UNIVERSITÀ DEGLI STUDI DI MILANO

*PhD course in Pharmaceutical Sciences XXXVIII*

*Department of Pharmaceutical Sciences*

TESI DI DOTTORATO DI RICERCA

# DESIGN AND DEVELOPMENT OF RNA-BASED DRUGS USING MOLECULAR DYNAMICS SIMULATIONS AND ARTIFICIAL INTELLIGENCE METHODS

R29

Progetto finanziato con il contributo dell'Unione europea - Next Generation EU

CUP n.G43C22001320007, CN3, spoke 7, National Center for Gene Therapy and Drugs based on RNA Technology

Marco Albani

Matr. R14147

ORCID n. 0009-0003-0093-0945

TUTOR: Prof. Giovanni Grazioso

CO-TUTOR: Prof. Gabriella Roda

COORDINATORE DEL DOTTORATO: prof. Giulio Vistoli

A.A.

(ANNO ACCADEMICO 2025/2026)

# Index

# 1. INTRODUCTION

RNA-based therapeutics have emerged as a powerful class of drugs that offer significant advantages for targeting pathways and targets with high precision, previously impossible with common small molecules. RNA-derived drugs have shown promise in treating multiple conditions, including genetic disorders, cancers, and infectious diseases [1]. With significant structural and functional discoveries and technological advances, RNA-based drugs are achieving their true therapeutic potential [2]. RNA molecules as therapeutic agents are more easily screened, synthesized, and modified for precise adjustments, if necessary, thus demonstrating significant potential in modern therapeutics [3].

The advent of RNA-based drugs has introduced numerous advantages over traditional pharmaceuticals, including cost-effective screening and synthesis processes, which substantially expand the range of targets. These drugs offer flexibility in design and customization for personalized therapeutics, enabling rapid screening and development of sequences against emerging pathogens [4].

Key milestones include the development of antisense oligonucleotides (ASOs) in the early 1980s and the proposal of RNA interference (RNAi) in the 2000s [2, 5], as well as the first-ever approval of a CRISPR-Cas9-based therapy, Casgevy. The recent approval of an aptamer, Izervay (avacincaptad pegol), further underscores the versatility and expanding horizons of NA therapeutics. These advancements collectively highlight the rapid evolution and transformative impact of NA-based treatments in modern medicine [4].

Numerous RNA-based therapeutic molecules have received approval from the US Food and Drug Administration (FDA) and have progressed into clinical trials, as outlined in **Table 1**.

**Table 1**: List of approved Nucleic acid drugs

| RNA type | Drug name(s) | Target profile | Chemical modification and carrier | Year approved | Reference |
|---|---|---|---|---|---|
| ASO | Fomivirsen (Vitravene, ISIS 2922) | Eye, cytomegalovirus (CMV) retinitis: Bind and degrade UL123 mRNA and inhibit CMV protein IE2 | 2′-H | 1998, FDA, EMA<br>Withdrawn 2002 | 6 |
| | Mipomersen (Kynamro™, ISIS 301012) | Liver, homozygous familial hypercholesterolemia, binds and degrades ApoB-100 mRNA | Gapmers, 2′-MOE | 2013 | 7 |
| | Nusinersen(Spinraza, ISIS 396443) | CNS, spinal muscular atrophy: splice the pre-mRNA of SMN2 | 2′-MOE | 2016 | 8 |
| | Eteplirsen (Exondys 51, AVI-4658) | Muscles, Duchenne muscular dystrophy: Splice exon 52 of DMD | 2′-MOE, PMO | 2016 | 9 |
| | Inotersen (Tegsedi, ISIS 420915) | Liver, familial amyloid polyneuropathy: Inhibit translation of TTR mRNA translation | 2′-MOE, | 2018 | 10 |
| | Milasen | CNS, Mila Makovec's CLN7 gene is associated with Batten disease: Splice the CLN7 mRNA | – | 2018, Approved for personalized use | 3 |
| | Golodirsen (Vyondys 53™, SRP-4053) | Muscles, Duchenne muscular dystrophy: Splice exon 53 of DMD | 2′-MOE, PMO | 2019 | 11 |
| | Viltolarsen (Viltepso, NS-065, NCNP-01) | Muscles, Duchenne muscular dystrophy: Splice exon 53 of DMD | 2′-MOE, PMO | 2020 | 12 |
| | Casimersen (SRP-4045, Amondys 45™) | Muscles, Duchenne muscular dystrophy: Splice exon 45 of DMD | PMO | 2021 | 13 |
| | Eplontersen (Wainua) | Muscles, polyneuropathy of hereditary transthyretin-mediated amyloidosis | GalNAc-conjugated | 2023 | 14 |
| siRNA | Patisiran (Onpattro, ALN-TTR02, ONPATTRO™) | Liver, polyneuropathy: Inhibit translation of TTR mRNA translation | PS, 2′-O-Me, 2′-F, LNP | 2018 | 15 |
| | Givosiran (Givlaari, ALN-AS1) | Liver, acute hepatic porphyria: Targeting ALAS1 mRNA and downregulation | PS, 2′-O-Me, 2′-F, GalNAc | 2019 | 16 |

| | | | | | |
|---|---|---|---|---|---|
| | Lumasiran (OXLUMO, ALNGO1) | Liver, primary hyperoxaluria type 1: Targeting HAO1 mRNA and downregulation of glycolate oxidase | PS, 2′-O-Me, 2′-F, GalNAc | 2020 | 17 |
| | Inclisiran (LEQVIO, ALN-PCSSC) | Liver, atherosclerotic cardiovascular disease: Targeting PCSK9 mRNA | PS, 2′-O-Me, 2′-F, GalNAc | 2021 | 18 |
| | Vutrisiran (ALN-TTRSC02) | Liver, hATTR treatment, TTR mRNA | PS, 2′-O-Me, 2′-F, GalNAc | 2022 | 18 |
| | Nedosiran (RIVFLOZA) | Primary hyperoxaluria | PS, 2′-O-Me, 2′-F, GalNAc | 2023 | 19 |
| Aptamers | Pegaptanib (Macugen) | Eye, macular degeneration: Binding with the hairpin loop of VEGF165 and subsequent inactivation | Pegylated, all PO, 2′-F, and 2′-OMe; G and A methylated | 2004 | 20 |
| | Izervay (avacincaptad pegol) | Eye, complement C5 inhibitor to treat geographic atrophy (GA) | PEGylated oligonucleotide | 2023 | 21 |
| mRNA | BNT162b2 | Immune system, COVID-19 vaccine: SARS-CoV-2 S antigens' expression | Nucleosides-modified, Lipid nanoparticle-formulated | 2020 | 22 |
| | mRNA-1273 | Immune system, COVID-19 vaccine: SARS-CoV-2 S antigens' expression | Lipid nanoparticle-formulated | 2020 | 23 |
| CRISPR-Cas9 | Casgevy (exagamglogene autotemcel) | Blood, editing of CD34+ HSPCs using CRISPR-Cas9 | – | 2023 | 24 |

## 1.1 Classification of RNA-Based Therapeutics

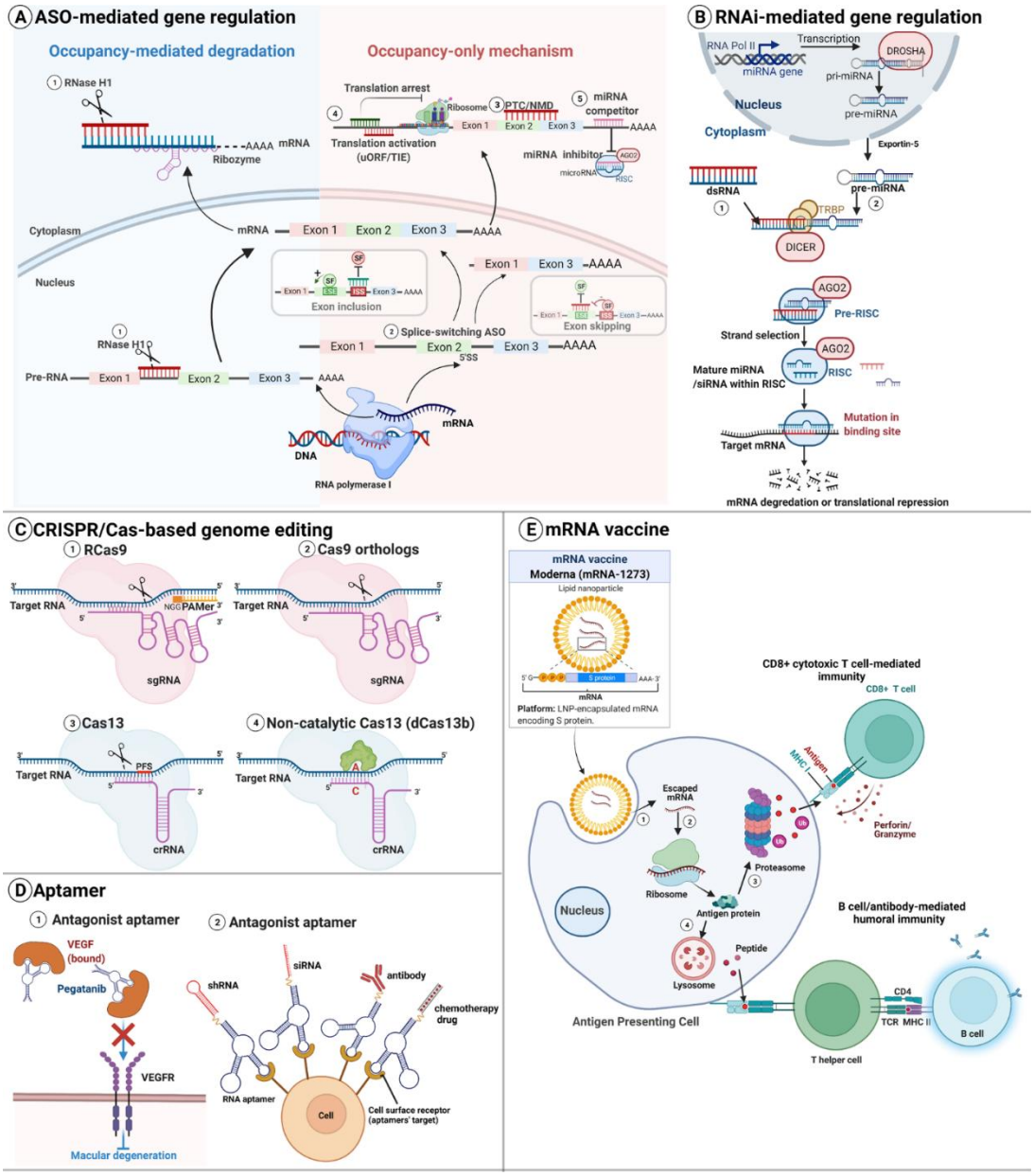



**Figure 1: Scheme of RNA-based drugs: A)** Antisense oligonucleotides (ASOs), **B)** RNA interference (RNAi). **C)** CRISPR/Cas-based RNA editing system, **D)** RNA aptamers, and **E)** mRNA vaccine [2].

