and aspartate production and determines cancer cell sensitivity to metformin. *Cell Metab.* **2016**, *24*, 716-727.

[29] Granchi, C.; Paterni, I.; Rani, R.; Minutolo, F. Small-molecule inhibitors of human Ldh5. *Future Med. Chem.* **2013**, *5*, 1967-1991.

[30] Yu, Y.; Deck, J.A.; Hunsaker, L.A.; Deck, L.M.; Royer, R.E.; Goldberg, E.; Vander Jagt D.L. Selective active site inhibitors of human lactate dehydrogenases A_4, B_4, and C_4. *Biochem. Pharmacol.* **2001**, *62*, 81-89.

[31] Kotlyar, A.B.; Randazzo, A.; Honbo, N.; Jin, Z.-Q.; Karliner, J.S.; Cecchini, G. Cardioprotective activity of a novel and potent competitive inhibitor of lactate dehydrogenase. *FEBS Lett.* **2010**, *584*, 159-165.

[32] Friberg, A.; Rehwinkel, H.; Nguyen, D.; Pütter, V.; Quanz, M.; Weiske, J.; Eberspächer, U.; Heisler, I.; Langer, G. Structural evidence for isoform-selective allosteric inhibition of lactate dehydrogenase A. *ACS Omega* **2020**, *5*, 13034-13041.

[33] Hermann, R.; Jaenicke, R.; Rudolph, R. Analysis of the reconstitution of oligomeric enzymes by crosslinking with glutaraldehyde: kinetics of reassociation of lactic dehydrogenase. *Biochemistry* **1981**, *20*, 5195-5201.

[34] Wang, X.; Ni, D.; Liu, Y.; Lu, S. Rational design of peptide-based inhibitors disrupting protein-protein interactions. *Front. Chem.* **2021**, *9*, 682675.

[35] Jafary, F.; Ganjalikhany, M.R.; Moradi, A.; Hemati, M.; Jafari, S. Novel peptide inhibitors for lactate dehydrogenase A (LDHA): a survey to inhibit LDHA activity via disruption of protein-protein interaction. *Sci. Rep.* **2019**, *9*, 4686.

[36] Thabault, L.; Brisson, L.; Brustenga, C.; Martinez Gache, S.A.; Prévost, J.R.C.; Kozlova, A.; Spillier, Q.; Liberelle, M.; Benyahia, Z.; Messens, J.; Copetti, T.; Sonveaux, P.; Frédérick, R. Interrogating the lactate dehydrogenase tetramerization site using (stapled) peptides. *J. Med. Chem.* **2020**, *63*, 4628-4643.

[37] Thabault, L.; Liberelle, M.; Koruza, K.; Yildiz, E.; Joudiou, N.; Messens, J.; Brisson, L.; Wouters, J.; Sonveaux, P.; Frédérick, R. Discovery of a novel lactate dehydrogenase tetramerization domain using epitope mapping and peptides. *J. Biol. Chem.* **2021**, *296*, 100422.

[38] Nadal-Bufi, F.; Mason, J.M.; Chan, L.Y.; Craik, D.J.; Kaas, Q.; Henriques, S.T. Designed β-hairpins inhibit LDH5 oligomerization and enzymatic activity. *J. Med. Chem.* **2021**, *64*, 3767-3779.

[39] Girg, R.; Rudolph, R.; Jaenicke, R. Limited proteolysis of porcine-muscle lactic dehydrogenase by thermolysin during reconstitution yields dimers. *Eur. J. Biochem.* **1981**, *119*, 301-305.

[40] Girg, R.; Jaenicke, R.; Rudolph, R. Dimers of porcine skeletal muscle lactate dehydrogenase produced by limited proteolysis during reassociation are enzymatically active in the presence of stabilizing salt. *Biochem. Int.* **1983**, *7*, 433-441.

[41] Kolappan, S.; Shen, D.L.; Mosi, R.; Sun, J.; McEachern, E.J.; Vocadlo, D.J.; Craig, L. Structures of lactate dehydrogenase A (LDHA) in apo, ternary and inhibitor-bound forms. *Acta Crystallogr. D Biol. Crystallogr.* **2015**, *71*, 185-195.

[42] Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

[43] Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926-935.

[44] Glenn, M.P.; Kelso, M.J.; Tyndall, J.D.A.; Fairlie, D.P. Conformationally homogeneous cyclic tetrapeptides: useful new three-dimensional scaffolds. *J. Am. Chem. Soc.* **2003**, *125*, 640-641.

[45] Trausch, G. On the determination of the molecular weight of lactate dehydrogenase by gel filtration on Sephadex G-200. *Biochimie* **1972**, *54*, 531-533.

[46] Chan, W.W.C.; Mosbach, K. Effects of subunit interactions on the activity of lactate dehydrogenase studied in immobilized enzyme systems. *Biochemistry* **1976**, *15*, 4215-4222.

[47] Pettit, S.M.; Nealon, D.A.; Henderson, A.R. Purification of lactate dehydrogenase isoenzyme-5 from human liver. *Clin. Chem.* **1981**, *27*, 88-93.

[48] Gori, A.; Lodigiani, G.; Colombarolli, S.G.; Bergamaschi, G.; Vitali, A. Cell penetrating peptides: classification, mechanisms, methods of study, and applications. *ChemMedChem* **2023**, *18*, e202300236.

[49] Green, M.; Loewenstein, P.M. Autonomous functional domains of chemically synthesized human immunodeficiency virus tat *trans*-activator protein. *Cell* **1988**, *55*, 1179-1188.

[50] Joliot, A.; Pernelle, C.; Deagostini-Bazin, H.; Prochiantz, A. Antennapedia homeobox peptide regulates neural morphogenesis. *Proc. Natl. Acad. Sci. USA,* **1991**, *88*, 1864-1868.

[51] Nadal-Bufi, F.; Chan, L.Y.; Mohammad, H.H.; Mason, J.D.; Salomon, C.; Lai, A.; Thompson, E.W.; Craik, D.J.; Kaas, Q.; Henriques, S.T. Peptide-based LDH5 inhibitors enter cancer cells and impair proliferation. *Cell. Mol. Life Sci.* **2022**, *79*, 606.

[52] Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248-254.

[53] Bernofsky, C.; Wanda, S.C. Formation of reduced nicotinamide adenine dinucleotide peroxide. *J. Biol. Chem.* **1982**, *257*, 6809-6817.

[54] Schevchenko, A.; Thomas, H.; Havliš, J.; Olsen, J.V.; Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **2007**, *1*, 2856-2860.

[55] Conte, E.; Vincelli, G.; Schaaper, R.M.; Bressanin, D.; Stefan, A.; Dal Piaz, F.; Hochkoeppler, A. Stabilization of the *Escherichia coli* DNA polymerase III ε subunit by the θ subunit favors *in vivo* assembly of the Pol III catalytic core. *Arch. Biochem. Biophys.* **2012**, *523*, 135-143.

[56] Rossi, V.; Govoni, M.; Di Stefano, G. Lactate can modulate the antineoplastic effects of doxorubicin and relieve the drug's oxidative damage on cardiomyocytes. *Cancers* **2023**, *15*, 3728.

Chapter VI. Expedient Recycling of Peptide Organocatalysts Using a Nanocrystalline Hydroxyapatite Catching System

6.1 Introduction

Asymmetric organocatalysis using small to medium-sized organic molecules is complementary and alternative to the chemistry of metal complexes and enzymes, and represents nowadays a mature methodology with numerous applications in the production of enantiomerically enriched molecules, including bioactive compounds, drugs, natural products.[1-3] Besides, catalysis is one of the design principles of Green Chemistry,[4] and small organic molecules are generally considered more environmentally friendly as compared to, for instance, catalysts based on transition metals.[5]

In reality, not all contemporary molecular scaffolds employed in catalysis are attractive for industrial applications. Recent examples seem so complex to be prepared that their employment on a large scale appears impractical. Besides, their preparations can paradoxically require transition metals, toxic reagents, or harsh reaction conditions.[6,7]

In contrast, amino acids as catalysts have the advantage of being cheap and can be used without any specialized approaches to perform catalytic reactions in a highly enantioselective manner.[8] They are available in both enantiomeric forms and different derivatives can provide complimentary selectivity. Their use dates back to the 1970s, with Hajos and Wiechert reporting the use of proline for the Robinson annulation in 1970. However, the use of proline in reactions involving enamine intermediates boomed in 2000, when List et al. reported its use in the direct asymmetric aldol reaction.[9,10]

On the other hand, short peptides and peptide-like molecules proved to be more versatile as compared to simple amino acids,[8,11,12] thanks to the modularity and tunability of their peptide backbones, well-defined secondary structures and often higher selectivity and activity. In fact, apart from the biological activity,[13] oligopeptides have found applications in a range of asymmetric reactions. For example, in the late 1970s, Inoue used diketopiperazines in the asymmetric Strecker reaction.[14] In the early 80s, Juliá and Colonna reported the epoxidation of chalcone catalyzed by poly-L-alanine, yielding epoxides with high enantiomeric excess (>90% e.e.). The result indicates that the use of peptides as “synthetic enzymes” can be a good option for organocatalysis.[15] Since then, a variety of peptide-catalyzed reactions have been described, including aldol condensations,[16] conjugate/Michael addition,[17-19] conjugate addition with *N*-heterocyclic substituents,[20,21] Mannich reaction,[22] epoxidation,[23] desymmetrization of diphenols or diols by acylation[24,25] or by silylation,[26] atroposelective bromination,[27] site-selective polyene oxidation,[28] enantioselective [2+2] photocycloaddition,[29] as far as the light-driven deracemization of cyclic ureas via excited-state electron transfer, a multi-step process consisting of a series of electron, proton, and H-atom transfer in the presence of three distinct molecular catalysts, one of which is a cysteine-containing H-atom donor peptide.[30]

Besides, peptides and peptide-lipid conjugates can self-assemble into nanostructures such as fibrils, nanotubes, micelles, etc. and these supramolecules can be used as platforms to present functional residues in order to catalyze a diversity of reactions.[31,32] Finally, peptides have been also utilized as scaffolds for the preparation of metal-peptide hybrid catalysts.[33]

Unfortunately, despite recent effort to improve the environmental sustainability,[34,35] the standard protocols for peptide synthesis[36,37] rise much concern in terms of atom economy, economic cost, and production of wastes,[38] especially toxic organic solvents.[39] Consequently, the recovery and recycle of peptides are highly attractive in a Green Chemistry perspective. To improve the greenness of catalysis by amino acids or peptides, immobilization on insoluble solid supports is expected to allows easy recovery and reuse (Figure 6.1 A).[40-43] Among the preferred materials there are cross-linked hydrophobic polystyrene, hydrophilic TentaGel, polyethylene glycol (PEG)-polyacrylamide, mesoporous silica, chitosan, etc. The immobilized sequence can be directly reused after e.g. simple filtration from the reaction mixture. For catalysts supported onto magnetic nanoparticles, recovery is done by magnetic separation. Dendrimer-supported catalysts can be easily separated from the reaction mixture by precipitation, membrane or nanofiltration methods.[10,41] This strategy consents in principle a scale-up of reactions or, instead, the use in continuous flow system.[44,45]

However, while the immobilization of proline derivatives appears more direct to plan (for selected examples from the recent literature, see Refs[10,46-49]), the success in the immobilization of a peptide can be challenging. Indeed, the modification can prevent the peptide from adopting the correct secondary structure as required for promoting the expected reaction, mainly because the solid support can interfere with the catalyst. Also, the deactivation of the organocatalyst can be a major issue, as it prevents its reuse by limiting the number of catalytic cycles. Deactivation can involve chemical modification or blockage of the catalyst’s active sites.

As an alternative to immobilization, the amino acid organocatalyst can be equipped with a perfluorinated tag, which consents recovery by fluorous extraction,[50] or it can be utilized in an environment of alternative media, e.g. ionic liquids, deep eutectic solvents, PEG, supercritical CO_2, etc. allowing a significant improvement of stability and recycling of the catalyst.[51] However, these alternatives have been applied to peptides very seldom.[1]

Another opportunity to prepare non-covalent heterogeneous catalysts based on amino acids/peptides is represented by their inclusion or adsorption on appropriate supports. Szőllősi et al. prepared chiral solid materials by adsorption of Pro or other amino acids on the solid surface of inorganic oxides. The catalysts could be reutilized several times in asymmetric Micheal additions.[52,53] In particular, the cation-exchanger layered magnesia-silicate laponite and the aluminum silicate bentonite gave the best results in terms of e.e., with inversion of the enantioselectivity as compared to the homogeneous reaction.

Juaristi and co-workers utilized a mechanochemical approach to prepare hybrid catalysts based on the intercalation of α-amino acids or α,β-dipeptides in reconstructed Mg/Al-hydrotalcite. This material was designed with the expectation that reactions could provide an augmented catalytic effect in the confined space rendered by the layered material. Remarkably, excellent results were obtained in the addition of branched aldehydes to maleimides even under solvent free conditions, and the hybrid catalysts could be recycled and reused also in large-scale experiments.[54]

In this context, we conceived to utilize hydroxyapatite crystals (HAp) as a peptide-catching system for its expedient capture and reuse from the reaction mixture, without the need of immobilization by covalent binding (Figure 6.1 B). We came up with this hypothesis during our recent studies on the synthesis of protected oligopeptides under Green Chemistry conditions, using nanocrystalline HAp as solid, bio-compatible, reusable inorganic base in minimal amounts of solvents.[55,56] Hydroxyapatite belongs to the apatites group with the formula $Ca_{10}(PO_4)_6(OH)_2$ that crystallizes in a hexagonal system. It displays the presence of surface Ca^{2+} and PO_4^{3-} ions that serve as effective adsorption sites, and the adsorption is facilitated by the ionic nature of the HAp surface. Moreover, HAp represents a convenient material in a Green Chemistry perspective. HAp is the primary inorganic component of bones and teeth; hence, it is a biocompatible, biodegradable, bioactive, nonimmunogenic, noninflammatory material. HAp also shows a relatively mild basic character that may prevent unwanted reactions. It can be utilized as a base to promote condensations of aldehydes (Knoevenagel, Claisen-Schmidt), but only with active methylene compounds.[57]

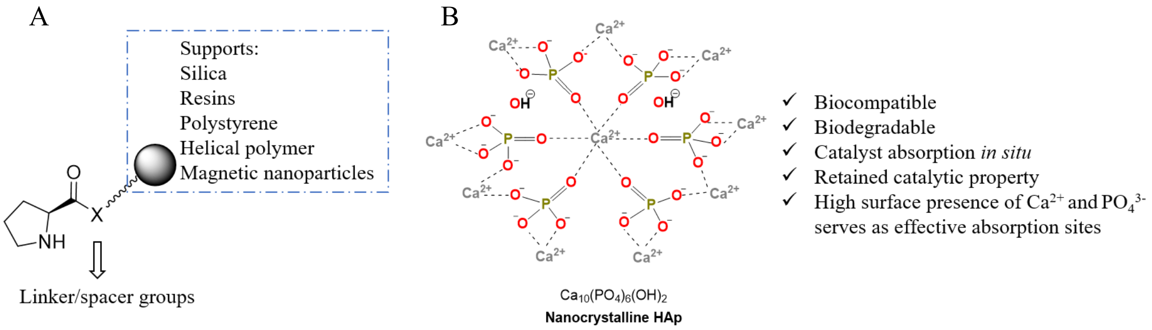


Figure 6.1 (A) Conventional strategies for covalent immobilization of Proline. (B) Structure of nanocrystalline HAp and its advantages as a solid support for absorption of catalysts.

As a result, HAp has great affinity for aspartic acid or glutamic acid,[58,59] and for charged peptides, due to electrostatic interactions between the ionic groups of the amino acids and the charged groups on the crystal surface. While peptides with positive net charge show modest affinity for HAp, the highest affinity is observed for zwitterionic or negatively charged peptides.[60,61] Carboxylic acids also bind strongly to P-OH via hydrogen bonding.[62] Consistently, high coating concentrations of the cell-adhesion peptide arginine-glycine-aspartate (RGD) onto the surface of hydroxyapatite have been obtained by simple incubation of the material in aqueous solutions of the peptide.[63] In addition, HAp forms composite materials with poly-L-aspartic acid or polyacrylic acid in which the polymers coat the crystal faces.[64]

6.2 Results and Discussion

The HAp nanocrystals were prepared using a co-precipitation method.[59,64] Their dimensions were observed by transmission electron microscopy (TEM), and the crystal phase was confirmed by powder X-ray diffraction (PXRD). In particular, TEM micrographs revealed the presence of plate-shaped HAp crystals with mean dimensions of about 200 nm × 40 nm (Figure 6.2 A), and the X-ray diffraction patterns showed the presence of the typical diffraction peaks indicating the presence of HAp as unique crystalline phase (Figure 6.2 B).

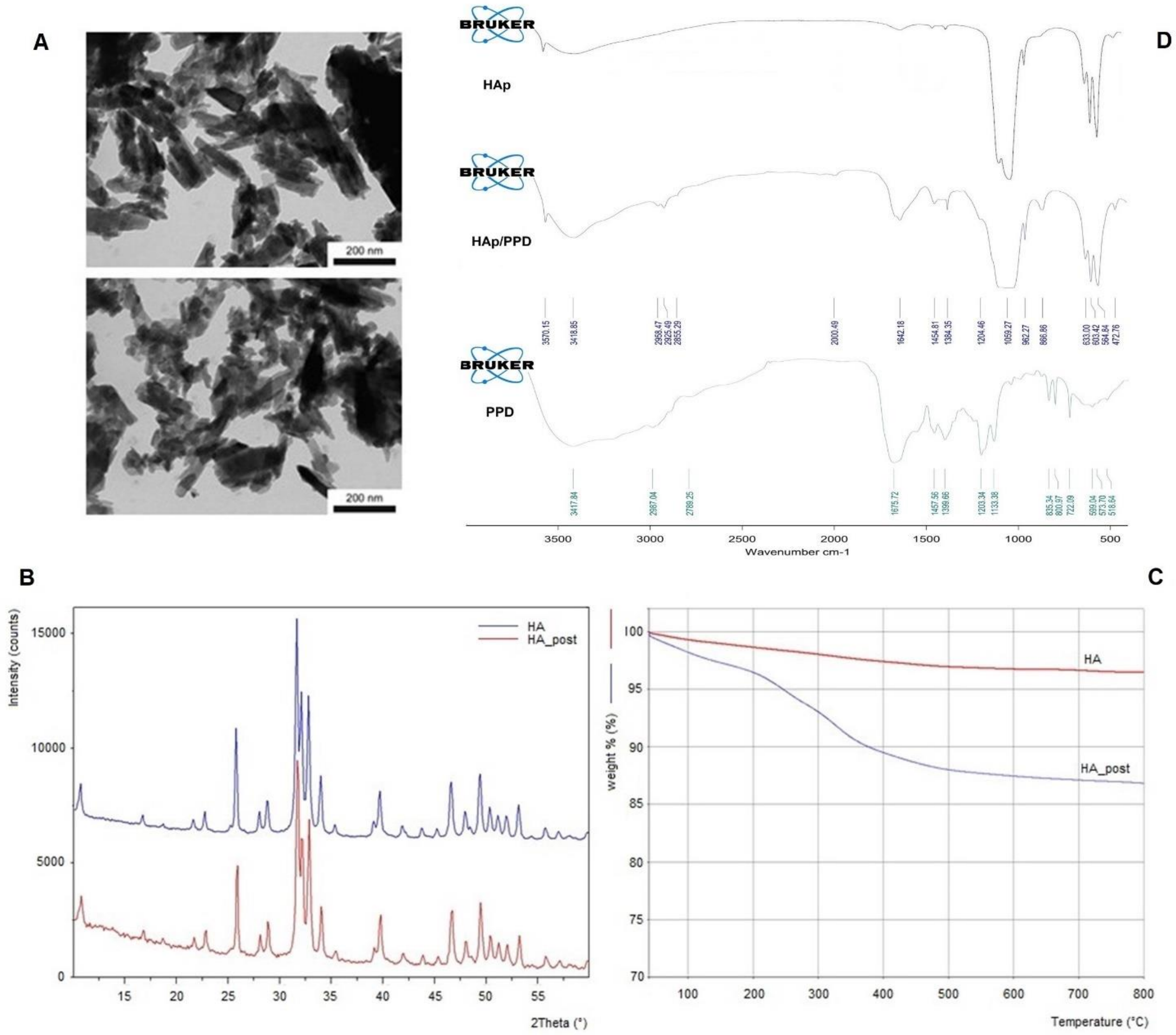
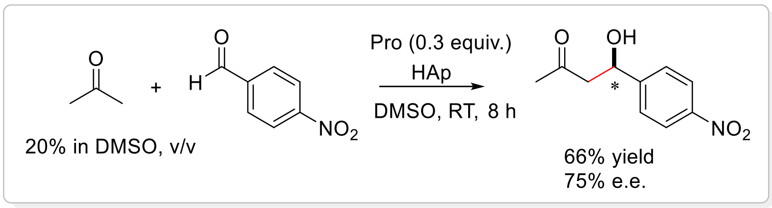


Figure 6.2 Characterization of HAp nanocrystals and recycled HAp/peptide blend. (A) TEM image of pristine HAp nanocrystals (top) and of recycled HAp/peptide blend (bottom); bar = 200 nm. (B) PXRD. (C) TGA patterns of pristine HAp and recycled HAp/peptide. (D) FT IR of HAp, HAp/PPD, and PPD (TFA salt).

In a preliminary experiment, we examined the potential of HAp for the recovery and reuse of proline in a classic direct asymmetric aldol reaction.[9] The reaction between 4-nitrobenzaldehyde and a large excess of acetone catalyzed by 0.3 equiv. of Pro was conducted under the same conditions reported in the literature, but in the presence of HAp (10 mg/mg Pro) (Scheme 6.1). After the scheduled time, the mixture was concentrated at reduced pressure, the residue was suspended in EtOAc by sonication, and the crystal were collected by centrifugation and washed with the same solvent, while the washes were collected and analysed. Remarkably, no traces of leaked Pro were detected by ^1H NMR or HPLC-MS analysis of the crude organic layer, suggesting that proline remained firmly adsorbed to the HAp support.



Scheme 6.1 Aldol reaction of acetone to 4-nitrobenzaldehyde, catalyzed by Pro in the presence of HAp.

As it turned out, the reaction proceeded as expected, giving the aldol product with 66% yield and 75% e.e., comparable to the values reported in the literature,[9] as determined after isolation by flash chromatography. Furthermore, the HAp/Pro mixture recovered by centrifugation was reutilized for a new reaction between the same partners, giving a similar yield of 62% after isolation, and 74% e.e. (Appendix D: Figure S4).

Encouraged by these results with Pro, we turned our attention to oligopeptide catalysts. Wennemers, and others, proposed the use of Pro-tripeptides as catalysis of not only 1,2- but also 1,4-addition reactions.[18-21] The distance between the secondary amine and the carboxylic acid is greater by approximately 3 Å than in simple amino acids. This additional distance, they argued, could facilitate the formation of stable complexes with longer electrophiles (Figure 6.3).

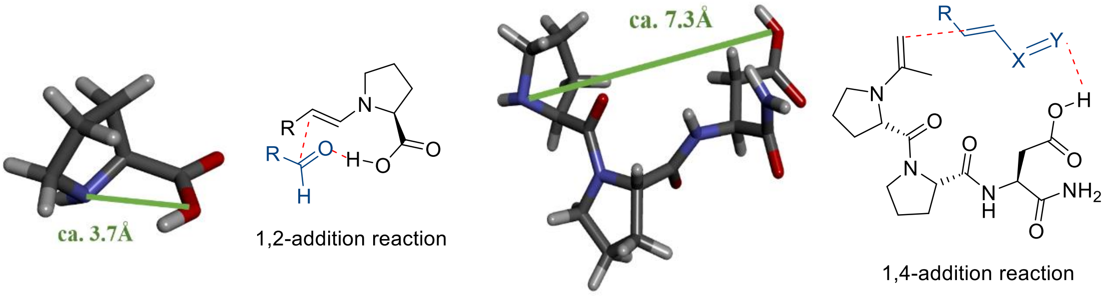


Figure 6.3 Schematic transition state of aldol reactions catalyzed by proline and its tripeptide analogue.

Asymmetric conjugate addition reactions of carbon nucleophiles to electron poor olefins are among the most useful carbon-carbon bond forming reactions, giving versatile intermediates for further transformations. As prototypic examples, we tested the 1,4-addition of aldehydes to nitroolefins or maleimide.

We selected the tripeptide Pro-Pro-AspNH2 (PPD) as a model sequence carrying both free amino and carboxylic groups (Figure 6.4). PPD was designed to promote direct aldol reactions, but also conjugate addition reactions of aldehydes to nitroolefins to give the corresponding γ-nitroaldehydes in good yields and stereoselectivities.[18] On the other hand, we also selected the peptide D-Pro-Pro-AsnNH2 (pPN) as another model sequence carrying the amino but not the free carboxylic group (Figure 6.4). This peptide was originally introduced for the stereoselective conjugate addition reactions of aldehydes to maleimide.[19]

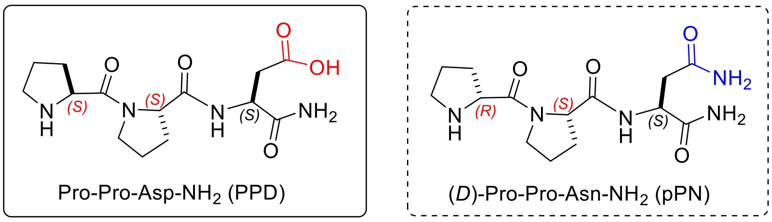
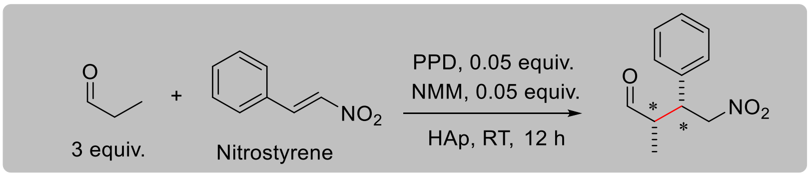


Figure 6.4 Structures of the tripeptides Pro-Pro-AspNH2 (PPD) and D-Pro-Pro-AsnNH2 (pPN).

The peptides were prepared as TFA salts in solid-phase on a Rink amide resin under standard conditions, using 9-fluorenylmethoxycarbonyl (Fmoc)-protected amino acids. Coupling reactions were performed with DCC/HOBt as activating agents. Fmoc deprotection was done with 20% v/v piperidine in DMF. Eventually, the peptides were detached from the resin by treatment with TFA in the presence of scavengers, and their identity was assessed by RP HPLC-ESI MS (Appendix D: Figures S1-S2).



Scheme 6.2 Conjugate addition reaction of propanal to nitrostyrene catalyzed by PPD tripeptide in the presence of HAp nanocrystals. For full details about solvents, e.e. and d.r., see Table 6.1.

The conjugate reaction between freshly distilled propanal and nitrostyrene was conducted at RT using the same amounts as reported,[18] i.e. 3 equiv. of aldehyde and 0.01 equiv. of *N*-methylmorpholine (NMM) in 9:1 $CHCl_3$/iPrOH, and the results were in line with the expectations (Table 6.1, Entry 1). The same reaction was performed also in the presence of a 10:1 w/w mixture of HAp and PPD containing 0.01 equiv. of peptide (Table 6.1, Entry 2). After 12 h, the mixture was concentrated, the residue was suspended in EtOAc, the crystals were collected by centrifugation and washed twice with the same solvent. The analysis of the collected organic layers revealed a modest conversion. Thereafter, the reaction was repeated upon increasing the amount of peptide and NMM, i.e. with 10:1 w/w HAp/PPD containing 0.05 equiv. of peptide, and 0.05 equiv. of NMM (Scheme 6.2). After 12 h, the mixture was concentrated, and work up was conducted as described above. The analysis of the collected organic layers revealed almost quantitative conversion, while the presence of the peptide was excluded, suggesting an efficient capture by HAp. The product was isolated by flash chromatography over silica-gel, and reaction yield, diastereomeric ratio (d.r.) and e.e. (Table 6.1, Entry 3) were similar to those reported in Entry 1, i.e. 95% yield, syn/anti 10:1, e.e. 83%.

Since HAp is a relatively mild base, we repeated the reaction without adding NMM. Indeed, the latter is utilized to liberate the secondary amine at the N-terminus of the peptide. In the absence of TFA and NMM using the 'desalted' peptide the addition products can be obtained with the same yields and stereoselectivity.[65] However, in our hands the reaction gave the product only in traces (Table 6.1, Entry 4), demonstrating that the basicity of HAp is insufficient to effectively desalt the peptide TFA salt, highlighting the importance of NMM.

The HAp/PPD hybrid material isolated from Entry 3 was recycled and reused for subsequent reactions between the same reaction partners. Each reaction provided the same product, albeit in gradually decreasing yields (Appendix D: Table S1), so that the fourth repeat gave 75% yield after isolation, while d.r. and e.e. remained practically the same (Table 6.1, Entry 5). On prosecuting the reuse of the catalyst beyond, the yield showed a significant drop after the sixth repeat (Appendix D: Table S1, Entry 6). Interestingly enough, also these reactions using recycled HAp/PPD required the presence of NMM.

After six repeats, the activity of the peptide/HAp system could be restored only partially by addition of 0.02 equiv. of fresh peptide, with a reaction yield of 85%, with the same d.r. and e.e. Since no leaching of the peptide was detected, the activity decrease may be explained by modification of the material by adsorbed reactants, products and side-products.

It is worth pointing out that the complete recovery of the peptide is not a mere consequence of its scarce solubility in the organic layer. Indeed, low percentages of peptides such as PPD can be solubilized and utilized to promote reactions in diverse organic solvents of moderate polarity such as THF, $CHCl_3$, EtOAc.[18,65]

In fact, the reaction was successfully performed in EtOAc, both in the absence (Table 6.1, Entry 6) and in the presence of HAp (Table 6.1, Entry 7), giving results similar to the literature.[18,65] The HAp/peptide mixture was reutilized in the same solvent, with tolerable loss of efficiency (Table 6.1, Entry 8, and Appendix D: Table S1). Similar good results have been observed in EtOH (Table 6.1, Entry 9 without HAp, Entry 10 with HAp). Also in this case the recycled hybrid catalyst could be reutilized with an acceptable drop of efficiency (Table 6.1, Entry 11). Remarkably, for their low toxicity and possible preparation from renewable raw materials, EtOAc and EtOH are currently classified among the greener solvents.[66] In contrast, when the reaction was conducted in $CHCl_3$, the formation of the product was observed as expected in the absence of HAp (Table 6.1, Entry 12), but not in the presence of HAp (Table 6.1, Entry 13).

Table 6.1 Reagents and conditions for the peptide-catalyzed stereoselective conjugate addition reactions of aldehydes (3 equiv.) to α, β-unsaturated acceptors.

Entry	Peptide ·TFA	Peptide (equiv.)	HAp w/w	Aldehyde	α, β-unsat. acceptor	Solvent	NMM (equiv.)	Yield (%) [a]	d.r.[b]	e.e. (%) [c]
1	PPD	0.01	-	Propanal	nitrostyrene	$CHCl_3$/iPrOH	0.01	95	10:1	84
2	PPD	0.01	10:1	Propanal	nitrostyrene	$CHCl_3$/iPrOH	0.01	45[d]	n.d.	n.d.
3	PPD	0.05	10:1	propanal	nitrostyrene	$CHCl_3$/iPrOH	0.05	95	10:1	83
4	PPD	0.05	10:1	propanal	nitrostyrene	$CHCl_3$/iPrOH	-	<10	n.d.	n.d.
5	PPD [e]	0.05	10:1 [e]	propanal	nitrostyrene	$CHCl_3$/iPrOH	0.05	75[e]	9:1	80
6	PPD	0.01	-	propanal	nitrostyrene	EtOAc	0.01	93	9:1	75
7	PPD	0.05	10:1	propanal	nitrostyrene	EtOAc	0.05	93	9:1	75
8	PPD [e]	0.05	10:1 [e]	propanal	nitrostyrene	EtOAc	0.05	78[e]	8:1	72
9	PPD	0.01	-	propanal	nitrostyrene	EtOH	0.01	95	9:1	74
10	PPD	0.05	10:1	propanal	nitrostyrene	EtOH	0.05	94	9:1	75
11	PPD [e]	0.05	10:1 [e]	propanal	nitrostyrene	EtOH	0.05	80 [e]	8:1	70
12	PPD	0.01	-	propanal	nitrostyrene	$CHCl_3$	0.01	92	11:1	83
13	PPD	0.05	10:1	propanal	nitrostyrene	$CHCl_3$	0.05	<5	n.d.	n.d.
14	pPN	0.05	-	propanal	maleimide	$CHCl_3$/iPrOH	0.05	85	1:1	n.d.
15	pPN	0.05	10:1	propanal	maleimide	$CHCl_3$/iPrOH	0.05	90	4:6	n.d.
16	pPN	0.05	-	HCA	maleimide	$CHCl_3$/iPrOH	0.05	80	9:1	97
17	pPN	0.05	10:1	HCA	maleimide	$CHCl_3$/iPrOH	0.05	80	4:6	92
18	pPN	0.05	-	HCA	maleimide	EtOH	0.05	93	9:1	84
19	pPN	0.05	10:1	HCA	maleimide	EtOH	0.05	95	4:6	n.d.

[a] After isolation by flash chromatography. [b] Determined by ^{1}H NMR spectroscopy. [c] Determined by chiral-phase HPLC analysis after isolation by flash chromatography. [d] Estimated from the analysis of the crude reaction mixture. [e] After 4 repeats, i.e. the initial reaction with fresh HAp/PPD + 3 subsequent reactions using HAp/PPD recycled from previous reactions. n.d. not determined. HCA, hydrocinnamaldehyde.

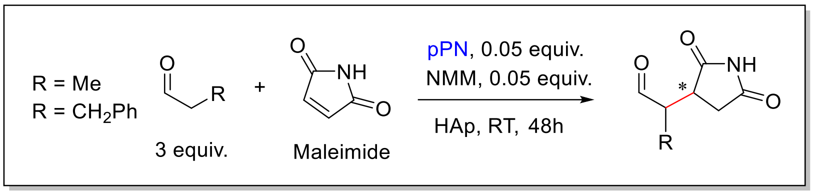
The recycled HAp/peptide mixture was analyzed by transmission electron microscopy (TEM), thermo-gravimetric analysis (TGA), powder X-ray diffraction, and Fourier transform infrared spectroscopy (FT IR) (Figure 6.2). TEM of pristine HAp nanocrystals and of recycled HAp/peptide blend confirmed that the morphology of the crystals was maintained (Figure 6.2A).

From TGA (Figure 6.2 C), the total loss of the HAp/peptide sample was 11.8%. Since HAp alone loses about 2-3% by weight, it can be estimated that the HAp/peptide sample contained about 10 ± 1 % by weight of organic component. This supports that all of the initial peptide mass was still present in the recycled material.

From the comparison of the X-rays diffraction patterns (Figure 6.2 B), it can be perceived that the crystal phase and the crystallinity of the HAp/peptide sample do not change respect to pristine HAp, and that the lattice of HAp is not altered. Therefore, it is possible to infer that the peptide interaction with the apatite crystals occurs on the surface of the inorganic particles through electrostatic interaction. Previously, a specific interaction between the carboxylic groups of amino acids and the Ca ions of HAp has been proposed.[58]

Figure 6.2 D reports the FT IR analyses of pristine HAp, of HAp/PPD peptide, and of PPD peptide alone. The spectrum of HAp/PPD clearly shows peaks attributable to the peptide, not present in the spectrum of HAp: 1204, 1384, 1454, 1642, 1670, 2855, 2925, 2958, 3418 cm^{-1}. However, small but significant differences can be observed in the wavenumbers, suggestive of direct peptide-material interactions. In particular, the shift of the major carboxylic C=O stretching band to lower wavenumbers, from 1675 cm^{-1} in PPD to 1642 cm^{-1} in the hybrid material, can be explained by an increase of the C-O bond length, attributable to its grafting to Ca ions on the surface of the hydroxyapatite, in agreement with previous results reported for apatites modified by amino-acids.[67] In addition, the comparison with the spectrum of the peptide PPD alone shows for HAp/PPD the disappearance of several peaks, including peaks attributable to the secondary amine of Pro, as well as of the strong and broad band covering the range between 2800 and 3500 cm^{-1} for the O-H stretch of carboxylic acid.

Then we moved to the conjugate addition reaction of aldehydes to maleimide catalyzed by pPN (Scheme 6.3). We utilized 3 equiv. of propanal and 0.05 equiv. of both pPN and NMM, in 1:1 $CHCl_3$/iPrOH, and the reaction proceeded almost quantitatively in the absence (Table 6.1, Entry 14) or in the presence of HAp (10:1 w/w respect to the peptide, Table 6.1, Entry 15). However, after product isolation by flash chromatography, ^1H NMR analysis showed null (Entry 14) or very modest (Entry 15), circa 40:60 diasteroselectivity (Appendix D: Figure S7).



Scheme 6.3 Conjugate addition reaction of aldehydes to maleimide catalyzed by pPN tripeptide in the presence of HAp nanocrystals. For full deteils about solvents, e.e. and d.r., see Table 6.1.

Since a similarly scarce diastereoselectivity was reported in the literature[68] for the reaction of propanal to *N*-phenyl maleimide, we repeated the reaction with hydrocinnamaldehyde (HCA, Table 6.1, Entry 16 without HAp, Entry 17 with HAp), for which the literature reports quantitative yield, 88:13 d.r., and 98% e.e.[19] In the absence of HAp the reaction proceeded as expected (Entry 16). However, also in this case the analysis of the reaction with HAp/pPN revealed a diasteromeric ratio of only 40:60, albeit a certain enantioselectivity was maintained (92 % e.e. Appendix D: Figures S8-S9). As a greener alternative to chloroform/isopropanol,[66] the reaction was performed in EtOH (Table 6.1, Entry 18 without HAp[19], Entry 19 with HAp), and again the results of the reaction catalyzed by HAp/pPN were disappointing in terms of diastereoselectivity (Table 6.1, Entry 19).

In addition to the modest stereocontrol, the use of the peptide pPN as organocatalyst in the presence of HAp was also hampered by its leakage into the organic layers, as confirmed by HPLC-MS and ^1H NMR analyses of the crude mixtures, which precluded a possible recycle.

The evidences discussed above led us to believe that peptide PPD is absorbed on the surface of HAp mainly via the salt bridge between the negatively charged aspartate and Ca^{2+} ions of HAp, while the secondary amine of proline seems way less important. Pro-peptides are requested to have a "desalted" secondary amine to exert catalytic activity, therefore the contribution of the amine to the interactions with HAp is expected to be inferior. HAp surface provides a specific environment that presents a high number of ions in repeating geometry, a feature which can induce alternative binding poses of the peptides. Recently, the interaction and folding of peptides onto the surface of HAp has been the object of much interest.[60,69] These studies support that charged residues and conformation have a role in the adhesion to the solid surface. For instance, computational studies conducted on Hyp-Pro-Gly peptide (Hyp, hydroxyproline) are suggestive of strong interactions and a turn conformation also in water.[70] Also, state-of-the-art calculations predicted the interaction between a designed peptide motif Glu^1-Ser-Gln-Glu^4-Ser^5 and the HAp solid. In the simulated structures, the peptide adopts an arch-like form at the surface, and the negatively charged Glu^1 and Glu^4 strongly contribute by providing electrostatic stabilizations with the calcium ions.[71]

Plausibly, the lack of a carboxylate group in pPN relates to its poor affinity for HAp. Nevertheless, the scarce diastereoselectivity of pPN in the conjugate additions of aldehydes to maleimide suggests that some interactions with the surface of the nanocrystals do occur. Besides the secondary amine at the N-terminus, a well-defined turn conformation is crucial for high catalytic activity and selectivity.[65,72] If peptide conformation changes upon interactions with HAp surfaces, then the functions of the peptides may be altered. In this perspective, the scarce diastereoselectivity of pPN could be the result of peptide misfolding, including a modification of the *trans*/*cis* amide bond ratio.[73]

6.3 Conclusion

Trying to mimic the catalytic action of the selective enzymes, over the last decade, small organic molecules have been successfully used as promoters for enantioselective organic transformations.

In this study, it has been shown that a peptide-acid organocatalyst can be expediently anchored on HAp as a crystalline support by simple adsorption *in situ* without pre-preparation of the solid catalyst. The peptide/HAp mixture has been utilized for performing conjugate addition reactions, and the hybrid material can be readily recovered by centrifuge, plausibly correlated to the sufficient affinity attained between the peptide and the charged Ca^{2+} and PO_4^{3-} groups on the surface of the nanocrystalline material HAp. The peptide-acid PPD was recovered without loss of either the organic component or catalytic performance, while the peptide-amide pPN was captured very poorly by HAp, preventing from its recycle and reuse.

6.4 Experimental Section

General methods. All commercially available reagents purchased from Merck (Darmstadt, Germany) were used without further purification, with the notable exception of aldehydes which were distilled prior to use. Purities were assessed by analytical reverse phase high-performance liquid chromatography (RP HPLC) on a 1200 series apparatus (Agilent, CA, USA), using a XSelect Peptide CSH C18 column (Waters, Milford, MA, USA), 4.6 mm × 100 mm, 130 Å, 3.5 μm; DAD 210 nm; mobile phase: from 8:2 water/CH_3CN containing 0.5% HCOOH, to 2:8 water/CH_3CN containing 0.5% HCOOH, in 8 min, at a flow rate of 0.5 ml min^{-1}, followed by 10 min at the same composition. ESI-MS was done on a MS single quadrupole HP 1100 MSD detector (Agilent). ^1H NMR was performed at 600 MHz on a Bruker600 in 5 mm tubes in $CDCl_3$ at RT; chemical shifts are reported as δ values relative to residual H peak ($\delta H = 7.26$ ppm). Chiral stationary phase (CSP)-HPLC analyses were performed on an Agilent Technologies Series 1200 instrument (Agilent, Santa Clara, CA, United States) using Daicel®chiral columns (DAICEL CHIRALCEL AD-H, 4.6 mm×250 mm, DAICEL CHIRALCEL OD-H, 4.6 mm×250 mm, DAICEL CHIRALCEL OJ-H, 4.6 mm×250 mm) and n-hexane/2-propanol (n-Hex/IPA) mixtures (Daicel, Osaka, Japan). Powder X-ray diffraction patterns were recorded using a PANalytical-X'Pert PRO powder diffractometer (Malvern Panalytical-Spectris, Malvern, UK) equipped with a fast X'Celerator (Malvern Panalytical-Spectris) detector ($\lambda = 0.154$ nm, 40 mA, 40 kV). For phase identification, the 2θ range was investigated from 10 to 60 2θ° with a step size of 0.1° and time/step of 100 s. The peptide content of the substrate was measured through thermogravimetric analysis using a Perkin Elmer TGA-7 (Perkin Elmer, Milan, Italy). Samples (5-10 mg) were heated in a platinum crucible in air flow (20 cm^3 min^{-1}) at a rate of 10 °C min^{-1} up to 800 °C. Transmission Electron Microscopy investigations were performed using a Philips CM 100 transmission electron microscope operating at 80 kV. To prepare samples for TEM observations, a drop of the suspension in ethanol was transferred onto holey carbon foils supported on conventional copper micro-grids. Fourier transform IR (FT IR) analysis was done on a Cary 630 FT IR spectrophotometer (Agilent CA, USA); 1 mg of the powdered HAp was mixed with 250 mg of KBr (infrared grade) and pelletized at the pressure of 10 tons for 2 min. The pellets were analyzed in the range 4000-400 cm^{-1}, 32 scans, resolution 4 cm^{-1}.

General Protocol for the Synthesis of Peptides. The peptides were prepared in solid-phase on a Rink amide MBHA resin (100-200 mesh, Fmoc protected, linker loading 0.4-0.9 mmol/g) following the standard protocol for Fmoc/^{t}-Bu peptide synthesis. Prior to use, the resin was deprotected from Fmoc following the procedure reported below. For peptide bond formation, Fmoc-protected amino acid (0.6 mmol, 2 equiv. calculated on an estimated resin loading of 0.6 mmol/g), and HOBt (0.081g, 0.6 mmol), were dissolved in DMF, and the mixture was stirred for 10 min. Then the mixture was added to a pre-swollen suspension of the amino-free resin (0.5 g) in DCM (3 mL), followed by DCC (0.12 g, 0.6 mmol), and the mixture was shaken at RT for 3 h. The resin was filtered and subsequently washed 3 times with DMF, MeOH and DCM (4 mL). Coupling efficacy was monitored for the absence of unreacted resin's amino groups by Kaiser test.

General Procedure for Fmoc Deprotection. 20% v/v piperidine in DMF was added to the resin and the reaction mixture was shaken for 30 min, drained and the piperidine treatment repeated for another 30 min. The resin was then subsequently washed with DMF, MeOH and DCM (3x each). Both couplings and deprotections were monitored by qualitative Kaiser (primary amines) and chloranil tests (secondary amines).

General Procedure for Peptide Cleavage. The release of peptides from the resin support was performed by shaking the resin in a mixture of TFA/TIPS/water/phenol (88:5:5:2) for 2.5 h. The mixture was filtered, and resin was then washed with Et_2O containing 5% TFA, with DCM, then with Et_2O (10 mL each). After filtration of the resin, the solvents and the volatiles were removed by blowing nitrogen, and the residue was suspended in ice-cold Et_2O (10 mL). The peptide which precipitated as a TFA salt was collected by centrifugation (10 min, 6,000 rpm), and the precipitate was resuspended in ice-cold Et_2O (10 mL) and centrifugated again. The peptide salts so collected were analyzed by RP-HPLC and ESI-MS spectroscopy.

