UNIVERSITÀ DEGLI STUDI DI MILANO

CORSO DI DOTTORATO IN SCIENZE FARMACEUTICHE (XXXV ciclo)

DIPARTIMENTO DI SCIENZE FARMACEUTICHE (DISFARM)

COMPUTATIONAL DESIGN OF PEPTIDES AND PEPTIDOMIMETICS AS POTENTIAL THERAPEUTIC AGENTS TO SATISFY PRESSING MEDICAL NEEDS

ENRICO MARIO ALESSANDRO FASSI

Matricola: R12492

TUTOR: Prof. GIOVANNI GRAZIOSO

COORDINATORE DEL DOTTORATO: Prof. GIULIO VISTOLI

A.A. 2021/2022

TABLE OF CONTENTS

1. INTRODUCTION

Peptides play critical roles in human physiology, acting as neurotransmitters, hormones, growth factors and antibacterial agents [1]. They have significantly influenced the development of the modern pharmaceutical industry and contributed to the progress of biological and chemical research [2]. In particular, all the physiological studies made in the first half of the 20th century on peptide hormones such as insulin, oxytocin, gonadotropin-releasing hormone (GnRH), and vasopressin, determined significant advances in the fields of pharmacology, biology, and chemistry [3]. The first commercially available peptide drug is represented by insulin (1923) for treating diabetes (**Figure 1**) and it is still on the "World Health Organization's List of Essential Medicines", the most effective and safe medicines needed in the health system [4].

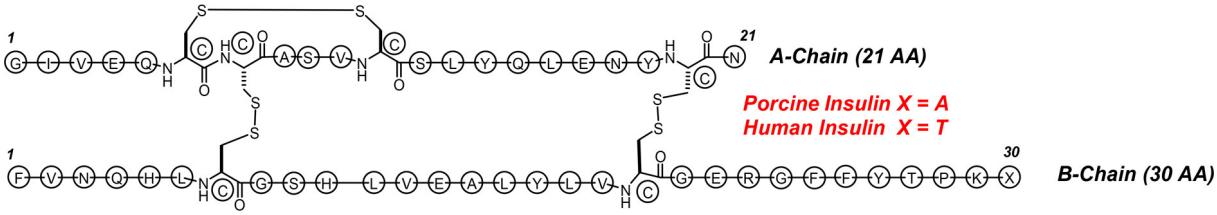


Figure 1. 2D structure representation of insulin [5].

Nevertheless, peptide drug discovery has been considered only a niche area until the end of the 20th century mainly due to the perception by pharmaceutical companies that peptides' drawbacks outweigh their advantages [5]. However, over the past two decades, peptide drug discovery showed a dramatic increase of interest, since many pharmaceutical companies started to appreciate the fundamental role played by the peptides for the treatment of unsatisfied medical needs, probably due to the recent scientific findings and achievements made by the academia and private researches on the improvement of the peptide pharmacokinetic profile [6]. In fact, over the last years, the number of FDA-approved peptide-based drugs has dramatically increased, leading to more than 100 peptide drug approvals until July 2022 [7].

The use of peptides as drug candidates has some advantages and disadvantages:

(1) peptides are the natural physiological messengers for most endocrine signalling pathways; therefore, it is possible to develop extremely potent and selective drugs;

(2) peptides are unable to cross membranes, consequently, their application is restricted to the extracellular compartment and transmembrane targets;

(3) peptides are also generally unable to cross the blood-brain barrier (BBB), which, from the toxicology point of view, it could be considered as an advantage but, in concrete, it precludes reaching targets in the Central Nervous System (CNS);

(4) peptides are metabolically unstable and, as a consequence, they have short plasma half-lives because, after they are produced in response to a biological signal and performed their functions, they are rapidly metabolized (i.e., the signal is turned off). In addition, being cleared primarily through proteolytic degradation and renal filtration, peptides are endowed with suboptimal pharmacokinetic properties (in particular, poor oral bioavailability), needing considerable efforts to discover new analogs with suitable plasma exposure. However, the proteolytic instability and the rapid clearance suggest that they cannot be accumulated in tissues;

(5) hepatic metabolism of peptides is not worthy of note and therefore drug-drug interactions and non-mechanistic-based toxicology are rarely observed. These issues, instead, are frequently observed in small molecules [5].

The small molecules dominate the global drug market showing several advantages compared to peptides, such as the low cost and the final price (the HPLC purification cost of peptides is still high in comparison with small molecules), the possibility of oral intake and their capacity (not in all cases) to penetrate membranes. These features allow small molecules to reach a wide range of biological targets [8]. However, peptides can show some advantages compared to small molecules in the optimization stage:

(1) in the most of cases, peptides from natural source or affinity matured hits show a high affinity and potency, which must be merely conserved during the optimization;

(2) the polymeric nature of peptides allows the optimization of each residue. Of course, the ADMET properties of peptides are restricted.

Nevertheless, despite the poor pharmacokinetic properties of peptides, this class of compounds can even represent a valuable alternative to small molecules and biological

treatments. However, in the majority of cases, they need several optimization processes regarding their chemical structures, in order to be used as potential drug candidates.

Among them, the most promising strategies are:

(1) the protection from degradation at the *C*- and *N*- terms by carboxy- and amino-peptidases by capping process (acetylation and amidation, respectively) [9];

(2) the proteolytic stabilization through directed backbone modifications (substitution by *D*-aa, α-Methylation, *N*-Methylation or using β-aa) [10,11];

(3) the proteolytic stabilization or solubility improvement through the side chain modification (for example, aromatic residues can be substituted by different unnatural heterocycles) [12];

(4) the chain cyclization, preventing the degradation of peptides by proteases and limiting the access of enzymes to the backbone amide bonds [13].

In particular, cyclic peptides possess a degree of structural pre-organisation which is not observed in linear peptides, significantly restricting the conformational flexibility [14]. If the cyclic peptide is pre-organised into the correct bioactive conformation, it is able to efficiently reduce the entropic cost of binding to the biological target, compared to the linear analogue [15].

The most recent findings regarding peptide delivery also determined the current growing interest in the peptide drug discovery process. In fact, parenteral administrations are the only canonical delivery methods for therapeutic peptides with all the issues related and the low compliance by patients [16]. Therefore, many drug carrier systems are currently being developed with the aim at increasing the oral bioavailability. Among them, the Nanoparticles (NPs), which are solid particles with sizes ranging from 10 to 1000 nm [17] encapsulate peptides or proteins into a polymeric matrix, protecting against the hydrolysis and the enzymatic degradation [17]. NPs are often constituted by PLA (polylactic acid), PLGA (polylactic-co-glycolic acid) (**Figure 2**), chitosan, poly-alkyl-cyanoacrylates or gelatin, which are non-toxic, non-thrombogenic, non-immunogenic, non-inflammatory, stable in blood and highly biodegradable [16].

Significant clinical successes were also reached using the slow-release subcutaneous polymeric depot technologies, such as PLGA microspheres [18], or the osmotic pumps, which are able to release the drug with a controllable zero-order kinetic for up to a year [19].

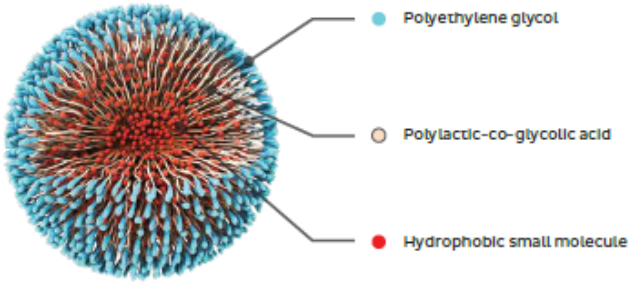


Figure 2. Representation of a PLGA (polylactic-co-glycolic acid) nanoparticle.

On the other hand, it is possible to mimic the biological effects, the conformational properties of the most active peptides by designing ***peptidomimetics***. In fact, if properly designed, they are capable to reproduce the secondary structure and the pharmacological effects of the parent peptides, without containing the fleeting amide bonds [20]. The conformation assumed by peptidomimetics represents their key feature, not only to enhance target binding affinity/selectivity and to confer cell-membrane permeability for targeting protein-protein interactions in cells, but also to ameliorate the pharmacokinetic drawbacks of peptides [21]. The classification of peptidomimetics has evolved together with the progress made over the years. The modern classification proposed by Grossmann refers to four classes (from A to D) of peptidomimetics, based on their similarity to the native substrate (**Figure 3**) [22].

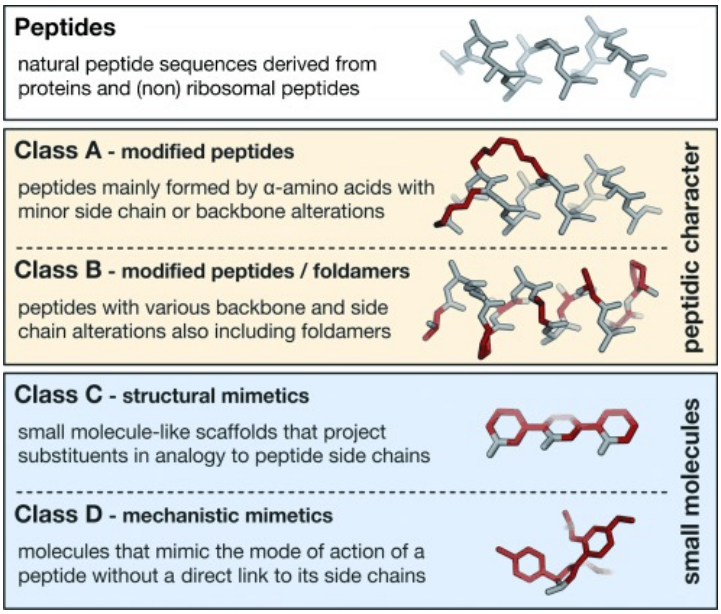


Figure 3. Grossmann classification of peptidomimetics. As an example, an α-helical peptide and corresponding helix mimetics are shown. Modifications are highlighted in red [22].

The mimicry of secondary structure features, such as β-turn, γ-turn, α-helix, or β-strand, represents one of the most well-investigated strategies in order to develop peptidomimetics targeting peptide-protein or peptide-receptor interactions [23]. This is mainly because the bioactive peptides often shape these kind of secondary structure conformations to be recognized by the targeted proteins. For example, receptors such as G-protein coupled receptors (GPCRs) often recognize a β-turn conformation of their endogenous peptide ligands [24] while the proteolytic enzymes are likely to recognize an extended β-strand conformation of their target peptides [25]. A clear and exhaustive review about various peptidomimetic scaffolds is available in literature [26]. Examples of β-turn mimics developed by Hirschmann and Smith are shown in **Figure 4**.

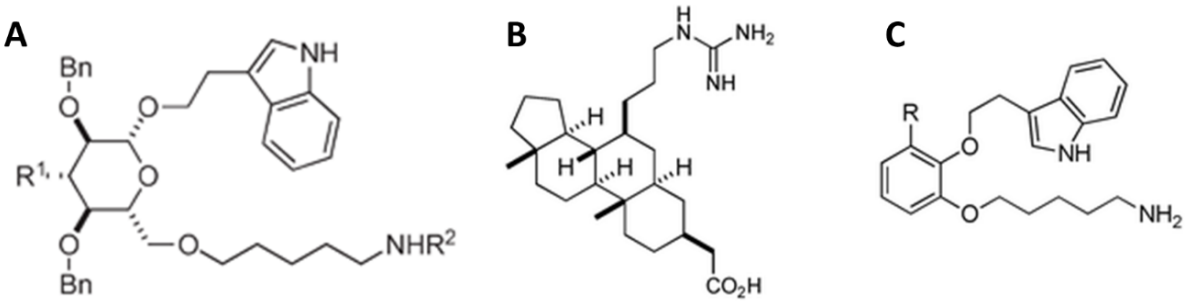


Figure 4. Examples of minimalist peptidomimetics designed by Hirschmann and Smith. The β-turn mimetics in which sugar (A) [27], steroid (B) [28] or catechol (C) [29] backbones were used to hold relevant sidechains at appropriate distances.

2. STRATEGIES FOR PEPTIDE DRUG DESIGN

Drug discovery is a very expensive and time-consuming process. Traditional approaches are based on a systematic synthesis and screening of large number of compounds to identify potential candidates. Nevertheless, over the last decade, increasing efforts have been dedicated to the application of computational weaponry, combined with biological experiments, in the field of drug discovery, design, development, and optimization [30]. The available computational methods play a crucial role in the clarification of the molecular recognition events between the protein target and candidate hits leading, with the final goal to design lead compounds active on the identified target [31]. As a consequence, the "Computer-Aided Drug Design" (CADD) approaches have been widely employed in Lead Identification and Lead Optimization stages of the drug development process. Compared to the most traditional drug discovery methods, CADD are green methods that dramatically

reduce the time and cost necessary for the drug development process. Consolidated molecular modeling techniques are used to predict the ligand binding affinity, the position, and the orientation of a ligand in a ligand-protein complex [32].

CADD methods mainly involve two types of approaches:

1) Ligand Based Drug Design (LBDD). This approach is applied when the three-dimensional (3D) structure of biological target is unknown. It is considered an indirect approach because it relies on the investigation of known molecules interacting with the biological target of interest. The known ligands are used to develop 3D Quantitative Structure Activity Relationships (3D-QSAR) or pharmacophore models, which represent the most important and widely used tools in LBDD. In fact, they are able to offer predictive models suitable for the lead identification and optimization processes [33].
2) Structure Based Drug Design (SBDD). This approach is applied when the structures of the drug target have been experimentally deciphered by X-ray diffraction [34], cryogenic electron microscopy (cryo-EM) [35], Nuclear Magnetic Resonance (NMR) [36], or they are predicted by homology modeling techniques. The latter approach can be applied only if a high sequence homology exists between the target protein and at least one solved structure [37]. Homology modelling can be divided in the following steps: (1) identification of homologous protein with at least one known 3D structure used as template; (2) sequence alignment of target and template proteins; (3) generation of models for the target based on the 3D structure of the template and the alignment; (4) models refinement, selection, and validation [38]. In the last decades, homology modeling has become the principal option to obtain the 3D structure of the drug target in the absence of crystal structures. In the SBDD approach, candidate drugs are designed through a detailed analysis of the protein binding sites, with the final goal to design ligands with high affinity and selectivity. Many successful applications have been reported in the field of molecular docking-based virtual screening [39,40].

In this PhD research thesis, only SBDD computational techniques have been adopted since in all the case studies the solved crystal structures of the drug targets were available. The computational protocols applied in this study are explained in detail in the next paragraphs.

2.1. Computational protocols applied

System selection and preparation

The first step of a SBDD project involves the acquisition of the target structure model, which commonly consists of a large biological molecule (protein, DNA, or RNA) [41]. The structures of these macromolecules can be easily retrieved from the Protein Data Bank (PDB) [42], which provides access to the atomic coordinates obtained by experimental methods. However, it is not unusual that the experimental 3D structure of the studied target is not complete, since many protein loops cannot be solved due to their high conformational mobility. To overcome this issue, computational prediction methods, such as homology modeling, can be used to obtain the complete the 3D structure of target macromolecules [43].

The "*Protein Preparation Wizard*" of Maestro (Schrödinger, LLC, New York, NY, 2021), was used to convert the raw PDB structure into all-atom, fully prepared protein model readable by the most common computational software. For example, it was used to add the missing hydrogen atoms, to correct the metal ionization states, to ensure the proper formal charge and force field assignment, and to add the residues missing sidechains. Then, the system was optimized by determining the most likely ligand protonation state, as well as the energy penalties associated with alternate protonation states, also determining the optimal protonation states for protein histidine residues. Finally, the system was subjected to a restraint minimization allowing hydrogen atoms to be energy minimized, while allowing for sufficient heavy-atom movement to relax strained bonds, angles, and clashes.

Once optimized the target structure, molecular docking of ligands (i.e., peptides or small molecules) in the active site can be accomplished.

Molecular Docking

Molecular docking is a key component of the CADD toolbox, since it is a computational procedure aiming at predicting the favored binding modes of a molecule (i.e., peptides or small molecules) within the putative active site of a macromolecular target (receptor). Since nowadays more and more protein structures are experimentally determined the usage of molecular docking in the last decades is dramatically increased as a valuable tool in the drug discovery process [44].

Numerous docking programs are available. Among them, the most commonly used are AutoDock, DOCK, GOLD, MOE-Dock, HADDOCK, GLIDE, and FLEXX [44].

However, all the different programs share two fundamental stages:

1) Pose generation: in which different orientations and conformations of the ligands within the protein binding site are predicted;

2) Ranking step: in which the ligand-protein binding affinity of the predicted poses are estimated using several scoring functions. They must recognize the best binding poses in the pool of predicted poses by assigning them the highest docking score.

The reliability of docking can be evaluated by analyzing the accuracy of the generated poses by comparing the best ranked solutions with the resolved complex structures.

In all the case studies reported in this PhD research activity, we decided to use GLIDE software [45] (Schrödinger, LLC, New York, NY, 2021), since it demonstrated to be one of the best predictive tool, as reported by many comparison studies, conducted to evaluate the relative performance of different docking programs [46,47,48]. In addition, GLIDE permits the tuning of the docking parameters to be adapted on the different systems, to obtain the best balance between speed and accuracy. As a further demonstration of the GLIDE algorithm accuracy, the paper first reporting on Glide [45] is the second most cited paper in the *Journal of Medicinal Chemistry.* In addition, many citing references report on the application of GLIDE in drug discovery projects.

Grid-based Ligand Docking with Energetics (GLIDE) software

The purpose of GLIDE software is to explore different conformations and orientations of the ligands within the target binding site. It is characterized by three docking modes which provide an array of options in the balance of speed *vs.* accuracy for most situations. High-Throughput Virtual Screening (HTVS) mode to dock compounds at a rate of about 2 seconds/compound. Standard Precision (SP) mode performs exhaustive sampling and is the recommended balance between speed and accuracy, requiring about 10 seconds/compound. GLIDE Extra Precision (XP) mode employs a different sampling approach making it the most accurate method; however it can dock compounds at a rate of about 2 minutes/compound [49].

In a standard docking protocol, the ligand poses generated by the GLIDE algorithm are refined in the binding site by the OPLS3 force field [50], in case of SP or XP calculations, or by the OPLS2005 force field during HTVS calculations. GLIDE uses Emodel [45] scoring function to pick the best pose of each ligand (pose selection), and then uses GlideScore (*Gscore*) [49] function to rank all these best poses in order to clarify which compounds strongly bind to the protein target (i.e., are endowed with the lowest *Gscores*).

Before performing docking calculations, the *docking grid* must be generated. The grid set up phase involves the generation of a grid that represents the shape and properties of the target using the OPLS2005 force field. The grid center can be defined by the centroid of a ligand in complex with the target in the X-ray structure or by the centroid of a set of selected residues that best represent the binding site. Practically, it is necessary to define two types of grids: the bounding box and enclosing box (**Figure 5**). The first (green in Figure 5) encompasses the area within the centroid of the ligand is free to move and rotate during the docking calculation phase. The second one (purple in Figure 5) contains the bounding box, representing the physical space in which the ligand atoms can be confined to during the calculations. It should be big enough to contain all residues involved in protein-ligand interaction.

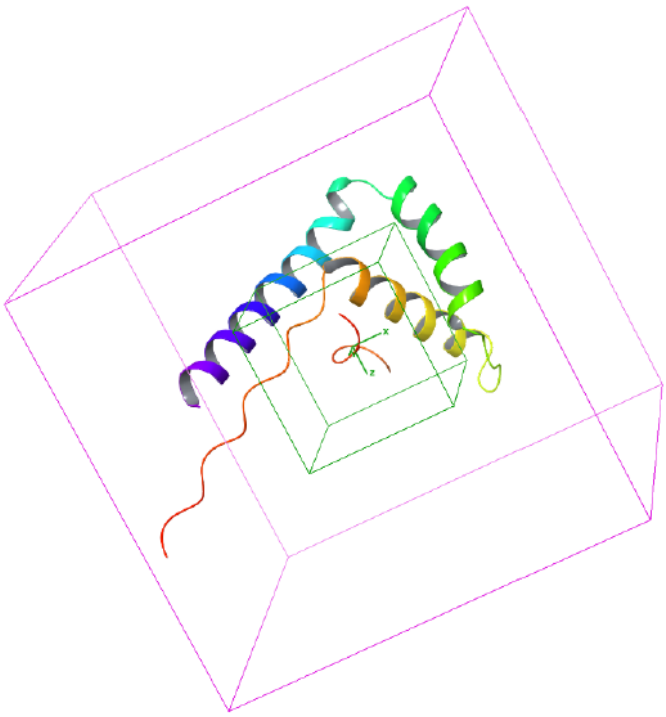


Figure 5. Example of a receptor grid generation step using GLIDE, in which it is observable the bounding box (green) and the enclosing box (purple).

The ligand docking phase involves the application of a series of hierarchical filters (**Figure 6A**). Firstly, GLIDE performs a full conformational search within the binding site, producing a multitude of ligand conformations. Then, the software performs a check of the generated conformations and eliminates those that are not suitable for binding with the receptor (e.g., those endowed with an excessively high energy). At this stage each ligand is considered to be composed of a central part (*core*) and a series of rotational groups (**Figure 6B**). Each of the possible conformation of the core and the various rotational groups are then rebuilt and subjected to a further docking step. For each conformation assumed by the ligand, determined through the previous phase, the distances between the center of mass of the ligand and its surface are calculated.

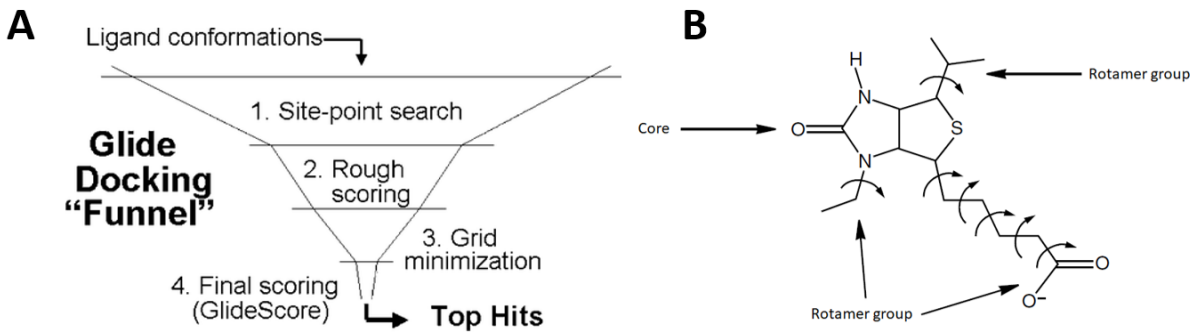


Figure 6. (A) Representation of the GLIDE docking funnel showing the application of the various hierarchical filters. (B) Example of a molecule divided into core and rotamer groups.

At this point the steric analysis of the binding site takes place and all those conformations that have a high number of steric clashes (i.e., too close physical contacts with the receptor) are discarded. Subsequently, the conformations that have been considered acceptable are moved around the diameter of the ligand and evaluated on the basis of a score taking into account the number of H-bonds, ligand-metal interactions, and the presence of steric clashes with the target. Then, the top scoring poses are re-scored by a "refinement" procedure, in which the whole ligand is allowed to rigidly move by ±1 Å in all the cartesian axes of space (x, y, or z). At this point, a very limited number of refined conformations (typically 100-400) will then undergo a minimization within the grid including the binding site.

Finally, the minimized poses are re-scored using the Schrödinger's GlideScore scoring function (*Gscore*) [45], and the best docking structure is selected among the ones showing the lowest *Gscore* value, expressed in kcal/mol.

Molecular Dynamics (MD)

MD is a computational technique that simulates the time dependent behavior of a biological system such as a protein or a nucleic acid, providing accurate and predictive information on the possible conformations that the system can assume. The result of the MD simulations allows to obtain a trajectory, composed by several snapshots (*frames*), which describes the variation of the vectors position, velocity and acceleration of the atoms on the system, over the simulation time. MD simulations typically allow the observation of events that occur in a time scale ranging from *femtoseconds* (fs, 10^{-15} s) to tens of *microseconds* (μs, 10^{-6} s) [51]. However, thanks to recent advances in computer simulations and in the GPU development, full-atoms MD simulations can reach millisecond (ms, 10^{-3} s) timescales, making MD an emerging and powerful method successfully complementing biological experiments [52]. In fact, different bio-phenomena occur on average in a time scale ranging from *femtoseconds* to *milliseconds* (**Table 1**), therefore, it is necessary that MD simulations' duration is long enough to cover the time scale corresponding to the phenomenon under investigation [53].

Table 1. Time scale of the main biophysical phenomena.

Motion patterns	Functional examples	Time scale
Local motions Atom fluctuations Side chain motions	Ligand flexibility on docking; Diffusion motions of molecules.	fs – ps (10^{-15} – 10^{-12} s)
Mid-scale motions Loop fluctuations N- and C-terminal motions Rigid body movement	Conformational adaptation of the active site; Binding specificity.	ns – μs (10^{-9} – 10^{-6} s)
Long-scale motions Domain fluctuations Subunit motions	Intra-domain motions; Allosteric transitions.	μs – ms (10^{-6} – 10^{-3} s)

MD simulations unveil principles on the base of a large variety of fundamental processes by mimicking the real-life perpetual motion of atoms triggered by interatomic forces. In MD, these forces are expressed through analytical functions and associated parameters, which are commonly referred to as the *force field*. The forces' effect on the movement of atoms is determined by Newton's second law [54].

$$\vec{F} = m\vec{a} \qquad \textbf{(eq.I)}$$

where (F) is the force acting on an atom of mass (m) with acceleration (a).

Acceleration is the second derivative of position, so the equation of motion can be written as:

$$\frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} = \frac{F_{x_i}}{m_i} \qquad \textbf{(eq.II)}$$

According to eq.II, an atom with mass (m_i) is moving along (x_i) direction under a force (F_{xi}) due to the interactions with other atoms. The force acting on an atom is the derivative of the potential energy with respect to the change in the atom's position.

$$F_i = m_i a_i = -\nabla_i V = \frac{-\mathrm{d}V}{\mathrm{d}r_i} = m_i \frac{\mathrm{d}^2 r_i}{\mathrm{d}t^2} \qquad \textbf{(eq.III)}$$

The potential energy (V) comes from the force field (FF) parameters. During the MD simulation, initial positions and velocities of all atoms are updated from time (t) to time ($t + \Delta t$) and the time is partitioned into small time steps (in the order of femtoseconds, typically 2). At each step, the forces acting on atoms are combined with the current positions and velocities to generate new positions and new velocities. Therefore, MD trajectories that describe positions and velocities of atoms as function of the simulation time are obtained by integrating the equations of motion.

Generally speaking, starting from a ligand/receptor complex, previously generated by docking, a MD algorithm like AMBER20 [55] is used to perform the MD simulations, and the typical protocol steps are:

1) *Selection of the force fields, solvation, and neutralization of the system:* the TLEAP module of AmberTools20 [55] was used to load the proper force fields parameters. In particular, the ff14SB [56] force field parameters were used to describe the protein, while the TIP3P [57] model and the parameters proposed by Joung *et al.* [58] were used to describe water and counter ions, respectively. Then, the system was solvated in a box of water with a minimum distance of 10 Å from the protein surface. This box is replicated in all directions using the Periodic Boundary Conditions (PBC). PBC allow to avoid problems with boundary effects because when an atom leaves the simulation box it re-appears on the opposite side with the same velocity. Finally, the net charge of the system was neutralized by adding a proper number of ions.

2) *Minimization and equilibration step:* The solvated system was relaxed by a two-step protocol to remove atomic clashes [59]. First, we performed an energy minimization for 10,000 steps, or until the energy gradient of 0.2 kcal/mol/Å was reached, restraining the atomic coordinates of backbone with harmonic potential (k = 20 kcal/mol/Å^2). This first phase was followed by an energy minimization for 100,000 steps or until an energy gradient of 0.0001 kcal/mol/Å was reached, without any restraint. After minimization, the temperature of the system was gradually increased to 300 K over 40 ps under constant volume condition (NVT) constraining the backbone coordinates in the first 20 with a harmonic potential (k = 20 kcal/mol/Å^2). Finally, the system was equilibrated at 300 K for 20 ps under constant pressure conditions (NPT, 1 atm).
3) *MD production*: in this step, long time-scale (usually 500 ns-long) MD simulation is carried out under constant pressure and temperature values, using the Berendsen barostat and thermostat, respectively [60]. Electrostatic interactions were treated with PME [61] with a cut-off of 9 Å. During these calculations, all bonds involving hydrogen atoms were constrained with the SHAKE [62] algorithm. All calculations were performed using the PMEMD of AMBER code in the GPU accelerated version [63] with a time step of 2 fs.
4) *Trajectory analysis:* in order to extract useful information from MD simulations, it is necessary to evaluate a series of parameters, among them the most important studied in this PhD thesis are the Root-Mean-Square Deviation (RMSD), the estimation of the binding free energy (ΔG) and the identification of the most representative structure through cluster analysis.

Binding Pose Metadynamics (BPMD)

In the BPMD tool, available in the Maestro suite (Schrödinger, LLC, New York, NY, 2021), the simulating system free-energy landscape is sampled by a history-dependent bias on a small set of collective variables (CVs). Then, monitoring the system free-energy values as function of the CVs variation, and by adopting the well-tempered algorithm, the simulating systems explore their free-energy landscapes escaping from the free-energy minima in which they could be trapped. Essentially, the ligand is automatically obliged to move in the binding site and the observed mobility under the biasing potential is considered indicative of the predicted

binding mode stability or instability. Applying this method, we were able to reliably discriminate between the correct ligand binding poses generated with the docking procedure. This method was applied in *case study 2* (see Chapter 4.2.2. for details) allowing us to discriminate the most favorable pose between the two best docking poses of each generated peptidomimetic, significantly reducing the overall computational time cost by reducing the amount of MD simulations to be performed. In particular, the best two docking poses of each ligand were subjected to twenty independent metadynamics simulations (each of 10 ns), using as CV the RMSD value of the ligand heavy atoms relative to their starting position. Before running the simulations, the systems were solvated by water molecules, followed by multiple minimization and restrained MD steps which guaranteed the system to slowly reach the final temperature of 300 K, as well as to remove the atomic clashes which could be present in the initial starting structures. Among the different scores provided by BPMD, that are related to the ligand stability during the course of the metadynamics simulations, we used the *PoseScore* (i.e., a value indicative of the average RMSD from the starting pose) in order to select which of the two best docking poses to be simulated by MD.

Root-Mean-Square Deviation (RMSD)

In bioinformatics, the RMSD of atomic positions is the measure of the average distance between the atoms of superimposed proteins. Notably, RMSD calculation can be also applied to other, non-protein molecules, such as small organic molecules [64].

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2} \qquad \textbf{(eq.IV)}$$

where δ_i is the distance between atom *i* and either a reference structure or the mean position of the N equivalent atoms.

In this PhD research work, it was calculated for the backbone (C, N, O, and Cα) or only Cα atoms. RMSD represents a formidable tool in the computational drug discovery process, in fact, it is possible to calculate, in an unbiased way, how the ligand remains stable in the binding site pocket during the MD simulations time. It can be done by plotting the RMSD of the ligand backbone atoms (peptide) *vs.* the simulation time. An example of RMSD plot in which the peptide ligand remains stable during the entire MD simulation is shown in **Figure 7**.

Estimation of the binding free energy (ΔG): MM-GBSA method

Free energy methods permit to estimate the difference of Gibbs free energy (ΔG) between two different molecular structures. In the case of drug design, it is usually calculated for the ligand, and the target, both in the free form, and for the ligand/protein complex form [65]. Then, the ligand binding free energy can be easily calculated by eq.V:

$$\Delta G_{bind} = G_{complex} - (G_{receptor} - G_{ligand}) \qquad \textbf{(eq.V)}$$

In this work, we have applied the Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) approach, which is able to calculate the binding free energy of any kind of ligands (proteins, peptides, ions, small molecules, and others) to the biological counterpart. The binding free energy ΔG_{bind} is related to the change in enthalpy (ΔH) and entropy (ΔS) by the eq.VI:

$$\Delta G_{bind} = \Delta H - T\Delta S \qquad \textbf{(eq.VI)}$$

In particular, in the MM-GBSA calculation, the free energy of each Gx state (where x is ligand, receptor, or the complex) is estimated by eq. VII:

$$Gx = E_{bond} + E_{el} + E_{vdW} + G_{pol} + G_{np} - TS \qquad \textbf{(eq.VII)}$$

where the first three terms are represented by standard molecular mechanics energy terms from bonded (bond, angle and dihedral), electrostatic and van der Waals (vdW) interactions. G_{pol} term is equal to the polar contributions to the solvation free energies and is obtained by adopting the Generalized Born (GB) model (giving the MM-GBSA approach), while the non-polar term (G_{np}) is estimated by a linear relation to the solvent accessible surface area (SASA).

To improve the accuracy of the binding free energy calculations, snapshots (*frames*) in which the ligand display the highest stability during the MD simulation time, according to RMSD analysis, were selected (**Figure 7**).

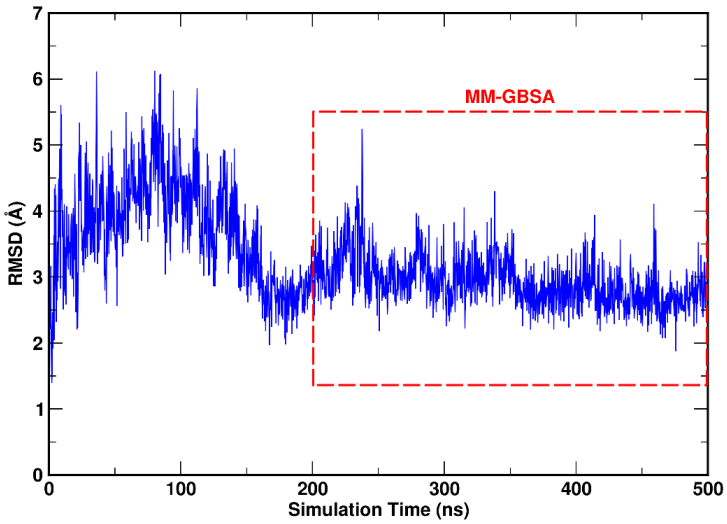


Figure 7. Example of RMSD plot in which we computationally performed the MM-GBSA analysis (highlighted in red) when the peptide ligand is more stable during the simulation time (range 200-500 ns).

In the majority of the case studies reported in this PhD thesis, we used the AMBER20 package and the Prime (Maestro Schrödinger suite) tool to compute the ligand binding free energy (ΔG) values, through the MM-GBSA approach.

Root-Mean-Square Fluctuation (RMSF)

In bioinformatics, the RMSF of a structure is the time average of the RMSD. It is calculated according to the below equation, where $\mathbf{x}_i$ is the coordinates of particle *i*, and $\langle \mathbf{x}_i \rangle$ is the ensemble average position of *i*.

$$\rho_i = \sqrt{\left\langle (\mathbf{x}_i - \langle \mathbf{x}_i \rangle)^2 \right\rangle} \quad \textbf{(eq.VIII)}$$

Where the RMSD quantifies how much a structure diverges from a reference over the time, the RSMF can reveal which areas of the simulating system display the highest conformational mobility. An area of the structure displaying high RMSF values frequently diverges from the average, indicating high instability. RMSF analysis is typically performed considering the Cα atoms, in fact, these are more characteristic of conformational changes compared to sidechains, which are more flexible. In this PhD work, RMSF analysis was done only in the case study involving the GABARAP protein (*case study 3*) in order to demonstrate how the cyclic peptides, that we have computationally designed by establishing an intramolecular disulfide bond, are endowed with lower conformational flexibility compared to the initial linear peptides (see Chapter 4.3).

Cluster analysis

Cluster analysis represents the task of grouping a set of similar objects in the same group (named *cluster*), while those which are significantly different are grouped into other *clusters*. In bioinformatics, clustering methods have been widely used to group together similar conformational states of biomolecules in solution, over MD simulations. Two main clustering protocols were adopted in the case studies considered in this PhD thesis. In the first, the MD trajectory frames were analyzed by clustering the conformations adopted by the peptide backbone atoms in complex with the protein target, using the *cpptraj* module [66] of AMBER20 [55]. This clustering method was applied in the design of ligands targeting PCSK9 (*case study 2*) and GABARAP (*case study 3*). By this algorithm, the MD frames were divided into clusters by the complete average linkage algorithm, however, this method automatically selects the RMSD cut-off threshold value used to divide the frames into the different *clusters*. For this reason, in the other case studies, in which we wanted to manually control the clustering procedure, we applied the *gromos* algorithm developed by Daura *et al*. [67], available in the GROMACS software package (version 5.0.7) [68]. In fact, by *gromos*, after several attempts and accurate visual inspection of the outputs, an appropriate RMSD cut-off value was used to obtain the optimal threshold to discriminate between the different peptide conformations sampled by MD.

Alanine scanning protocol

Alanine scanning is a site-directed mutagenesis technique used to determine the impact of single amino acids contribution to the interaction with a given protein or peptide to another [69]. Alanine residue is used because of its chemical properties (non-bulky, chemically inert, methyl functional group that nevertheless mimics the secondary structure preferences that many of the other amino acids possess).

Practically, we performed *in silico* alanine scanning calculation by mutating each residue of the native peptide sequence, one by one, by an alanine (**Figure 8**) in order to identify the “hot” and “non-hot” spots of the peptide/target interaction using the AMBER software package. Then, long-time scale MD simulations were performed on each mutated peptide in complex with the protein target. Finally, MM-GBSA calculations was carried out to estimate the binding affinity (ΔG) of the mutated peptides. This procedure highlights which residues can be

considered as *hotspots* (i.e., the residues fundamental for the interaction) and which are the ones poorly contributing to the binding energy (*non-hotspots*).

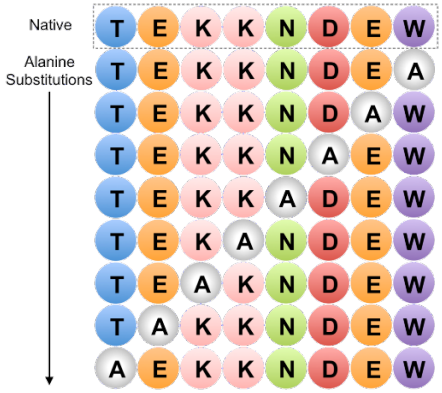


Figure 8. Example of mutated peptides generated by alanine substitution of each residue composing the native peptide.

Accordingly with these results, the *non-hotspots* residues can be systematically and suitably replaced by new amino acids (natural or non-natural) to improve the complementarity between the new peptides and the protein target by the application of the "affinity maturation protocol".

Affinity Maturation protocol

By these simulations, the *non-hot spots* were mutated into all possible combinations of natural amino acids by using the "Residue Scanning Calculation" tool, implemented in the BioLuminate module of Maestro (Schrödinger, LLC, New York, NY, 2021) [70]. This tool is able to sample the side-chain rotamers of the mutated residue, also optimizing the geometry of the resulting peptides. The mutant peptides are then ranked by ΔAffinity and/or ΔStability (values calculated by Prime MM-GBSA in implicit solvent). In particular, ΔAffinity is defined by the change in binding affinity of the mutated peptide, treated as the ligand, to the rest of the system, treated as the receptor, while ΔStability is defined as the difference in free energy between the folded state and the unfolded state of the peptide, due to the mutation. In both cases, negative values mean that the mutant peptides bind better than the native protein, or they are more stable, respectively.

The affinity maturation tool, together with the *in silico* alanine scanning, was applied in all the case studies and allowed us to successfully fine-tune the peptide/target docking model thus identifying mutated peptides with improved theoretical binding affinity compared to the native peptide.

Machine learning

Artificial Intelligence (AI) is an area of computer science capable of simulating the structures and operating principles of the human brain. AI methods had shown tremendous potential in different areas such as computer vision, speech recognition, natural language processing and others. Recently, they had also been successfully applied in drug discovery, they are able to accelerate the drug development pipeline, both in preclinical and clinical stages [71,72].

Machine learning is a subfield of AI, which aims to develop and utilize algorithms that learn from raw data. Machine learning tasks are referred to as supervised learning tasks where a set of input-output paired data is provided for training. When the task is to predict from a set of potential outcomes, it is a classification task, and when the task is to predict a numeric value, it is a regression task. This prediction method had been successfully applied and reported in several scientific works published in literature. In fact, examples of classification tasks are virtual screening predictions [73] and disease diagnosis [74], while examples of regression tasks are drug efficacy [75] and absorption, distribution, metabolism, excretion, and toxicity (ADMET) predictions [76].

Deep learning is a subset of machine learning that uses artificial neural networks algorithms which mimic how the brain work, learning high-level abstractions for data [77]. Briefly, deep learning architectures are based on layers which train on a distinct set of features based on the previous layer's output. The further you advance into the neural net, the more complex the features your nodes can recognize, since they aggregate and recombine features from the previous layer. As a consequence, deep learning networks are potentially capable of handling very large data sets, and they have found an important place in drug discovery processes obtaining stunning results [78,79,80].

In this research activity, a machine learning approach was successfully applied to identify tetrapeptides potentially capable of inhibiting WWP1 (*case study 4*), using the Peptide QSAR and DeepChem/AutoQSAR (DC/AQ) tools, both implemented in BioLuminate (Schrödinger, LLC, New York, NY, 2021). Typically, QSAR is employed to measure the association between the structures of small drug-like compounds with experimental properties, like activity or cytotoxicity. If a predictive relationship exists, it is possible to apply the QSAR model to identify new molecules, within the same series, with improved biological activity.