### 1.1.1 Messenger RNA

Messenger RNAs (mRNAs) have an important intermediary role in the central dogma of molecular biology, carrying the blueprints of our genome. The mRNA therapeutics can be divided into three subgroups:

1. Replacement therapy: the mRNAs are administered to compensate for a defective gene or protein, or to supply therapeutic proteins.
2. Vaccination: the mRNA encoding a specific antigen is administered to trigger an immune response.
3. Cell therapy: the mRNA is transfected into cells ex vivo to alter cell phenotype or function, and these cells are then delivered into the patients [25].

The application and design of mRNA as drugs present several benefits, including transient expression without genomic integration (eliminating the risk of insertional mutagenesis [26]), which improves safety [27]; the ability to produce most of any functional protein or peptide [28, 29]; faster design and production compared to standard approaches [29]; cost-effectiveness and flexibility [28]; higher transfection efficiency and lower toxicity and compared to DNA-based drugs [28, 29].

This aspect guarantees that the mRNA is only active for a short period of time, reducing the potential burden on host homeostasis and decreasing off-target effects in gene editing [30]. The tunable aspect of mRNA allows for the rapid production of various proteins, making it a strategic and strong tool for precise medicine [30]. This versatility could permit the treatment of various diseases, including those previously considered unmanageable or genetic [31, 32]. Numerous mRNA vaccines are approved, starting with the mRNA rabies vaccine [33] and others targeting various infectious diseases, such as mRNA vaccines for COVID-19 [34- 36]. Synthetic mRNAs can stimulate an immune response or replace defective proteins in genetic disorders by encoding specific proteins [37]. After the huge success of COVID-19 vaccines (tozinameran from Pfizer–BioNTech, and elasomeran from Moderna), researchers around the world are working on the design and development of mRNA-based cancer vaccines [37].

These new drugs are part of personalized therapeutics, drugs designed to induce the immune system to recognize and eliminate cancer cells [38].

In fact, these vaccines can promote immune responses against tumor-specific neoantigens [39–42], with promising clinical benefits [43].

Now, no mRNA cancer vaccines have yet received regulatory approval, but several phase I–II clinical trials have shown encouraging results, even in poorly immunogenic tumors [43]. Moreover, mRNA vaccines could be designed to treat autoimmune diseases, where they could be highly selective in reducing pathological immune responses while preserving normal immune function [44].

Several optimization and modification processes can be applied to mRNA therapeutics to overcome numerous obstacles, including stability, immunogenicity, delivery, and scalability. In fact, unmodified mRNAs are easily targeted by cellular ribonucleases (RNases), causing their degradation, reducing their clinical efficacy, since they need to remain intact and unaltered to be translated into functional proteins or peptides. One of the main approaches is to insert unnatural nucleobases such as pseudouridine [45, 46], 5-methylcytidine [47, 48], and 1-methylpseudouridine [49] to improve their resistance to RNases [50, 51].

Additionally, to reduce immunogenic responses while improving stability, structural modifications at the 3', 5' ends, and untranslated regions (UTRs) can be applied [52]. The main problems are attaining high and sustained protein expression in specific cell types, shaping the need for repeated dosing, which can cause immune responses and reduced efficacy despite using modified mRNA and advanced delivery systems.

One of the main obstacles to mRNA therapeutics is the negative charge. In fact, the drug needs to cross cell membranes and reach the translation sites in the cytoplasm. To overcome this challenge, several techniques can be employed, specifically lipid nanoparticles (LNPs) [53], viral vectors [54], and classical electroporation techniques. Overall, these mechanisms can be applied to all RNA-based drugs as an optimization approach, improving half-life and therapeutic abilities.

Furthermore, several rare genetic diseases could be treated by mRNA therapeutics, and preclinical studies on rare disease models illustrate that [55], such as methylmalonic acidemia, glycogen storage disease type 1a, cystic fibrosis, Ornithine transcarbamylase deficiency, Phenylketonuria, Propionic acidemia, Hypercholesterolemia, Refractory hyperlipidemia, and α-1 antitrypsin deficiency. These overall astonishing results highlight the large expansion of interest in mRNA therapeutics.

To enhance the attractiveness of mRNA therapeutics, overall flexibility and cost are key factors [56, 57]. To achieve personalized therapy, a new generation of mRNA therapeutics needs to be developed. Many researchers are working to improve stability, delivery, and immune modulation to achieve that [57].

Additionally, mRNA vaccines have recently emerged as a promising strategy to generate anti-cancer immunity to magnify the effects of Immune checkpoint inhibitors (ICIs) [58–61] and determine the RNA therapeutics targeting infectious disease antigens as universal modulators of antitumour immunity [5]. These findings define vaccines as an available tool to enhance the efficacy of cancer immunotherapy. Furthermore, the next goal could be to design specific mRNA therapeutics to reset patient immune systems for enhanced response to immunotherapy [62].

### 1.1.2 Small Interfering RNAs

Small interfering RNA (siRNA) operates through the RNA interference (RNAi) pathway to silence specific genes [63-65] by binding complementary mRNA sequences, causing their degradation and allowing targeted gene knockdown [64]. They act by specifically hybridizing the targeted sense strand of RNA via Watson-Crick hydrogen bonding [65]. siRNA therapeutics offer several advantages, including targeting almost any gene with high precision [65] and silencing genes previously considered “undruggable” [66].

The siRNA duplexes have a length range between 30 and over 100 base pairs, in both natural and artificial precursor siRNAs. The enzyme Dicer processes the precursors into 21-bp-long siRNAs with two-base 3′ overhangs [67]. Thanks to this truncation process, a mature siRNA is produced that will interact with the RNA-induced silencing complex (RISC) to start the RNA interference (RNAi) process. Argonaute 2 (AGO2), an endonuclease and member of RISC, cleaves the sense strand of the duplex, leaving the antisense strand intact, then guides the active RISC to its target mRNA. AGO2 then attacks and destroys the phosphodiester backbone of the target mRNA. The antisense strand is usually fully complementary to the coding region of the target mRNA, enabling siRNA to knock down a specific target gene with high precision [68, 69].

siRNAs remain inactive until the transactivation-responsive RNA-binding protein (TRBP) activates them by loading them into AGO2 (while being double-stranded), which splits the sense strand, allowing the binding of the antisense strand to catalytic AGO2 [69].

The binding to the mRNA target can be perfect, in which the antisense strand has 100% complementarity to its target, or imperfect. The imperfect alignment leads to translational repression and potential off-target effects [70–72].

The efficiency of the siRNAs depends on different factors, such as target availability, binding site position, secondary mRNA structures, and the intrinsic siRNA characteristics. In fact, the success of RNAi strategies is mainly achieved through the optimization of the siRNA design [73].

Numerous modifications to standard siRNA designs have produced benefits such as reduced passenger strand activity and improved effectiveness. Part of these modified siRNAs are Dicer-substrate siRNAs [74], Small internally segmented siRNAs [75], Divalent siRNAs [76], Single-stranded siRNAs [77], and Self-delivering siRNAs (asymmetric and hydrophobic) [78].

Additionally, to improve the stability of siRNA, chemical modifications can be introduced to avoid its degradation by RNases while improving its half-life. Moreover, the clinical application of siRNA therapeutics has encountered four main problems: delivery, stability, specificity, and safety.

Furthermore, siRNAs have difficulty crossing the cell membranes, making it difficult to reach the target sites. This issue is related to their nature as double-stranded RNA sequences with a high negative charge. However, this difficulty can be resolved by different strategies. The easiest solutions are based on encapsulating the siRNA into nanoparticles, such as LNPs, polymer-based delivery systems such as polyethylenimine (PEI) and polyethylene glycol (PEG) [79]. Additionally, several approaches, such as antibody-vector-based [80] and organ/tissue-specific delivery, are utilized for efficient targeting [81]. For instance, GalNAc-siRNA conjugates represent a significant milestone in achieving targeted delivery of siRNA to the liver [82]. The FDA has approved many siRNA-based drugs: Patisiran [15, 83], Vutrisiran [84], Inclisiran [18], Lumasiran [85], Nedosiran [86], Fitusiran [87], Givosiran [16, 88], and others [89].

Fitusiran is a siRNA therapeutic designed to lower antithrombin (AT) and restore sufficient thrombin generation to rebalance homeostasis in hemophilia A or B (PwHA/B) and prevent bleeding [90].