HAp nanocrystals. The synthesis of HAp was carried out under N_2 atmosphere and using CO_2-free distilled water to avoid the uncontrolled insertion of carbonate ions into the crystal structure. A $(NH_4)_2HPO_4$ solution (50 ml, 0.65 M) at pH adjusted to 10 with NH_4OH was heated to 90 °C, afterwards 50 mL of 1.08 M $Ca(NO_3)_2{\cdot}4H_2O$, pH 10 adjusted with NH_4OH, was added dropwise while stirring. The suspension was stirred for 5 h at 90 °C, then separated by centrifugation at 10,000 rpm. The product was repeatedly washed with CO_2-free distilled water and dried overnight at 37 °C.

Pro-catalyzed aldol reaction of acetone with 4-nitrobenzaldehyde. A 10:1 w/w mixture of L-proline (3.4 mg, 0.03 mmol) and HAp (34 mg) was suspended in 4:1 DMSO/acetone (1.5 mL), and the suspension was stirred for 15 min at RT. Then, 4-nitrobenzaldehyde (15.1 mg, 0.1 mmol) was added while stirring at RT. After 8 h, the mixture was concentrated at reduced pressure. The residue was suspended in EtOAc by sonication, the crystals were collected by centrifugation (10,000 rpm) for 15 min, and washed again twice with the same solvent. The washes were collected and concentrated. The mixture was purified by flash chromatography on silica gel (eluant 4:1 hexane/EtOAc), giving the product as an oil (13.8 mg, 66% yield).

Conjugate addition reaction of propanal to nitrostyrene. A 10:1 w/w mixture of peptide Pro-Pro-$AspNH_2{\cdot}TFA$ (PPD, 4.0 mg, 0.009 mmol, 0.05 equiv) and HAp (40 mg) was suspended in the selected solvent (1 mL, 9:1 v/v mixture of $CHCl_3$/iPrOH, or EtOAc), then nitrostyrene (26.8 mg, 0.18 mmol, 1.0 equiv) and propanal (38.6 mL, 0.54 mmol, 3.0 equiv) were added under stirring. After 10 min, NMM (0.92 mg, 0.009 mmol, 0.05 equiv) was also added. The reaction mixture was stirred for 18 h at RT and then the solvent was evaporated. The mixture was suspended in EtOAc, and sonicated for 10 min. The crystals were collected by centrifugation (10,000 rpm, 15 min). The crystals were washed twice as describe, then the collected organic layers were concentrated. The residue was separated by flash chromatography (eluant 3:1 hexane/EtOAc), giving the product in almost quantitative yield (37.2 mg, 95%).

Conjugate addition of aldehydes to maleimide. A 10:1 w/w mixture of peptide D-Pro-Pro-AsnNH$_2$·TFA (pPN, 2.8 mg, 0.0064 mmol, 0.05 equiv) and HAp (40 mg) was suspended in the selected solvent (1.0 mL, 1:1 v/v CHCl3 and iPrOH or EtOH), then maleimmide (12.4 mg, 0.128 mmol, 1.0 equiv) and the aldehyde (HCA, 0.38 mmol, 3.0 equiv) were added while stirring. After 30 min *N*-methylmorpholine (0.6 mg, 0.0064 mmol, 0.05 equiv) was added. The reaction mixture was stirred for 48 h at RT, then the solvent was evaporated at reduced pressure, and the mixture was separated as described above. The product was isolated (23.6 mg, 80%) by flash chromatography over silica gel (eluant 2:1 hexane/EtOAc).

Note: The work described in this chapter has been published in ***Sustainable Chemistry and Pharmacy*** with slight modification and has been reproduced here with the permission of the copyright holder.

He, T.; Valagussa, B.; Boanini, E.; Gentilucci, L.* Expedient recycling of peptide organocatalysts using a nanocrystalline hydroxyapatite catching system. *Sustain. Chem. Pharm.* **2024**, *37*, 101383.

References

[1] Hernández, J.G.; Juaristi, E. Recent efforts directed to the development of more sustainable asymmetric organocatalysis. *Chem. Commun.* **2012**, *48*, 5396-5409.

[2] Han, B.; He, X.-H.; Liu, Y.-Q.; He, G.; Peng, C.; Li, J.-L. Asymmetric organocatalysis: An enabling technology for medicinal chemistry. *Chem. Soc. Rev.* **2021**, *50*, 1522-1586.

[3] Aukland, M.H.; List, B. Organocatalysis emerging as a technology. *Pure Appl. Chem.* **2021**, *93*, 1371-1381.

[4] Anastas, P.T.; Warner, J.C. Principles of green chemistry. *Green Chem.: Theor. Pract.* Oxford University Press, New York, **1998**, 29-56.

[5] Ávila-Ortiz, C.G.; Vega-Peñaloza, A.; Juaristi, E. Organocatalytic Reactions Under Nontraditional Conditions. In *Asymmetric Organocatalysis*, **2023**, pp 433-454.

[6] Antenucci, A.; Dughera, S.; Renzi, P. Green chemistry meets asymmetric organocatalysis: A critical overview on catalysts synthesis. *ChemSusChem* **2021**, *14*, 2785-2853.

[7] Xiang, S.-H.; Tan, B. Advances in asymmetric organocatalysis over the last 10 years. *Nat. Commun.* **2020**, *11*, 3786.

[8] Jarvo, E.R.; Miller, S.J. Amino acids and peptides as asymmetric organocatalysts. *Tetrahedron* **2002**, *58*, 2481-2495.

[9] List, B.; Lerner, R.A.; Barbas III, C.F. Proline-catalyzed direct asymmetric Aldol reactions. *J. Am. Chem. Soc.* **2000**, *122*, 2395-2396.

[10] Pasuparthy, S.D.; Maiti, B. Enantioselective Organocatalytic Michael Addition Reactions Catalyzed by Proline/Prolinol/Supported Proline based Organocatalysts: An overview. *ChemistrySelect* **2022**, *7*, e202104261.

[11] Wieczorek, R.; Adamala, K.; Gasperi, T.; Polticelli, F.; Stano, P. Small and random peptides: An unexplored reservoir of potentially functional primitive organocatalysts. The case of Seryl-Histidine. *Life* **2017**, *7*, 19.

[12] Metrano, A.J.; Chinn, A.J.; Shugrue, C.R.; Stone, E.A.; Kim, B.; Miller, S.J. Asymmetric catalysis mediated by synthetic peptides, version 2.0: Expansion of scope and mechanisms. *Chem. Rev.* **2020**, *120*, 11479-11615.

[13] Gentilucci, L.; Tolomelli, A.; Squassabia, F. Peptides and Peptidomimetics in Medicine, Surgery and Biotechnology. *Curr. Med. Chem.* **2006**, *13*, 2449-2466.

[14] Oku, J.-i.; Inoue, S. Asymmetric cyanohydrin synthesis catalysed by a synthetic cyclic dipeptide. *J. Chem. Soc., Chem. Commun.* **1981**, 229-230.

[15] Juliá, S.; Masana, J.; Vega, J.C. "Synthetic enzymes". Highly stereoselective epoxidation of chalcone in a triphasic Toluene-Water-Poly[(S)-alanine] system. *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 929-931.

[16] Krattiger, P.; Kovasy, R.; Revell, J.D.; Ivan, S.; Wennemers, H. Increased structural complexity leads to higher activity: Peptides as efficient and versatile catalysts for asymmetric Aldol reactions. *Org. Lett.* **2005**, *7*, 1101-1103.

[17] Avila-Ortiz, C.G.; Díaz-Corona, L.; Jiménez-González, E.; Juaristi, E. Asymmetric Michael addition organocatalyzed by α,β-dipeptides under solvent-free reaction conditions. *Molecules* **2017**, *22*, 1328.

[18] Wiesner, M.; Revell, J.D.; Wennemers, H. Tripeptides as efficient asymmetric catalysts for 1,4-addition reactions of aldehydes to nitroolefins-a rational approach. *Angew. Chem. Int. Ed.* **2008**, *47*, 1871-1874.

[19] Grünenfelder, C.E.; Kisunzu, J.K.; Wennemers, H. Peptide-catalyzed stereoselective conjugate addition reactions of aldehydes to maleimide. *Angew. Chem. Int. Ed.* **2016**, *55*, 8571-8574.

[20] Čmelová, P.; Vargová, D.; Šebesta, R. Hybrid Peptide-Thiourea Catalyst for Asymmetric Michael Additions of Aldehydes to Heterocyclic Nitroalkenes. *J. Org. Chem.* **2021**, *86*, 581-592.

[21] Möhler, J.S.; Beiersdörfer, L.K.; Masina, B.; Wechsler, P.; Wennemers, H. Tripeptide Organocatalysts for Stereoselective Conjugate Addition Reactions with N-Heterocyclic Substituents. *Adv. Synth. Catal.* **2022**, *364*, 3354-3359.

[22] Čmelová, P.; Šramel, P.; Zahradníková, B.; Modrocká, V.; Szabados, H.; Mečiarová, M.; Šebesta, R. Pro-Pro Dipeptide-Thiourea Organocatalyst in the Mannich Reaction between α-Imino Esters and Pyruvates. *Eur. J. Org. Chem.* **2022**, e202200106.

[23] Peris, G.; Jakobsche, C.E.; Miller, S.J. Aspartate-catalyzed asymmetric epoxidation reactions. *J. Am. Chem. Soc.* **2007**, *129*, 8710-8711.

[24] Lewis, C.A.; Gustafson, J.L.; Chiu, A.; Balsells, J.; Pollard, D.; Murry, J.; Reamer, R.A.; Hansen, K.B.; Miller, S.J. A case of remote asymmetric induction in the peptide-catalyzed desymmetrization of a bis(phenol). *J. Am. Chem. Soc.* **2008**, *130*, 16358-16365.

[25] Müller, C.E.; Wanka, L.; Jewell, K.; Schreiner, P.R. Enantioselective kinetic resolution of trans-cycloalkane-1,2-diols. *Angew. Chem. Int. Ed.* **2008**, *47*, 6180-6183.

[26] Zhao, Y.; Rodrigo, J.; Hoveyda, A.H.; Snapper, M.L. Enantioselective silyl protection of alcohols catalysed by an amino-acid-based small molecule. *Nature* **2006**, *443*, 67-70.

[27] Diener, M.E.; Metrano, A.J.; Kusano, S.; Miller, S.J. Enantioselective synthesis of 3-arylquinazolin-4(3H)-ones via peptide-catalyzed atroposelective bromination. *J. Am. Chem. Soc.* **2015**, *137*, 12369-12377.

[28] Lichtor, P.A.; Miller, S.J. Experimental lineage and functional analysis of a remotely directed peptide epoxidation catalyst. *J. Am. Chem. Soc.* **2014**, *136*, 5301-5308.

[29] Du, J.N.; Skubi, K.L.; Schultz, D.M.; Yoon, T.P. A dual-catalysis approach to enantioselective [2+2] photocycloadditions using visible light. *Science* **2014**, *344*, 392-396.

[30] Shin, N.Y.; Ryss, J.M.; Zhang, X.; Miller, S.J.; Knowles, R.R. Light-driven deracemization enabled by excited-state electron transfer. *Science* **2019**, *366*, 364-369.

[31] Hamley, I.W. Biocatalysts based on peptide and peptide conjugate nanostructures. *Biomacromolecules* **2021**, *22*, 1835-1855.

[32] Sinibaldi, A.; Della Penna, F.; Ponzetti, M.; Fini, F.; Marchesan, S.; Baschieri, A.; Pesciaioli, F.; Carlone, A. Asymmetric organocatalysis accelerated via self-assembled minimal structures. *Eur. J. Org. Chem.* **2021**, *2021*, 5403-5406.

[33] Lewis, J.C. Artificial metalloenzymes and metallopeptide catalysts for organic synthesis. *ACS Catal.* **2013**, *3*, 2954-2975.

[34] De Marco, R.; Tolomelli, A.; Greco, A.; Gentilucci, L. Controlled Solid Phase Peptide Bond Formation Using N-Carboxyanhydrides and PEG Resins in Water. *ACS Sustainable Chem. Eng.* **2013**, *1*, 566-569.

[35] Varnava, K.G.; Sarojini, V. Making Solid-Phase Peptide Synthesis Greener: A Review of the Literature. *Chem. Asian J.* **2019**, *14*, 1088-1097.

[36] Zompra, A.A.; Galanis, A.S.; Werbitzky, O.; Albericio, F. Manufacturing Peptides as Active Pharmaceutical Ingredients. *Future Med. Chem.* **2009**, *1*, 361-377.

[37] Gentilucci, L.; Tosi, P.; Bauer, A.; De Marco, R. Modern Tools for the Chemical Ligation and Synthesis of Modified Peptides and Proteins. *Future Med. Chem.* **2016**, *8*, 2287-2304.

[38] Rasmussen, J.H. Synthetic peptide API manufacturing: A mini review of current perspectives for peptide manufacturing. *Bioorg. Med. Chem.* **2018**, *26*, 2914-2918.

[39] Isidro-Llobet, A.; Kenworthy, M.N.; Mukherjee, S.; Kopach, M.E.; Wegner, K.; Gallou, F.; Smith, A.G.; Roschangar, F. Sustainability Challenges in Peptide Synthesis and Purification: From R&D to Production. *J. Org. Chem.* **2019**, *84*, 4615-4628.

[40] Bartók, M. Advances in immobilized organocatalysts for the heterogeneous asymmetric direct Aldol reactions. *Catal. Rev.* **2015**, *57*, 192-255.

[41] Franconetti, A.; de Gonzalo, G. Recent developments on supported hydrogen-bond organocatalysts. *ChemCatChem* **2018**, *10*, 5554-5572.

[42] García-Monzón, I.; Borges-González, J.; Martín, T. Solid-supported tetrahydropyran-based hybrid dipeptide catalysts for Michael addition of aldehydes to nitrostyrenes. *Adv. Synth. Catal.* **2022**, *364*, 2822-2829.

[43] Arakawa, Y.; Wiesner, M.; Wennemers, H. Efficient recovery and reuse of an immobilized peptidic organocatalyst. *Adv. Synth. Catal.* **2011**, *353*, 1201-1206.

[44] Arakawa, Y.; Wennemers, H. Enamine catalysis in flow with an immobilized peptidic catalyst. *ChemSusChem* **2013**, *6*, 242-245.

[45] Ötvös, S.B.; Mándity, I.M.; Fülöp, F. Asymmetric Aldol reaction in a continuous-flow reactor catalyzed by a highly reusable heterogeneous peptide. *J. Catal.* **2012**, *295*, 179-185.

[46] Zhao, Z.; Feng, D.; Xie, G.; Ma, X. Functionalized hollow double-shelled polymeric nano-bowls as effective heterogeneous organocatalysts for enhanced catalytic activity in asymmetric Michael addition. *J. Catal.* **2018**, *359*, 36-45.

[47] Sánchez-Antonio, O.; Romero-Sedglach, K.A.; Vázquez-Orta, E.C.; Juaristi, E. New mesoporous silica-supported organocatalysts based on (*2S*)-(1,2,4-triazol-3-yl)-proline: Efficient, reusable, and heterogeneous catalysts for the asymmetric aldol reaction. *Molecules* **2020**, *25*, 4532.

[48] An, J.-X.; Yang, F.-F.; Wang, P.; Gu, Z.-C.; Li, Y.; Chen, L.; Zhao, Y.-L.; He, B. A solid-supported organocatalyst for asymmetric Mannich reaction to construct C2-quaternary indolin-3-ones. *RSC Adv.* **2022**, *12*, 7040-7045.

[49] Zhang, J.; Li, S.; Liu, J.; Ma, X. Universal one-pot strategy for fabricating supported chiral organocatalysts via direct covalent immobilization upon hollow mesoporous polystyrene nanospheres. *Appl. Catal. A Gen.* **2023**, *649*, 118976.

[50] Zu, L.; Xie, H.; Li, H.; Wang, J.; Wang, W. Highly enantioselective Aldol reactions catalyzed by a recyclable fluorous (*S*) pyrrolidine sulfonamide on water. *Org. Lett.* **2008**, *10*, 1211-1214.

[51] Branco, L.C.; Faisca Phillips, A.M.; Marques, M.M.; Gago, S.; Branco, P.S. Recent advances in sustainable organocatalysis. *Recent advances in organocatalysis* **2016**, pp 141-182.

[52] Szőllősi, G.; Gombkötő, D.; Mogyorós, A.Z.; Fülöp, F. Surface-Improved Asymmetric Michael Addition Catalyzed by Amino Acids Adsorbed on Laponite. *Adv. Synth. Catal.* **2018**, *360*, 1992-2004.

[53] Kozma, V.; Szőllősi, G. Enantioselective Michael addition of aldehydes to maleimides catalysed by surface-adsorbed natural amino acids. *Catal. Sci. Technol.* **2022**, *12*, 4709-4726.

[54] Landeros, J.M.; Cruz-Hernández, C.; Juaristi, E. α-Amino Acids and α,β-Dipeptides Intercalated into Hydrotalcite: Efficient Catalysts in the Asymmetric Michael Addition Reaction of Aldehydes to *N*-Substituted Maleimides. *Eur. J. Org. Chem.* **2021**, *2021*, 5117-5126.

[55] Anselmi, M.; Stavole, P.; Boanini, E.; Bigi, A.; Juaristi, E.; Gentilucci, L. Green synthesis of bioactive oligopeptides promoted by recyclable nanocrystalline hydroxyapatite. *Future Med. Chem.* **2020**, *12*, 479-491.

[56] Santino, F.; Petruzzelli, R.; Zhao, J.; Boanini, E.; Gentilucci, L. Peptide bond formation using unprotected *N*-carboxyanhydrides under green chemistry conditions. *Sustain. Chem. Pharm.* **2021**, *24*, 100540.

[57] Fihri, A.; Len, C.; Varma, R.S.; Solhy, A. Hydroxyapatite: A review of syntheses, structure and applications in heterogeneous catalysis. *Coord. Chem. Rev.* **2017**, *347*, 48-76.

[58] Boanini, E.; Torricelli, P.; Gazzano, M.; Giardino, R.; Bigi, A. Nanocomposites of hydroxyapatite with aspartic acid and glutamic acid and their interaction with osteoblast-like cells. *Biomaterials* **2006**, *27*, 4428-4433.

[59] Boanini, E.; Fini, M.; Gazzano, M.; Bigi, A. Hydroxyapatite nanocrystals modified with acidic amino acids. *Eur. J. Inorg. Chem.* **2006**, *2006*, 4821-4826.

[60] Duanis-Assaf, T.; Hu, T.; Lavie, M.; Zhang, Z.; Reches, M. Understanding the adhesion mechanism of hydroxyapatite-binding peptide. *Langmuir* **2022**, *38*, 968-978.

[61] Ling, C.; Zhao, W.; Wang, Z.; Chen, J.; Ustriyana, P.; Gao, M.; Sahai, N. Structure-activity relationships of hydroxyapatite-binding peptides. *Langmuir* **2020**, *36*, 2729-2739.

[62] Tanaka, H.; Watanabe, T.; Chikazawa, M.; Kandori, K.; Ishikawa, T. TPD, FTIR, and molecular adsorption studies of calcium hydroxyapatite surface modified with hexanoic and decanoic acids. *J. Colloid Interface Sci.* **1999**, *214*, 31-37.

[63] Sawyer, A.A.; Hennessy, K.M.; Bellis, S.L. Regulation of mesenchymal stem cell attachment and spreading on hydroxyapatite by RGD peptides and adsorbed serum proteins. *Biomaterials* **2005**, *26*, 1467-1475.

[64] Bigi, A.; Boanini, E.; Gazzano, M.; Kojdecki, M.A.; Rubini, K. Microstructural investigation of hydroxyapatite-polyelectrolyte composites. *J. Mater. Chem.* **2004**, *14*, 274-279.

[65] Wiesner, M.; Neuburger, M.; Wennemers, H. Tripeptides of the type H-d-Pro-Pro-Xaa-NH_2 as catalysts for asymmetric 1,4-addition reactions: Structural requirements for high catalytic efficiency. *Chem. Eur. J.* **2009**, *15*, 10103-10109.

[66] Byrne, F.P.; Jin, S.; Paggiola, G.; Petchey, T.H.M.; Clark, J.H.; Farmer, T.J.; Hunt, A.J.; Robert McElroy, C.; Sherwood, J. Tools and techniques for solvent selection: Green solvent selection guides. *Sustain. Chem. Process.* **2016**, *4*, 7.

[67] ben Moussa, S.; Mehri, A.; Gruselle, M.; Beaunier, P.; Costentin, G.; Badraoui, B. Combined effect of magnesium and amino glutamic acid on the structure of hydroxyapatite prepared by hydrothermal method. *Mater. Chem. Phys.* **2018**, *212*, 21-29.

[68] Torregrosa-Chinillach, A.; Moragues, A.; Pérez-Furundarena, H.; Chinchilla, R.; Gómez-Bengoa, E.; Guillena, G. Enantioselective Michael addition of aldehydes to maleimides organocatalyzed by a chiral primary amine-salicylamide. *Molecules* **2018**, *23*, 3299.

[69] Rimola, A.; Aschi, M.; Orlando, R.; Ugliengo, P. Does adsorption at hydroxyapatite surfaces induce peptide folding? Insights from large-scale B3LYP calculations. *J. Am. Chem. Soc.* **2012**, *134*, 10899-10910.

[70] Almora-Barrios, N.; de Leeuw, N.H. Modelling the interaction of a Hyp-Pro-Gly peptide with hydroxyapatite surfaces in aqueous environment. *CrystEngComm.* **2010**, *12*, 960-967.

[71] Kato, K.; Fukuzawa, K.; Mochizuki, Y. Modeling of hydroxyapatite-peptide interaction based on fragment molecular orbital method. *Chem. Phys. Lett.* **2015**, *629*, 58-64.

[72] Rigling, C.; Kisunzu, J.K.; Duschmalé, J.; Häussinger, D.; Wiesner, M.; Ebert, M.-O.; Wennemers, H. Conformational properties of a peptidic catalyst: Insights from NMR spectroscopic studies. *J. Am. Chem. Soc.* **2018**, *140*, 10829-10838.

[73] Schnitzer, T.; Wennemers, H. Influence of the *trans/cis* conformer ratio on the stereoselectivity of peptidic catalysts. *J. Am. Chem. Soc.* **2017**, *139*, 15356-15362.

Chapter VII. Summary, Conclusions, and Future Perspectives

The research activities presented in this thesis focused on the design and synthesis of bioactive peptide-based therapeutic agents by adopting diverse peptidomimetic strategies, such as cyclization, the introduction of α/β hybrid amino acids, and D-configured amino acids, targeting integrin- and lactate dehydrogenase-related diseases. Moreover, by introducing appropriate linkers, these potent peptidic ligands/inhibitors were conjugated to tailor-made nanoparticles, facilitating their application in targeted therapies due to their abnormal expression in cancers and inflammatory sites. Lastly, the research also explored the application of peptides in organocatalysis, utilizing nanocrystalline hydroxyapatite as a recyclable support.

The **Introduction** briefly presented the structural and functional characteristics of the two targets involved in the thesis: integrin receptors and lactate dehydrogenase. Subsequently, peptidomimetic strategies and their applications in drug discovery and asymmetric catalysis were discussed.

Next, a special focus was given to the $\alpha_4\beta_1$ integrin subfamily in **Chapter 2**. We thoroughly reviewed the reported agonists/antagonists of integrin $\alpha_4\beta_1$. Additionally, we speculated on plausible models to explain agonism versus antagonism by comparison with the well-studied RGD-binding integrins and by analysis of computational simulations performed with homology or hybrid receptor structures, since nor the crystallographic or the cryo-EM structures of $\alpha_4\beta_1$ integrin are known. This chapter provides valuable structural information of $\alpha_4\beta_1$ integrin to guide the design of novel $\alpha_4\beta_1$ ligands in the subsequent chapters.

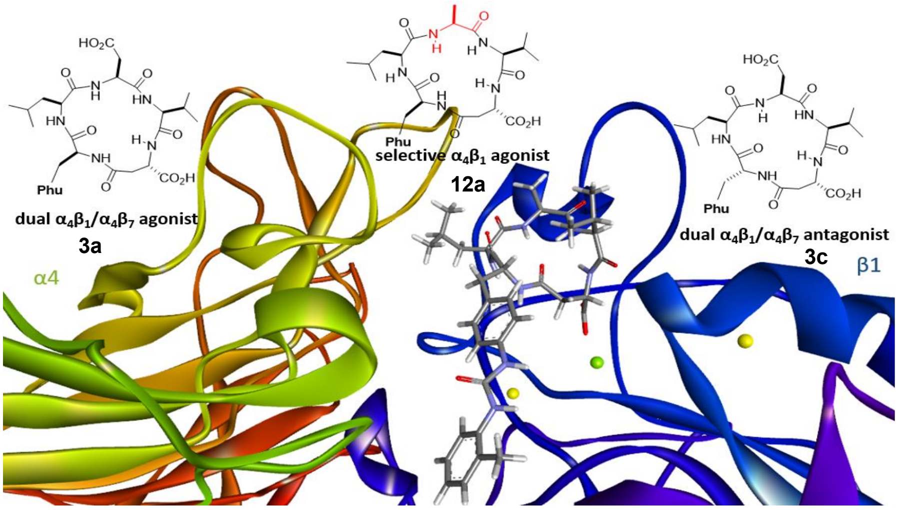


Figure 7.1 Overview of research activities carried out in **Chapter 3**.

In **Chapter 3**, we designed and constructed a minilibrary of LDV-containing α/β hybrid cyclopentapeptides (CPPs), which led to the discovery of diverse potent ligands against α_4 integrins. In particular, the CPP **3c**, c[*(R)*-Phu-LDV-*(S)*-isoAsp], resulted to be an antagonist of α_4 integrins with moderate potency, while **3a**, c[*(S)*-Phu-LDV-*(S)*-isoAsp], appeared to be a potent agonist capable to increase both $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrin-mediated cell adhesion. In addition, **12a**, c[*(R)*-Phu-LAV-*(S)*-isoAsp], was an agonist which selectively promoted the adhesion of $\alpha_4\beta_1$ with low nanomolar potency but not that of $\alpha_4\beta_7$ integrin-expressing cells. Furthermore, these molecules were used as 3D probes to investigate the preferred bioactive conformations, and to analyze $\alpha_4\beta_1$ integrin-ligand binding behavior. Eventually, the conformational analysis by NMR combined with quantum mechanics computations predicted distinct ligand-receptor interactions for antagonists and agonists, plausibly referable to receptor inhibition or activation.

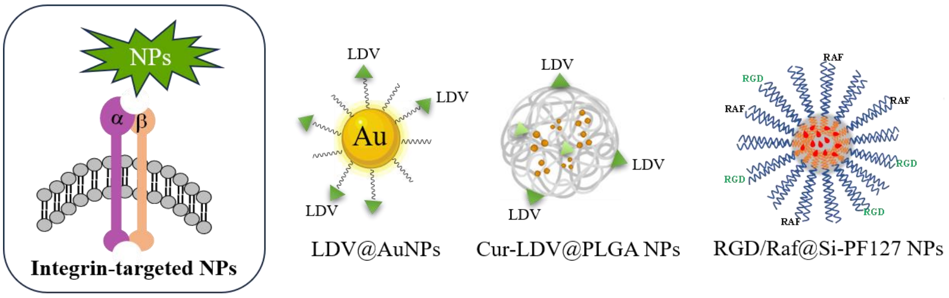


Figure 7.2 Overview of research activities carried out in **Chapter 4**.

In **Chapter 4**, firstly, building on the results obtained in Chapter 3, several potent LDV-containing integrin ligands decorated metallic gold nanoparticles were prepared. The functionalization was well-characterized by TEM, DLS, and UV-vis analysis. Given the unique optical and electronic properties of AuNPs, they hold great potential for targeted photothermal therapy. Subsequently, biocompatible polymeric PLGA nanoparticles coated with LDV peptides were fabricated. Furthermore, curcumin, an anticancer drug was loaded into the NPs with a drug loading capacity of 10%. Since curcumin is intrinsically fluorescent, this property can be advantageous for visualizing the NPs in cells. Lastly, we designed nanoplatforms with a tumor-homing RGD-based ligand and/or a cytotoxic peptide Raf analogue, based on inorganic fluorescent silica NPs coated with a biocompatible organic PEG shell. As a result, the nano-conjugates showed moderate to superior toxicity towards A549 and MCF7 cancer cells in the preliminary cell viability assays. These findings highlight the potential of integrin-targeted nanotherapeutics for cancer treatment. To further elucidate the roles that different nanoparticles and integrin ligands could play, in-depth biological evaluations are necessary and currently underway in our laboratory.

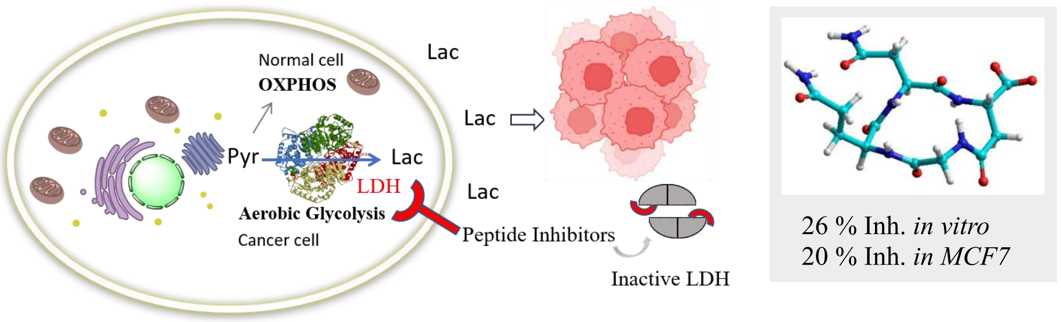


Figure 7.3 Overview of research activities carried out in **Chapter 5**.

In **Chapter 5**, taking into account that the activity of LDH is exclusively exerted by their tetrameric forms, we conceived to target the protein-protein interactions (PPIs) involved in the oligomerization of LDH 5, thereby preventing its enzymatic activity. To this end, a series of peptide inhibitors derived from the C-terminal of LDH-A were prepared. Furthermore, in order to identify their inhibitory effects in disruption of enzyme oligomerization, we exploited the unprecedented isolation of recombinant monomeric human LDH-A at neutral pH, and resulting monomer was utilized to study the assembly of corresponding catalytically-active LDH-A tetramer in vitro. Lastly, we performed the inhibitory effects of the peptides on lactate production in cellulo. By means of these assays we finally detected a selective action of linear and cyclic tetrapeptides, inhibiting lactate secretion in MCF7 cells only.

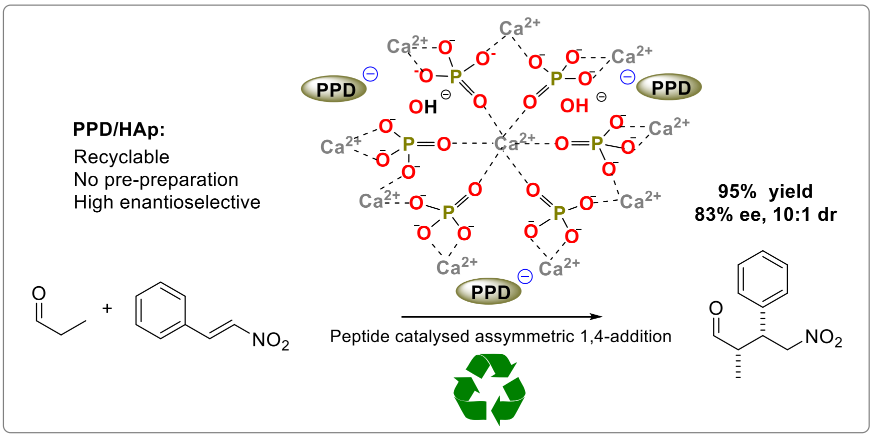


Figure 7.4 Overview of research activities carried out in **Chapter 6**.

In **Chapter 6**, we described the expedient use of nanocrystalline hydroxyapatite (HAp) as a biocompatible, inert catching system for the recovery and reuse of peptide catalysts. In detail, we used tripeptides in amalgam with HAp via electrostatic interaction as a blended recyclable catalytic system in asymmetric 1,4-conjugate addition reactions with high enantioselectivity. Additionally, the recycled HAp/peptide mixture was thoroughly characterized by TEM, TGA, X-ray diffraction, and FT-IR analysis. The blend catalyst can be recovered easily by centrifugation and reused for at least 4 cycles without compromising its catalytic performance.

In conclusion, the findings herein underscore the potential application of peptidomimetics in drug discovery, particularly in the realm of integrin ligands and lactate dehydrogenases. We believe that the constructed ligands, inhibitors and conjugates could find a wide range of applications in cancer and inflammatory-related diseases. Future developments could focus on, for example, synthesizing less-studied IDS-based sequences for targeting integrin $\alpha_4\beta_1$, as well as their nano-conjugates. Additionally, the development of inhibitors that target other critical sites involved in the association of lactate dehydrogenases, along with efforts to enhance the cell-penetrating ability of the inhibitors, will be crucial, given that the enzymes are located in the cytosol. Finally, exploring more innovative and sustainable ways for the synthesis and application of peptidomimetics represents a thriving and long-term research topic.

Appendix-A

Design and Synthesis of LDV-based $\alpha_4\beta_1$ Integrin Cyclopeptides

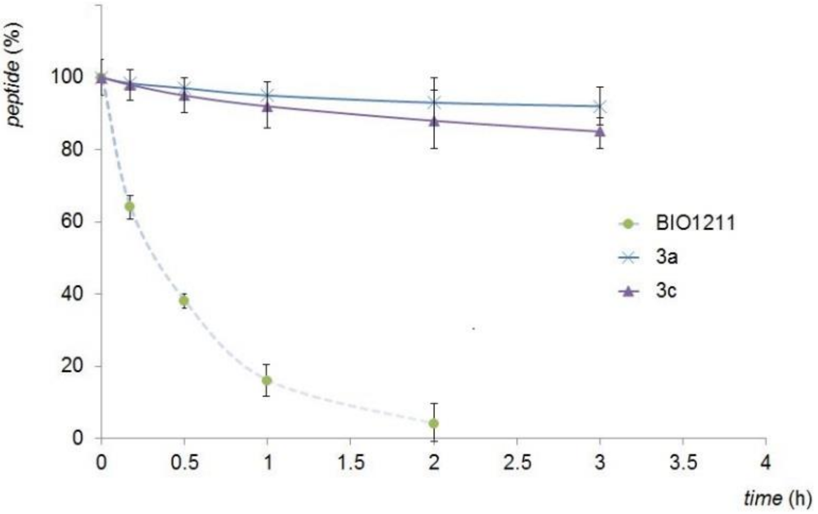


Figure S1. Degradation of BIO1211, **3a**, **3c** in mouse serum. Samples were collected from the incubation solution at the indicated times. Peptide stability was determined using an RP-HPLC ESI-MS analysis (described in the Methods). Values are presented as mean ± SD (n = 3).

Table S1. RP HPLC and ESI MS analyses of the linear precursors **9** and of the cyclopeptides **10**, and their yields.

9	Linear peptide sequence [a]	Yield (%) [b]	Purity (%) [c]	ESI MS [M+1]$^+$ found/calcd.[d]	10	Yield (%) [e]	Purity (%) [f]	ESI-MS [M+1]$^+$ found/calcd.[d]	CPP [g]
-	H-isoAsp(OBn)-Phu-Leu-Asp(OBn)-Val-OH	30	76	936.2/936.4 [h]	**a**	traces	nd	nd	**3a**
-	H-Asp(OBn)-Val-isoAsp(OBn)-Phu-Leu-OH	32	78	936.2/936.4 [h]	**a**	traces	nd	nd	**3a**
a	H-(*S*)-Phu-Leu-Asp(OBn)-Val-(*S*)-Asp-OBn	22	80	936.2/936.4 [h]	**a**	28	98	918.2/918.4 [i]	**3a**
b	H-(*S*)-Phu-Leu-Asp(OBn)-Val-(*R*)-Asp-OBn	24	81	936.2/936.4 [h]	**b**	19	97	918.2/918.4 [i]	**3b**
c	H-(*R*)-Phu-Leu-Asp(OBn)-Val-(*S*)-Asp-OBn	34	79	936.2/936.4 [h]	**c**	29	97	918.2/918.4[i]	**3c**
d	H-(*R*)-Phu-Leu-Asp(OBn)-Val-(*R*)-Asp-OBn	44	85	936.6/936.4 [h]	**d**	34	99	918.6/918.4 [i]	**3d**
e	H-(*S*)-Phu-Leu-Asp(OBn)-Val-(*R*)-β^3homoAla-OH	27	75	816.6/816.4 [j]	**e**	50	95	798.2/798.4 [k]	**11a**
f	H-(*S*)-Phu-Leu-Ala-Val-(*S*)-Asp-OBn	46	83	802.6/802.4 [l]	**f**	37	96	784.4/784.4 [m]	**12a**
g	H-(*S*)-Phu-Leu-Asp(OBn)-Val-(*S*)-Asp-nPr [a]	45	82	887.4/887.4 [n]	**g**	27	95	869.2/869.4 [o]	**13**
h	H-(*R*)-Phu-Leu-Asp(OBn)-Val-(*R*)-β^3homoAla-OH	47	78	816.6/816.4 [j]	**h**	33	98	798.2/798.4 [k]	**11c**
i	H-(*R*)-Phu-Leu-Ala-Val-(*S*)-Asp-OBn	40	79	802.6/802.4 [l]	**i**	15	97	784.2/784.4 [m]	**12c**
j	H-(*S*)-Phu-Phe-Asp(OBn)-Val-(*S*)-Asp-OBn	48	73	970.2/970.4 [p]	**j**	25	97	952.2/952.4 [q]	**14**
k	H-(*S*)-Phu-Phe-Ala-Val-(*S*)-Asp-OBn	36	85	836.2/836.4 [r]	**k**	48	98	818.2/818.4 [s]	**15**
l	H-(*R*)-Phu-Leu-Asp(OBn)-Phg-(*S*)-Asp-OBn	37	74	970.2/970.4 [p]	**l**	43	96	952.2/952.4 [q]	**16**

[a]isoAsp(OBn) corresponds to Asp-OBn; isoAsp(*n*Pr) corresponds to Asp-*n*Pr. [b]based on the estimated loading of the resin. [c]Determined by analytical RP HPLC (General methods) on a C18 RP column (100×3 mm, 3 μm, 110 Å), mobile phase from 9:1 $H_2O/CH_3CN/0.1\%$ HCOOH to 2:8 $H_2O/CH_3CN/0.1\%$ HCOOH in 20 min, flow rate of 1.0 mL min^{-1}. [d]MS single quadrupole HP 1100MSD detector. [e]Determined after semi-preparative RP HPLC (General methods) on a C18 RP column (21.2×150 mm, 7 μm 80 Å), mobile phase from 8:2 H_2O/CH_3CN to 100% CH_3CN in 10 min, flow rate 12 mL min^{-1} for CPPs **10a-i**, or on a C18 RP column (19×150 mm, 5 μm 130 Å), isocratic mobile phase 1:1 $H_2O/CH_3CN/0.1\%$ TFA in 8 min, followed by 100% CH_3CN in 5 min, flow rate 10 mL min^{-1} for CPPs **10j-l**, [f]Same stationary phase as for [c], mobile phase from 9:1 H_2O/CH_3CN to 2:8 H_2O/CH_3CN in 20 min, flow rate 1.0 mL min^{-1}. [g]Ester deprotection proceeded in nearly quantitative yield; purities are reported in Table 1 and 2. [h]Calcd for $[C_{50}H_{62}N_7O_{11}]^+$. [i]Calcd for $[C_{50}H_{60}N_7O_{10}]^+$. [j]Calcd for $[C_{43}H_{58}N_7O_9]^+$. [k]Calcd for $[C_{43}H_{56}N_7O_8]^+$. [l]Calcd for $[C_{42}H_{56}N_7O_9]^+$. [m]Calcd for $[C_{42}H_{54}N_7O_8]^+$. [n]Calcd for $[C_{46}H_{63}N_8O_{10}]^+$. [o]Calcd for $[C_{46}H_{61}N_8O_9]^+$. [p]Calcd for $[C_{53}H_{60}N_7O_{11}]^+$. [q]Calcd for $[C_{53}H_{58}N_7O_{10}]^+$. [r]Calcd for $[C_{45}H_{54}N_7O_9]^+$. [s]Calcd for $[C_{45}H_{52}N_7O_8]^+$. nd: not determined.

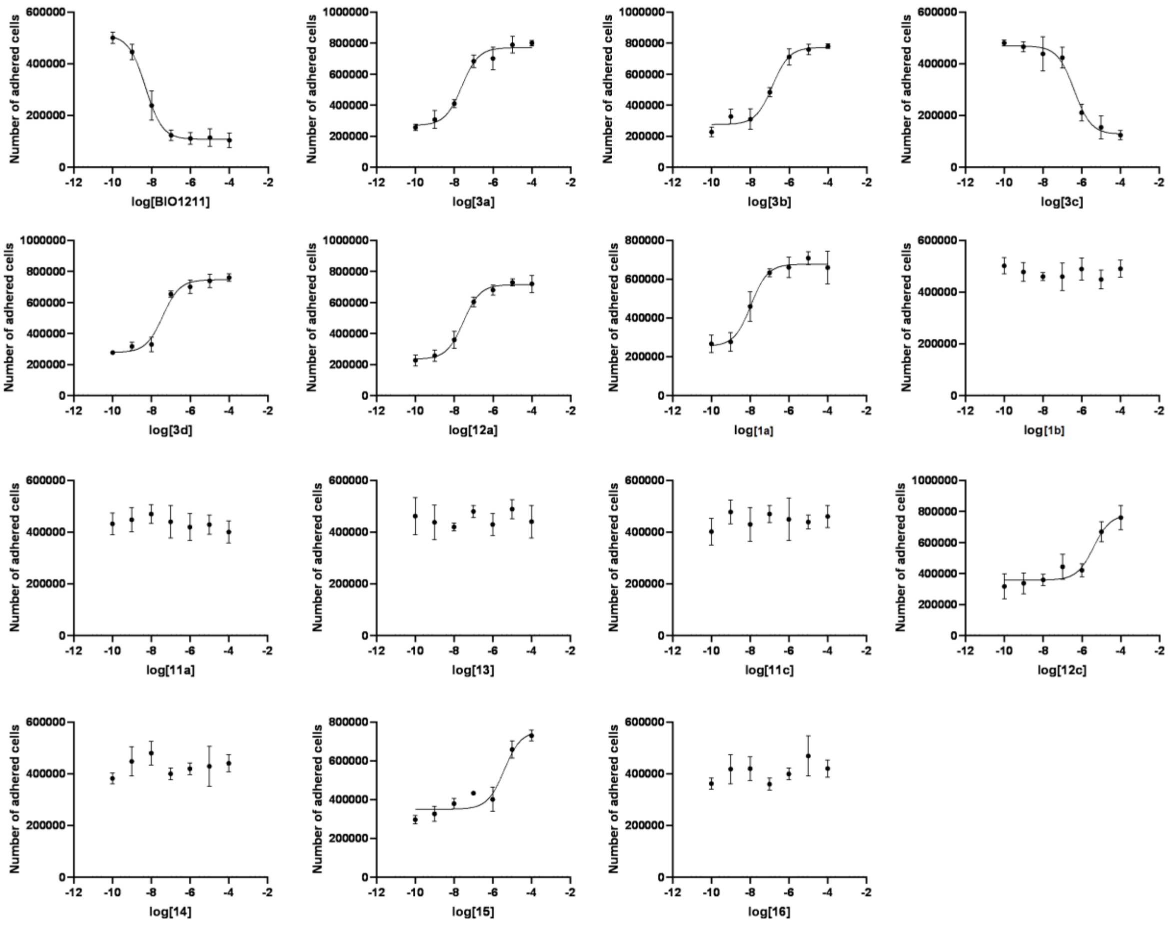


Figure S2. Concentration-response curves obtained from cell adhesion assays: evaluation of CPPs effects on $\alpha_4\beta_1$-mediated Jurkat cell adhesion to FN. Values represent the mean ± SD of three independent experiments carried out in quadruplicate.

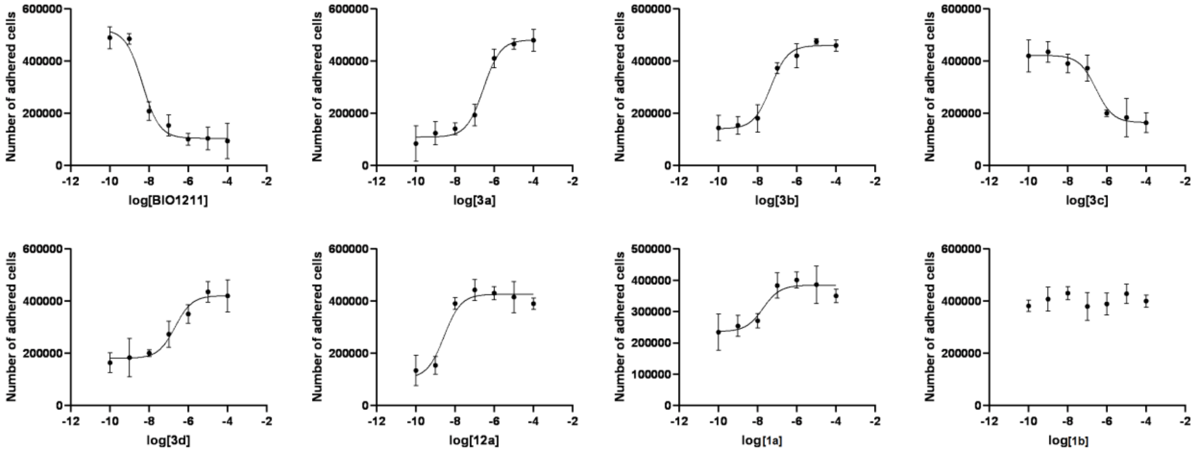


Figure S3. Concentration-response curves obtained from cell adhesion assays: evaluation of CPPs effects on $\alpha_4\beta_1$-mediated Jurkat cell adhesion to VCAM-1. Values represent the mean ± SD of three independent experiments carried out in quadruplicate.