A very similar approach can be applied to peptides. To this aim, the Peptide QSAR tool is able to predict a property of a peptide (e.g., binding affinity or solubility) relying only on amino acid sequences, as a function of given experimental data used to generate a prediction model. Two statistical method could be used to generate the QSAR model: Partial Least Squares (PLS) and Kernel-based PLS, with the number of factors user-defined depending on the size of the data.

DC/AQ tool employs cutting edge deep learning methods, enabling non-expert users to easily create high-performance QSAR models from large datasets containing even hundreds of thousands data points. It is based on an automated approach to neural-networks, by implementing the Graph-Convolutional Neural Networks (GCNN), similar to those involved in image and video processing. In fact, DC/AQ treats compounds as a graph, where nodes are atoms and edges are bonds, and the chemical features (such as valence, atom type, charge, etc.) are attached to each node in the graph. The convolution, in this case, is to apply filters to contiguous atoms, instead of neighboring pixels (in the case of image and video software).

Both methods were used to predict the docking score (*Gscore*) of a previously generated 28,000 tetrapeptide sequences library. In particular, we performed the docking calculation of only 500 random sequences, generating a prediction model which allowed us to dramatically reduce the computational time cost (for details, see Chapter 4.4).

2.2. Affinity assessment using biophysical methods

Biophysical methods, which can include a wide range of techniques aimed at acquiring information on the structure, properties, dynamics or function of biomolecules, have been extensively employed in the drug discovery process since their first introduction, in the early 1990s [81]. In the drug discovery process, the biophysical methods are applied for the quantification of the drug-target interaction, extremely useful to accomplish structure-activity relationship studies. The equilibrium dissociation constant (K_d) is the principal parameter used to evaluate the binding potency of a putative drug [82]. Therefore, the experimental measurement of the ligands K_d values is fundamental for the study of the drug-target complexes [83,84] and the lead optimization process. Currently, the most widely applied biophysical methods are the Isothermal Titration Calorimetry (ITC) [85], the Surface Plasmon Resonance (SPR) [86], the Fluorescence Energy Resonance Transfer (FRET) [87], and the relatively new Microscale Thermophoresis (MST) [88].

Isothermal Titration Calorimetry (ITC). Over the past decade ITC has been identified as one of the gold standard methods to directly measure the ligand binding affinity and thermodynamics equilibrium between the interacting species (see the Freyer and Lewis review [85]). ITC is a technique based on the reaction heat to quantify the interactions of various biomolecules (for the complete explained protocol see Holdgate review [89]). The experiments are conducted at (almost) constant temperature. A ligand solution, contained within the calorimeter cell, is titrated by the solution of the target by means of an injection syringe. Therefore, the ITC detector can measure any heat changes resulting from any biochemical reaction process. A schematic representation of the ITC machine and an example of ITC plot data is shown in **Figure 9**. Depending on the binding affinity and the amounts of available reagents, it is also possible to arrange the experimental conditions in order to obtain the affinity (K_d), the enthalpy (ΔH), entropy (ΔS) and the stoichiometry (n) of the binding interaction, performing only a single experiment [90].

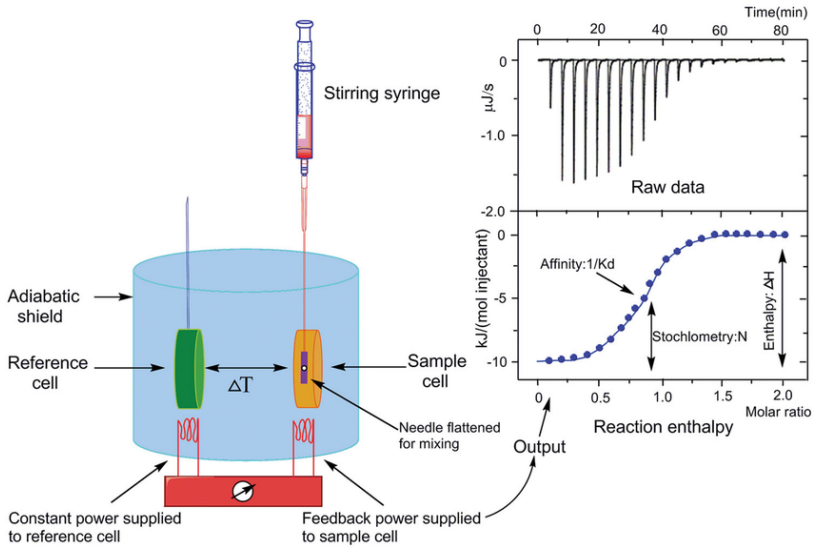


Figure 9. Basic principle of isothermal titration calorimetry. Schematic representation of the isothermal titration calorimeter (left) and a characteristic titration experiment (upper right) with its evaluation (lower right) [85].

ITC is widely used in molecular biology research, drug design, and mechanistic studies. However, the most useful application of ITC relies on the characterization of compounds produced by medicinal chemists during the lead optimization phase of the drug discovery process. Understanding the thermodynamics of a molecular interaction is fundamental in drug discovery, as it can confirm the expected binding target, it can permit to understand the structure-activity relationships, and provide the guidance for the selection of the candidate compounds [91].

Surface Plasmon Resonance (SPR). During the last two decades SPR biosensor technology has seen a rapid evolution starting with the launch of the first commercial BIAcore instrument by Pharmacia Biosensor in 1990 [92]. SPR is based on the principle that the incident light can resonate with the plasma on the metal surface during the total reflection. This enables the determination of the kinetic parameters and the equilibrium constants for a given system, i.e. on-rate, off-rate and apparent dissociation constants. SPR can measure changes in the mass of a dissolved material in the aqueous layer (biosensor surface) close to the metal film, which allows the interaction of proteins with other ligands to be monitored in real-time [93].

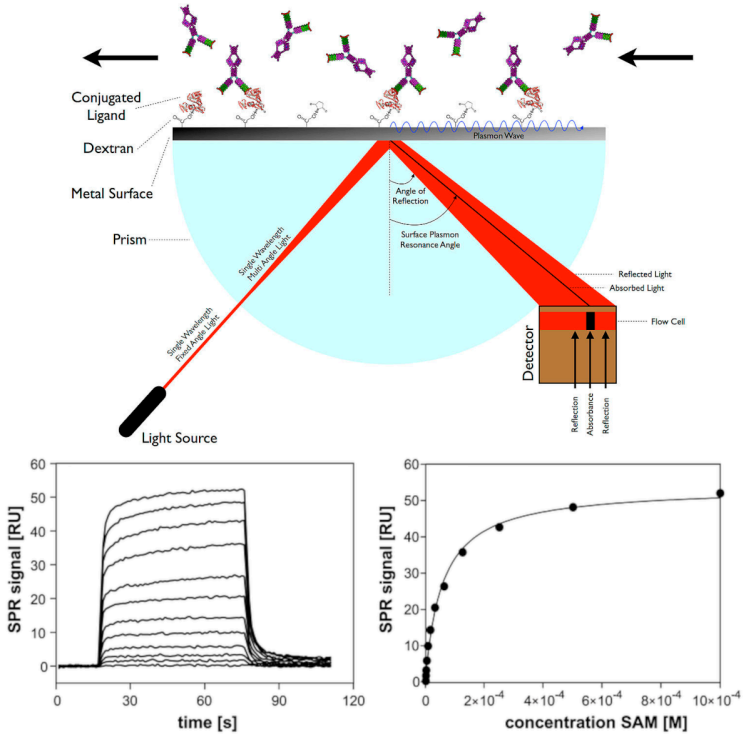


Figure 10. (Up) This figure shows the configuration of an SPR chip. The metal chip (silver or gold) is prepared with a dextran surface which can bind the amino end of protein to conjugate them to the metal surface. At the bottom a single wavelength laser beam enters a prism which results in many light angles striking the metal surface, all of them are reflected except for the angle in which the metal will absorb and turn its energy into a plasmon wave onto its outer surface, at this angle no light is reflected and thus appears with very little intensity on the detector. Since the plasmon wave propagates on the outer side of the metal, any interaction with the conjugated protein will change the SPR angle. (Bottom) Example of an SPR binding affinity plot using a Biacore S51 machine [97].

The most common approach is represented by the random-oriented immobilization through accessible primary amines on the protein surface by activating the carboxymethylated dextran-matrix (a complex and branched glucose polymer) with a mixture of EDC and NHS, to create NHS-esters that can react with amino-containing molecules.

SPR technique represents a sensitive and specific technique for the analysis of biomolecular interactions, and it is used to detect whether biological molecules interact with each other,

and to further explore the specificity of the interaction, kinetic parameters and ligand affinity (K_d) [94]. In the last decades, because of the high-throughput screening characteristic, SPR has been widely used in the identification of drug targets and the optimization of the lead compounds [95,96]. A schematic representation of the SPR method and an example of binding affinity plot data generated is shown in **Figure 10**.

Fluorescence Energy Resonance Transfer (FRET). Fluorescence spectroscopy is one of the most popular techniques both in biology and medicine research fields. The principle of FRET is represented by a nonradiative energy transfer between two fluorescent molecules that are located close to each other (less than 10 nm) [98]. Briefly, FRET occurs between a donor fluorophore and an acceptor fluorophore, which can be carried either by the ligand or by the target. In particular, two parameters affect FRET efficacy [99] (**Figure 11A**). FRET can be used to study receptor-ligand interactions, affinity constants, and receptor dimerization. Therefore, it is widely used in drug-target affinity studies under equilibrium condition, with no need to separate the free and combinative ligands [100]. FRET is a high sensitivity technique used to selectively study specific intermolecular interactions under physiological conditions (e.g., living cell states) [101].

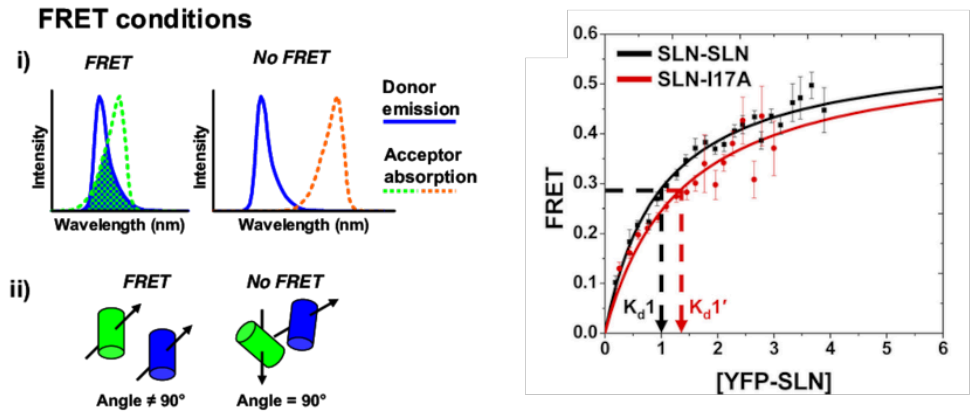


Figure 11. Principles of FRET. (Left) Two conditions are necessary for FRET: i) the overlapping of the donor emission and the acceptor excitation, and ii) the alignment of the fluorophores' dipoles [102]. (Right) Example of a binding affinity plot determined by FRET microscopy [103].

Microscale Thermophoresis *(MST)*. it is a relatively new biophysical analysis which is based on the fluorescence change induced by temperature on a target as a function of different concentrations of a non-florescent ligand. In detail, it is possible to measure the movement of molecules within a temperature gradient which induces changes in molecular properties of the studied molecules in terms of charge, size, hydration shell or conformation [104].

The MST technique is based on the directional movement of molecules along a temperature gradient. Practically, the MST instrument records fluorescence of the sample with a focal infrared (IR) laser during and after the laser is turned on. Affinity is quantified by monitoring the change in normalized fluorescence called Fnorm as a function of the concentration of the binding partner. The K_d model used by MST describes a 1:1 stoichiometry interaction according to the law of mass action and allows to derive a formula for the fraction bound in case of a binding event. The fraction bound is defined by the K_d and the concentration of the target molecule depends on the ligand concentration [105]. Since MST is a solution-based method, it avoids surface artifacts and immobilization protocols. An illustration of the MST method is shown in **Figure 12** [105].

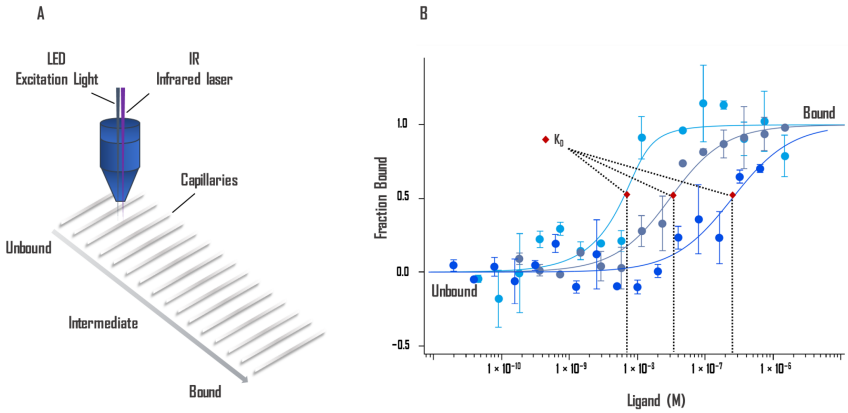


Figure 12. (A) Schematic representation of the optical system. Fluorescent molecules in the 16 capillaries are excited and the fluorescence detected. An IR laser heats up the middle part of each capillary, and thermophoresis of the fluorescent molecules across the temperature gradient is detected. (B) Example of MST binding affinity curves.

In this PhD research activity, the binding affinity (measuring the K_d) of a series of computationally designed peptides was determined by using MST and SPR biophysical methods. In fact, the very large amount of protein sample needed to perform ITC experiments makes this technique not feasible when the amount of available protein is limited. On the other hand, the FRET method is characterized by several disadvantages, among them the size of probes can introduce problems during the binding assay.

Conversely, MST and SPR experiments are relatively easy and fast, they are endowed with a good accuracy, even if a very small quantity of both ligand and protein is available. However, our preferred biophysical method is the MST, since we have conducted experiments on peptides, performing assays directly in solution also avoiding surface or immobilization artifacts.

3. AIMS OF THE PhD RESEARCH ACTIVITY

In this research project, I have applied advanced computational approaches, in combination with biochemical and biophysical studies, to discover, design, and optimize the structure of peptides and peptidomimetics capable of interacting with biological targets critically involved in several pathologies. Accordingly, the objective of this project is to discover novel promising peptide-based ligands which could represent potential drug candidates, paving the way to the identification of new innovative therapies against multiple disabling pathologies.

In particular, the peptides or peptidomimetics showing the most promising theoretical profile, as resulted by computational studies, have been synthesized (or acquired) and experimentally tested through *in vitro* biophysical experiments. Then, the ligands endowed with the most promising profile have been investigated by *in cell* experiments.

In this research activity, several computational approaches have been applied to different case studies with the objective at identifying potential inhibitors of HMGB1 protein, an emerging target for the development of anti-inflammatory drugs (*case study 1*). My attention was also focused on the design of innovative potential drugs endowed with a dual inhibitory activity against PCSK9 and HMG-CoAR protein targets, both crucial for the treatment of hypercholesterolemia (*case study 2*). In this case, I was also able to identify new promising peptidomimetics. In the *case study 3*, I used some computational weaponry in order to discover ligands capable of interfering with the autophagy machinery in which GABARAP protein is involved. Similarly, in this case peptides with potential anticancer activity were designed and biologically assayed. Finally, many efforts have been dedicated to the identification of ligands able to inhibit WWP1, a new promising target of anticancer drugs (*case study 4*) for which no drug candidates have been till now reported in literature. In this case study, machine learning algorithms were applied to achieve the goals.

In conclusion, if the pharmacological investigations will confirm the predicted activity profile, the structure of the most promising ligands will be optimized (by additional steps of computational studies) to improve their pharmacokinetic properties, using the most suitable medicinal chemistry strategies developed to this aim (see introduction chapter).

4. CASE STUDIES

4.1. HMGB1

High-mobility Group Box 1 (HMGB1) is an abundant chromatin-associated protein present in all mammalian cells. It is formed by 215 amino acids, divided into two structurally similar domains, called BoxA (Gly2-Ile79) and BoxB (Phe89-Arg163), connected by a nine amino acid loop, and a highly disordered negatively charged C-terminal tail. BoxA contains a pair of cysteines (Cys23 and Cys45) that can form a disulfide bond under oxidative conditions. In contrast, only one unpaired cysteine is present in BoxB (Cys106, **Figure 13**) [106].

Depending on its cellular localization, HMGB1 performs different functions. As a nuclear protein, it is involved in DNA repair, transcription, telomere maintenance, and genome stability [107,108,109], while during cellular death or inflammation, HMGB1 is released in the extracellular space where it functions as an alarmin [110,111].

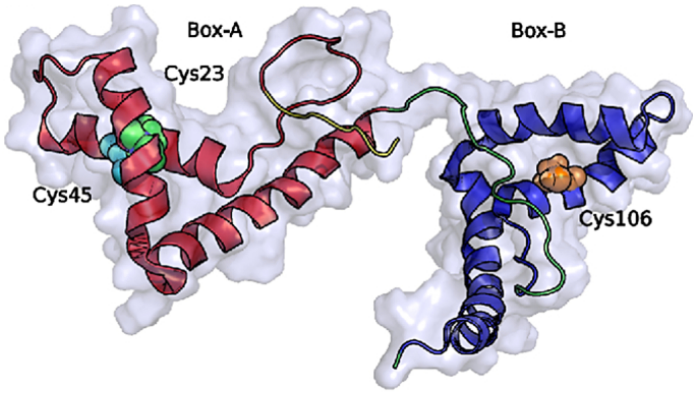


Figure 13. Structure of HMGB1 (PDB ID code 2YRQ) solved by NMR. Protein domains are presented in different colors: BoxA (red), BoxB (blue). The three cysteines (Cys23, Cys45 of BoxA and Cys106 of BoxB) are displayed as van der Waals balls.

According to several studies, the HMGB1 functions depend on its redox states [106,112]. In fact, the nuclear and cytosolic environments are characterized by a negative redox potential that maintains HMGB1 in the reduced form (fr-HMGB1). During the inflammatory process, the extracellular space, enriched in reactive oxygen species, leads to the formation of a disulfide bond between cysteines at positions 23 and 45 of BoxA (ds-HMGB1) [113]. The ds-HMGB1 is able to activate Toll-like Receptor 2 (TLR2) and 4 (TLR4) inducing the release of proinflammatory chemokines and cytokines activating innate and adaptive immune responses, while fr-HMGB1 is able to form an heterocomplex with the CXC ligand 12 (CXCL12) [114].

CXCL12 is a chemokine expressed in many tissues both under homeostatic and inflammatory conditions and that can stimulate the recruitment of inflammatory cells by activating the CXC chemokine receptor type 4 (CXCR4) [115]. While a structure of the heterocomplex is currently unavailable, NMR chemical shift mapping clearly showed an interaction between CXCL12 and the two domains of HMGB1 (BoxA and BoxB), separately [116]. Furthermore, the same experiments showed that the binding of CXCL12 to HMGB1 induces conformational changes in the *N*-terminal domain of CXCL12 which is required to trigger the activation of the CXCR4 receptor. Based on these data, it was hypothesized that the heterocomplex is formed by two CXCL12 molecules bound to HMGB1 (one to BoxA and one to BoxB), and that it would bind CXCR4 dimers [116].

A particular feature of the CXCL12/HMGB1 heterocomplex is that only fr-HMGB1 can form a complex with CXCL12, promoting CXCR4-induced response [114] as confirmed also by Microscale Thermophoresis (MST) experiments recently published [117].

4.1.1. HBP08 discovery

Despite the importance of this target, to date only few inhibitors (Salicylic Acid [118], Diflunisal [119] and 5,5'-Methylenedi-2,3-Cresotic Acid [120]) have been reported in literature, however showing a millimolar K_d on HMGB1. Only glycyrrhizin exerts its inhibitory activity in the high micromolar range binding to HMGB1-BoxA [121]. Starting from these data, we designed a novel nonapeptide (namely HBP08) directed on HMGB1-BoxA, aiming at inhibiting the formation of the CXCL12/HMGB1 heterocomplex and to abolish the synergistic effect on cell migration in CXCR4 transfected cells and in human monocytes [106].

Design of HBP08. Utilizing the nuclear magnetic resonance (NMR) chemical shift perturbation (CSP) studies [121], a model of the glycyrrhizin in complex with the BoxA of HMGB1, was generated. In this model, glycyrrhizin interacted with Gln20 and Arg23 and occupies the region at the junction between the two arms of the "L-shape" characterizing HMGB1. Then, a library of 40,000 nonapeptides with a random sequence was generated and docked in the binding site identified by the presence of glycyrrhizin in the HMGB1-BoxA. In particular, all the residues of BoxA having a distance smaller than 7.5 Å from glycyrrhizin carbon atoms were selected, obtaining a total of 17 amino acids. Then, the whole peptide library was docked to BoxA, guided by the set of 17 amino acids selected before. In the next step, only the structure

of each peptide (among the 25 generated) endowed with the smallest distance restraint violations has been selected and minimized, then ranked by binding energy. The best 100 ranking peptides were then selected, and, aiming at reducing the number of potential false positives, they were docked again in the BoxA using the "peptide docking" protocol of GLIDE software, leaving the algorithm free to search for the best binding site on the protein surface. Finally, the peptides resulting at the end of these calculations were visually inspected and only the best pose of the highest scored peptides (by *Gscore*), also mimicking the glycyrrhizin binding mode, was retained for further analysis. Several studies have shown that the free-energy methods, such as MM-GBSA [122,123], can be valuable tools in the selection of active peptides in the virtual screening investigations. Therefore, a 500 ns long MD simulation was performed on each of the 57 peptides resulted from the previous steps of docking calculations. Finally, based on the MM-GBSA predicted values, 13 different peptides were selected and then tested by *in vitro* assays (**Table 2** and **Annex 1**, Chapter 7).

Table 2. Peptides ranked accordingly with the calculated binding free energy (ΔG) values.

Peptide	Sequence	ΔG ± SE (kcal/mol)
HBP01	HEMYWEDEW	-52.8 ± 0.3
HBP02	IDLRFFMRQ	-52.0 ± 0.3
HBP03	FAFELIQTD	-51.7 ± 0.4
HBP04	CIPMMMHAW	-50.0 ± 0.3
HBP05	WISNWILMW	-45.8 ± 0.3
HBP06	TWNIHFADH	-45.6 ± 0.4
HBP07	HWTLANWCR	-45.2 ± 0.4
HBP08	GYHYERWIH	-45.1 ± 0.5
HBP09	QFMKNCEEM	-44.8 ± 0.4
HBP10	SINWHMYVN	-44.8 ± 0.3
HBP11	MYRENQPTR	-42.9 ± 0.4
HBP12	YHICWYGDY	-42.5 ± 0.5
HBP13	WLWYEWGWQ	-41.9 ± 0.3

Biological evaluation. The 13 peptides showing the lowest ΔG values (namely, HBP01-13) were synthesized and tested by *in vitro* chemotaxis assay on a murine cell line expressing the human CXCR4, to evaluate their efficacy as inhibitors of the CXCL12/HMGB1-induced migration. Our experiments showed that 4 out of 13 peptides efficaciously inhibited the enhanced migration induced by the formation of the CXCL12/HMGB1 heterocomplex (i.e., HBP05, HBP07, HBP08 and HBP12) (**Figure 14A**). Further experiments, performed using CXCL12 alone, showed that HBP07 and HBP08 did not affect the CXCL12-induced cell migration, for this reason they were

furtherly tested on primary human monocytes. Conversely, HBP05 and HBP12 inhibited also the migration induced by the chemokine alone, so they were discarded. On primary human monocytes, only HBP08 significantly blocked the activity of the heterocomplex, without altering the migration induced by CXCL12 alone, do not exhibiting any toxicity on both cell types. A dose-response curve of the migration induced by the heterocomplex in the presence of descending concentrations of the HBP08 peptide revealed that the 50% of inhibition (IC_{50}) was observed at the HBP08 concentration of 50 μM (**Figure 14B**).

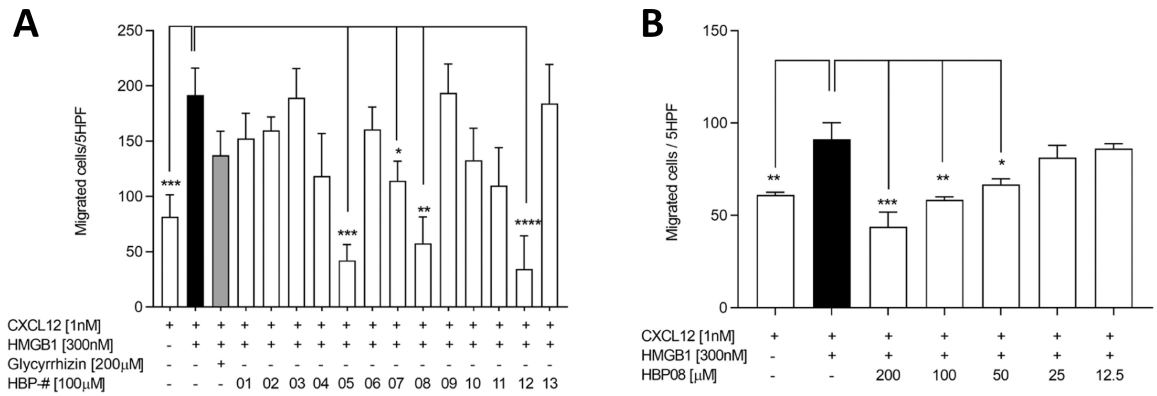


Figure 14. In vitro activity of the identified peptides. (A) Inhibition of cell migration in response to the CXCL12/HMGB1 heterocomplex was assessed on 300-19 Pre-B cells transfected with human CXCR4 using the identified peptides or glycyrrhizin. The numbers in the last horizontal row correspond to the different peptides. (B) Inhibition of cell migration in response to the CXCL12/HMGB1 heterocomplex was assessed on 300-19 Pre-B cells transfected with CXCR4 using scaling concentrations of HBP08.

In the extracellular space, the oxidized form of HMGB1, through the binding to TLR4, activates the NF-kB pathway inducing the transcription of several pro-inflammatory cytokines [114,124]. Thus, in order to establish if HBP08 was effective also on this pathway, the cytokine release assay on monocytes treated with HMGB1 alone or in the presence of HBP08, was performed. In these experiments, a significant release of IL-6 and TNF, which could be blocked by the treatment with a neutralizing antibody against TLR4 was observed. This suggested that HBP08 did not induce IL-6 or TNF release and did not block the HMGB1-mediated release of these cytokines. This clearly indicated that HBP08 inhibits selectively the CXCL12/HMGB1 heterocomplex activity, leaving HMGB1 able to trigger the TLR4 inflammatory pathway.

Biophysical evaluations. MST experiments were performed to measure the affinity of HBP08 to the recombinant HMGB1, attaining a K_d of 0.8 ± 0.4 μM (**Figure 15**). The affinity for HMGB1 of the identified peptide was, therefore, several orders of magnitude higher than the ones of

other HMGB1 inhibitors reported in the literature so far. In fact, it is known that glycyrrhizin has a K_d of ~ 150 µM [121], diflunisal ~ 1.6 mM [119], and 5,5-methylenedi-2,3-cresotic acid ~ 0.9 mM [120]. Overall, these results indicate that HBP08 is the most potent inhibitor of the CXCL12/HMGB1 heterocomplex formation.

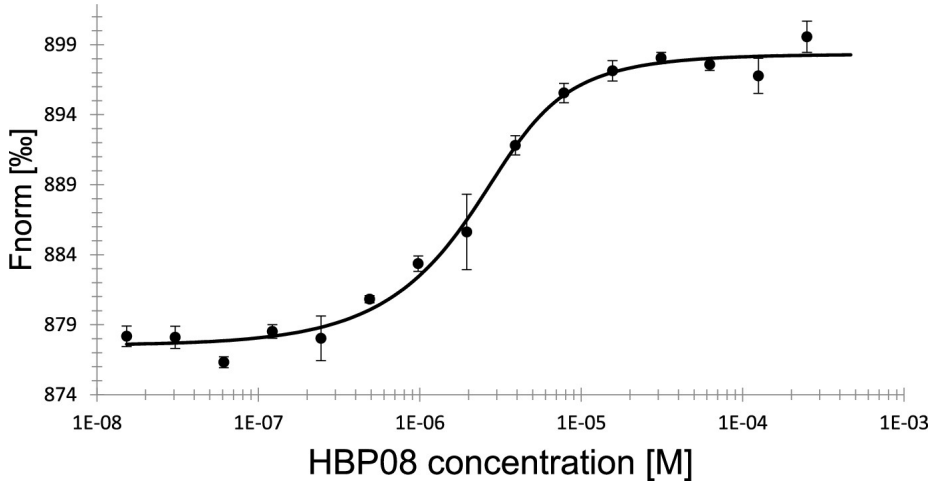


Figure 15. MST analysis of the interaction between HBP08 and HMGB1 (K_d = 0.8 ± 0.4 µM).

Additional MST experiments were conducted to measure the K_d values on two constructs containing only BoxA and BoxB of HMGB1. These experiments showed that HBP08 binds on HMGB1-BoxA with the same affinity as for the full-length protein (K_d = 0.8 ± 0.3 µM), while the affinity for HMGB1-BoxB was slightly lower (K_d = 17 ± 3.8 µM). Finally, NMR CSP experiments were accomplished to acquire atomistic details on the interaction between HBP08 and both HMGB1-BoxA and -BoxB.

Furthermore, to identify the HBP08 residues crucial for the binding with HMGB1, alanine scanning experiments were accomplished, and the affinity (K_d) of each mutant peptide to the full-length HMGB1 protein was measured through MST experiments (**Table 3**). The results of these experiments indicated that HBP08-Ala3, HBP08-Ala7, and HBP08-Ala9 are fundamental for the binding. The affinity of two peptides formed by the first (pentapept-1) or the last (pentapept-2) five residues of HBP08 were tested and, in agreement with the Ala-scan results, no binding was observed for pentapept-1 in the range of concentrations applied to the analysis of the other peptides. Otherwise, a K_d of 160 ± 80 µM was observed for pentapept-2, confirming the importance of the *C*-term in the binding.

Table 3. K_d observed the complexes between HMGB1 and the HBP08 mutated peptides, measured by MST.

Peptide	Sequence	K_d (μM)
HBP08	GYHYERWIH	0.8 ± 0.4
HBP08-Ala1	AYHYERWIH	8.6 ± 3.5
HBP08-Ala2	GAHYERWIH	5.8 ± 1.1
HBP08-Ala3	GYAYERWIH	26.2 ± 4.8
HBP08-Ala4	GYHAERWIH	9.9 ± 1.3
HBP08-Ala5	GYHYARWIH	0.8 ± 0.2
HBP08-Ala6	GYHYEAWIH	N.D. [1]
HBP08-Ala7	GYHYERAIH	22.0 ± 4.5
HBP08-Ala8	GYHYERWAH	1.9 ± 0.6
HBP08-Ala9	GYHYERWIA	>80
pentapept-1	GYHYE	no binding [2]
pentapept-2	ERWIH	160 ± 80
HBP08-RI	d-HIWREYHYG	14.0 ± 4.5

[1] Not determined due to poor solubility in PBS.

[2] No binding was detected in the explored concentration range.

As it is known, *L*-peptides are susceptible to the action of proteolytic enzymes such as peptidases, hindering their application *in vivo*. Conversely, *D*-peptides are less prone to the action of peptidases and to the acidic hydrolysis that occurs in the stomach, and these increase their oral bioavailability and the half-life in the blood circulation. Moreover, *D*-peptides have a low immunogenicity [125]. All these features make *D*-peptides suitable for drug development [126]. Consequently, we investigated the binding of the retro-inverso analogue of HBP08 (HBP08-RI), made by *D*-amino acids in the reversed order. The results of the binding experiments indicated that HBP08-RI has a lower but still good affinity for HMGB1 (K_d = 14.0 ± 4.5 μM, **Table 3**), therefore representing a good candidate for future drug development studies.

NMR CSP experiments. In analogy with previous investigations on the binding of proteins [116] or small molecules to HMGB1 [119,120], we performed NMR CSP experiments (**Figure 16BD**) to experimental characterize the interaction between HBP08 and both boxes of HMGB1 (BoxA and BoxB). This analysis enabled us to identify the protein residues involved in the peptide binding (**Table 4**). These data were then used to guide the docking calculations, by using HADDOCK v2.4 program in the webserver implementation [127], aiming at generating a reliable model of the HBP08/HMGB1-BoxA and HBP08/HMGB1-BoxB complexes (**Figure 16AC**). Finally, the poses acquiring the best HADDOCK score were selected and visually inspected, providing interesting information on the specific interactions displayed by the nonapeptide.

Table 4. Residues showing significant chemical-shift difference upon the HBP08 binding.

HMGB1 domain	Residues
BoxA	Y15, F17, V19, Q20, E25, K27, K28, K29, H30, S34, V35, E46
BoxB	D90, K95, A100, K113, G114, E115, G118, L119, D123, A125, G129, E130, M131, W132, N133

In fact, regarding the interaction with BoxA, the HBP08-His3 interacts with Asp66; HBP08-Trp7 is in contact with Arg23, Ser34, and Val35; and HBP08-His9, that MST experiments indicated as the more important residue for the formation of the complex, occupies a small cavity delimited by Tyr15, Phe17, and Gln20 (**Figure 16A**).

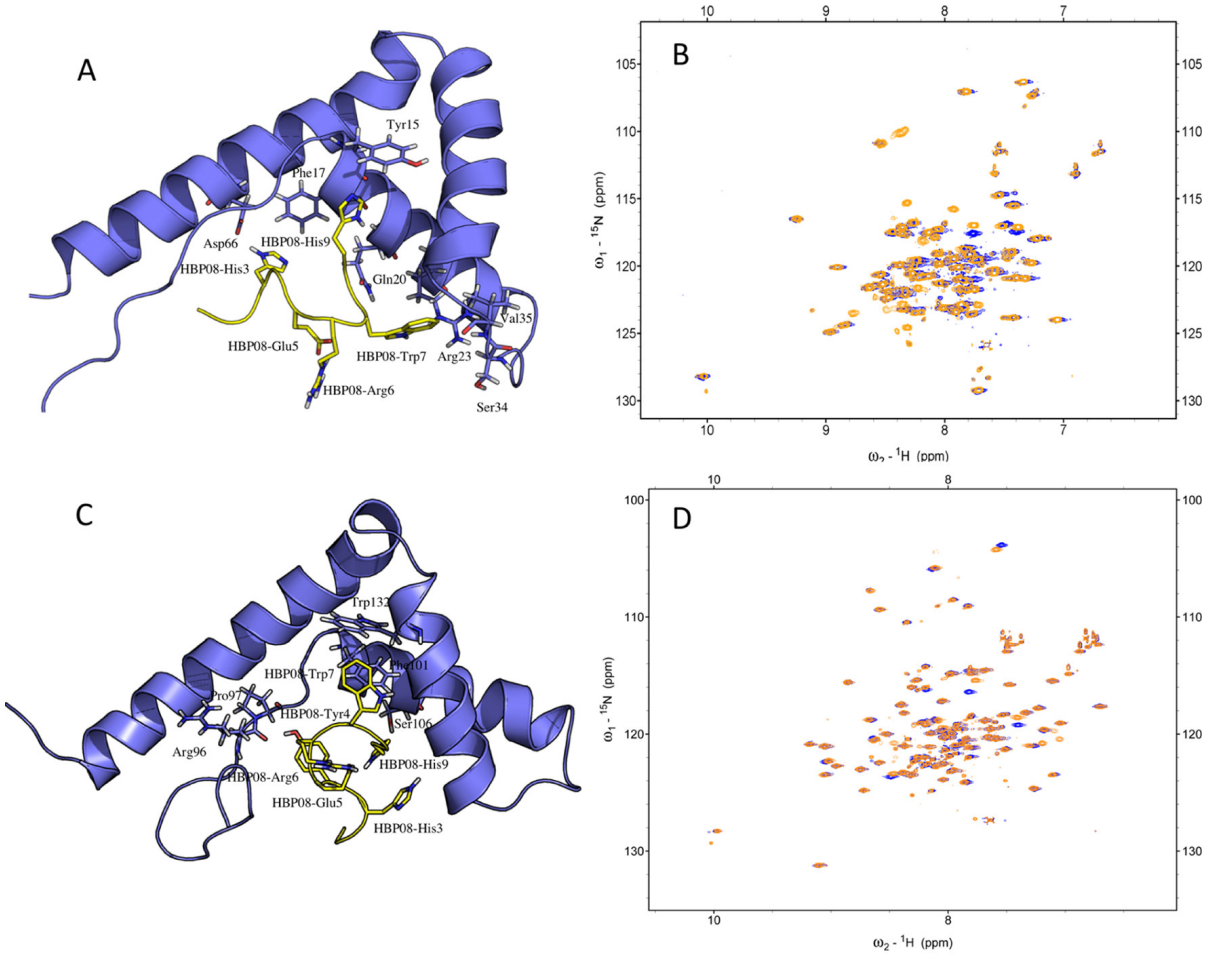


Figure 16. (A) Molecular model of the (A) HBP08-BoxA and (C) HBP08-BoxB complexes, obtained with HADDOCK. NMR spectra of (B) BoxA alone (blue) and in complex with HBP08 (orange) and (D) BoxB alone (blue) and in complex with HBP08 (orange).

At variance, regarding the interactions with BoxB, HBP08-Tyr4 forms a H-bond with the backbone of Arg96, HPB08-Trp7 was in contact with the aromatic rings of Phe101 and Trp132, and HBP08-His9 interacted with Ser106 by a H-bond (**Figure 16C**).

Finally, comparing the structure of the HBP08/BoxA complex with the one of the CXCL12/BoxA, obtained by previous NMR investigations (**Figure 17**) [116,117], it was noted that the binding site of HBP08 is in common with CXCL12, confirming the capability of the peptide to interfere with the creation of the CXCL12/HMGB1 heterocomplex.

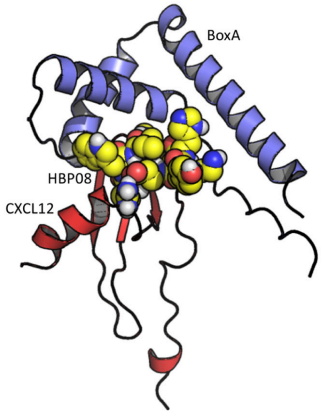


Figure 17. Comparison between the HBP08 binding mode and the structure of the CXCL12–BoxA complex obtained by docking in our previous study [128].

4.1.2. Design of HBP08 analogs with improved affinity on HMGB1-BoxB

So far, we have focused our attention on the pro-inflammatory effect induced by the formation of the fr-HMGB1/$(CXCL12)_2$ heterocomplex (**Figure 18B**), identifying a promising peptide named HBP08.

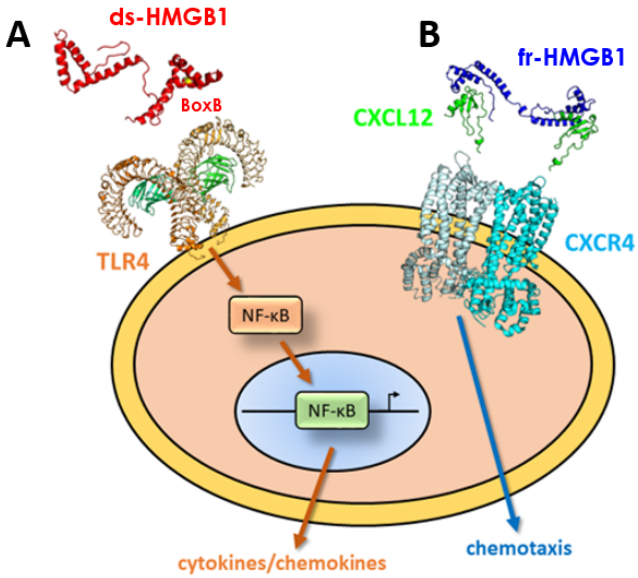


Figure 18. Representation of the pro-inflammatory pathways in which HMGB1 is involved. (A) the BoxB of the oxidized form of HMGB1 (ds-HMGB1) is able to interact with TLR4 receptor mediated by Cys106 of ds-HMGB1 [129]. This interaction induces the release of proinflammatory chemokines and cytokines able to activate innate and adaptive immune responses. (B) the reduced form of HMGB1 (fr-HMGB1) is able to form an heterocomplex with two CXCL12 chemokines whose N-terminals interact with the dimerized CXCR4 receptor. This interaction increases the leukocyte recruitment to the inflammatory site, triggering the inflammation.

Nevertheless, we showed that HBP08 selectively inhibits the CXCL12/HMGB1 heterocomplex activity, leaving HMGB1 able to trigger the TLR4 inflammatory pathway [128]. Consequently, in the second part of this project, the optimization of the HBP08 peptide sequence was planned, aiming at identifying new peptides with improved affinity on HMGB1-BoxB (**Figure 18A**).

Computational studies. Taking into account the NMR Chemical Shift data of the HBP08/HMGB1-BoxB complex [128], NMR-guided docking calculations of HBP08 on ds-HMGB1-BoxB, were accomplished. The structure representative of the most populated cluster of ds-HMGB1 conformations (which accounts the 54% of the total geometries) [117] has been selected as target for these calculations. The docking pose acquiring the highest Gscore was selected (**Figure 19A**) and the resulting complex was subjected to three independent 500 ns-long MD simulations. Then, considering the Cα-RMSD/time plot and the cluster analysis results, MM-GBSA calculations were performed on the snapshots extracted from the last 300 ns of each replica (**Table 5** and **Figure 19B**).