Plozasiran is a first-in-class GalNAc-conjugated siRNA therapy designed to reduce the production of APOC3, a key component of triglyceride-rich lipoproteins (TRLs) and regulator of triglyceride metabolism, causing a reduction of triglycerides and restoring normal lipid profiles [91].

siRNA-based therapeutics may have additional applications, and continued clinical exploration in this field is necessary to discover new uses [92, 93]. In conclusion, continuing improvements in delivery systems, stability via chemical modifications, and specificity through sequence optimization are together making siRNAs a respected tool against a wide range of diseases and viral infections like HBV.

### 1.1.3 MicroRNA

MicroRNAs (miRNAs) are small, non-coding RNA molecules of 20–24 nt that can regulate gene expression at the post-transcriptional level by targeting and binding to a complementary sequence present in the 3'-untranslated regions (UTRs) of target mRNAs, leading to either translational repression or mRNA degradation through RISC. To target the specific mRNA, the miRNA is guided by AGO2, permitting the gene silencing process [94].

Unlike siRNAs, miRNAs show partial complementarity to their target, permitting them to regulate multiple genes at the same time and offering potential advantages over traditional single-target therapies [95]. They have critical roles in various physiological processes, including development, cell division, differentiation, apoptosis, and immune responses [96-98].

In cancer, miRNAs can act as oncogenes, defined as oncomiRs, or tumor suppressors, and this identification made it possible to understand their therapeutic potential. The miR-34 family acts as a tumor suppressor by targeting genes involved in cell cycle regulation and apoptosis [99]. On the other hand, the miR-21 family promotes tumor progression by downregulating tumor suppressors acting as oncomiRs [100]. Additionally, recent evidence has shown that miRNAs can be developed into a clinical biomarker for hepatocellular carcinomas (HCC) [101].

miRNA-based therapeutics can be used to restore the function of downregulated miRNAs, acting as miRNA mimics or antagomiRs, to inhibit overexpressed miRNAs [102] and target multiple genes simultaneously [95]. They are also useful as an advanced approach in gene therapy [103, 104]. miRNA inhibitors are usually chemically modified, single-stranded oligonucleotide sequences designed to bind to and inhibit endogenous miRNAs [95]. In this way, the translation of target mRNA is upregulated, reducing the possible effects caused by negative miRNAs that are overexpressed in the cells [105]. They are highly specific [95] and can last long due to chemical modifications [103]. At the current state, many different miRNA-based drugs are under investigation, promising an advancement in gene therapy and cancer treatments [106, 107].

### 1.1.4 Non-Coding RNAs

Non-coding RNAs (ncRNAs) have a huge impact on the regulation of cellular processes and gene expression. Specifically, they have specific roles in modulating post-transcriptional gene expression and transcriptional control, plus they can influence epigenetic regulation [108-112]. Moreover, they have a pivotal role in several cellular processes, such as signal transduction pathways, nuclear structure maintenance, and RNA splicing [109, 113]. Additionally, during RNA splicing, ncRNAs contribute to a precise elimination of introns and the joining of exons in pre-mRNA transcripts [114, 115]. This process is critical to generate mature mRNA that can translate into functional proteins. ncRNAs aid the organization of the chromatin, enabling powerful interactions between different genomic regions [114]. Thanks to that, the gene expression and cellular functions are well sustained and can be executed properly.

Furthermore, ncRNAs facilitate extracellular signals to intracellular responses and other transduction pathway signaling [116–118]. Thanks to their influence over these pathways, they can modulate several important processes, such as stress and immune responses, and metabolism. They also have a strong role in the regulation of other signaling pathways that control essential cellular functions, such as apoptosis, differentiation, and proliferation [109].

Furthermore, ncRNAs have a role in controlling the expression of pro-apoptotic and anti-apoptotic factors [119], regulating apoptosis. This action has a key role in establishing whether a cell will experience programmed cell death or survive under stress conditions.

Their contribution to many disease mechanisms has significant consideration due to their implications in several pathologies, including cancer, neurological disorders, and cardiovascular diseases [110, 111]. They are also tangled in regulatory processes in cardiovascular diseases, such as cardiac hypertrophy and vascular remodeling [120, 121].

In summary, ncRNAs are fundamental for an organism's health [119]. Highlighting their huge role in preserving homeostasis and understanding the roles of ncRNAs in these diseases provides insights into their crucial mechanisms. In the proximate future, it will be possible to develop targeted therapies that hold the unique properties of these regulatory nucleic acids.

### 1.1.5 Long Non-Coding RNAs

Long non-coding RNAs (lncRNAs) are good targets and tools in gene therapy, and they have vast potential as drug candidate models [122]. lncRNAs are usually longer than 200 nucleotides with important roles in several cellular processes, modulating complex genetic networks and genetic regulation [122]. Specifically, lncRNAs can influence transcriptional and post-transcriptional processes, modulate chromatin structure, and interact with proteins and other RNAs, making them fundamental to maintaining a balance in cellular functions and homeostasis [122].

Thanks to their large spectrum of actions, they could be used as gene therapy to offer a different solution as treatment for complex genetic disorders and cancers [109]. Targeting lncRNAs permits massive advantages, starting with a tissue/cell type-specific expression pattern [109], scaffolds for protein complexes, enhancers of gene expression, or even inhibitors of oncogenic pathways [122]. Additionally, they are useful as diagnostic and prognostic biomarkers [109].

lncRNAs are involved in numerous diseases. In fact, recent studies have shown that targeting specific oncogenic lncRNAs can inhibit tumor growth and metastasis in preclinical models [30]. Working on the restoration of lost or downregulated lncRNAs presents an exciting opportunity for therapeutic development [122].

### 1.1.6 Antisense Oligonucleotides

Antisense Oligonucleotides (ASOs) are short, usually 18–30 nucleotide long, synthetic nucleic acids that can bind complementary mRNA target [21]. Base pairing through Watson-Crick hydrogen bonds between nucleobases is the simplest yet most effective strategy to manipulate gene expression [123].

ASOs can reduce the expression of disease-causing proteins by targeting specific mRNAs [66], acting as valuable tools in treating genetic disorders. Several ASOs have received FDA and EMA approval for clinical use in the treatment of cytomegalovirus retinitis polyneuropathy, Duchenne muscular dystrophy, hereditary transthyretin amyloidosis, homozygous familial hypercholesterolemia, and spinal muscular atrophy [67].

ASOs can modulate gene expression or modify pre-mRNA splicing depending on their chemistry, target sites, and binding sequences. One mechanism of action is determined after binding between the ASO and the mRNA target, in which degradation is achieved by RNase H-mediated cleavage (gapmers) (method defined as RNase H1-dependent. The other mechanism of action is through steric hindrance (mixmers) via high-affinity binding to complementary targets [124], as an RNase H1-independent method [22]. Thanks to these pathways, ASOs can modulate splicing and, effectively, alter protein production [109, 125].

The endogenous RNase H1 is an enzyme well distributed in the nucleus and cytoplasm, permitting it to target long non-coding RNAs and immature pre-mRNAs [126-128].

Moreover, ASOs can be manufactured to bind the AUG start site and nearby sequences or UTR sequences, using steric hindrance to block the binding of RNA-binding protein complexes, like RNA-binding proteins and ribosomal subunits, causing the inhibition of the translation of target mRNA [129]. ASOs can also target the upstream open reading frames (uORFs) to enhance protein translation and inhibit the expression of the main ORFs [130].

Another method by which ASOs can eliminate abnormal mRNAs is through the activation of the endogenous cellular surveillance mechanism. When ASOs target pre-mRNAs, they induce premature termination codons in mRNAs, and they can be degraded via nonsense-mediated mRNA decay [131]. Furthermore, ASOs can regulate the translation process by inhibiting or activating various mechanisms. One of the methods is the inhibition of 5′-cap formation [132], modifications in polyadenylation [133], and translational arrest [134].

The synthesis and design of ASOs is well established and can be easily scaled up. Moreover, many different optimization methods are available to improve affinity, stability, and pharmacokinetic/pharmacodynamic properties. Between them, chemically modified gapmer and mixmer ASOs can contain scattered nucleotides linked by phosphorothioate bonds. Moreover, several other modifications can be used [135-141].

Moreover, to improve efficiency and tissue-specific delivery, the conjugation of ASOs with antibodies, aptamers, and cell-penetrating peptides is a common practice [142, 143]. Additionally, advancements have been made for efficient and tissue-specific delivery strategies by employing conjugation approaches, such as coupling ASOs with cell-penetrating peptides.

One of the main issues related to ASOs, and shared with other RNA-based drugs, is the presence of a high density of negative charges, mainly due to phosphate groups, which makes them difficult to cross the hydrophobic cytoplasmic membrane. Moreover, exogenous RNAs can be very immunogenic, causing cell toxicity and reducing their effectiveness as therapeutic drugs.

The FDA gave regulatory approval for RNase H-competent ASOs: fomivirsen [144], mipomersen [145], volanesorsen [146], inotersen [147], tofersen [148], eplontersen [14], and olezarsen [149]. Additionally, various splice-switching ASOs have received FDA approval, including Eteplirsen [150], Golodirsen [151], Viltolarsen [34], Casimersen [152], and Nusinersen [153].

Olezarsen is an ASO used to treat familial chylomicronemia syndrome (FCS), acting as an adjunct to diet by reducing triglyceride levels in adults [154]. Next, Donidalorsen is a GalNAc3-conjugated ASO that binds prekallikrein (PKK) mRNA in the liver, reducing its expression. In fact, PKK is a key enzyme involved in activating inflammatory mediators associated with acute attacks of hereditary angioedema (HAE) [155]. Last, Ulefnersen is an ASO designed for the treatment of patients with amyotrophic lateral sclerosis (ALS) caused by a mutation of the fused in sarcoma (FUS) gene [156].