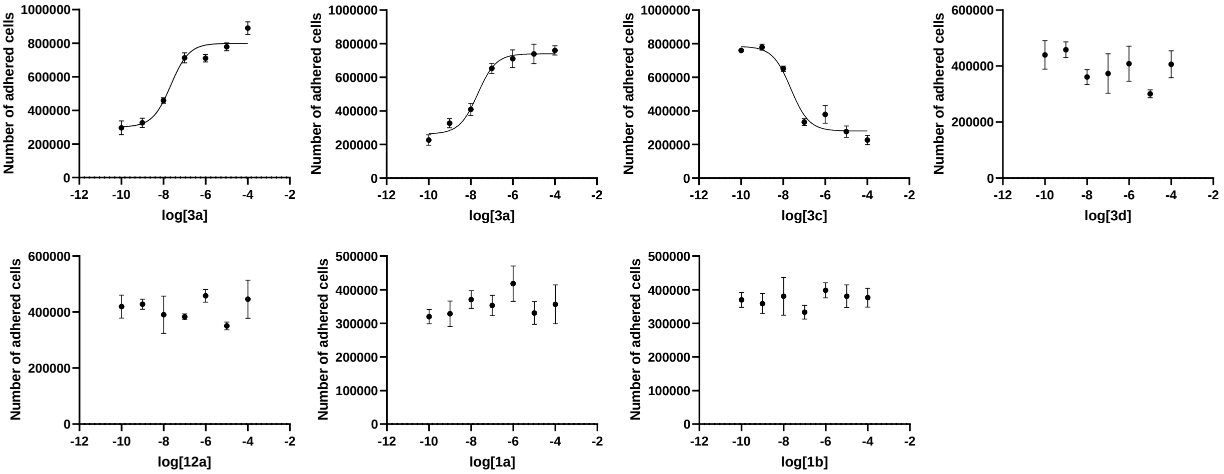


Figure S4. Concentration-response curves obtained from cell adhesion assays: evaluation of CPPs effects on $\alpha_4\beta_7$-mediated RPMI8866 cell adhesion to MAdCam-1. Values represent the mean ± SD of three independent experiments carried out in quadruplicate.

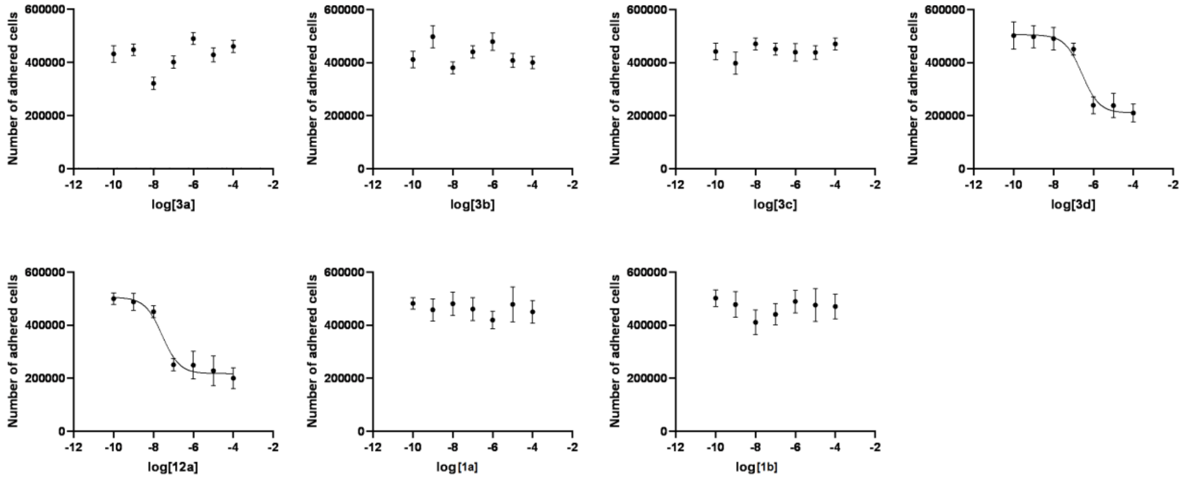


Figure S5. Concentration-response curves obtained from cell adhesion assays: evaluation of CPPs effects on $\alpha_M\beta_2$-mediated HL60 cell adhesion to FN. Values represent the mean ± SD of three independent experiments carried out in quadruplicate.

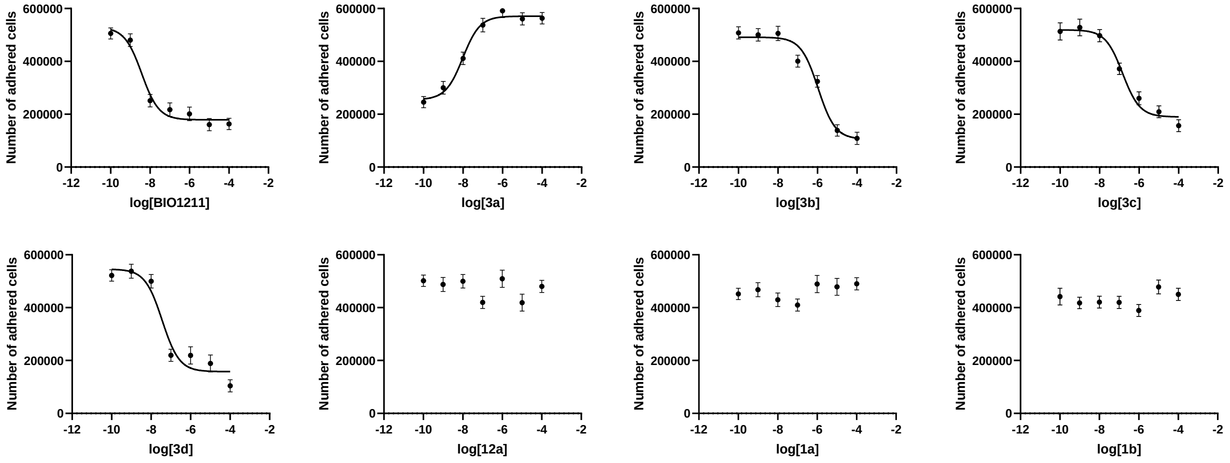


Figure S6. Concentration-response curves obtained from cell adhesion assays: evaluation of CPPs effects on $\alpha_L\beta_2$-mediated Jurkat cell adhesion to ICAM-1. Values represent the mean ± SD of three independent experiments carried out in quadruplicate.

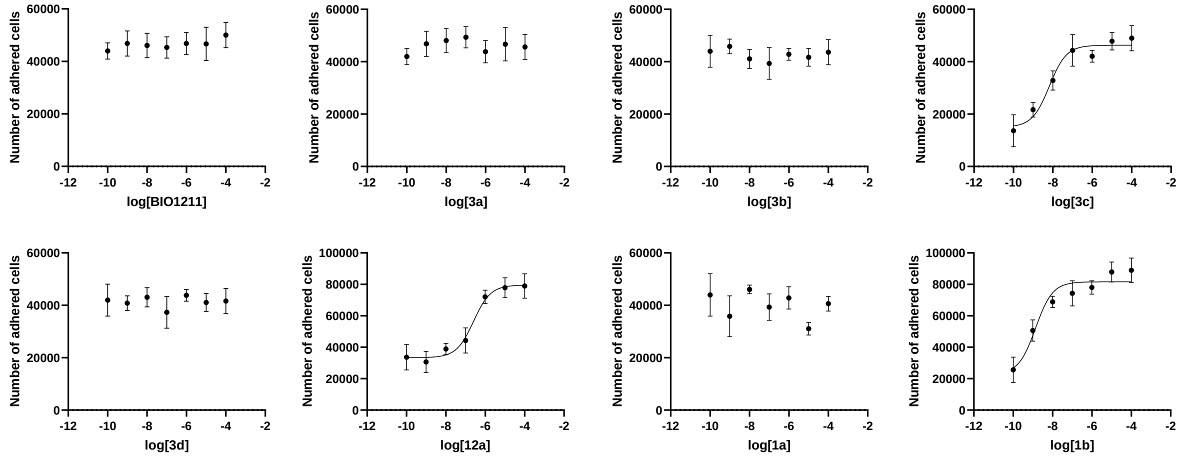


Figure S7. Concentration-response curves obtained from cell adhesion assays: evaluation of CPPs effects on $\alpha_5\beta_1$-mediated K562 cell adhesion to FN. Values represent the mean ± SD of three independent experiments carried out in quadruplicate.

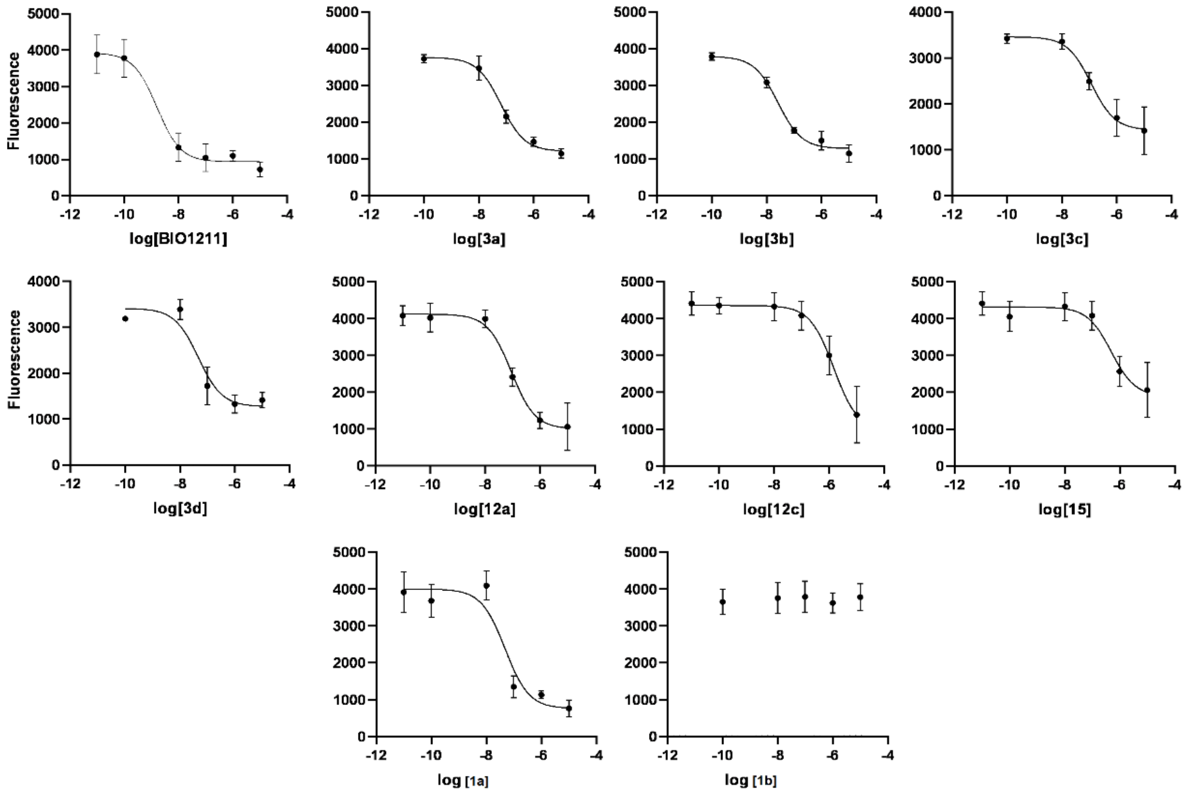


Figure S8. Binding assay curves for $\alpha_4\beta_1$/FN in presence of increasing concentrations of LDV CPPs. Values represent the mean ± SD of three independent experiments carried out in triplicate.

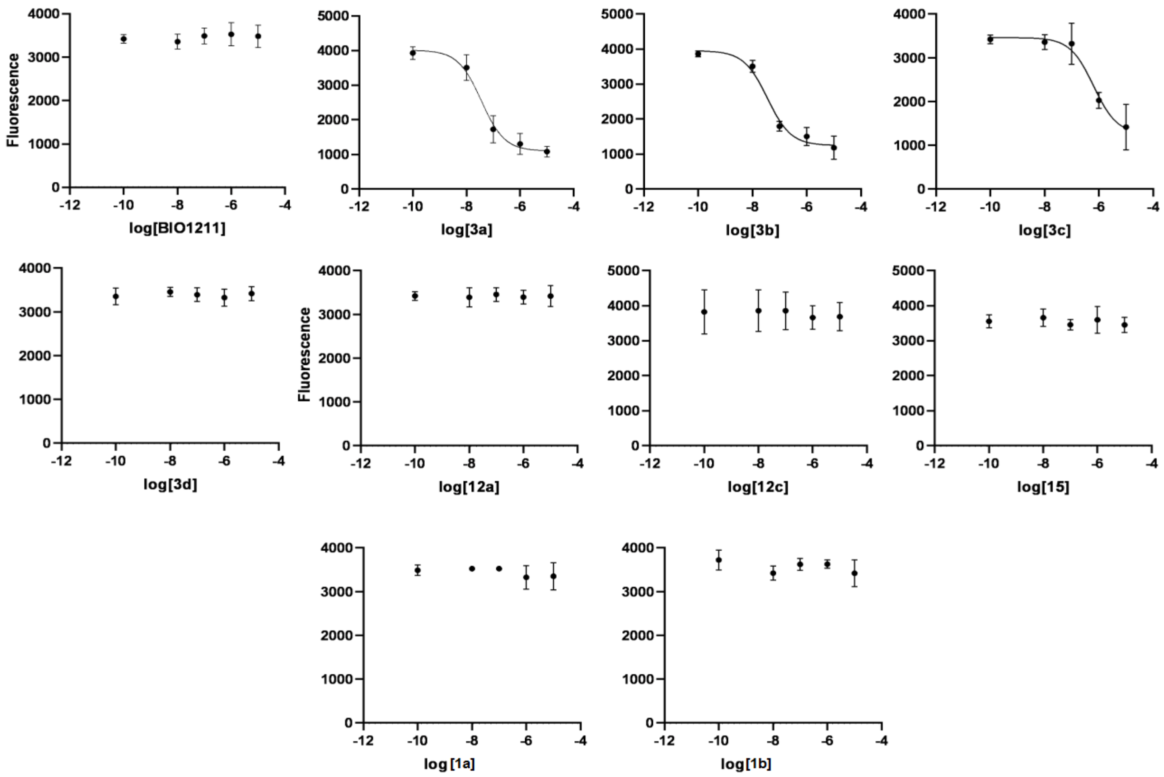


Figure S9. Binding assay curves for $\alpha_4\beta_7$/MAdCAM-1 in presence of increasing concentrations of LDV CPPs. Values represent the mean ± SD of three independent experiments carried out in triplicate.

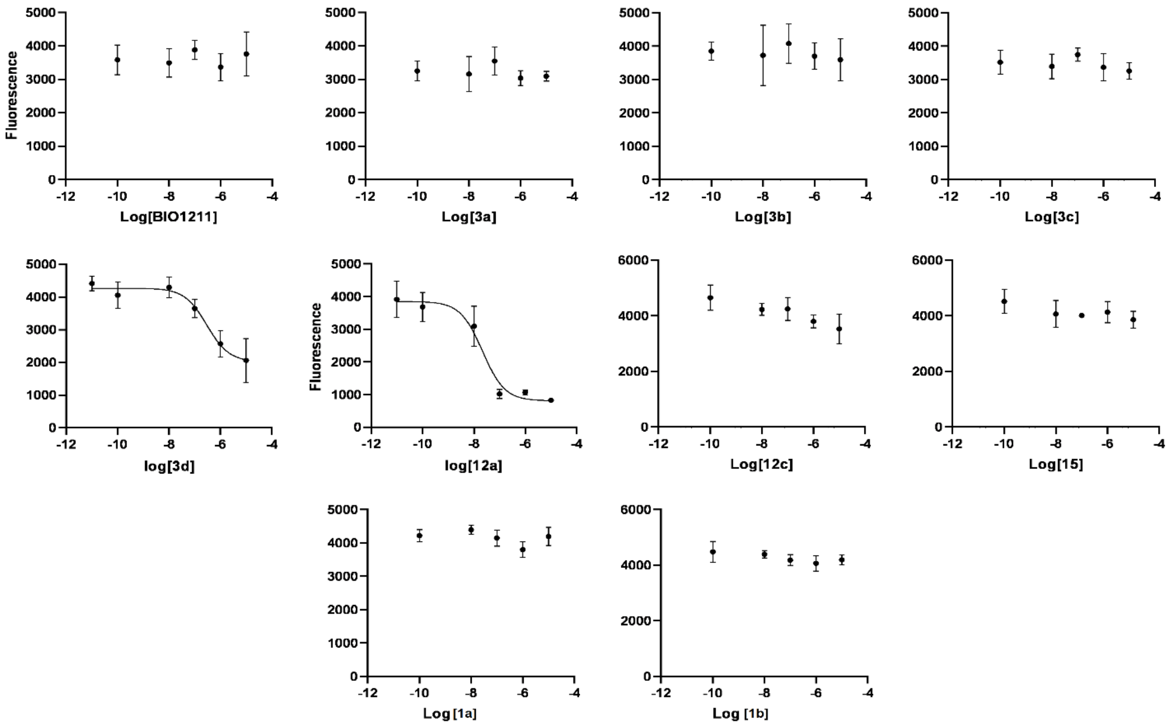


Figure S10. Binding assay curves for $\alpha_M\beta_2$/fibrinogen in presence of increasing concentrations of LDV CPPs. Values represent the mean ± SD of three independent experiments carried out in triplicate.

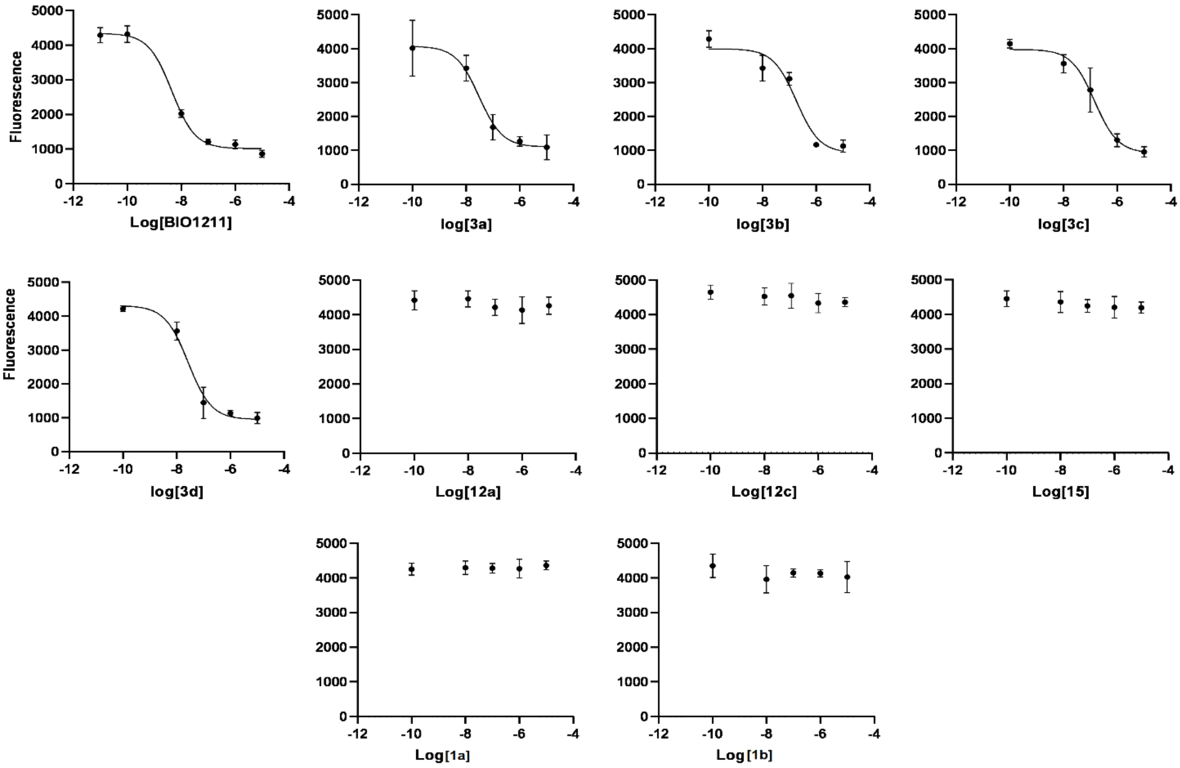


Figure S11. Binding assay curves for $\alpha_L\beta_2$/ICAM-1 in presence of increasing concentrations of LDV CPPs. Values represent the mean ± SD of three independent experiments carried out in triplicate.

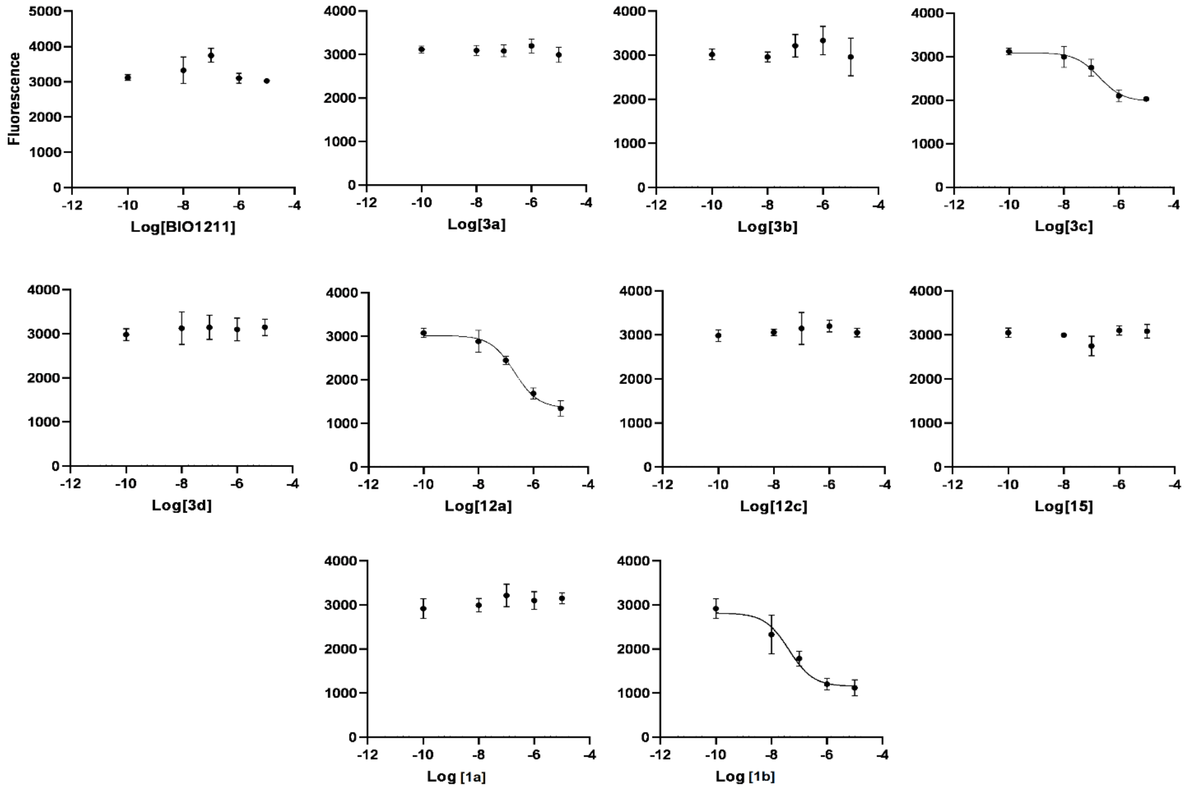


Figure S12. Binding assay curves for $\alpha_5\beta_1$/FN in presence of increasing concentrations of LDV CPPs. Values represent the mean ± SD of three independent experiments carried out in triplicate.

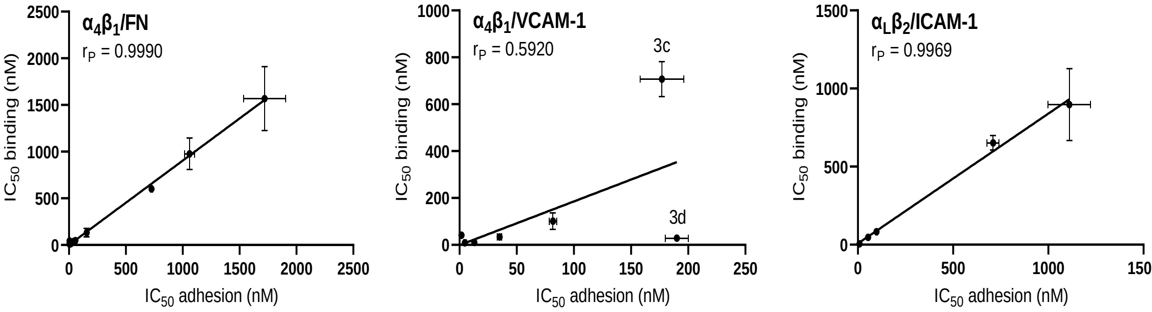


Figure S13. Relation between experimentally determined potencies (IC_{50}/EC_{50} deriving from cell adhesion assays) and ligand binding affinities (IC_{50} deriving from solid-phase binding assays) of the LDV CPPs. The IC_{50}/EC_{50} values ± S.E. are obtained by Tables 1-2. Correlation was measured using Pearson (r_P) correlation coefficient.

Table S2. Δδ/Δt values (ppb K^{-1}) of amide protons for LDV CPPs (CPPs) **3a-d**, and of the correlated **11a, 11c, 12a, 12c, 13-16** by VT NMR spectroscopy, determined at 400 MHz (**3a-d**, **11a, 11c, 12a, 12c, 13**) or at 600 MHz (**14-16**) in DMSO-d_6/H_2O (8:2) over the range 298-318 K.

CPP	**Sequence**	**Phu^{1}NH**	**AA^{2}NH**	**AA^{3}NH**	**AA^{4}NH**	**isoAsp^{5}NH**
3a	*c*[(*S*)-Phu-LDV-(*S*)-isoAsp]	-5.0	-7.0	-4.0	-0.8	-0.5
3b	*c*[(*S*)-Phu-LDV-(*R*)-isoAsp]	-4.9	-3.5	-3.7	-1.9	-2.2
3c	*c*[(*R*)-Phu-LDV-(*S*)-isoAsp]	-5.5	-5.3	-0.3	-7.9	-5.5
3d	*c*[(*R*)-Phu-LDV-(*R*)-isoAsp]	-2.1	-2.9	+0.4	-3.0	-3.5
11a	*c*[(*S*)-Phu-LDV-(*R*)-βAla^5]	-4.5	-4.1	-2.2	-0.6	-1.1
12a	*c*[(*S*)-Phu-LAV-(*S*)-isoAsp^5]	-5.4	+1.6	-4.4	-0.5	0.0
13	*c*[(*S*)-Phu-LDV-(*S*)-isoAsp(NHPr)]	-1.7	-4.8	-4.9	-0.9	-0.2
11c	*c*[(*R*)-Phu-LDV-(*R*)-βAla]	-6.0	-3.9	-0.7	-6.4	-2.5
12c	*c*[(*R*)-Phu-LAV-(*S*)-isoAsp]	-3.7	-4.2	-2.2	-6.0	-6.0
14	*c*[(*S*)-Phu-FDV-(*S*)-isoAsp]	-3.5	-4.8	-3.7	-0.8	0.0
15	*c*[(*S*)-Phu-FAV-(*S*)-isoAsp]	-4.2	-5.8	-3.8	-0.6	+0.6
16	*c*[(*R*)-Phu-LD-Phg-(*S*)-isoAsp]	-5.7	-3.9	-2.6	-4.6	-4.6

Table S3. ROESY cross peaks for **3a** in 8:2 DMSOd_6/H_2O; vs = very strong, s = strong, m = medium, w = weak.

Crosspeak	Intensity	Crosspeak	Intensity
PhuNHb-PhuMe	w	PhuArH2',6'-LeuHγ	w
PhuNHb-LeuHα	w	PhuArH2',6'-PhuMe	w
LeuNH-LeuMe	m	PhuArH2',6'-isoAspH$\beta_{2.6}$	w
LeuNH-LeuHγ	m	PhuArH2',6'-PhuH$\beta_{2.7}$	w
LeuNH-LeuH$\beta_{1.5}$	w	PhuArH2',6'-PhuH$\beta_{2.9}$	w
LeuNH-LeuH$\beta_{1.7}$	m	PhuArH2',6'-PhuHα	w
LeuNH-PhuH$\beta_{2.9}$	m	PhuArH3',5'-3,5-LeuMe	m
LeuNH-LeuHα	vs	PhuArH3',5'-3,5-LeuHγ	w
LeuNH-PhuHα	s	PhuArH3',5'-3,5-LeuHβ	w
LeuNH-PhuArH3',5'	m	PhuArH3',5'-3,5-PhuMe	w
LeuNH-AspNH	s	PhuArH3',5'-3,5-PhuHα	vs
AspNH-LeuMe	w	PhuArH3',5'-3,5-isoAspH$\beta_{2.6}$	w
AspNH-LeuHγ	w	isoAspNH-ValMe	m
AspNH-LeuH$\beta_{1.5}$	s	isoAspNH-PhuMe	w
AspNH-LeuH$\beta_{1.7}$	s	isoAspNH-ValHβ	w

AspNH-AspHβ$_{2.9}$	w	isoAspNH-isoAspHβ$_{2.6}$	m
AspNH-AspHβ$_{3.0}$	w	isoAspNH-isoAspHβ$_{2.7}$	w
AspNH-LeuHα	s	isoAspNH-ValHα	s
AspNH-ValHα	w	isoAspNH-AspHα	w
AspNH-AspHα	vs	isoAspNH-isoAspHα	s
AspNH-ValNH	s	isoAspHα-ValMe	w
PhuNH-isoAspHβ$_{2.6}$	s	isoAspHα-isoAspHβ$_{2.6}$	m
PhuNH-isoAspHβ$_{2.7}$	s	isoAspHα-isoAspHβ$_{2.7}$	w
PhuNH-PhuHβ$_{2.7}$	vs	PhuHα-PhuHβ$_{2.7}$	m
PhuNH-PhuHβ$_{2.9}$	m	PhuHα-PhuHβ$_{2.9}$	s
PhuNH-PhuHα	m	AspHα-AspHβ$_{2.9}$	m
PhuNH-isoAspHα	m	AspHα-AspHβ$_{3.0}$	m
PhuNH-isoAspNH	m	ValHα-ValMe	vs
PhuNH-PhuArH3',5'	s	ValHα-ValHβ	s
PhuNHa-LeuMe	w	ValHα-AspHβ$_{2.9}$	w
PhuNHa-LeuHβ$_{1.7}$	w	LeuHα-LeuMe	vs
PhuNHa-LeuHα	w	LeuHα-LeuHγ	m
PhuNHa-PhuHα	w	LeuHα-LeuHβ$_{1.5}$	s
PhuNHa-PhuMe	vs	LeuHα-LeuHβ$_{1.7}$	m
PhuNHa-PhuNHb	vs	PhuHβ$_{3.0}$-LeuHγ	m
ValNH-ValMe	vs	PhuHβ$_{3.0}$-LeuHβ$_{1.5}$	w
ValNH-LeuHβ$_{1.7}$	w	AspHβ$_{2.8}$-LeuHβ$_{1.5}$	w
ValNH-ValHβ	m	AspHβ$_{2.9}$-LeuMe	m
ValNH-AspHβ$_{2.9}$	m	PhuHβ$_{2.7}$-LeuHγ	w
ValNH-ValHα	s	PhuMe-LeuMe	m
ValNH-AspHα	m	PhuMe-LeuHγ	w
ValNH-isoAspHα	w	PhuMe-LeuHβ$_{1.5}$	w
ValNH-isoAspNH	s	PhuMe-LeuHβ$_{1.7}$	w
PhuArH2',6'-LeuMe	m		

Table S4. ROESY cross peaks for **3b** in 8:2 DMSOd_6/H$_2$O; vs = very strong, s = strong, m = medium, w = weak.

Crosspeak	Intensity	Crosspeak	Intensity
PhuNHb-LeuMe	w	isoAspNH-isoAspHα	m/s
PhuNHb-PhuHα	w	PhuNHa-ValMe	w
PhuNHb-PhuMe	w	PhuNHa-PhuMe	s/vs
AspNH-Val/LeuMe	w	PhuArH6-LeuMe	w
AspNH-LeuHγ/LeuHβ$_{1.5}$	s	PhuArH2',6'-LeuMe	w/m
AspNH-LeuHβ$_{1.7}$	m	PhuArH2',6'-LeuHγ/LeuHβ$_{1.7}$	w
AspNH-AspHβx2	m	PhuArH2',6'-LeuHβ$_{1.5}$	vw
AspNH-LeuHα	w	PhuArH2'6'-PhuHβ$_{2.7}$	w
AspNH-AspHα	vs	PhuArH2',6'-Phuβ$_{3.0}$	vw
AspNH-ValNH	w	PhuArH2',6'-Phuα	vw
AspNH-LeuNH	w	PhuArH3',5',3,5-LeuMe	m/w
LeuNH-LeuMe	m	PhuArH3',5',3,5-LeuHγ/LeuHβ$_{1.7}$	w
LeuNH-LeuHγ/LeuHβ$_{1.5}$	m/s	PhuArH3',5',3,5-LeuHβ$_{1.5}$	w
LeuNH-LeuHβ$_{1.7}$	m	PhuArH3',5',3,5-PhuHβ$_{2.7}$	s
LeuNH-PhuHβ$_{2.7}$	w	PhuArH3'5',3,5-PhuHβ$_{3.0}$	s
LeuNH-PhuHβ$_{3.0}$	w/m	PhuArH3'5',3,5-PhuHα	s
LeuNH-LeuHα	s	AspHα-LeuHγ/LeuHβ$_{1.7}$	w
LeuNH-PhuHα	m	AspHα-AspHβx2	vs
PhuNH-isoAspHβ$_{2.2}$	vw	AspHα-ValHα	w
PhuNH-PhuHβ$_{2.7}$	s/vs	PhuHα-PhuHβ$_{2.7}$	m
PhuNH-isoAspHβ$_{2.8}$	s/vs	PhuHα-PhuHβ$_{3.0}$	vs
PhuNH-PhuHβ$_{3.0}$	w	isoAspHα-isoAspHβ$_{2.4}$	vs
PhuNH-PhuHα	m	isoAspHα-isoAspHβ$_{2.8}$	w
PhuNH-PhuAr3',5'	m/s	isoAspHα-ValHα	w
ValNH-ValMe	s	LeuHα-LeuMe	vs
ValNH-LeuHγ/LeuHβ	w,w	LeuHα-LeuHβ$_{1.7}$/Hγ	s/vs
ValNH-ValHβ	m	LeuHα-LeuHβ$_{1.5}$	w/m
ValNH-AspHβ	m	ValHα-LeuMe	vs
ValNH-ValHα	m	ValHα-ValHβ	vs
ValNH-AspHα	m/w	PhuHβ$_{3.0}$-LeuHβ$_{1.7}$/Hγ	w
isoAspNH-ValMe	w	PhuHβ$_{3.0}$-LeuMe	w
isoAspNH-ValHβ	m/s	PhuHβ$_{2.7}$-LeuMe	w
isoAspNH-isoAspHβ$_{2.8}$	m/s	AspHβx2-LeuHβ$_{1.7}$/Hγ	w
isoAspNH-ValHα	w/m	AspHβx2-LeuHβ$_{1.5}$	w

Table S5. ROESY cross peaks for **3c** in 8:2 DMSOd_6/H$_2$O; vs = very strong, s = strong, m = medium, w = weak.

Crosspeak	Intensity	Crosspeak	Intensity
PhuNHb-LeuMe$_{0.6}$	w	PhuNHa-LeuMe$_{0.7}$	w
PhuNHb-LeuMe$_{0.7}$	m	PhuNHa-Phu(2-Me)	vs
PhuNHb-ValMe$_{0.8}$	w	PhuNHa-ValHα	w
PhuNHb-Phu(2-Me)	w	PhuArH6-LeuMe$_{0.6}$	w
PhuNHb-LeuNH	w	PhuArH6-LeuMe$_{0.7}$	w
PhuNHb-PhuNH	w	PhuArH6-Phu(2-Me)	w
PhuNHb-ValNH	w	AspNH-LeuHβ	m
PhuNH-ValMe$_{0.8}$	w	AspNH-AspHβ$_{2.6}$	m
PhuNH-isoAspHβ	vs	AspNH-AspHβ$_{2.8}$	m
PhuNH-PhuHβ	vs	AspNH-LeuHα	s
PhuNH-PhuHα	s	AspNH-PhuHα	w
PhuNH-isoAspHα	w	AspNH-AspHα	s
PhuNH-PhuArH3',5'	m	PhuArH2',6'-LeuMe$_{0.6}$	m
PhuNH-AspNH	w	PhuArH2',6'-LeuMe$_{0.7}$	w
PhuNH-ValNH	w	PhuArH2',6'-ValMe$_{0.8}$	w
isoAspNH-ValMe$_{0.8}$	w	PhuArH2',6'-LeuHγ	m
isoAspNH-ValHβ	m	PhuArH2',6'-PhuHα	w
isoAspNH-isoAspHβ	w	PhuArH3',5'-LeuMe$_{0.6}$	m
isoAspNH-ValHα	s	PhuArH3',5'-LeuMe$_{0.7}$	w
isoAspNH-isoAspHα	m	PhuArH3',5'-ValMe$_{0.8}$	w
isoAspNH-AspHα	w	PhuArH3',5'-LeuHγ	m
isoAspNH-LeuHα	w	PhuArH3',5'-LeuHβ	w
LeuNH-LeuMe$_{0.6}$	w	PhuArH3',5'-ValHβ	w
LeuNH-LeuMe$_{0.7}$	w	PhuArH3',5'-isoAspHβ	w
LeuNH-LeuHγ	s	PhuArH3',5'-LeuHα	w
LeuNH-LeuHβ	vs	PhuArH3',5'-isoAspHα	w
LeuNH-AspHβ$_{2.6}$	m	PhuArH3',5'-PhuHα	s
LeuNH-PhuHβ	m	AspHα-ValMe$_{0.7}$	w
LeuNH-ValHα	w	AspHα-AspHβ$_{2.6}$	s
LeuNH-LeuHα	s	AspHα-AspHβ$_{2.8}$	s
LeuNH-isoAspHα	w	AspHα-ValHα	w
LeuNH-PhuHα	vs	AspHα-LeuHα	w
LeuNH-PhuArH3',5'	w	PhuHα-PhuHβ	vs
LeuNH-AspNH	s	isoAspHα-isoAspHβ	m
ValNH-ValMe$_{0.8}$	m	LeuHα-LeuMe$_{0.6}$	vs
ValNH-ValMe$_{0.9}$	w	LeuHα-LeuMe$_{0.7}$	m
ValNH-ValHβ	m	LeuHα-LeuHγ	m

ValNH-AspHβ$_{2.6}$	w	LeuHα-LeuHβ	vs
ValNH-AspHβ$_{2.8}$	m	ValHα-ValMe$_{0.7}$	s
ValNH-ValHα	vs	ValHα-ValMe$_{0.8}$	s
ValNH-AspHα	s	ValHα-ValHβ	m
ValNH-AspNH	w	PhuHβ-LeuHγ	w

Table S6. ROESY cross peaks for **3d** in 8:2 DMSOd$_6$/H$_2$O; vs = very strong, s = strong, m = medium, w = weak.

Crosspeak	Intensity	Crosspeak	Intensity
PhuNHb-LeuMe$_{0.6}$	w	PhuNHa-PhuMe	vs
PhuNHb-LeuMe$_{0.7}$	w	PhuNHa-PhuArH6	vs
ValNH-ValMe	s	AspNH-LeuMe$_{0.6}$	m
ValNH-ValHβ	w	AspNH-LeuMe$_{0.7}$	m
ValNH-AspHβ$_{2.8}$	w	AspNH-LeuHγ	m
ValNH-ValHα	s	AspNH-AspHβ$_{2.5}$	w
ValNH-AspHα	s	AspNH-AspHβ$_{2.8}$	w
ValNH-AspNH	w	AspNH-LeuHα	w
ValNH-isoAspNH	w	AspNH-AspHα	m
PhuNH-LeuHγ	w	PhuArH2',6'-PhuHα	w
PhuNH-isoAspHβ$_{2.7}$	s	PhuArH3',5'-LeuMe$_{0.6}$	w
PhuNH-PhuHβ$_{2.8}$	s	PhuArH3',5'-LeuHγ	w
PhuNH-PhuHα	m	PhuArH3',5'-LeuHα	w
PhuNH-PhuArH3',5'	w	PhuArH3',5'-PhuHα	s
PhuNH-PhuNHa	w	AspHα-LeuHγ	w
PhuNH-LeuNH	m	AspHα-AspHβ$_{2.5}$	s
PhuNH-isoAspNH	m/s	AspHα-AspHβ$_{2.8}$	m
isoAspNH-ValHβ	w	AspHα-LeuHα	w
isoAspNH-isoAspHβ$_{2.6}$	m	isoAspHα-isoAspHβ$_{2.6}$	m
isoAspNH-ValHα	w	isoAspHα-isoAspHβ$_{2.7}$	m
isoAspNH-isoAspHα	m	PhuHα-PhuArH3',5'	m
isoAspNH-ValMe	w	PhuHα-LeuNH	vs
isoAspNH-AspNH	w	LeuHα-LeuMe$_{0.6}$	s
isoAspNH-LeuNH	w	LeuHα-LeuMe$_{0.7}$	w
LeuNH-LeuMe$_{0.6}$	w	LeuHα-LeuHγ	w
LeuNH-LeuMe$_{0.7}$	w	LeuHα-LeuHβ$_{1.4}$	m
LeuNH-LeuHγ	m	LeuHα-ValHα	w
LeuNH-LeuHβ$_{1.3}$	m	ValHα-ValMe	vs
LeuNH-LeuHα	m	LeuHβ$_{1.4}$-LeuMe$_{0.6}$	w
LeuNH-PhuHα	vs	LeuHβ$_{1.4}$-LeuMe$_{0.7}$	w
LeuNH-AspNH	s	LeuHβ$_{1.3}$-LeuMe$_{0.6}$	w
PhuNHa-PhuNHb	vs		

Table S7. ROESY cross peaks for **12a** in 8:2 DMSOd_6/H_2O; vs = very strong, s = strong, m = medium, w = weak.

Crosspeak	Intensity	Crosspeak	Intensity
LeuNH-LeuMe$_{0.5}$	m	PhuArH3',5'- LeuHβ$_{1.8}$	w
LeuNH-LeuHα	s	PhuArH3',5'-PhuHβ$_{2.6}$	vs
LeuNH-PheHα	vs	PhuArH3',5'- PhuHβ$_{2.8}$	s
PhuNHb-LeuMe$_{0.7}$	w	PhuArH3',5'- PhuHα	vs
PhuNH-ValMe$_{1.1}$	w	isoAspNH-isoAspHα	s
PhuNH-isoAspHβ$_{2.6}$	m	isoAspNH-ValHα	m
PhuNH-PheHβ$_{2.8}$	s	ValHα-ValMe$_{0.9}$	s
PhuNH-isoAspHβ$_{2.9}$	s	ValHα-ValMe$_{1.1}$	m
PhuNH-PhuHα	m	ValHα-ValHβ	vs
AlaNH-LeuHβ$_{1.2}$	s	isoAspHα-isoAspHβ$_{2.6}$	m
AlaNH-AlaMe	s	isoAspHα-isoAspHβ$_{2.9}$	vs
AlaNH-LeuHβ$_{1.8}$	s	PheHα-PheHβ$_{2.6}$	m
AlaNH-AlaHα	m	PheHα-PheHβ$_{2.8}$	s
AnaNH-ValNH	m	LeuHα-LeuMe$_{0.5}$	vs
PhuNHa-Phu(2-Me)	vs	LeuHα-LeuMe$_{0.7}$	m
ValNH-ValMe$_{0.9}$	m	LeuHα-LeuHβ$_{1.2}$	vs
ValNH-ValMe$_{1.1}$	s	PhuHβ$_{2.8}$-LeuHβ$_{1.2}$	w
ValNH-AlaHα	m	LeuHβ$_{1.8}$-LeuMe$_{0.5}$	s
ValNH-ValHα	m	LeuHβ$_{1.8}$-LeuMe$_{0.7}$	s
ValNH-isoAspNH	vs	LeuHβ$_{1.8}$-LeuHγ	w
PhuArH2',6'-LeuMe$_{0.5}$	m	AlaMe-LeuMe$_{0.5}$	w
PhuArH2',6'-LeuMe$_{0.7}$	s	AlaMe-LeuMe$_{0.7}$	w
PhuArH3',5'-LeuMe$_{0.5}$	m	LeuHβ$_{1.2}$-LeuMe$_{0.5}$	m
PhuArH3',5'-LeuMe$_{0.7}$	w	LeuHβ$_{1.2}$-LeuMe$_{0.7}$	s
PhuArH3',5'-LeuHβ$_{1.2}$	w	LeuHβ$_{1.2}$-LeuHγ	m

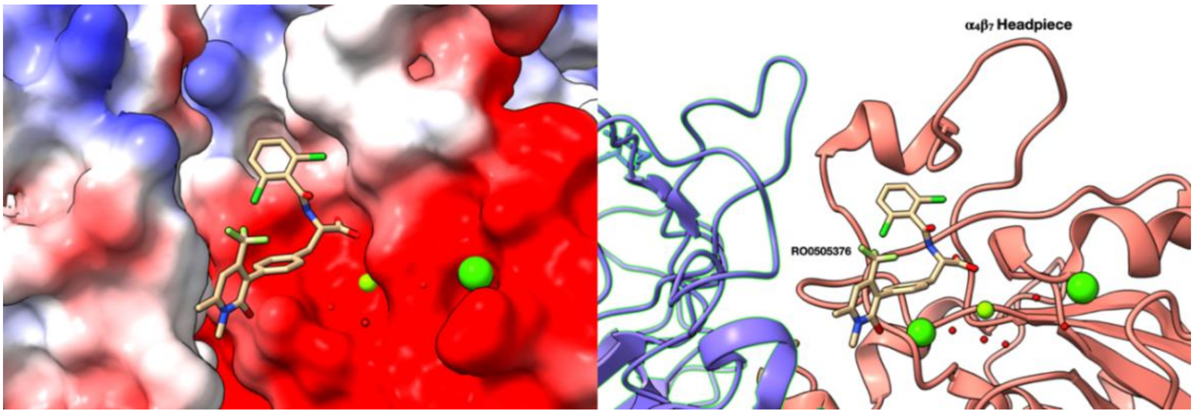


Figure S14. PDB 3V4V, crystal structure of $\alpha_4\beta_7$ headpiece complexed with RO0505376; left, the integrin is represented as molecular solvent-excluded surfaces (SES) and colored using Coulombic electrostatic potential (ESP). RO0505376 is rendered in sticks, while metal ions were rendered as green spheres.