Table 5. Binding free energy (ΔG) values of each independent replica of the HBP08/HMGB1-BoxB complexes.

HBP08	ΔG ± SE (kcal/mol)
MD replica1	-31.9 ± 0.4
MD replica2	-31.2 ± 0.3
MD replica3	-36.4 ± 0.5

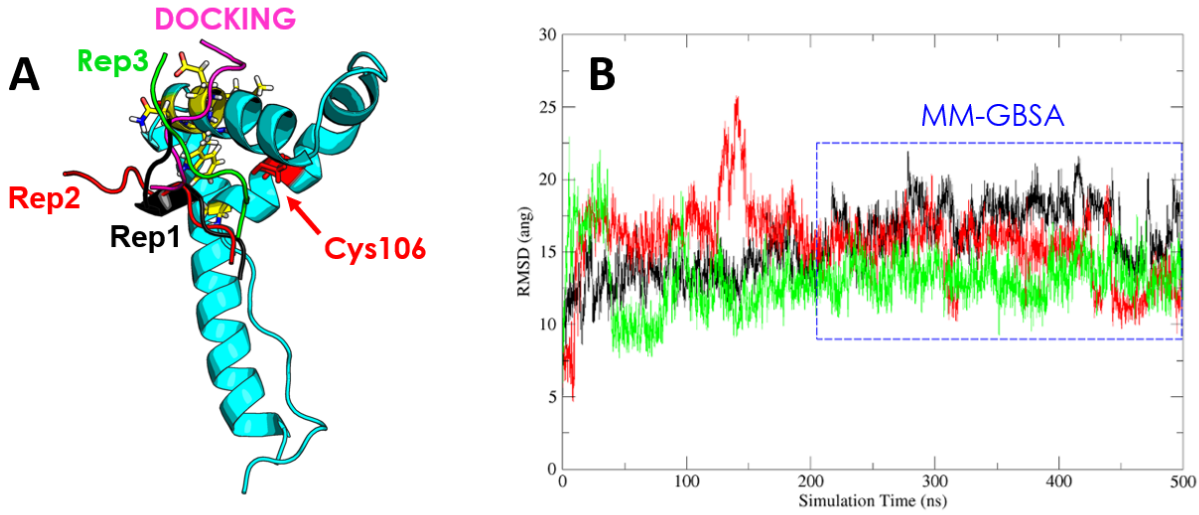


Figure 19. (A) The best docking pose (magenta), the most representative cluster conformations (black, green and red cartoons) found in each MD replica, and the NMR key residues (yellow sticks) are represented. (B) HBP08 RMSD analysis of replica1 (black), replica2 (red) and replica3 (green), the blue broken line highlighted the simulation time considered in the MM-GBSA analysis.

The attained results suggested that HBP08 tends to slightly shift from the initial docking pose to a more stable conformation, which is then retained for the remaining MD simulations. MM-GBSA calculations on HBP08 showed an average ΔG value of -33.2 kcal/mol. The conformational stability of the HMGB1/HBP08 complex was additionally ascertained by performing three independent 500 ns-long MD simulations on the third replica, the one in which HBP08 showed the lowest ΔG value (-36.4 kcal/mol). MM-GBSA calculations were performed on the snapshots extracted from about 300 ns of each replica and the results of the analysis is showed in **Figure 19B**.

The Cα-RMSD of HBP08 confirmed again the high stability of the obtained complex, in fact, the peptide remained stable over the whole simulation time. Only the replica3-3 displayed visible a slight variation of the binding mode after 400 ns of MD simulations, but it was noted that a binding mode close to the starting one was again adopted after only 50 ns of MD simulations (**Figure 20B**)**.** The HBP08 calculated ΔG values over the three replicas confirmed the good quality of the replica3. In fact, the average value obtained considering all the three independent replica was -36.5 kcal/mol (**Figure 20B**), which is very similar to the reference structure (-36.4 kcal/mol).

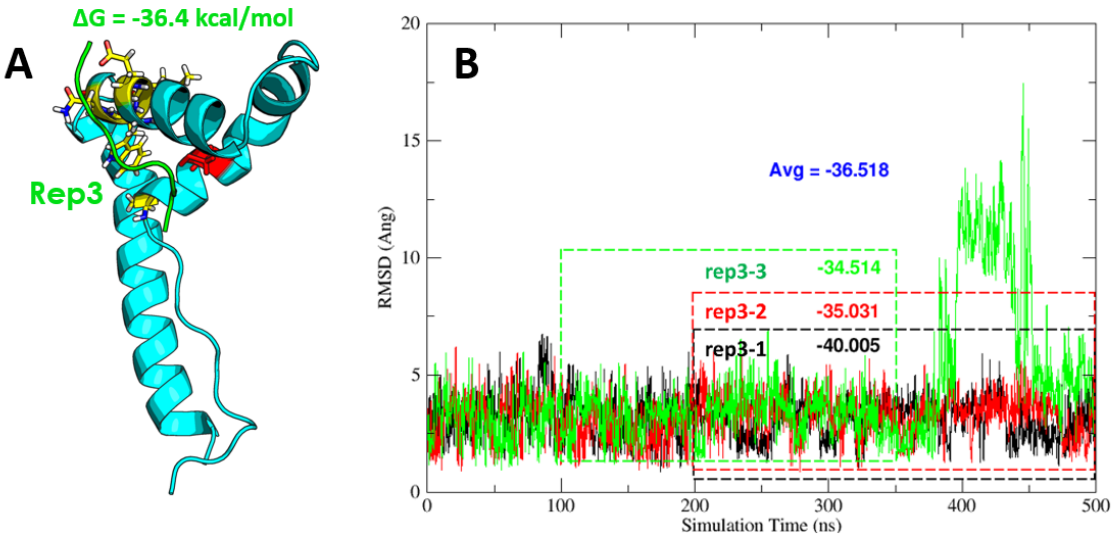


Figure 20. (A) The most representative cluster conformation of HBP08 replica3 (green), and the NMR key residues (yellow sticks) are represented. (B) RMSD analysis of replica3-1 (black), replica3-2 (red) and replica3-3 (green), in broken lines are highlighted the snapshots considered in the MM-GBSA analysis. The HBP08 ΔG values calculated by MM-GBSA for each replica and the average value are also reported.

Subsequently, in order to design new HBP08 analogs with improved affinity on HMGB1-BoxB, alanine scanning calculations were performed. Long-time scale MD simulations (500 ns) were

performed on each mutant peptide in complex with HMGB1-BoxB. Then, MM-GBSA calculations were again carried out to estimate the binding affinity of each mutant (**Table 6**).

Table 6. Binding free energy (ΔG) of the mutated peptides derived from the alanine scanning calculations.

Peptide	Sequence	ΔG ± SE (kcal/mol)
HBP08-rep3	GYHYERWIH	-36.4 ± 0.5
HBP08-Ala1	AYHYERWIH	-39.2 ± 0.3
HBP08-Ala2	GAHYERWIH	-35.9 ± 0.5
HBP08-Ala3	GYAYERWIH	-37.6 ± 0.4
HBP08-Ala4	GYHAERWIH	-33.8 ± 0.4
HBP08-Ala5	GYHYARWIH	-36.9 ± 0.3
HBP08-Ala6	GYHYEAWIH	-38.8 ± 0.4
HBP08-Ala7	GYHYERAIH	-26.5 ± 0.3
HBP08-Ala8	GYHYERWAH	-37.2 ± 0.5
HBP08-Ala9	GYHYERWIA	-32.6 ± 0.3

The attained results suggested that the Trp7 could be fundamental for the interaction of HBP08 to HMGB1-BoxB. In fact, a loss of about 10 kcal/mol in the theoretical binding free energy was observed if compared to the parent peptide. In addition, also Tyr4 and His9 appeared to be critical for the interaction, in fact it was observable a small loss of affinity for HMGB1-BoxB of 2.6 and 3.8 kcal/mol, respectively. Therefore, Tyr4, Trp7 and His9 were considered *hot spots* and were retained in the sequence of the new designed peptides. Conversely, the remaining residues that poorly contribute to the binding energy (*non-hot spots*) were systematically replaced by different amino acids capable of improving the complementarity between the new peptides and HMGB1-BoxB, through the "affinity maturation protocol". In particular, we decided to perform the affinity maturation on the Tyr2 and Arg6 residues, obtaining a total of 20^2 (i.e., 400) mutant peptides, finally ranked by the algorithm depending on the ΔAffinity and ΔStability values. The best three peptides by ΔAffinity, ΔStability, and both parameters (Mixed group) were subjected to 500 ns-long MD simulations (total = 9 peptides), also calculating the peptides ΔG values by MM-GBSA method (**Table 7**). Accordingly, the mutated peptide containing Arg2 and Met6 was the most promising one, since it showed a predicted ΔG value more than 3 kcal/mol lower than the parent peptide HBP08.

Table 7. Binding free energy (ΔG) of the mutated peptides derived from the affinity maturation protocol.

Group	Mutation	Sequence	ΔAffinity [1]	ΔStability [1]	ΔG ± SE [1]
HBP08-rep3	/	GYHYERWIH	/	/	-36.4 ± 0.5
Affinity	Y2P + R6F	GPHYEFWIH	-5.07	+93	-36.1 ± 0.4
	Y2N + R6Y	GNHYEYWIH	-5.04	+13	-33.9 ± 0.3
	Y2P + R6Y	GPHYEYWIH	-4.98	+93	-31.8 ± 0.4
Stability	Y2R + R6L	GRHYELWIH	-1.61	-2.91	-36.9 ± 0.5
	Y2M + R6L	GMHYELWIH	-1.72	-1.71	-37.3 ± 0.3
	Y2F + R6L	GFHYELWIH	-1.53	-1.53	-35.9 ± 0.3
Mixed	Y2R + R6M	G**R**HYE**M**WIH	-2.97	-1.47	-39.5 ± 0.3
	Y2R + R6I	GRHYEIWIH	-2.93	-0.94	-38.3 ± 0.5
	Y2R + R6V	GRHYEVWIH	-2.62	-0.26	-35.0 ± 0.5

[1] (kcal/mol).

For this reason, the peptide Y2R+R6M was additionally subjected to two independent 500 ns-long MD simulations replica, in order to better sample the conformational spaces of the complex and to acquire a more accurate prediction of the peptide ΔG value (**Table 8**).

Table 8. Binding free energy (ΔG) of the mutated peptide Y2R+R6M compared to the parent peptide HBP08.

Peptide	Sequence	ΔG Rep1 [1]	ΔG Rep2 [1]	ΔG Rep3 [1]	Average ΔG [1]
HBP08-rep3	GYHYERWIH	-40.0	-35.0	-34.5	-36.5
Y2R+R6M	G**R**HYE**M**WIH	-39.5	-39.5	-34.9	-38.0

[1] (kcal/mol).

The MD simulations replicas further confirmed that the peptide Y2R+R6M shows a slightly lower ΔG value. Thus, the representative structure of the most populated cluster of Y2R+R6M peptide (that accounts for 83% of conformational ensembles explored, **Figure 21**) was selected for an additional step of affinity maturation.

In this attempt, the remaining 4 *non-hotspot* residues of the Y2R+R6M peptide sequence (namely, Gly1, His3, Glu5 and Ile8) were mutated applying two runs of affinity maturation. Firstly (run 1), His3 and Ile8 were simultaneously mutated to all the possible natural amino acids, but we did not find any peptide showing improved theoretical binding affinity to HMGB1-BoxB (data not shown).

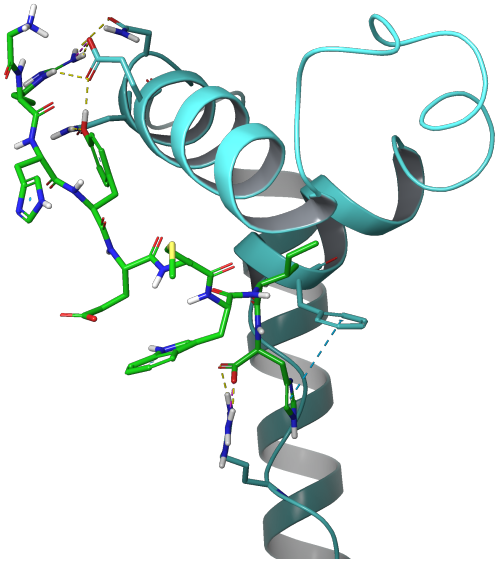


Figure 21. The representative structure of the most populated cluster of Y2R+R6M peptide considering 1.5 μs of MD simulations (that accounts for 83% of conformational ensembles explored) which was used as starting complex structure for the following affinity maturation steps.

As a consequence, both His3 and Ile8 were retained in the final sequence since they appeared to be fundamental for the interaction. Then (run 2), Gly1 and Glu5 were simultaneously mutated, and the best three mutant peptides ranked by ΔAffinity, ΔStability, and considering both parameters (Mixed group), were subjected to 500 ns-long MD simulations, and the peptides binding free energy vales were computationally estimated (**Table 9**).

Table 9. Binding free energy (ΔG) of the mutated peptides derived from the affinity maturation protocol.

Group	Mutation	Sequence	ΔAffinity [1]	ΔStability [1]	ΔG ± SE [1]
Y2R+R6M	/	GR**M**WIH	/	/	-38.0
Affinity	G1K + E5M	KRHYMMWIH	-10.46	+9.61	-40.6 ± 0.3
	G1K + E5F	KRHYFMWIH	-7.70	+10.98	-39.5 ± 0.2
	G1K + E5I	KRHYIMWIH	-7.67	+12.97	-39.9 ± 0.2
Stability	G1P + E5W	PRHYWMWIH	-2.31	-18.14	-39.8 ± 0.3
	G1P + E5Q	PRHYQMWIH	-3.40	-17.82	-41.0 ± 0.2
	G1P + E5R	PRHYRMWIH	-1.25	-16.33	-40.2 ± 0.2
Mixed	G1P + E5M	PRHYMMWIH	-6.41	-16.08	-41.3 ± 0.2
	G1N + E5M	NRHYMMWIH	-5.89	-4.35	*unbound*
	G1P + E5I	PRHYIMWIH	-3.61	-11.71	-40.3 ± 0.2

[1] (kcal/mol).

Interestingly, all mutant peptides displayed ΔG values lower that the parent peptide (Y2R+R6M). The peptides showing ΔG values more than 2 kcal/lower were selected for further computational studies. In particular, they were subjected to two additional, and independent,

500 ns-long MD simulations, in order to increase the statistical significance of the results. The results are resumed in **Table 10**.

Table 10. Binding free energy (ΔG) of the mutant peptides compared to the parent peptide Y2R+R6M.

Peptide	Sequence	ΔG Rep1 [1]	ΔG Rep2 [1]	ΔG Rep3 [1]	Average ΔG [1]
Y2R + R6M	GRHYEMWIH	-39.5	-39.5	-34.9	-38.0
G1K + E5M	KRHYMMWIH	-40.6	-39.7	-38.3	-39.5
G1P + E5Q	PRHYQMWIH	-41.0	-42.2	-40.7	-41.3
G1P + E5R	PRHYRMWIH	-40.2	-43.0	-40.3	-41.1
G1P + E5I	PRHYIMWIH	-40.3	-39.7	-40.6	-40.2
G1P + E5M	PRHYMMWIH	-41.3	-41.2	-43.3	-41.9

[1] (kcal/mol).

The MD results highlighted that the mutated peptides G1P+E5M (named HBP08pep1) and G1P+E5Q (named HBP08pep2), which differs by only one amino acid from the template peptide, were endowed with the lowest predicted ΔG values. Thus, considering these data, both peptides were purchased and tested by MST experiments using recombinant HMGB1.

MST experiments. Performing experiments following the protocol previously described, HBP08pep2 displayed a K_d on HMGB1-BoxB of 13.7 ± 2.2 nM, which is more than 1000-fold lower than HBP08 (17 µM) as shown in **Figure 22A**.

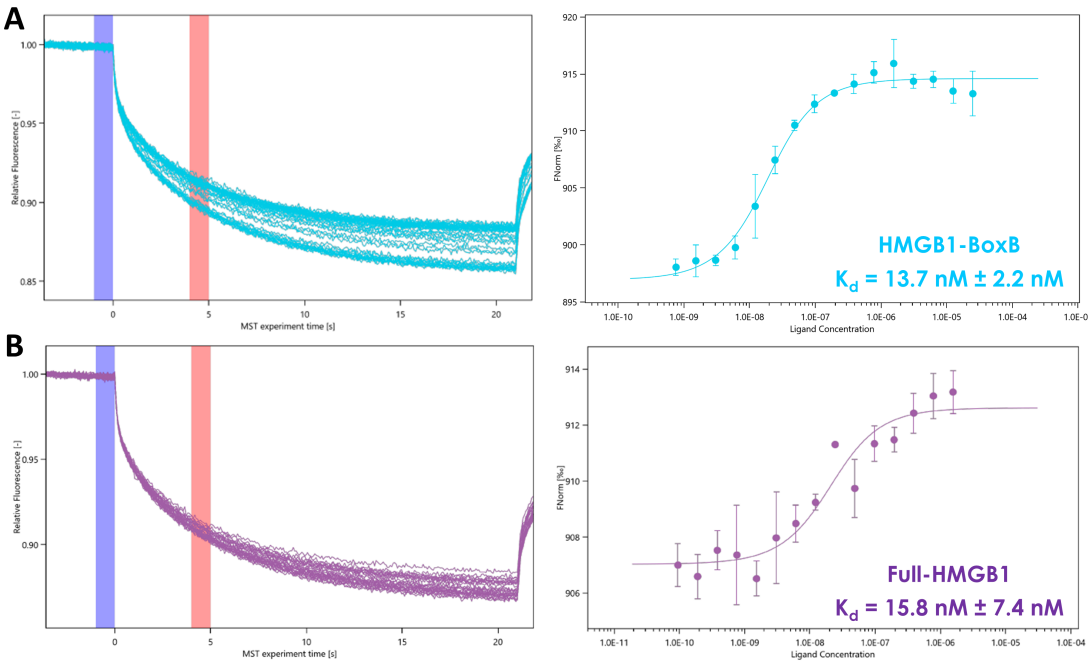


Figure 22. MST experiment of HBP08pep2 to (A) HMGB1-BoxB and (B) full sequence HMGB1 protein in three independent replicas. On the left it is observable the MST traces, while on the right the K_d curve.

This value was further confirmed by repeating the experiment using the full version of HMGB1 (K_d = 15.8 ± 7.4 nM) as represented in **Figure 22B**. Conversely, a K_d of 3.85 ± 0.41 µM was measured for HMGB1-BoxA, a value about 5 times higher than the one reported for HBP08 (0.8 µM). MST experiments on HBP08pep1 showed similar results to the one reported for HBP08pep2 but since it displayed a lower solubility in the PBS buffer, the further *in vitro* chemotaxis assays were performed only on HBP08pep2, in order to evaluate the peptide efficacy as inhibitor of the CXCL12/HMGB1-induced migration.

Biological evaluation. Biological assays were accomplished by the research group of Prof. M.G. Uguccioni (Institute for Research in Biomedicine, IRB, Bellinzona, Switzerland). The attained results revealed that HBP08pep2 efficiently inhibited the enhanced migration induced by the CXCL12/HMGB1 heterocomplex, showing an IC_{50} of about 2.5 µM, which is 20-fold lower than the value observed for the template peptide HBP08 (50 µM) [128] (**Figure 23**).

Further *in vitro* experiments are still ongoing, with the aim at clarifying if HBP08pep2 is also able to block the interaction between HMGB1 and TLR4, i.e. interrupting the pro-inflammatory cascade determined by this pathway. Finally, NMR CSP experiments are also currently ongoing to acquire atomistic details of the interaction between HBP08pep2 and HMGB1-BoxB.

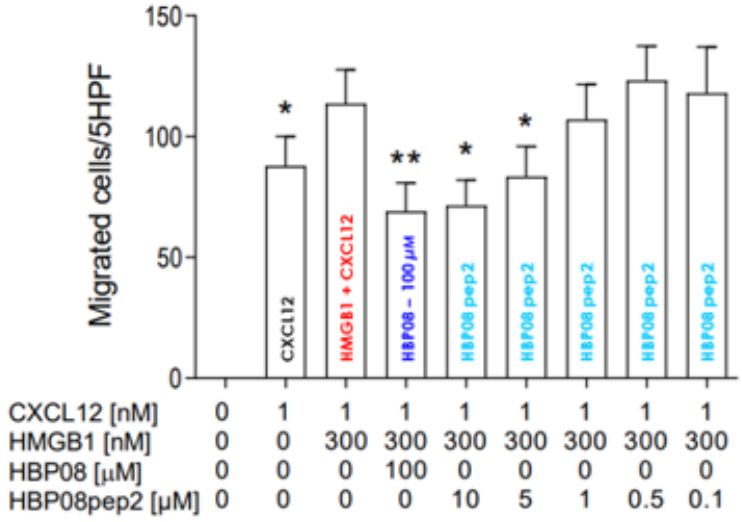


Figure 23. Inhibition of cell migration in response to the CXCL12/HMGB1 heterocomplex was assessed on 300-19 Pre-B cells transfected with CXCR4 using scaling concentrations of HBP08pep2.

4.2. PCSK9 and HMG-CoAR

Proprotein convertase subtilisin/kexin 9 (PCSK9) regulates the plasma low-density lipoprotein cholesterol (LDL-c) levels by the direct interaction with the LDL receptor (LDLR) [130]. Their protein-protein interaction (PPI) on the surface of the hepatocytes directs the LDLR to degradation via the lysosome pathway. As a consequence, the physiological activity of PCSK9 leads to an increment in the circulating cholesterol levels by decreasing the LDLR population on liver cell membranes. For this reason, the PCSK9 inhibition represents a valuable therapeutic strategy for the treatment of hypercholesterolemia and consequent coronary heart disease [131].

The most common drugs for hypercholesterolemia treatment are statins, which inhibit 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoAR), the rate-limiting enzyme in cholesterol biosynthesis. This enzyme lowers intracellular cholesterol levels, leading to an increased expression of LDL receptors (LDLR) on cell surfaces and a reduction of serum LDL-cholesterol (LDL-C) via the activation of the sterol-regulatory element-binding protein (SREBP)-2 transcription factor pathway. Although this approach is considered an efficient way to reduce circulating LDL-C, cardiovascular events still occur in some patients. Moreover, statins induce known and serious side effects, such as headache, muscle and joint pain, and a higher risk of developing diabetes. For this reason, PCSK9 (discovered in 2003) has been recognized as one of the most promising targets for counteracting hypercholesterolemia and atherosclerotic cardiovascular diseases [131].

The expression of PCSK9 is also controlled by the activity of SREBP-2 as well as a specific transcriptional activator hepatocyte nuclear factor-1α (HNF-1α) [132], which is a liver-enriched transcription factor regulating many target [133] genes in the liver and intestine. In contrast, the ability of SREBP-2 to co-stimulate the PCSK9 and LDLR expression limits the therapeutic efficacy of statins which are known to produce their effects via SREBP-2 activation. Indeed, it is well documented that statins improve the PCSK9 protein level production through the augmentation of the intracellular HNF-1α levels [134].

Hence, in the last two decades, academia and pharmaceutical companies have financed considerable research studies on the development of compounds capable of target PCSK9 developing different strategies including siRNA, anti-sense oligonucleotides (ASOs), and monoclonal antibodies (mAbs).

In fact, Evolocumab [135] and Alirocumab [136], two mAbs endowed with PCSK9 inhibitory activity, have been released on the market to treat statin-resistant hypercholesterolemia. However, these therapeutic tools are characterized by high cost and low patient compliance. Therefore, both industry and academic research activities are focused on the development of alternative therapeutic strategies like the usage of small molecules [137], but also peptides [138], or peptidomimetics [139], capable of impairing the PCSK9/LDLR PPI.

In this field, Lammi *et al*. sorted out the most potent natural peptide (LILPKHSDAD, P5) derived from the peptic lupin (Lupinus A.) protein hydrolysate, which impairs the PCSK9-LDLR interaction endowed with an IC_{50} of 1.6 μM [140]. In parallel, P5 reduces the catalytic activity of HMG-CoAR with an IC_{50} value of 147.2 μM, through the inhibition of the enzyme activity [140]. The computational PCSK9/P5 model was obtained by performing a molecular docking of the most favorable conformation of peptide P5, selected after 40 ns of MD simulations, into the binding site of PCSK9, which was recognized by the presence of the EGF-A domain of LDLR in the PCSK9 crystal structure (PDB accession code 4NE9) [141]. Finally, the system was simulated by MD for 50 ns, to optimize the small peptide geometries.

Moreover, P5 increases the LDLR protein level in HepG2 cells through the activation of SREBP-2 and, through a downregulation of HNF-1α, it reduces the PCSK9 protein levels and secretion in the extracellular environment. This unique synergistic multi-target inhibitory behaviour of P5 determines the improved ability of HepG2 cells to uptake extracellular LDL, with a final hypocholesterolemic effect. P5 was successfully transported by differentiated human intestinal Caco-2 cells through transcytosis, and, during transport, it was partially metabolized in a breakdown fragment (LPKHSDAD, P5-met), which retained the biological activity of the parent peptide [142]. In facts, we have demonstrated that P5-met reduces PCSK9-LDLR binding with a dose-response trend and an IC_{50} of 1.7 μM and inhibits the HMG-CoAR with an IC_{50} of 175.3 μM [142].

At a cellular level, such as P5, P5-met improves the LDLR and reduces PCSK9 levels, through the modulation of both SREBP-2 and HNF-1α, respectively [142]. Therefore, since P5-met displayed the same activity and behaviour of the parent peptide, P5, our results indicated that the first two amino acid residues (LI) do not play a key role in the interaction with both PCSK9 and HMG-CoAR target.

These data clearly indicate that P5, with its dual-inhibitory activity, represents a new alternative strategy to the use of single classical PCSK9 and HMG-CoAR inhibitors. Notably, the strategy in which dual-inhibitors are employed may be more effective in overcoming the deficits attributed to the classical use of statins (adverse effects and co-stimulation of PCSK-9 and LDLR via a common transcriptional activator, i.e., SREBP-2, in statin-treated patients limited the efficacy of these classical HMG-CoAR inhibitors) or PCSK9 inhibitors (including expensiveness, low compliance of the patients, repeated administrations, and injection site irritations) on health and to meet the desired health goals and public priorities in terms of safety and cost-related issues.

4.2.1. Design of P5 analogs with improved hypocholesterolemic activity

Considering all the previous observations, the overall aim of the present study is the identification of new P5 analogs able to target both PCSK9 and HMG-CoAR, therefore displaying an improved and dual hypocholesterolemic activity. Hence, the theoretical study was validated and confirmed by performing a detailed biological investigation on the most promising P5 analogs.

Computational studies. To obtain a robust hypothesis on the peptide P5 binding mode, the PCSK9/P5 complex model was optimized by performing 1 µs-long MD simulations. At the end of these simulations, the MD trajectory frames were grouped using a cluster analysis algorithm to determine which was the most preferred P5 conformation in complex with PCSK9. The PCSK9/P5 complex conformation representative of the most populated cluster (78%) suggested that peptide P5 could bind to PCSK9, as illustrated in **Figure 24A**.

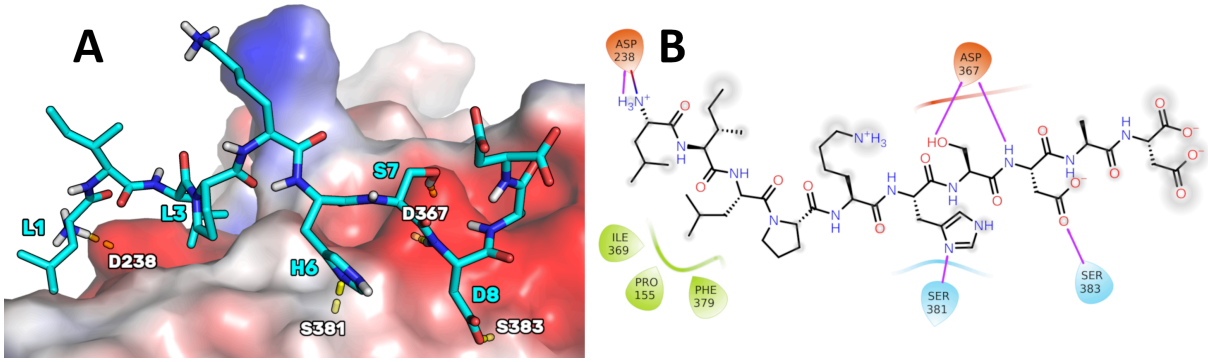


Figure 24. (A) The expected P5 (cyan sticks) binding mode on the PCSK9 surface after 1 µs-long MD simulations. Yellow dotted lines highlight the H-bond network. The enzyme solvent-accessible surface is depicted accordingly by the partial charge of the residues: blue for positive areas and red for negative areas. (B) 2D representation of the predicted P5 binding mode in complex with PCSK9. H-bonds and salt bridges are showed in purple and blue/red lines, respectively. Hydrophobic residues interacting with P5 are colored in green.

In particular, P5 could bind to PCSK9 through (1) a salt bridge between the charged NH term of P5-Leu1 and the side chain of Asp238, (2) an H-bond between the imidazole ring of the P5-His6 and the NH group of Ser381, (3) an H-bond network between the side chain of P5-Ser7, the NH of P5-Asp8, and the side chain of Asp367, and (4) an H-bond shaped by the side chain of P5-Asp8 and the side chain of Ser383. The side chain of P5-Leu3 was deeply inserted into a hydrophobic basin sized by the PCSK9 residues Phe379, Pro155, and Ile369, creating van der Waals interactions (**Figure 24B**).

Then, to identify new P5 analogs endowed with improved PCSK9 affinity, we designed new peptides by performing a computational alanine-scanning mutagenesis analysis (widely described in Chapter 2.1). Specifically, nine 3D models, in which PCSK9 was in complex with each alanine-mutated peptide P5, were simulated by 200 ns-long MD simulations, and the estimation of the binding free energy values (ΔG, **Table 11**) were achieved using MM-GBSA approach.

Table 11. Estimated ΔG values of the peptides under investigation, as calculated using the MM-GBSA approach.

Peptide/Mutation	Sequence	ΔG value ± SE [1]	ΔΔG value [1]
P5	LILPKHSDAD	-18.9 ± 0.5	0
L1A	AILPKHSDAD	-13.5 ± 0.5	+5.4
I2A	LALPKHSDAD	-20.9 ± 0.5	-1.0
L3A	LIAPKHSDAD	-6.6 ± 0.7	+12.3
P4A	LILAKHSDAD	-23.0 ± 0.7	-4.1
K5A	LILPAHSDAD	-14.1 ± 0.6	+4.8
H6A (P5-H6A)	LILPKASDAD	-1.2 ± 0.5	+17.7
S7A (P5-S7A)	LILPKHADAD	-19.3 ± 0.3	-0.4
D8A	LILPKHSAAD	-21.9 ± 0.3	-3.0
D10A	LILPKHSDAA	-19.6 ± 0.5	-0.7

[1] (kcal/mol).

The attained results suggest that positions 3 and 6 can be considered *hotspots*, as their mutation into alanine led to P5 analogs endowed with a considerable reduction of the predicted binding affinity (ΔΔG higher than 10 kcal/mol). Specifically, P5-His6 seemed crucial for peptide binding because its substitution led to a dramatic drop in the peptide binding interaction energy. MD simulations suggested that the substitution of the alkaline side chain of His6 with a methyl group led to a change in the peptide binding mode due to a lack of an H-bond between the PCSK9-Ser282 amide group and the imidazole ring of P5-His6. For this reason, the [H6A] peptide P5 was unbound from the PCSK9 surface after the initial steps of

the MD simulations. Similarly, the removal of the side chain of P5-Leu3 led to a peptide incapable of maintaining the P5 initial binding mode, as the hydrophobic contacts engaged by the Leu3 isobutyl group with the hydrophobic crevice sized by the PCSK9 residues Leu159, Pro156, Ala240, and Ile370 were missing.

The alanine mutation of Leu1 and Lys5 led to peptides with a calculated binding affinity that was slightly lower than that of P5 (ΔΔG close +5 kcal/mol). However, given the inaccuracy of the MM-GBSA calculations and the observation that the side chains of Leu1 and Lys5 fluctuating in a solvent environment do not stably bind to PCSK9 during the MD simulations, positions 1 and 5 cannot be considered strong hotspots similar to positions 3 and 6. Conversely, the substitution with alanine of Ile2, Ser7, and Asp10 of P5 led to peptides with a predicted binding affinity close to that predicted for the template peptide. Therefore, they can be considered non-hotspots and can potentially be substituted with different amino acids. However, the predicted data on Leu1 and Ile2 are in accordance with our recent experimental data, which show that a metabolite of peptide P5 that does not contain the first two residues (P5-met, LPKHSDAD) displays an IC_{50} value close to the parent peptide P5 [142].

Conversely, the P4A and D8A mutant peptides showed a higher affinity to PCSK9 than P5. However, as the gain in the ΔG value was not extremely high, the synthesis and biological evaluation of these peptides was not considered suitable.

The alanine-scanning study showed that the positions 1, 2, 7, and 10 on the P5 sequence could be considered *non-hotspots*. Moreover, the alanine in position 9 should be considered a *non-hotspot*, as alanine is already present in the natural P5 sequence. Nevertheless, the P5-Ser7 OH group could create an H-bond with PCSK9-Asp367, and the P5-Asp10 side chain could be involved in the fold of the peptide, as an internal H-bond could be shaped with the side chain of P5-Ser7. Thus, we decided to mutate the residues in positions 2 and 9 to develop novel P5 analogs with improved PCSK9 binding affinity.

Accordingly, with this assumption, 20^2 peptide P5 analogs were computationally designed through the systematic substitution of positions 2 and 9 with all natural amino acids. Their theoretical affinity for PCSK9 was preliminary evaluated by the Prime algorithm, which can estimate the peptide binding free energy using the MM-GBSA approach. PCSK9 in complex with the 10 top-ranking P5 analogs (i.e., those with the lowest ΔAffinity values, **Table 12**) again underwent MD simulations. The ΔG values were estimated using the MM-GBSA protocol,

which allowed for the acquisition of ΔG values comparable with those previously attained for P5 and other P5 alanine mutants.

Table 12. Estimated binding affinity values of newly designed P5 analogs (columns 1-2), calculated by Prime software (ΔAffinity, column 3) and standard MD/MM-GBSA calculations (ΔG values, column 4).

Peptide	Sequence	ΔAffinity [1]	ΔG ± SE [1]
P5	LILPKHSDAD	/	-18.9 ± 0.5
[I2P-A9R] P5	LPLPKHSDRD	-23.9	-26.3 ± 0.6
[I2M-A9R] P5	LMLPKHSDRD	-23.5	-26.3 ± 0.6
[I2R-A9R] P5	LRLPKHSDRD	-21.5	-30.3 ± 0.7
[I2Q-A9R] P5	LQLPKHSDRD	-21.0	-24.8 ± 0.8
[I2L-A9R] P5	LLLPKHSDRD	-20.7	-25.4 ± 0.6
[I2Y-A9R] P5 (P5-Best)	LYLPKHSDRD	-20.4	-41.7 ± 0.7
[I2H-A9R] P5	LHLPKHSDRD	-20.4	-24.6 ± 0.8
[I2T-A9R] P5	LTLPKHSDRD	-19.5	-28.0 ± 0.5
[I2F-A9R] P5	LFLPKHSDRD	-19.4	-19.2 ± 0.3
[I2E-A9R] P5	LELPKHSDRD	-19.1	-24.5 ± 0.6

[1] (kcal/mol).

The Prime calculations (third column of **Table 12**) suggested that the peptides acquiring an improved predicted binding energy were the ones containing arginine in position 9. At variance, the substitutions in position 2 did not considerably affect the affinity of the resulting peptides (differences in the ΔAffinity values followed in the range of 5 kcal/mol). Subsequently, using the AMBER20/MM-GBSA calculations, the resulting ΔG values spanned from -19 to -42 kcal/mol. This allowed us to assess that the peptide [I2Y-A9R]P5 (i.e., P5-Best) was endowed with the highest predicted PCSK9 binding affinity. In fact, the ΔG value of P5-Best was two times the value predicted for the template peptide P5, suggesting that P5-Best could show an affinity to PCSK9 appreciably lower than P5. Our simulations showed that, as indicated in the conformation representative of the most populated cluster (**Figure 25A**), P5-Best could acquire an ameliorated PCSK9 complementarity because of the possibility of creating two salt bridges: the first between the new arginine in position 9 and the side chains of Glu366 and Asp367, and the second between the side chain of P5-Best-Asp10 and the side chain of Lys222. These interactions were also enforced by the presence of an H-bond between the P5-Best-Asp10 and the OH group of Ser225. Moreover, the P5-Best-Ile3 side chain was in contact with the hydrophobic pocket sized by Phe379, while the phenol ring of the new residue P5-Best-Tyr2 was located close to Arg194. The NH groups of P5-Best-Tyr2 and -Ile3 created two H-bonds with the side chain of Asp238.

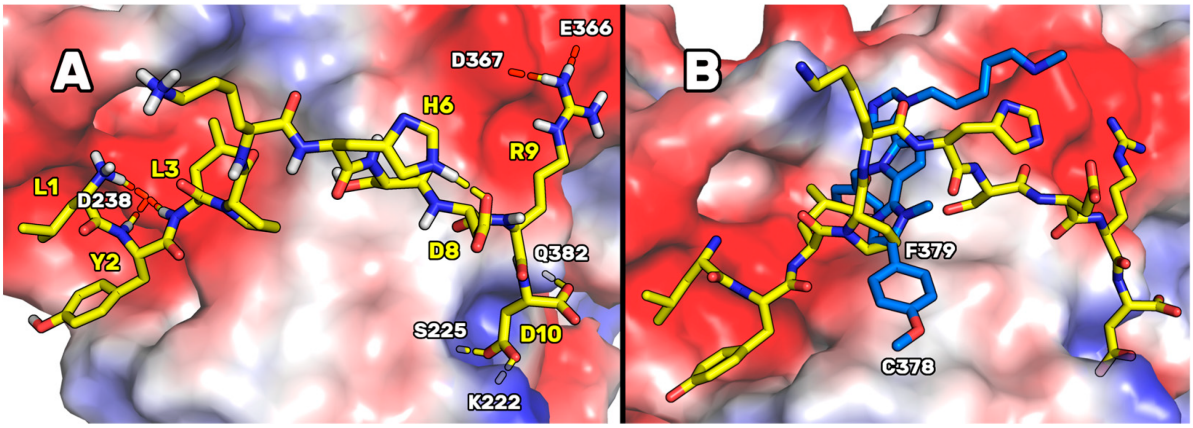


Figure 25. (A) Expected binding mode of [I2Y-A9R]P5 (i.e., P5-Best) on the PCSK9 surface resulting from the MD simulations and cluster analysis. Yellow dotted lines depict the H-bond network. The enzyme solvent-accessible surface is colored according to the partial charge of the residues: blue for positive areas and red for negative areas. P5-Best is represented as yellow sticks. (B) Superimposition of P5-Best and Rim13, both bound on the PCSK9 surface.

These results were also compared to the computational data attained designing the poly-imidazole derivatives capable of inhibiting PCSK9 [143]. In fact, in our previous paper, by applying a computational approach such as the one here applied, we designed and biologically evaluated two poly-imidazole derivatives endowed with PCSK9 inhibiting activity. The biological evaluation of the most interesting poly-imidazoles, named Rim13 and Rim14, allowed us to report on their ability to modulate the LDLR expression on the human hepatic HepG2 cell surface, and their capacity to increase the extracellular uptake of LDL by the same cells. Here, structurally aligning the P5-Best and the Rim13 hypothetical binding modes, we noted that the backbone atoms of the peptide residues Pro4 and Lys5 were mimicked by the first two imidazole rings of Rim13 (**Figure 25B**). Moreover, the benzyl chain of the second imidazole ring of Rim13 was projected in the same hydrophobic cleft shaped by Phe379 and occupied by the side chain of P5-Best-Leu3, creating van der Waals interactions. Furthermore, the negatively charged area created by the PCSK9 residues Asp367 and Glu366 were in contact with the side chain of the P5-Best-Arg9 and the amino-methyl chain of Rim13. Since they bind similarly, creating contacts with the same PCSK9 residues, this alignment could help in the design of new poly-imidazole derivatives. In fact, aiming at designing more potent poly-imidazoles derivatives, the benzyl moiety of Rim13 could be substituted by alkyl chains (linear or not), to reproduce the interactions played by the P5-Best-Leu3 residue. Conversely, regarding the design of new P5 analogs, the Pro4 of P5-Best could be replaced by aromatic residues such as Phe, Tyr or Trp, in order to reproduce the interactions played by the p-

methoxyphenyl ring of Rim13. However, the oral pharmacokinetic (PK) properties of peptides remain strongly limited by the presence of degrading enzymes in the gastrointestinal tract. Nevertheless, the research efforts are still devoted to solving this limitation. In fact, active peptides could be orally administered together with penetration enhancers, within hydrogels or in combination with digestive enzyme inhibitors. As alternative, they can be suitably coated by acid-stable polymers or administered through intestinal patches [145]. By means of one of these innovative delivery strategies, even peptides active in the high micromolar range could be successfully employed for the treatment of several pathologies. Actually, numerous peptides are in phase III of clinical trials but, until now, only desmopressin is available in the market, and used in the clinic [145].