### 1.1.7 Circular RNA Molecules

Circular RNAs (circRNAs) are highly abundant and conserved, with tissue- and developmental stage-specific expression patterns. circRNAs are critical molecules because of their regulatory roles in various biological processes [157]. The presence of covalent bonds shapes them as a loop without having 5' caps or 3' poly(A) tails, making them highly resistant to nucleases and very stable in cellular environments [158]. Their development starts with precursor mRNA (pre-mRNA), where the downstream 5' splice site is joined to an upstream 3' splice site, defining this process as back-splicing [159]. They enable their function through a process defined as sponging miRNAs [160, 161]. During this process, circRNAs harbor multiple miRNA response elements (MREs) and act as competing endogenous RNAs (ceRNAs), thereby sequestering miRNAs and regulating the expression of miRNA target genes [161]. Moreover, they can modulate the activity and localization of RBPs [162], and some of them can be translated into peptides with functional activities [161, 163].

circRNAs are promising and attractive candidates for the development of innovative RNA-based drugs with multiple applications, thanks to their high stability and low immunogenicity [164]. Moreover, pattern recognition receptors such as Toll-like receptors (TLRs) can not trigger the innate immune responses when attacked by circRNAs [165]. Besides their innate properties, the bigger obstacle is the efficient delivery of circRNA therapeutics. circRNA-based drugs in several preclinical studies have demonstrated their potential but overall are still in the early stages of development [166].

Overall, circRNAs are promising RNA-based drugs that could be designed to treat pathologies from cancer to viral infections, demonstrating a large landscape of targets [167-174].

### 1.1.8 Aptamers

Aptamers are short single-stranded DNA or RNA molecules (ssDNA/ssRNA) with specific tertiary structures, which allow them to bind to a large variety of targets, such as proteins, peptides, carbohydrates, organelles, and even cells. The Systematic Evolution of Ligands by EXponential enrichment (SELEX) was the protocol identified and finalized in 1990 for the first-ever synthesis of aptamers [175-177], and is a process created to screen aptamers from large databases of random sequences that can contain over 1 trillion different sequences. These sequences are brought to bind candidate target sequences under specific physicochemical conditions. Once the molecules are bound to their target, they are isolated through several processes [178]. Hence, another selection approach is performed, and after multiple cycles, a few high-efficiency molecules are identified to hit the desired target. Aptamers are very easy to separate thanks to their specific three-dimensional folding and in vitro replication preference [177, 179], and they can be used for screening ssDNA or RNA aptamers [180]. The purified-protein-based SELEX has some issues; in fact, it cannot mimic with high accuracy the native conformations of target proteins in their natural cellular environment [181]. The cell-based SELEX screening can overcome this issue but is designed to be efficient at targeting proteins on the cell surface, in opposition to what most therapeutic target molecules are likely to reside inside the cell [181, 182].
A recent technique called CRISmers is a CRISPR-based RNA aptamer screening platform. Thanks to this technique, numerous RNA aptamers designed to target the receptor-binding domain (RBD) of the SARS-CoV-2 spike glycoprotein were successfully identified, showing highly sensitive binding and the ability to neutralize SARS-CoV-2 virus and its Delta and Omicron variants [4].

Aptamers are one of the most versatile nucleic acid-based drugs because they can bind to a wide variety of targets with high specificity and affinity. Thanks to these properties, they can be defined as antibody-like structures. In fact, they can adopt well-defined three-dimensional structures, such as hairpins, loops, or pseudoknots, which are important for their interaction with target molecules [4]. These properties define aptamers as valuable tools for improving targeted drug delivery systems and molecular recognition elements in therapeutic and diagnostic applications, as well as for modulating protein function [183]. Aptamers can function as antagonists, avoiding the interactions between ligand and receptor [183] or carriers, where they can deliver specific therapeutic drugs to cells or tissues with high precision [184-186].

Aptamers can modulate gene expression by acting as multidirectional riboswitches. Specifically, when the aptamer interacts with one of its ligands, it induces conformational changes in its tertiary structure, adapting to the specific situation, a property that enables the design of ligand-induced aptamer switches for multidirectional expression systems [187, 188]. Natural and synthetic aptamer switches have been identified to modulate almost every process of mRNA turnover [189, 190].

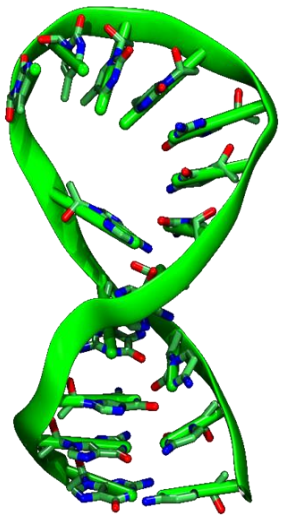


**Figure 2**: Structure example of an aptamer.

Besides all the positive aspects, there are a few issues related to aptamers and their druggability. The first one is that unmodified aptamers are quickly degraded by nucleases, such as DNases and RNases, in the cell system in seconds to an hour, depending on their length and tertiary conformation in space [191, 192]. Structure-wise, aptamers are very light, with a weight range of 6–30 kDa, and with a small size (<5 nm) [193], determining their fast renal clearance while being prone to leak through renal filtration [194]. To reduce the impact of these problems, improvements must be made to permit their clinical application; massive modifications are necessary.

Part of these modifications are applied through conjugation with additional moieties, such as Polyethylen glycol (PEG), improving aptamers' half-life while severely reducing the elimination by renal filtration. In fact, heavier compounds, such as cholesterols [195], peptides [196], PEG [197], liposomes [198], and nanomaterials [199, 200] can be used in these situations. Moreover, aptamers can be synthesized via chemical synthesis, are easy to modify, and have less immunogenicity compared with antibodies [201].

Only two aptamers received approval from the FDA. The first is Pegaptanib, which specifically binds to vascular endothelial growth factor-A (VEGF-A), inhibiting its interaction with receptors and preventing angiogenesis [20]. The second one is Izervay (avacincaptad pegol), a complement C5 inhibitor designed for the treatment of geographic atrophy (GA) caused by advanced dry age-related macular degeneration (AMD). Thanks to their overall properties, aptamers are extensively studied in immunology, oncology, metabolism, neurodegenerative diseases, and cardiovascular diseases [21]. Several aptamers are under clinical phases [202-204].

### 1.1.9 Ribozymes

RNA molecules can have the ability to act as enzymes, performing catalysis of reactions; in fact, it was proven that the catalytic skill of purified ribonucleases from prokaryotes is entirely dependent on the RNA moieties without the interference of proteins and enzymes, demonstrating the pivotal role of RNA molecules [205, 206].

Available in nature are two different ribozymes: the hammerhead ribozymes were discovered in 1987, and later, hairpin ribozymes in 1990, with the latter being well investigated and used in therapeutics [207, 208].

The hammerhead ribozyme consists of three helices, which flank and bind the antisense strand of target RNA, and a catalytic core that mediates the cleavage reaction [209]. The cleavage is performed on the RNA at a specific site, and the two resulting cleaved RNAs are prone to degradation, while the hammerhead ribozymes stand away, ready to target another copy of such sequences, intending to suppress their target sequences [210]. Several clinical trials have studied ribozyme therapeutics, but most of them were abandoned due to insufficient efficiency (e.g., angiozyme) [211, 212]. Ribozymes have shown limited stability in biological environments, mainly because they are easily targeted by nucleases.

The development of chemically modified ribozymes or improved delivery systems is essential to enhance their stability and overall performance. The application and study of 2’ modified pyrimidine nucleosides and phosphonothioates of hammerhead ribozymes showed significantly augmented stability and activity [213, 214]. To overcome the challenges associated with intracellular environments and the need for high accuracy, several improvements must be made in RNA chemistry, delivery vectors, and technologies to enable their successful therapeutic application [4].

### 1.1.10 RNA-Based Drug Delivery

Delivery systems are essential tools for RNA-based therapies and are divided into two main categories: viral vectors (e.g., Adeno-associated viruses, AAVs) and non-viral vectors. The application of viral vectors permits the transduction of target cells while facilitating the long-term expression of therapeutic RNAs, making them a valuable tool to treat several chronic diseases and genetic disorders [215].

AAVs are favored for their low immunogenicity and ability to transduce both dividing and non-dividing cells [215]. They can be used in a large spectrum of tissues, including those that are difficult to target with other delivery methods [215]. AAVs have an exceptional safety profile and versatility, allowing them to gain popularity for gene therapy [216-220]. Over 1000 variants and 12 AAV serotypes have been discovered, each with the potential to target different cell types without being pathogenic [220-224].

Another positive aspect regarding AAVs is related to their episomal transgene expression, determining a low risk of insertional mutagenesis [225-227]. The outgoing presence of recombinant adeno-associated viruses (rAAVs) allowed for alleviating concerns relative to the integration of the RNA material into the host genome, making them safer [227, 228].

Lentiviruses can perform long, stable transgene expressions while infecting both dividing and non-dividing cells, including stem cells and neurons [229, 230]. They can integrate their DNA into the host cell's genome, permitting long-term gene expression, especially in dividing cells [231, 232]. Latter technology includes the development and advancement of non-integrating lentiviral vectors (NILVs). NILVs contain the advantages of lentiviruses while severely dropping the risk of insertional mutagenesis. To produce them, mutations are introduced in the viral enzyme integrase or by modifying the viral DNA that integrase recognizes [233].