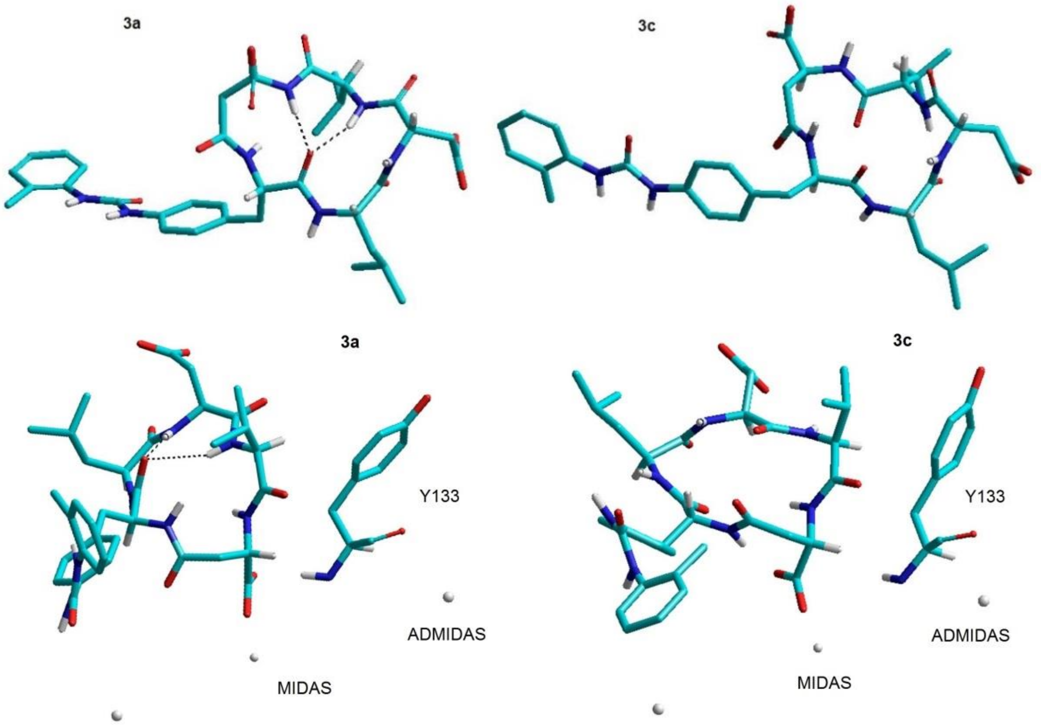


Figure S15. In-solution (top) and receptor-bound (bottom) structures of **3a**, **3c** showing hydrogen-bonded secondary structure elements (dotted lines). The structures differ only by the inversion of stereochemistry at Phu^1. β_1: Tyr^{133}and cations of the adhesion sites are shown for better comparison.

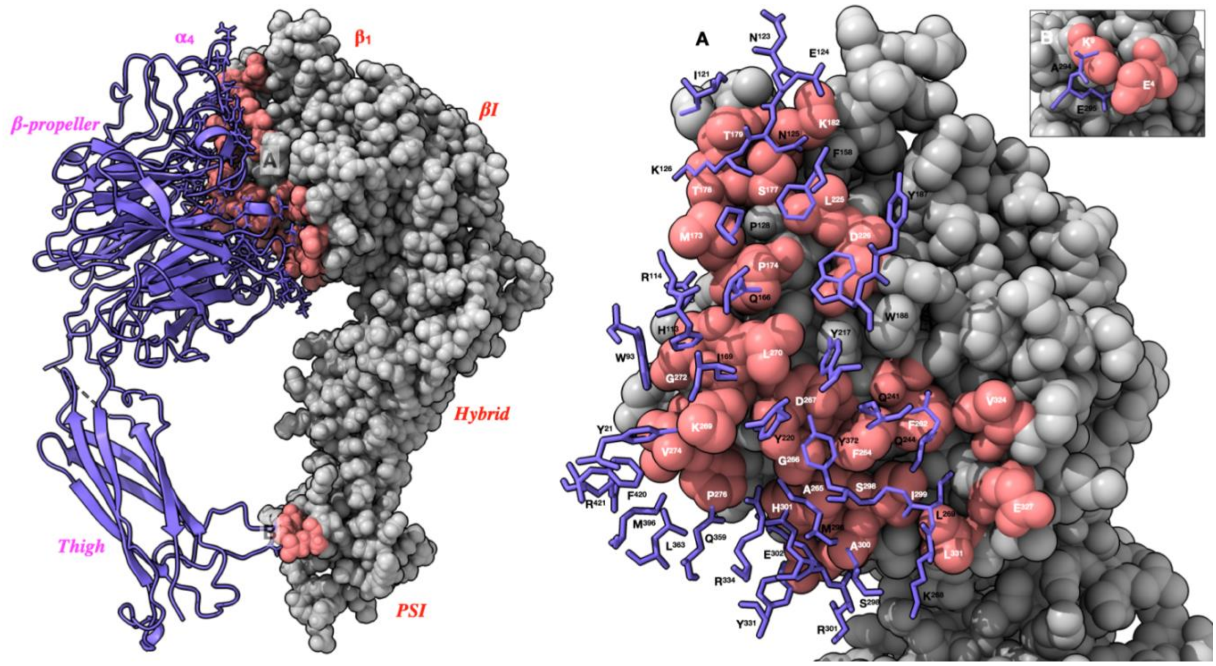


Figure S16. Left. Rendering of $\alpha_4\beta_1$ integrin model, with evidenced the β-propeller (α_4), Thigh (α_4), βI (β_1), Hybrid (β_1) and PSI (plexin-semaphorin-integrin, β_1) domains. The α_4 subunit is colored in blue and represented as solid ribbon, with key residues at the α_4/β_1 interface in stick. The β_1 subunit is represented as gray CPK, with key residues at the α_4/β_1 interface in pink. (A) Close look of the interface between the β-propeller (α_4) and βI (β_1) domains. The α_4 residues are shown in stick representation and colored in blue (black labels), while the β_1 residues are represented in CPK, with key residues colored in pink (white labels). (B) Particular of the interface between Thigh (α_4) and PSI (β_1) domains. Molecular graphics and analyses performed with UCSF ChimeraX.

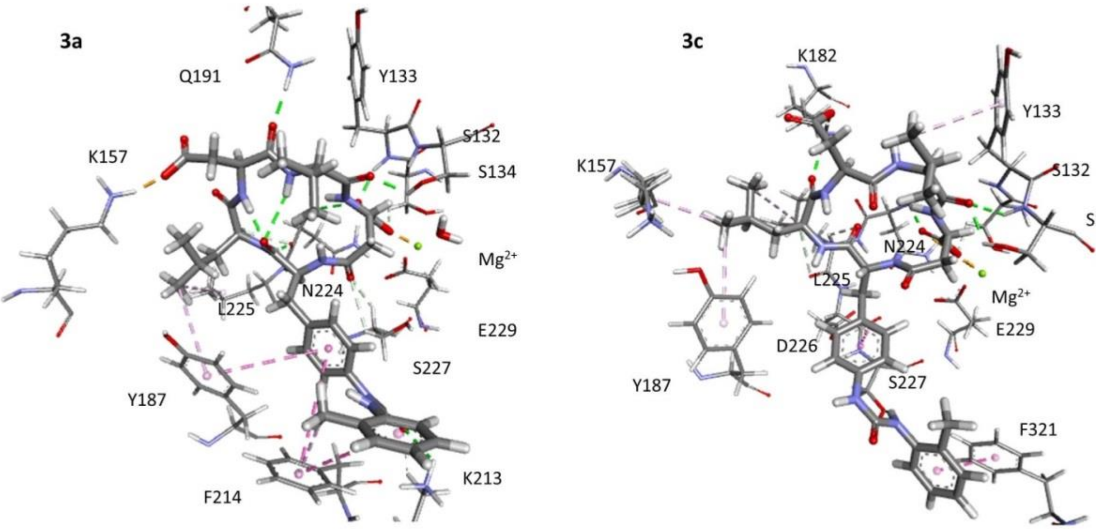


Figure S17. The agonist **3a**, c[(*S*)-Phu-LDV-(*S*)-isoAsp], makes many interactions with α_4 subunit. Phu^{1}MePh group interacts with α_4: F^{214} (π-π stacking), Phu^{1}Ph interacts with α_4: F^{214} (π-π T-shaped interaction), and with α_4: Y^{187} (π-π stacking). Urea C=O makes a hydrogen bond with α_4: K^{213} (2.31 Å). Phu^1 C=O establishes intramolecular hydrogen bonds with Asp^3 NH (1.72 Å) and Val^4 NH (2.30 Å). The hydrophobic side chain of Leu^2 is included within a cavity delimited by α_4: K^{157}, β_1: L^{225}, and α_4: Y^{187}, while Leu^{2}NH group makes a hydrogen bond with β_1: N^{224}O (1.84 Å). Asp^{3}COO$^-$ makes a salt bridge with α_4: K^{157}Nζ^+ (1.63 Å), while Asp^3 C=O gives rise to a hydrogen bond with β_1: Q^{191} CONH$_2$ (1.78 Å). Val^4 C=O is hydrogen bonded to β_1: S^{134} OH (1.97 Å). The isopropyl of Val^4 is pseudo axial, therefore it does not get in touch with any residues of β_1 subunit. Finally, isoAsp^5 COO$^-$ is coordinated to Mg^{2+} in MIDAS, with Mg^{2+} that exhibits a slightly distorted coordination geometry (RMSD 0.12). Two hydrogen bonds are formed with water (W) W^1 (2.86 Å) and W^3 (2.67 Å), a hydrogen bond with β_1: Y^{133} NH (1.97 Å) and a hydrogen bond with β_1: N^{224} NH (2.75 Å).

The antagonist **3c**, c[(*R*)-Phu-LDV-(*S*)-isoAsp], differs from the all *S*-configured **3a** for the inverted stereochemistry at Phu^1. The most noticeable effect is represented by a greater interaction with β_1 subunit. Phu^{1}MePh makes a π-π stacking interaction with β_1: F^{321}. Urea NH makes a hydrogen bond with β_1: S^{227} OG (2.11 Å), and Phu^{1}Ph interacts with the peptide bond between β_1: D^{226} and β_1: S^{227} (amide-π stacking). The side chain of Leu^2 is included within the cavity formed by α_4: K^{157}, α_4: Y^{187}, β_1: L^{225}. Leu^2 C=O makes an hydrogen bond with β_1: K^{182} NHζ^+ (1.77 Å), which in turn interacts with Asp^3 COO$^-$ by a salt bridge (1.62 Å). Val^4 C=O makes two hydrogen bonds with β_1: S^{134} OH (1.97 Å) and β_1: S^{134} NH (1.52 Å). In contrast to **3a**, the isopropyl of Val^4 is in contact with β_1: Tyr^{133} (π-alkyl, 4.76 Å). Finally, isoAsp^{5}COO$^-$ is coordinated with Mg^{2+}. In this complex, the cation shows a distorted octahedral coordination geometry (RMSD 0.17); isoAsp^{5}COO$^-$ makes also other interacts: with β_1: S^{132} (O-HC, 2.46 Å), two hydrogen bonds with W^1 (2.60 Å) and W^3 (2.66 Å), hydrogen bond with β_1: N^{224} NH (1.92 Å).

Molecular graphics and detailed analysis of the interactions between ligands and integrin $\alpha_4\beta_1$ receptor model. The following figures show the best binding conformation of the ligands within $\alpha_4\beta_1$ integrin binding site, represented as molecular solvent-excluded surfaces (partially transparent), colored using Coulombic electrostatic potential (ESP), with default coloring ranging from red for negative potential through white to blue for positive potential. Ligands are rendered in sticks, while metal ions belonging to MIDAS and ADMIDAS are rendered as green colored spheres. Key receptor residues in thick sticks. Molecular graphics and analyses were performed with Biovia Discovery Studio visualizer.

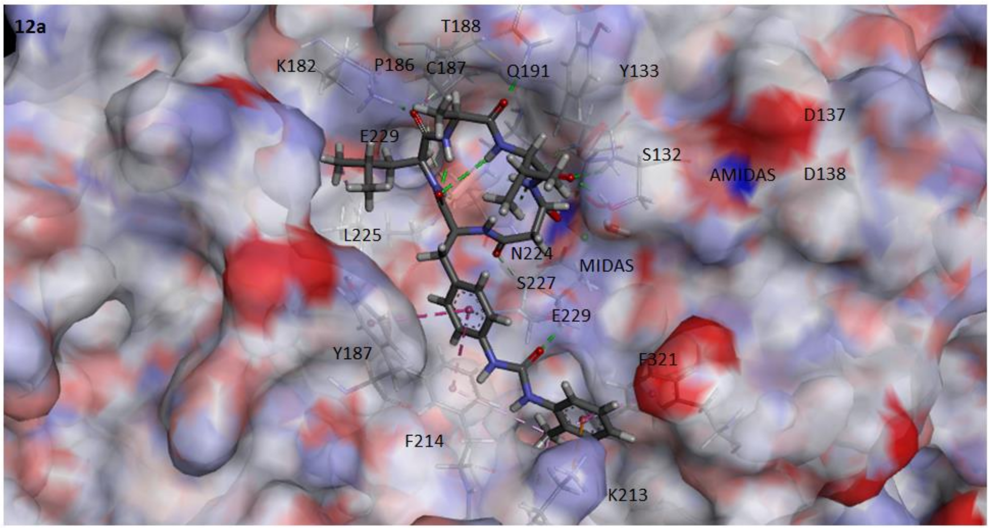


Figure S18. The CPPs **12a**, *c*[(*S*)-Phu-LAV-(*S*)-isoAsp], maintains the same stereochemistry as **3a**, but Asp^3 is replaced by Ala^3. Phu^{1}MePh group is in contact with β_1: F^{321} through a π-π stacking interaction and α_4: K^{213}NHζ^+ through a π-cation interaction (2.12 Å). Urea C=O makes a conventional hydrogen bond with β_1: S^{227}OH (2.21 Å). Phu^{1}Ph interacts with α_4: F^{214} through π-π T-shaped interaction and with α_4: Y^{187} through a stacked π-π interaction. Phu^1 C=O makes two intramolecular hydrogen bonds with Ala^3 NH (1.78 Å) and Val^4 NH (2.91 Å). Leu^2 NH give rise to a hydrogen bond with β_1: N^{224}O (1.76 Å), while Leu^2 C=O makes a hydrogen with β_1: K^{182}NHζ^+ (2.01 Å). In addition to the intramolecular bond with Phu^1 C=O, Ala^3 shows an additional hydrogen bond between Ala^3 C=O and β_1: N^{191}Hε (1.89 Å). The methyl of Val^3 nicely packs against methyl of β_1: T^{188}. Val^4 C=O makes two hydrogen bonds, with β_1: S^{134}NH (2.95 Å) and β_1: S^{134}OH (1.69 Å). isoAsp^5 residue coordinates Mg^{2+} (RMSD 0.11). Furthermore, isoAsp^{5}COO$^-$ is involved in a hydrogen bonds network with W^1 (2.83 Å), W^3 (2.72 Å), β1: Y^{133} NH (2.45 Å), β_1: N^{224} NH (2.08 Å) and β_1: N^{224} NH (2.55 Å).

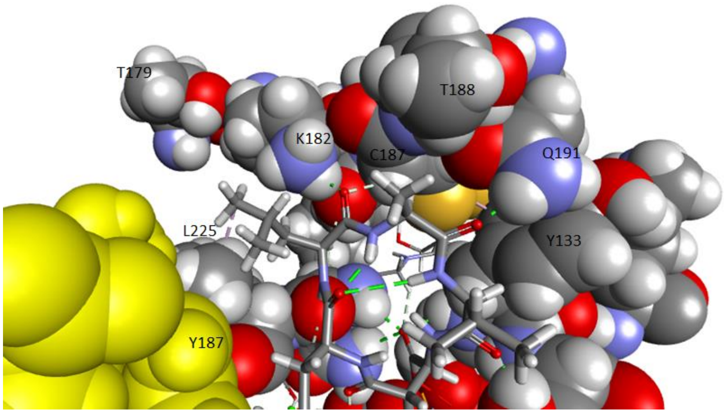


Figure S19. Detail of the upper region of the binding site of $\alpha_4\beta_1$ integrin hosting **12a**, *c*[(*S*)-Phu-LAV-(*S*)-isoAsp]. Receptor residues in contact with the cyclopeptide are rendered as CPK models; residues of α_1 subunit are highlighted in yellow. The sequence A^{181}-K-L-R-N-P-C-T^{188} of the β_1 subunit is replaced by S^{191}-K-L-R-H-P-C-P^{198} in β_7 subunit. In the $\alpha_4\beta_1$-**12a** complex, the side chain of Thr^{188} of β_1 subunit invades the space of Val^3, so that the methyl of the latter nicely packs against the methyl of Thr^{188}. In the β_7 subunits, this residue is mutated for Pro^{198}.

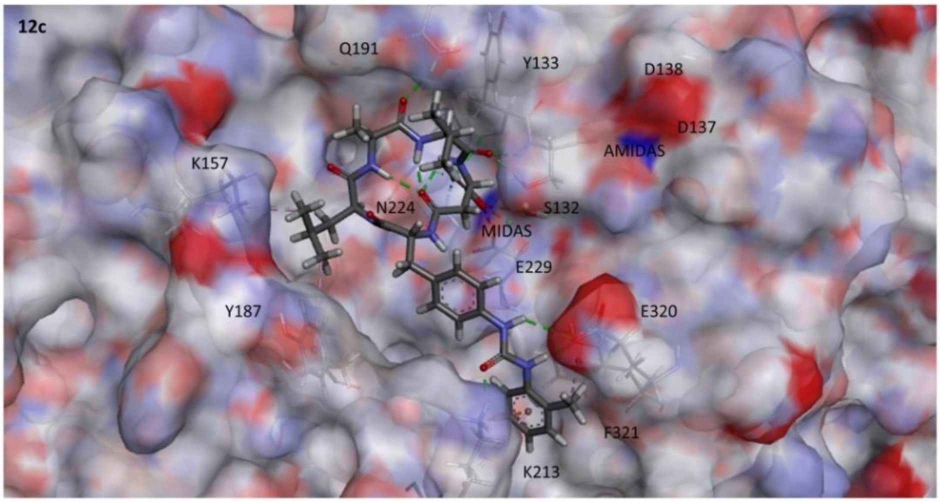


Figure S20. The CPPs **12c**, *c*[(*R*)-Phu-LAV-(*S*)-isoAsp], is an analogue of **3c**, sharing the same stereochemistry array, in which Asp^3 is replaced by Ala^3. MePh group of Phu^1 is in contact with α_4: K^{213} (π-cation interaction with NHζ^+, 3.76 Å, plus π-alkyl, 3.76 Å) and with β_1: F^{321} (π-alkyl, 4.37 Å). Urea C=O is hydrogen bonded to α_4: K^{213}NHζ^+ (1.84 Å), while both urea NHs are hydrogen bonded to β_1: Glu^{320}COO$^-$ (1.94 Å, 2.90 Å). The side chain of Leu^2 is inserted within the cavity delimited by β_1: L^{225}, α_4: K^{157}, α_4: Y^{187}, making hydrophobic interactions. Ala^3 NH is involved in a intramolecular hydrogen bond with isoAsp C=O (2.03 Å), Ala^3 C=O makes a hydrogen bond with β_1: N^{191} CONH$_2$ (2.19 Å), while Ala^{3}methyl is close to β_1: T^{188}methyl (alkyl interaction). Val^4 NH is involved in an intramolecular hydrogen bond with isoAsp C=O (1.81 Å), Val^4 C=O is hydrogen bonded to β_1: S^{134}NH (2.79 Å), but Val^4 isopropyl is directed perpendicularly above the molecular plane, therefore making very little contacts with any residues of β_1 subunit. Finally, isoAsp^5 COO$^-$ is coordinated to Mg^{2+} in MIDAS.

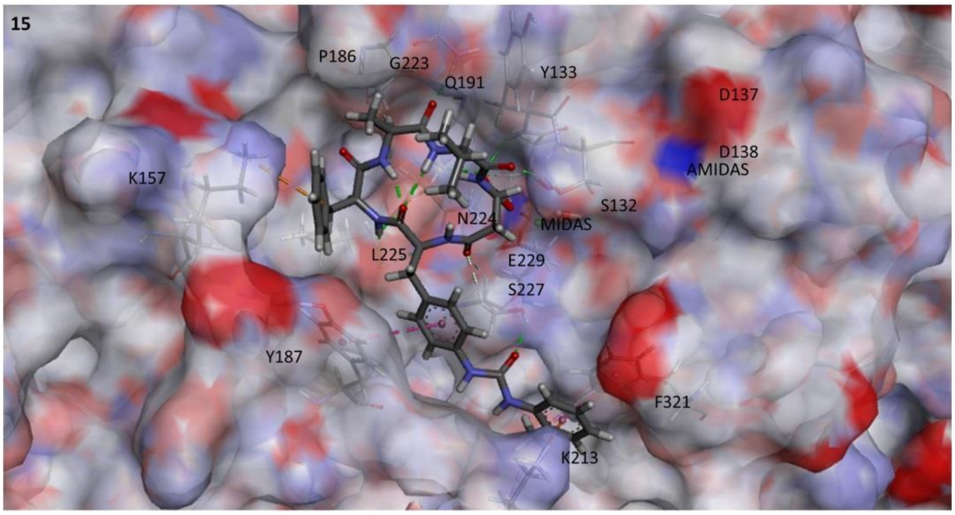
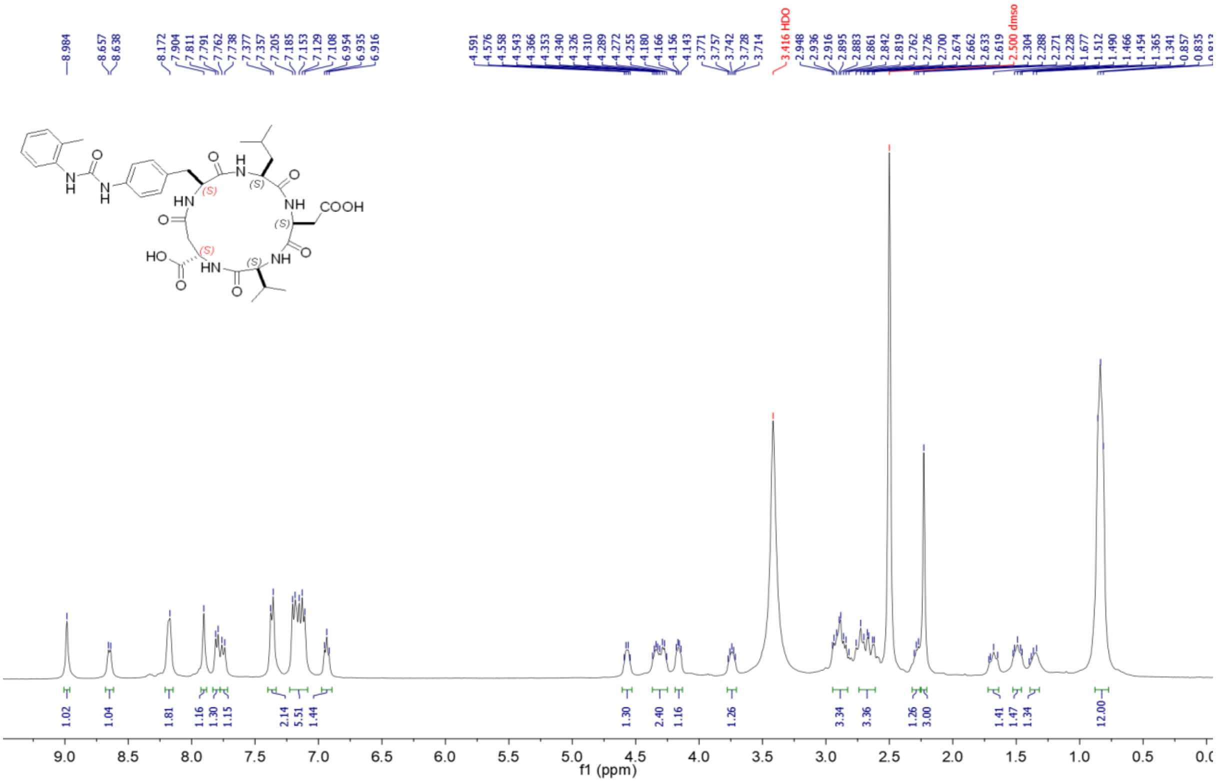


Figure S21. The CPP **15** was obtained from **3a** by replacing Leu^2 with Phe^2 and Asp^3 with Ala^3. The Phu^1 MePh group is located between the residues β_1: F^{321} and α_4: K^{213}, with which it interacts giving rise to a π-π stacking interaction and a π-cation interaction (3.19 Å). Urea C=O makes a hydrogen bond with β_1: S^{227}OH (2.05 Å), while Phu^{1}Ph makes a π-π stacking interactions with α_4: Y^{187}. Phu^1 C=O makes two hydrogen bonds with Ala^3 NH (1.72 Å) and Val^4 NH (2.15 Å). Phe^2 NH makes two hydrogen bonds with β_1: N^{224}O (2.87 Å) and with β_1: L^{225}O (2.14 Å), while Phe^2 Ph is hold in position by a π-cation interaction with β_1: K^{182}Nζ^+ (4.01 Å). In addition to the intramolecular hydrogen bonds with Phu^1, Ala^3 residue makes a hydrogen bond between C=O and β_1: Gln^{191}Hε (1.89 Å), while the Val^4 C=O forms a hydrogen bond with β_1: S^{134}OH (1.85 Å). The carboxyl group of the isoAsp^5 residue makes two water-hydrogen bonds with W^1 (2.84 Å) and W^3 (2.70 Å), a hydrogen bond with β_1: Y^{133}NH (2.12 Å) and a hydrogen bond with β_1: N^{224}NH (2.30 Å). isoAsp^5 COO$^-$ coordinates the Mg^{2+} in MIDAS.

NMR spectra:



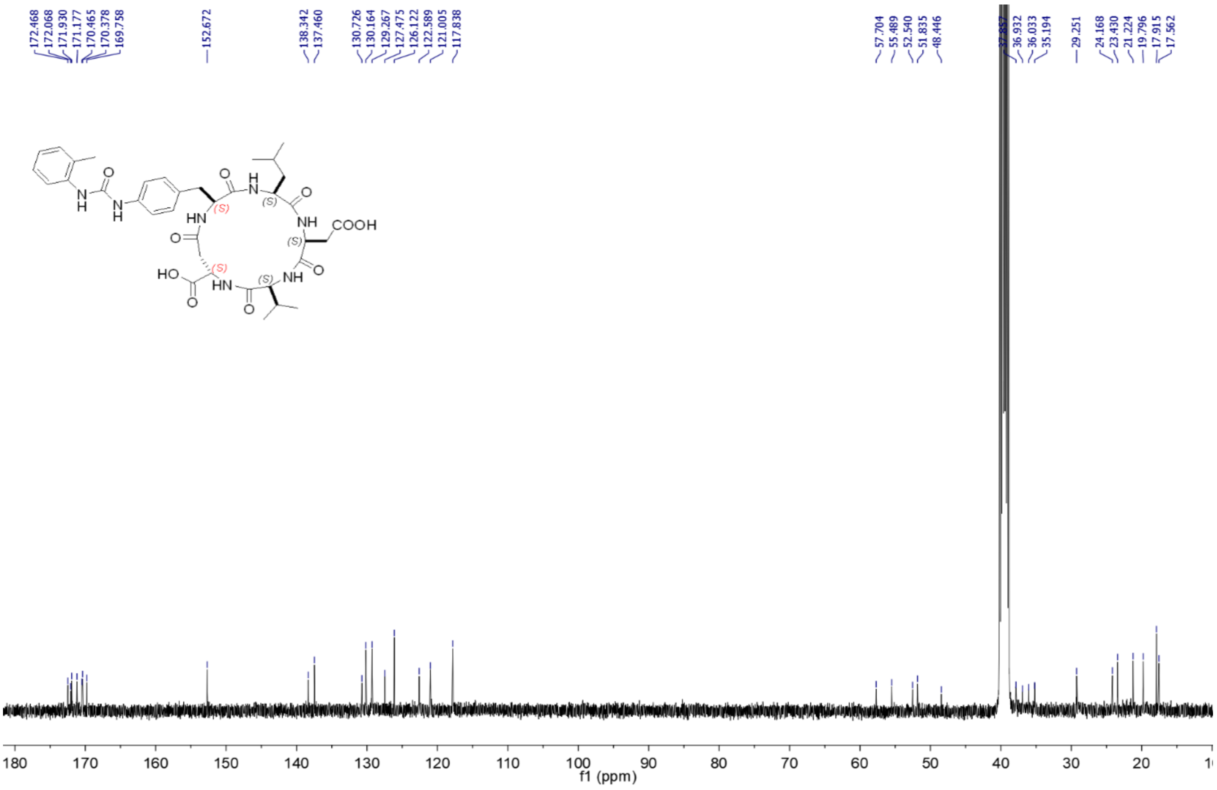


Figure S22. ^{1}H NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C NMR (DMSOd$_6$, 100 MHz) of **3a**.

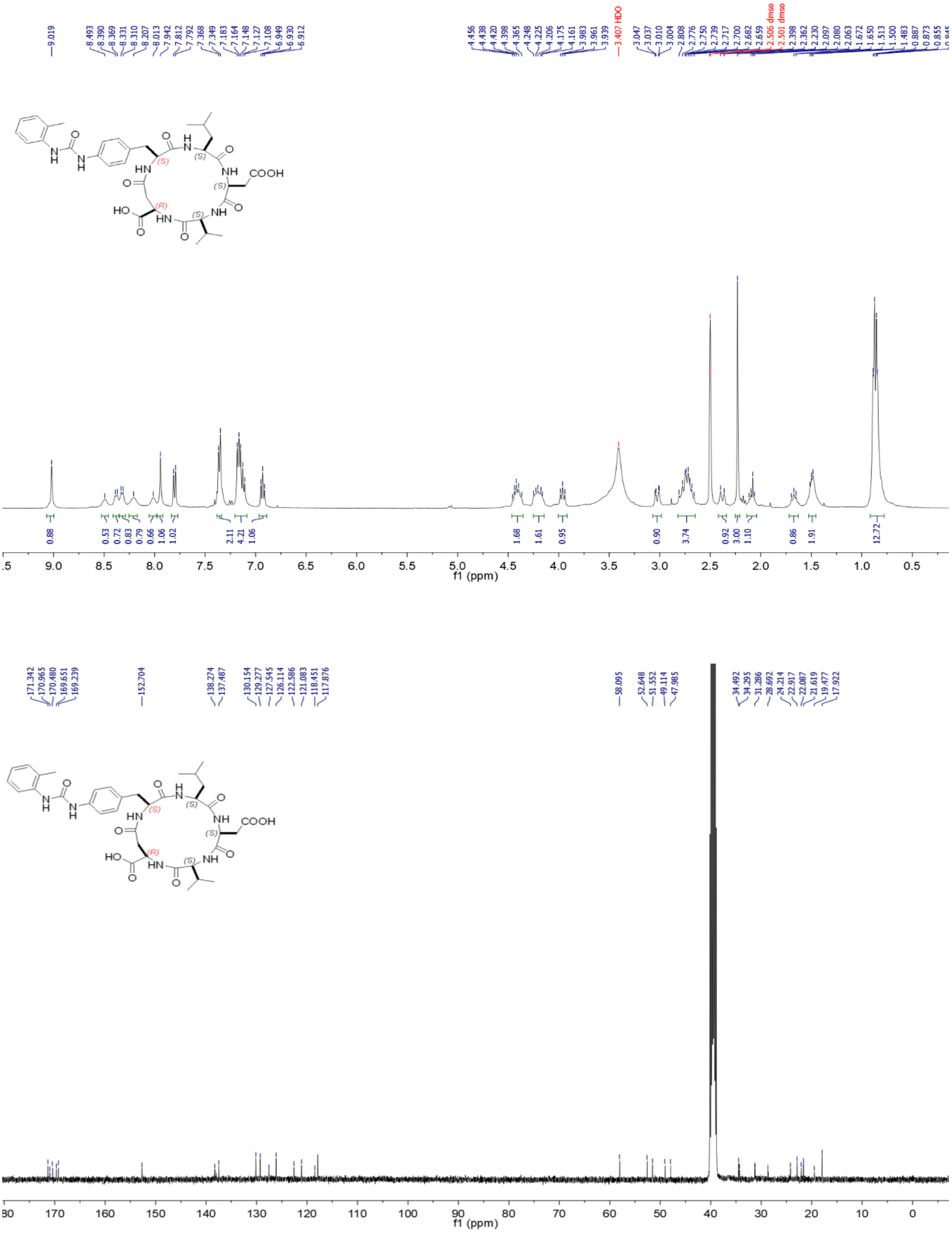


Figure S23. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **3b**.

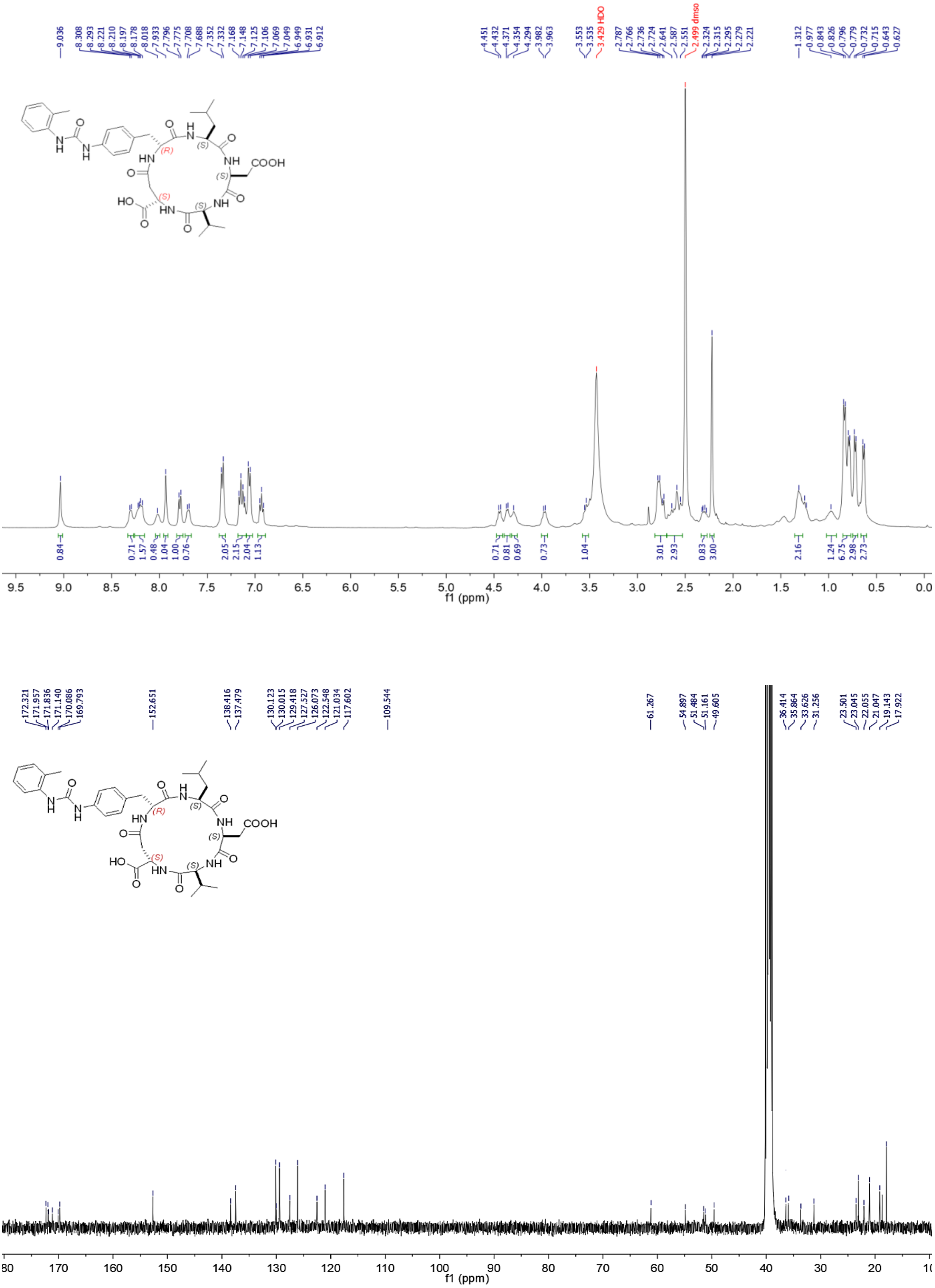


Figure S24. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **3c**.

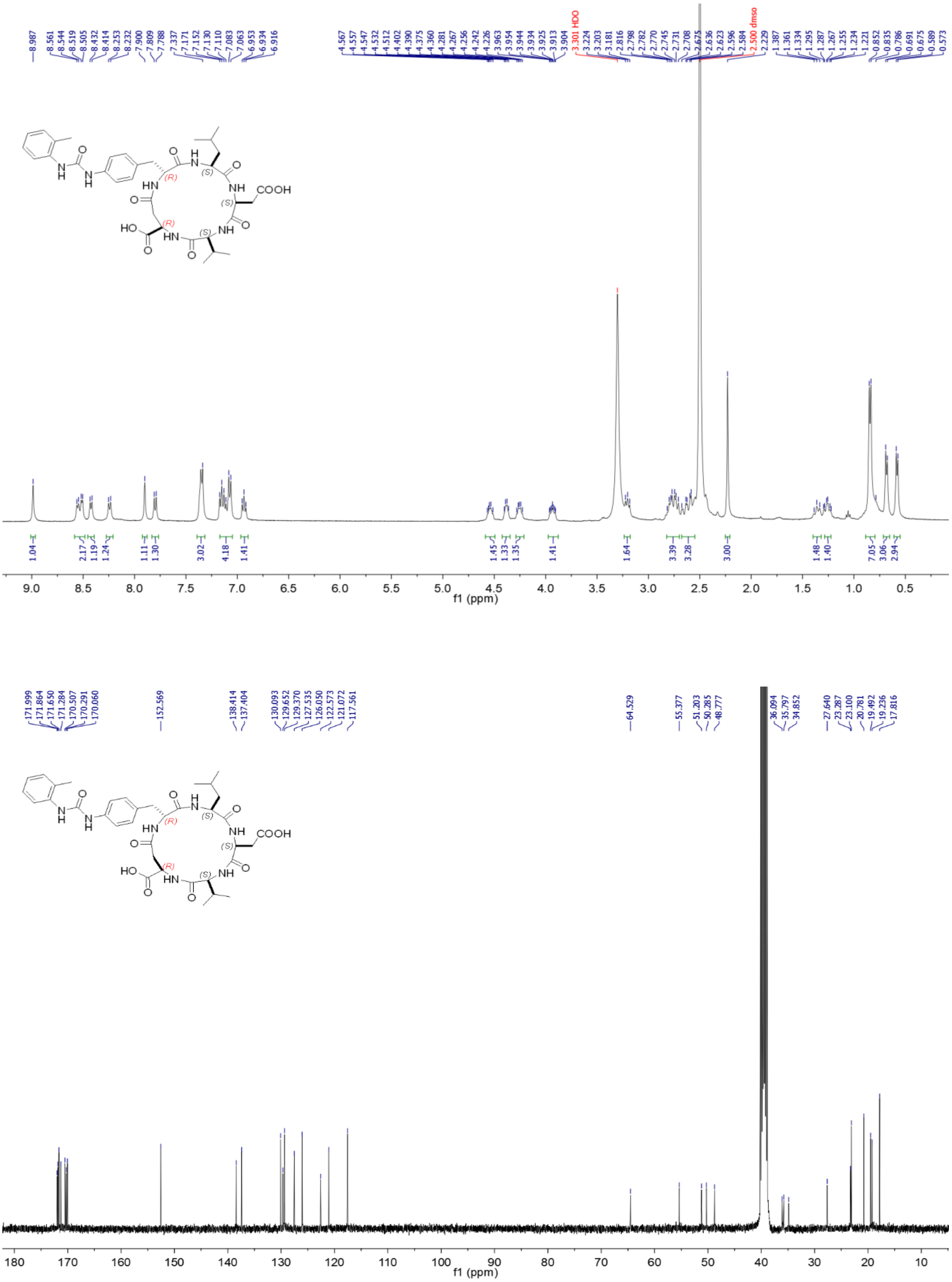


Figure S25. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **3d**.

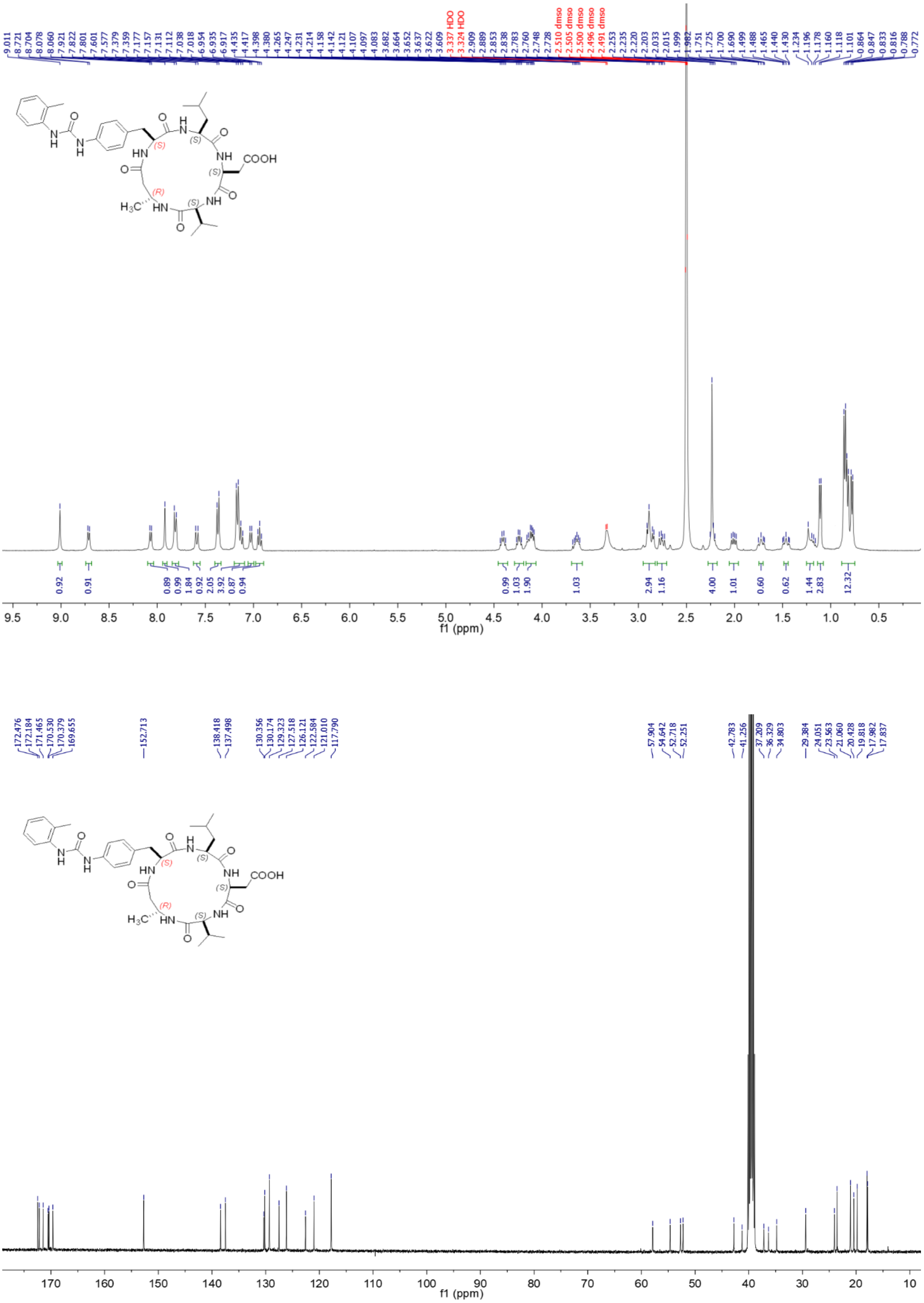


Figure S26. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **11a**.