In light of these theoretical studies, empirical assays were performed on the [H6A] peptide P5 (i.e., P5-H6A) because position 6 was recognized as a *hotspot* (**Table 11**), on P5-Best because of its lowest predicted binding free energy value, and on [S7A]P5 (i.e., P5-S7A) because it represents one of the peptides for which the alanine mutation did not remarkably alter the predicted binding free energy value. Conversely, the mutation of P5-Asp10 into alanine affected peptide folding (as shown by MD simulations) and the water solubility of the peptide, as the negatively charged side chain of Asp10 should be substituted with the aliphatic methyl group of alanine. Thus, the peptides P5-H6A, P5-S7A, and P5-Best were purchased by GenScript and biochemically evaluated by *in vitro* experiments.

Biological experiments. In order to verify whether P5 derivatives impair the PPI between PCSK9 and LDLR, dedicated biochemical experiments were accomplished in the group of Prof. C. Lammi (University of Milan). The results showed that P5-Best, P5-H6A and P5-S7A reduced the PCSK9-LDLR binding with a dose-response trend and IC_{50} values of 0.7, 9.0, and 1.45 μM, respectively (**Figure 26A**). The results confirmed that the new P5 derivatives were more active than P5 (1.6 μM). These data are in line with the computational predictions. In fact, P5-Best peptide showed the lowest ΔG value calculations (ΔG= -41.7 kcal/mol) indicating it as the most active peptide, while P5-S7A displayed binding affinity in the same range of P5 (ΔG values of -19.3 and -18.9 kcal/mol, respectively). However, by our computational studies, P5-H6A should be do not active since the side chain of H6 plays a crucial role in the stabilization of the peptide on the PCSK9 surface. In our view, the lack of linearly between the binding affinity

experimental data (IC_{50} values) and the computational predictions, could be due to the omitted calculation of the entropic contribution to binding free energy value. In fact, the calculation of this contribution is highly computationally demanding, and the error associated with the estimation is very often greater than the value itself. Moreover, the data attained by the further biological investigations on these peptides cannot be compared with the computational predictions, since it is very difficult to discuss the *in silico* results in comparison with the biological data obtained from the HepG2 cells. In fact, the molecular modeling studies have been performed on a PCSK9 model immersed in a box of water molecules and the biological experiment capable of reproducing these conditions is only the one in which the recombinant PCSK9 is in contact with the LDLR (e.g., the binding assays displayed in **Figure 26**). Conversely, when the biological properties of peptides are assessed in complex experimental conditions, such as the one in which cells are involved, the molecular modeling results cannot be linearly compared with the experimental data. In fact, the effects of membranes, extracellular or intracellular enzymes cannot be considered by our calculations.

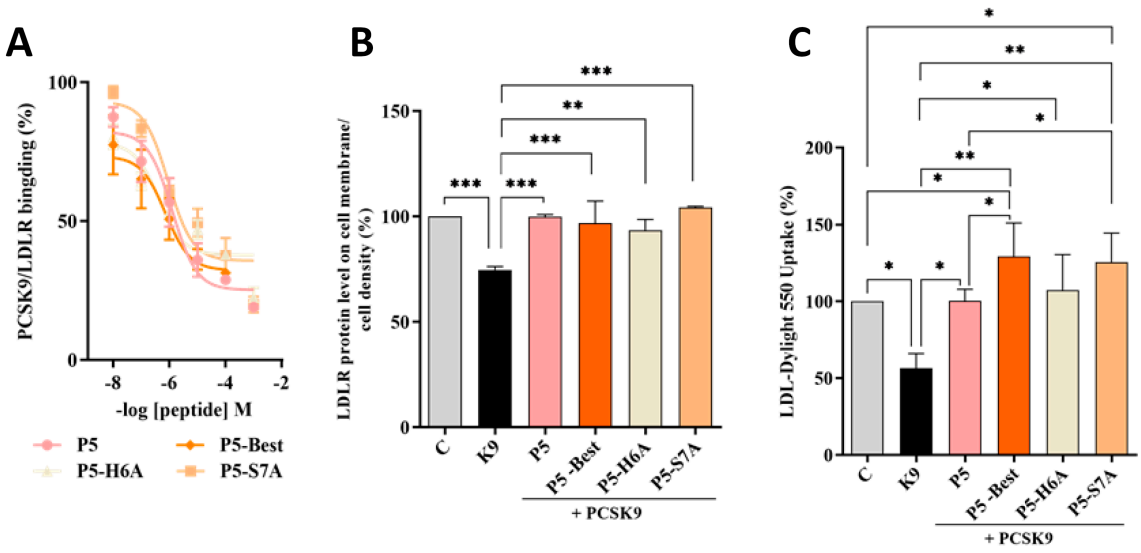


Figure 26. Inhibition of the PPI between PCSK9 and LDLR. (A) Impairment of the protein-protein interaction between PCSK9 and LDLR. (B) The treatment of HepG2 cells with PCSK9 (K9 in the graphs, 4 µg/mL) reduced the active LDLR protein levels localized on the surface of the cells, which were restored by P5 and P5 analogs (50 µM). (C) The decreased functional ability of HepG2 cells to absorb LDL from the extracellular space observed after incubation with PCSK9 (4 µg/mL) improved after treatment with P5 and P5 analogs (50 µM). The data points represent the average ± SD of four independent experiments performed in duplicate.

Furthermore, the ability of these P5 analogs to modulate the levels of LDLR localized on HepG2 surfaces was investigated in the presence of PCSK9 (4 µg/mL) using an in-cell western (ICW) assay. The results showed that the LDLR levels decreased in the presence of PCSK9 alone by

25.4 ± 1.6% compared with the untreated control cells, and that P5-Best, P5-H6A, and P5-S7A could significantly restore the LDLR levels to 96.9 ± 10.4%, 93.4 ± 5.2%, and 104.4 ± 0.4%, respectively, when co-incubated with PCSK9 (**Figure 26B**). Finally, functional cell-based assays were performed to investigate the ability of HepG2 cells to uptake extracellular LDL in the presence of PCSK9. HepG2 cells incubated with PCSK9 alone showed a 43.6 ± 9.6% reduction in the uptake of fluorescent LDL compared with the untreated control cells. This result is in agreement with the reduction of active LDLR population on the cell surface, which was observed by ICW. After co-incubation with PCSK9, P5-Best, P5-H6A, and P5-S7A completely restored the LDLR function, increasing the LDL uptake to 129.2 ± 21.9%, 107.4 ± 23.0%, and 125.4 ± 19.0% (**Figure 26C**), respectively.

P5 analogs demonstrated to be more active than peptide Pep2-8 (TVFTSWEEYLDWV) [145] and its analogs [Y9A]Pep2-8 and [T4R,W12Y]Pep2-8 [146] as PPI inhibitors of PCSK9. In details, at the fixed concentration of 100 µM, Pep2-8 impaired the PCSK9-LDLR binding by -36.5% *vs.* the control, whereas [Y9A]Pep2-8 and [T4R,W12Y]Pep2-8 by -69.8% and -93.0%, respectively [146]. Indeed, the IC_{50} value of [Y9A]Pep2-8 was equal to 27.12 ± 1.2 µM and that of [T4R,W12Y]Pep2-8 equal to 14.50 ± 1.3 µM [146], clearly indicating that P5-Best is about 30- and 20-fold more potent than the Pep2-8 mutant peptides. In addition, at cellular levels, Pep2-8 and both Pep2-8 analogs were less efficient than P5 analogs to restore the LDLR protein levels and the functional ability of hepatocytes to absorb LDL from the extracellular environment [146]. On the contrary, P9-38, a cyclized Pep2-8 analogue, demonstrated to be 35-fold more potent than P5-Best in impairing the PPI between PCSK9/LDLR displaying and IC_{50} equals to 20 nM, and it was 1000-fold more potent to restore the LDLR level and functionality in HepG2 cells [147].

Finally, P5-Best is slightly more potent than the poly-imidazole Rim13 which inhibit the interaction between PCSK9 and LDLR by an IC_{50} equals to 1.4 µM, a value similar to the reference peptide P5. In the same range of concentration of Rim13, P5 analogs successes in the restoring the functional activity of LDLRs on the surface of hepatocytes preventing their degradation [147].

Although all P5 analogs successfully restored the level of LDLR protein similar to peptide P5, statistical analysis revealed that from a functional point of view, both P5-Best and P5-S7A not only restored the ability of hepatic cells to uptake LDL from the extracellular environment but also improved this capability against untreated cells. These results suggest that the

hypocholesterolemic effect occurs with a dual mechanism of action involving the modulation of HMG-CoAR activity and protein levels. To assess this aspect and deepen this behaviour, further HMG-CoAR activity assay and western blot experiments were performed.

P5 Analogs Modulate the Hepatic LDLR Pathway by Inhibiting HMG-CoAR Activity. To better characterize the dual inhibitory activity of all P5 analogs, a biochemical investigation was conducted to assess their effect on the modulation of HMG-CoAR activity. The results suggested that P5-Best, P5-H6A, and P5-S7A inhibited enzyme activity with an IC_{50} of 88.9, 74.4, and 73.8 µM, respectively, showing more effective inhibitory activity than P5 (147.2 µM) (**Figure 27A**), but they are still less active than statins. In facts, the IC50 values for the inhibition of HMG-CoAR activity for pravastatin simvastatin, atorvastatin, and rosuvastatin are equals to 44.1, 11.2, 8.2, and 5.4 nM, respectively [148,149].

Even though, P5 and P5 analogs are less active than statins as HMG-CoAR inhibitors and their clinical implication is still too far, they display the unique feature to inhibit both HMG-CoAR and PCSK9 targets, making them lead compounds for developing new peptidomimetic and/or small molecules endorsed by improved activity on both targets involved in the control of the circulating cholesterol level.

Further experiments were performed to verify the ability of these P5 analogs to modulate the LDLR pathway in HepG2 cells. Similar to P5, P5-Best and P5-H6A induced an upregulation of the SREBP-2 transcription factor level up to 118.6 ± 17.7% and 115.6 ± 10.1% (**Figure 27B**), respectively, resulting in an augmentation of the LDLR protein levels up to 148.4 ± 23.4% and 143.5 ± 24.0%, respectively (**Figure 27C**). Interestingly, although P5-S7A caused a slight reduction of the SREBP-2 protein level to 96.9 ± 16.1% (**Figure 27B**), it led to an increase in the LDLR protein level up to 126.5 ± 13.6% (**Figure 27C**).

Thus, in contrast to P5, P5-Best, and P5-H6A, the upregulation of the LDLR protein level and activity induced by P5-S7A was not through SREBP-2 pathway activation.

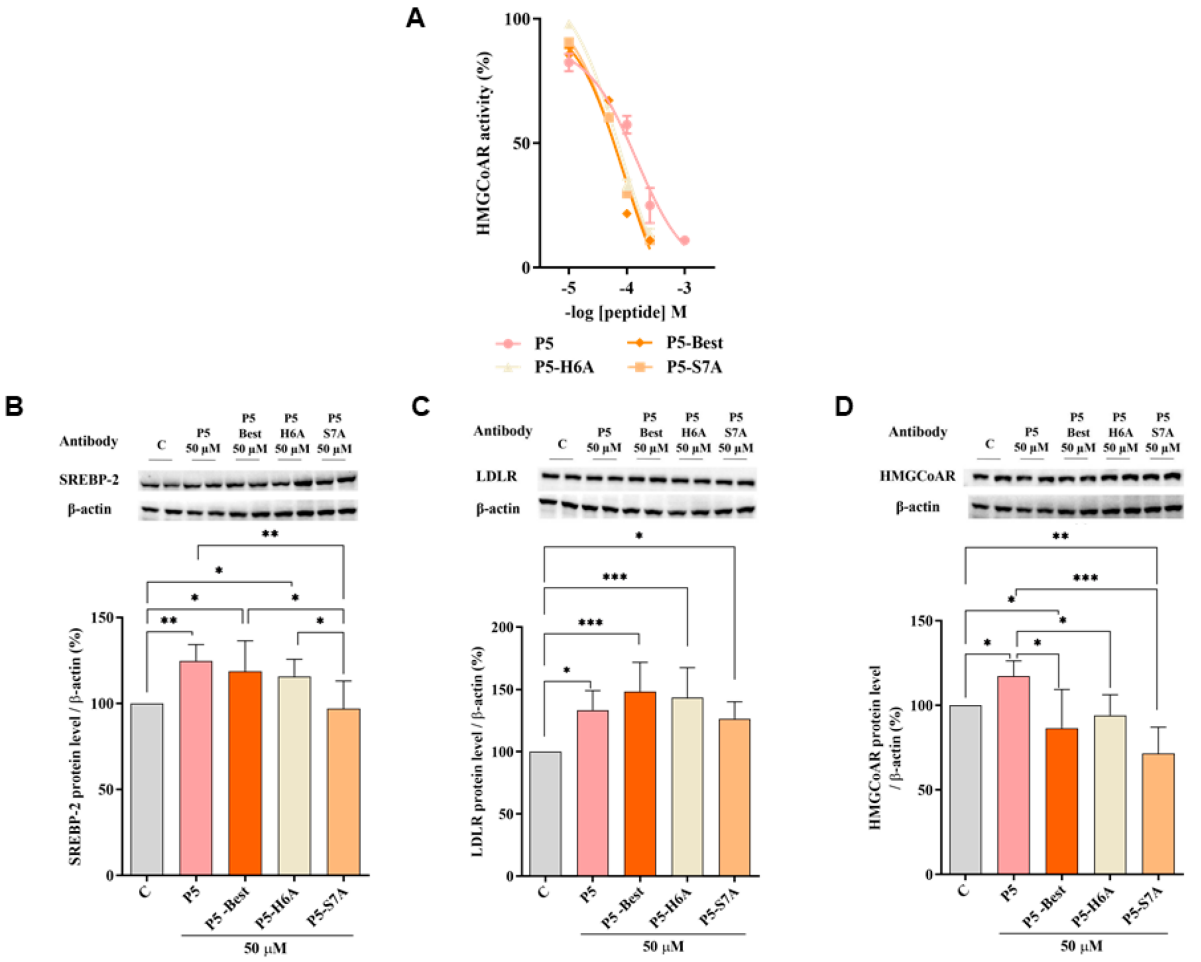


Figure 27. Modulation of the LDLR pathway in HepG2 cells treated with P5 and P5 analogs. (A) *In vitro* inhibition of HMG-CoAR activity. (B) The effect on SREBP-2 protein levels, (C) LDLR protein levels, and (D) HMG-CoAR protein levels after the treatment of HepG2 cells with P5 and P5 analogs, respectively. Data points represent the average ± SD of four independent experiments performed in duplicate.

Notably, P3 (YDFYPSSTKDQQS), a peptide from lupin protein that inhibits HMG-CoAR activity, leads to an increase in the LDLR protein levels without SREBP-2 activation but through the compensatory upregulation of the SREBP-1 [150]. Therefore, it was hypothesized that P5-S7A could possess the same effect as P3 on LDLR protein production through the regulation of the SREBP-1 pathway. However, unlike P5, P5-Best, P5-H6A, and P5-S7A decreased the HMG-CoAR levels up to 86.5 ± 22.8%, 94.1 ± 12.2%, and 71.6 ± 15.5%, respectively (**Figure 27D**), indicating that the P5 analogs were more active as HMG-CoAR inhibitors than P5. Interestingly, P5-Best, P5-H6A, and P5-S7A which are about two-fold more potent than P5 as both HMG-CoAR and PCSK9/LDLR PPI inhibitors, respectively, are also more efficient in the reduction of the HMG-CoAR protein levels with a direct effect in the intracellular cholesterol homeostasis. Indeed, overall P5-Best and P5-S7A can improve the functional ability of hepatic cells to absorb extracellular LDL (**Figure 27D**).

P5 Analogs Modulate the Hepatic PCSK9 Pathway. P5-Best, P5-H6A, and P5-S7A decreased the PCSK9 protein levels by 21.8 ± 11.8%, 28.2 ± 17.5%, and 25.8 ± 17.9%, respectively (**Figure 28A**). Moreover, also the HNF1-α protein levels were decreased by 1.2 ± 15.4%, 10.3 ± 2.4%, and 18.7 ± 7.6% (**Figure 28B**), respectively.

Although the ability to reduce the secretion of mature PCSK9 was weaker than that of P5, P5-Best, P5-H6A, and P5-S7A could also induce a slight reduction by 2.7 ± 1.9%, 7.4 ± 2.6%, and 5.1 ± 1.7%, respectively (**Figure 28C**).

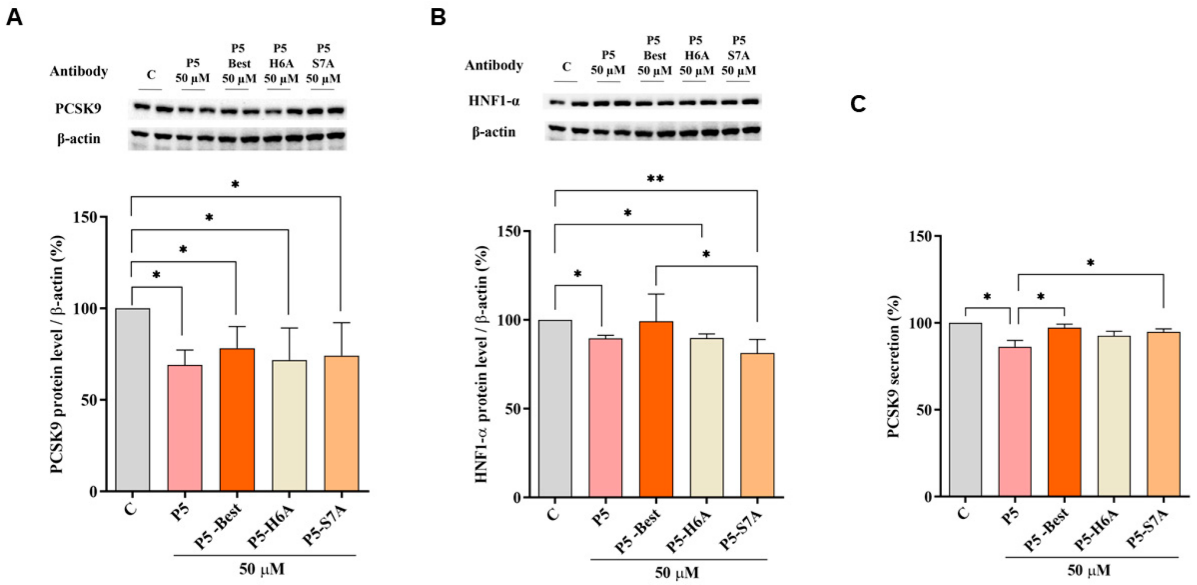


Figure 28. Modulation of the PCSK9 pathway in HepG2 cells. (A) Effects on the PCSK9 protein levels, (B) effects on the HNF1-α protein level, and (C) effects on mature PCSK9 secretion. Data points represent the average ± SD of six independent experiments performed in duplicate.

P5 Analogs Increase the Expression of LDLR Localized in the Cellular Membranes and Modulate LDL Uptake in HepG2 Cells. In accordance with the above results, P5-Best, P5-H6A, and P5-S7A increased the LDLR levels localized in the cellular membranes of HepG2 cells by 156.3 ± 12.1%, 158.9 ± 12.0%, and 140.2 ± 15.1% at 50 µM, respectively (**Figure 29A**). Experiments were also performed in parallel, with P5 as the reference compound, which increased the membrane LDLR protein levels by 153.6 ± 16.4% at the same concentration of 50 µM. Consequently, the functional capability of HepG2 cells to uptake extracellular LDL after treatments with P5, P5-Best, P5-H6A, and P5-S7A was observed, leading to an increased ability of 203.8 ± 40.67%, 254.3 ± 16.4%, 229.8 ± 27.9%, and 211.1 ± 40.1%, respectively (**Figure 29B**).

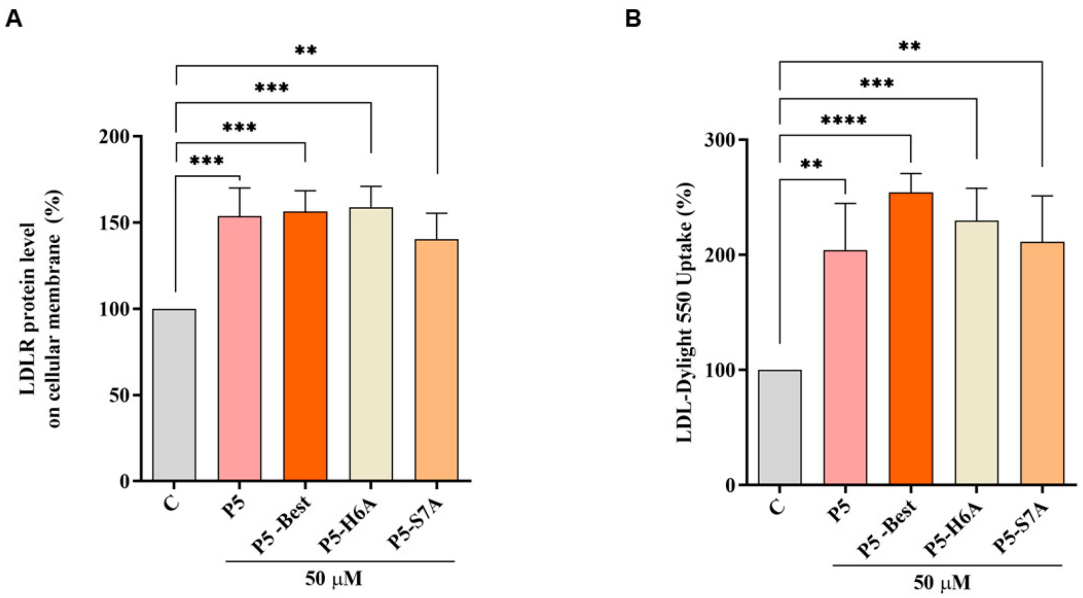


Figure 29. Modulation of the LDLR pathway in HepG2 cells. (A) The effect on the LDLR localized on the surface of HepG2 cells after the treatment of HepG2 cells with P5 and P5 analogs, respectively. (B) Enhancement of the functional ability of HepG2 cells to uptake LDL from the extracellular environment. Data points represent the average ± SD of four independent experiments performed in duplicate.

Docking of P5-S7A and MD Simulations on HMG-CoAR. The experimental assays on the purchased peptides highlighted the improvement in the dual inhibitory activity of the P5 mutant peptides. Specifically, P5-S7A showed the lowest IC_{50} value for HMG-CoAR. Thus, docking and MD simulations were conducted to acquire atomistic details on the putative binding mode of P5-S7A in complex with HMG-CoAR. This study can pave the way for the design of more dual-active peptides. P5-S7A was docked to the statin binding site of HMG-CoAR using the GLIDE software, and the best docking pose (Gscore = −9.88 kcal/mol) was selected for further 500 ns-long MD simulations in explicit water solvent. As the enzyme was in the dimeric state, the statin present in the other binding sites was not deleted to preserve the overall folding of the simulating system. At the end of the MD simulations, RMSD/time plot of the peptide was analyzed, and the peptide conformations sampled during the MD production run were clustered. The results showed that only one cluster was mainly populated, representing 73.1% of the peptide conformations. The structure representative of this cluster is depicted in **Figure 30**.

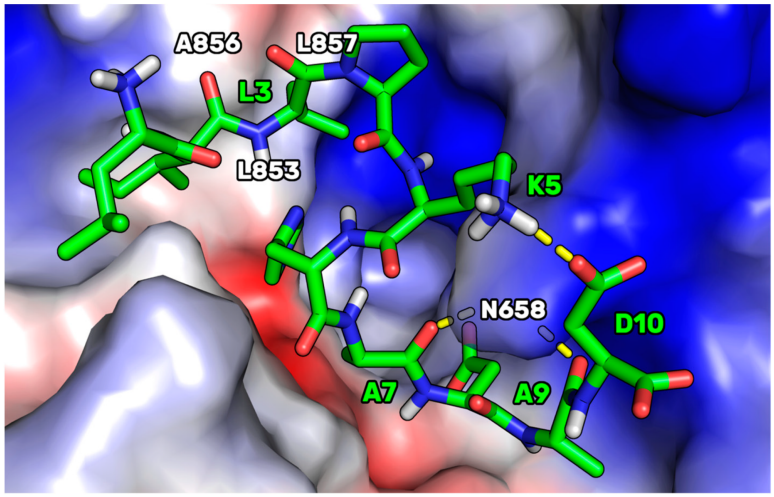


Figure 30. The representative structure of the most populated cluster of P5-S7A (green sticks) bound to HMG-CoAR (electrostatic surface). The small hydrophobic pocket (Leu853, Ala856, and Leu857) interacting with P5-S7A-Leu3 is highlighted. The H-bonds are represented by yellow dashed lines. Only polar hydrogens are shown in the figure.

This HMGCoAR/P5-S7A complex showed the presence of a H-bond network between the P5-S7A-Ala7 and -Ala9 backbone atoms, with a side chain of HMG-CoAR-Asn658. P5-S7A-Leu3 projected its side chain in a small hydrophobic pocket sized by HMGCoAR-Leu853, -Ala856, and -Leu857. Interestingly, the presence of an intramolecular H-bond between the side chains of P5-S7A-Lys5 and -Asp10 improved the overall peptide conformational stability. Moreover, the supposed binding mode of P5-S7A was consistent with the binding affinity data, indicating that the IC_{50} of P5-H6A on HMG-CoAR was close to that of P5-S7A. Both residues could point their side chains to an effectively empty pocket sized by HMG-CoAR-Leu853, -Ala856, and -Leu857, which did not create any interactions with the HMG-CoAR counterpart. Thus, the substitution of positions 6 and 7 with alanine did not elicit any strong variation in the experimental binding affinity. This hypothesis paves the way for the design of new P5 analogs in which positions 6 and 7 can be mutated by unnatural amino acids capable of creating stronger interactions with HMG-CoAR.

The binding mode supposed for P5-S7A was then compared to the one of P5 in complex with HMG-CoAR, to understand the possible reasons on the base of the improved binding affinity displayed by the mutant peptide. In our previous article [141] we have reported on the results of docking calculations of P5. Here, performing MD simulations starting from the P5 docking pose, we noted that, in the complex conformation representative of the most populated cluster (70%), P5 adopted a binding mode in which the side chain of P5-S7 created two intramolecular H-bonds with the NH groups of P5-A9 and P5-D10 (**Figure 31**).

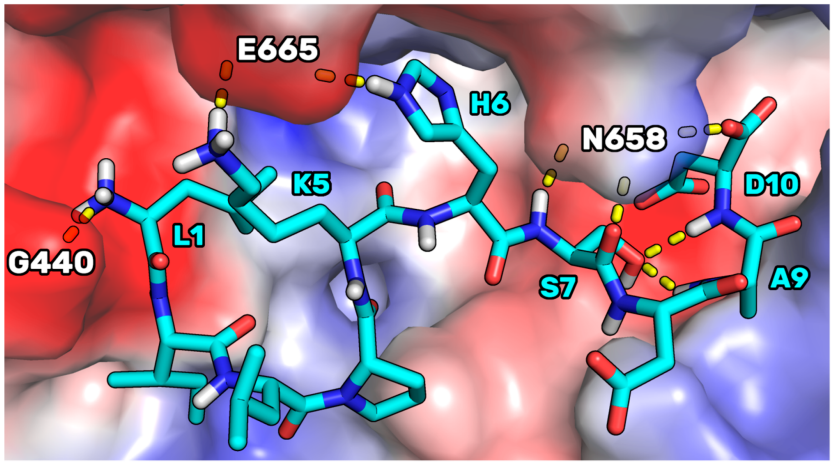


Figure 31. The representative structure of the most populated cluster of P5 (cyan sticks) bound to HMG-CoAR (electrostatic surface). H-bonds are represented by yellow dashed lines.

In the mutant peptide P5-S7A, these internal bonds cannot be created for the absence of the OH group in position 7. This, in our opinion, led to a peptide endowed with an increased conformational freedom, leaving the C-terminal residues to adopt a cyclic conformation in which an internal salt bridge can be shaped between the side chains of P5-K5 and P5-D10. This conformation could be more prone to create remodeled and ameliorated interactions with the enzyme.

Finally, the binding mode supposed for P5-S7A was also compared to that of atorvastatin in complex with HMG-CoAR (as reported in the PDB, accession code 1HWK [151]). The structural alignment of both complexes (**Figure 32A**) permitted to us suppose that the first four residues of P5-S7A essentially reproduce the contact played by the three aromatic substituents of the atorvastatin pyrrole ring.

In particular, the aniline is mimicked by the P5-S7A-Ile2 side chain, the P5-S7A-Leu3 was overlapped to the phenyl ring of the statin, and the *p*-F-phenyl ring of atorvastatin was spatially close to the P5-S7A-Pro4 (**Figure 32**). Unfortunately, the remaining moiety of the peptide pointed to an enzyme area different from the one in which the 3,5-dihydroxyl-heptanoic acid moiety was bound in the HMG-CoAR/atorvastatin complex.

This portion is considered essential for the biological activity of the statins and could explain the reason on the base of the low affinity displayed by the mutant peptide. More efforts should be made to design peptides capable of mimicking such interactions and occupying the HMG-CoAR pocket sized by Lys735, Ser684, Arg590, Lys 691, Asn755, and Glu559 residues (**Figure 32B**).

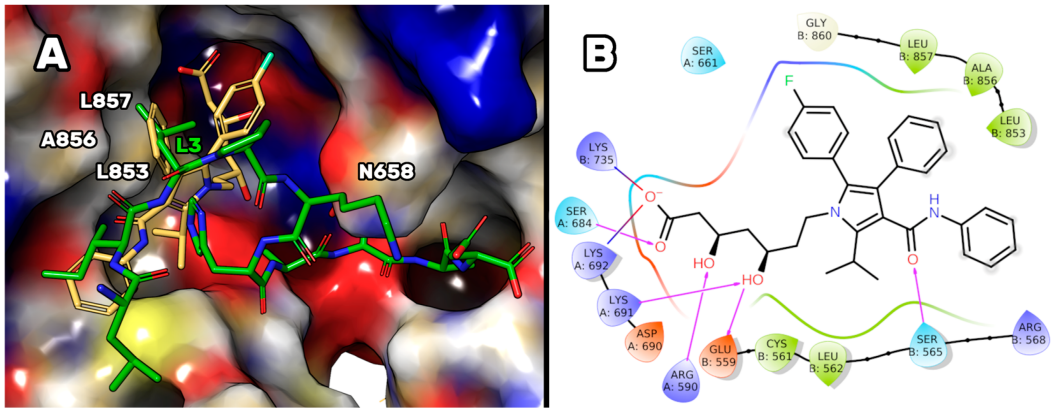


Figure 32. (A) Crystallographic pose of atorvastatin (yellow sticks), as was found in the X-ray structure available in the PDB (accession code 1HWK [151]), superimposed on the representative structure of the most populated cluster of P5-S7A (green sticks) in complex with HMG-CoAR. (B) 2D representation of the protein-ligand contacts displayed by atorvastatin (cut-off = 3.00 Å) in complex with HMG-CoA, as in the X-ray structure. H-bonds are highlighted by purple arrows and the hydrophobic residues involved in the protein-ligand interactions are colored in green.

4.2.2. Optimization of the β-sheet peptidomimetic Rim13

In these years, considerable resources have been dedicated by academia and pharmaceutical companies to the identification of compounds capable of inhibiting PCSK9. Few years ago, the release on the market of two monoclonal antibodies (mAbs), Evolocumab (Repatha® by Amgen) [135], and Alirocumab (Praulent® by Sanofi) [136], proved that the PCSK9 inhibition is a successful therapeutic approach for the treatment of statin-resistant hypercholesterolemia. Additionally, Novartis developed the first siRNA drug (Inclisiran) [152] capable of interrupting the liver transcription of PCSK9, leading to a persistent hypocholesterolemic effects on the treated patients. Nevertheless, these drugs are expensive and do not certainly meet the patient compliance since they are subcutaneously administered.

For these reasons, both pharmaceutical companies and academia are greatly interested in the clinical development of orally bioavailable small molecules, as it was demonstrated by the high number of patents applications in this field [137]. Among the most known PCSK9-LDLR interaction inhibitors, peptides have also their high value since numerous research studies have been reported in literature [153,154]. In fact, peptides, or peptidomimetics, constitutes useful starting point for the identification of new drugs [155]. On the other hand, numerous small molecules have been reported in literature, for example Cpd13 [156], CB36 [157], 3f [158], and Rim13 [143], or in patents (**Figure 33**). Remarkably, some of them are in advanced clinical stages.

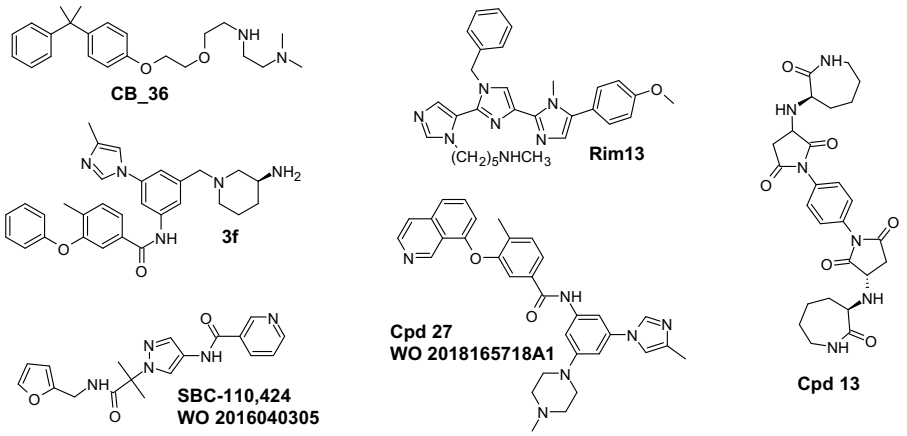


Figure 33. Structure of selected PCSK9 inhibitors reported in literature and patents.

In our previous studies, taking inspiration from the β-strand of the LDLR EGF-A domain in complex with PCSK9 in the X-ray structure [159], we supposed that minimalist peptidomimetic poly-imidazoles could represent promising PCSK9-LDLR interaction inhibitors [160]. As a prove of this, the simplest tetra-imidazole MeIm displayed IC_{50} value in the low micromolar range [160]. Then, optimizing the substitution pattern of the imidazole rings by computational studies, a tri-imidazole derivative (Rim13, **Figure 33**), displayed PCSK9 IC_{50} value close to 1 µM [143].

In this attempt, the MeIm poly-imidazole structure has been further refined by designing novel di-imidazole derivatives considering the high synthetic feasibility and a higher affinity expected on PCSK9. Indeed, by applying computational techniques, new PCSK9 inhibitors were designed, and a selected library of compounds was synthesized. Then, their biological activity was fully investigated by performing assays ranging from cell viability tests to the study of the modulation of the cholesterol pathway on HepG2 cells, highly influenced by the dual inhibitory activity of some compounds. Finally, the pharmacokinetic properties of the most promising compounds were also determined, and their antiplatelet activity was also investigated.

Computational studies. Based on the PCSK9 computational model previously developed [143], new poly-imidazole analogs were designed estimating their binding free energy after docking calculations, pose selection by metadynamics simulations (to improve the accuracy of the binding pose selection), MD simulations, and MM-GBSA calculations. In particular, all compounds were docked in the PCSK9 area depicted by the presence of EGF-A in the X-ray

crystal structure. Then, the most probable docking poses attained by GLIDE software were additionally investigated by BPMD simulations, permitting to choose the most accurate binding pose (widely described in the Chapter 2.2). Only the data of the 6 molecules synthesized is shown in the **Figure 34**. As it is observable, the BPMD method allowed us to discriminate the most plausible pose (endowed with the lowest *PoseScore*) between the two best docking poses of each molecule (as it can be clearly seen in the case of CS179 and CS188), significantly reducing the MD simulations runs to be performed.

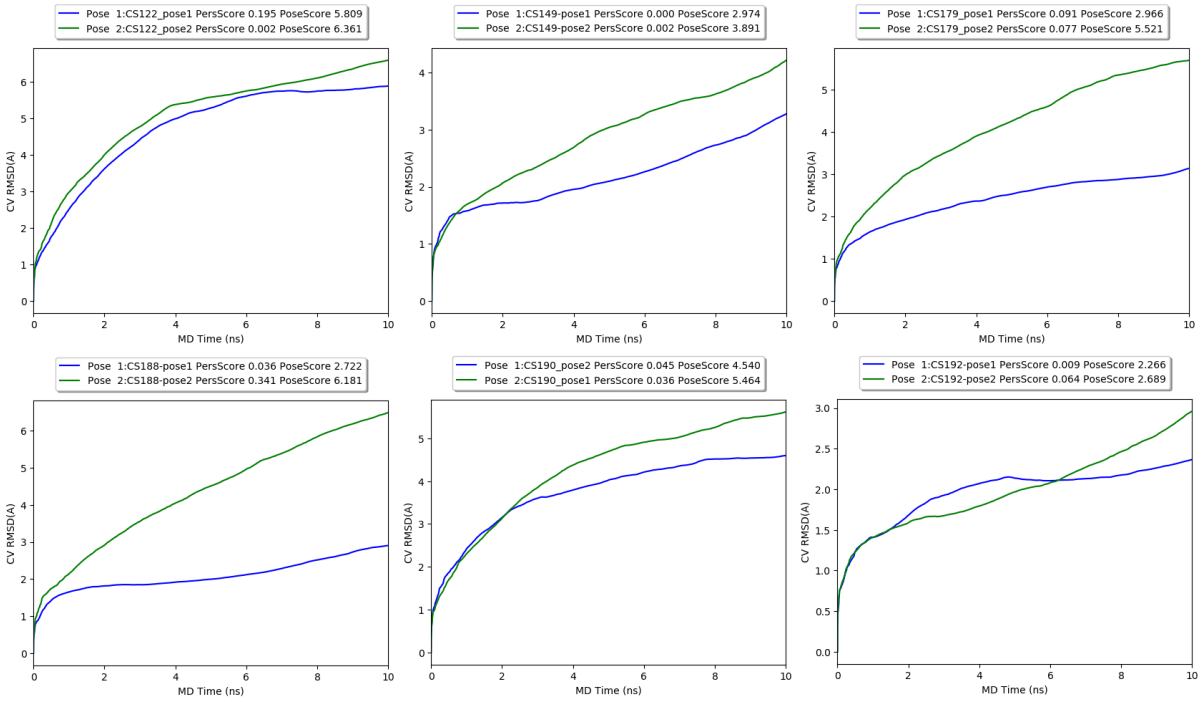
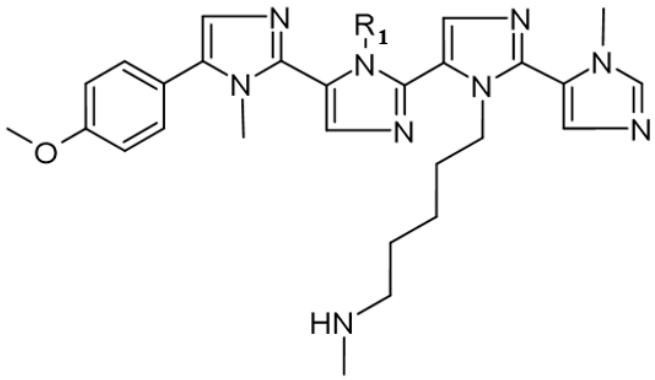


Figure 34. Binding Pose Metadynamics (BPMD) simulations of the two best docking poses of the 6 molecules synthesized and assayed.

Consequently, only the best ligand binding pose resulting from BPMD simulations, for each designed compound, were chosen to build the final PCSK9/ligand complexes, then optimized by MD simulations. The attained trajectory frames were deeply examined by visual inspection and by plotting the ligand not-hydrogen atom RMSD *vs* the simulations time. Subsequently, the frames corresponding to 50 ns of MD simulation length, in which the ligands displayed the lowest conformational freedom in the bonding site, were exploited for the estimation of the ligand binding free energy values (ΔG), by the application of the MM-GBSA approach. Finally, a selected list of compounds endowed with the lowest ΔG values, together with the best synthetic feasibility and the lowest cost of the reactants, were selected for the synthesis and the further biological assays.

Design of new poli-imidazole analogs. In the previous paper of Lammi *et al.* [143], we have scored 13 compounds, aiming at optimizing the substituents capable of interacting with the negative charged areas shaped by the PCSK9 residue Asp367. In this attempt, starting from the general tetra-imidazolyl structure of MeIm and aiming at refining the substituent capable of occupying the PCSK9 hydrophobic pocket shaped by to Ile369, Pro155, Ala239, Phe379, and Ala371, 13 new poli-imidazoles were designed (**Table 13**). Then, calculating their ΔG values, the attained results suggested that compound Tetra10, bearing the -$(CH_2)_3$ *t*-Bu group as R_1 displayed the highest estimated affinity on PCSK9.

Table 13. Chemical structure and ΔG estimation of the new series of compound derived from Rim13 β-sheet peptidomimetic.