Non-viral delivery systems have gained an important distinction in RNA-based gene therapy, mainly through LNPs [234]. LNPs can simplify cellular uptake and improve the endosomal escape, and are designed to encapsulate and protect RNA molecules, avoiding extracellular ribonucleases [235, 236]. Thanks to these properties, the RNA-based therapeutics can reach their target within the cell, maximizing their efficacy. Moreover, their natural composition permits an efficient cellular uptake; in fact, the ionizable lipids allow for a high encapsulation efficiency, promoting an effective cell entry [237].

Additionally, LPNs have low immunogenicity compared to viral vectors, showing a safer profile for clinical applications [238]. Moreover, LNPs enhance endosomal escape, allowing the release of RNA from the cytosol through interactions with endosomal membranes [237].

Furthermore, LNPs can be used to encapsulate several different types of RNA and can be designed and engineered to target specific tissues [238]. Polymeric nanoparticles are another type of non-viral vector, which are synthetic carriers valuable for delivering RNA molecules, composed of biodegradable polymers to encapsulate and protect RNA [239, 240].

This type of vector can offer top-tier stability and protection, defending the RNAs from degradation; moreover, it furnishes efficient cellular uptake and increased half-life in the bloodstream. Several types of vectors are available, such as dendrimers, polymeric micelles, and polymer-drug conjugates [240] as part of the nanoparticle platforms. The development of cationic polymers allowed the formation of complexes with highly negatively charged RNA structures, enabling improved delivery into cells [239, 241]. The silica nanoparticles, a delivery system that can be easily personalized on the surface, thanks to their intrinsic properties, can be functionalized to target ligands or drugs, permitting a good delivery of RNA molecules [239] while enhancing their therapeutic efficacy [242].

Silica nanoparticles can be engineered to perform a controlled release, and they have a high surface area for cargo encapsulation [243]. Specifically, Mesoporous silica nanoparticles (MSNs) have shown good results in carrying siRNA, causing the downregulation of genes correlated with osteoporosis-related diseases [244, 245]. Another recent and available option is the design of gold nanoparticles (AuNPs), which can be used to create self-assembled capsules [246].

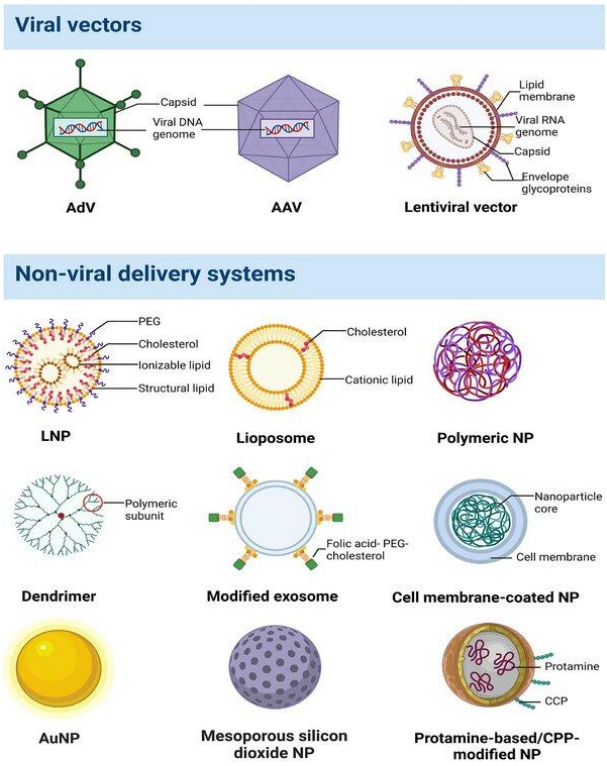



**Figure 3**: Illustration of the available viral and non-viral carriers for RNA-based drug delivery [247].

### 1.1.11 Chemical Modifications in RNA Molecules

One of the main issues related to RNA molecules is that they are unstable in the cellular environment because they are easily and rapidly attacked by RNases available in cells, destroying them as they are recognized as extraneous objects. Associated with that, the nature of RNA's 2'-OH group makes RNA fragile and unstable. This instability poses a huge challenge for the development of RNA-based drugs [4].

To improve the RNA's stability, a large set of chemical modifications can be used, including:

1) Modification and optimization of the sugar-phosphate backbone
2) Application of modified bases
3) Insertion of nucleotide modifications.

The main goals of these modifications are to improve the resistance to nucleases, enhance affinity to the target, improve pharmacokinetics, and reduce pro-inflammatory reactions or immunogenic responses.

These modifications need to be applied to a specific site, specifically for siRNA and ASO, where the number of nucleotides is limited. In fact, researchers around the globe can design site-specific chemical modifications [4]. Hereinafter, a list of chemical modifications applied to each type of RNA-based drug is provided.

### 1.1.12 Chemical Modifications in mRNA

Several methods can be applied to optimize mRNA vaccines: sequence optimization, nucleobase modification, or sequence substitution to increase translational capacity [248]. Chemical modifications like 5-methyluridine ($m^5U$), N1-methylpseudouridine, 2-thiouridine, and pseudouridine are added into mRNA sequences to reduce the innate immune activation [249].

### 1.1.13 Chemical Modifications in ASOs

The first ever modification applied to the backbone of nucleic acids was the phosphorothioate backbone, which was introduced in ASOs and siRNAs [250] The insertion of phosphorothioate oligodeoxynucleotides (PS-ODNs) can be accomplished in several ways, and the phosphorothioate linkage is formed by adding a sulfur atom to replace a non-bridged oxygen, which increases the resistance to nuclease degradation [251-253]. The formation of new hydrogen bonds by the sulfur atom, specifically in the PS linkage, can stabilize the local structure in the RNA, improving its resistance against nucleolytic degradation, as was demonstrated with Fomivirsen, in which the half-life of ASOs increased in serum [254-256]. However, the improved binding to serum proteins, such as albumin, can cause off-target effects and toxicity [257, 258]

To improve safety, efficacy, stability, and half-life in vivo, the 2'-OH of the ribose can be modified, and several methods are available, such as 2'-O-methyl (2'-OMe), 2'-fluoro (2' -F), 2'-O-methoxy-ethyl (2'-MOE), and 2'-O-aminopropyl (2'-O-AP) [259]. Additionally, the 2'-OH modification can improve the binding affinity of RNA molecules to their target (like mRNA or miRNA), thanks to an increased flexibility and conformational range while reducing off-target effects [259].

To improve the affinity to target mRNAs, while adding resistance to degradation by nucleases and peptidases, other modifications in the furanose ring, along with the PS, riboses, and nucleotide, are applied [260].

Other backbones that can be applied, part of the Xeno Nucleic Acid, are Locked nucleic acids (LNAs), peptide nucleic acids (PNAs), and phosphorodiamidate morpholino oligomers (PMOs). All of them are valuable options to enhance RNA resistance and stability [261-263]. Thanks to these chemical modifications and optimizations, the application of ASOs as therapeutics has improved, specifically for genetic disorders and viral infections [264]. Another key aspect related to these modifications is that they can be used independently or in combination.

### 1.1.14 Chemical Modifications in siRNA

The modifications that are applied to ASOs motivated the application and use of siRNA therapeutics using the same techniques [265]. In fact, PS, PMO, and PNAs are the most prevalent changes to the phosphate backbone for siRNAs [266]. These molecules show extraordinary uptake in many tissues, including lymph nodes, adipocytes, bone marrow, liver, spleen, and kidney. Thanks to these types of modifications, PS-modified RNA-based drugs can be stable with a half-life of 1–4 weeks [266, 267].

Another modification is the implementation of modified linkages in the RNA, such as Methylphosphonate and borano phosphate (BP), that are generally categorized as neutral or cationic internucleoside linkages [268, 269].

However, PS-modified molecules exhibit reduced sequence-specific binding affinity [270]. A peculiar problem of siRNAs with PS bonds is that, if more than half of the phosphodiester bonds are replaced with PS bonds, the overall toxicity is increased [271]. This problem is related to the fact that this type of modification impairs RISC activity [272]. BP increases lipophilicity more than PS analogs and offers twice the resistance to nuclease degradation while causing low toxicity. On the other hand, this modification cannot be performed in the center of the guide strand, since it is not well tolerated, limiting its application [273]. PMOs and PNAS don’t present these issues and are largely used [274]. 2'-MOE, 2'-OMe, and 2'-F are present in all siRNA drugs that are undergoing clinical trials [275].

Other types of modifications are Tricyclo-DNA oligonucleotides (tc-DNA), constrained ethyl oligonucleotides (cEt), and LNAs [276, 277]. Moreover, the main nucleobase modifications include carboxyl substitution, cytosine methylation, deamination, hydroxymethylation, guanine oxidation, and adenine methylation [278]. Additionally, several unnatural nucleobases can be implemented in the starting sequence.

The modification of nucleobases is performed less often than sugar-phosphate. When nucleobase analogs are added, certain functions can be promoted, such as molecular size or charge distribution [278, 279].

### 1.1.15 Chemical Modifications of Aptamers

Several downsides related to aptamers are their shortage of chemical diversity [280], high hydrophilicity, and negative charge [281]. They primarily form ionic interactions, hydrogen bonds, stacking, and polar interactions with their targets. The application of large chemical moieties in aptamers results in a promising option [280].