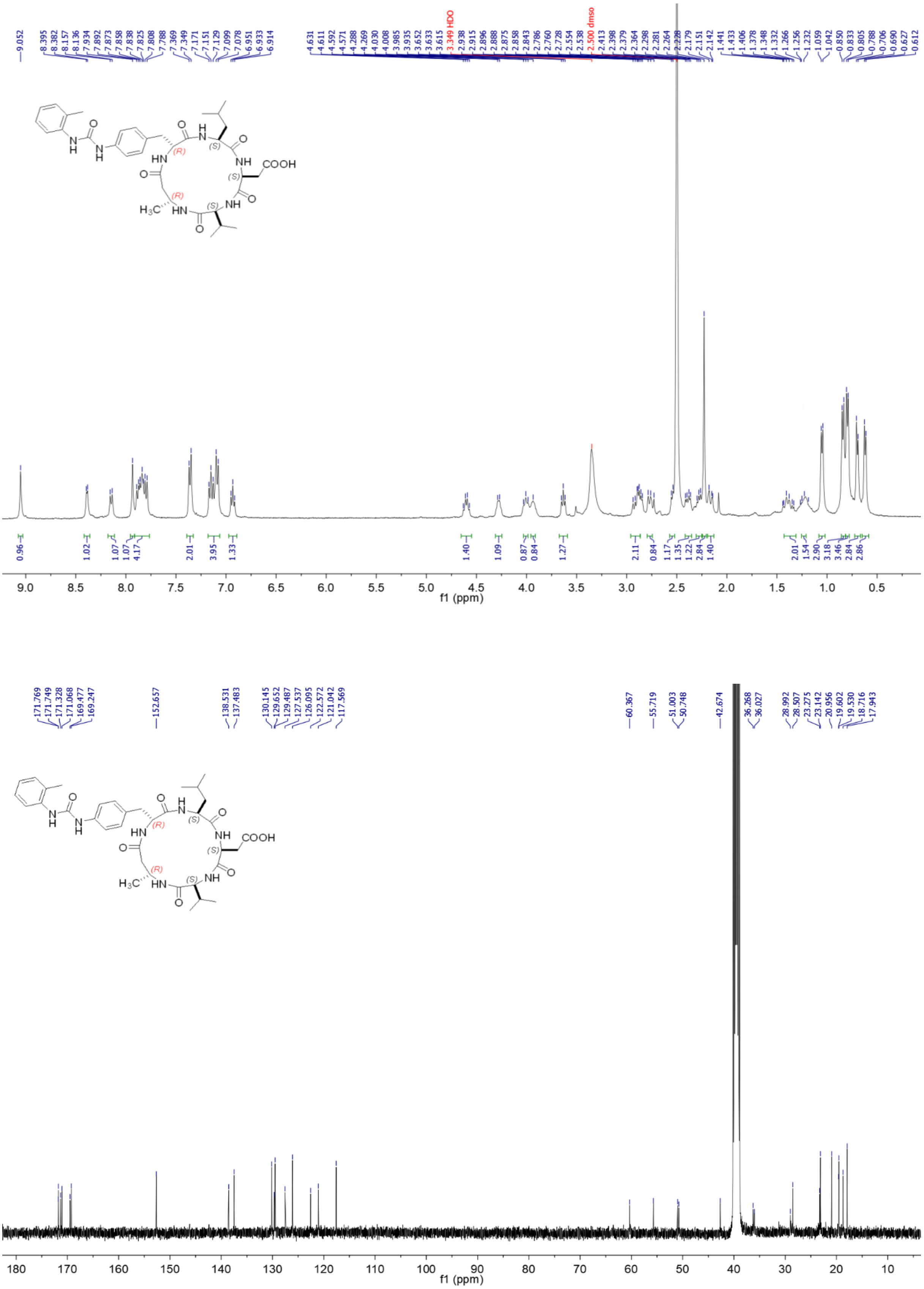


Figure S27. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **11c**.

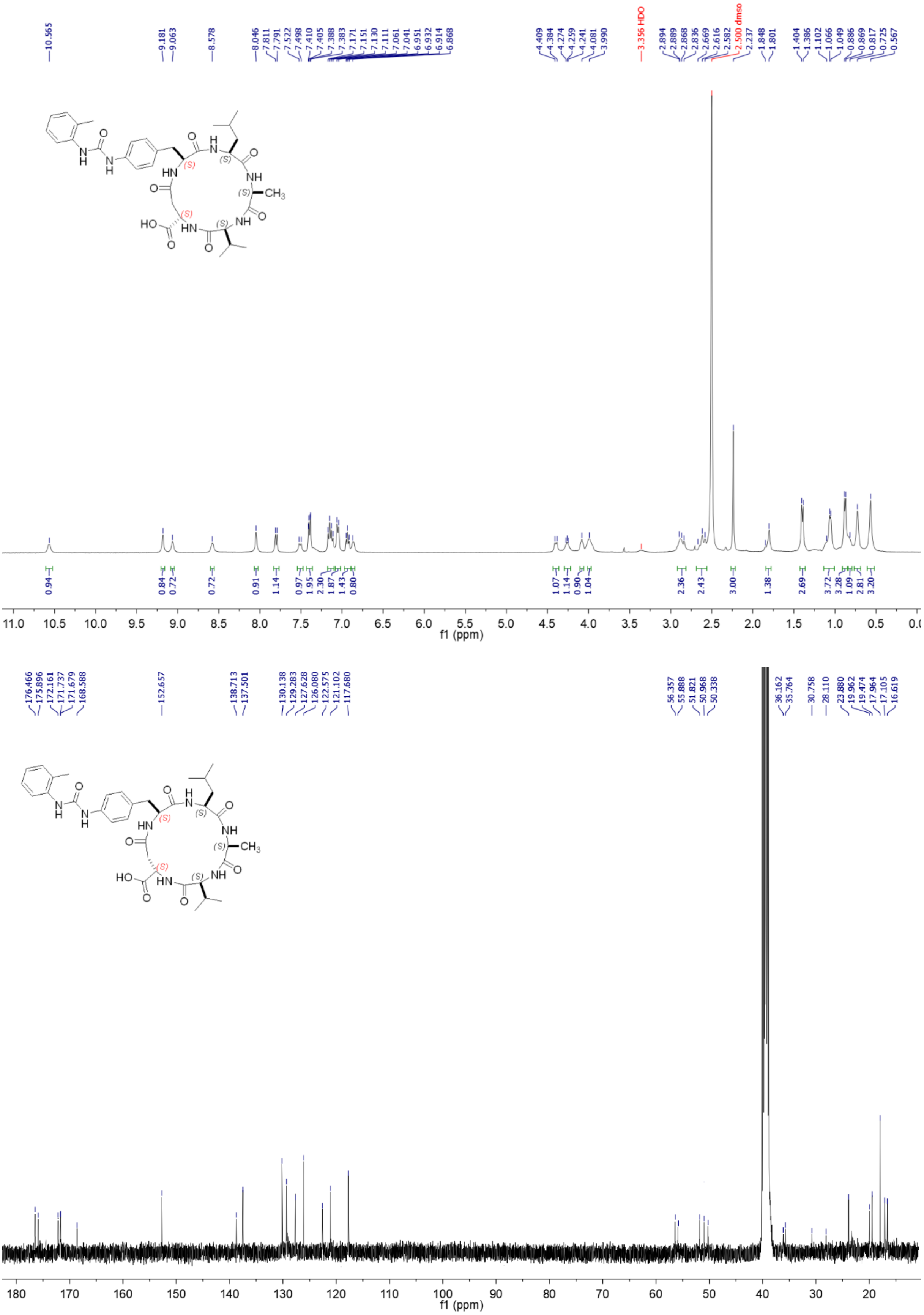


Figure S28. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **12a**.

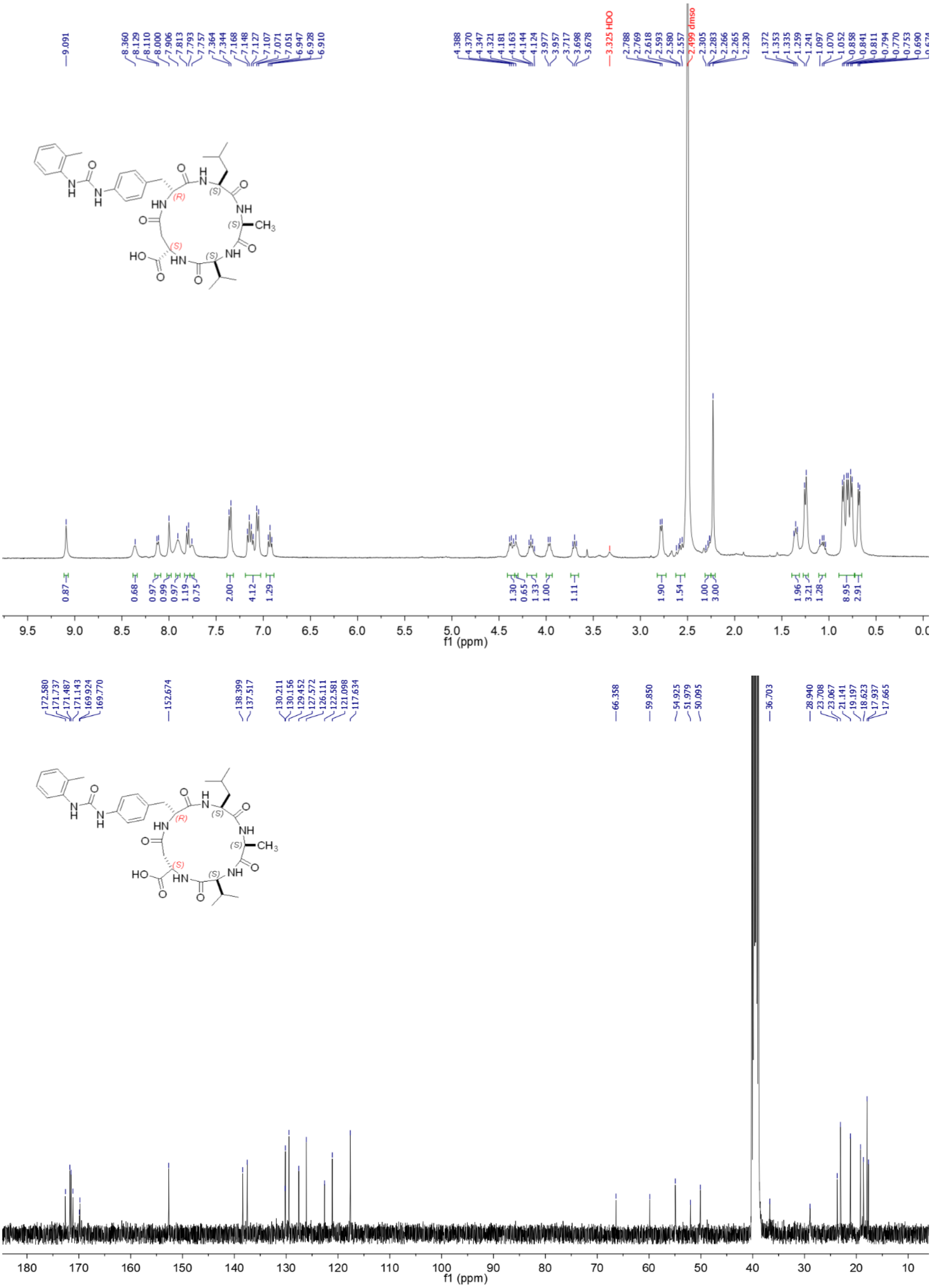


Figure S29. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **12c**.

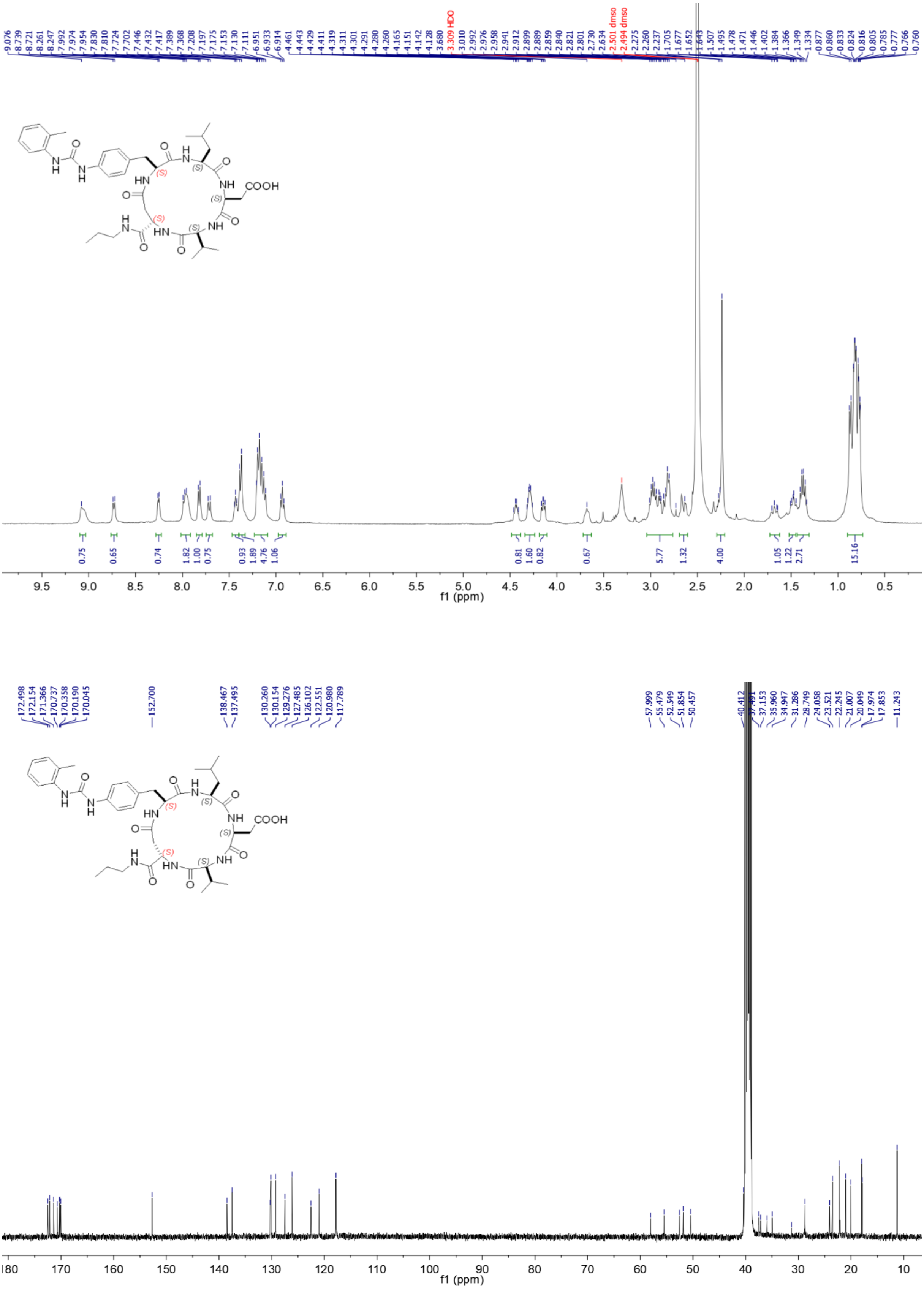


Figure S30. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **13**.

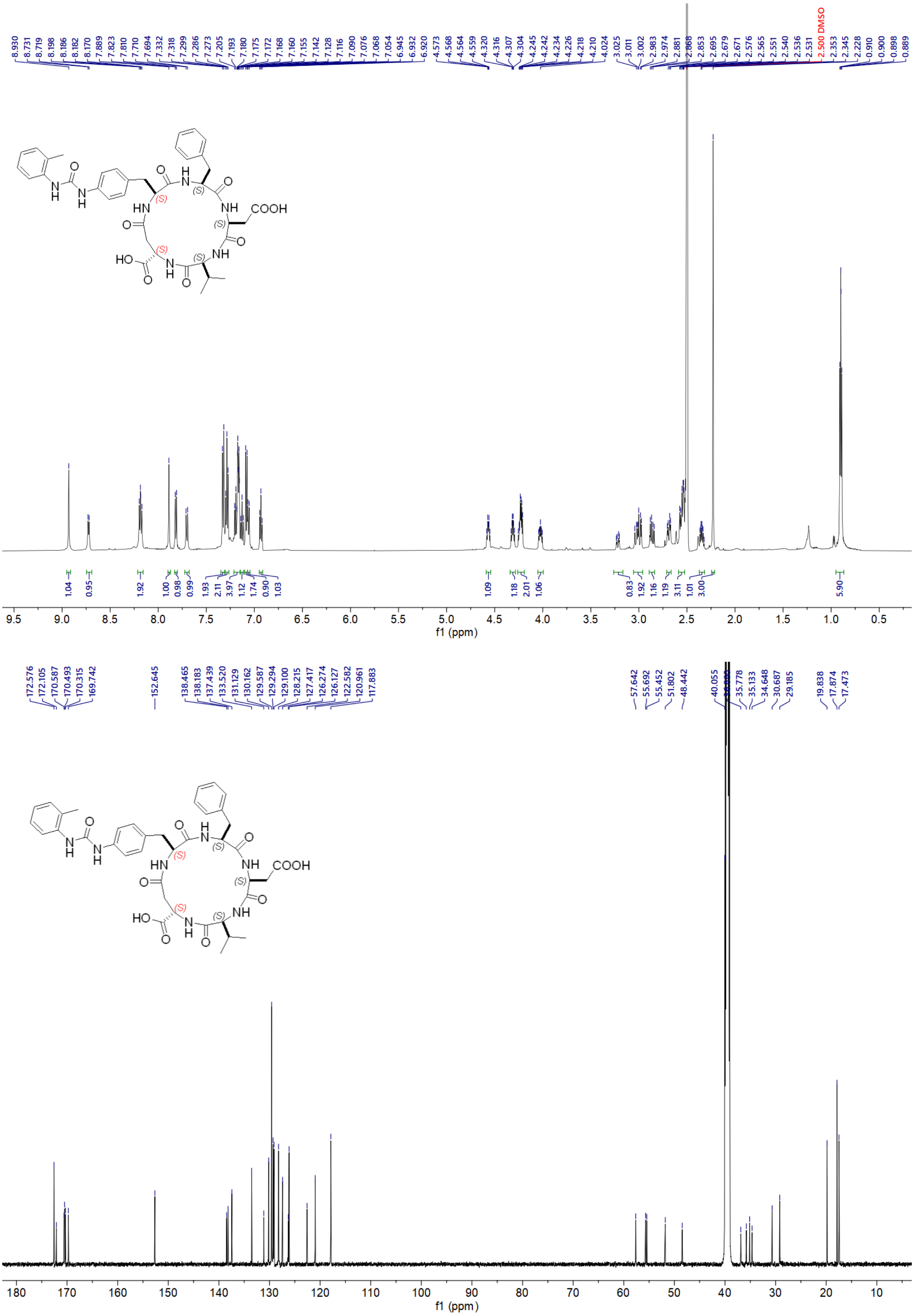


Figure S31. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 600 MHz) and ^{13}C-NMR (DMSOd$_6$, 150 MHz) of **14**.

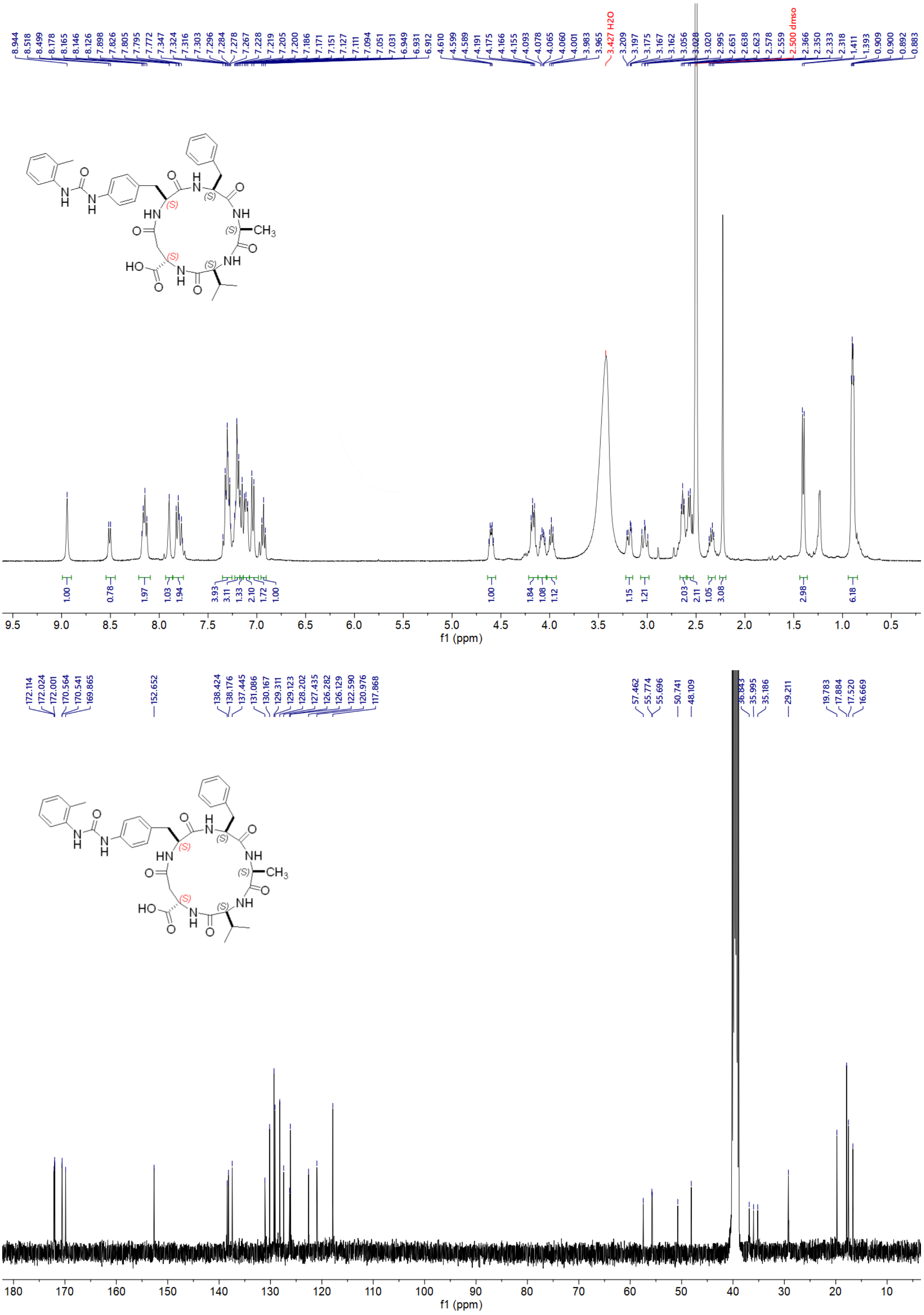


Figure S32. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **15**.

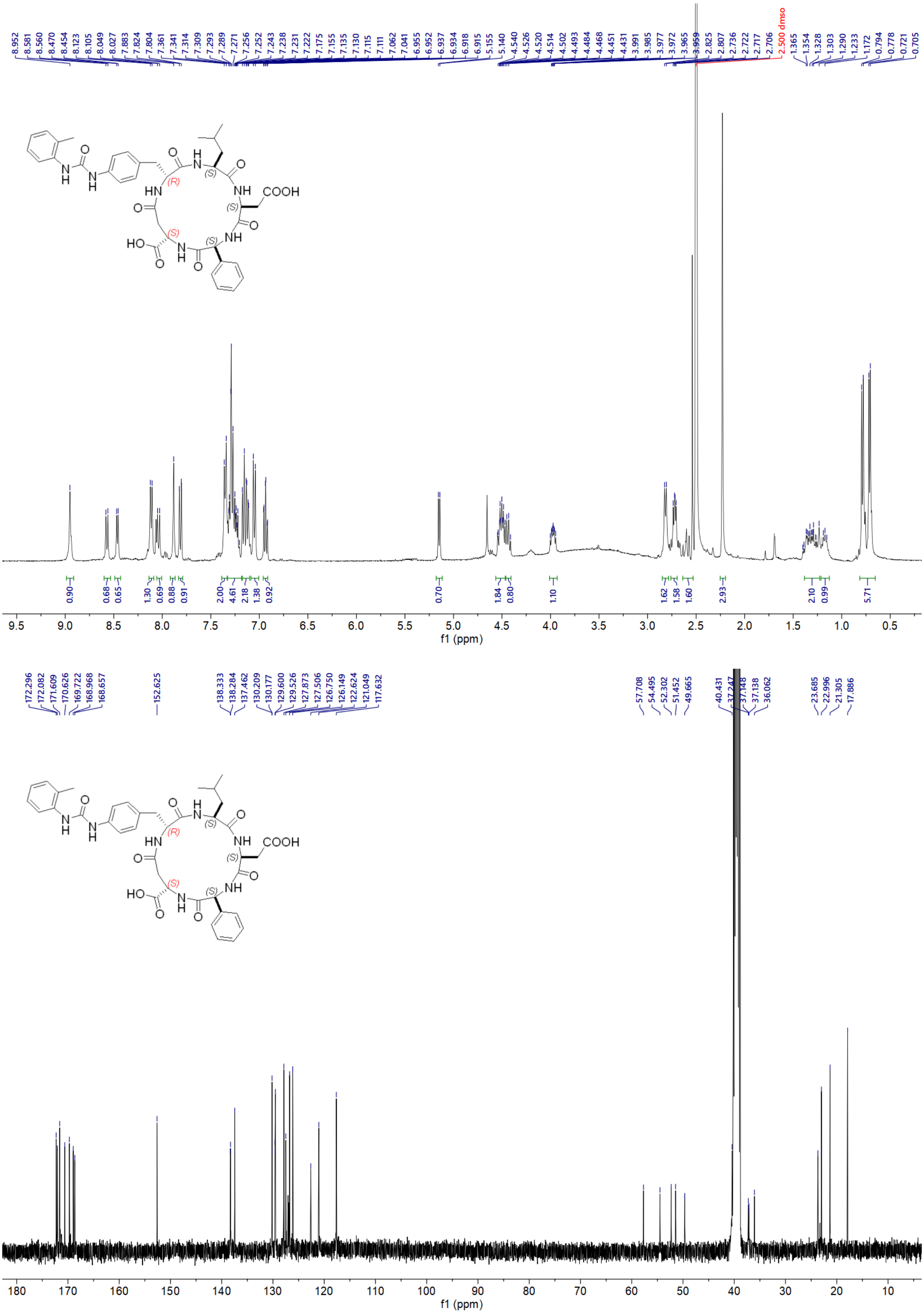
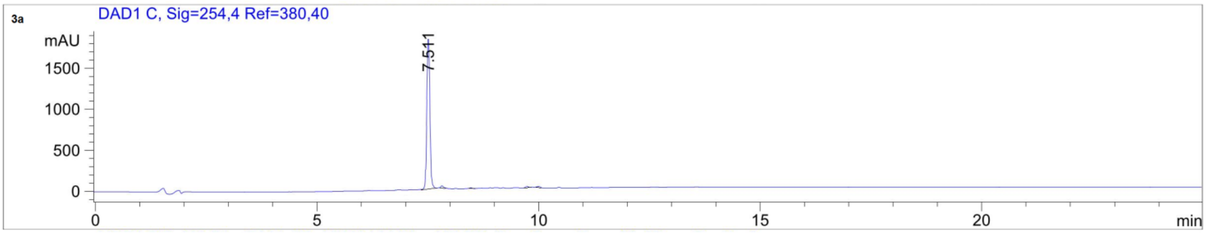


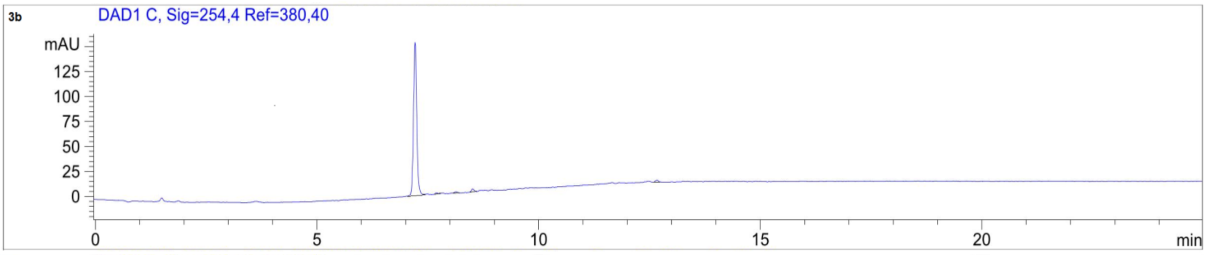
Figure S33. ^{1}H-NMR (8:2 DMSOd$_6$/H$_2$O at 400 MHz) and ^{13}C-NMR (DMSOd$_6$, 100 MHz) of **16**.

RP HPLC chromatograms:

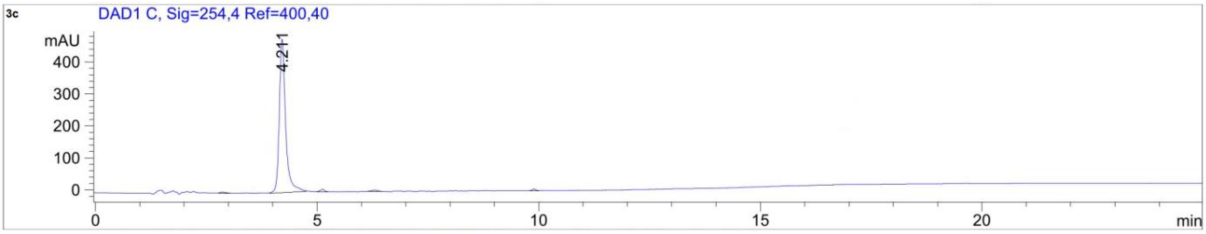
3a



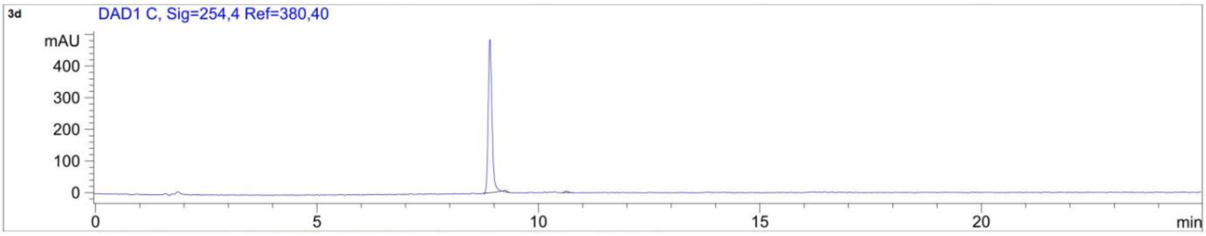
3b



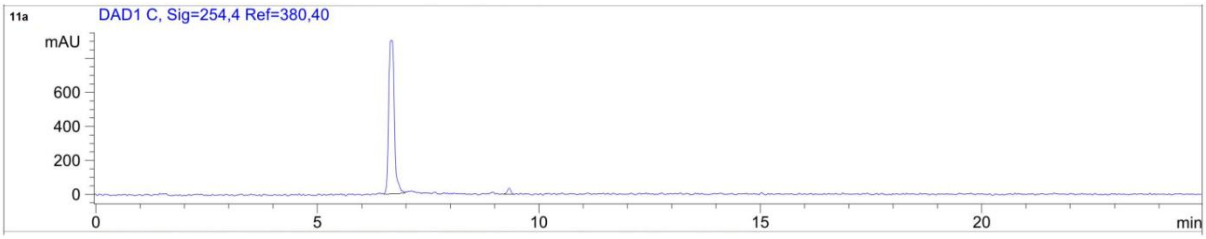
3c



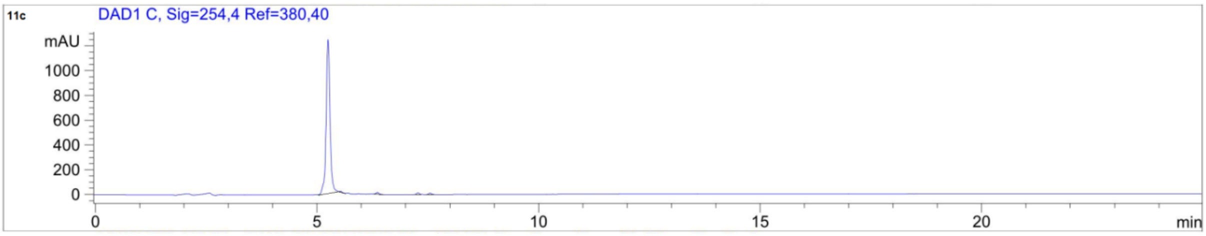
3d



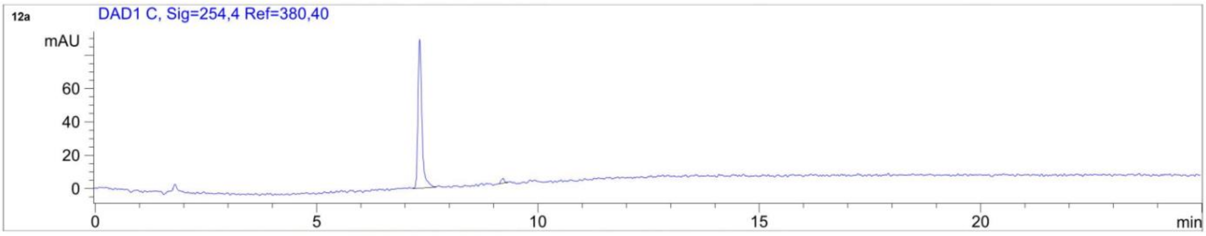
11a



11c



12a



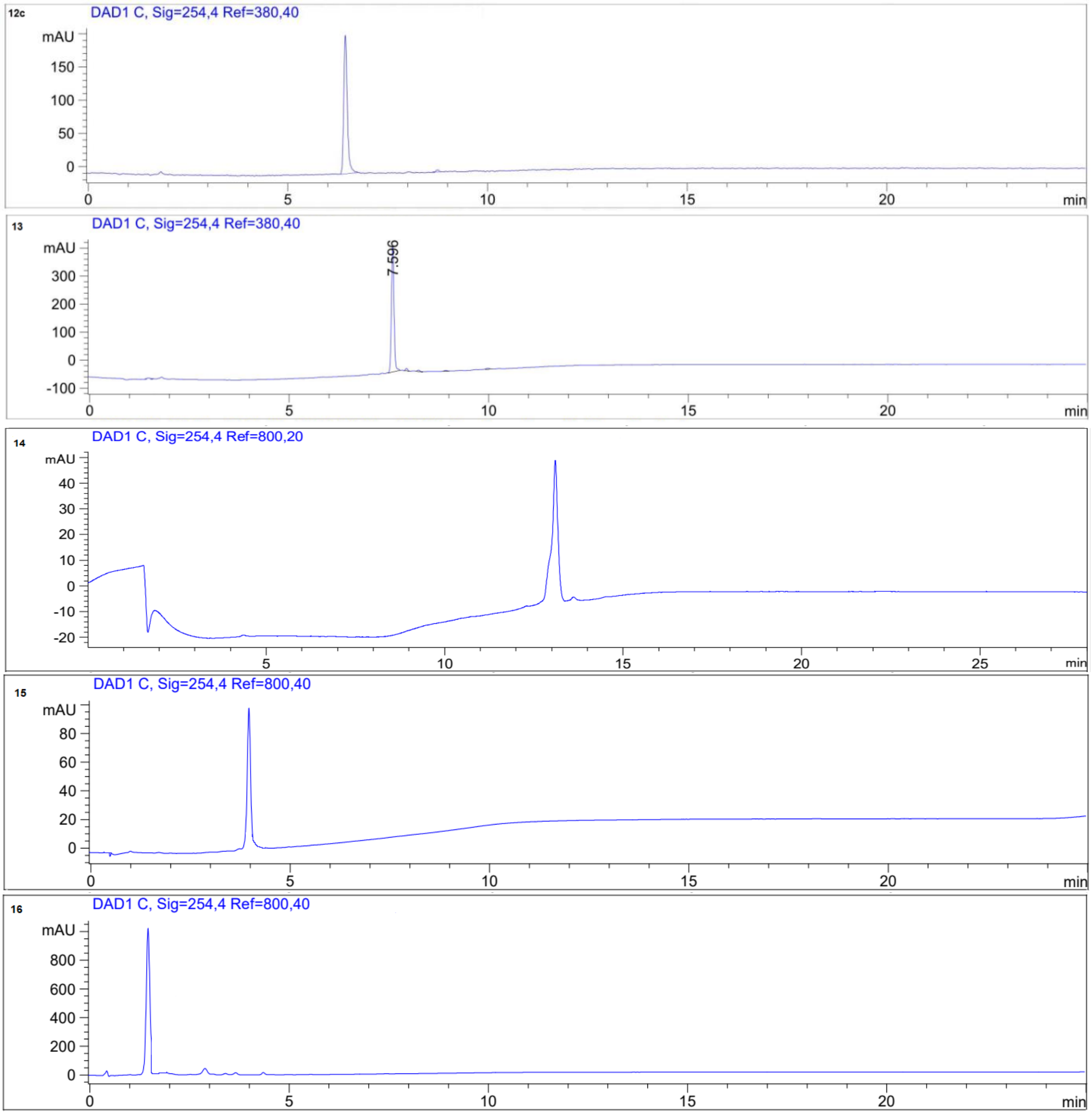


Figure S34. RP HPLC analyses performed on a column Phenomenex mod. Gemini 3 μm C_{18} 110 Å 100 × 3.0 mm; mobile phase from 9:1 $H_2O/CH_3CN/0.1\%$ HCOOH to 2:8 $H_2O/CH_3CN/0.1\%$ HCOOH in 20 min, flow rate of 1.0 mL min^{-1}. DAD 254 nm, unless otherwise specified. For peptide **14**, the analytical flow rate is 0.5 mL min^{-1}.

Appendix-B

Fabrication and Characterization of Versatile Integrin-targeted Nanoplatforms for Diagnostic and Therapeutic Applications

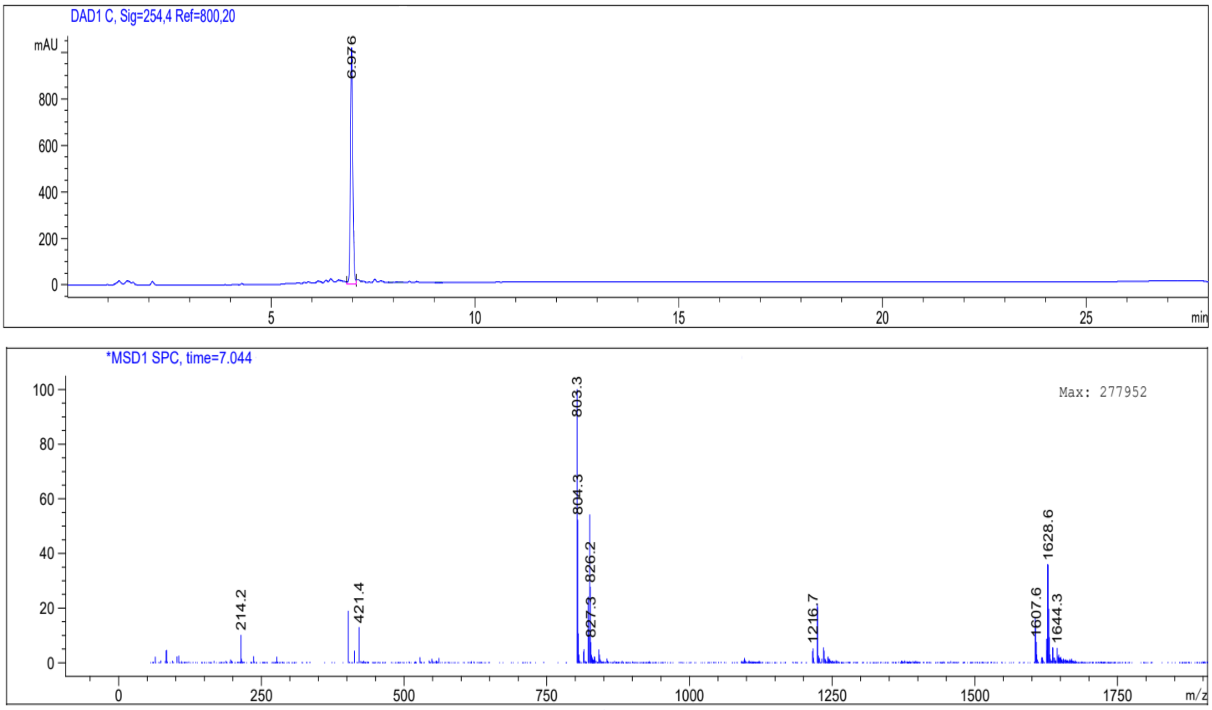


Figure S1. RP HPLC and ESI-MS of LDV Pep **1**. MS-ESI m/z calcd for $[C_{41}H_{55}N_8O_9]^+$ 803.4, found 803.3.

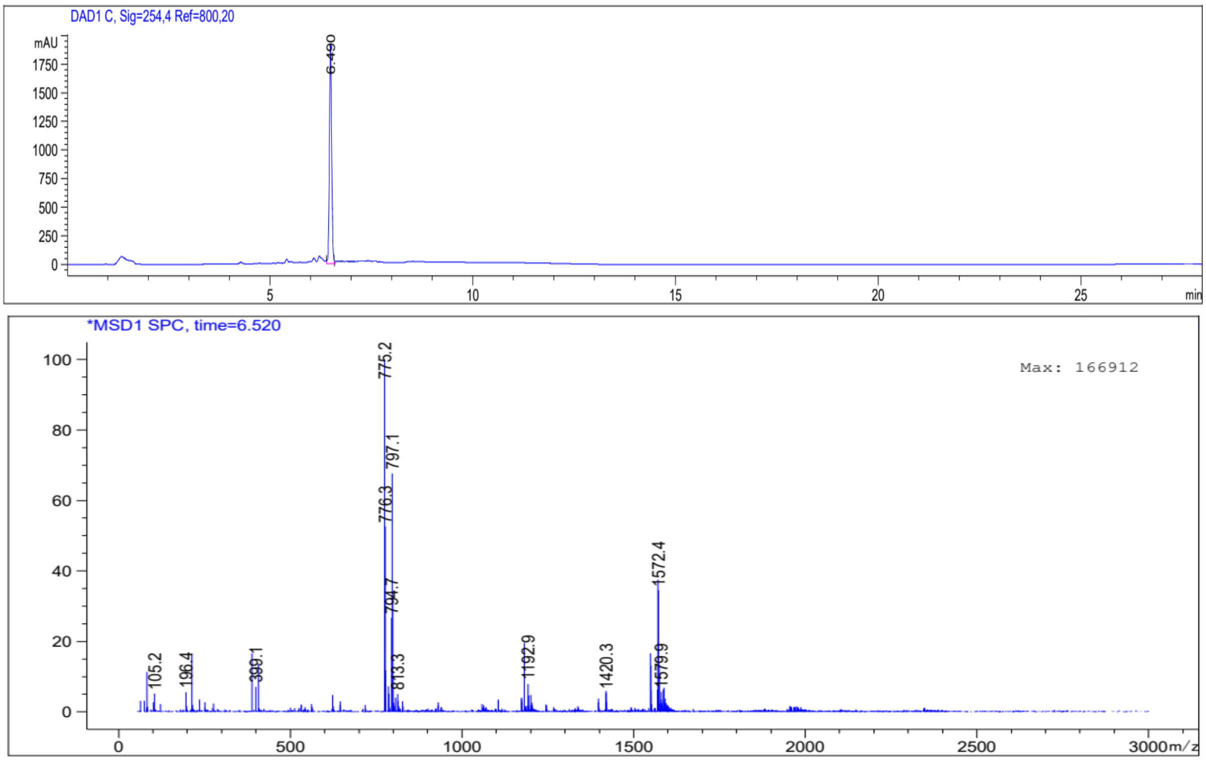


Figure S2. RP HPLC and ESI-MS of LDV Pep **2**. MS-ESI m/z calcd for $[C_{39}H_{51}N_8O_9]^+$ 775.4, found 775.2.