Compound	R_1	ΔG ± SE [1]
Tetra1	-$CH_2C_5H_9$	-30.4 ± 0.4
Tetra2	-$(CH_2)_4CH_3$	-24.4 ± 0.4
Tetra3	-$CH_2C_4H_7$	-25.5 ± 0.4
Tetra4	-$CH_2C_3H_5$	-26.0 ± 0.3
Tetra5	-$(CH_2)_3CH_3$	-26.1 ± 0.4
Tetra6	-CH_2-$CH(CH_3)_2$	-22.7 ± 0.5
Tetra7	-CH_2-CH(Et)(CH_3)	-23.2 ± 0.4
Tetra8	-CH_2(*t*-Bu)	-24.1 ± 0.5
Tetra9	-$(CH_2)_2$ *t*-Bu	-26.5 ± 0.4
Tetra10	-$(CH_2)_3$ *t*-Bu	-36.6 ± 0.6
Tetra11	-$(CH_2)_3CH(CH_3)_2$	-27.4 ± 0.5
Tetra12	-$(CH_2)_2$Cy	-30.2 ± 0.4
Tetra13	-CH_2Cy	-30.8 ± 0.5

[1] (kcal/mol)

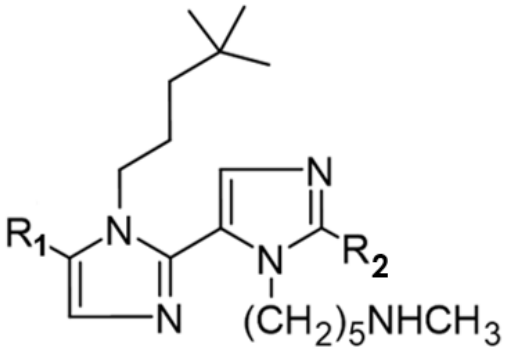
Furthermore, aiming at simplifying the chemical structure of the compounds, improving the synthetic feasibility of the compounds as well, we tried to fuse the benzene and the first imidazole ring into a naphthalene ring capable of mimicking the π electrons conjugation between both rings. The resulting compounds (Dim1, **Table 14**) displayed an increased

predicted binding affinity on PCSK9, since the calculated ΔG was about 5 kcal/mol lower than Tetra10 although, unfortunately, the ligand unbound from the enzyme surface within the initial 100 ns of MD simulations. Thus, to overcome this issue and to evaluate the influence of the third imidazole ring on the predicted ΔG of the compounds, we additionally simplified the chemical structure of Dim1 by displacing the R_2 group by a H atom (Dim2) and by electron rich groups among the classical or non-classical bioisosteres [161,162] of the imidazole ring.

In particular, the presence of alkynes, alkenes, trifluoromethyl, or halogens (Dim3-Dim20, **Table 14**) in the chemical structure of the compounds was investigated. Interestingly, Dim2 was stably bound on PCSK9 over 200 ns of MD simulations and displayed a ΔG value of -35.9 kcal/mol, a value similar to the one of Tetra10 (-36.6 kcal/mol), indicating that the structural simplification did not greatly impact on the binding affinity of the resulting compound.

In the case of the alkynyl series of compounds (Dim3-Dim8), the calculated ΔG values suggested that the applied change was nearly successful, since a gain in the ΔG close to 2 kcal/mol was attained for Dim3, in the respect of Dim2. Compounds Dim4-Dim8 were designed aiming at additional obtaining an advantage in the predicted affinity by decorating the naphthalene ring. Among them, Dim8, bearing the 6-Br-Napht-2-yl as R_1, displayed the lowest predicted ΔG value. Then, compounds Dim9-Dim12 were designed to prove the effect of the R_2 moiety on Dim8, but the attained results suggested that the R_2 substituent must not be greater than the ethynyl.

Table 14. Chemical structure and ΔG estimation of the new series of Dim peptidomimetics derived from Tetra10.



Compound	R_1	R_2	ΔG ± SE [1]
Tetra10	*for comparison*	*for comparison*	-36.6 ± 0.6
Dim1	Napht-2-yl	-*N*[1]-Me-imidazol-5-yl	-41.0 ± 0.3
Dim2	Napht-2-yl	-H	-35.9 ± 0.4
Dim3	Napht-2-yl	-C≡C-H	-38.0 ± 0.4
Dim4	6-Me-Napht-2-yl	-C≡C-H	-40.8 ± 0.5
Dim5	1-OH,6-Me-Napht-2-yl	-C≡C-H	-39.4 ± 0.3
Dim6	6-Me,8-OH-Napht-2-yl	-C≡C-H	-38.6 ± 0.5
Dim7	6,8-diMe-Napht-2-yl	-C≡C-H	-40.0 ± 0.4
Dim8	6-Br-Napht-2-yl	-C≡C-H	-41.9 ± 0.4
Dim9	6-Br-Napht-2-yl	-C≡C-CH_3	-42.6 ± 0.5
Dim10	6-Br-Napht-2-yl	-C≡C-CH_2CH_3	-39.2 ± 0.3*
Dim11	6-Br-Napht-2-yl	-C≡C-$CH(CH_3)_2$	-40.6 ± 0.4
Dim12	6-Br-Napht-2-yl	-*t*(CH=CH)-CH_3	-38.6 ± 0.5
Dim13	6-Br-Napht-2-yl	-CF_3	-38.7 ± 0.3
Dim14	6-Br-Napht-2-yl	-Cl	-38.8 ± 0.3
Dim15	6-Br-Napht-2-yl	-I	-41.0 ± 0.4
Dim16	Napht-2-yl	-I	-39.6 ± 0.2
Dim17	Ph	-I	-34.6 ± 0.4
Dim18	-CH_3	-I	-36.5 ± 0.4
Dim19	-CH_2CH_3	-I	-27.9± 0.5
Dim20	-$(CH_2)_2CH_3$	-I	-31.1± 0.5

* unbound within 150 ns.

In fact, although Dim9 showed the lowest predicted ΔG value it also showed a high ligand RMSD fluctuation along the MD simulation time (**Figure 35A**). Similarly, Dim10, bearing a -C_2Et as R_2, unbound from the PCSK9 surface within the initial 150 ns of MD simulations. Conversely, alkenes Dim11 and Dim12 displayed a high stability on PCSK9 binding site although their ΔG values were not lower than the one of Dim8.

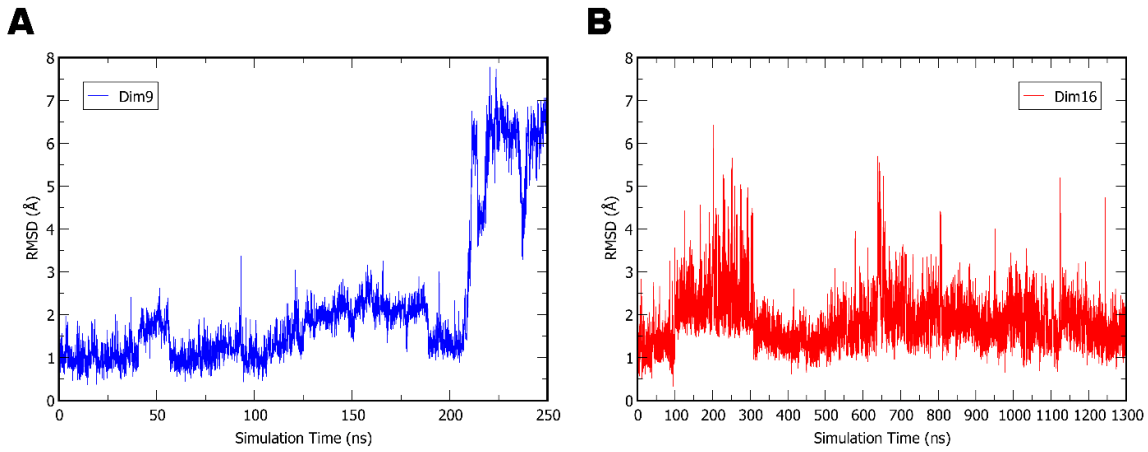
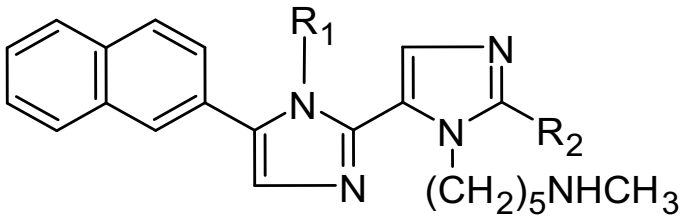


Figure 35. RMSD vs simulation time plot of Dim9 (A) and Dim16 (B).

Compounds Dim13-Dim15 were designed to test the effect of the presence of halogens as R_2 group on Dim8. Remarkably, Dim15 displayed a ΔG value very close to the one of Dim8, together with a great stability on the PCSK9 surface. Finally, compounds Dim16-Dim20 were designed to investigate the effect of the presence of the 6Br-napht-2yl on Dim15. The attained results suggested that removal of the bromine atom, as in the compound Dim16, was not extremely detrimental, since a ΔG value similar to the one of Dim15 was attained. For this compound, MD simulations were extended to 1300 ns, to better sample the conformational space of the complex (the backbone RMSD *vs.* simulation time plot is reported in **Figure 35B**), and the attained results confirmed the high stability of the compound on the PCSK9 surface (average RMSD = 1.84 Å, standard deviation = 0.62). Conversely, the ΔG values calculated for compounds Dim17-Dim20 suggested that a benzene ring or linear alkyl chains as R_1 in this series of compounds did not lead to compounds more promising than Dim15 or Dim16, although they retained a residual predicted affinity on PCSK9 (**Table 14**).

Compounds selection for synthesis and biologically evaluation. Considering the results on **Table 14**, compounds Dim8 and Dim15 could be considered the most promising ones, since they displayed the lowest predicted ΔG values. Nevertheless, pondering on the synthetic feasibility and the cost of the reactants, Dim3 and Dim16 (do not containing the bromine) were chosen for the synthesis and the biological evaluation. In addition to them, to experimentally prove the effect of the R_2 substituent on the biological activity of the compounds, Dim2, the simplest derivative containing the H atom as R_2, was also selected for the synthesis. Moreover, since the -CH_2-Cy resulted second ranked as R_2 moiety in the Tetra series (**Table 13**), we designed the di-imidazoles Dim21, Dim22, and Dim23 (**Table 15**), in which the -CH_2-Cy replaces the -$(CH_2)_3$ *t*Bu moiety of compounds Dim2, Dim3 and Dim16. These compounds, which are not more promising than Dim15, considering their predicted ΔG values (**Table 14**), were synthesized as negative control for the validation of the applied computational design protocol.

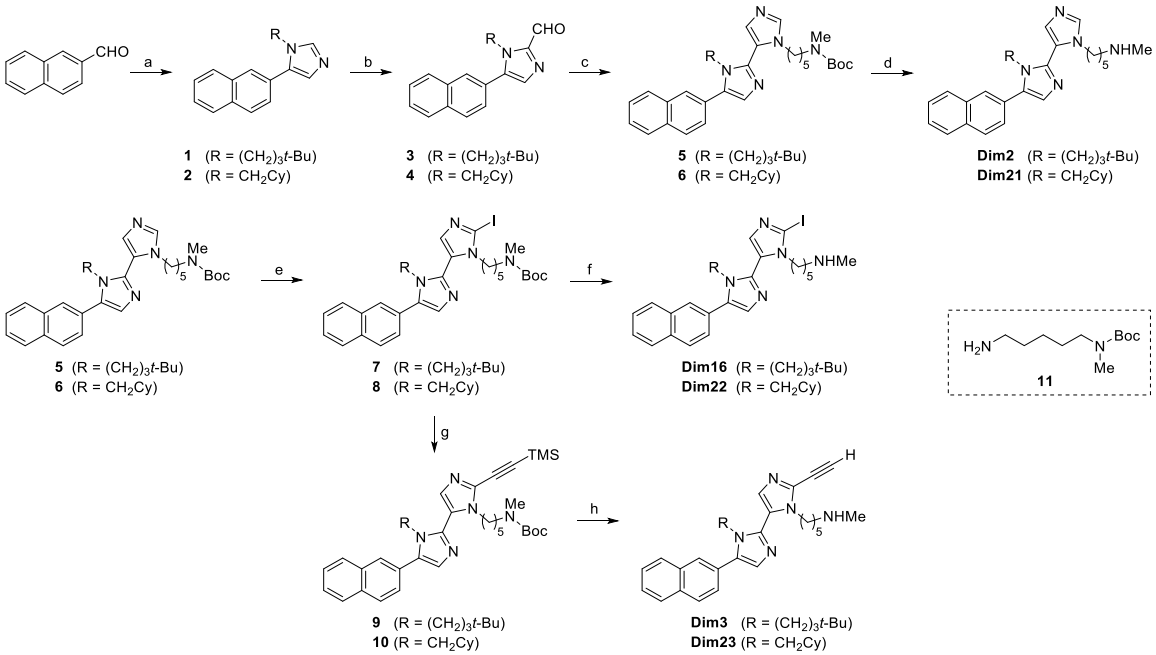
Table 15. Chemical structure and predicted ΔG values of the compounds chosen for synthesis and biological evaluation.



Compound	R_1	R_2	ΔG ± SE (kcal/mol)
Dim2	-$(CH_2)_3$ *t*-Bu	-H	-35.9 ± 0.4
Dim3	-$(CH_2)_3$ *t*-Bu	-C≡C-H	-38.0 ± 0.4
Dim16	-$(CH_2)_3$ *t*-Bu	-I	-39.6 ± 0.2
Dim21	-CH_2Cy	-H	-29.8 ± 0.4
Dim22	-CH_2Cy	-I	-32.6 ± 0.5
Dim23	-CH_2Cy	-C≡C-H	-33.8 ± 0.5

Chemical synthesis. The six target compounds were synthesized in the group Prof. A. Silvani (University of Milan). The synthesis consisted in the twice repeated van Leusen three-component reaction (vL-3CRs) as key reaction process (**Scheme 1**).

Scheme 1. Synthesis of target compounds [a].



[a] **Reagents and conditions:** a) amine, DMF, 70 °C, 2h; then TosMIC, K_2CO_3, overnight (95% for **1**, 84% for **2**). b) *n*-BuLi, THF, from -78 °C to -30 °C, 2h; then DMF, rt, overnight (76% for **3**, 77% for **4**). c) Amine **11**, DMF, 70 °C, 2h; then TosMIC, K_2CO_3, overnight (66% for **5**, 83% for **6**). d) 4N HCl in AcOEt, from 0 °C to rt, 2h; then $NaHCO_3/CH_2Cl_2$ (quant. yield for both **Dim2** and **Dim21**). e) *n*-BuLi, THF, from -78 °C to -30 °C, 2h; then I_2, rt, overnight (62% for **7**, 75% for **8**). f) see d) (quant. yield for both **Dim16** and **Dim22**). g) trimethylsilylacetylene, $Pd(PPh_3)_2Cl_2$, CuI, THF/Et_3N, 60 °C, 3h (32% for **9**, 44% for **10**). h) K_2CO_3, MeOH/THF, 2h, rt; then $BF_3{\cdot}OEt_2$, CH_2Cl_2, 4 Å molecular sieves, -40 °C, 2h (quant. yield for both **Dim3** and **Dim23**).

Di-imidazole Analogs Impair the PCSK9-LDLR PPI and the HMG-CoAR activity. To evaluate the inhibitory ability of Dim analogs, dedicated experiments were carried out in the group of Prof. C. Lammi (University of Milan), with the aim at verifying whether they are able to impair the PPI between PCSK9 and LDLR and to drop the HMG-CoAR activity. The attained results indicated that Dim2, Dim3, Dim16, Dim22, and Dim23 reduced the PCSK9-LDLR binding with a dose response trend and an IC_{50} values of 1.99 ± 1.65 µM, 0.009 ± 0.01 µM, 0.0008 ± 0.001 µM, 1.99 ± 2.86 µM and 1.18 ± 1.06 µM, respectively (**Table 16**). Results indicated that Dim3 and Dim16 are more active than the other analogs (**Figure 36A**). Additionally, biochemical investigation was carried out for assessing the ability of Di-imidazole analogs to modulate the *in vitro* HMGCoAR activity. Results suggested that Dim2, Dim16, Dim21 and Dim22 inhibited the enzyme with a dose-response trend and an IC_{50} of 40.48 ± 15.24 µM, 146.8 ± 75.09 µM, 38.4 ± 12.71 µM, and 36.21 ± 5.98 µM, respectively. Specially, Dim2, Dim21, and Dim22 displayed activity in the micromolar range (**Figure 36B**), whereas Dim3 and Dim23 were not active, as reported in the **Table 16**.

Table 16. Results of the biological experiments.

Compound	R_1	R_2	ΔG ± SE [1]	PCSK9/LDLR binding IC_{50} [2]	HMGCoAR activity IC_{50} [2]
Dim2	-$(CH_2)_3$ *t*-Bu	-H	-35.9 ± 0.4	1.99 ± 1.65	40.48 ± 15.24
Dim3	-$(CH_2)_3$ *t*-Bu	-C≡C-H	-38.0 ± 0.4	0.009 ± 0.01	N/A
Dim16	-$(CH_2)_3$ *t*-Bu	-I	-39.6 ± 0.2	0.0008 ± 0.001	146.8 ± 75.09
Dim21	-CH2Cy	-H	-29.8 ± 0.4	4.50 ± 0.50	38.4 ± 12.71
Dim22	-CH2Cy	-I	-32.6 ± 0.5	1.99 ± 2.86	36.21 ± 5.98
Dim23	-CH2Cy	-C≡C-H	-33.8 ± 0.5	1.18 ± 1.06	N/A

[1] (kcal/mol); [2] (µM).

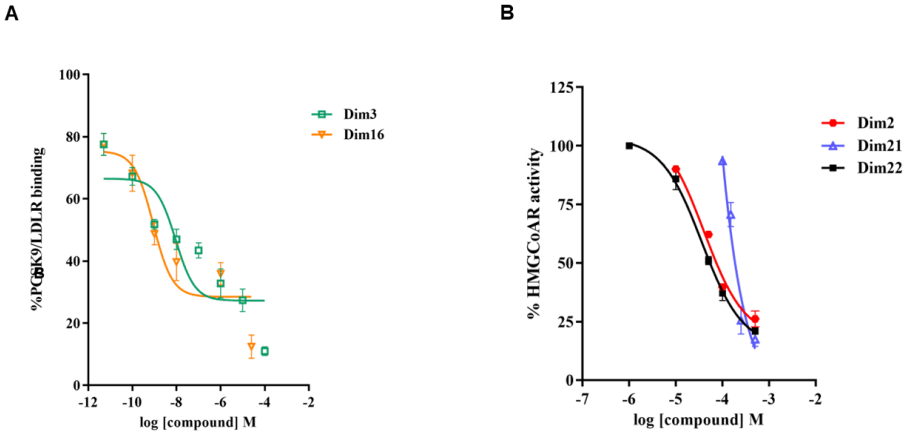


Figure 36. A) Inhibition of the protein–protein interaction between PCSK9 and LDLR. B) Inhibition of HMGCoAR activity. The data points represent the mean ± s.d. of three independent experiments.

Effect of Dim3 and Dim16 on the HepG2 Cell Vitality. Considering that Dim3 and Dim16 were the most active compounds inhibiting the PCSK9 ability to bind the LDLR in the biochemical system and that Dim16 also showed the capability to modulate the HMG-CoAR activity, cell-based experiments were realized with the aim at characterizing the molecular and functional behavior of both Dim analogs, using human hepatic HepG2 cells. Hence, preliminary cellular viability experiments were carried out for excluding any potential effects of the treatment with Dim3 and Dim16 on the HepG2 cell's vitality. After the 48-h treatment, any reduction of hepatic cell vitality was observed up to 10 µM versus control cells, indicating that Dim3 and Dim16 were safe for HepG2 cells in this dose range (data not shown).

Di-imidazole Analogs Increase the Expression of the LDLR Localized on the Cellular Membranes. In addition, the ability of these Di-imidazole analogs to modulate the levels of LDLR localized on HepG2 surfaces was investigated in the presence of PCSK9 (4 µg/mL). Results indicated that LDLR levels decreased in the presence of PCSK9 alone by 39.71 ± 2.05% compared to untreated control cells, whereas Dim3 and Dim16 can significantly restore LDLR levels to 77.87 ± 3.04% and 101.1 ± 15.06% (**Figure 37A**), 91.1 ± 2.22% and 87.17 ± 7.42% when co-incubated with PCSK9 (**Figure 37B**) at 1 nM and 10 nM, respectively.

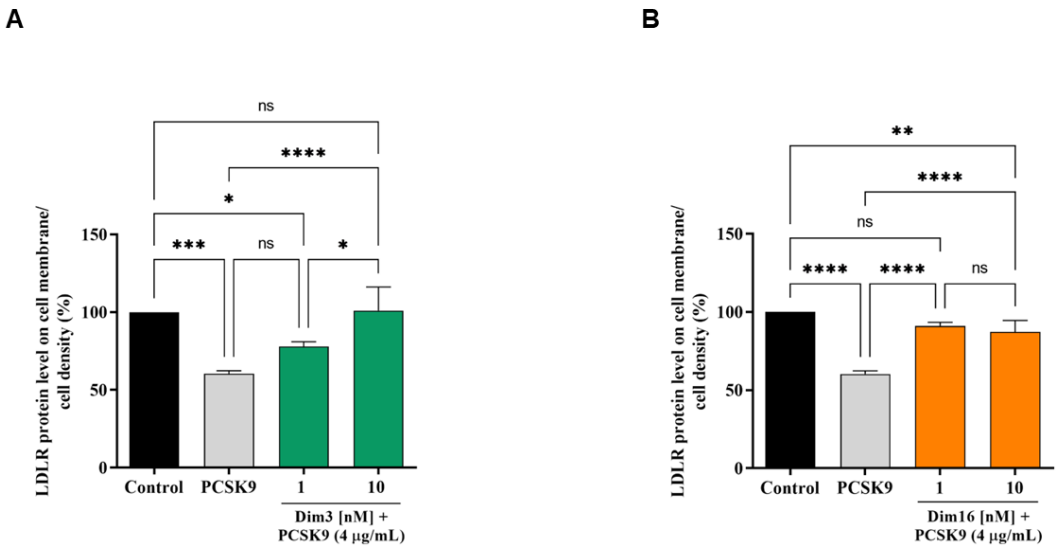


Figure 37. The treatment of HepG2 cells with PCSK9 (4 µg/mL) reduced active LDLR protein levels localized on the surface of cells, which were restored by Dim3 (A) and Dim16 (B), inducing an increase of LDLR protein level on HepG2 cell surface at 1 nM and 10 nM, respectively. The data points represent the mean ± s.d. of three independent experiments.

Di-imidazole Analogs Modulate the LDL-uptake in HepG2 cells. Finally, functional cell assays were performed to verify the capacity of HepG2 cells to uptake extracellular LDL in the presence of PCSK9 (4 µg/mL). HepG2 cells incubated with PCSK9 alone displayed a 51.69 ± 15.30% reduction in the uptake of fluorescent LDL compared to untreated control cells, indicating reduced LDLR function. After co-incubation with PCSK9 at 1 or 10 nM, Dim3 and Dim16 restored LDLR function, increasing LDL uptake up to 94.12 ± 10.95% and 103.47 ± 7.34% (**Figure 38A**), 81.87 ± 7.45% and 136.47 ± 8.81% (**Figure 38B**), respectively.

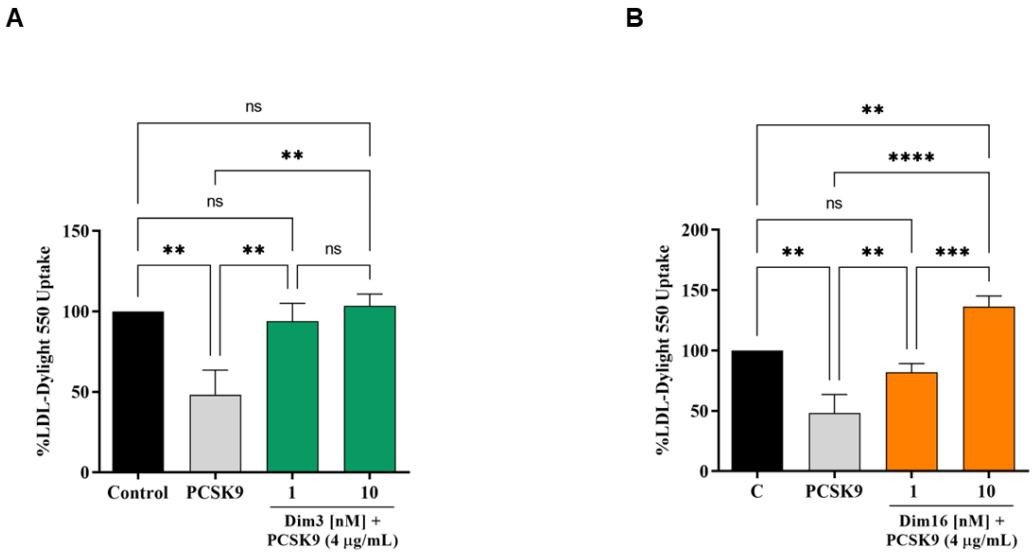


Figure 38. The decreased ability to uptake LDL from the extracellular space by HepG2 cells induced by PCSK9 is prevented by Dim3 (A) and Dim16 (B), inducing an improved ability of HepG2 cells to absorb LDL at 1 nM and 10 nM, respectively. The data points represent the mean ± s.d. of three independent experiments.

To conclude, in this study, starting from our studies on the simplest tetraimidazole (MeIm), we have designed new PCSK9 inhibitors endowed with a di-imidazole scaffold, which had shown the lowest PCSK9 IC_{50} value (0.8 nM) reported in literature by us [143]. Considering the theoretical and the experimental studies on the series of tetra-imidazoles (**Table 13**), tri-imidazoles, di-imidazoles (**Table 14**), we can advance the hypothesis that potent PCSK9 inhibitors, endowed with poli-imidazole general structure reported in **Figure 39B**, must need at least four substituents:

- R_1: a planar aromatic group capable of interacting with the zone of PCSK9 created by the residues close to the disulphide bridge Cys375-Cys378;
- R_2: a branched alkyl chain capable of filling the hydrophobic pocket sized by Ile369, Pro155, Ala239, Phe379, and Ala371 of PCSK9;
- R_3: a basic chain with the optimal length to reach the PCSK9 area close to Asp367;

- R_4: an electron rich group, like a halogen or an alkyne. These features are required to interact with the positively charged amino term of Ser153, resulting from the autocatalytic maturation of the enzyme.

These structural features can be found in the structure of Dim16, whose supposed binding mode on PCSK9 is reported on **Figure 39A**.

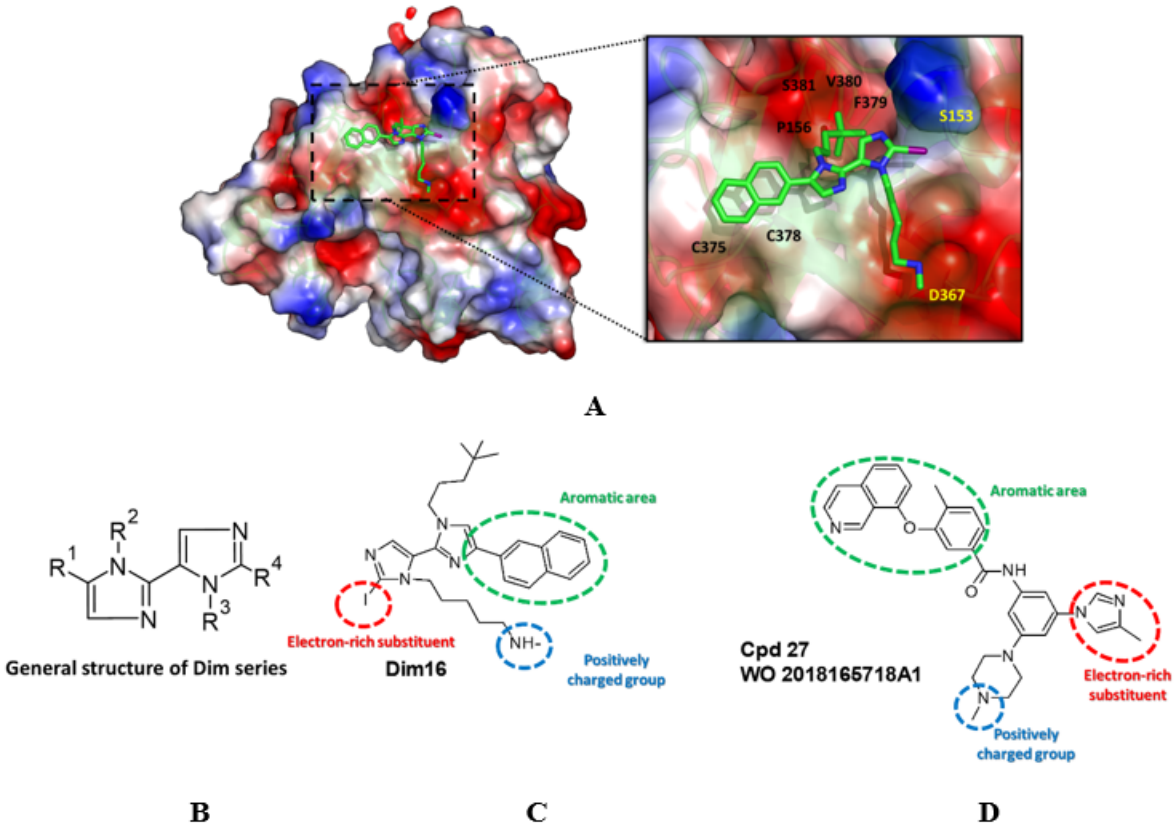


Figure 39. (A) Supposed binding mode of Dim16, as resulted by docking, metadynamics, and MD simulations. PCSK9 is represented by ribbons revealing the secondary structure. The solvent-accessible surface of PCSK9 is colored depending on the partial charge of the atoms: positive areas are depicted as blue, while red areas suggest the presence of positively charged residues. The carbon atoms of Dim16 are represented as green sticks. (B) General structure of the Dim series reported in this paper. (C) Molecular formula of Dim16. (D) molecular formula of Cpd27, as resulted from the Nyrada patent WO2018165718A1. The red, green, and cyan areas depicted in panels C and D highlight the common structural features of Dim16 and Cpd27.

The lack of one of these structural requirements leads to compounds less active on PCSK9. In fact, Dim2, which contains H as R_4, displayed a IC_{50} value 220 times higher than Dim3, holding the ethyne, and about 2500 times higher than Dim16, which contains the iodine as R_4. The importance of the proper alkyl chain as R_2 substituent is demonstrated by the data of Dim21-23. In fact, in all cases high IC_{50} values were obtained when compared to their analogs Dim2, Dim3, and Dim16. The importance of the alkaline chain such as R_3 had been discussed in a

previous paper published by our research group [143] while an aromatic ring as R_1 seems essential to obtain theoretically active compounds, as demonstrated by the ΔG values calculated for compounds Dim17-20.

Interestingly, some of the structural features of Dim16 can be also found in the Cpd27 (**Figure 33**, patent deposited by Nyrada). In fact, comparing the Dim16/Cpd27 chemical structures (**Figure 39CD**) can be easily noted that both contain:

- a planar skeleton bearing some substituents: the di-imidazole scaffold of Dim16 and the carboxamido-phenyl moiety in Cpd27;
- the presence of an area rich in aromatic substituents: the naphthyl scaffold of Dim16 and the isoquinoline moiety of Cpd27 (green area in the **Figure 39CD**);
- electron-rich substituents: the iodine atom of Dim16 (or the ethyne of Dim2) and the 3-methyl-imidazole of Cpd27 (red area in the **Figure 39CD**);
- a positively charged moiety: the *N*-methylpentan-1-amine of the Dim series and the N_4-methyl-piperazine of Cpd27 (blue area in the **Figure 39CD**), both protonated at physiologic pH.

Finally, we can suppose that Dim16, bearing also an additional branched alkyl chain as R_2, may have all the structural features that justify the low-nanomolar affinity (IC_{50} = 0.8 nM, **Table 16**).

Docking and MD simulations on HMG-CoAR. In order to predict the binding mode of Dim2, Dim3, and Dim16 on HMG-CoAR, rationalizing their structure-activity relationships (**Table 16**), docking calculations and MD simulations were performed (**Figure 40A**). The best docking poses of the compounds explained how the substitution, in the R_2 position, by hydrogen (Dim2) or iodine (Dim16) atoms to a huge bulky group, such as the alkyne of Dim3 (**Figure 40A**), strongly influence the predicted binding mode of the compounds. In fact, the -H and -I groups of Dim2 and Dim16, respectively, are positioned in a small hydrophobic pocket and their substitution inevitably causes steric hindrance.

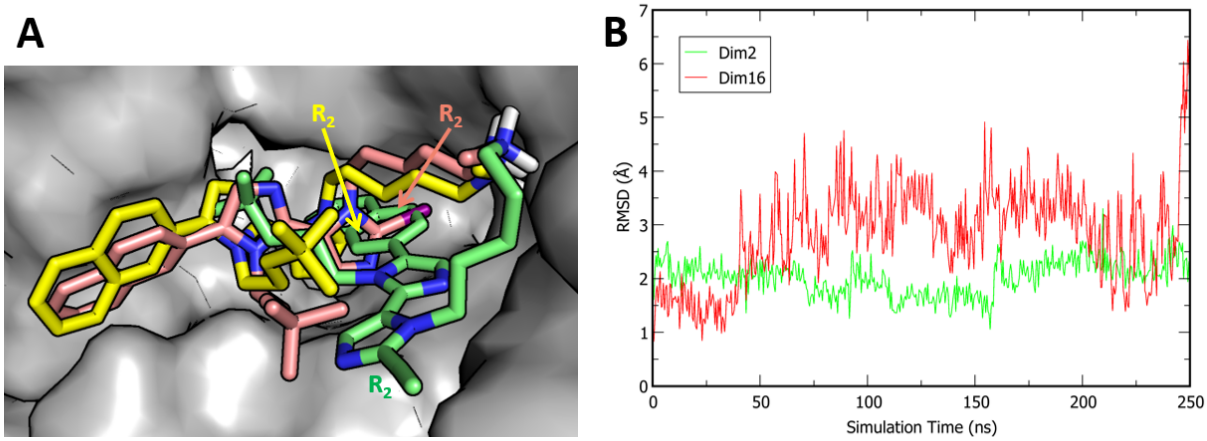


Figure 40. (A) The best docking pose of Dim2 (yellow sticks), Dim16 (salmon sticks) and Dim3 (green sticks) compounds. The R_2 position for each compound is highlighted by the arrow. (B) Cα-RMSD plot of Dim2 (green) and Dim16 (red) during a 250 ns-long MD simulation.

The predicted binding modes were further inspected by 250 ns-long MD simulations, and the binding free energy (ΔG) were estimated only for Dim2 and Dim16. In fact, Dim3 unbound from the active site of the target after few nanoseconds of MD simulations (data not shown). At variance, both Dim2 and Dim16 remained well anchored on the HMG-CoAR binding site for the whole simulation length, showing an average Cα RMSD value of 2.0 Å (std. dev. = 0.3 Å) and 2.8 Å (std. dev. = 0.9 Å), respectively (**Figure 40B**). Dim2 and Dim16 displayed a ΔG value of -42.6 ± 0.5 kcal/mol and -37.9 ± 0.6 kcal/mol, respectively, further confirming the experimental data obtained.

4.3. GABARAP

Autophagy plays a fundamental role in cellular, tissue, and cellular homeostasis. Metabolic stress (induced by starvation or hypoxia) or the presence of dangerous cellular components, including dysfunctional organelles, intracellular microbes, and pathogenic proteins, can activate this pathway. Briefly, a multistep process, starting with the assembly of the phagophore, mediates the sequestration of organelles and proteins into the autophagosome. Its subsequent fusion with a lysosome leads to the formation of the autolysosome, in which the autophagosome content is degraded by lysosomal hydrolases [163].

More than 50 proteins (called mAtgs) are involved in the mammalian autophagy machinery, but those responsible for the formation of the autophagosome and for cellular trafficking are members of the mAtg8 family. In mammals, the mAtg8 proteins are further divided into two

subfamilies: GABARAP (GABA-A receptor-associated protein) and MAP1LC3 (microtubule-associated protein 1 light chain 3), or simply LC3. The former comprises GABARAP, GARAPL1, and GABARAPL2, while the latter includes LC3A (with the two splicing variants LC3Aα and LC3Aβ), LC3B, LC3B2, and LC3C [164]. Members belonging to the same subfamily share a high-sequence homology and play similar physiological roles at the intracellular level. The GABARAP subfamily seems fundamental for the closure of the autophagosome and for the recruitment of the autophagy players, while LC3 proteins appear to be mainly involved in the cargo recruitment process. A resume scheme of the LC3/GABARAP family proteins roles in autophagy-related processes is represented in **Figure 41** [165].

In detail, the autophagy initiation involves the formation of a phagophore. The nucleation step (**Figure 41A**) starts with the formation of ULK1 heterocomplex, by binding GABARAP/GABARAPL1. The phagophore can swallow the content selectively by the LC3/GABARAP family members which are able to bind specific linker proteins (**Figure 41B**). LC3B is the main family protein involved in the elongation step (**Figure 41C**), which requires the delivery of membrane components (orange) to the growing phagophore (purple). In the phagophore closure step (**Figure 41D**), both LC3B and GABARAPL2 are the protein family members mainly involved to obtain a closed double-membraned vesicle. The subsequent fusion of the autophagosome with a lysosome (**Figure 41E**) is necessary to acquire the degradative enzymes and the lumen acidification needed for the subsequent cargo degradation. This last step requires GABARAP and GABARAPL2. Aside autophagy, GABARAP family proteins are also involved in the receptor trafficking (**Figure 41F**).

Disfunctions in the autophagy machinery have been identified in various pathological conditions, including neurodegenerative diseases, cancer, and inflammation; however, the role of each mAtg8 component in cancer is still undetermined and controversial. In the early stage of tumorigenesis, a high level of autophagy proteins is considered a good prognosis factor, since GABARAP is downregulated in renal and breast cancers, and hepatocarcinoma [164]. Conversely, high levels of GABARAP have been detected in colorectal and thyroid cancers. Moreover, cancer cells use autophagy to survive to several antitumor drugs. Some reports indicate that the efficacy of radio- and chemo-therapy is strongly influenced by the effective modulation of the autophagy process [166,167]. Additionally, it has been demonstrated that the expression level of mAtg8 proteins is strictly related to the tumor development, stage, and type [164].

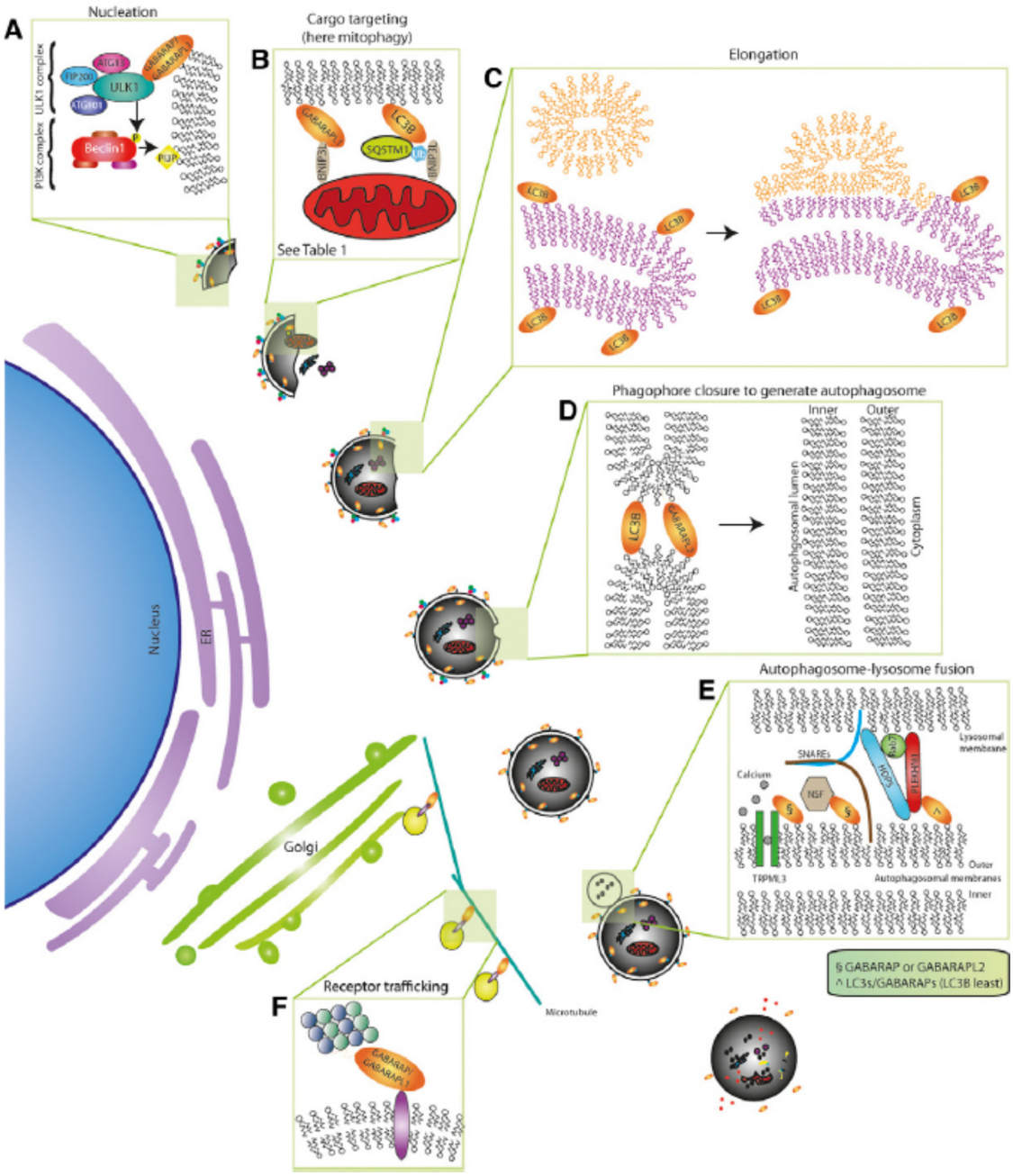


Figure 41. Resume scheme of the LC3/GABARAP family proteins roles in autophagy-related processes [165].