When we consider natural chemistry aptamers, constituted by DNA and RNA backbones, they mainly rely on shape complementarity and polar or electrostatic interactions while they bind to their targets. With the modern advancement of synthetic chemistry, new protocols have emerged to improve nucleotide structures, including ribose modifications [282, 283] and the introduction of new chemical moieties [284-286], and lastly, the expansion of the genetic alphabet [287–289]. Another problem related to the synthesis of the aptamers is associated with the introduction of these chemically modified nucleotides, which need to be compatible with polymerase-mediated synthesis and convenient in the SELEX protocol for the identification of modified aptamers [290, 291]. Moreover, several chemical modifications can be applied, such as 5'-end PEGylation [292], inverted thymidine capping [293], and sugar ring substitutions. These modifications improve aptamers 'binding affinity and pharmacokinetic properties [294]. As mentioned before regarding the other RNA-based drugs, for the aptamers, several chemical modifications can be introduced: at the nucleobase level, phosphodiester backbone, and ribose sugar. The most common base modifications include 5'-methyluridine, pseudouridine, and dihydrouridine. Base modifications are frequently exploited for the selection of DNA aptamers with protein-like functionality, such as SOMAmers [295, 296]. Numerous sugar modifications can be exploited to improve aptamer stability, including 2'-OMe, 2'-F, 2'-amino ribose (2'-NH2), cyclohexane nucleic acid (CeNA), threose nucleic acid (TNA), arabinonucleic acid (ANA), 2'-fluoroarabino nucleic acid (FANA), LNA, and PNA. Several backbone modifications are available regarding aptamers, including the replacement of the phosphodiester linkage with methylphosphonate, triazole on the α-phosphorus, and phosphorothioate [297–299]. PS is a valuable modification for the formation of G-quadruplex oligonucleotides, by increasing their thermal stability [300].

The introduction of post-SELEX chemical modifications (incorporation of 2'-F, 2'-OMe, 2'-NH2, CeNA, LNA, FANA, ANA, TNA, L-DNA, and L-RNA) can occasionally harmfully impact the function of aptamers. One way to avoid this problem is to perform SELEX experiments directly with chemically modified nucleotides with the support of an engineered polymerase [301]. A low percentage of current aptamers contain modified nucleic acids (MNAs) compared to conventional RNA and DNA, but the only FDA-approved aptamers are chemically modified. However, the introduction of MNA-based aptamers is likely to rise as modified nucleotides become more commercially accessible [302]. The presence of experimental and computational evidence reveals that structural complexity improves the functional potential of nucleic acids [303-309].

### 1.1.16 Ribose Modifications

The ribose modifications confer improved stability to aptamers by reducing the influence of nucleases in intracellular and extracellular environments [309]. The main changes on the 2' OH include halogenation [310], amination [311], and alkoxyl group substitution [312].

Pegaptanib, the FDA-approved aptamer for the treatment of age-related macular degeneration (AMD), consists of 13 2' F pyrimidine nucleotides, and 12 out of 14 purine nucleotides are 2' OMe [313-314] while severely improving the half-life [314]. The 2'OMe group of the ribose induces steric hindrance, causing a reduction in affinity between the polymerase and the 2'OMe nucleotides, establishing nuclease resistance. To work with these ribose modifications, advanced polymerases are required to perform correctly during SELEX protocols.

### 1.1.17 Nucleobase Modifications

Another pivotal method to modify aptamers is by adding new features through chemical moieties. Since nucleobases have a crucial role in epitope recognition, their modification needs to be performed with high precision [315]. The most favorable positions to apply the modifications are C5 of pyrimidines and C8 of purines [280]. Moreover, protein-like or hydrophobic moieties are largely used as chemical appendages [316–318].

Several novel nucleobases are available to overcome the few naturally available options. These available unnatural base pairs do not follow the conventional Watson-Crick pairing [319].

The expanded genetic alphabet is a powerful tool in diversifying the functionalities of aptamers [320–322]. This method is defined as artificially expanded genetic information systems (AEGIS), and it was used to design potent modified aptamers against proteins [323] and cancer cell lines [324, 325]. Another available method is the enzymatic construction of metal base pairs [326], in which two synthetic nucleotides serve as ligands for metal coordination [327-331].

### 1.1.18 Backbone and Tag Modifications

Aptamers have a small molecular size and fast renal clearance. To avoid this, the phosphate backbone can be modified by adding tags [332, 333], which can slow down renal clearance [334]. Cholesterol or PEG are the main tag moieties applied on the terminus of the aptamers; they increase their molecular size while reducing renal clearance [335]. This modification is also visible on Pegaptanib, which is PEGylated to increase its half-life [335].

### 1.1.19 Synthetic Approaches

By adding these unnatural nucleosides, new problems came to play. In fact, they are less accessible by polymerases in SELEX approaches and in solid-phase synthesis (SPS). These methods usually have a substandard yield and a higher cost [336]. To overcome these issues, the biocatalytic synthesis, which uses enzymes to catalyze chemical reactions [337], can be used. Another available synthetic approach is through click chemistry [338], which allows the acquisition of noncanonical nucleosides at lower cost and higher efficiency.

### 1.1.20 RNA-Based Gene Therapy Applications and Future Perspectives

The RNA-based gene drugs offer advanced strategies for cancer treatment, such as silencing oncogenes, restoring tumor suppressor functions, and modulating immune responses [5].

SiRNAs and ASOs are available for cancer treatment [339]. Specifically, siRNAs can be designed to target mRNAs that translate into oncogenic proteins with the goal of silencing these genes that drive tumor growth [66]. ASOs can be designed to perform the same action, restoring standard gene expression [340]. Moreover, mRNA-based therapies are being investigated for their application as an advanced therapy against cancer [28, 29]. These treatments encode for proteins that can boost immune responses against tumors or restore regular cellular functions [28].

One of the methods is designing mRNA that encodes for tumor-associated antigens (TAAs), new types of mRNA vaccines that can elicit strong immune responses.

Thanks to this approach, the delivery of multiple antigens can be achieved at the same time, improving humoral and cell-mediated immunity [28, 29].

Additionally, RNA therapies could be useful for direct antiviral effects in addition to vaccines. Specifically, siRNAs that can target viral genes effectively stop viral replication by silencing vital genes and significantly reducing viral loads in patients [341]. Furthermore, RNA-based treatments are viable options for illnesses related to malfunctioning proteins [342].

The recent discoveries and advancements in technology will improve the specificity of RNA-based therapeutics while reducing the impact of natural instability, the possibility of off-target effects, possible immunogenicity, and poor performance with traditional delivery systems.

The groundbreaking introduction of artificial intelligence tools is increasingly being used in RNA therapy for several purposes, such as RNA structure prediction [343], and sequence optimization by the analysis of large datasets, etc [344]. Another advanced technique is the Guide RNA-based CRISPR-Cas system, which provides genome editing precision, a useful treatment of genetic disorders [345] while improving the specificity of gene manipulation. All things considered, the development of RNA-based therapeutics as a potentially significant therapeutic approach is closely linked to advancements in several crucial areas, including improving stability in the native cellular environment, guaranteeing cost-effectiveness, and accomplishing effective and targeted delivery to the target site. To fully realize the potential benefits of NA therapeutics in the future, significant improvements in these areas are essential.

# 2. COMPUTATIONAL CHEMISTRY

The drug discovery and development processes are composed of three main steps: drug discovery, preclinical development, and clinical trials.

Drug discovery begins with identifying a hit molecule, a compound that produces a desired effect in screening assays [346, 347]. After that, the structure of the hit molecule is optimized and modified to improve affinity and selectivity, decreasing toxicity, increasing water and lipid solubility, and enhancing ADME properties. Thanks to this long process, the hit molecule is converted into a lead molecule. The further optimization of the lead molecule delivers the drug candidate.

High investments and time are necessary for the discovery of new drugs, but the success rate after all the phases is only 13% [348]. In most cases (40-60%), the failure of the drug identification is determined by poor pharmacokinetic properties (ADME/Tox) [349]. In this critical situation, computer-aided drug discovery (CADD) techniques are advanced tools that improve and accelerate drug discovery and development by reducing costs and increasing success rates [350]. An integral part of CADD is rational drug design; this approach provides knowledge to understand and investigate the binding affinity and molecular interaction between the target protein and ligand [351].

Furthermore, the advancement in technology related to pharmaceutical sciences has permitted aiding the lead identification by the construction and development of supercomputing facilities, advanced programs, clusters, algorithms, etc [352]. Moreover, the innovative recent improvements in artificial intelligence (AI) and machine learning (ML) methods have significantly improved the analysis and investigation of pharmaceutical-related big data in the drug discovery process [353].

Besides the advancement in AI, several other methods can be applied for the identification of new drugs, such as chemical databases, molecular docking, pharmacophore modeling, quantum mechanics, quantitative structure-activity relationship (QSAR), and statistical learning methods.

When we consider the CADD methods, these are divided into structure-based (SBDD) and ligand-based drug design (LBDD) protocols.

The SBDD methods are based on the availability of a 3D structure of the target protein (receptor) plus information associated with its active sites. This information is necessary to investigate and understand the molecular interaction between the receptor and ligand [354].

On the other hand, LBDD methods rely on well-known and studied ligands that can interact with the target receptor [355].