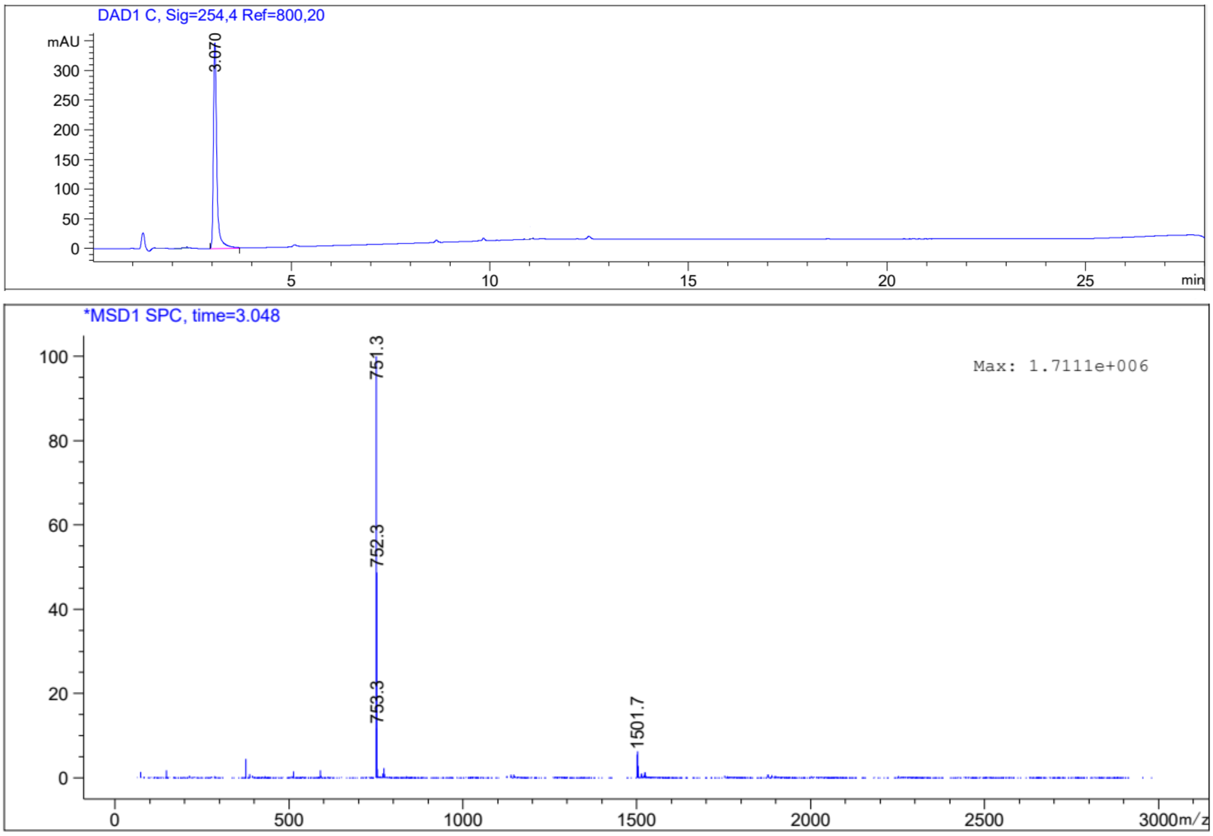


Figure S3. RP HPLC and ESI-MS of LDV Pep **3**. MS-ESI m/z calcd for $[C_{38}H_{55}N_8O_8]^+$ 751.4, found 751.3.

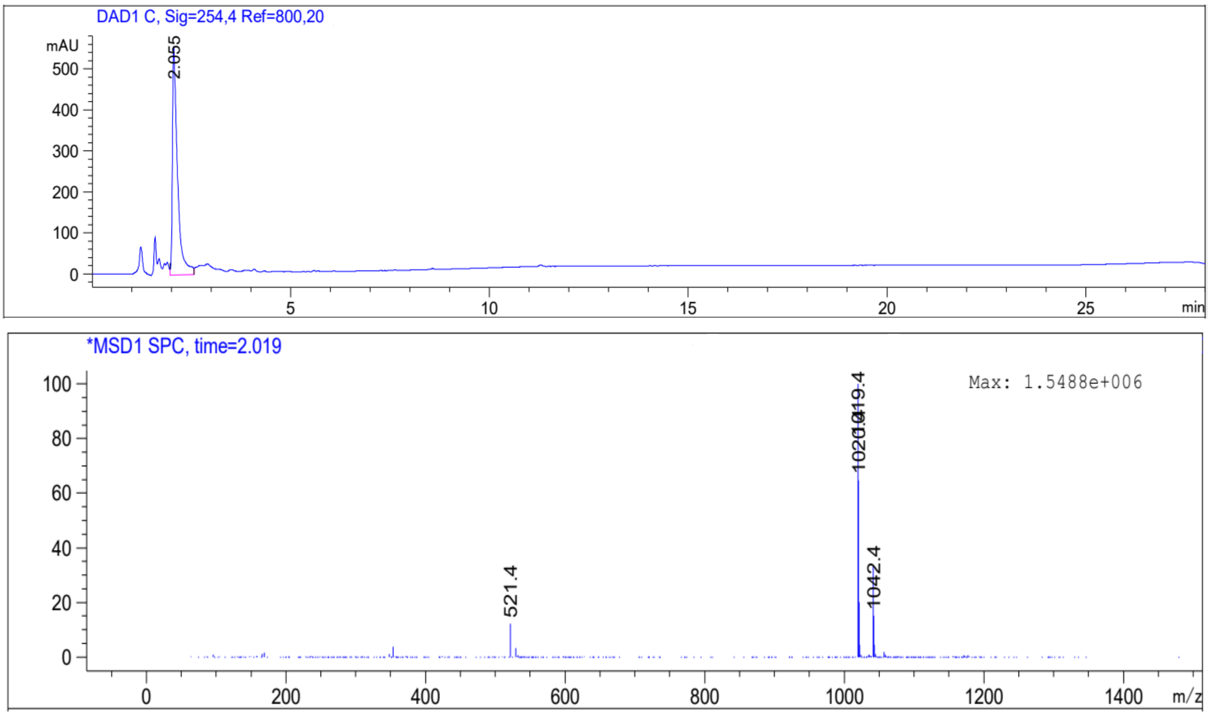
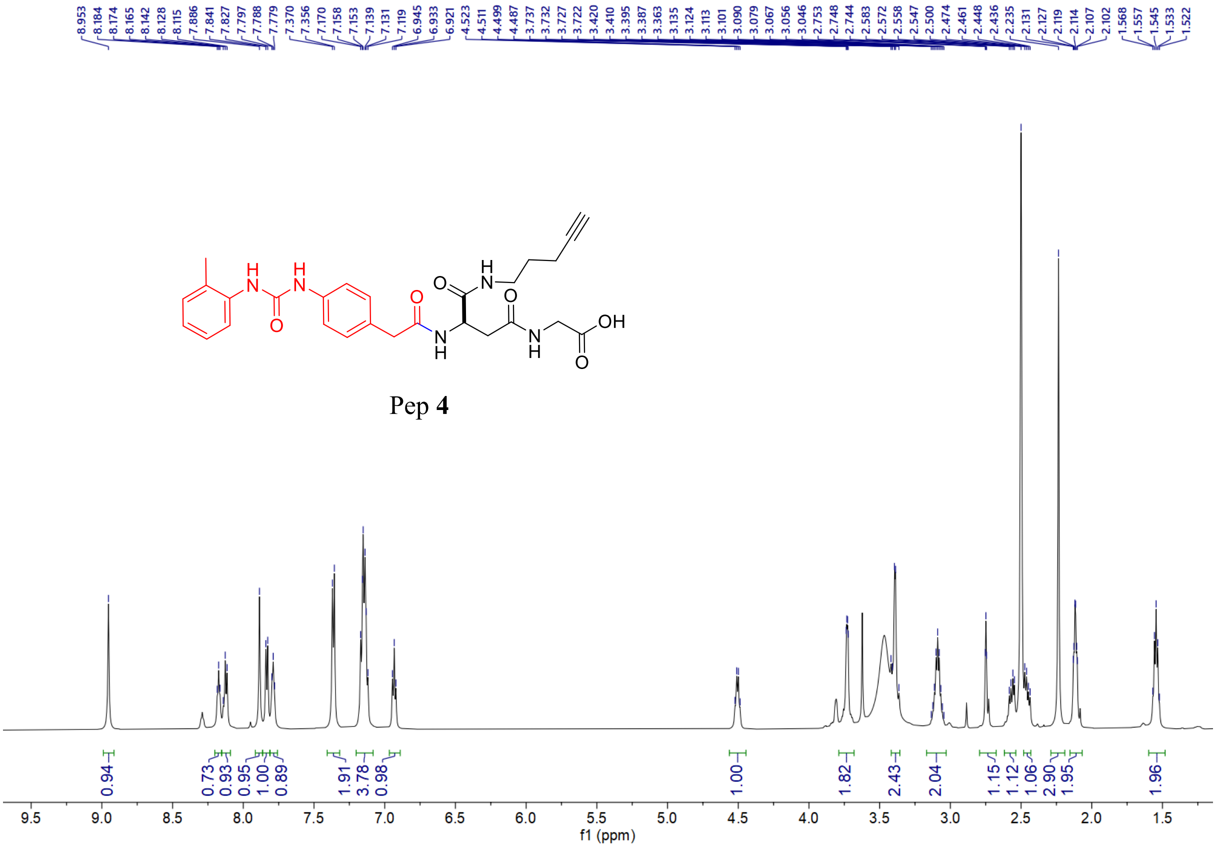


Figure S4. RP HPLC and ESI-MS of **Raf-alkyne** Peptide. MS-ESI m/z calcd for $[C_{49}H_{71}N_{12}O_{12}]^+$ 1019.5, found 1019.4.



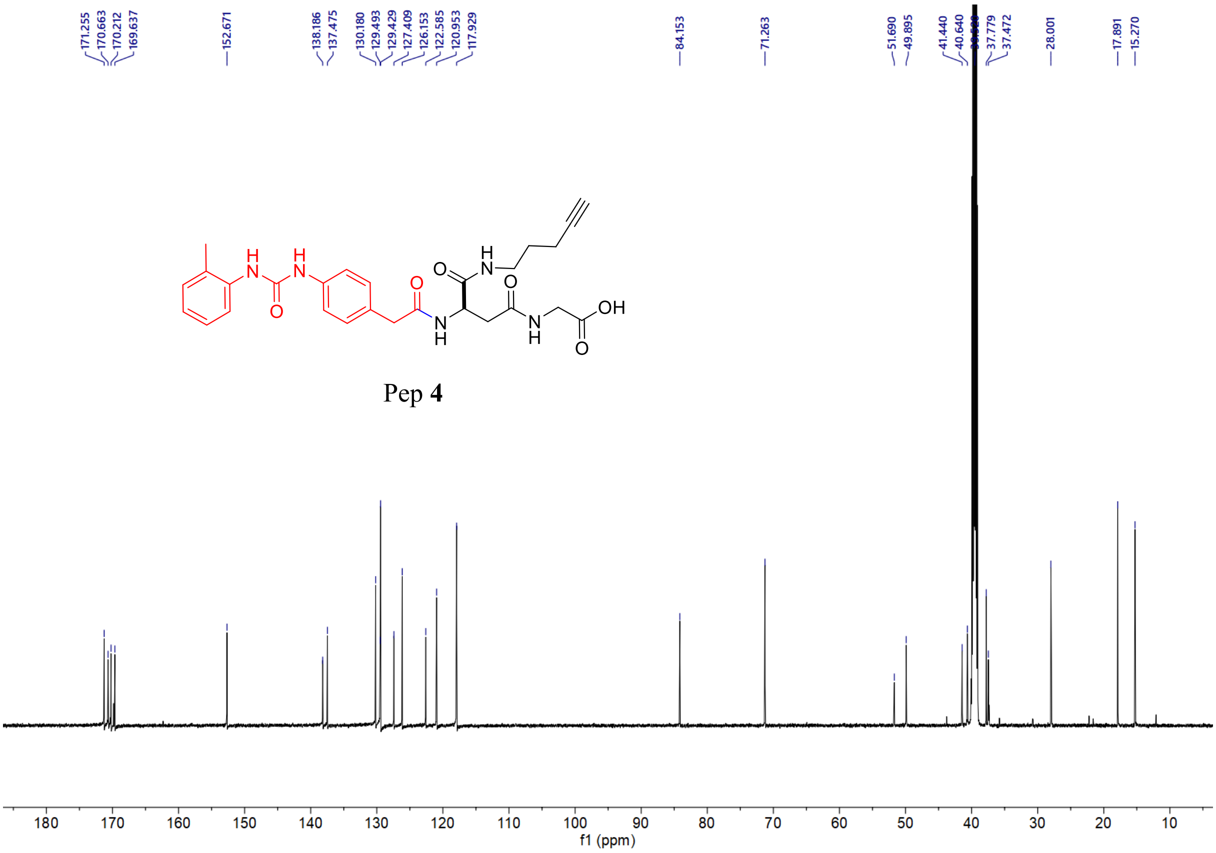


Figure S5. NMR spectra of Peptide **4**.

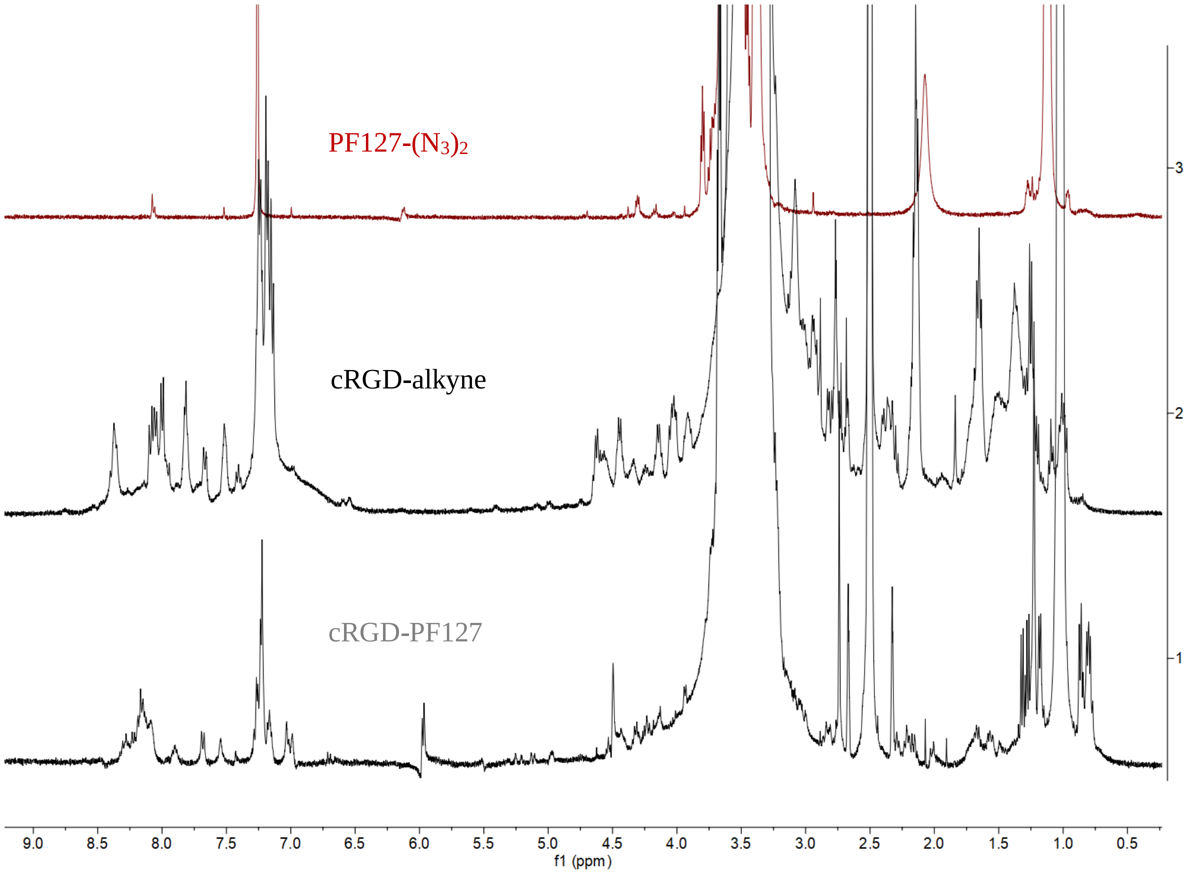


Figure S6. Stack NMR spectra of **PF127-(N_3)$_2$**, **cRGD-alkyne** peptide, and **cRGD-PF127**.

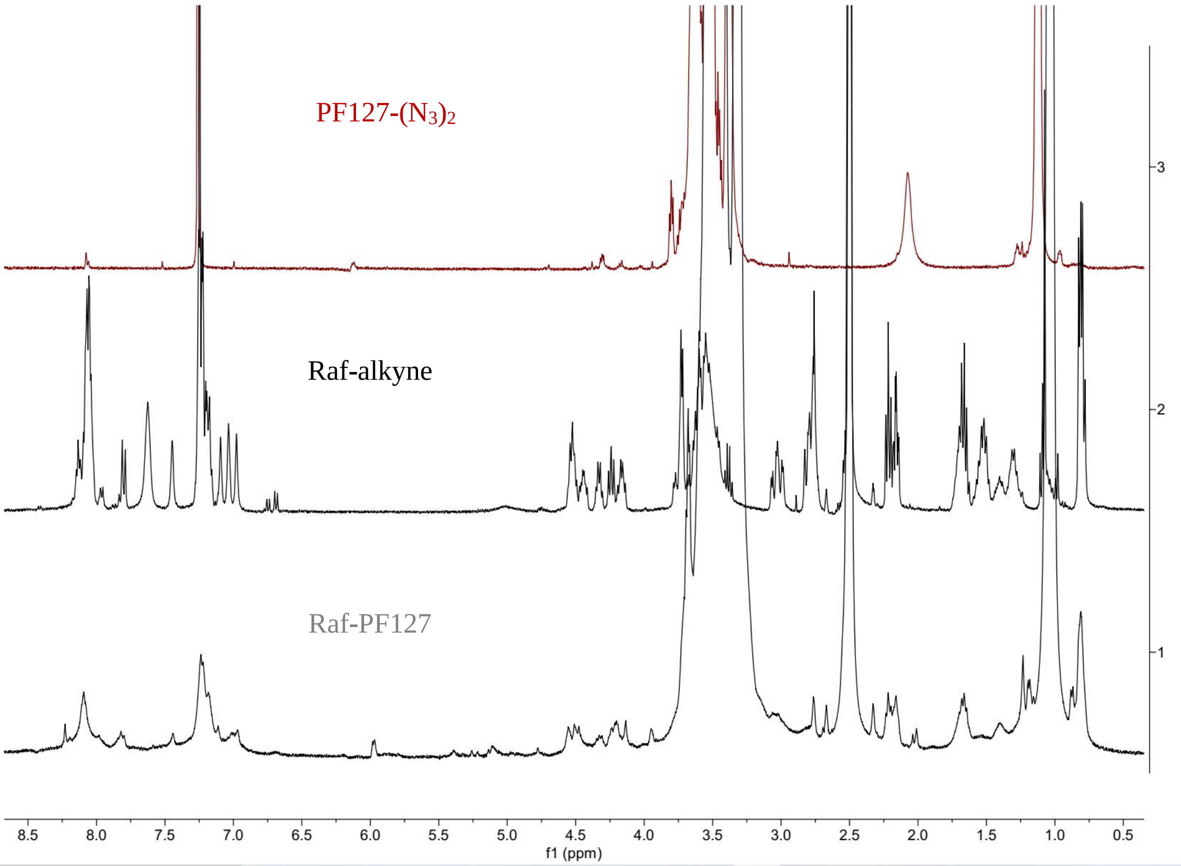


Figure S7. Stack NMR spectra of **PF127-(N_3)$_2$**, **Raf-alkyne** peptide, and **Raf-PF127**.

Appendix-C

Synthesis of Peptides Inhibiting the Assembly of Monomeric Human L-lactate Dehydrogenase into Catalytically-active Homotetramer

CATATGGCGACCCTGAAAGACCAACTGATTTACAACCTGCTGAAAGAGGAACAGACCCCGCAAAA
CAAGATTACCGTGGTGGGCGTGGGCGCGGTTGGTATGGCGTGCGCGATCAGCATTCTGATGAAGGA
CCTGGCGGATGAACTGGCGCTGGTGGACGTTATCGAAGATAAGCTGAAAGGCGAGATGATGGACC
TGCAGCACGGCAGCCTGTTCCTGCGTACCCCGAAGATTGTGAGCGGCAAAGATTACAACGTTACCG
CGAACAGCAAGCTGGTGATCATTACCGCGGGTGCGCGTCAGCAAGAAGGCGAGAGCCGTCTGAAC
CTGGTGCAACGTAACGTTAACATCTTCAAGTTCATCATTCCGAACGTGGTTAAGTACAGCCCGAAC
TGCAAACTGCTGATCGTGAGCAACCCGGTTGACATTCTGACCTATGTTGCGTGGAAGATCAGCGGT
TTCCCGAAAAACCGTGTGATTGGTAGCGGCTGCAACCTGGATAGCGCGCGTTTTCGTTATCTGATG
GGTGAACGTCTGGGCGTTCATCCGCTGAGCTGCCATGGTTGGGTTCTGGGCGAGCATGGTGACAGC
AGCGTGCCGGTTTGGAGCGGTATGAACGTGGCGGGCGTTAGCCTGAAAACCCTGCACCCGGATCTG
GGCACCGACAAGGATAAAGAACAGTGGAAAGAGGTGCACAAACAAGTGGTTGAAAGCGCGTACG
AGGTGATCAAGCTGAAAGGTTATACCAGCTGGGCGATTGGCCTGAGCGTGGCGGACCTGGCGGAG
AGCATCATGAAGAACCTGCGTCGTGTGCACCCGGTTAGCACCATGATCAAGGGTCTGTACGGCATT
AAAGACGATGTGTTTCTGAGCGTTCCGTGCATCCTGGGTCAGAACGGCATTAGCGATCTGGTGAAA
GTTACCCTGACCAGCGAGGAAGAGGCGCGTCTGAAAAAGAGCGCGGACACCCTGTGGGGCATCCA
AAAGGAACTGCAATTCTAA**GGATCC**

MATLKDQLIYNLLKEEQTPQNKITVVGVGAVGMACAISILMKDLADELALVDVIEDKLKGEMMDLQH
GSLFLRTPKIVSGKDYNVTANSKLVIITAGARQQEGESRLNLVQRNVNIFKFIIPNVVKYSPNCKLLIVSN
PVDILTYVAWKISGFPKNRVIGSGCNLDSARFRYLMGERLGVHPLSCHGWVLGEHGDSSVPVWSGMN
VAGVSLKTLHPDLGTDKDKEQWKEVHKQVVESAYEVIKLKGYTSWAIGLSVADLAESIMKNLRRVHP
VSTMIKGLYGIKDDVFLSVPCILGQNGISDLVKVTLTSEEEARLKKSADTLWGIQKELQF

Figure S1. Synthetic gene coding for human L-lactate dehydrogenase A (SwissProt Code: P00338). Sequence of the synthetic gene, optimized for the *Escherichia coli* codon usage, coding for human L-lactate dehydrogenase A. Start and stop codons are underlined. *Nde*I and *BamH*I sites are marked in green and blue, respectively. The corresponding translation is also shown.

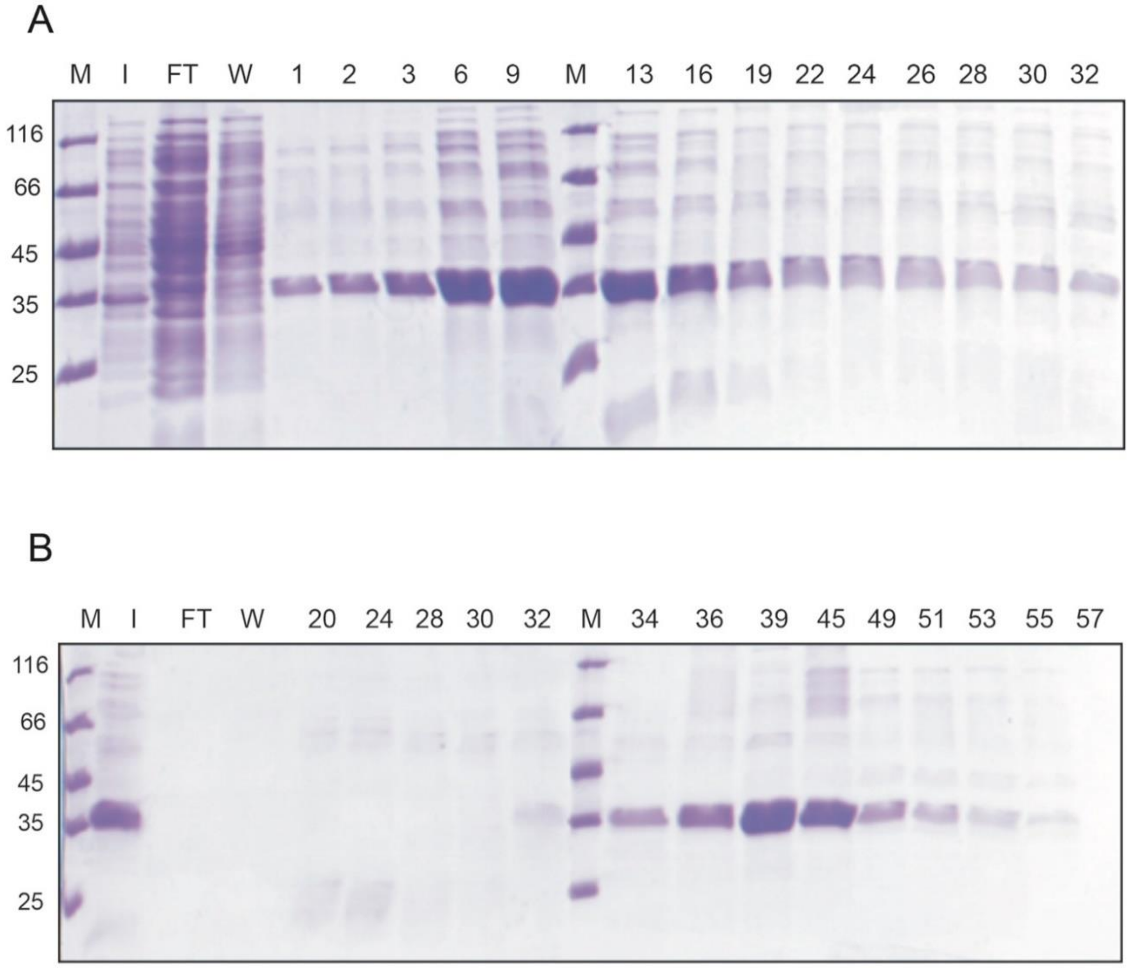


Figure S2. SDS-PAGE analysis of the purification of monomeric human LDH-A by means of affinity and hydrophobic interaction chromatography. (A, B) Electrophoretic analysis of fractions eluted from the Cibacron Blue (A) and HiTrap Phenyl (B) columns used to perform the first two purification steps of monomeric human LDH-A. M: molecular mass markers (their M_r is indicated in kDa at the left); I: input; FT: flow through; W: wash; the fraction numbers are indicated on the top.

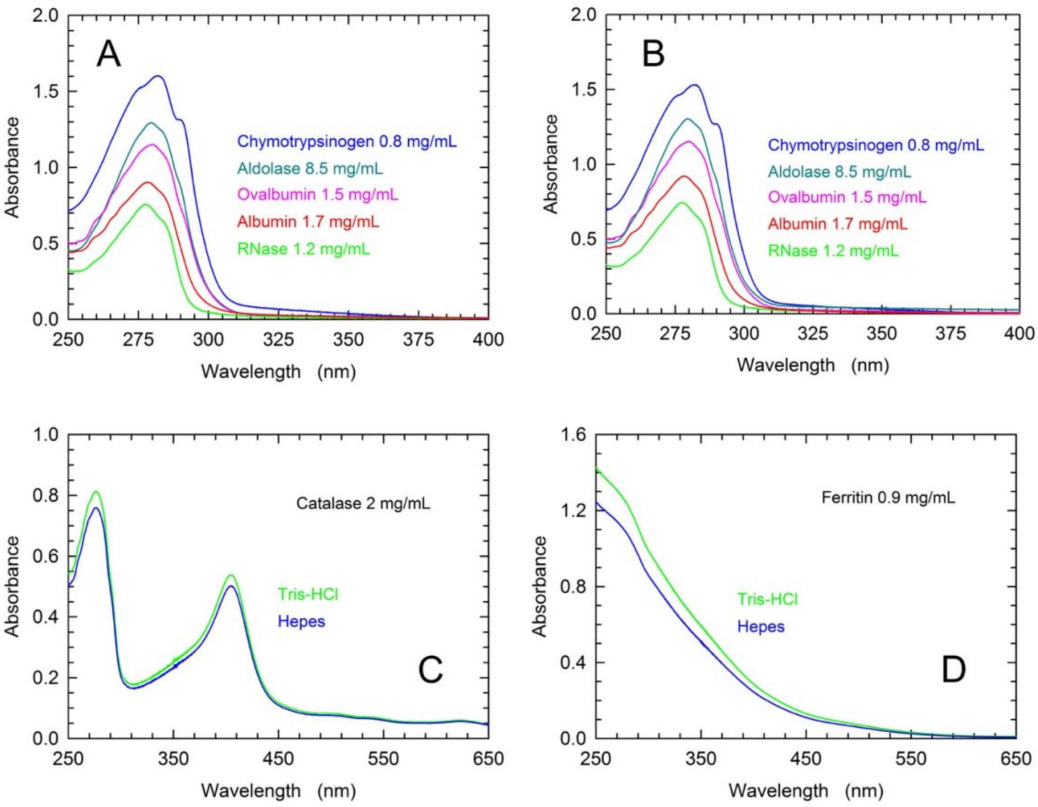


Figure S3. Absorption spectra of the molecular mass markers used to calibrate the Superdex 200 column utilized for gel filtration experiments (see Figure 5.1).

(A, B) Absorption spectra of the indicated molecular mass markers dissolved in 10 mM Tris-HCl (A) or in 10 mM HEPES (B). Both buffers were poised at pH 7.5. (C, D) Absorption spectra of catalase (C) and ferritin (D) dissolved in 10 mM TrisHCl (green lines) or in 10 mM HEPES (blue lines). Both buffers were poised at pH 7.5. The spectra of ferritin are in agreement with those reported by May and Fish (May ME, Fish WW, 1978, The UV and visible spectral properties of ferritin. Arch Biochem Biophys 190:720-725).

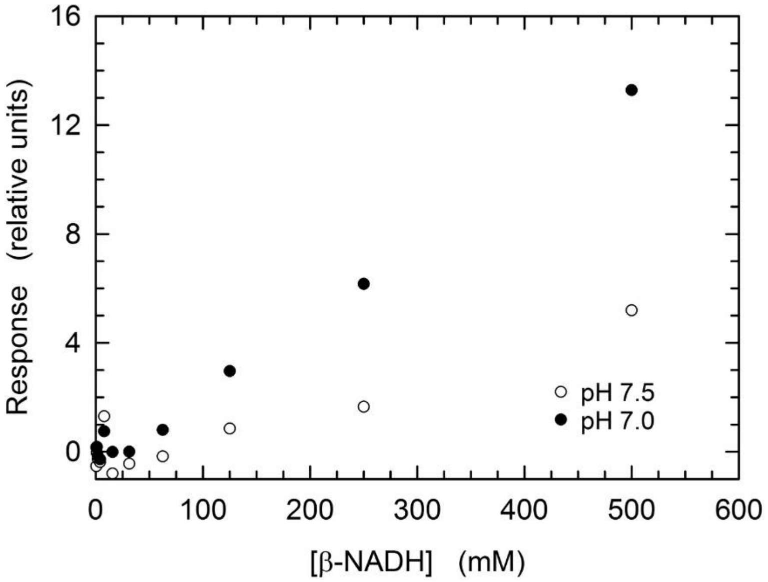


Figure S4. Binding of β-NADH to monomeric human LDH-A.

The association of β-NADH to human LDH-A in monomeric form was evaluated by Surface Plasmon Resonance (SPR) at pH 7.5 (white circles) and 7.0 (black circles). Other conditions as in Figure 5.4.

Table S1. RP HPLC and ESI MS analyses of the linear precursors and of the cyclopeptides considered.

Protected linear peptides[a]				
Compound	Sequence	Yield (%)[e]	Purity (%)[c]	ESI-MS $[M+1]^+$ [d]
TH6	GQND-OBn	78	98	523.2/523.2
TH7	GQN-(*R*)-D-OBn	75	97	523.2/523.2
TH8	Fmoc-KSD(OBn)L	80[b]	89	773.8/774.4
TH9	Fmoc-KSD(OBn)-(*R*)-L	77[b]	85	773.8/774.4
TH14	GQND(OBn)	78	97	523.2/523.2
Protected cyclic peptides				
Compound	Sequence	Yield (%)[e]	Purity (%)[c]	ESI-MS $[M+1]^+$ [d]
TH10	c[GQN-isoD(OBn)]	68	96	505.2/505.2
TH11	c[GQN-(*R*)-isoD(OBn)]	64	96	505.2/505.2
TH12	c[isoK(Fmoc)-SDL]	59	82	666.3/666.3
TH13	c[isoK(Fmoc)-SD-(*R*)-L]	55	80	666.3/666.3
TH15	c[GQND(OBn)]	68	96	505.2/505.2
Cyclic peptides				
Compound	Sequence	Yield (%)[e]	Purity (%)[c]	ESI-MS $[M+1]^+$ [d]
TH2	c[GQN-isoD]	62	97	415.4/415.4
TH3	c[GQN-(*R*)-isoD]	55	96	415.4/415.4
TH4	c[isoKSDL]	55	96	444.1/444.2
TH5	c[isoKSD-(*R*)-L]	47	95	444.1/444.2

[a] isoAsp corresponds to Asp-OBn. [b] based on the estimated loading of the resin. [c] Determined by analytical RP HPLC (General Methods). [d] Experimental/Expected. [e] Determined after semi-preparative RP HPLC.

Table S2. RP HPLC and ESI MS analyses of the fully deprotected linear peptides.

Protected peptides[a]		**Fully deprotected peptides**				
Compound	Sequence	Compound	Sequence	Yield (%)[e]	Purity (%)[c]	ESI-MS $[M+1]^+$[d]
TH6	GQND-OBn	TH16	GQND	72	96	432.9/433.2
TH7	GQN-(*R*)-D-OBn	TH17	GQN-(*R*)-D	69	97	432.9/433.2
-	-	TH1	GQNGISDL	85	97	803.3/803.4
TH8	Fmoc-KSD(OBn)L	TH18	KSDL	73	97	462.0/462.2
TH9	Fmoc-KSD-(*R*)-L	TH19	KSD-(*R*)-L	72	96	462.0/462.2

[a] isoAsp corresponds to Asp-OBn. [b] Determined after semi-preparative RP HPLC. [c] Determined by analytical RP HPLC. [d] Experimental/Expected. [e] Determined after semi-preparative RP HPLC.

NMR spectra:

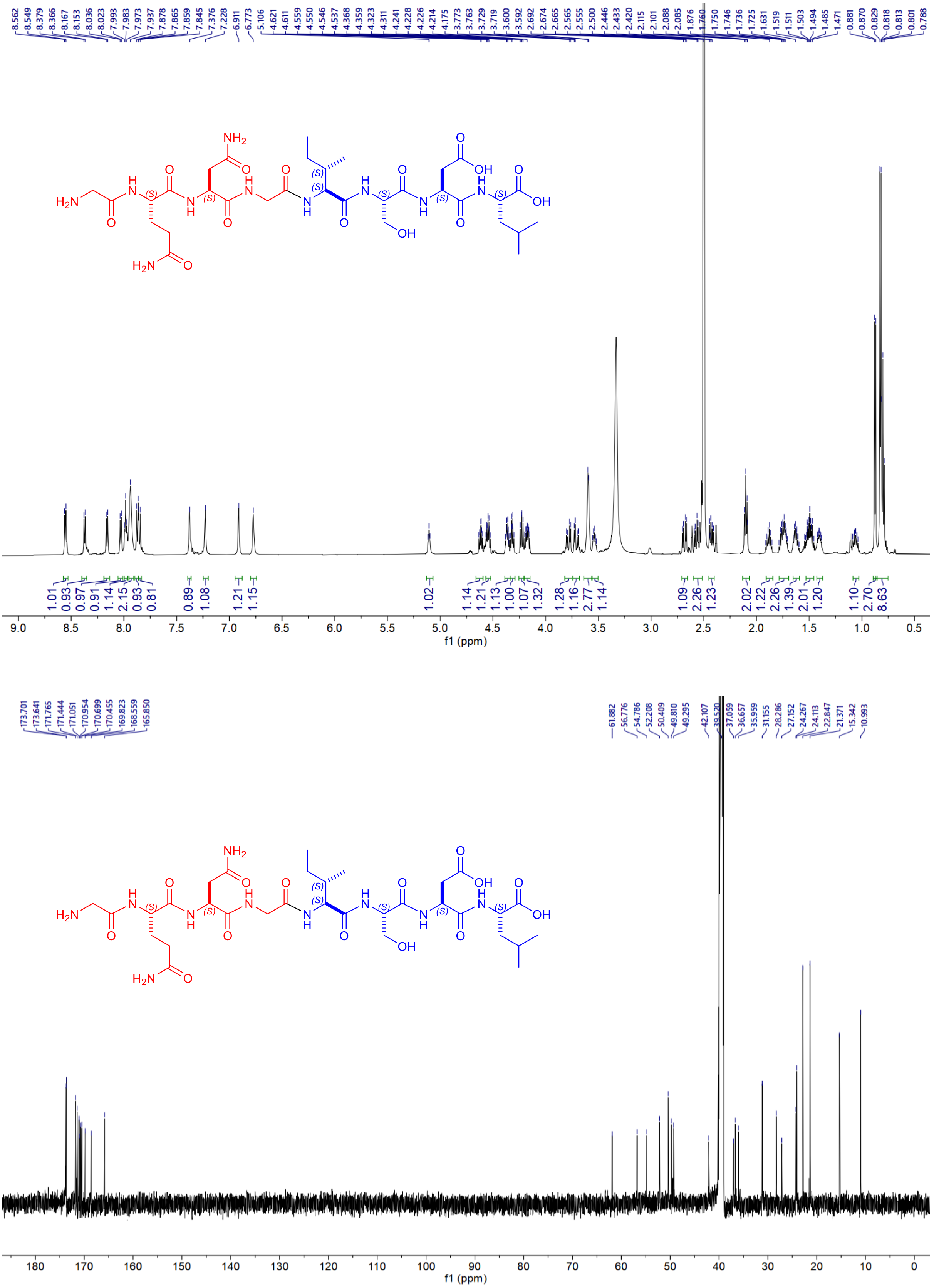


Figure S5A. NMR spectra of the purified **TH1** octapeptide.

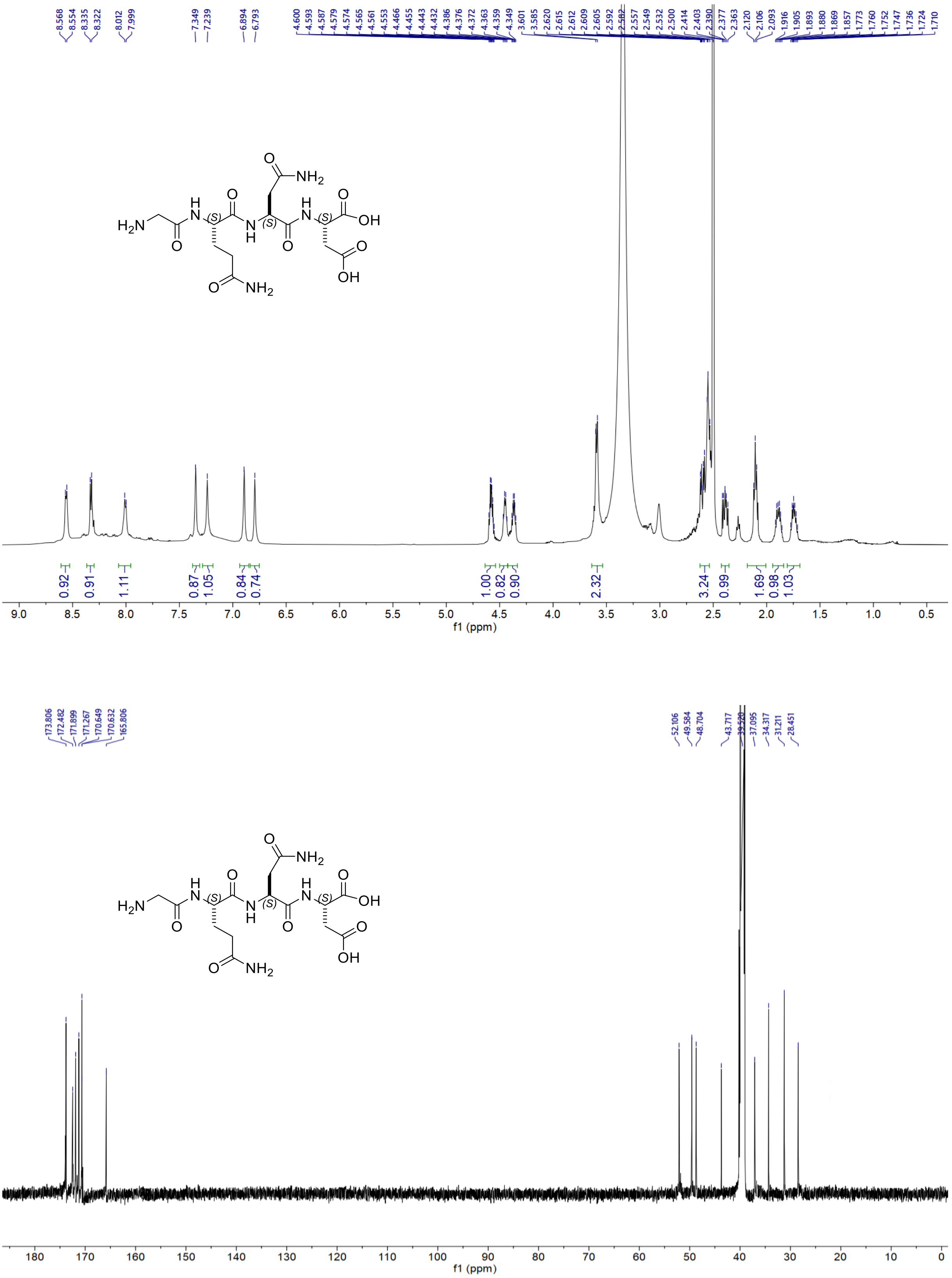


Figure S5B. NMR spectra of the purified **TH16** tetrapeptide.

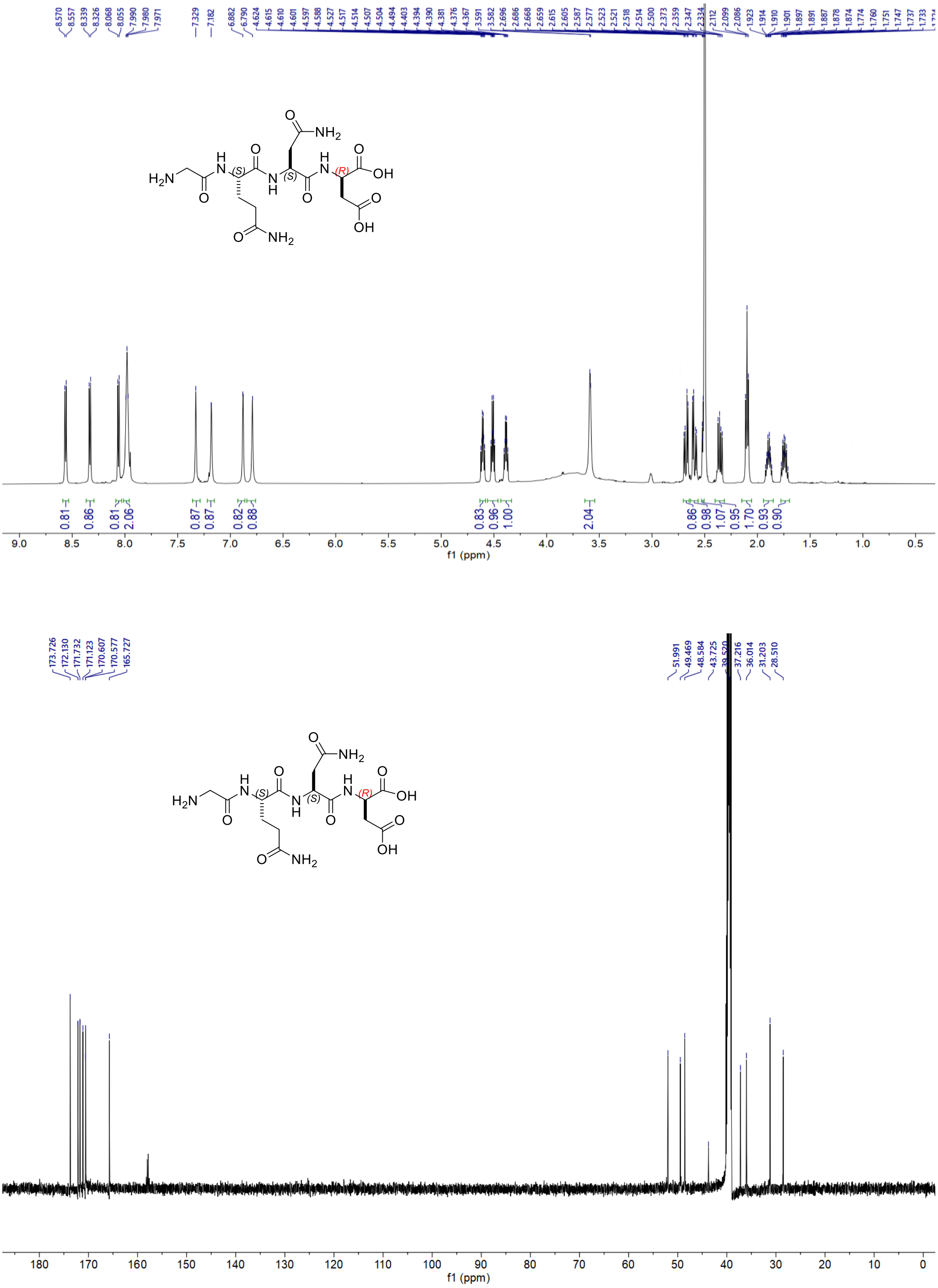


Figure S5C. NMR spectra of the purified **TH17** tetrapeptide.