Furthermore, it has been shown that the autophagy machinery is often inefficient in prostate cancer cells, due to a reduction in catabolic pathways [168]. In 2012, He *et al*. [169] suggested that the apoptotic effects of some agents, such as the TNF-related apoptosis-inducing ligand (TRAIL), was enhanced by pharmacologically inhibiting autophagy [170]. Therefore, these processes seem to play a pivotal role in the regulation of the death/survival balance in prostate cancer cells. In 2016, Engedal *et al*. proved that GABARAP-subfamily proteins are strongly involved in autophagy mechanisms in prostate cancer [171], supporting the idea that GABARAP inhibitors could be valuable tools to fight the progression of this disease.

To the best of our knowledge, the known GABARAP inhibitors are essentially proteins or peptides of various sizes. Among them, the small synthetic peptide K1 (DATYTWEHLAWP) is one of the most active candidates, showing a K_d value close to 390 nM (data measured by SPR) [172]. Additionally, an interesting natural GABARAP binder is AnkirinB (AnkB), a 440 kDa neural-specific protein expressed in unmyelinated axons. Similar to other proteins involved in the autophagy machinery, AnkB has an LC3 interacting region (LIR), a small area containing four conserved amino acids. These residues can be briefly represented as a sequence of "X_0–X_1–X_2–X_3", in which X_0 is an aromatic residue (Trp/Phe/Tyr), X_1 and X_2 can be any amino acids (often acidic or hydrophobic residues), and X_3 is a large hydrophobic residue like Leu, Val, or Ile [173]. The main target of AnkB-LIR is GABARAP, since it was demonstrated that the peptide EEWVIVSDEEIEEARQKA binds to it with a K_d value of 0.27 nM. However, despite its potency, AnkB cannot be considered a selective mAtg8-binding peptide, because it interacts with all members of the Atg8 family, displaying a high affinity [174]. The atomic details of the AnkB-LIR/GABARAP interaction were disclosed through X-ray studies by Li *et al.*, who demonstrated that GABARAP inhibitors successfully block autophagy in cultured cells [174].

The results of this investigation paved the way to the design of GABARAP binders as potential tools for the development of anticancer drugs. While it is known that peptides are endowed with poor PK properties (because of their low resistance to intestinal degradation), they can still be valuable tools for the study of the physio-pathological pathway in which their biological counterparts are involved. Moreover, they constitute valuable templates for the design of novel small molecules or peptidomimetics.

4.3.1. Design of novel cyclic peptides (WC8 e WC10)

In this case study, with the aim at identifying new peptides endowed with inhibitory activity against GABARAP, we start from the AnkB-LIR/GABARAP X-ray complex and, by applying a computational approach, we identify new peptides with low micromolar affinity for the target. Experimental assays were carried out to measure the peptide K_d values by MST and SPR, and to evaluate their activity on prostate cancer cells. Remarkably, two of them displayed anticancer effects on PC-3 cells. Considering that prostate cancer is the second most diagnosed malignancy in men worldwide, we are confident that this study will open new avenues to identify the chemical entities endowed with significant therapeutic effects on this widespread disease.

Computational studies. Initially, the GABARAP/AnkB-LIR computational model was generated, starting from the three-dimensional data of the complex, available in the PDB. Computational studies were then accomplished to design new peptides endowed with high affinity for GABARAP. The following procedure was adopted:

1) Identification of the minimal AnkB-LIR sequence (*core*) responsible for the interaction with GABARAP;
2) Mutation of the core sequence aimed at improving the theoretical affinity of the resulting peptides;
3) Rigidification of the most promising peptides by disulphide bonds
4) Optimization of the peptide sequence by the addition of residues potentially occupying supplementary GABARAP binding pockets;
5) Assessment of the binding affinity of the peptides by biophysical experiments;
6) Evaluation of the killing effect on prostate cancer cells, exerted by the most promising candidates.

Setup of the GABARAP computational model and identification of the AnkB-LIR core sequence. Firstly, the GABARAP/AnkB-LIR complex model (**Figure 42A**) was retrieved from the PDB (accession code 5YIR) [174] and then refined by energy minimization and MD simulations. The AnkB-LIR peptide rapidly reached the geometrical stability over the 500 ns-long MD simulations, as demonstrated by the protein Cα RMSD plot (**Figure 42B**).

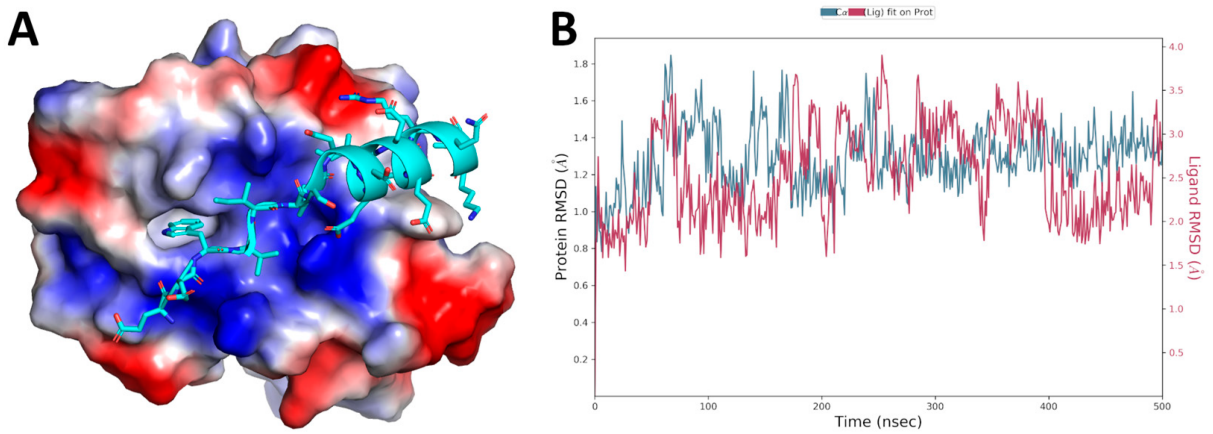


Figure 42. (A) 3D representation of the GABARAP/AnkB-LIR complex, as derived from the X-ray structure (PDB accession code 5YIR). The protein surface is colored depending on the atomic partial charges of the protein residues: blue for positive and red for negative charges. The AnkB-LIR peptide is represented as cyan sticks. (B) Plot of the protein and ligand (AnkB-LIR) Cα atoms RMSD over the simulation time.

As expected, and verified by inspecting the MD trajectory frames, the LIR domain (sequence WVIV of AnkB-LIR) created numerous contacts with GABARAP (**Figure 43**). Moreover, also the SDEE residues were involved in productive contacts, including the electrostatic interactions between the side chains of the peptide glutamates and the positively charged area of GABARAP close to K46 and R47. Conversely, the remaining *C*-terminal residues were mainly involved in internal contacts, stabilizing the α helix shaped by the DEEIEEARQKA sequence.

Then, to exactly define the minimal portion of AnkB-LIR with the highest affinity for GABARAP, the peptides AnkB-LIR and AnkB-core (sequence WVIVSDEE) were subjected to MD simulations and MM-GBSA calculations for the estimation of their binding free energy. Desmond and Prime tools of Maestro were employed for these computations, which predicted ΔG values of -135.1 and -107.9 kcal/mol for AnkB-LIR and AnkB-core, respectively (**Table 17**). These results indicates that the 8 amino acids belonging to the AnkB-core contribute 75% of the overall interaction energy of the full AnkB-LIR peptide (composed of 20 residues). Thus, to further define the contact area, MD simulations and MM-GBSA calculations were performed on the GABARAP/WVIV complex model, in which only the LIR motif (AnkB-wviv peptide) was simulated.

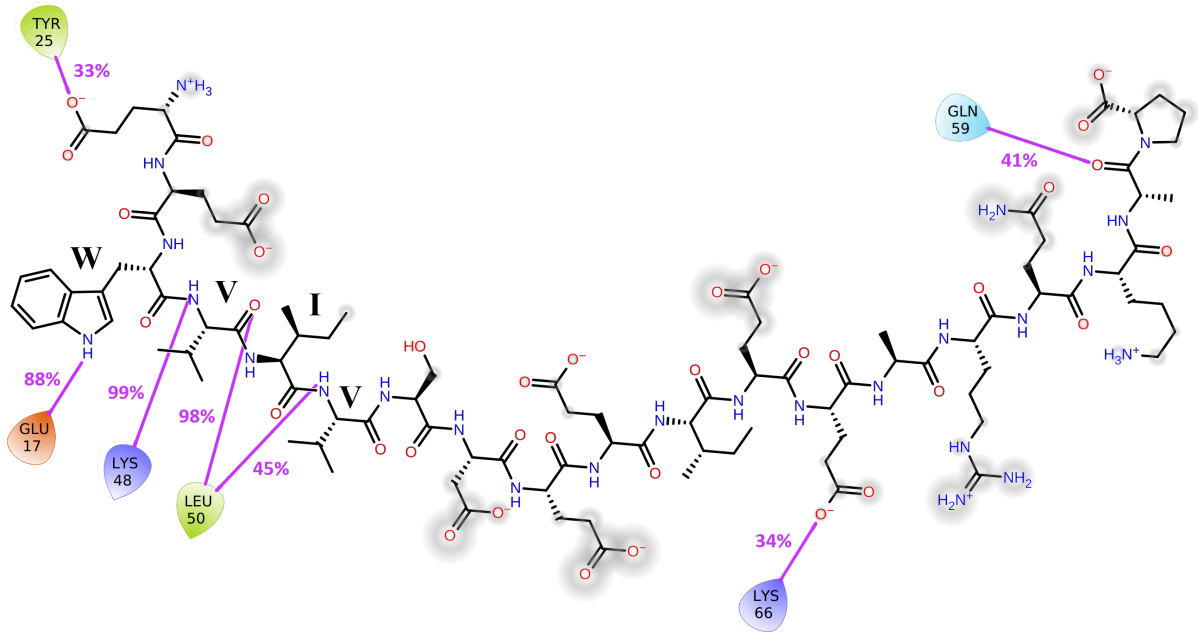


Figure 43. GABARAP H-bond interactions with AnkB-LIR ligand (EEWVIVSDEEIEEARQKA) monitored throughout the simulation. The key residues of AnkB-LIR (WVIV) interacting with GABARAP protein are highlighted. Only the interactions that occur more than 30% of the simulation time are shown.

As reported in **Table 17**, residues WVIV contribute 65% of the overall binding free energy. This outcome confirms that the LIR motif, shared by all proteins involved in the autophagy machinery, displays the highest complementarity with the GABARAP-binding site and is responsible for the most significant protein-protein contacts. Subsequently, the same computational protocol (MD simulations and MM-GBSA calculations) was applied to study the interaction of peptide K1. Considering that its experimentally determined K_d lies in the nanomolar range (390 nM), this peptide could be considered as a reference inhibitor of GABARAP, together with AnkB-LIR.

We also predicted the ΔG of the peptide K1, attaining a value of-118.9 kcal/mol (**Table 17**), a value slightly higher than that of AnkB-LIR (-135.1 kcal/mol). This result is in line with the K_d values reported for the two peptides (0.27 and 390 nM, respectively).

Table 17. Sequence and calculated ΔG values of AnkB analogs and K1 peptide.

Peptide	Sequence	ΔG Prime [1]	SD [2]
AnkB-LIR	EEWVIVSDEEIEEARQKA	−135.1	10.3
AnkB-*core*	WVIVSDEE	−101.4	5.9
AnkB-wviv	WVIV	−88.5	3.9
Peptide K1	DATYTWEHLAWP	−118.9	10.2

[1] (kcal/mol); [2] Standard deviation (kcal/mol).

Computational Design of AnkB-Core Analogs. Considering that the WVIVSDEE (AnkB-core) sequence accounts for 75% of the GABARAP/AnkB-LIR contacts, we proceeded to the design of small peptides endowed with high affinity for GABARAP using AnkB-core as a template. In this attempt, only the residues of the LIR domain (WVIV, positions 2-5 of AnkB-LIR) of AnkB-core (WVIVSDEE) were mutated, because of their direct involvement in the interaction with GABARAP. In this challenging effort, we tried to optimize the peptide sequence, also shared by other GABARAP binders, to improve the ligand/protein complementarity and selectivity. To this end, the affinity maturation protocol was utilized to mutate the VIVS residues into all possible natural amino acid combinations. To avoid the combinatorial explosion, the Monte Carlo optimization option was selected. By this option, 2000 peptides were randomly generated by Monte Carlo algorithm and the peptides with a maximum of three simultaneous mutations were accepted to create the output file containing 100 solutions. Then, the Prime module was also employed to establish whether the mutations led to a more favourable

interaction with the biological counterpart, by calculating the mutant peptides binding free-energy (ΔAffinity) values.

At the end of these calculations, we visually inspected the results for the first 100 peptides with the highest predicted affinity. Among the predicted peptides, we noted that 7 of them displayed ΔAffinity values lower than 2 kcal/mol with respect to the initial template (AnkB-core). In these peptides, position 2 was substituted by Arg, Glu, or Ile; positions 3 and 4 contained only Ile; while position 5 included only alkaline residues, such as His and Arg. Among them, only one candidate, WEIIHDEE, named Pep-sol4, was further investigated by MD simulations and MM-GBSA calculations, because it presented an interesting Glu in position 2. Through this acidic amino acid, the peptide could potentially interact with the positive area shaped by GABARAP-K46 and GABARAP-R67 (two conserved residues among Atg8 proteins). Moreover, GABARAP-K46 is considered to be a universal gate-keeper, regulating the entrance of ligands interacting through the LIR motif [175]. The structural alignment of the GABARAP/Pep-sol4 complex to the GABARAPL2/UBA5 NMR structure (PDB accession code 6H8C) [176] confirmed that the glutamate in position 2 of Pep-sol4 could reproduce the interaction displayed by E15 (GAMEIIHEDNEWGIELVSE) of the "ubiquitin-like modifier activating enzyme 5" (UBA5) with GABARAP-K46.

By applying the computational protocol previously adopted for the reference inhibitors, the binding free-energy value of Pep-sol4 was calculated to be slightly lower than that of AnkB-core (-103.3 *vs.* -101.4 kcal/mol), suggesting that the new peptide could bind GABARAP with a similar binding affinity (**Table 18**).

Table 18. Sequence and calculated ΔG values of AnkB-core analogs.

Peptide	MW [1]	Sequence	ΔG Prime [2]	SD [3]
AnkB-*core*	976.1	WVIVSDEE	-101.4	5.9
Pep-sol4	1070.1	WEIIHDEE	-103.3	7.9
Pep-sol4cc	1032.1	WEC*IHDEC*	-103.7	4.4
WC8	1042.1	WEC*IFDEC*	-115.7	4.3
YC10	1262.4	YGWEC*IFDEC*	-100.4	5.9
WC10	1285.4	WGWEC*IFDEC*	-122.0	5.7

[1] Molecular weight (g/mol); [2] (kcal/mol); [3] Standard deviation (kcal/mol); * Cys involved in disulfide bonds.

Moreover, the visual inspection of the GABARAP/Pep-sol4 MD trajectory and the RMSF plot of the ligand atoms over the simulation time suggested that the C-terminal portion of the peptide was not firmly bound to the GABARAP surface, thus preventing a stable and productive interaction with the target (**Figure 44**). Consequently, considering that the side chains of I3 and D8 were spatially close in the binding mode adopted by Pep-sol4, we designed a cyclic peptide in which I3 and D8 were mutated into two Cys residues bound by a disulphide bond. This modification aimed at reducing the conformational flexibility of the ligand, generating a more stable binding mode on the GABARAP surface. The resulting peptide (named Pep-sol4cc, WECIHDEC) was again analyzed in the complex with GABARAP by MD simulations and MM-GBSA calculations. At the end of these calculations, the estimated ΔG of Pep-sol4cc was -103.7 kcal/mol, a value comparable to that of Pep-sol4 (-103.3 kcal/mol). This information led us to conclude that the structural rigidification did not affect the affinity of the peptide; however, as demonstrated by the ligand RMSF plot (**Figure 44**), an improvement of the conformational stability was successfully achieved.

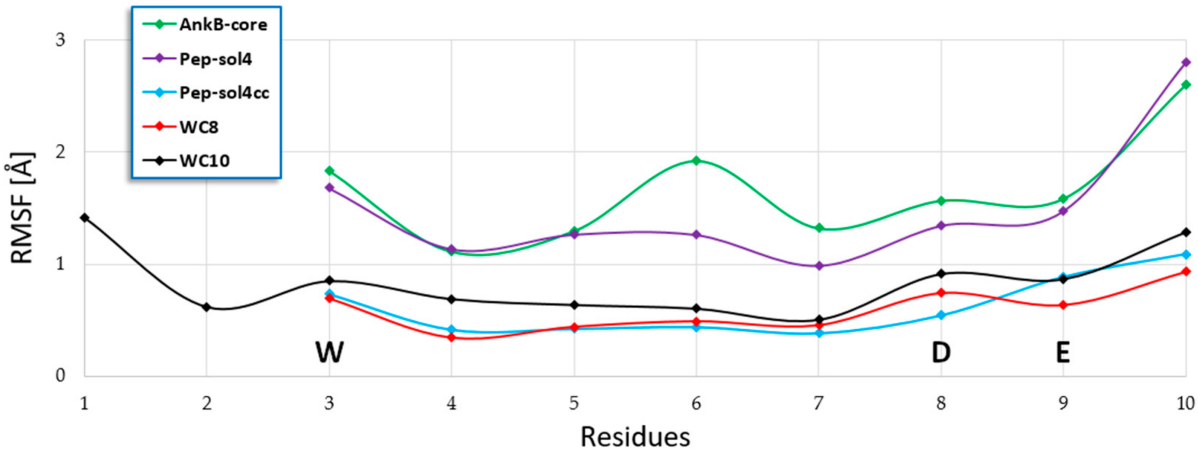


Figure 44. RMSF plots of AnkB-core analogs. Backbone atoms were considered in these calculations. The residues shared by all peptides are highlighted by capital letters.

Computational design of the WC8 and WC10 peptides. Then, with the aim of improving the theoretical binding affinity of Pep-sol4cc, H5 was mutated into a Phe, supposing that it could better fill the hydrophobic area sized by P52, L55, and Q59. The resulting peptide (WC8, sequence WECIFDEC) was analyzed by MD simulations and MM-GBSA calculations, which led to a ΔG value of 12 kcal/mol, lower than that of the originator (**Table 18**). The cluster analysis performed on the MD trajectory frames displayed that, in the structure representative of the most populated cluster of GABARAP/WC8 conformations (accounting for 79% of

conformational ensembles explored), the ligand was stably bound on the GABARAP surface (see **Figure 44** for the RMSF plot), forming numerous interactions (**Table 19** and **Figure 45A**).

In detail, several H-bonds were established, and salt bridges formed between WC8-E2 and the side chains of GABARAP-K46 and -R67, and between the C-term of WC8-C8 and the side chain of GABARAP-R28. Regarding the hydrophobic contacts, the indole ring of WC8-W1 was positioned in a pocket formed by residues I23, I32, P30, K48, and F104, while the side chain of WC8-I4 pointed toward another pocket delimited by Y49, V51, F60, L63, and I64, establishing van der Waals interactions. Finally, similar hydrophobic contacts were also observed between WC8-F5 and the GABARAP area shaped by P52, L55, and L63.

Table 19. List of the interactions established by the GABARAP/WC8 complex during MD simulations.

WC8	GABARAP (H-Bonds)	GABARAP (Hydrophobic)
W_1(NH)	E_{17}(COO^-)	I_{21}, I_{32}, P_{30}, K_{48}, F_{104}
E_2(COO^-)*	K_{46}($NH3^+$)*, R_{67}(=$NH2^+$)*	none
E_2(C=O)	L_{50}(NH)	none
I_4(NH)	L_{50}(C=O)	Y_{49}, V_{51}, F_{60}, L_{63}, I_{64}
F_5	none	P_{52}, L_{55}, L_{63}
C_8(COO^-_{ter})*	R_{28}(=$NH2^+$)*	none

* Salt bridges.

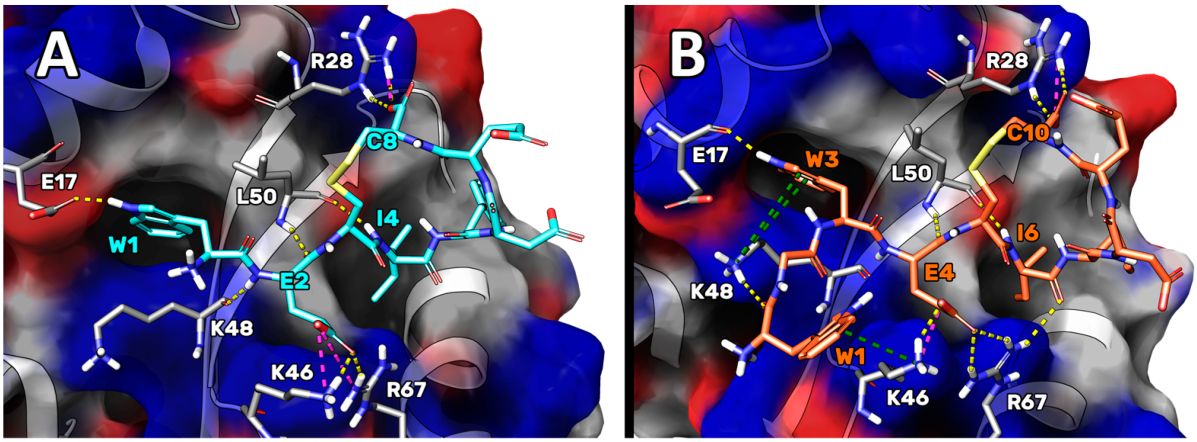


Figure 45. Depiction of the representative structure of the most populated cluster of conformations assumed by WC8 (A) and WC10 (B) in the complex with GABARAP. The GABARAP solvent-accessible surface is shown accordingly by residue charges: blue for positive and red for negative residues, respectively. The interactions between complexes are represented in colored dashes: yellow for H-bonds, purple for salt bridges, and green for cation-π.

WC8 exhibited an estimated ΔG value close to that of K1; hence, with the aim of designing a more potent peptide, we included two additional N-terminal residues. This hypothesis was supported by the fact that the AnkB-LIR peptide (EEWVIVSDEEIEEARQKA), used as a template, contains two glutamate residues before the AnkB-core (WVIVSDEE). For this reason, we speculated that the homologation of WC8 on the N-term could lead to a more potent peptide, considering that the new atoms could create additional bonds. Our objective was to reach the region sized by I32, Y5, and K47, close to the W site, on the GABARAP surface (the yellow area in **Figure 46A**). Therefore, to find the optimal N-terminal sequence, two glycines were initially added to WC8 (GGWECIFDEC); then, the application of the "affinity maturation protocol" on the first Gly residue led to the identification of Tyr (YC10, **Table 18**) and Trp (WC10, **Table 18**) as the most suitable substitutions. In this attempt, the glycine in position 2 was not mutated to allow a certain conformational mobility on the N-terminal tail of the new peptide. Interestingly, the N-terminal residues (WG) and the Glu in position 4 (E4) of WC10 (**WG**W**E**CIFDEC) reproduced the interactions displayed by UBA5 (GAMEIIHEDNE**WG**I**E**LVSE) in the complex with GABARAP and GABARAPL2 [176] (**Figure 46B**).

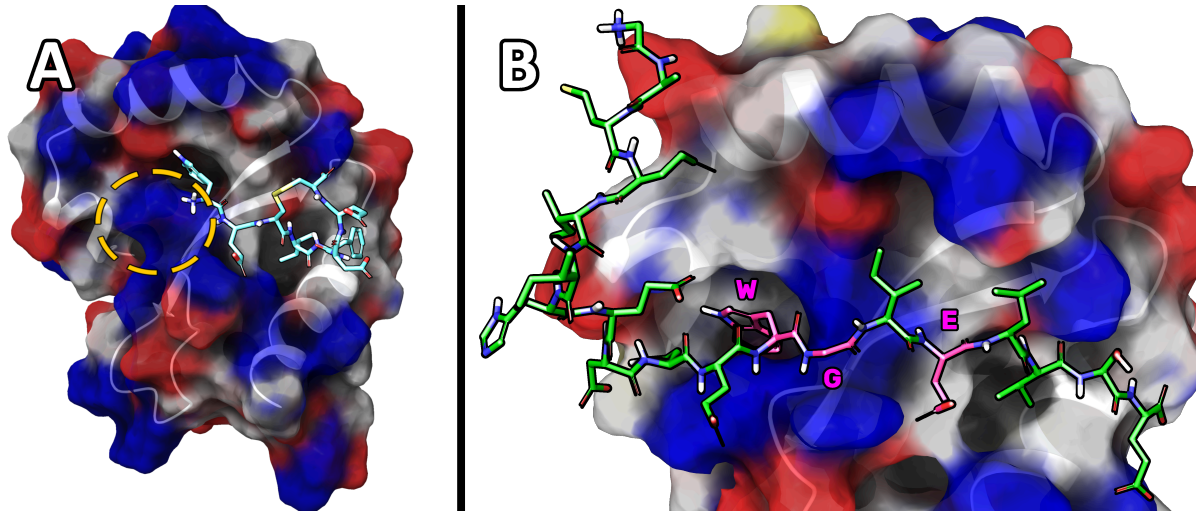


Figure 46. (A) GABARAP/WC8 most representative structure. The yellow circle highlights the protein area in which the two N-terminal residues added on WC8 could be inserted to improve theoretical binding affinity. (B) GABARAPL2/UBA5 NMR structure (PDB accession code 6H8C). In pink sticks are highlighted the residues of UBA5 (GAMEIIHEDNEWGIELVSE) in complex with GABARAP and GABARAPL2 whose interactions are reproduced by peptide WC10 (WGWECIFDEC). The GABARAP solvent-accessible surface is showed accordingly by residue charges: blue for positive and red for negative residues.

MD simulations and MM-GBSA calculations on the GABARAP/YC10 and GABARAP/WC10 complexes revealed that the latter possessed the highest affinity, with a predicted ΔG value almost 7 kcal/mol lower than that of WC8 (**Table 18**). Cluster analysis was then performed on

the GABARAP/WC10 MD trajectory frames; the representative structure of the most populated cluster of conformations (which accounts for the 88% of total conformational ensemble explored) is represented in **Figure 45B**. Notably, the visual inspection of the GABARAP/WC10 most representative structure highlighted that the side chain of the newly added residue (W1) formed a cation-π interaction with GABARAP-K46, while the carbonyl group of the same residue established an H-bond with the side chain of GABARAP-K48 (**Figure 45B**). Surprisingly, the side chain of W1 did not occupy the expected region of GABARAP, but the new additional cation-π interaction greatly contributed to the calculated binding affinity of the peptide. In addition, WC10 (1) shares all the interaction networks established by the GABARAP/WC8 complex, (2) is able to orientate GABARAP-K48 in order to establish additional cation–π interactions with WC10-W3, and (3) is able to shape an additional H-bond interaction between the I6(C=O) and GABARAP-R67(=NH_2^+) (**Figure 45B**).

To conclude, we designed two new cyclic peptides (WC8 and WC10) endowed with a reduced conformational mobility and calculated binding free-energy values in a lower range than those estimated for the reference peptides, AnkB-core and K1. In light of these data, WC8 and WC10 could exhibit higher experimental affinities compared to the reference peptides. Nevertheless, it must be considered that our computations did not account for the entropic contributions to the binding free energy; hence, they should be regarded as a starting point for further experimental studies.

Biophysical experiments. Based on the results of the computational study, the K1, AnkB-core, WC8, and WC10 peptides were purchased by Proteogenix for the experimental investigations. In detail, MST and SPR assays were conducted on the peptides displaying a sufficient stability in water and PBS buffer. Then, the anticancer potential of the most promising candidates was investigated. Initially, we verified that the peptides were water soluble and stable in the buffer in which the recombinant GABARAP protein was solved. Unfortunately, AnkB-core was not soluble in water; thus, it was impossible to use this peptide as reference. Conversely, K1, WC8, and WC10 displayed an excellent solubility and stability in water and PBS. In detail, MST and SPR experiments were carried out with the aim of measuring their K_d values on GABARAP. As a preliminary step, the K_d of the reference peptide K1 was determined in order to (1) check the experimental procedure and verify the result against data reported in the literature by

Weiergräber *et al*. (K_d = 390 nM) [172], and (2) obtain a reference value to compare the K_d measured for the new peptides. MST experiments were conducted on a Monolith NT.115 instrument (**Figure 47A**), while SPR analyses were performed using a BIAcore 8K system, applying the protocol reported by Weiergräber *et al.* [172] (**Figure 47B**).

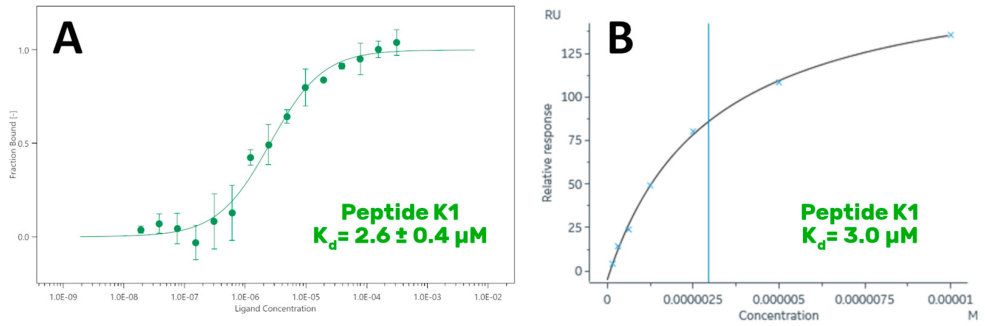


Figure 47. Binding of K1 peptide on GABARAP. (A) MST curve and (B) steady-state analysis obtained by fitting the proper form of the Scatchard equation for the plot of the bound RU at equilibrium *vs.* the ligand concentration in solution during SPR experiments.

Surprisingly, the K_d measured for K1 was close to 3 µM, a value 7 times higher than the one reported in the literature. However, all the techniques employed in this study agreed on this value. The data obtained for WC8 revealed a K_d of 22 µM (**Figure 48AB**), consistent among the different biophysical approaches. Remarkably, WC10 displayed a K_d in the same range of the reference peptide K1, with a value close to 4 µM, obtained by both MST and SPR (**Figure 48CD**).

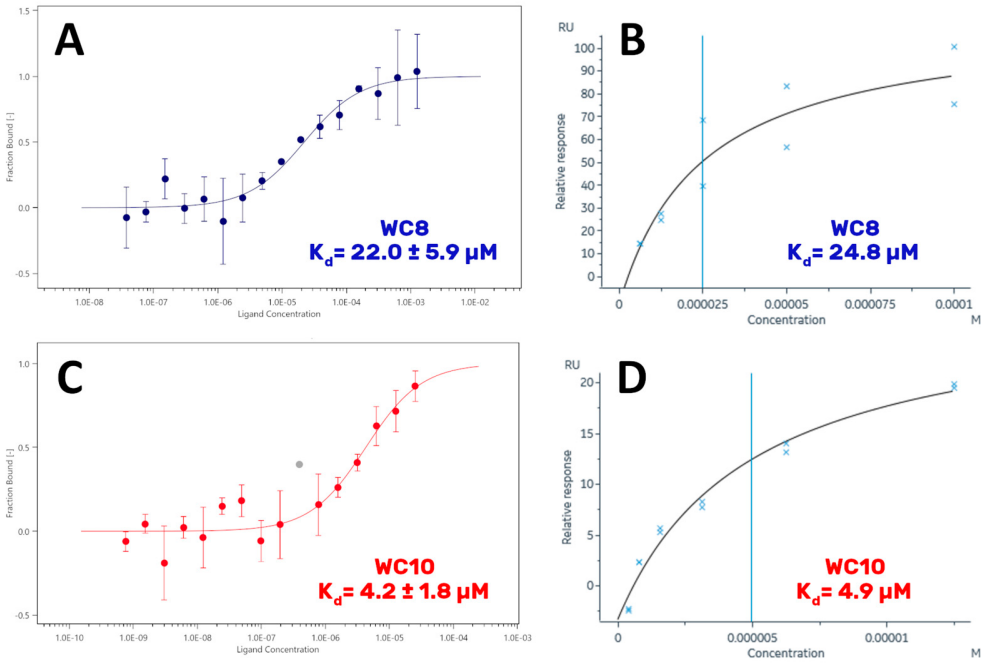


Figure 48. Binding of WC8 and WC10 to GABARAP. MST and SPR curves acquired by recombinant GABARAP incubated with different concentrations of WC8 (A,B) and WC10 (C,D) peptides. In the MST plot referred to WC10 (C), the point corresponding to a concentration of 391 nM (evidenced in gray) appears to be a clear outlier, also considering the other experiments; hence, it was discarded and not included in the calculation of the K_d value.

Since the K_d value of peptide K1 proved to be higher than the one reported in the literature, we decided to validate our data by repeating the MST-binding affinity experiments using another Monolith instrument (Monolith NT.115Pico), located in a different laboratory. The new results confirmed our previous findings, with all the peptides displaying K_d values consistent with those determined earlier (**Figure 49**).

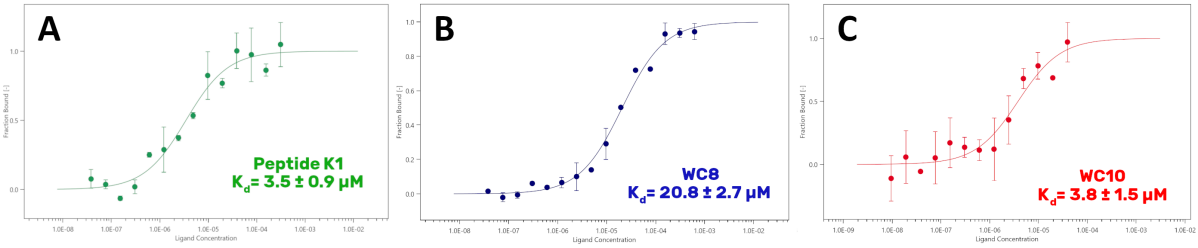


Figure 49. MST curves acquired by GABARAP recombinant protein incubated with different concentrations of K1 (A), WC8 (B) and WC10 (C) peptides using the Monolith NT.115Pico instrument.

Accordingly with the theoretical predictions, WC10 should be more active than K1 (ΔG = -122.0 *vs.* -118.9 kcal/mol, respectively), and WC8 less active than the others (ΔG = -115.7 kcal/mol), as shown in **Table 18**. Considering the confidence range of the experimental K_d and the omission of the entropic contribution to the estimated binding free-energy values, the computations predicted the affinity trend of the selected peptides well.

Biological experiments. The assays were performed by Dr. M. Garofalo (University of Padova). Finally, K1, WC8, and WC10 were tested *in vitro* on PC-3 prostate cancer cells, to evaluate their potential antitumor effects (**Figure 50**). Prostate cancer represents the second most commonly diagnosed malignancy in men worldwide.

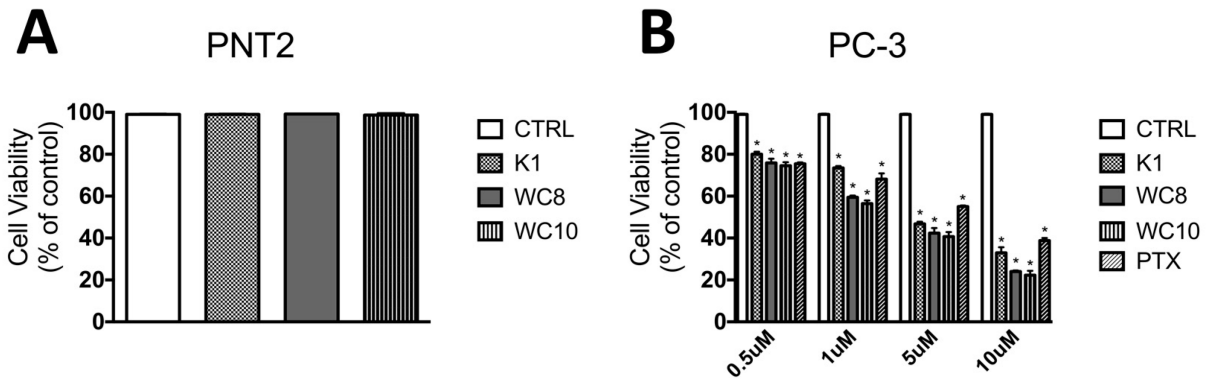


Figure 50. Effect of K1, WC8, and WC10 on cell viability. Cell viability was determined by MTS assay on PNT2 (A) and PC-3 cell lines (B) 96 h post-treatment. The absorbance was measured with a 96-well-plate spectrophotometer (Varioskan Flash Multimode Reader) at 490 nm.

Considering that the probability of developing the disease during a man's lifetime is 15% and that prostate tumor cells can also spread to the lungs and bones via angiogenesis [177], we decided to evaluate the biological activity of the peptides *in vitro* on a prostate cancer model. PC-3 cells were chosen for the screening due to their highly metastatic nature, effectively mimicking an aggressive form of the disease. Notably, it has recently been demonstrated that prostate cancer models show a significant upregulation of autophagy [169,171,178].

Therefore, the biological activity of different concentrations of K1, WC8, and WC10 (from 0.5 to 10 μM) was evaluated with an MTS cell viability assay on PC-3 cells and non-cancerous PNT2 prostate cells (**Figure 50**). The results reported in **Figure 50A** show that, 96 h post-treatment, none of the tested samples displayed a significant cytotoxicity (cell availability > 90%), confirming the excellent biocompatibility and potential pharmacological selectivity for tumor cells. Indeed, as shown in **Figure 50B**, a reduction in cell viability (expressed as percentage of viable cells) was observed in PC-3 cells treated with K1, WC8, and WC10 compared to the untreated control.

Interestingly, the treatments of PC-3 cells with WC8 and WC10 (from 1 to 10 μM) display high efficacy, when compared to Paclitaxel (**Figure 50B**). The *in vitro* data demonstrate that the compounds exhibited a considerable anticancer activity, especially at the highest tested concentration (cell viability 27.16% for K1, 24.06% for WC8, and 22.5% for WC10). The biological data on PC-3 cells indicate that all peptides possess IC_{50} values close to 5 μM, consistent with the K_d estimated by the biophysical experiments. Surprisingly, WC8 displayed a better activity profile than the reference peptide K1. Based on this finding, we may speculate that some other biochemical mechanism or additional activity on different mAtg8 subfamilies could improve the activity of the new peptides [171]. Nevertheless, since the work presented here is a proof-of-concept study, the peptides have been preliminary tested in a non-cancerous and subsequently in a cancer cell line, to exclude possible off-target cytotoxicity and perform an initial pilot study to evaluate the *in vitro* anti-cancer efficacy. However, we are planning to extend the screening to other cancer cell lines in the upcoming further evaluation of the peptides and their antineoplastic mechanism of action. Furthermore, to shed light on the possible secondary targets, we have planned biological and biophysical experiments on LC3B to evaluate if our peptides could show any affinity to it. Further biological assays are needed to unveil the mechanism by which these peptides trigger cell death.

4.4. WWP1

In the fourth case study, I aimed at applying different computational approaches for the rational design of peptides capable of efficiently inhibiting this new promising target, which is involved in the development of cancer. WWP1 is an E3 ubiquitin ligase which has been recently discovered to trigger nondegradative K27-linked polyubiquitination of PTEN to suppress its dimerization and therefore its tumor-suppressive functions. WWP1 is genetically amplified and overexpressed in multiple cancers, including breast, prostate and liver which may lead to the inactivation of the PTEN tumor-suppressor [179].

Drug candidates capable to inhibit WWP1 will pave the way toward a long-requested innovative "tumor suppression reactivation approach" against these kinds of cancer [179]. The unique ligand known for WWP1 is the Indole-3-Carbinol (I3C), a natural compound available in vegetables such as broccoli or cabbage. The literature reports on its high affinity to WWP1, in fact a K_d of 446 nM is known [179]. Interestingly, it derives from the degradation of tryptophan, as a consequence our goal was to design peptides (and then peptidomimetics) having the tryptophan in the primary structure (**Figure 51 - Right**).

To achieve this goal, firstly the X-ray structure of the WWP1 have been retrieved from the protein data bank (PDB accession code 5HPS) [180] (**Figure 51 - Left**). Unfortunately, this structure is not complete, so the missing loops were completely built by homology modeling techniques, using as template the structure of another E3 ligase protein (NEDD4 HECT, PDB code 5C91) [181].

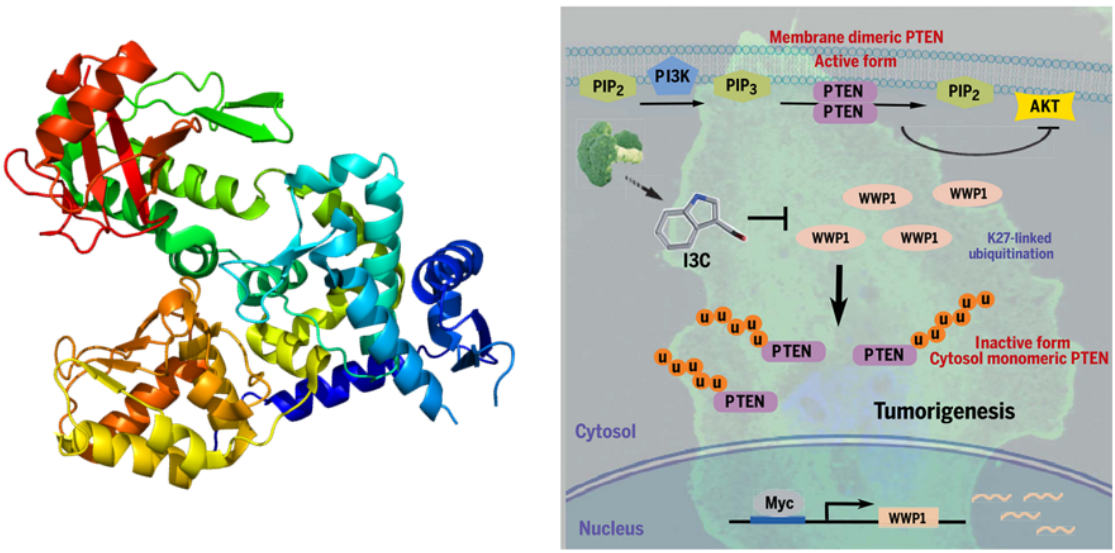


Figure 51. (Left) X-ray crystal structure of WWP1, PDB ID code: 5HPS [180]. (Right) Model for WWP1-mediated PTEN K27-linked polyubiquitination in tumor development and progression [179].