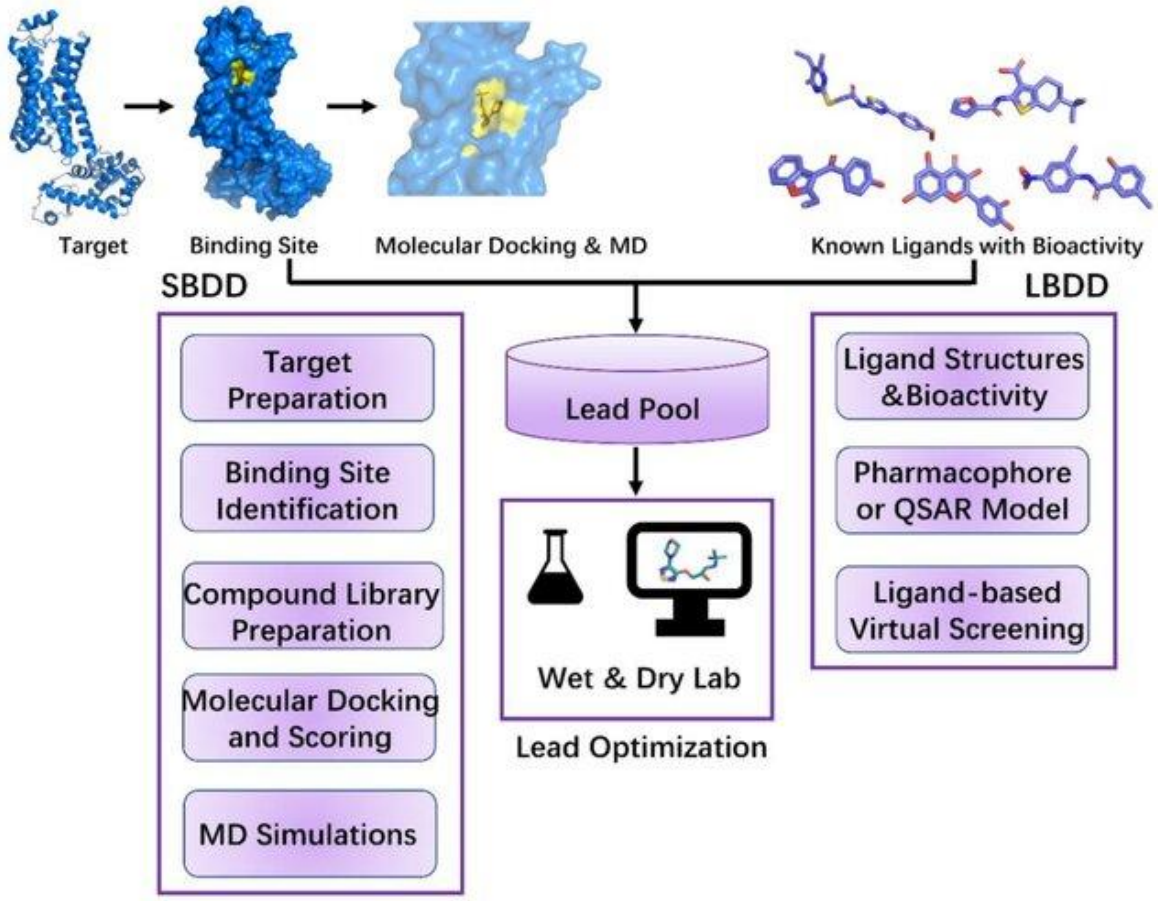



**Figure 4**: Workflow of SBDD and LBDD [356]

## 2.1 Structure-Based Drug Design

The SBDD method is useful for the identification and optimization of lead molecules. Since the structure of the receptor is available, this method is helpful to understand the interactions at the molecular level [357]. The main techniques used in SBDD methods are structure-based virtual screening (SBVS), molecular docking, and molecular dynamics (MD) simulations. These methods are valuable for evaluating binding energetics, conformational changes in the receptor after binding with a ligand, and protein-ligand interactions [358].

A standard SBDD protocol consists of numerous steps:

A) Preparation of the target structure

B) Identification of the ligand binding site

C) Compound library preparation

D) Molecular docking and scoring functions

E) MD simulation

F) Binding free energy calculation

The availability of structural elucidation techniques such as X-ray and NMR has led to a quick advancement in the resolution and quality of the structures deposited and available in the protein data bank (PDB), but still, several protein structures have not been solved yet [359]. Even with these issues, the more modern computational techniques, such as ab initio modeling [360], comparative homology modeling [361], and threading [362], have allowed and transformed the protein structures from their sequences.

- Homology modeling: a recent computational technique for determining with high accuracy the three-dimensional structure of a protein from its amino acid sequence. This method relies on a suitable template structure [363]. It consists of several steps: The identification of a template, followed by sequence alignments, construction and refinement of the model of the target, and lastly the model validation [361]. Several tools are available, such as MODELLER [364], SWISS-MODEL [365], MODBASE [366], ProModel [367], etc.

- Protein threading: a valuable technique for protein structure prediction that is usually applied when a target protein has a low sequence or structural similarity with other proteins available in the PDB (<25% sequence identity). To improve the accuracy of the method, protein threading considers structural information such as secondary structure and solvent accessibility [368]. Several tools are available, such as I-TASSER [369], FUGUE [370], mGENthreader [371], Phyre [372], etc.

- Ab initio modeling: a method used if the target protein does not have any template structures available in the databases, like PDB [360]. This method reaches for the global or near-global minimum potential energy to identify the dihedral angle values for a specific protein structure, which relates to the structure's stability [373]. A valuable tool to perform this is ROSETTA [374].

Numerous drugs available in the market were developed thanks to SBDD methods, such as Amprenavir [375, 376], Raltitrexed [377], Norfloxacin [354], Dorzolamide [378], Isoniazid [379], and Flurbiprofen [380, 381].

The identification of the ligand-binding site is a prerequisite for performing targeted docking. This information can be gathered from X-ray crystallographic structures of proteins co-crystallized with ligands, substrates, or inhibitors [382] or site-directed mutagenesis studies. When a protein is without these types of information, several software and webservers can aid in predicting the binding sites, such as CASTp [383], DoGSite Scorer [384], NSiteMatch [385], DEPTH [386], MSPocket [387], MetaPocket [388], and Q-SiteFinder [389].

During the identification of the ligands, numerous ligands that cannot fit well into the binding site pocket are eliminated during the lead identification step.

## 2.2 Compound Library Preparation

Different drug-like compound libraries are available online [390, 391], but usually a custom-made library is required. Libraries may be divided into:

a) Libraries containing many generic compounds, useful for virtual high-throughput screenings (vHTS),
b) Libraries containing highly chemically diverse compounds,
c) Libraries containing target-oriented compounds,
d) Libraries containing compounds with specific molecular property diversity and property profiles (i.e., solubility, lipophilicity, etc.),
e) Libraries containing natural compounds
f) Libraries containing drug clinics or in the clinical stages of discovery, useful for a drug repurposing campaign.

Several chemical databases are available in which numerous chemical compounds can be selected, such as PubChem (N=111 million pure and characterized chemical compounds) [392], ZINC (N=230 million purchasable compounds) [393], MCULE (N=122 million synthetically accessible compounds) [394], ChemSpider ( N=25 million unique chemical compounds) [395], ChEMBL (>1.6 million distinct compounds) [396], and DrugBank (N=14528 drug molecules) [397].

Molecular docking is accomplished with drug-like compounds, which are selected and filtered using Lipinski's rule of five plus their ADMET parameters. Several other risk parameters can be used to improve the selection of drug-like compounds, such as carcinogenicity, hepatotoxicity, acute rat toxicity, serum glutamic oxaloacetic transaminase elevation, and inhibition of 3A4 oxidation of midazolam [359]. According to Lipinski's rule of five, a drug-like compound is orally bioactive if its physicochemical properties are between specific limits, such as molecular weight MW $\leq$ 500, partition coefficient between n-n-octanol and water logP $\leq$ 5, number of hydrogen bond donors HBD $\leq$ 5, and number of hydrogen bond acceptors HBA $\leq$ 10 [398].

Some normally used ADMET properties can be predicted, which comprise P-glycoprotein (P-gp) inhibition, human gastrointestinal absorption (HIA), cytochromes P450 (CYP) inhibition, blood-brain barrier (BBB) permeation, and plasma protein binding [399]. Another important property to consider is the synthetic accessibility of these compounds.

The compounds analysis can be filtered using several online tools, such as Chembioserver [400] and FAF-Drugs2 server; The latter one can filter compounds by their ADMET and physicochemical properties, while identifying key functional or undesirable moieties [401].

Additionally, the compounds in analysis must be without fragments such as counter-ions, metals, and solvent molecules. Furthermore, all compounds require assigned bond order and filled valences, partial charges, a right protonation state at physiological pH or at the pH of interest, and proper tautomeric states [402, 403]. This process is usually accomplished by a Ligand preparation process. This phase is usually performed by the docking programs themselves, such as Autodock Tools [404], LigPrep [405], MOE [406], MAPS platform [407], DISI [408], Pipeline Pilot [409], or Hyperchem [410].

### 2.3 Molecular Docking

Several software have been developed to "dock" chemical compounds into a biological target while predicting the geometry of the bound state and how strongly each compound can bind the binding site of the target. Molecular docking methods are largely applied in drug discovery, biological research, and many other fields. They are used to determine novel molecules that can strongly bind to proteins and nucleic acids. Moreover, they can be used to reveal how a ligand can influence the structural conformation and adaptability of the target [411]. Molecular docking approaches permit more efficient ligand optimization, a key part of drug discovery in which stronger binding, improved efficacy, reduced toxicity, and fewer side effects are required [411]. Docking is mainly used to predict interactions between proteins and small molecules, compounds with fewer than 100 atoms, but it can also be applied to perform protein-protein docking, protein-nucleic acid (DNA and RNA), or peptides and macrocycles [412, 413].

Molecular docking predicts the binding pose of the compound with the most energetically favorable geometry when bound to the target, determined by the 3D coordinates of each atom regarding the compound and the target. Moreover, docking estimates an associated binding affinity, which quantifies the degree of binding by the compound to the target [411].

The docking methods require two input files: a 3D structure of the target and the chemical structure of the compound/ligand.

In many cases, an approximate binding site needs to be indicated to the software, often specified by a 3D box within which the compounds are placed. This method, defined as targeted docking, allows the discovery of new ligands or gathers knowledge about where a specific ligand is likely to bind [411].