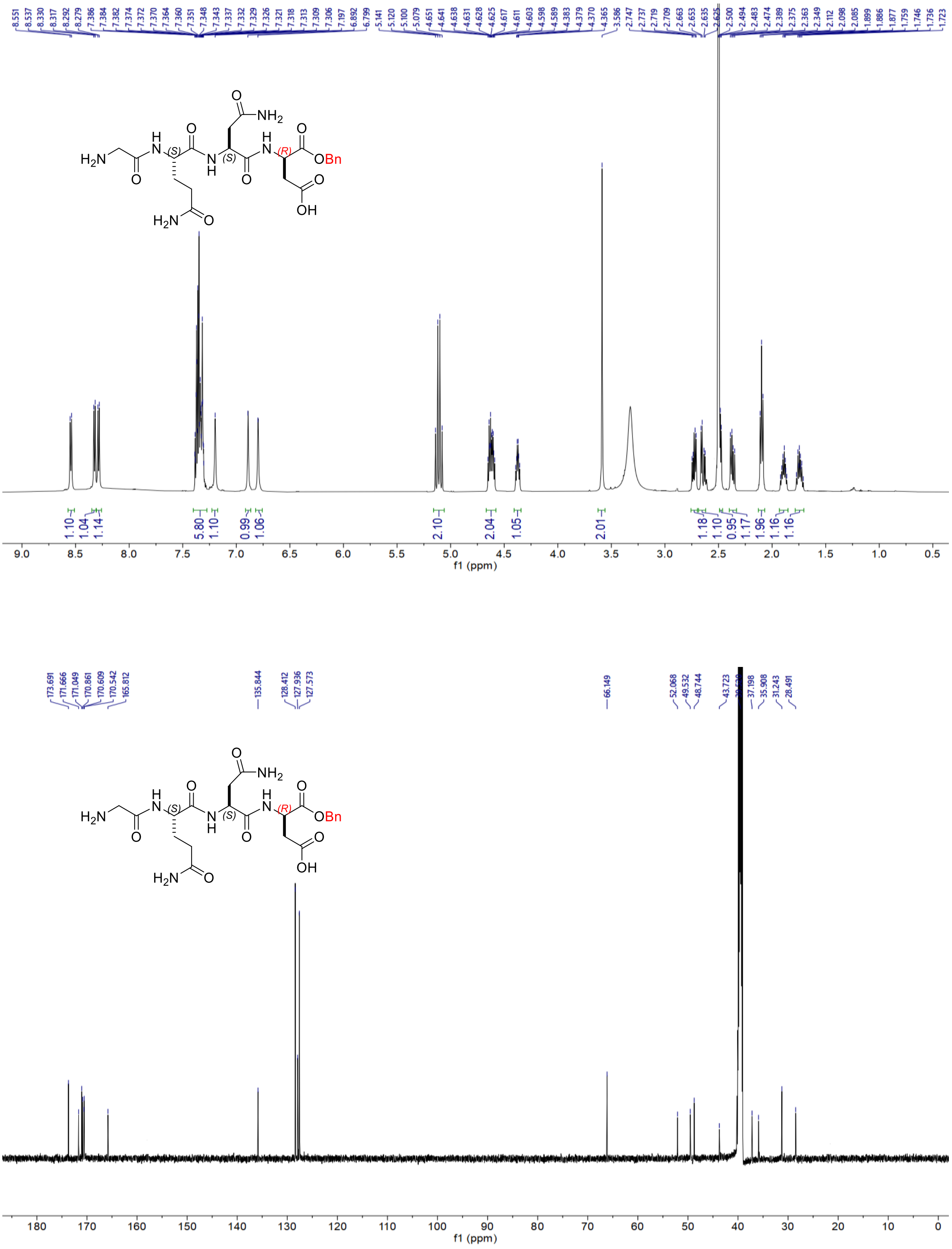


Figure S5D. NMR spectra of the purified **TH7** tetrapeptide.

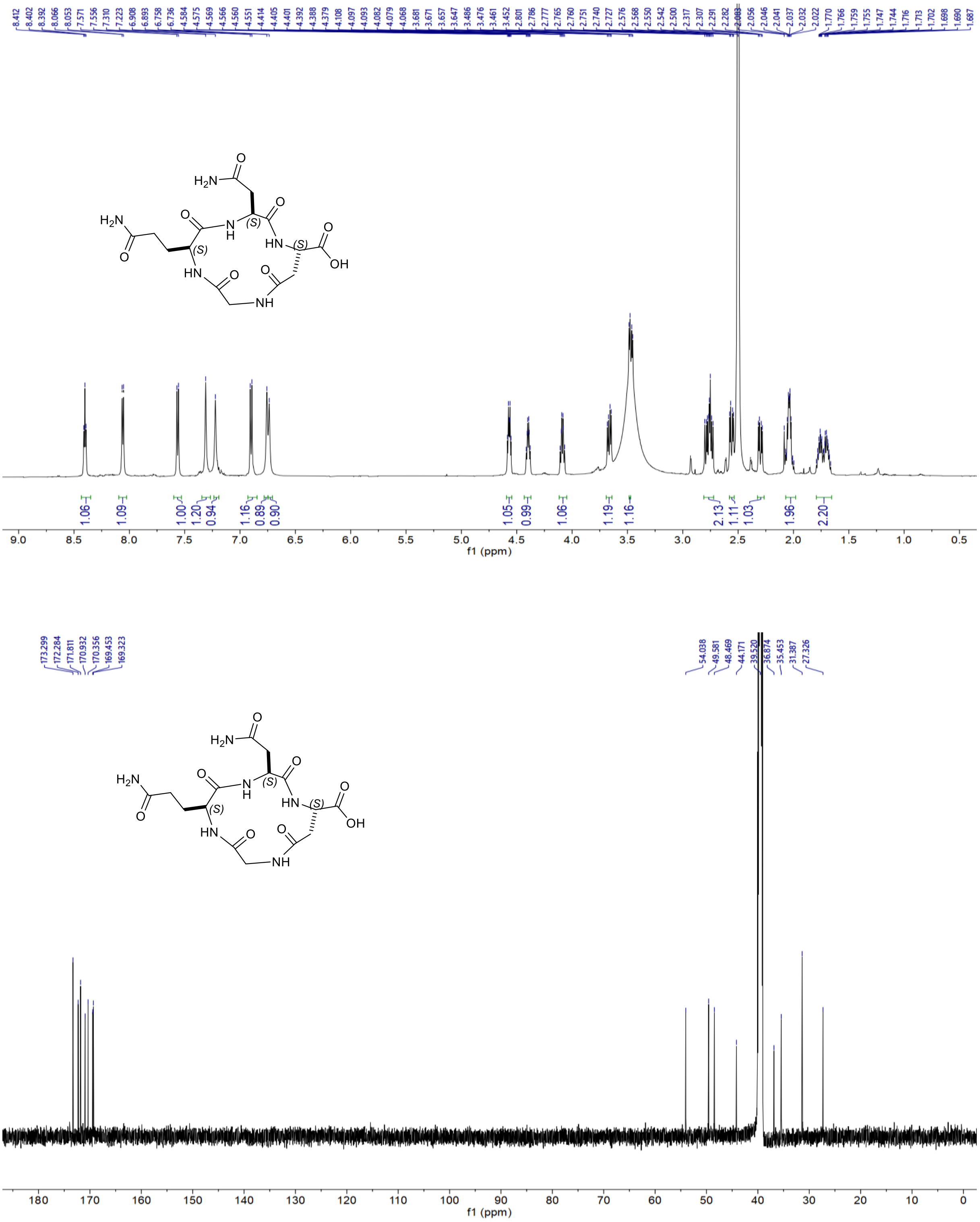


Figure S5E. NMR spectra of the purified **TH2** tetrapeptide.

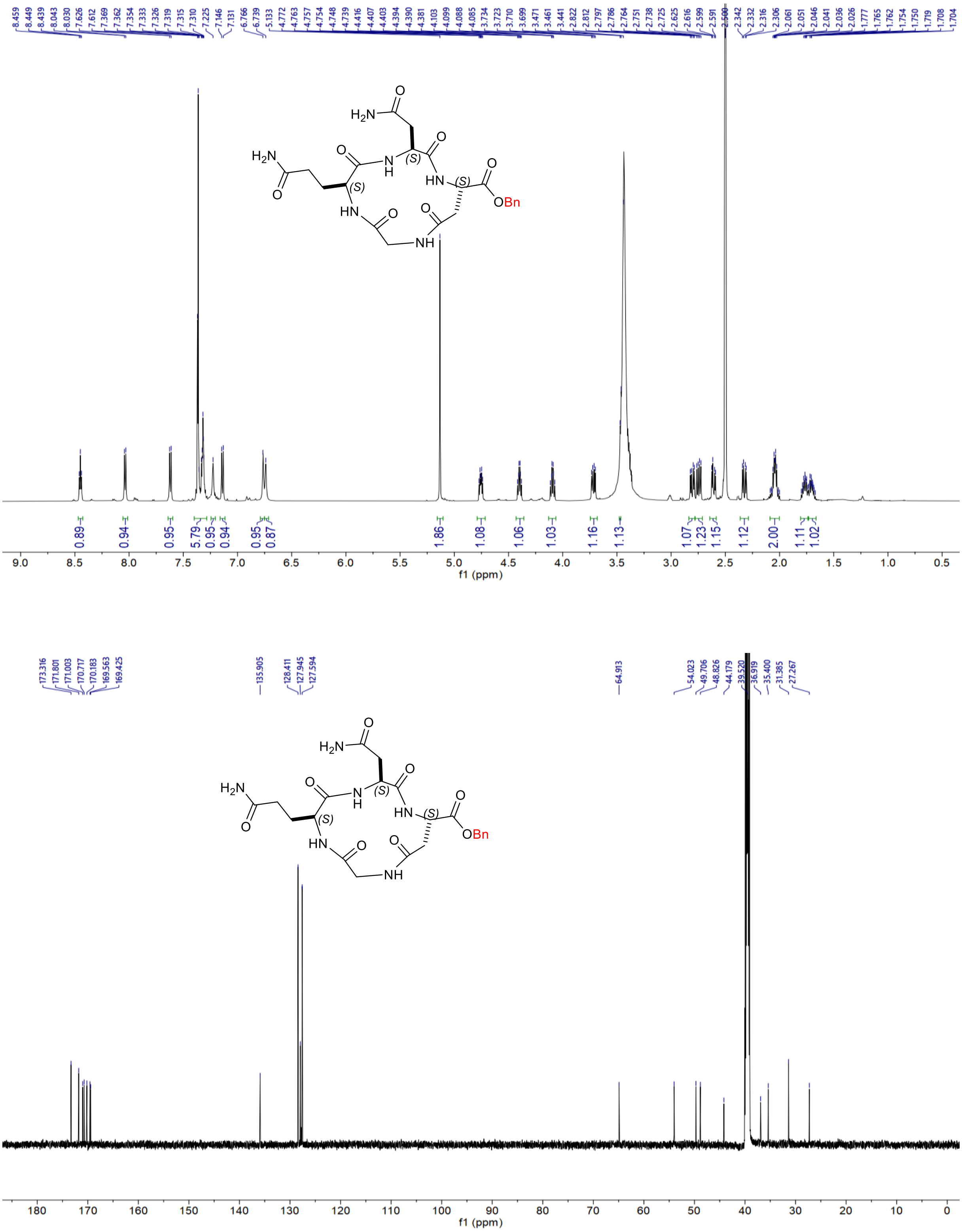


Figure S5F. NMR spectra of the purified **TH10** tetrapeptide.

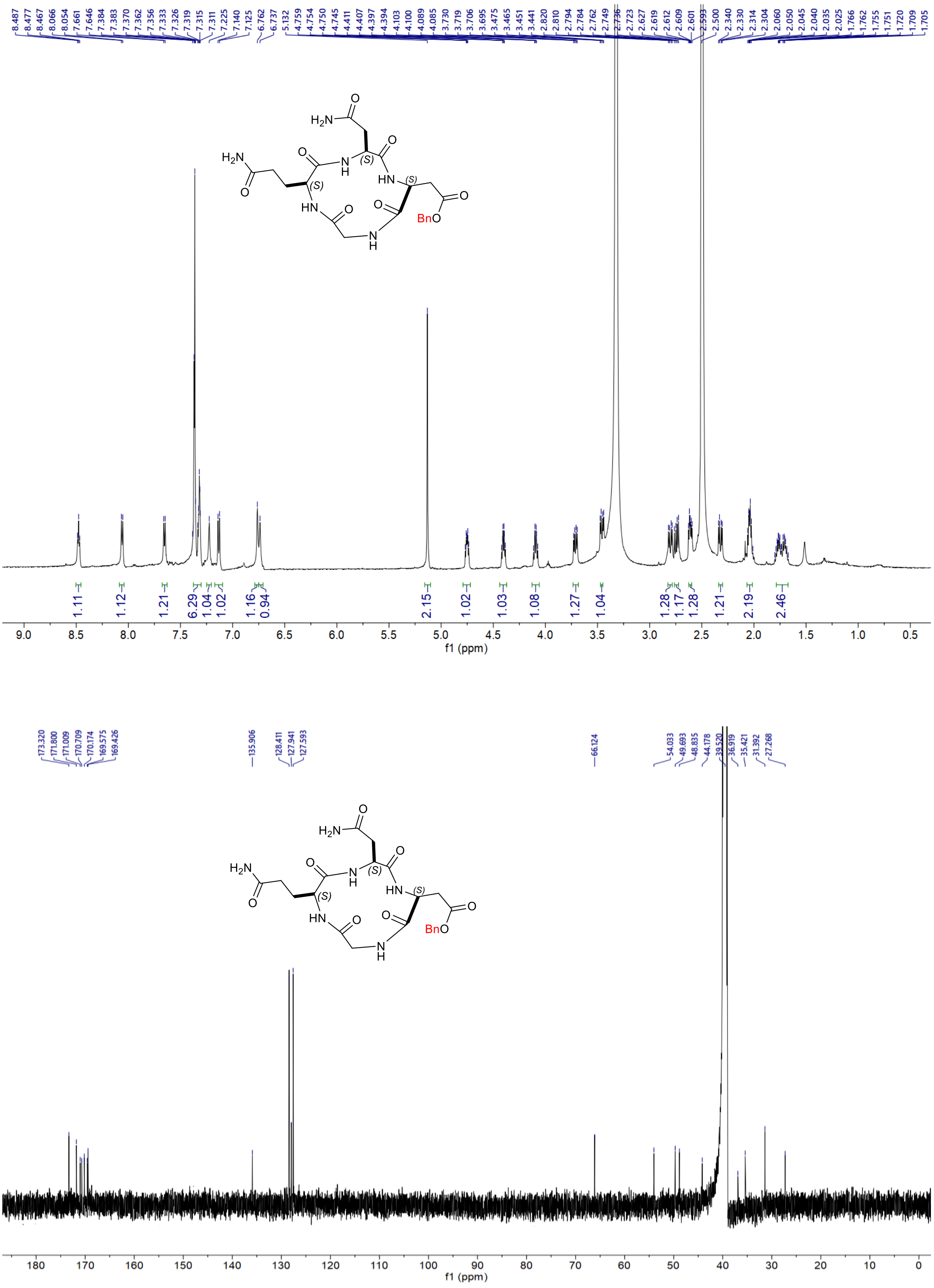


Figure S5G. NMR spectra of the purified **TH15** tetrapeptide.

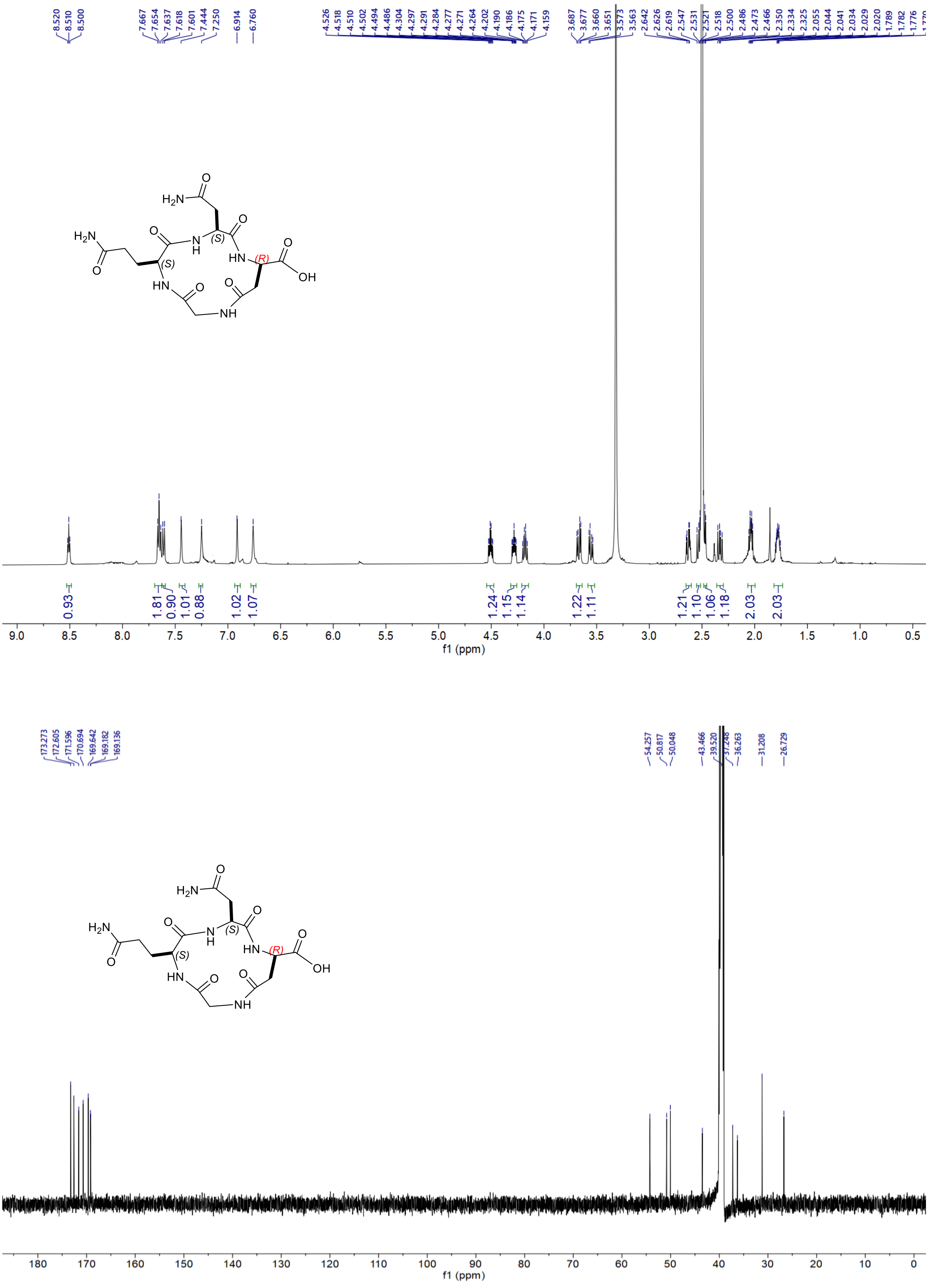


Figure S5H. NMR spectra of the purified **TH3** tetrapeptide.

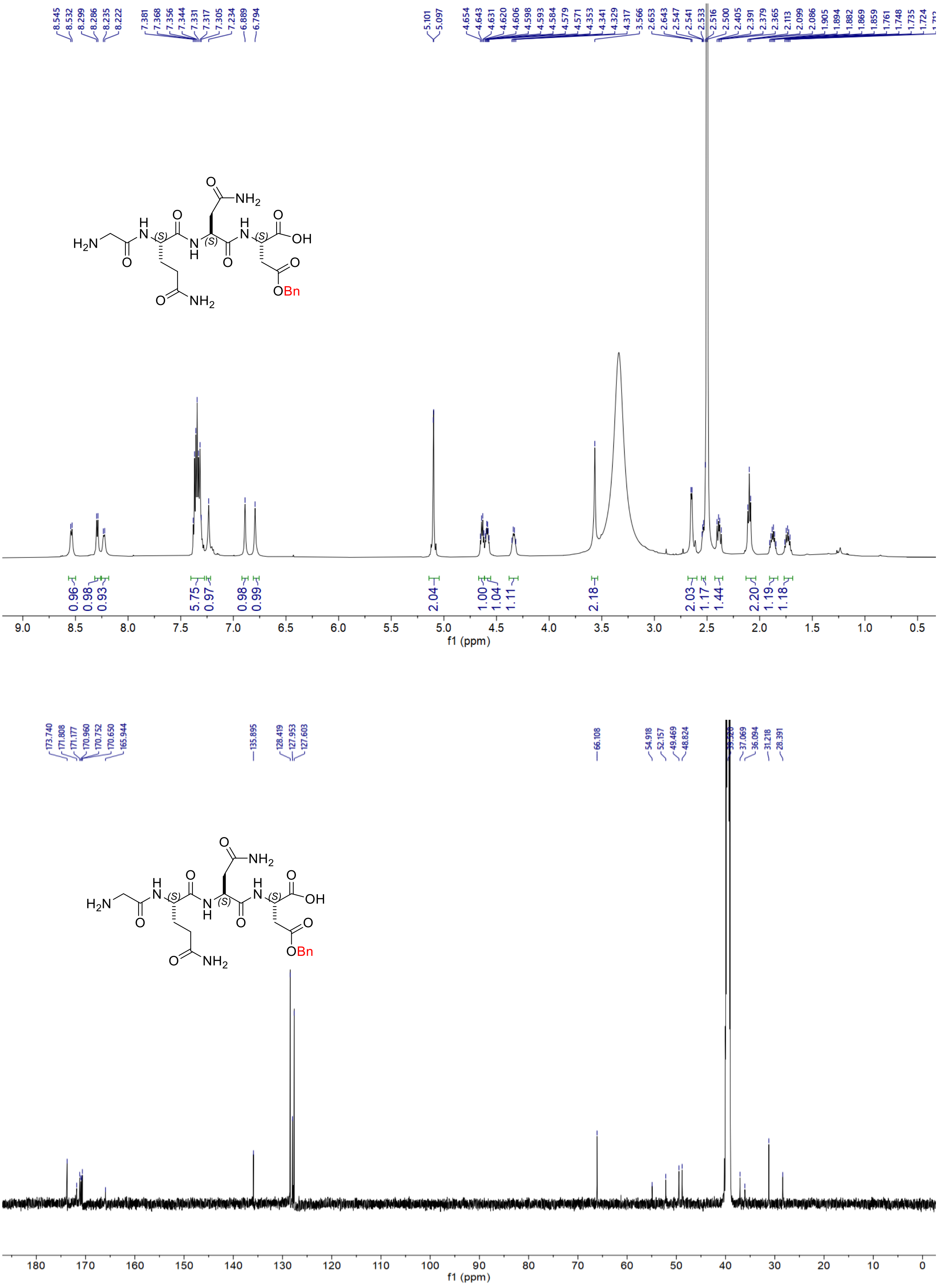


Figure S5I. NMR spectra of the purified **TH14** tetrapeptide.

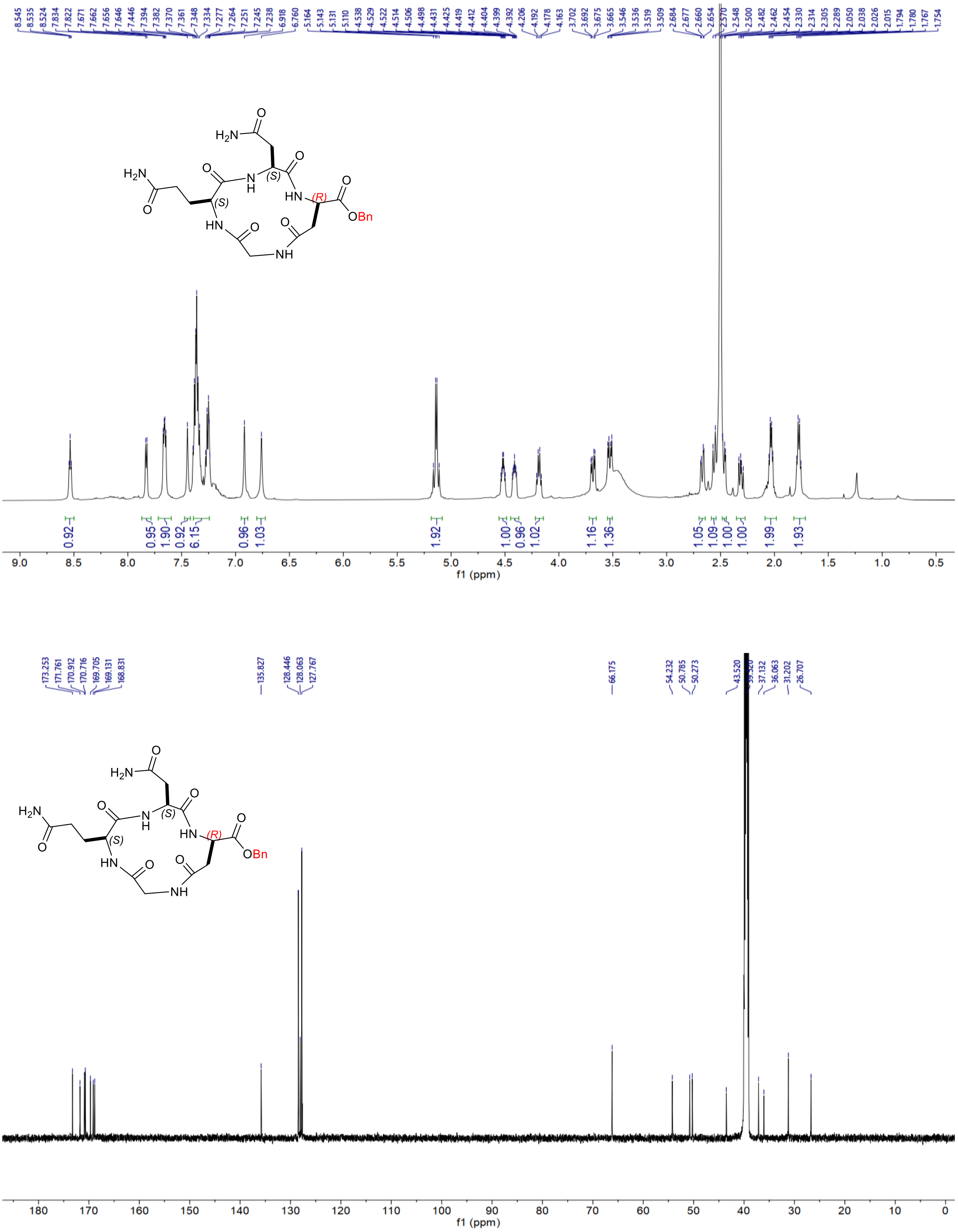


Figure S5L. NMR spectra of the purified **TH11** tetrapeptide.

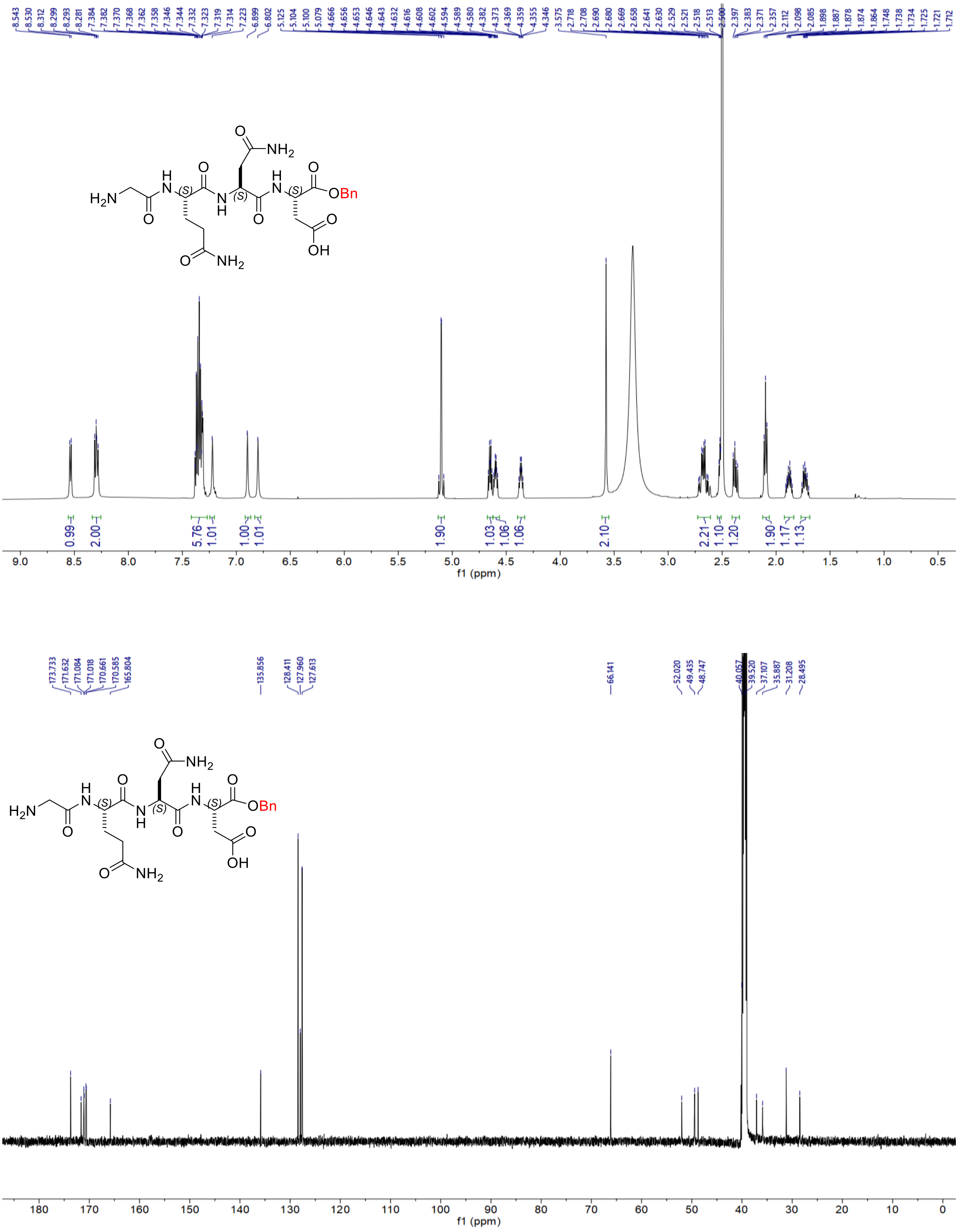


Figure S5M. NMR spectra of the purified **TH6** tetrapeptide.

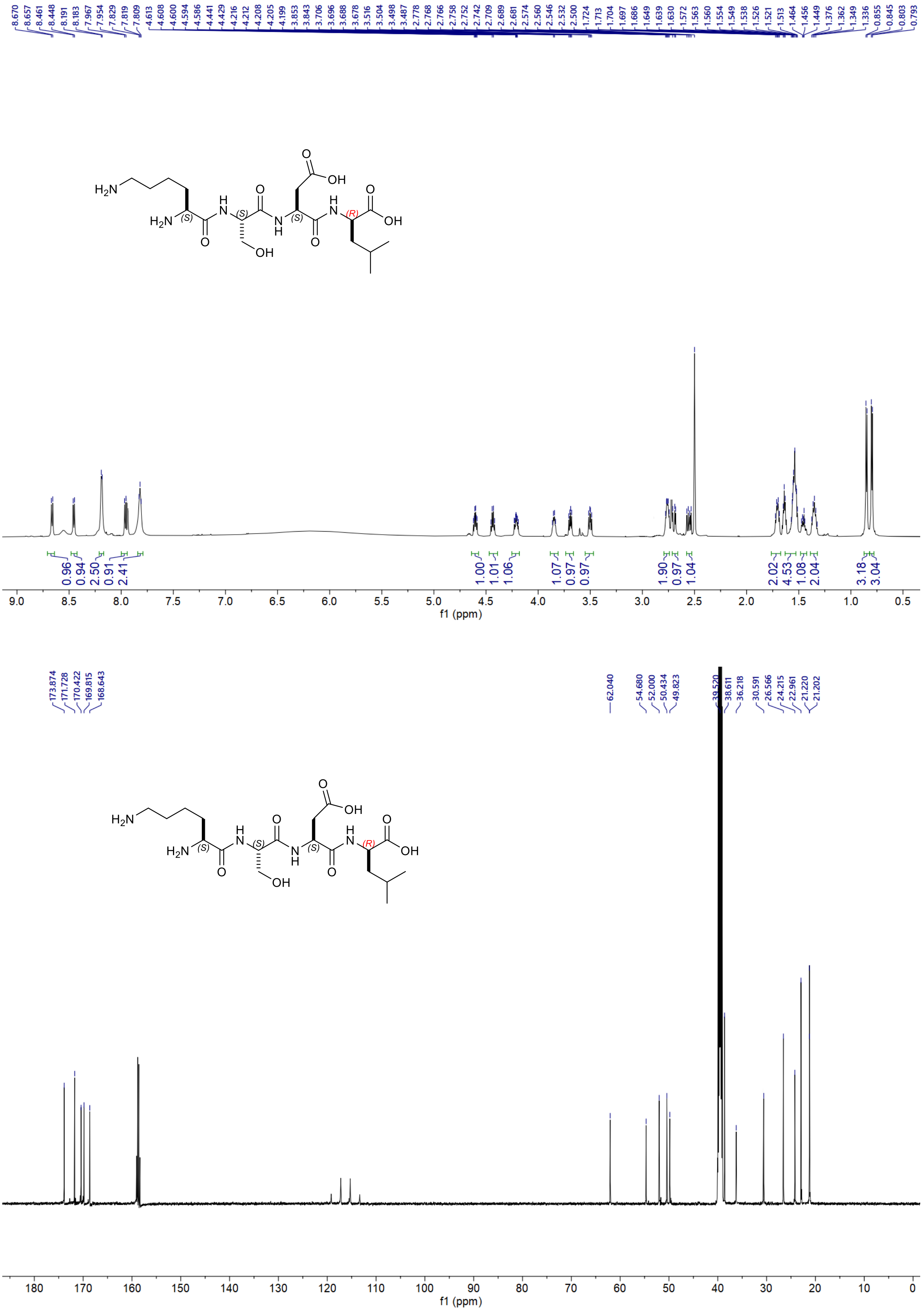


Figure S5N. NMR spectra of the purified **TH19** tetrapeptide.

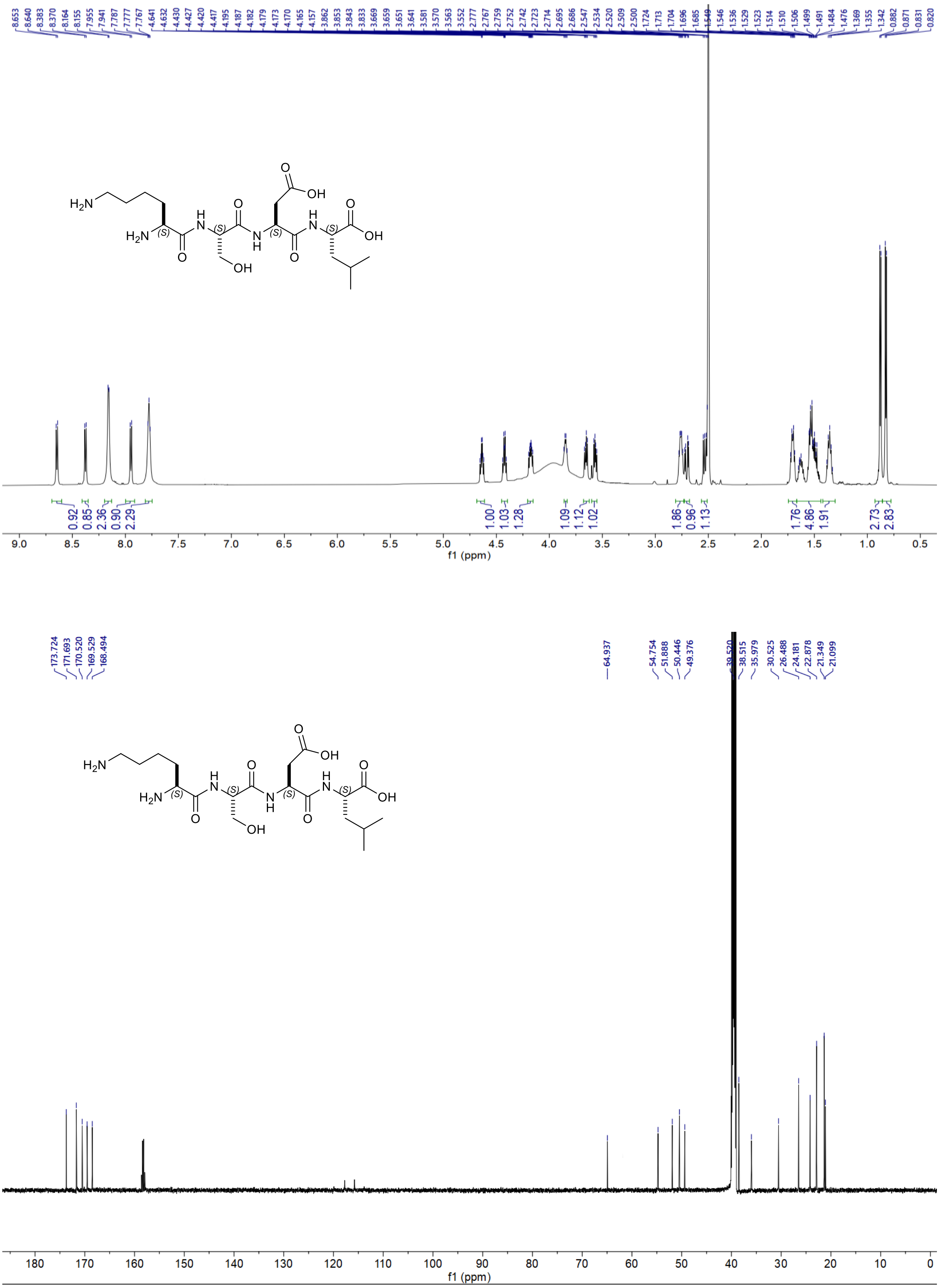


Figure S5O. NMR spectra of the purified **TH18** tetrapeptide.

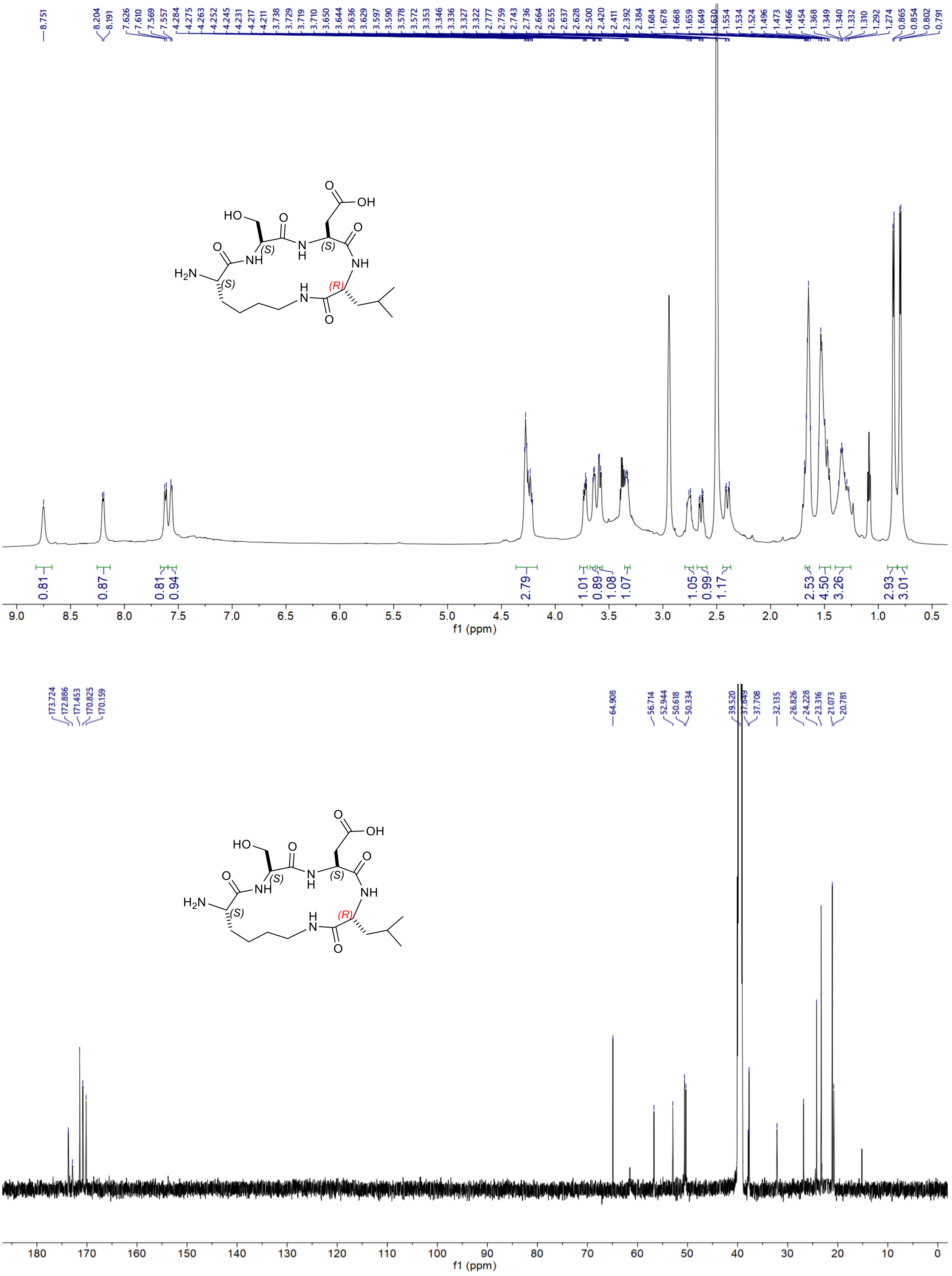


Figure S5P. NMR spectra of the purified tetrapeptide **TH5**.