4.4.1. Design of random peptides targeting WWP1

In order to find new potential peptide-based WWP1 inhibitors, we decided to apply the same protocol previously applied for the discovery of the HBP08 peptide for HMGB1-BoxA. At variance, in this case, we wanted to produce random 4-aa peptides containing at least a tryptophan, because its side chain mimics the I3C chemical structure. To this aim, I have developed a python script enclosing features capable to reduce the number of possible generated peptides. In particular, the script is able to generate sequences with at least a Trp (but no more than two and do not consecutive) and three different amino acids those cannot be repetitive and consecutive. These rules allowed us to dramatically reduce the number of generated peptides from 160,000 to 28,481 peptides. For each peptide, both *N*- and *C*-terminals were protected by acetylation and amidation, respectively. This was needed to avoid intramolecular bonds, reducing the possibility of self-cyclization too.

Computational studies. The whole peptide library was docked into the WWP1 putative binding site, by means of the "peptide docking protocol" of GLIDE software. The best 100 peptides ranked by docking score have been selected and subjected to 250 ns-long MD simulations. Then, MM-GBSA calculations were performed to estimate the ΔG values, considering the MD frames in which the peptides displayed the highest stability, according to the RMSD/time plot (for the complete list, see **Annex 2**, Chapter 7). The 11 most promising peptides showing the lowest ΔG values were then selected, and their MD simulations were extended to 250 ns. Next, the peptides showing ΔG values < −40 kcal/mol were selected (five peptides) and subjected to additional independent 500 ns-long MD simulations, in order to increase the statistical significance of the computational studies (**Table 20**). However, only two peptides (namely, 014 and 043) showed a noticeable stability during the additional simulations, in which the atoms velocities were reassigned. These two peptides were then purchased and experimentally tested by *in vitro* experiments.

On parallel to this classical computational pipeline, we decided to apply three different design protocols, characterized by a significant lower computational time cost. In these approaches, the initial docking of a low number of sequences (randomly selected from the list of 28,481 peptides previously generated by our python script) were accomplished.

Table 20. Binding free energy (ΔG) of the 11 most promising peptides out of all the 100 peptides simulated.

Peptide	ΔG 250 ns [1]	ΔG 500 ns [1]	ΔG Rep2 [1]	Average ΔG [1]
011	-32.0 ± 0.2	-30.7 ± 0.3	/	/
012	-30.3 ± 0.1	-30.6 ± 0.2	/	/
014	-37.7 ± 0.3	-41.2 ± 0.2	-36.8 ± 0.3	-39.0
019	-34.9 ± 0.3	-42.1 ± 0.1	*unbound*	/
037	-39.3 ± 0.3	-31.3 ± 0.3	/	/
043	-40.3 ± 0.2	-40.2 ± 0.2	-41.1 ± 0.3	-40.7
064	-30.5 ± 0.2	*unbound*	/	/
072	-39.2 ± 0.4	-44.3 ± 0.4	*unbound*	/
087	-30.3 ± 0.3	*unbound*	/	/
098	-32.4 ± 0.4	-34.2 ± 0.3	*unbound*	/
099	-38.4 ± 0.2	-43.2 ± 0.2	*unbound*	/

[1] (kcal/mol)

The results attained for these peptides were then used to generate a model for the prediction of the activity of the whole peptide library (composed of 28,481 peptides).

Protocol 1: application of a deep learning algorithm, implemented in the DEEPCHEM tool of Maestro Schrödinger Suite. In this case, the training set used to generate the prediction model was made by the best and worst 150 peptides (ranked by docking score) of 500 peptides randomly selected and docked on WWP1. This set was trained for 4 hours using the random split method (correlation: r^2=0.54) and the Gscore of the whole peptide library of 28,481 was predicted. Then, 200 peptides acquiring the highest predicted Gscore were docked in the WWP1 target protein using GLIDE. Finally, 10 of the most promising peptides were simulated in complex with the target by 500 ns-long MD simulations, and the ligand ΔG values were calculated using MM-GBSA approach (**Table 21**).

Applying this deep learning algorithm, we were able to identify two promising peptides which showed ΔG values of -35.8 and -42.2 kcal/mol, respectively. These peptides (namely DC3 and DC4) were then acquired and experimentally tested *in vitro*.

Table 21. Binding free energy (ΔG) of the 10 most promising peptides using the DeepChem-based approach.

Peptide	Docking Gscore [1]	ΔG ± SE [1]
DC1	-8.571	*unbound*
DC2	-8.550	*unbound*
DC3	-8.547	-35.8 ± 0.3
DC4	-8.450	-42.2 ± 0.4
DC5	-8.450	*unbound*
DC6	-8.436	*unbound*
DC7	-8.434	-19.9 ± 0.1
DC8	-8.416	*unbound*
DC9	-8.389	-21.4 ± 0.2
DC10	-8.372	*unbound*

[1] (kcal/mol)

Protocol 2: generation of a predictive model by the application of the Peptide QSAR tool on 500 random peptides, having at least a **Trp** residue in their sequence (out of the 28,481 peptides generated before). In this protocol, two different set of peptides were used to generate the prediction model:

1) The best and worst 50 peptides by docking score (r^2=0.44 and q^2=0.60).

2) The best and worst 25 peptides selected by docking score, MD simulations and binding free energy estimation (r^2=0.633 and q^2=0.697).

Table 22. ΔG values of the 10 most promising Trp-peptides using the Peptide QSAR approach.

Peptide	Pred.Gscore [1]	ΔG Rep1 ± SE [1]	ΔG Rep2 ± SE [1]	Average ΔG [1]
aPQ1	-8.08	*unbound*	*unbound*	/
aPQ2	-7.95	*unbound*	*unbound*	/
aPQ3	-7.90	-28.0 ± 0.3	-30.7 ± 0.3	-29.3
aPQ4	-7.89	*unbound*	*unbound*	/
aPQ5	-7.87	-25.5 ± 0.3	-23.2 ± 0.4	-24.2
bPQ1	-8.51	-38.7 ± 0.3	-41.9 ± 0.3	-40.3
bPQ2	-8.24	*unbound*	*unbound*	/
bPQ3	-8.04	-31.3 ± 0.3	-24.7 ± 0.3	-28.0
bPQ4	-7.96	-36.7 ± 0.3	-25.3 ± 0.3	-31.0
bPQ5	-7.91	-19.0 ± 0.3	*unbound*	/

[1] (kcal/mol)

Using both sets, the best 5 peptides were selected to be further analyzed through two independent 500 ns-long MD simulations replicas, and the peptide ΔG values were calculated (**Table 22**). Two peptides for each method (namely, aPQ3, aPQ5, bPQ1 and bPQ5) were bought and experimentally tested *in vitro*.

Protocol 3: generation of a predictive model by the application of the Peptide QSAR tool on 500 random peptides, having at least a **Tyr** residue in their sequence. In this protocol, the best and worst 50 peptides by docking score were used, attaining r^2=0.70 and q^2=0.74. The five peptides acquiring the highest predicted Gscore were simulated on complex with WWP1 through 500 ns-long MD simulations (**Table 23**). The two most promising peptides (namely, yPQ1 and yPQ5) were synthesized and experimental assayed.

Table 23. Binding free energy (ΔG) of the 10 most promising Tyr-peptides using the Peptide QSAR approach.

Peptide	Pred.Gscore [1]	ΔG Rep1 ± SE [1]	ΔG Rep2 ± SE [1]	Average ΔG [1]
yPQ1	-8.99	-25.9 ± 0.2	-30.9 ± 0.4	-28.4
yPQ2	-8.65	-23.8 ± 0.3	-23.6 ± 0.3	-23.7
yPQ3	-8.47	*unbound*	*unbound*	/
yPQ4	-8.27	*unbound*	*unbound*	/
yPQ5	-8.08	-39.8 ± 0.4	-17.6 ± 0.2	-28.7

[1] (kcal/mol)

To resume, applying the 3 protocols described before, 10 peptides were purchased for the further biophysical experiments. In particular, DSF, MST and SPR experiments were carried out using the recombinant WWP1-HECT domain.

Biophysical experiments. Firstly, DSF experiments showed that peptide bPQ1 was able to significantly reduce the temperature of melting (Tm) of WWP1 (**Figure 52A**), meaning that this peptide could be (probably) able to establish a covalent bond with the target, while for the other peptides, no significant differences in the WWP1 Tm were observed. Then, by performing "binding check" MST experiments, all peptides were tested at a fixed concentration and, interestingly, all of them showed fluorescence signals different from the one observed for WWP1 alone, indicating that at the fixed concentration of 100 µM, they can create meaningful interactions with the target. Consequently, SPR experiments were conducted to estimate the K_d values, and the attained results confirmed the MST outcomes. In fact, the SPR data highlighted that all peptides were able to bind WWP1 with K_d values falling in the low micromolar range. However, a more accurate observation of the SPR curves showed a nonspecific binding for all the peptides, except for the *peptide 043*, which showed a perfect SPR curve exhibiting a K_d value of 6.4 µM (**Figure 52B**).

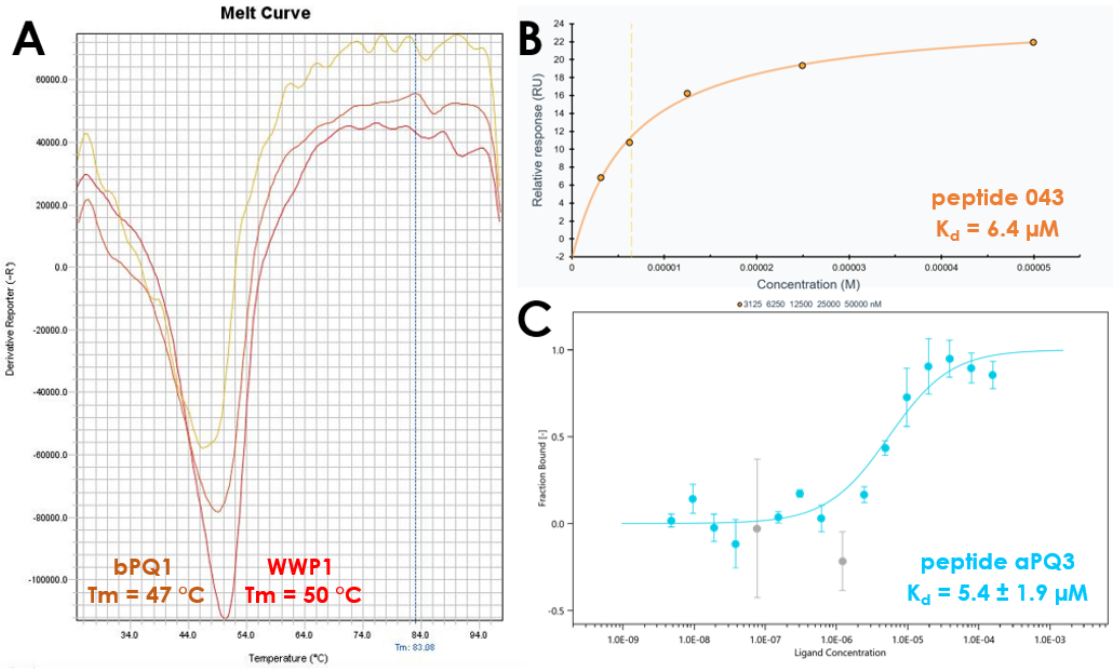


Figure 52. (A) DSF experiment showing only bPQ1 and WWP1 melting curves, (B) SPR plot of peptide 043 by affinity (K_d = 6.4 µM), (C) MST experiment performed in three independent replicas of the peptide aPQ3.

Interestingly, also the SPR data of *peptide bPQ1* confirmed the covalent binding suggested by DSF experiments. Finally, MST experiments were carried out on all the peptide library in order determine their exact K_d values. These experiments highlighted that only the *peptide aPQ3* possessed a significant K_d curve, revealing a value of 5.4 ± 1.9 µM (**Figure 52C**).

Design of new analogs endowed with improved affinity on WWP1. Considering the promising affinity displayed by aPQ3 and 043 peptides, additional computational investigations were performed to identify new mutant peptides with improved affinity on WWP1. In particular, the sequence of both aPQ3 and 043 peptides were extended by two Gly residues in the C-terminal region. Subsequently, both Gly residues were systematically replaced by all the natural amino acids applying the affinity maturation protocol, and ranked by the ΔAffinity values. The most promising hexapeptides were then simulated through two independent 500 ns-long MD simulations, and their ΔG values were estimated by MM-GBSA method (**Table 24**).

The attained results suggested that the peptide 043-B showed a ΔG value 10 kcal/mol lower than its parent peptide 043, while the ΔG values of the new aPQ3 analogs aPQ3-B and -E were 11.7 and 15.5 kcal/mol lower than the one of aPQ3, respectively. Consequently, 043-B, aPQ3-B and aPQ3-E were synthesized and biophysically examined through MST and SPR experiments.

Table 24. ΔG of the mutated 043 and aPQ3 hexapeptides derived from the affinity maturation protocol.

Peptide	ΔAffinity [1]	ΔStability [1]	ΔG Rep1 ± SE [1]	ΔG Rep2 ± SE [1]	Average ΔG [1]
043	/	/	-40.2 ± 0.2	-41.1 ± 0.3	-40.7
043-A	-9.58	+0.43	*unbound*	*unbound*	/
043-B	-4.36	+1.58	-52.0 ± 0.5	-49.5 ± 0.3	-50.7
043-C	-2.86	+0.38	-30.6 ± 0.2	*unbound*	/
aPQ3	/	/	-28.0 ± 0.3	-30.7 ± 0.3	-29.3
aPQ3-A	-32.33	-10.16	*unbound*	-33.2 ± 0.4	/
aPQ3-B	-27.81	-6.38	-49.5 ± 0.3	-32.5 ± 0.2	-41.0
aPQ3-C	-27.47	-0.87	-40.8 ± 0.2	-28.6 ± 0.3	-34.7
aPQ3-D	-11.56	-6.57	*unbound*	*unbound*	/
aPQ3-E	-10.03	+5.49	-42.7 ± 0.2	-46.9 ± 0.2	-44.8
aPQ3-F	-4.60	+1.52	*unbound*	*unbound*	/
aPQ3-G	-2.01	-3.22	*unbound*	*unbound*	/

[1] (kcal/mol).

Biophysical experiments, run 2. SPR experiments showed clear affinity curves by for both aPQ3-B (K_d = 920 nM, **Figure 53A**) and aPQ3-E (K_d = 625 nM, **Figure 53C**). However, only for aPQ3-B it is observable a clear kinetic plot (K_d = 1300 nM, **Figure 53B**). Conversely, no binding curve was detected for 043-B. Further MST experiments confirmed the SPR results, showing a similar K_d for both aPQ3-B (reported in **Figure 53D**) and aPQ3-E (data not shown), while no binding was again detected for 043-B.

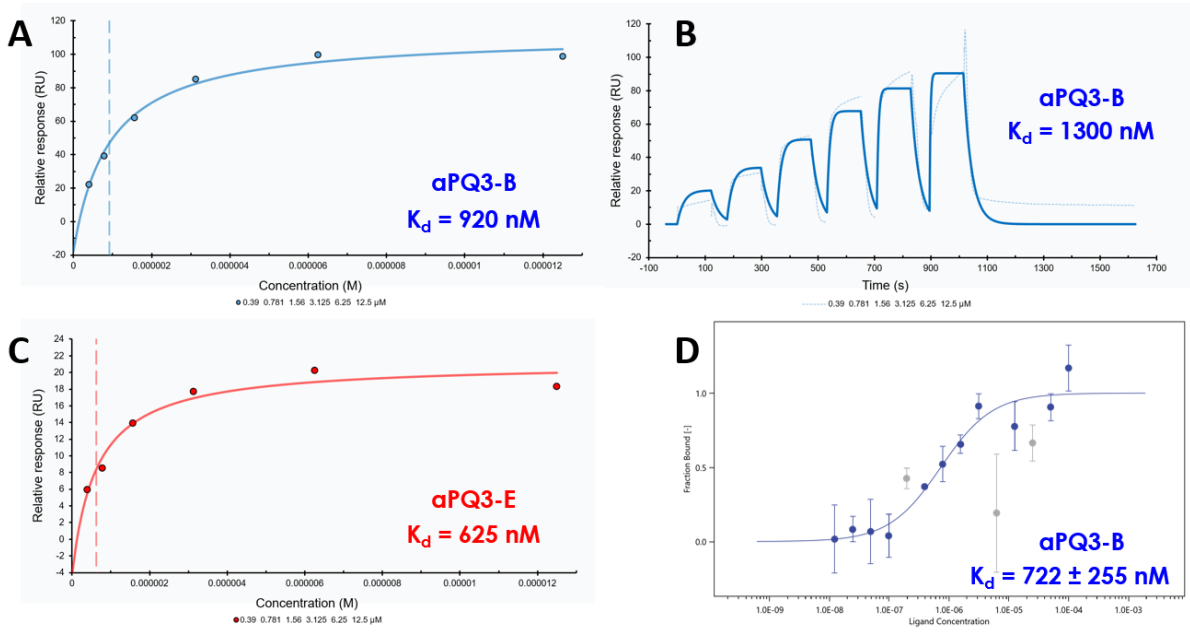


Figure 53. (A) SPR affinity plot of aPQ3-B. (B) SPR kinetic plot of aPQ3-B. (C) SPR affinity plot of aPQ3-E. (D) MST experiment of aPQ3-B on WWP1-HECT domain in two independent replicas, the grey points were discarded since they could be considered as outliers.

Accordingly with these data, aPQ3-B and bPQ1 peptides were selected for the further evaluation of their potential inhibitory activity of the WWP1 mediated poly-ubiquitination. In fact, even if aPQ3-E showed a higher binding affinity compared to aPQ3-B, its relative response in the SPR affinity curve (**Figure 53C**) is significantly lower than the one reported for aPQ3-B (**Figure 53A**), suggesting a lower reliability of the obtained results. In addition, aPQ3-E did not show a kinetic curve and the MST curve was not significant.

5. CONCLUSIONS AND FUTURE PERSPECTIVES

Since HMGB1 has been identified as one of the main mediators in both acute and chronic inflammation, playing a significant role in several pathological conditions [113] including rheumatoid arthritis [182,183], systemic lupus erythematosus [184], ankylosing spondylitis [185] and other autoimmune diseases [186,187], we believe that our research activity could give a significant contribution to the pharmacological treatment of these chronic inflammatory diseases, and in particular to rheumatoid arthritis (RA). In fact, in the past decade, it has been demonstrated that the CXCL12/HMGB1 heterocomplex perpetuates the chronic inflammation observed in RA [183]. The lack of full remission in a portion of RA patients, and the evidence that the composition of the synovial tissue correlates with the response to the available treatments, calls for the identification of novel drug targets and the development of selective therapies [188,189]. In fact, to date, the currently available drugs for the treatment of RA can only slightly reduce the symptoms and the progression of the disease, producing several side effects for their non-specific mechanism of action. Therefore, small molecules or peptides able to hinder the formation of the CXCL12/HMGB1 heterocomplex could be useful as novel therapeutic strategies.

Applying several computational approaches we were able to discover the HBP08 peptide, which represents the first potent peptide inhibitor of the CXCL12/HMGB1 PPI. In particular, HBP08 is able to selectively block the activity of the HMGB1/CXCL12 heterocomplex ($IC_{50} \simeq 50$ μM), without altering the migration induced by CXCL12 alone, and leaving HMGB1 capable of interacting with TLR4. In addition, HBP08 is able to bind on HMGB1 in the low micromolar range (K_d = 0.8 ± 0.4 μM), representing the ligand with the highest affinity reported in literature so far. The results of this work have been recently published in the *Journal of Medicinal Chemistry* [128]. Then, with the aim to improve the affinity of HBP08 to HMGB1-BoxB, we optimized the peptide structure by the application of additional computational approaches, such as alanine scanning and affinity maturation protocols. In this way, we identified the HBP08-pep2 peptide, which selectively binds on HMGB1-BoxB in the low nanomolar range (K_d = 13.7 ± 2.2 nM), confirmed also by repeating the experiment to *full*-HMGB1 (K_d = 15.8 ± 7.4 nM), while it is able to bind HMGB1-BoxA showing a K_d of 3.85 ± 0.41 μM.

In addition, preliminary chemotaxis assays showed that HBP08-pep2 is able block the activity of the HMGB1/CXCL12 heterocomplex ($IC_{50} \simeq 2.5$ μM), which it means it is about 20-fold more active than HBP08. Further biological experiments are currently ongoing in order to better clarify the action mechanism of the peptide.

While the use of peptides as therapeutics remains challenging, we believe that these peptides can be exploited for therapeutic intervention while being immediately useful as a tool for cell biologists to further dissect the inflammatory pathways triggered by the CXCL12/HMGB1 heterocomplex. Moreover, our biophysical and structural biology studies indicated that the *C*-terminal moiety of the peptide is crucial for the interaction with both BoxA and BoxB of HMGB1, providing important information for the design of novel peptidomimetic anti-inflammatory drugs. Peptidomimetics are characterized by less pharmacokinetic troubles being, they are more suitable for further *in vivo* studies and could represent a new class of anti-inflammatory drugs with an innovative mechanism of action, potentially free of the typical side-effects displayed by the currently anti-inflammatory drugs. This new class of drugs could significantly determine a strong impact on the quality of life for all people affected by these disabling pathological ailments in which HMGB1 protein is greatly involved.

In the second case study, based on the promising data regarding the dual hypocholesterolemic activity of the lupin peptide P5 [141], we computationally designed new analogs endowed with improved PCSK9 and HMG-CoAR inhibitory activities. The attained experimental data confirmed the theoretical studies, revealing that the mutant peptide P5-Best showed the lowest PCSK9 IC_{50} value of 0.7 μM. Further biological assays demonstrated that all mutant peptides that maintained the dual PCSK9/HMG-CoAR inhibitory activity also improving the ability of HepG2 cells to absorb extracellular LDL by up to 254% (P5-Best data). The results of this study were recently published on *Pharmaceutics* journal [154]. Doubtless, peptide P5 and its analogs displayed activity in the micromolar range suggesting that still their exploitation in the clinical application is challenging. Therefore, more efforts have to be pursued in order to improve their dual-inhibitory activity. However, evidence support the fact that P5 and its analogs can be considered as promising lead compounds for the development of a new class of hypocholesterolemic drugs endowed with dual-inhibitory activity of both PCSK9 and HMG-CoAR targets. Indeed, the dual and synergistic activity may be useful for better achieving the

biological effect than compound actives on one of those targets. This study confirms that a multidisciplinary approach in the design of new peptides is successful in identifying peptides endowed with hypocholesterolemic effects, offering a promising starting point for the design of peptidomimetics that lack the bioavailability problems of peptides. In fact, in the second part of this research project, based on the published data regarding the β-sheet peptidomimetic RIm13, which exhibits an inhibitory activity on the PCSK9/LDLR PPI ($IC_{50} \simeq 1.0$ μM), we computationally designed new poly-imidazole derivatives endowed with improved PCSK9 and HMG-CoAR inhibitory activities. In particular, we refined the chemical structure of RIm13 in order to fully occupy a hydrophobic pocket found on the PCSK9 surface. Then, aiming at further simplifying the chemical structure of the compounds, and improving their synthetic feasibility as well, we tried to fuse the benzene and the first imidazole ring into different aromatic rings, capable of mimicking the π electrons conjugation between both rings. Finally, aiming at improving the stability of the compound on the PCSK9 surface and to evaluate the influence of the third imidazole ring on the predicted ΔG of the compounds, we additionally modified the chemical structure by displacing the R group on the third imidazole ring to others. In this way, we identified a di-imidazole derivate (namely, Dim16) able to inhibit the PCSK9-LDLR binding with a dose response trend in the high picomolar range (IC_{50} = 0.8 ± 1.0 nM), and to inhibit the HMG-CoAR activity with an IC_{50} of 146.8 ± 75.09 μM. The identified β-sheet peptidomimetic Dim16 could represent a new class of drugs for the treatment of hypercholesterolemia, potentially free of the typical side-effects displayed by the currently statins targeting HMG-CoAR. Further experiments will be performed to evaluate the intestinal stability and propensity of Dim16 to be trans-epithelial transported by mature Caco-2 cells.

In the third case study, we were able to create an affordable GABARAP/AnkB-LIR complex computational model, starting from the GABARAP/AnkB-LIR X-ray crystal structure. This was utilized to investigate the role played by different regions of the AnkB-LIR sequence. Then, by applying integrated computational techniques we designed two cyclic peptides (namely, WC8 and WC10) endowed with theoretical affinities in line with the ones predicted for the reference peptides AnkB-core and K1. The experimental measurement of the K_d values led us to prove that WC10 (2 residues shorter and more rigid than K1) displays a biological activity like that of K1. MST, SPR, and *in vitro* assays on PC-3 cells confirmed this observation. This work has been recently published on the *International Journal of Molecular Sciences* [190].

Further biological assays are needed to unveil the mechanism by which these peptides trigger cell death, however, we believe this study has the potential to open new avenues of research towards the design of novel anticancer compounds, employing WC10 as a structural template. Additionally, our results confirm that a suitable interference with the autophagy process of cancer cells can represent an innovative and viable therapeutic strategy. Consequently, we are confident that the discovery of new potent and specific autophagy modulators will become increasingly important in the treatment of cancer [191].

Finally, in the last case study, we applied both classical and machine learning-based computational approaches in order to predict the theoretical binding affinity of more than 28,000 tetrapeptides to WWP1, a promising anticancer target. In fact, it is shown that WWP1 is overexpressed in several type of cancers, such as breast, prostate and liver. In this way we were able to identify a promising hexapeptide (namely, aPQ3-B) able to bind the WWP1 protein in the high nanomolar range, as confirmed by both SPR and MST biophysical experiments. Further biological experiments are currently ongoing in order to clarify if this peptide also possess an inhibitory activity on WWP1.

In conclusion, since it is known that the structure of active peptides can be used as template for the discovery of new peptidomimetics, we are going to perform a virtual screening of peptidomimetics library with the application of machine learning algorithms. In fact, machine learning techniques have shown tremendous potential in areas such as computer vision, speech recognition, and natural language processing. Very recently, machine learning algorithms have also been successfully applied in drug discovery programs obtaining stunning results. In fact, novel promising compounds endowed with strong binding affinity to targeted proteins have been identified. In addition, machine learning approaches will also be used to predict key property values of these new compounds, to prioritize them for the follow-up screening, and to gain insight into their structure-activity studies (SARs) [192].

6. REFERENCES

[1] Nasedkin A, Cerveny S, Swenson J. *Molecular Insights into Dipole Relaxation Processes in Water-Lysine Mixtures*. J Phys Chem B. **2019** Jul 18;123(28):6056-6064.

[2] Wieland T, Bodanszky M. *The World of Peptides: A Brief History of Peptide Chemistry.* **1991**. ISBN 978-3-642-75850-8.

[3] Russell JA. *Fifty Years of Advances in Neuroendocrinology*. Brain Neurosci Adv. **2018** Nov 16;2:2398212818812014.

[4] *"WHO Model List of Essential Medicines (19th List)" (PDF).* World Health Organization. **2015**. Archived (PDF) from the original on 13 December 2016. Retrieved 8 December 2016.

[5] Henninot A, Collins JC, Nuss JM. *The Current State of Peptide Drug Discovery: Back to the Future?* J Med Chem. **2018** Feb 22;61(4):1382-1414.

[6] Levine PM, Balana AT, Sturchler E, Koole C, Noda H, Zarzycka B, Daley EJ, Truong TT, Katritch V, Gardella TJ, Wootten D, Sexton PM, McDonald P, Pratt MR. *O-GlcNAc Engineering of GPCR Peptide-Agonists Improves Their Stability and in Vivo Activity*. J Am Chem Soc. **2019** Sep 11;141(36):14210-14219.

[7] Sharma K, Sharma KK, Sharma A, Jain R. *Peptide-based drug discovery: Current status and recent advances*. Drug Discov Today. **2022** Dec 5:103464.

[8] Liu RH, He YH, Yu W, Zhou B, Han B. *Silver-Catalyzed Site-Selective Ring-Opening and C-C Bond Functionalization of Cyclic Amines: Access to Distal Aminoalkyl-Substituted Quinones*. Org Lett. **2019** Jun 21;21(12):4590-4594.

[9] Hibbitts A, O'Connor AM, McCarthy J, Forde ÉB, Hessman G, O'Driscoll CM, Cryan SA, Devocelle M. *Poly(ethylene glycol)-Based Peptidomimetic "PEGtide" of Oligo-Arginine Allows for Efficient siRNA Transfection and Gene Inhibition*. ACS Omega. **2019** Jun 10;4(6):10078-10088.

[10] Werner HM, Cabalteja CC, Horne WS. *Peptide Backbone Composition and Protease Susceptibility: Impact of Modification Type, Position, and Tandem Substitution*. Chembiochem. **2016** Apr 15;17(8):712-8.

[11] Taylor M, Moore S, Mayes J, Parkin E, Beeg M, Canovi M, Gobbi M, Mann DM, Allsop D. *Development of a proteolytically stable retro-inverso peptide inhibitor of beta-amyloid oligomerization as a potential novel treatment for Alzheimer's disease.* Biochemistry. **2010** Apr 20;49(15):3261-72.

[12] Frey V, Viaud J, Subra G, Cauquil N, Guichou JF, Casara P, Grassy G, Chavanieu A. *Structure-activity relationships of Bak derived peptides: affinity and specificity modulations by amino acid replacement*. Eur J Med Chem. **2008** May;43(5):966-72.

[13] Bird GH, Madani N, Perry AF, Princiotto AM, Supko JG, He X, Gavathiotis E, Sodroski JG, Walensky LD. *Hydrocarbon double-stapling remedies the proteolytic instability of a lengthy peptide therapeutic*. Proc Natl Acad Sci USA. **2010** Aug 10;107(32):14093-8.

[14] Buckton LK, Rahimi MN, McAlpine SR. *Cyclic Peptides as Drugs for Intracellular Targets: The Next Frontier in Peptide Therapeutic Development*. Chemistry. **2021** Jan 21;27(5):1487-1513.

[15] Sindhikara D, Borrelli K. *High throughput evaluation of macrocyclization strategies for conformer stabilization*. Sci Rep. **2018** Apr 26;8(1):6585.

[16] Bruno BJ, Miller GD, Lim CS. *Basics and recent advances in peptide and protein drug delivery*. Ther Deliv. **2013** Nov;4(11):1443-67.

[17] Zielińska A, Carreiró F, Oliveira AM, Neves A, Pires B, Venkatesh DN, Durazzo A, Lucarini M, Eder P, Silva AM, Santini A, Souto EB. *Polymeric Nanoparticles: Production, Characterization, Toxicology and Ecotoxicology*. Molecules. **2020** Aug 15;25(16):3731.

[18] Chan YP, Meyrueix R, Kravtzoff R, Nicolas F, Lundstrom K. *Review on Medusa:a polymer-based sustained release technology for protein and peptide drugs*. Expert Opin Drug Deliv. **2007** Jul;4(4):441-51.

[19] Chen C, Zheng H, Xu J, Shi X, Li F, Wang X. *Sustained-release study on Exenatide loaded into mesoporous silica nanoparticles: in vitro characterization and in vivo evaluation*. Daru. **2017** Sep 4;25(1):20.

[20] Perez JJ. *Designing Peptidomimetics*. Curr Top Med Chem. **2018**;18(7):566-590.

[21] Mizuno A, Matsui K, Shuto S. *From Peptides to Peptidomimetics: A Strategy Based on the Structural Features of Cyclopropane*. Chemistry. **2017** Oct 17;23(58):14394-14409.

[22] Pelay-Gimeno M, Glas A, Koch O, Grossmann TN. *Structure-Based Design of Inhibitors of Protein-Protein Interactions: Mimicking Peptide Binding Epitopes*. Angew Chem Int Ed Engl. **2015** Jul 27;54(31):8896-927.

[23] Whitby LR, Ando Y, Setola V, Vogt PK, Roth BL, Boger DL. *Design, synthesis, and validation of a β-turn mimetic library targeting protein-protein and peptide-receptor interactions*. J Am Chem Soc. **2011** Jul 6;133(26):10184-94.

[24] Ruiz-Gómez G, Tyndall JD, Pfeiffer B, Abbenante G, Fairlie DP. *Update 1 of: Over one hundred peptide-activated G protein-coupled receptors recognize ligands with turn structure*. Chem Rev. **2010** Apr 14;110(4):PR1-41.

[25] Madala PK, Tyndall JD, Nall T, Fairlie DP. *Update 1 of: Proteases universally recognize beta strands in their active sites*. Chem Rev. **2010** Jun 9;110(6):PR1-31.

[26] Ko E, Liu J, Burgess K. *Minimalist and universal peptidomimetics*. Chem Soc Rev. **2011** Aug;40(8):4411-21.

[27] Hirschmann R, Nicolaou KC, Pietranico S, Salvino J, Leahy EM, Sprengeler PA, Furst G, Strader CD, Smith AB. *Nonpeptidal peptidomimetics with β-D-glucose scaffolding. A partial somatostatin agonist bearing a close structural relationship to a potent, selective substance P antagonist*. J Am Chem Soc. **1992** Nov 1; 114(23):9217-9218.

[28] Hirschmann R, Sprengeler PA, Kawasaki T, Leahy JW, Shakespeare WC, Smith AB. *The first design and synthesis of a steroidal peptidomimetic. The potential value of peptidomimetics in elucidating the bioactive conformation of peptide ligands*. J Am Chem Soc. **1992** Nov 1; 114(24):9699-9701.

[29] Mowery BP, Prasad V, Kenesky CS, Angeles AR, Taylor LL, Feng JJ, Chen WL, Lin A, Cheng FC, Smith AB, Hirschmann R. *Catechol: A minimal scaffold for non-peptide peptidomimetics of the i + 1 and i + 2 positions of the beta-turn of somatostatin*. Org Lett. **2006** Sep 28;8(20):4397-400.

[30] Kapetanovic IM. *Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach*. Chem Biol Interact. **2008** Jan 30;171(2):165-76.

[31] Shaikh SA, Jain T, Sandhu G, Latha N, Jayaram B. *From drug target to leads--sketching a physicochemical pathway for lead molecule design in silico*. Curr Pharm Des. **2007**;13(34):3454-70.

[32] Aparoy P, Reddy KK, Reddanna P. *Structure and ligand based drug design strategies in the development of novel 5- LOX inhibitors*. Curr Med Chem. **2012**;19(22):3763-78.

[33] Acharya C, Coop A, Polli JE, Mackerell AD Jr. *Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach*. Curr Comput Aided Drug Des. **2011** Mar;7(1):10-22.

[34] Bunaciu AA, Udriştioiu EG, Aboul-Enein HY. *X-ray diffraction: instrumentation and applications*. Crit Rev Anal Chem. **2015**;45(4):289-99.

[35] Earl LA, Falconieri V, Milne JL, Subramaniam S. *Cryo-EM: beyond the microscope*. Curr Opin Struct Biol. **2017** Oct;46:71-78.

[36] Campagne S, Gervais V, Milon A. *Nuclear magnetic resonance analysis of protein-DNA interactions*. J R Soc Interface. **2011** Aug 7;8(61):1065-78.

[37] Kroemer RT. *Structure-based drug design: docking and scoring*. Curr Protein Pept Sci. **2007** Aug;8(4):312-28.

[38] Cavasotto CN, Phatak SS. *Homology modeling in drug discovery: current trends and applications*. Drug Discov Today. **2009** Jul;14(13-14):676-83.

[39] Finn J. *Application of SBDD to the discovery of new antibacterial drugs*. Methods Mol Biol. **2012**;841:291-319.

[40] Soni M, Pratap JV. *Development of Novel Anti-Leishmanials: The Case for Structure-Based Approaches*. Pathogens. **2022** Aug 22;11(8):950.

[41] Yu W, MacKerell AD Jr. *Computer-Aided Drug Design Methods*. Methods Mol Biol. **2017**;1520:85-106.

[42] Burley SK, Berman HM, Kleywegt GJ, Markley JL, Nakamura H, Velankar S. *Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive*. Methods Mol Biol. **2017**;1607:627-641.

[43] Wells JN, Bergendahl LT, Marsh JA. *Computational Modelling of Protein Complex Structure and Assembly*. Methods Mol Biol. **2018**;1764:347-356.

[44] Pagadala NS, Syed K, Tuszynski J. *Software for molecular docking: a review*. Biophys Rev. **2017** Apr;9(2):91-102.

[45] Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. *Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy*. J Med Chem. **2004** Mar 25;47(7):1739-49.

[46] Cross JB, Thompson DC, Rai BK, Baber JC, Fan KY, Hu Y, Humblet C. *Comparison of several molecular docking programs: pose prediction and virtual screening accuracy*. J Chem Inf Model. **2009** Jun;49(6):1455-74.

[47] Plewczynski D, Łaźniewski M, Augustyniak R, Ginalski K. *Can we trust docking results? Evaluation of seven commonly used programs on PDBbind database*. J Comput Chem. **2011** Mar;32(4):742-55.

[48] Li X, Li Y, Cheng T, Liu Z, Wang R. *Evaluation of the performance of four molecular docking programs on a diverse set of protein-ligand complexes*. J Comput Chem. **2010** Aug;31(11):2109-25.

[49] Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL. *Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening*. J Med Chem. **2004** Mar 25;47(7):1750-9.

[50] Harder E, Damm W, Maple J, Wu C, Reboul M, Xiang JY, Wang L, Lupyan D, Dahlgren MK, Knight JL, Kaus JW, Cerutti DS, Krilov G, Jorgensen WL, Abel R, Friesner RA. *OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins*. J Chem Theory Comput. **2016** Jan 12;12(1):281-96.

[51] Doniach S, Eastman P. *Protein dynamics simulations from nanoseconds to microseconds*. Curr Opin Struct Biol. **1999** Apr;9(2):157-63.

[52] Hollingsworth AS, Dror RO. *Molecular Dynamics Simulation for All*. Neuron, volume 99, issue 6, **2018**, pages 1129-1143, ISSN 0896-6273.

[53] Dror RO, Jensen MØ, Borhani DW, Shaw DE. *Exploring atomic resolution physiology on a femtosecond to millisecond timescale using molecular dynamics simulations*. J Gen Physiol. **2010** Jun;135(6):555-62.

[54] Hernández-Rodríguez M, Rosales-Hernández MC, Mendieta-Wejebe JE, Martínez-Archundia M, Basurto JC. *Current Tools and Methods in Molecular Dynamics (MD) Simulations for Drug Design*. Curr Med Chem. **2016**;23(34):3909-3924.

[55] Case DA, Belfon K, Ben-Shalom IY, Brozell SR, Cerutti DS, Cheatham TE, Cisneros GA, Cruzeiro VWD, Darden TA, Duke RE, Giambasu G, Gilson MK, Gohlke H, Goetz AW, Harris R, Izadi S, Izmailov SA, Kasavajhala K, Kovalenko A, Krasny R, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Man V, Merz KM, Miao Y, Mikhailovskii O, Monard G, Nguyen H, Onufriev A, Pan F, Pantano S, Qi R, Rahnamoun A, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling CL, Skrynnikov NR, Smith J, Swails J, Walker RC, Wang J, Wolf RM, Wu X, Xiong Y, Xue Y, York DM, Zhao S, and Kollman PA (**2020**), AMBER 2020, University of California, San Francisco.

[56] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. *ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB*. J Chem Theory Comput. **2015** Aug 11;11(8):3696-713.

[57] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein LM. *Comparison of simple potential functions for simulating liquid water*. J Chem Phys. **1983**;79:926-935.

[58] Joung IS, Cheatham TE 3rd. *Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations*. J Phys Chem B. **2008** Jul 31;112(30):9020-41.

[59] Chiarelli LR, Mori M, Barlocco D, Beretta G, Gelain A, Pini E, Porcino M, Mori G, Stelitano G, Costantino L, Lapillo M, Bonanni D, Poli G, Tuccinardi T, Villa S, Meneghetti F. *Discovery and development of novel salicylate synthase (MbtI) furanic inhibitors as antitubercular agents*. Eur J Med Chem. **2018** Jul 15;155:754-763.

[60] Eslami H, Mojahedi F, Moghadasi J. *Molecular dynamics simulation with weak coupling to heat and material baths*. J Chem Phys. **2010** Aug 28;133(8):084105.

[61] Darden TDY, Pedersen L. *Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems*. J Chem Phys. **1993**;98:10089-10092.

[62] Ryckaert JP, Ciccotti G, Berendsen HJC. *Numerical integration of the Cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes*. J Comput Phys. **1977**;23:327-341.

[63] Le Grand S, Götz AW, Walker RC. *SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations*. Comput Phys Commun. **2013**;184:374-380.