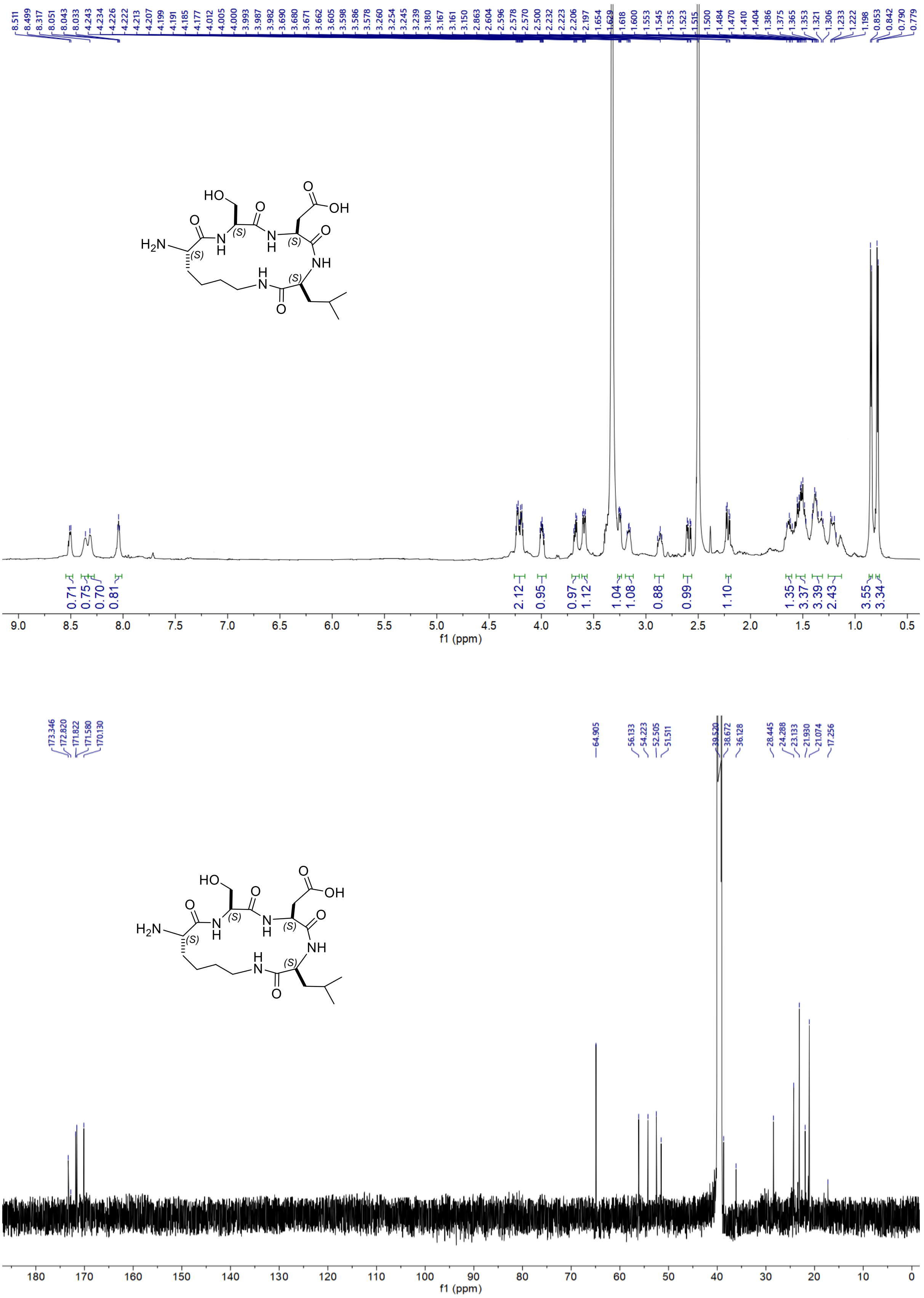
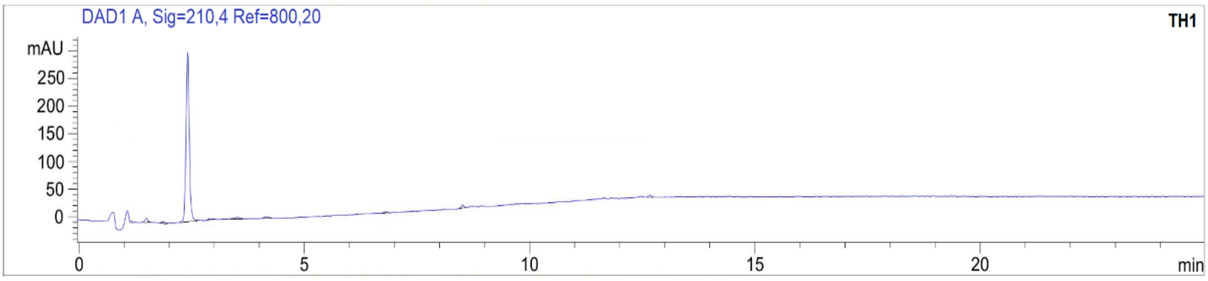
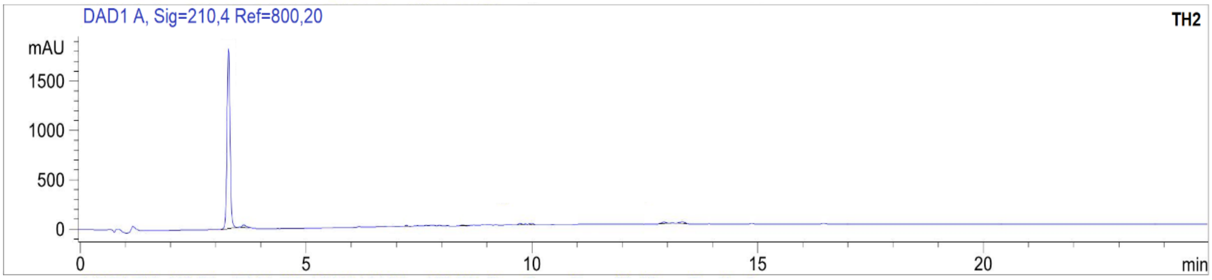
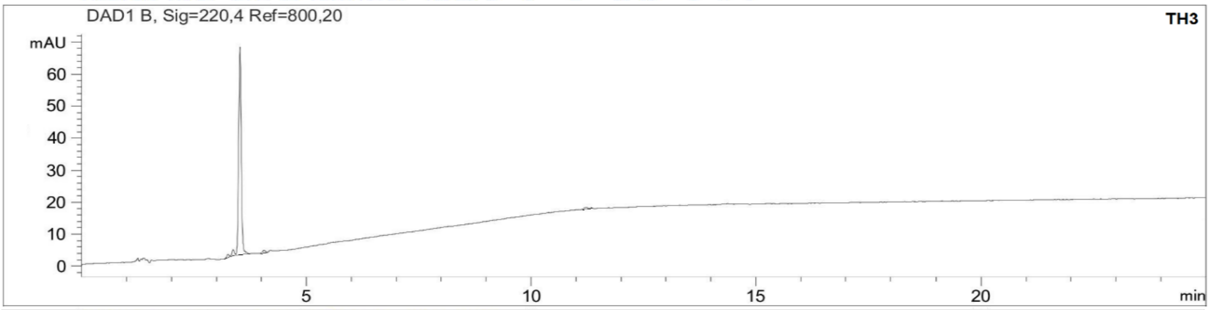


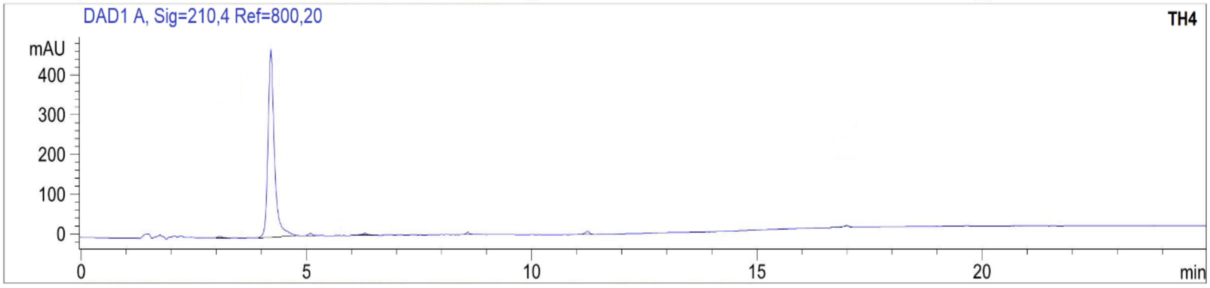
Figure S5Q. NMR spectra of the purified tetrapeptide **TH4**.

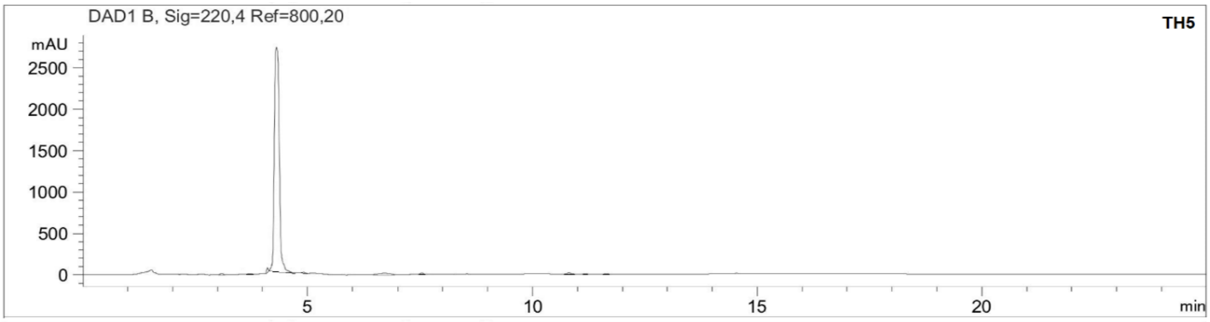
RP HPLC chromatograms:

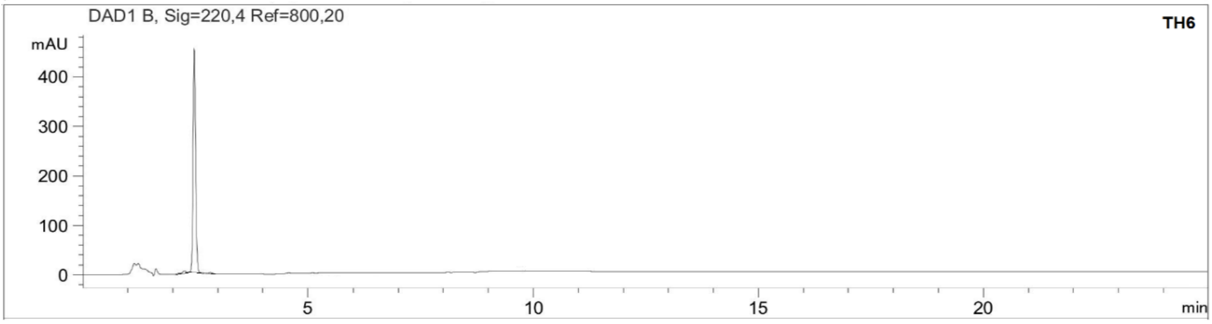

TH1

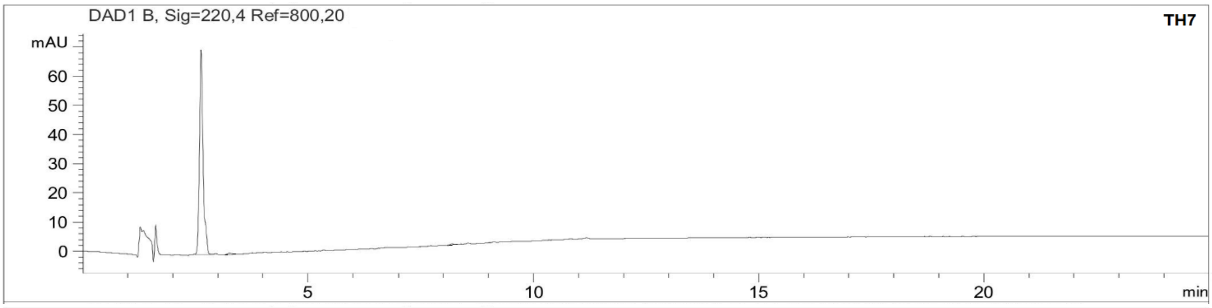

TH2

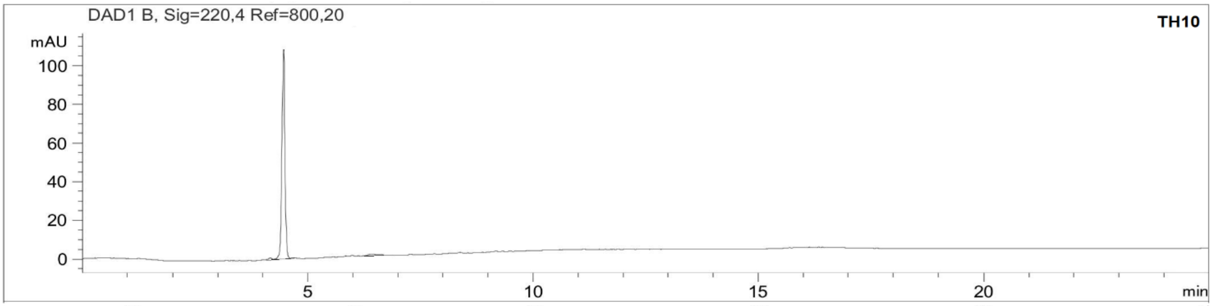

TH3

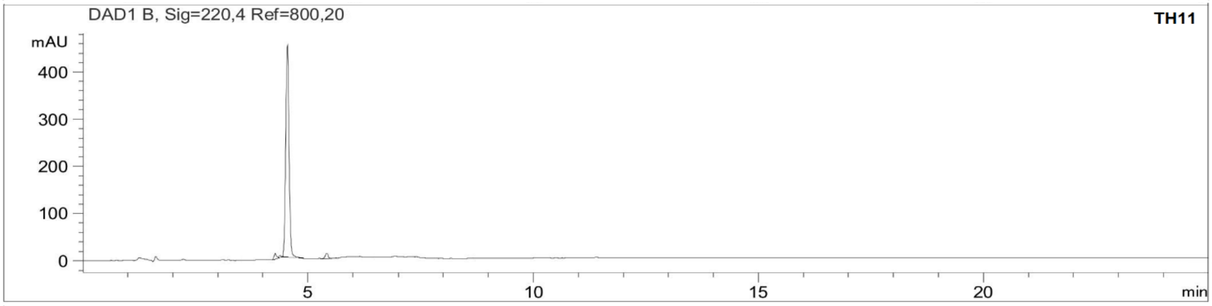

TH4

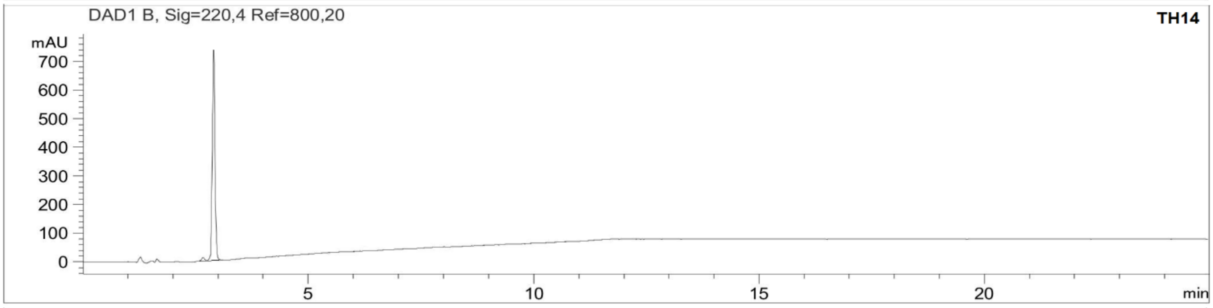

TH5

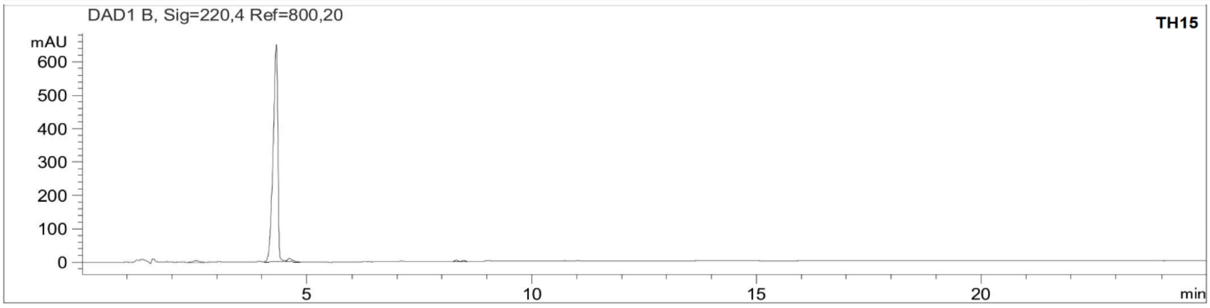

TH6

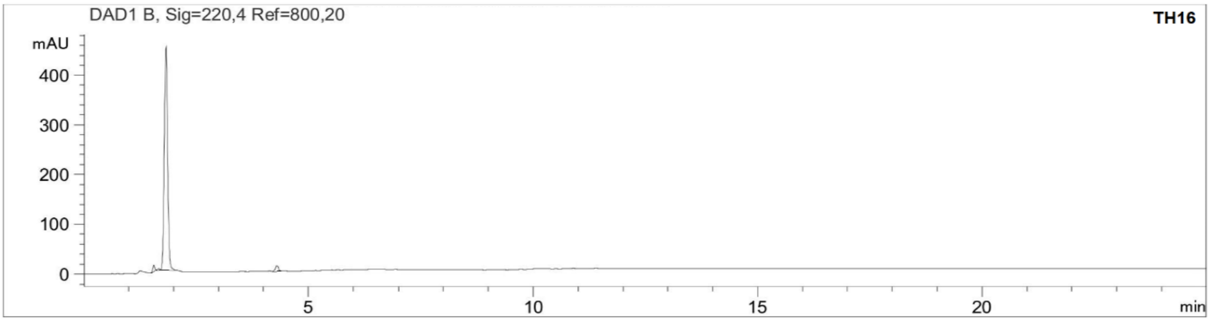

DAD1 B, Sig=220,4 Ref=800,20
TH7
mAU
60
50
40
30
20
10
0
5
10
15
20
min


DAD1 B, Sig=220,4 Ref=800,20
TH10
mAU
100
80
60
40
20
0
5
10
15
20
min


DAD1 B, Sig=220,4 Ref=800,20
TH11
mAU
400
300
200
100
0
5
10
15
20
min


DAD1 B, Sig=220,4 Ref=800,20
TH14
mAU
700
600
500
400
300
200
100
0
5
10
15
20
min


DAD1 B, Sig=220,4 Ref=800,20
TH15
mAU
600
500
400
300
200
100
0
5
10
15
20
min


DAD1 B, Sig=220,4 Ref=800,20
TH16
mAU
400
300
200
100
0
5
10
15
20
min

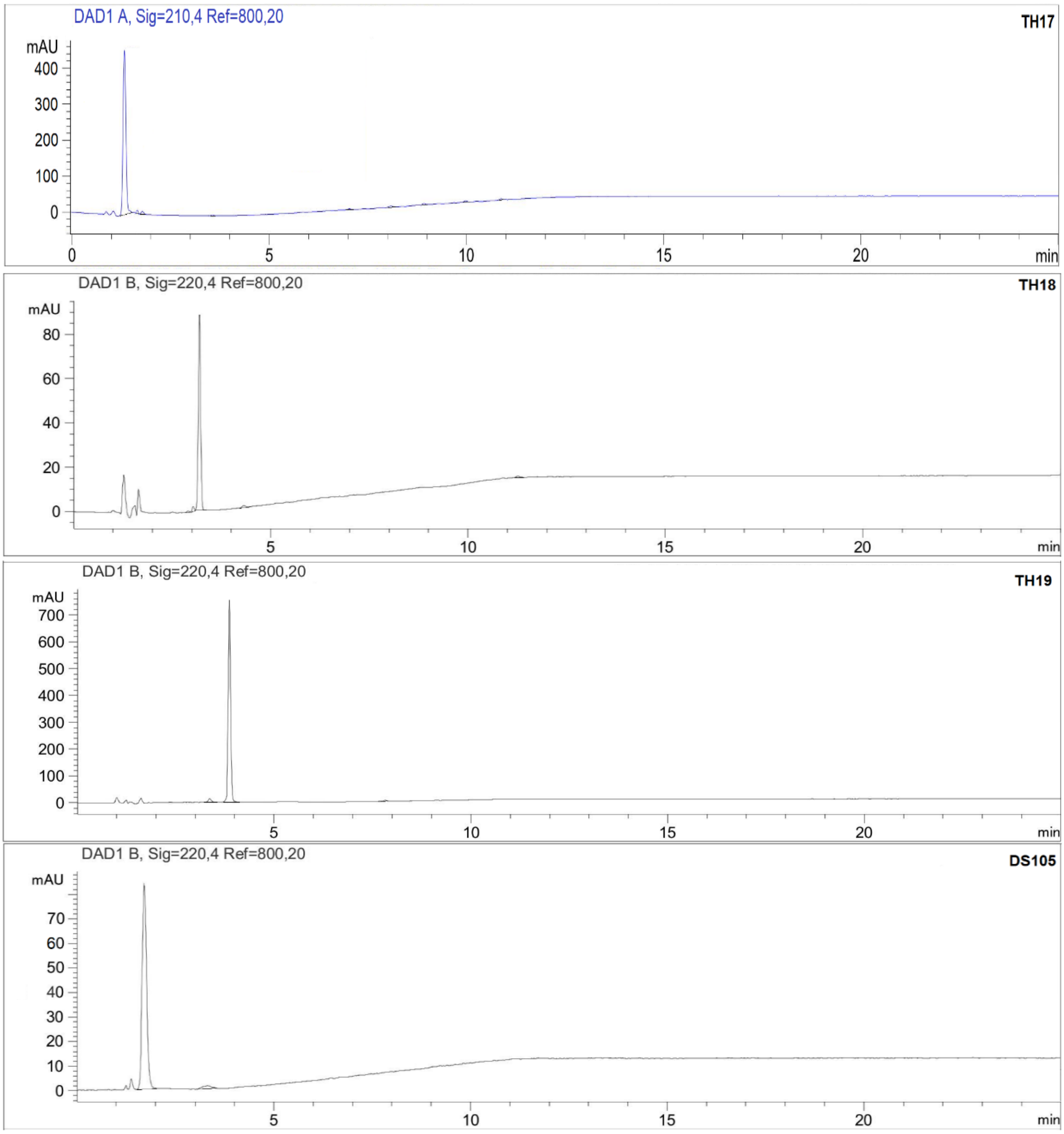


Figure S6. Analytical HPLC chromatograms of the indicated peptides.

Appendix-D

Expedient Recycling of Peptide Organocatalysts using a Nanocrystalline Hydroxyapatite Catching System

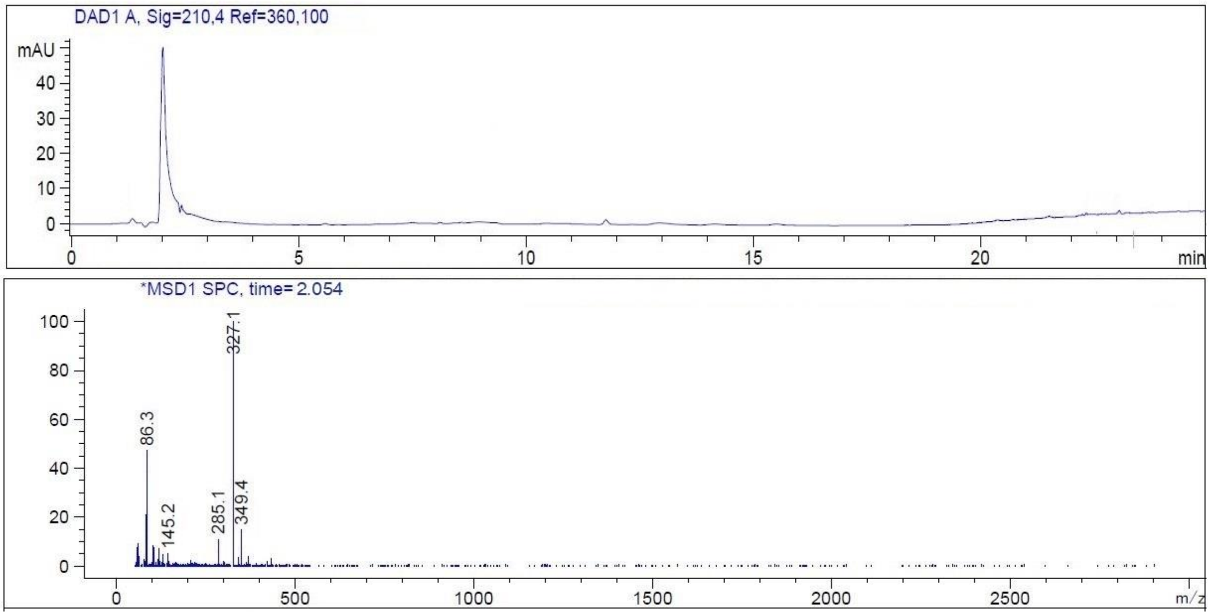


Figure S1. RP HPLC (XSelect Peptide CSH C18 column) and ESI-MS of peptide PPD. MS-ESI m/z calcd for $[C_{14}H_{23}N_4O_5]^+$ 327.2, found 327.1.

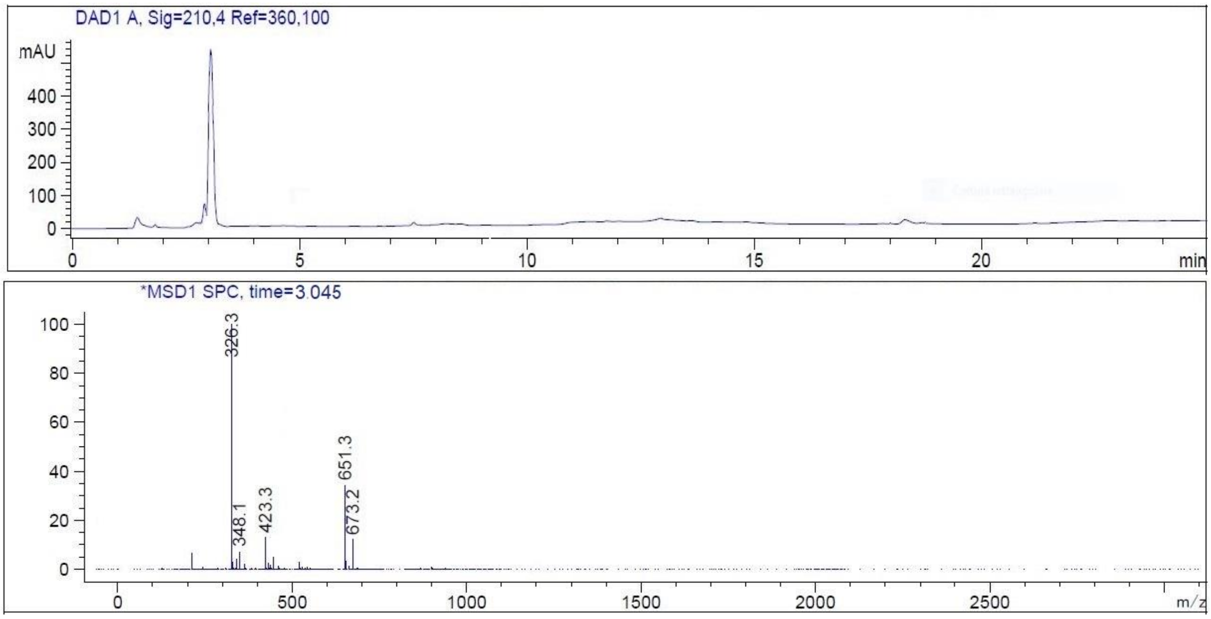


Figure S2. RP HPLC (XSelect Peptide CSH C18 column) and ESI-MS of peptide pPN. MS-ESI m/z calcd for $[C_{14}H_{24}N_5O_4]^+$ 326.2, found 326.3.

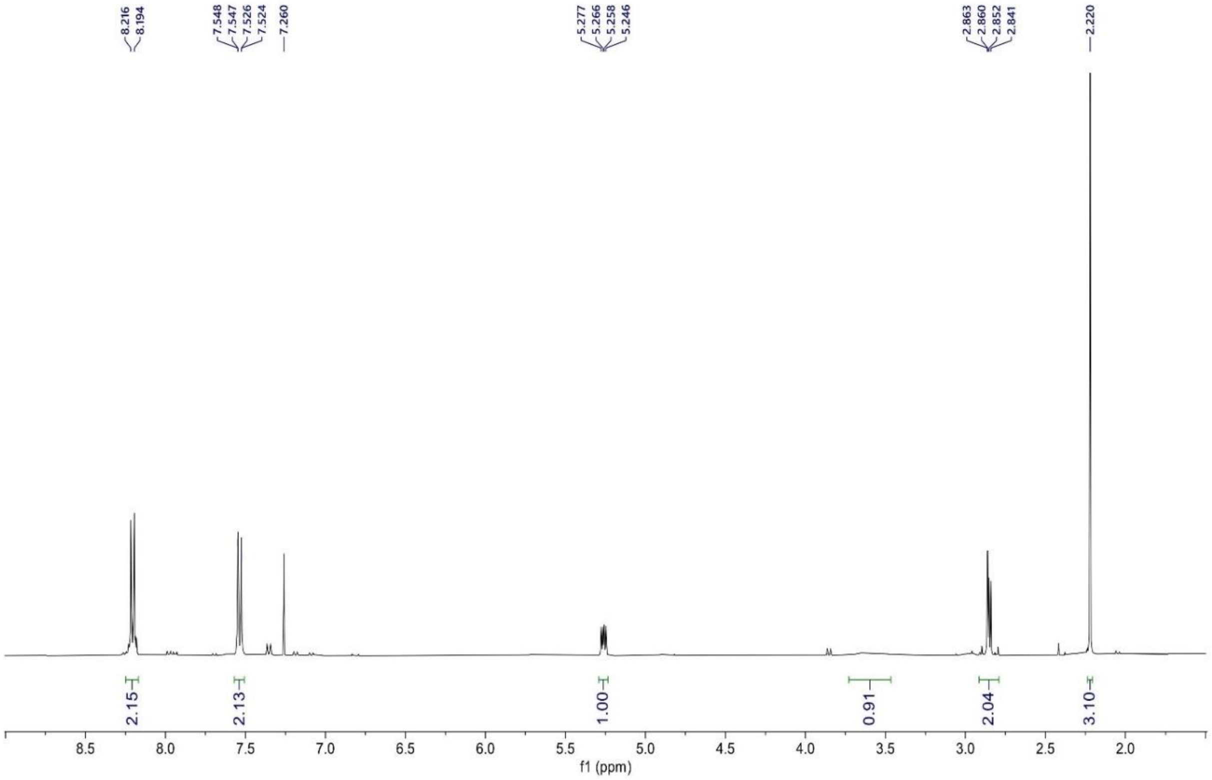


Figure S3. ^{1}H NMR analysis (600 MHz, CDCl$_3$) of Aldol reaction after isolation by flash chromatography.

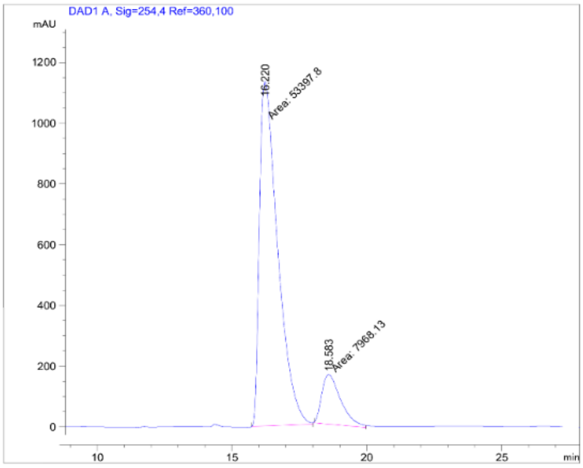


Figure S4. Chiral HPLC analysis of Aldol reaction. Conditions: OJ-H 80/20 *n*-Hex/IPA, flow rate 1 mL/min at 25 °C; λ 254 nm. Peak areas expressed as mAu x s.

Table S1. Results of the conjugate addition reactions of propanal (3 equiv.) using recycled 10:1 w/w HAp/PPD (PPD, 0.05 equiv.) in the presence of NMM (0.05 equiv.), at RT for 12 h.

Repeat	**HAp/PPD**	**Solvent**	**Yield (%)** [a]	**d.r.**[b]	**e.e. (%)** [c]
1	fresh	$CHCl_3$/iPrOH	95	10:1	83
2	recycled	$CHCl_3$/iPrOH	92	10:1	83
3	recycled	$CHCl_3$/iPrOH	86	9:1	84
4	recycled	$CHCl_3$/iPrOH	75	9:1	80
5	recycled	$CHCl_3$/iPrOH	66	9:1	80
6	recycled	$CHCl_3$/iPrOH	56	9:1	79
1	fresh	EtOAc	93	9:1	75
2	recycled	EtOAc	89	9:1	75
3	recycled	EtOAc	84	8:1	74
4	recycled	EtOAc	78	8:1	72
5	recycled	EtOAc	54	8:1	72

[a]After isolation by flash chromatography. [b]Determined by ^{1}H NMR spectroscopy. [c]Determined by chiral-phase HPLC analysis after isolation by flash chromatography.

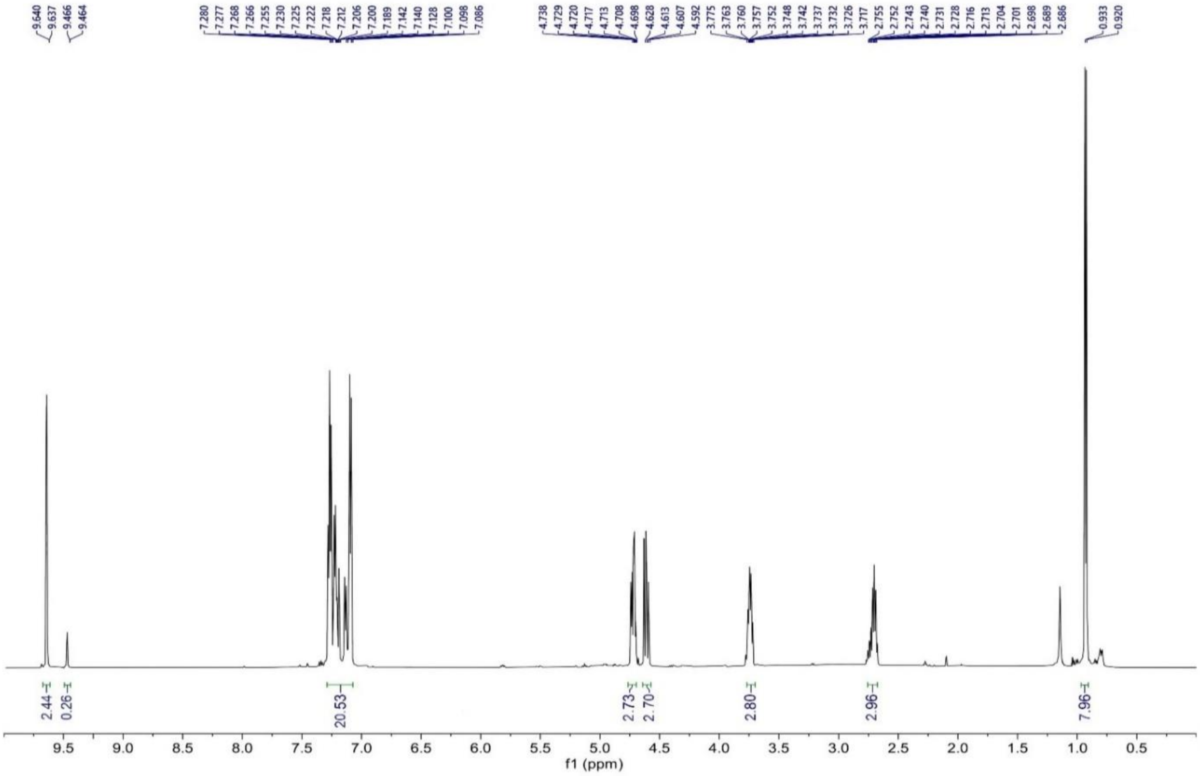


Figure S5. ^{1}H NMR analysis of the conjugate addition of propanal to nitrostyrene after isolation by flash chromatography.

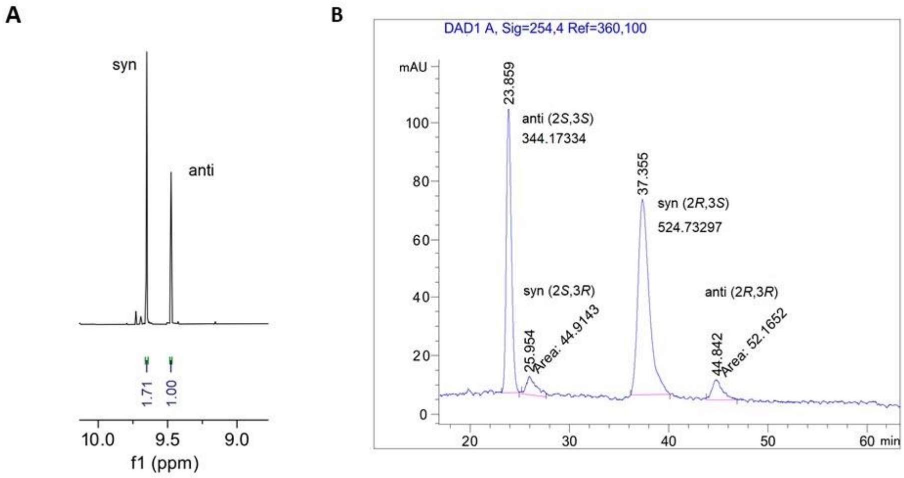


Figure S6. Conjugate addition of propanal to nitrostyrene; the analyses refer to a distinct fraction isolated by of flash chromatography, whose d.r. can be calculated from (A) the inset of ^{1}H NMR showing the diastereomeric aldehyde protons. (B) Chiral HPLC analysis. Conditions: OD-H 90/10 *n*-Hex/IPA, flow rate 1 mL/min at 25 °C; λ 254 nm. Peak Areas expressed as mAu x s.

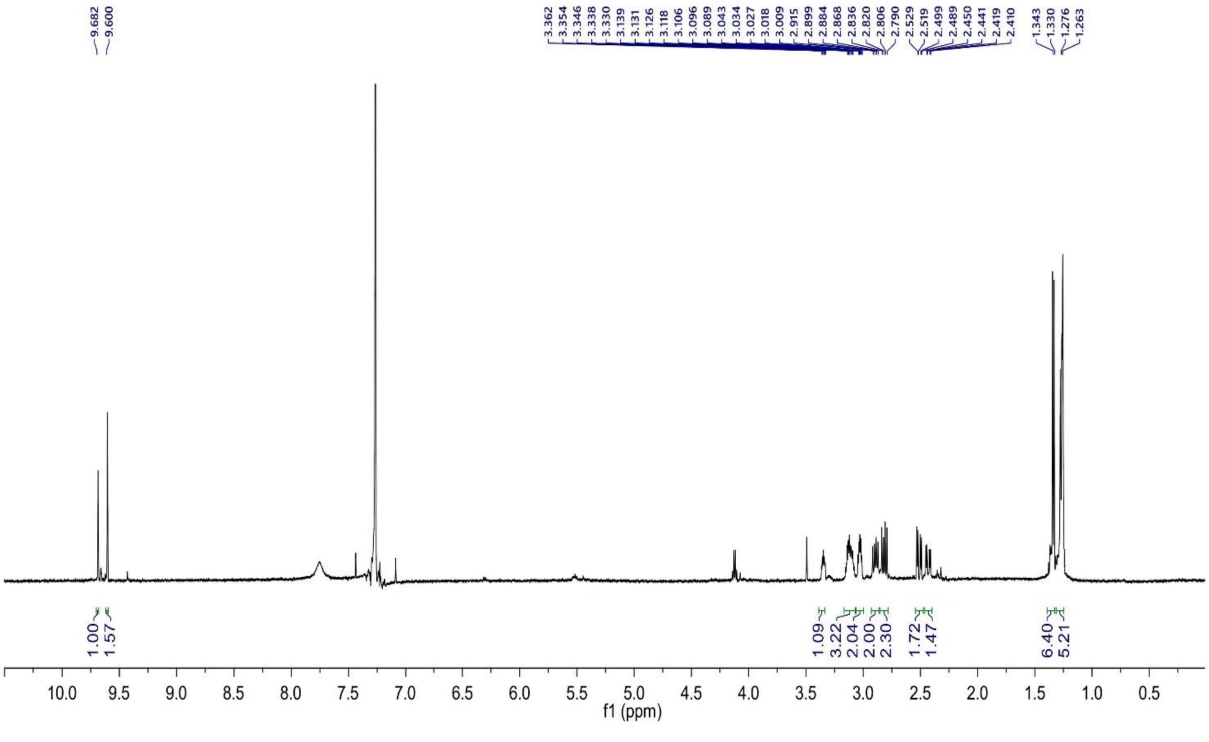


Figure S7. ^{1}H NMR analysis (600 MHz, $CDCl_3$) of the conjugate addition of propanal to maleimide, after isolation by flash chromatography.

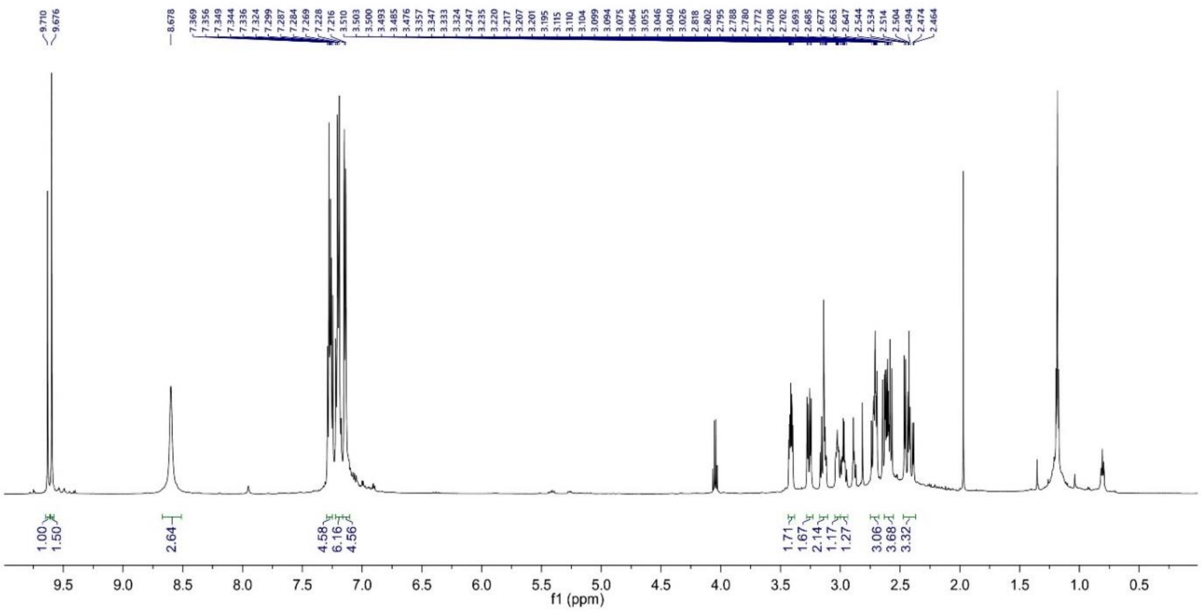


Figure S8. ^1H NMR analysis (600 MHz, $CDCl_3$) of the conjugate addition of hydrocinnamaldehyde to maleimide, after isolation by flash chromatography.

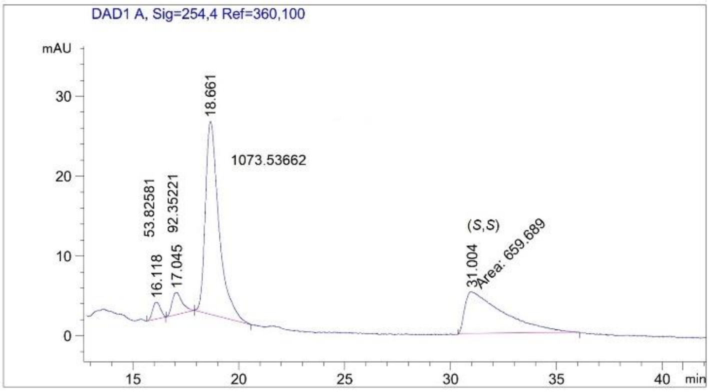


Figure S9. Chiral HPLC analysis of the conjugate addition of hydrocinnamaldehyde to maleimide. Conditions: AD-H 60/40 *n*-Hex/IPA, flow rate 0.5 mL/min at 25 °C; λ 254 nm. Peak areas expressed as mAu x s.

Acknowledgements

First and foremost, I would like to express my deepest appreciation to my supervisor, Prof. Luca Gentilucci, for your expertise and patient guidance throughout my PhD journey. Your insightful academic comments and timely feedback have been incredibly helpful. Moreover, your positive attitude and sense of humor have made this learning experience much more enjoyable. I am sincerely grateful for the countless hours you devoted to my development. Without your support, I would not have been able to complete this thesis. Thank you for everything; I have learned so much from you—not only in terms of academic knowledge but also in understanding the importance of approaching the surroundings with anticipation and a positive mindset.

I would also like to extend my sincere thanks to Prof. Jesús Martínez de la Fuente during my research mobility in BionanoSurf Group. Thanks for you provide me such a wonderful scientific family: Carlos, África, Adrián, Natalia, Maria, Maura, Bea, and Antonio for always being kind and helpful to me. I also appreciate Prof. Rafael and Valeria, along with all the group members, I really appreciate the time and efforts they invested in sharing their knowledge and expertise, which contribute a lot to my research project in Chapter 4, part 1.

Moreover, I also want to take this opportunity to thank everyone, who has helped, supported, and contributed to my research. Your collaboration and contributions have been extremely helpful, and without any of you, this thesis would not have been possible. Thank you for being part of this beautiful journey:

Prof. Daria Giacomini, Prof. Alessandra Tolomelli and Prof. Monica Baiula for your precious contribution in Chapter 2;

Dr. Michele Anselmi, Prof. Monica Baiula, Prof. Santi Spampinato, and Dr. Roberto Artali for your indispensable contribution in Chapter 3;

Prof. Mayra Paolillo, Prof. Luca Prodi, and Prof. Enrico Rampazzo for your help and work in Chapter 4, part 2.

Prof. Alejandro Hochkoeppler, Alessandra Stefan, Francesca Ruffolo, Valentina Rossi, Sofia Sordi, Prof. Giuseppina di Stefano, Federica Santino, Maurizio Brigotti, Claudia Scotti, Luisa Iamele, Hugo de Jonge, Fabrizio Dal Piaz, and Danilo Rocco Santarcangelo for your contribution and collaboration in Chapter 5;

Prof. Elisa Boanini, and Beatrice Valagussa for your efforts in Chapter 6;

I also want to thank my lovely lab partners: Marco, Danilo, Alice, Bea, Davide, Luisa, Francesco, Giusep, and Matia for their companion and support. I am very happy to spend time with them. With their companion, I feel relaxed and energized, making even the challenging days enjoyable. The positive atmosphere we created together made a significant difference in my research experience, and I truly appreciate each moment shared with them.

I also would like to express my gratitude to the Department of Chemistry "Giacomo Ciamician" for providing me a well-equipped research platform and community that facilitated my research endeavors. The assistance I received from the faculty members, administrative staffs, and technician has been invaluable.

For the PhD grant, I appreciate the China Scholarship Council (CSC).

Lastly, I want to express my heartfelt appreciation to my friends and family. Their unwavering support and encouragement have been my anchor throughout this journey. They provided comfort during challenges, reminding me of the importance of resilience and perseverance. Their belief in my abilities has motivated me to strive for excellence, and I am truly fortunate to have them by my side. Thanks for being with me all the time.

轻舟已过万重山，未来加油！

TINGTING HE
08/02/2025