[64] Coutsias EA, Seok C, Dill KA. *Using quaternions to calculate RMSD*. J Comput Chem. **2004** Nov 30;25(15):1849-57.

[65] Miller BR 3rd, McGee TD Jr, Swails JM, Homeyer N, Gohlke H, Roitberg AE. *MMPBSA.py: An Efficient Program for End-State Free Energy Calculations*. J Chem Theory Comput. **2012** Sep 11;8(9):3314-21.

[66] Roe DR, Cheatham TE 3rd. *PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data*. J Chem Theory Comput. **2013** Jul 9;9(7):3084-95.

[67] Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. *Peptide Folding: When Simulation Meets Experiment*. Angew. Chem. Int. Ed. **1999**; 38:236-240.

[68] Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E. *GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics*. **2013** Apr 1;29(7):845-54.

[69] Morrison KL, Weiss GA. *Combinatorial alanine-scanning*. Curr Opin Chem Biol. **2001** Jun;5(3):302-7.

[70] Zhu K, Day T, Warshaviak D, Murrett C, Friesner R, Pearlman D. *Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction*. Proteins. **2014** Aug;82(8):1646-55.

[71] Merk D, Friedrich L, Grisoni F, Schneider G. *De Novo Design of Bioactive Small Molecules by Artificial Intelligence*. Mol Inform. **2018** Jan;37(1-2):1700153.

[72] Zhang Y, Zhang G, Shang Q. *Computer-Aided Clinical Trial Recruitment Based on Domain-Specific Language Translation: A Case Study of Retinopathy of Prematurity*. J Healthc Eng. **2017**;2017:7862672.

[73] Cano G, Garcia-Rodriguez J, Garcia-Garcia A, Perez-Sanchez H, Benediktsson JA, Thapa A, Barr A. *Automatic selection of molecular descriptors using random forest: Application to drug discovery. Expert Systems with Applications*, volume 72, **2017**, pages 151-159, ISSN 0957-4174.

[74] Gunčar G, Kukar M, Notar M, Brvar M, Černelč P, Notar M, Notar M. *An application of machine learning to haematological diagnosis*. Sci Rep. **2018** Jan 11;8(1):411.

[75] Li H, Zhang W, Chen Y, Guo Y, Li GZ, Zhu X. *A novel multi-target regression framework for time-series prediction of drug efficacy*. Sci Rep. **2017** Jan 18;7:40652.

[76] Maltarollo VG, Gertrudes JC, Oliveira PR, Honorio KM. *Applying machine learning techniques for ADME-Tox prediction: a review*. Expert Opin Drug Metab Toxicol. **2015** Feb;11(2):259-71.

[77] Schmidhuber J. *Deep learning in neural networks: an overview*. Neural Netw. **2015** Jan;61:85-117.

[78] Goh GB, Hodas NO, Vishnu A. *Deep learning for computational chemistry*. J Comput Chem. **2017** Jun 15;38(16):1291-1307.

[79] Gawehn E, Hiss JA, Brown JB, Schneider G. *Advancing drug discovery via GPU-based deep learning*. Expert Opin Drug Discov. **2018** Jul;13(7):579-582.

[80] Ghasemi F, Fassihi A, Pérez-Sánchez H, Mehri Dehnavi A. *The role of different sampling methods in improving biological activity prediction using deep belief network*. J Comput Chem. **2017** Feb 5;38(4):195-203.

[81] Holdgate G, Embrey K, Milbradt A, Davies G. *Biophysical methods in early drug discovery*. ADMET DMPK. **2019** Dec 11;7(4):222-241.

[82] Ma W, Yang L, He L. *Overview of the detection methods for equilibrium dissociation constant KD of drug-receptor interaction*. J Pharm Anal. **2018** Jun;8(3):147-152.

[83] Ferruz N, De Fabritiis G. *Binding Kinetics in Drug Discovery*. Mol Inform. **2016** Jul;35(6-7):216-26.

[84] Fang Y. *Ligand-receptor interaction platforms and their applications for drug discovery*. Expert Opin Drug Discov. **2012** Oct;7(10):969-88.

[85] Freyer MW, Lewis EA. *Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions*. Methods Cell Biol. **2008**;84:79-113.

[86] Olaru A, Bala C, Jaffrezic-Renault N, Aboul-Enein HY. *Surface plasmon resonance (SPR) biosensors in pharmaceutical analysis*. Crit Rev Anal Chem. **2015**;45(2):97-105.

[87] Zadran S, Standley S, Wong K, Otiniano E, Amighi A, Baudry M. *Fluorescence resonance energy transfer (FRET)-based biosensors: visualizing cellular dynamics and bioenergetics*. Appl Microbiol Biotechnol. **2012** Nov;96(4):895-902.

[88] El Deeb S, Al-Harrasi A, Khan A, Al-Broumi M, Al-Thani G, Alomairi M, Elumalai P, Sayed RA, Ibrahim AE. *Microscale thermophoresis as a powerful growing analytical technique for the investigation of biomolecular interaction and the determination of binding parameters*. Methods Appl Fluoresc. **2022** Jul 28;10(4).

[89] Holdgate GA. *Making cool drugs hot: isothermal titration calorimetry as a tool to study binding energetics*. Biotechniques. **2001** Jul;31(1):164-170.

[90] Falconer RJ. *Applications of isothermal titration calorimetry - the research and technical developments from 2011 to 2015*. J Mol Recognit. 2016 Oct;29(10):504-15.

[91] Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, Ji XL, Liu SQ. *Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods*. Int J Mol Sci. **2016** Jan 26;17(2):144.

[92] Jönsson U, Fägerstam L, Ivarsson B, Johnsson B, Karlsson R, Lundh K, Löfås S, Persson B, Roos H, Rönnberg I, et al. *Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology*. Biotechniques. **1991** Nov;11(5):620-627.

[93] Huber W, Mueller F. *Biomolecular interaction analysis in drug discovery using surface plasmon resonance technology.* Curr Pharm Des. **2006**;12(31):3999-4021.

[94] Patching SG. *Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery*. Biochim Biophys Acta. **2014** Jan;1838(1 Pt A):43-55.

[95] Su Z, Chen L, Liu Y, He X, Zhou Y, Xie Q, Yao S. *35 MHz quartz crystal microbalance and surface plasmon resonance studies on the binding of angiotensin converting enzyme with lisinopril*. Biosens Bioelectron. **2011** Mar 15;26(7):3240-5.

[96] Anraku K, Fukuda R, Takamune N, Misumi S, Okamoto Y, Otsuka M, Fujita M. *Highly sensitive analysis of the interaction between HIV-1 Gag and phosphoinositide derivatives based on surface plasmon resonance*. Biochemistry. **2010** Jun 29;49(25):5109-16.

[97] Schmidt J, Müsken M, Becker T, Magnowska Z, Bertinetti D, Möller S, Zimmermann B, Herberg FW, Jänsch L, Häussler S. *The Pseudomonas aeruginosa chemotaxis methyltransferase CheR1 impacts on bacterial surface sampling*. PLoS One. **2011** Mar 22;6(3):e18184.

[98] Hirata E, Kiyokawa E. *Future Perspective of Single-Molecule FRET Biosensors and Intravital FRET Microscopy*. Biophys J. **2016** Sep 20;111(6):1103-1111.

[99] Vogel SS, Thaler C, Koushik SV. *Fanciful FRET*. Sci STKE. **2006** Apr 18;331:re2.

[100] Shrestha D, Jenei A, Nagy P, Vereb G, Szöllősi J. *Understanding FRET as a research tool for cellular studies*. Int J Mol Sci. **2015** Mar 25;16(4):6718-56.

[101] Lu S, Wang Y. *Fluorescence resonance energy transfer biosensors for cancer detection and evaluation of drug efficacy*. Clin Cancer Res. **2010** Aug 1;16(15):3822-4.

[102] Cottet M, Faklaris O, Zwier JM, Trinquet E, Pin J-P, Durroux T. *Original Fluorescent Ligand-Based Assays Open New Perspectives in G-Protein Coupled Receptor Drug Screening*. Pharmaceuticals. **2011**; 4(1):202-214.

[103] Autry JM, Rubin JE, Pietrini SD, Winters DL, Robia SL, Thomas DD. *Oligomeric interactions of sarcolipin and the Ca-ATPase*. J Biol Chem. **2011** Sep 9;286(36):31697-706.

[104] Seidel SA, Dijkman PM, Lea WA, van den Bogaart G, Jerabek-Willemsen M, Lazic A, Joseph JS, Srinivasan P, Baaske P, Simeonov A, Katritch I, Melo FA, Ladbury JE, Schreiber G, Watts A, Braun D, Duhr S. *Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions*. Methods. **2013** Mar;59(3):301-15.

[105] Magnez R, Bailly C, Thuru X. *Microscale Thermophoresis as a Tool to Study Protein Interactions and Their Implication in Human Diseases*. Int J Mol Sci. **2022** Jul 12;23(14):7672.

[106] Venereau E, Schiraldi M, Uguccioni M, Bianchi ME. *HMGB1 and leukocyte migration during trauma and sterile inflammation.* Mol Immunol. **2013** Aug;55(1):76-82.

[107] Venereau E, De Leo F, Mezzapelle R, Careccia G, Musco G, Bianchi ME. *HMGB1 as biomarker and drug target*. Pharmacol Res. **2016** Sep;111:534-544.

[108] Naglova H, Bucova M. *HMGB1 and its physiological and pathological roles*. Bratisl Lek Listy. **2012**;113(3):163-71.

[109] He SJ, Cheng J, Feng X, Yu Y, Tian L, Huang Q. *The dual role and therapeutic potential of high-mobility group box 1 in cancer*. Oncotarget. **2017** May 16;8(38):64534-64550.

[110] Andersson U, Yang H, Harris H. *High-mobility group box 1 protein (HMGB1) operates as an alarmin outside as well as inside cells*. Semin Immunol. **2018** Aug;38:40-48.

[111] Andersson U, Tracey KJ. *HMGB1 is a therapeutic target for sterile inflammation and infection*. Annu Rev Immunol. **2011**;29:139-62.

[112] Janko C, Filipović M, Munoz LE, Schorn C, Schett G, Ivanović-Burmazović I, Herrmann M. *Redox modulation of HMGB1-related signaling*. Antioxid Redox Signal. **2014** Mar 1;20(7):1075-85.

[113] Bianchi ME, Crippa MP, Manfredi AA, Mezzapelle R, Rovere Querini P, Venereau E. *High-mobility group box 1 protein orchestrates responses to tissue damage via inflammation, innate and adaptive immunity, and tissue repair*. Immunol Rev. **2017** Nov;280(1):74-82.

[114] Venereau E, Casalgrandi M, Schiraldi M, Antoine DJ, Cattaneo A, De Marchis F, Liu J, Antonelli A, Preti A, Raeli L, Shams SS, Yang H, Varani L, Andersson U, Tracey KJ, Bachi A, Uguccioni M, Bianchi ME. *Mutually exclusive redox forms of HMGB1 promote cell recruitment or proinflammatory cytokine release*. J Exp Med. **2012** Aug 27;209(9):1519-28.

[115] Schrader AJ, Lechner O, Templin M, Dittmar KE, Machtens S, Mengel M, Probst-Kepper M, Franzke A, Wollensak T, Gatzlaff P, Atzpodien J, Buer J, Lauber J. *CXCR4/CXCL12 expression and signalling in kidney cancer*. Br J Cancer. **2002** Apr 22;86(8):1250-6.

[116] Schiraldi M, Raucci A, Muñoz LM, Livoti E, Celona B, Venereau E, Apuzzo T, De Marchis F, Pedotti M, Bachi A, Thelen M, Varani L, Mellado M, Proudfoot A, Bianchi ME, Uguccioni M. *HMGB1 promotes recruitment of inflammatory cells to damaged tissues by forming a complex with CXCL12 and signaling via CXCR4*. J Exp Med. **2012** Mar 12;209(3):551-63.

[117] Fassi EMA, Sgrignani J, D'Agostino G, Cecchinato V, Garofalo M, Grazioso G, Uguccioni M, Cavalli A. *Oxidation State Dependent Conformational Changes of HMGB1 Regulate the Formation of the CXCL12/HMGB1 Heterocomplex*. Comput Struct Biotechnol J. **2019** Jun 21;17:886-894.

[118] Yang H, Pellegrini L, Napolitano A, Giorgi C, Jube S, Preti A, Jennings CJ, De Marchis F, Flores EG, Larson D, Pagano I, Tanji M, Powers A, Kanodia S, Gaudino G, Pastorino S, Pass HI, Pinton P, Bianchi ME, Carbone M. *Aspirin delays mesothelioma growth by inhibiting HMGB1-mediated tumor progression*. Cell Death Dis. **2015** Jun 11;6(6):e1786.

[119] De Leo F, Quilici G, Tirone M, De Marchis F, Mannella V, Zucchelli C, Preti A, Gori A, Casalgrandi M, Mezzapelle R, Bianchi ME, Musco G. *Diflunisal targets the HMGB1/CXCL12 heterocomplex and blocks immune cell recruitment*. EMBO Rep. **2019** Oct 4;20(10):e47788.

[120] De Leo F, Quilici G, De Marchis F, Mantonico MV, Bianchi ME, Musco G. *Discovery of 5,5'-Methylenedi-2,3-Cresotic Acid as a Potent Inhibitor of the Chemotactic Activity of the HMGB1·CXCL12 Heterocomplex Using Virtual Screening and NMR Validation*. Front Chem. **2020** Nov 26;8:598710.

[121] Mollica L, De Marchis F, Spitaleri A, Dallacosta C, Pennacchini D, Zamai M, Agresti A, Trisciuoglio L, Musco G, Bianchi ME. *Glycyrrhizin binds to high-mobility group box 1 protein and inhibits its cytokine activities*. Chem Biol. **2007** Apr;14(4):431-41.

[122] Rastelli G, Del Rio A, Degliesposti G, Sgobba M. *Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA*. J Comput Chem. **2010** Mar;31(4):797-810.

[123] Grazioso G, Pomè DY, Matera C, Frigerio F, Pucci L, Gotti C, Dallanoce C, De Amici M. *Design of novel alpha7-subtype-preferring nicotinic acetylcholine receptor agonists: application of docking and MM-PBSA computational approaches, synthetic and pharmacological studies*. Bioorg Med Chem Lett. **2009** Nov 15;19(22):6353-7.

[124] Yang H, Wang H, Ju Z, Ragab AA, Lundbäck P, Long W, Valdes-Ferrer SI, He M, Pribis JP, Li J, Lu B, Gero D, Szabo C, Antoine DJ, Harris HE, Golenbock DT, Meng J, Roth J, Chavan SS, Andersson U, Billiar TR, Tracey KJ, Al-Abed Y. *MD-2 is required for disulfide HMGB1-dependent TLR4 signaling*. J Exp Med. **2015** Jan 12;212(1):5-14.

[125] Arranz-Gibert P, Ciudad S, Seco J, García J, Giralt E, Teixidó M. *Immunosilencing peptides by stereochemical inversion and sequence reversal: retro-D-peptides*. Sci Rep. **2018** Apr 24;8(1):6446.

[126] Liu M, Li X, Xie Z, Xie C, Zhan C, Hu X, Shen Q, Wei X, Su B, Wang J, Lu W. *D-Peptides as Recognition Molecules and Therapeutic Agents*. Chem Rec. **2016** Aug;16(4):1772-86.

[127] van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bonvin AMJJ. *The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes*. J Mol Biol. **2016** Feb 22;428(4):720-725.

[128] Sgrignani J, Cecchinato V, Fassi EMA, D'Agostino G, Garofalo M, Danelon G, Pedotti M, Simonelli L, Varani L, Grazioso G, Uguccioni M, Cavalli A. *Systematic Development of Peptide Inhibitors Targeting the CXCL12/HMGB1 Interaction*. J Med Chem. **2021** Sep 23;64(18):13439-13450.

[129] Yang H, Hreggvidsdottir HS, Palmblad K, Wang H, Ochani M, Li J, Lu B, Chavan S, Rosas-Ballina M, Al-Abed Y, Akira S, Bierhaus A, Erlandsson-Harris H, Andersson U, Tracey KJ. *A critical cysteine is required for HMGB1 binding to Toll-like receptor 4 and activation of macrophage cytokine release*. Proc Natl Acad Sci U S A. **2010** Jun 29;107(26):11942-7.

[130] Nelson RH. *Hyperlipidemia as a risk factor for cardiovascular disease*. Prim Care. **2013** Mar;40(1):195-211.

[131] Goldstein JL, Brown MS. *Regulation of the mevalonate pathway*. Nature. **1990** Feb 1;343(6257):425-30.

[132] Alvi SS, Ansari IA, Ahmad MK, Iqbal J, Khan MS. *Lycopene amends LPS induced oxidative stress and hypertriglyceridemia via modulating PCSK-9 expression and Apo-CIII mediated lipoprotein lipase activity*. Biomed Pharmacother. **2017** Dec;96:1082-1093.

[133] Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, Volkert TL, Schreiber J, Rolfe PA, Gifford DK, Fraenkel E, Bell GI, Young RA. *Control of pancreas and liver gene expression by HNF transcription factors*. Science. **2004** Feb 27;303(5662):1378-81.

[134] Dong B, Singh AB, Shende VR, Liu J. *Hepatic HNF1 transcription factors control the induction of PCSK9 mediated by rosuvastatin in normolipidemic hamsters*. Int J Mol Med. **2017** Mar;39(3):749-756.

[135] Seidah NG, Awan Z, Chrétien M, Mbikay M. *PCSK9: a key modulator of cardiovascular health*. Circ Res. **2014** Mar 14;114(6):1022-36.

[136] Kwon HJ, Lagace TA, McNutt MC, Horton JD, Deisenhofer J. *Molecular basis for LDL receptor recognition by PCSK9*. Proc Natl Acad Sci USA. **2008** Feb 12;105(6):1820-5.

[137] Xu S, Luo S, Zhu Z, Xu J. *Small molecules as inhibitors of PCSK9: Current status and future challenges*. Eur J Med Chem. **2019** Jan 15;162:212-233.

[138] Xia XD, Peng ZS, Gu HM, Wang M, Wang GQ, Zhang DW. *Regulation of PCSK9 Expression and Function: Mechanisms and Therapeutic Implications*. Front Cardiovasc Med. **2021** Oct 15;8:764038.

[139] Xie H, Yang K, Winston-McPherson GN, Stapleton DS, Keller MP, Attie AD, Smith KA, Tang W. *From methylene bridged diindole to carbonyl linked benzimidazoleindole: Development of potent and metabolically stable PCSK9 modulators*. Eur J Med Chem. **2020** Nov 15;206:112678.

[140] Lammi C, Zanoni C, Aiello G, Arnoldi A, Grazioso G. *Lupin Peptides Modulate the Protein-Protein Interaction of PCSK9 with the Low Density Lipoprotein Receptor in HepG2 Cells*. Sci Rep. **2016** Jul 18;6:29931.

[141] Schroeder CI, Swedberg JE, Withka JM, Rosengren KJ, Akcan M, Clayton DJ, Daly NL, Cheneval O, Borzilleri KA, Griffor M, Stock I, Colless B, Walsh P, Sunderland P, Reyes A, Dullea R, Ammirati M, Liu S, McClure KF, Tu M, Bhattacharya SK, Liras S, Price DA, Craik DJ. *Design and synthesis of truncated EGF-A peptides that restore LDL-R recycling in the presence of PCSK9 in vitro*. Chem Biol. **2014** Feb 20;21(2):284-94.

[142] Lammi C, Aiello G, Bollati C, Li J, Bartolomei M, Ranaldi G, Ferruzza S, Fassi EMA, Grazioso G, Sambuy Y, Arnoldi A. *Trans-Epithelial Transport, Metabolism, and Biological Activity Assessment of the Multi-Target Lupin Peptide LILPKHSDAD (P5) and Its Metabolite LPKHSDAD (P5-Met)*. Nutrients. **2021** Mar 5;13(3):863.

[143] Lammi C, Sgrignani J, Arnoldi A, Lesma G, Spatti C, Silvani A, Grazioso G. *Computationally Driven Structure Optimization, Synthesis, and Biological Evaluation of Imidazole-Based Proprotein Convertase Subtilisin/Kexin 9 (PCSK9) Inhibitors*. J Med Chem. **2019** Jul 11;62(13):6163-6174.

[144] Drucker DJ. *Advances in oral peptide therapeutics*. Nat Rev Drug Discov. **2020** Apr;19(4):277-289.

[145] Zhang Y, Eigenbrot C, Zhou L, Shia S, Li W, Quan C, Tom J, Moran P, Di Lello P, Skelton NJ, Kong-Beltran M, Peterson A, Kirchhofer D. *Identification of a small peptide that inhibits PCSK9 protein binding to the low density lipoprotein receptor*. J Biol Chem. **2014** Jan 10;289(2):942-55.

[146] Lammi C, Sgrignani J, Arnoldi A, Grazioso G. *Biological Characterization of Computationally Designed Analogs of peptide TVFTSWEEYLDWV (Pep2-8) with Increased PCSK9 Antagonistic Activity*. Sci Rep. **2019** Feb 20;9(1):2343.

[147] Tombling BJ, Lammi C, Lawrence N, Gilding EK, Grazioso G, Craik DJ, Wang CK. *Bioactive Cyclization Optimizes the Affinity of a Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Peptide Inhibitor*. J Med Chem. **2021** Mar 11;64(5):2523-2533.

[148] Olsson AG, McTaggart F, Raza A. *Rosuvastatin: a highly effective new HMG-CoA reductase inhibitor*. Cardiovasc Drug Rev. **2002** Winter;20(4):303-28.

[149] McTaggart F, Buckett L, Davidson R, Holdgate G, McCormick A, Schneck D, Smith G, Warwick M. *Preclinical and clinical pharmacology of Rosuvastatin, a new 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitor*. Am J Cardiol. **2001** Mar 8;87(5A):28B-32B.

[150] Lammi C, Zanoni C, Arnoldi A, Aiello G. *YDFYPSSTKDQQS (P3), a peptide from lupin protein, absorbed by Caco-2 cells, modulates cholesterol metabolism in HepG2 cells via SREBP-1 activation*. Journal of Food Biochemistry. **2019**;43(1):e12757.

[151] Istvan ES, Deisenhofer J. *Structural mechanism for statin inhibition of HMG-CoA reductase*. Science. **2001** May 11;292(5519):1160-4.

[152] Lamb YN. *Inclisiran: First Approval*. Drugs. **2021** Feb;81(3):389-395.

[153] Tombling BJ, Zhang Y, Huang YH, Craik DJ, Wang CK. *The emerging landscape of peptide-based inhibitors of PCSK9*. Atherosclerosis. **2021** Aug;330:52-60.

[154] Lammi C, Fassi EMA, Li J, Bartolomei M, Benigno G, Roda G, Arnoldi A, Grazioso G. *Computational Design and Biological Evaluation of Analogs of Lupin Peptide P5 Endowed with Dual PCSK9/HMG-CoAR Inhibiting Activity*. Pharmaceutics. **2022** Mar 18;14(3):665.

[155] Tucker TJ, Embrey MW, Alleyne C, Amin RP, Bass A, Bhatt B, Bianchi E, Branca D, Bueters T, Buist N, Ha SN, Hafey M, He H, Higgins J, Johns DG, Kerekes AD, Koeplinger KA, Kuethe JT, Li N, Murphy B,

Orth P, Salowe S, Shahripour A, Tracy R, Wang W, Wu C, Xiong Y, Zokian HJ, Wood HB, Walji A. *A Series of Novel, Highly Potent, and Orally Bioavailable Next-Generation Tricyclic Peptide PCSK9 Inhibitors*. J Med Chem. **2021** Nov 25;64(22):16770-16800.

[156] Sun H, Wang J, Liu S, Zhou X, Dai L, Chen C, Xu Q, Wen X, Cheng K, Sun H, Yuan H. *Discovery of Novel Small Molecule Inhibitors Disrupting the PCSK9-LDLR Interaction.* J Chem Inf Model. **2021** Oct 25;61(10):5269-5279.

[157] Min DK, Lee HS, Lee N, Lee CJ, Song HJ, Yang GE, Yoon D, Park SW. *In silico Screening of Chemical Libraries to Develop Inhibitors That Hamper the Interaction of PCSK9 with the LDL Receptor*. Yonsei Med J. **2015** Sep;56(5):1251-7.

[158] Evison BJ, Palmer JT, Lambert G, Treutlein H, Zeng J, Nativel B, Chemello K, Zhu Q, Wang J, Teng Y, Tang W, Xu Y, Rathi AK, Kumar S, Suchowerska AK, Parmar J, Dixon I, Kelly GE, Bonnar J. *A small molecule inhibitor of PCSK9 that antagonizes LDL receptor binding via interaction with a cryptic PCSK9 binding groove*. Bioorg Med Chem. **2020** Mar 15;28(6):115344.

[159] Bottomley MJ, Cirillo A, Orsatti L, Ruggeri L, Fisher TS, Santoro JC, Cummings RT, Cubbon RM, Lo Surdo P, Calzetta A, Noto A, Baysarowich J, Mattu M, Talamo F, De Francesco R, Sparrow CP, Sitlani A, Carfí A. *Structural and biochemical characterization of the wild type PCSK9-EGF(AB) complex and natural familial hypercholesterolemia mutants*. J Biol Chem. **2009** Jan 9;284(2):1313-23.

[160] Stucchi M, Grazioso G, Lammi C, Manara S, Zanoni C, Arnoldi A, Lesma G, Silvani A. *Disrupting the PCSK9/LDLR protein-protein interaction by an imidazole-based minimalist peptidomimetic*. Org Biomol Chem. **2016** Oct 18;14(41):9736-9740.

[161] Talele TT. *Acetylene Group, Friend or Foe in Medicinal Chemistry*. J Med Chem. **2020** Jun 11;63(11):5625-5663.

[162] Wilcken R, Zimmermann MO, Bauer MR, Rutherford TJ, Fersht AR, Joerger AC, Boeckler FM. *Experimental and Theoretical Evaluation of the Ethynyl Moiety as a Halogen Bioisostere*. ACS Chem Biol. **2015** Dec 18;10(12):2725-32.

[163] Xie Z, Klionsky DJ. *Autophagosome formation: core machinery and adaptations*. Nat Cell Biol. **2007** Oct;9(10):1102-9.

[164] Jacquet M, Guittaut M, Fraichard A, Despouy G. *The functions of Atg8-family proteins in autophagy and cancer: linked or unrelated?* Autophagy. **2021** Mar;17(3):599-611.

[165] Schaaf MB, Keulers TG, Vooijs MA, Rouschop KM. *LC3/GABARAP family proteins: autophagy-(un)related functions*. FASEB J. **2016** Dec;30(12):3961-3978.

[166] Wright TJ, McKee C, Birch-Machin MA, Ellis R, Armstrong JL, Lovat PE. *Increasing the therapeutic efficacy of docetaxel for cutaneous squamous cell carcinoma through the combined inhibition of phosphatidylinositol 3-kinase/AKT signalling and autophagy*. Clin Exp Dermatol. **2013** Jun;38(4):421-3.

[167] Rosenfeld MR, Ye X, Supko JG, Desideri S, Grossman SA, Brem S, Mikkelson T, Wang D, Chang YC, Hu J, McAfee Q, Fisher J, Troxel AB, Piao S, Heitjan DF, Tan KS, Pontiggia L, O'Dwyer PJ, Davis LE, Amaravadi RK. *A phase I/II trial of hydroxychloroquine in conjunction with radiation therapy and concurrent and adjuvant temozolomide in patients with newly diagnosed glioblastoma multiforme*. Autophagy. **2014** Aug;10(8):1359-68.

[168] Kaini RR, Sillerud LO, Zhaorigetu S, Hu CA. *Autophagy regulates lipolysis and cell survival through lipid droplet degradation in androgen-sensitive prostate cancer cells*. Prostate. **2012** Sep 15;72(13):1412-22.

[169] He W, Wang Q, Xu J, Xu X, Padilla MT, Ren G, Gou X, Lin Y. *Attenuation of TNFSF10/TRAIL-induced apoptosis by an autophagic survival pathway involving TRAF2- and RIPK1/RIP1-mediated MAPK8/JNK activation*. Autophagy. **2012** Dec;8(12):1811-21.

[170] Gabriele F, Martinelli C, Comincini S. *Prostate cancer cells at a therapeutic gunpoint of the autophagy process*. J Cancer Metastasis Treat. **2018**;4:17.

[171] Engedal N, Seglen PO. *Autophagy of cytoplasmic bulk cargo does not require LC3*. Autophagy. **2016**;12(2):439-41.

[172] Weiergräber OH, Stangler T, Thielmann Y, Mohrlüder J, Wiesehan K, Willbold D. *Ligand binding mode of GABAA receptor-associated protein*. J Mol Biol. **2008** Sep 19;381(5):1320-31.

[173] Johansen T, Lamark T. *Selective Autophagy: ATG8 Family Proteins, LIR Motifs and Cargo Receptors*. J Mol Biol. **2020** Jan 3;432(1):80-103.

[174] Li J, Zhu R, Chen K, Zheng H, Zhao H, Yuan C, Zhang H, Wang C, Zhang M. *Potent and specific Atg8-targeting autophagy inhibitory peptides from giant ankyrins*. Nat Chem Biol. **2018** Aug;14(8):778-787.

[175] Suzuki H, Tabata K, Morita E, Kawasaki M, Kato R, Dobson RC, Yoshimori T, Wakatsuki S. *Structural basis of the autophagy-related LC3/Atg13 LIR complex: recognition and interaction mechanism*. Structure. **2014** Jan 7;22(1):47-58.

[176] Huber J, Obata M, Gruber J, Akutsu M, Löhr F, Rogova N, Güntert P, Dikic I, Kirkin V, Komatsu M, Dötsch V, Rogov VV. *An atypical LIR motif within UBA5 (ubiquitin like modifier activating enzyme 5) interacts with GABARAP proteins and mediates membrane localization of UBA5*. Autophagy. **2020** Feb;16(2):256-270.

[177] Siegel RL, Miller KD, Fuchs HE, Jemal A. *Cancer Statistics, 2021*. CA Cancer J Clin. **2021** Jan;71(1):7-33.

[178] Farrow JM, Yang JC, Evans CP. *Autophagy as a modulator and target in prostate cancer*. Nat Rev Urol. **2014** Sep;11(9):508-16.

[179] Lee YR, Chen M, Lee JD, Zhang J, Lin SY, Fu TM, Chen H, Ishikawa T, Chiang SY, Katon J, Zhang Y, Shulga YV, Bester AC, Fung J, Monteleone E, Wan L, Shen C, Hsu CH, Papa A, Clohessy JG, Teruya-Feldstein J, Jain S, Wu H, Matesic L, Chen RH, Wei W, Pandolfi PP. *Reactivation of PTEN tumor suppressor for cancer treatment through inhibition of a MYC-WWP1 inhibitory pathway*. Science. **2019** May 17;364(6441):eaau0159.

[180] Zhang W, Wu KP, Sartori MA, Kamadurai HB, Ordureau A, Jiang C, Mercredi PY, Murchie R, Hu J, Persaud A, Mukherjee M, Li N, Doye A, Walker JR, Sheng Y, Hao Z, Li Y, Brown KR, Lemichez E, Chen J, Tong Y, Harper JW, Moffat J, Rotin D, Schulman BA, Sidhu SS. *System-Wide Modulation of HECT E3 Ligases with Selective Ubiquitin Variant Probes*. Mol Cell. **2016** Apr 7;62(1):121-36.

[181] Kathman SG, Span I, Smith AT, Xu Z, Zhan J, Rosenzweig AC, Statsyuk AV. *A Small Molecule That Switches a Ubiquitin Ligase From a Processive to a Distributive Enzymatic Mechanism.* J Am Chem Soc. **2015** Oct 7;137(39):12442-5.

[182] Park SY, Lee SW, Kim HY, Lee WS, Hong KW, Kim CD. *HMGB1 induces angiogenesis in rheumatoid arthritis via HIF-1α activation*. Eur J Immunol. **2015** Apr;45(4):1216-27.

[183] Cecchinato V, D'Agostino G, Raeli L, Nerviani A, Schiraldi M, Danelon G, Manzo A, Thelen M, Ciurea A, Bianchi ME, Rubartelli A, Pitzalis C, Uguccioni M. *Redox-Mediated Mechanisms Fuel Monocyte Responses to CXCL12/HMGB1 in Active Rheumatoid Arthritis*. Front Immunol. **2018** Sep 19;9:2118.

[184] Lu M, Yu S, Xu W, Gao B, Xiong S. *HMGB1 Promotes Systemic Lupus Erythematosus by Enhancing Macrophage Inflammatory Response*. J Immunol Res. **2015**;2015:946748.

[185] Oktayoglu P, Em S, Tahtasiz M, Bozkurt M, Ucar D, Yazmalar L, Nas K, Yardımeden I, Cevik F, Celik Y, Mete N. *Elevated serum levels of high mobility group box protein 1 (HMGB1) in patients with ankylosing spondylitis and its association with disease activity and quality of life*. Rheumatol Int. **2013** May;33(5):1327-31.

[186] Wan Z, Zhang X, Peng A, He M, Lei Z, Wang Y. *TLR4-HMGB1 signaling pathway affects the inflammatory reaction of autoimmune myositis by regulating MHC-I*. Int Immunopharmacol. **2016** Dec;41:74-81.

[187] Sun Y, Chen H, Dai J, Zou H, Gao M, Wu H, Ming B, Lai L, Xiao Y, Xiong P, Xu Y, Gong F, Zheng F. *HMGB1 expression patterns during the progression of experimental autoimmune encephalomyelitis*. J Neuroimmunol. **2015** Mar 15;280:29-35.

[188] Pitzalis C, Kelly S, Humby F. *New learnings on the pathophysiology of RA from synovial biopsies*. Curr Opin Rheumatol. **2013** May;25(3):334-44.

[189] Smolen JS, Aletaha D, Barton A, Burmester GR, Emery P, Firestein GS, Kavanaugh A, McInnes IB, Solomon DH, Strand V, Yamamoto K. *Rheumatoid arthritis*. Nat Rev Dis Primers. **2018** Feb 8;4:18001.

[190] Fassi EMA, Garofalo M, Sgrignani J, Dei Cas M, Mori M, Roda G, Cavalli A, Grazioso G. *Focused Design of Novel Cyclic Peptides Endowed with GABARAP-Inhibiting Activity*. Int J Mol Sci. **2022** May 3;23(9):5070.

[191] Amaravadi RK, Kimmelman AC, Debnath J. *Targeting Autophagy in Cancer: Recent Advances and Future Directions*. Cancer Discov. **2019** Sep;9(9):1167-1181.

[192] Lo YC, Rensi SE, Torng W, Altman RB. *Machine learning in chemoinformatics and drug discovery*. Drug Discov Today. **2018** Aug;23(8):1538-1546.

7. ANNEXES

Annex 1. Results of the affinity prediction for the 57 nonapeptides of HMGB1 selected after docking calculations. The unbound states indicate that the peptide did not conserve its original binding mode over the simulation time. The peptides are ranked by ΔG.

Code	Sequence	ΔG ± SE (kcal/mol)	Code	Sequence	ΔG ± SE (kcal/mol)
HBP01	HEMYWEDEW	-52.8 ± 0.3	44	FVGMRWKFL	*unbound*
HBP02	IDLRFFMRQ	-52.0 ± 0.3	45	WQIPDHRDH	*unbound*
HBP03	FAFELIQTD	-51.7 ± 0.4	46	QCFHPSFED	*unbound*
HBP04	CIPMMMHAW	-50.0 ± 0.3	47	VPSSAKNRD	*unbound*
HBP05	WISNWILMW	-45.8 ± 0.3	48	KHMTKCEQW	*unbound*
HBP06	TWNIHFADH	-45.6 ± 0.4	49	ETYQFRPNK	*unbound*
HBP07	HWTLANWCR	-45.2 ± 0.4	50	WNCHRDRPK	*unbound*
HBP08	GYHYERWIH	-45.1 ± 0.5	51	KHMTKCEQW	*unbound*
HBP09	QFMKNCEEM	-44.8 ± 0.4	52	KCVVFHYDP	*unbound*
HBP10	SINWHMYVN	-44.8 ± 0.3	53	PTFEEFAAF	*unbound*
HBP11	MYRENQPTR	-42.9 ± 0.4	54	QCFHPSFED	*unbound*
HBP12	YHICWYGDY	-42.5 ± 0.5	55	EWLYRQEYH	*unbound*
HBP13	WLWYEWGWQ	-41.9 ± 0.3	56	QDYAPRASN	*unbound*
14	DYCWKIMTQ	-41.9 ± 0.3	57	KDKAFKNVS	*unbound*
15	WCHFFFPHW	-41.6 ± 0.4			
16	MKSSDCCLE	-39.7 ± 0.5			
17	EWFVMKHLN	-39.0 ± 0.4			
18	MIRDQILHN	-38.9 ± 0.4			
19	WHQLTEHWI	-38.2 ± 0.5			
20	HDHDFWAWY	-37.5 ± 0.2			
21	WQWHQFQGR	-35.9 ± 0.3			
22	VMASWQHGL	-34.9 ± 0.5			
23	LDNFLGDHW	-34.7 ± 0.4			
24	PRMGWEKPE	-34.1 ± 0.4			
25	WICVWHHAS	-33.9 ± 0.3			
26	IRWCVDARY	-30.0 ± 0.6			
27	WNAMSFCCS	-28.9 ± 0.4			
28	IFHIMTEMW	-28.7 ± 0.2			
29	FDRPRYRTT	-28.6 ± 0.4			
30	QIEDMPTSK	-28.3 ± 0.4			
31	FDCMMDMTK	-28.0 ± 0.3			
32	NTVALKLRD	-27.8 ± 0.4			
33	YHYHMLMQS	-27.7 ± 0.3			
34	NITHNVWHR	-27.6 ± 0.3			
35	DRNLEVEQI	-26.7 ± 0.3			
36	HYNKWKHQE	-25.6 ± 0.4			
37	ICMPPNTKN	-24.6 ± 0.3			
38	SMIPVQEAS	-24.5 ± 0.3			
39	YQRNELEYL	-24.4 ± 0.2			
40	HYFDMLHFH	-21.0 ± 0.4			
41	SHYFKHSNF	-19.8 ± 0.4			
42	FIKQMEEST	-18.5 ± 0.3			
43	KYQWMHYTP	-16.9 ± 0.6			

Annex 2. Results of the affinity prediction for the 100 tetrapeptides of WWP1 selected after docking calculations. The unbound states indicate that the peptide did not conserve its original binding mode over the simulation time.

Peptide	ΔG ± SE (kcal/mol)	Peptide	ΔG ± SE (kcal/mol)	Peptide	ΔG ± SE (kcal/mol)
01	-27.5 ± 0.3	46	-24.5 ± 0.2	91	*unbound*
02	-20.5 ± 0.2	47	-23.5 ± 0.2	92	*unbound*
03	*unbound*	48	*unbound*	93	-23.3 ± 0.3
04	-15.4 ± 0.5	49	*unbound*	94	-22.9 ± 0.4
05	*unbound*	50	*unbound*	95	*unbound*
06	-23.9 ± 0.3	51	*unbound*	96	*unbound*
07	-28.0 ± 0.2	52	-20.1 ± 0.2	97	-22.5 ± 0.2
08	-17.6 ± 0.2	53	-24.2 ± 0.3	98	-32.4 ± 0.4
09	-26.4 ± 0.3	54	-19.9 ± 0.4	99	-38.4 ± 0.2
10	*unbound*	55	*unbound*	100	-21.2 ± 0.2
11	-32.0 ± 0.2	56	-20.3 ± 0.3		
12	-30.3 ± 0.1	57	*unbound*		
13	-22.0 ± 0.3	58	-17.4 ± 0.4		
14	-37.7 ± 0.3	59	-18.2 ± 0.3		
15	-21.0 ± 0.2	60	-22.0 ± 0.5		
16	*unbound*	61	-22.2 ± 0.2		
17	*unbound*	62	*unbound*		
18	-29.3 ± 0.4	63	*unbound*		
19	-34.9 ± 0.3	64	-30.5 ± 0.2		
20	-8.3 ± 0.2	65	*unbound*		
21	-17.6 ± 0.2	66	*unbound*		
22	-26.3 ± 0.3	67	-13.3 ± 0.3		
23	-17.3 ± 0.2	68	-20.2 ± 0.3		
24	*unbound*	69	*unbound*		
25	-17.7 ± 0.2	70	-25.2 ± 0.3		
26	*unbound*	71	-21.3 ± 0.4		
27	-18.5 ± 0.3	72	-39.2 ± 0.4		
28	*unbound*	73	-23.2 ± 0.3		
29	-15.9 ± 0.2	74	-20.7 ± 0.5		
30	*unbound*	75	-23.2 ± 0.2		
31	*unbound*	76	-24.3 ± 0.4		
32	-17.8 ± 0.2	77	-26.2 ± 0.3		
33	*unbound*	78	*unbound*		
34	-22.4 ± 0.2	79	-24.5 ± 0.2		
35	*unbound*	80	*unbound*		
36	*unbound*	81	-26.5 ± 0.3		
37	-39.3 ± 0.3	82	-15.2 ± 0.3		
38	*unbound*	83	-20.3 ± 0.2		
39	-19.9 ± 0.3	84	*unbound*		
40	-23.9 ± 0.2	85	*unbound*		
41	-27.7 ± 0.3	86	-23.6 ± 0.3		
42	*unbound*	87	-30.3 ± 0.3		
43	-40.3 ± 0.2	88	-22.6 ± 0.3		
44	-19.4 ± 0.1	89	-17.6 ± 0.3		
45	-21.4 ± 0.2	90	*unbound*		