Docking methods are developed to predict binding poses that would match an equivalent structure resolved experimentally by crystallography or cryogenic electron microscopy (cryo-EM), and since our predictions are, of course, less accurate than experimental methods, but it is way faster and cheaper, and it does not require the synthesis of every compound under investigation [411]. Notably, the results of docking depend on the accurate structure of the target that is used, and the predicted binding pose and binding energy are estimations of the binding pose and binding energy for the specific target structure under analysis.

Several molecular docking tools are avaialable to perform protein-ligand interaction studies, such as Autodock [414], AutoDock Vina [415], GOLD [416], CDOCKER [417], FlexX [418], Surflex [419], GLIDE [420], DOCK6 [421], SwissDock [422], and HDOCK [423].

Molecular docking methods have two main components:

1) A scoring function that estimates the binding energy of the compound in a pose

2) A sampling procedure that looks at the space of potential poses to investigate the pose that has the most favorable score predicted by the scoring function.

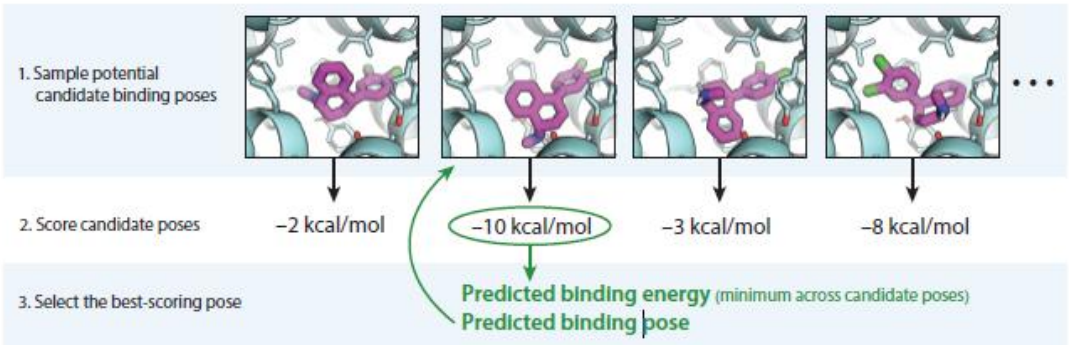



**Figure 5**: Description of scoring function and sampling procedure.

The pose with the most favorable score is classified as the predicted pose, and the score associated with the pose is the predicted binding energy.

### 2.3.1 Scoring Functions

Scoring functions assess a candidate binding pose, and they evaluate the energetic profile of the ligand in that pose. Most molecular docking software uses empirical scoring functions [424], in which specific terms describe several protein–ligand interactions that are important for determining the binding energy, and the parameters associated with these interactions are tuned to mirror experimental measurements. Moreover, most scoring functions include penalty scores to penalize overlapping atoms and electrostatic repulsion. In contrast, other interactions, such as electrostatic attraction, hydrogen bonding, and favorable van derWaals interactions, are rewarded.

Furthermore, scoring functions usually include ligand-only terms defined as ligand strain:

A) The entropic penalty is associated with a constraint that keeps the ligand in the binding pocket.

B) The enthalpic penalty of the ligand is associated with the adoption of a shape specified by the candidate binding pose.

These scoring functions are empirical and contain fit parameters to agree with experimental data. Specifically, scoring functions are adjusted to maximize the accuracy regarding predicting binding poses, binding energies, or the best binders between a set of compounds [425, 426]. The overall score is calculated considering all individual terms, and many other scoring functions are available [427].

### 2.3.2 Sampling Procedures

The sampling procedure evaluates the candidate binding poses to identify the one that should be predicted as most favorable by the scoring function. Particularly, the sampling procedure explores several degrees of freedom:

1. The position and orientation of the ligand relative to the target

2. The shape of the ligand itself in the pose, which could change based on internal rotations around single bonds

3. Optionally, local reorganization of the target structure

The sampling procedure catalogues all possible candidate poses while assessing the scoring function on each ligand to identify the most favorable score. Additionally, most advanced molecular docking tools use scoring functions to aid in steering the sampling procedure. For example, search algorithms like Markov chain Monte Carlo [428], exhaustive search using approximate scoring functions [429], or advanced genetic algorithms [430].

Usually, the sampling procedures employed in molecular docking do not reproduce the actual binding of the ligand. In fact, performing the actual binding on large compound databases would be extremely slow, specifically since a ligand can bind and dissociate from the target many times before remaining in its most favorable pose. The sampling procedure quickly goes through the space of possible binding poses to discover the global minimum. One important trade-off when choosing a docking method is the comparison between runtime and thoroughness in the search protocol. In fact, many molecular docking software are developed to run quickly without performing a thorough sampling for individual ligands [431]. This trade-off is critical for screening large compound libraries; In fact, when a high accuracy selection is required, the running time of the pose for every single ligand is much less of a concern. Some docking software packages, such as Glide from Maestro ®, allow the researcher to adjust the thoroughness of sampling and so the runtime [420].

### 2.3.3 Ligand Optimization

Binding poses can be predicted by docking and then used to lead to an optimization of a ligand's binding affinity by structure modifications, such as chemical additions, deletions, or substitutions to make the ligand more affinative to the binding site. The ligand affinity can be estimated by calculating the binding free energy values. These, predicted and obtained from molecular docking, can be used to compare chemical modification of the same ligand, though with low accuracy [432]. Moreover, poses predicted by molecular docking are mainly used as starting points for further simulations, including MD simulation–based methods, aimed at increasing the accuracy of the ligand binding free energy values.

In other words, molecular docking and predicted binding poses can offer valuable insights into the structural mechanisms underlying how ligands modulate target function. Specifically, when both agonist and antagonist ligands are discovered for a target protein, the variations in their docked poses can highlight the interaction profile of each molecule [433, 434].

Furthermore, docked poses can furnish information about how receptors can recognize their endogenous ligands and how those ligands influence receptor function [435, 436]. Finally, molecular docking can provide a starting pose for MD simulations [437–441]. This is a powerful combination, because molecular docking can constitute a static pose, without analyzing structure rearrangements, while MD simulations can provide information related to the movements of ligand-target complexes [442], simulating the real induced-fit effects.

### 2.3.4 Molecular Dynamics Simulations Approaches

A major limitation of classic molecular docking methods, as well as flexible receptor docking, is that each score is computed based on a single ligand pose and target conformation. To evaluate the motion of the ligand and the receptor, MD simulations, which simulate the motions of all atoms in the molecules of interest [442], can drastically increase the accuracy of docking, at the cost of demanding more computational resources. MD simulations are becoming more accessible, thanks to the improved availability of graphics processing units (GPUs), and more accurate, thanks to upgrades in the quality of the molecular force fields [443, 444].

The best way to use MD simulations to predict a binding pose and energy is by running a long simulation of the ligand-target complex to obtain a stable bound state of the ligand in the binding pocket of the target. The alchemical MD simulations method can estimate the difference in binding free energy between a pair of similar ligands [445]. This technique requires information regarding the binding poses, and it can predict the binding energies spread around a family of structurally similar ligands with higher accuracy than traditional docking scoring functions [446].

The estimation of the absolute binding free energies of individual ligands can be accomplished for each binding pose [447, 448]. These methods rank chemically diverse ligands in virtual screening, but they are substantially less accurate and more computationally intensive than those used to compute relative binding energies of similar ligands. MD simulation–based methods suitable for binding pose prediction have been proposed [449, 450]. These methods use customized simulation procedures that are much less computationally expensive than alchemical methods to assess the stability of several candidate binding poses.

### 2.3.5 Machine Learning Approaches

Recently, Machine learning (ML) methods have been used to improve the accuracy and applicability of docking. The application of Machine learning scoring functions (MLSFs) improves the accuracy of the empirical scoring functions used in classical docking [451–456] because several new parameters are included in the scoring functions compared to empirical scoring functions.

MLSFs learn directly from the data by using a simple representation of a candidate binding pose to grasp information, such as 3D coordinates and the elements of the atoms [457], and can consider receptor flexibility into the evaluation implicitly [458, 459]. MLSFs' development has been aided by training and benchmarking datasets [460].

Another method to accelerate VS approaches is by approximating the molecular docking process by applying ML models [461–463].

Moreover, various generative modeling protocols and genetic algorithms can predict output compounds compatible with a target with favorable docking scores [464, 465].

The advancement of this technology could break through classic CADD methods, improving the results while saving computational time.

## 2.4 Binding Free Energy Prediction Methods

The equation that describes the Ligand-Receptor complex interactions is:

$$\boldsymbol{L + R \leftrightarrow RL}$$

**L**: Ligand, **R:** Receptor, **RL:** Ligand-receptor complex

The binding strength or affinity can be measured to evaluate the stability of the RL complex relative to the unbound states associated with L and R on their own. The binding strength is determined by the binding free energy, usually called $\Delta G_{bind}$ [466]. The relationship between the binding free energy and the equilibrium dissociation constant $K_d$ is described by the logarithmic formula:

$$\Delta G_{bind} = RT \ln K_d$$

**R:** The universal gas constant (1.987 cal/mol K).

**T:** Absolute temperature in Kelvin.

**K:** The concentration of ligand at which 50% of the receptors are occupied.

A more negative $\Delta G_{bind}$ value is associated with a more stable complex and a higher binding affinity, corresponding to a smaller equilibrium dissociation constant ($K_d$). The binding free energy relies on two forces:

$$\Delta G_{bind} = \Delta H - T\Delta S$$

**ΔH:** Enthalpy

**T**: Absolute temperature in Kelvin

**ΔS**: Entropy

The enthalpy (**ΔH)** represents the heat released or absorbed by the system when the ligand interacts with the target while forming specific bonds with the receptors, such as hydrogen bonds, van der Waals forces, and electrostatic interactions.

When the enthalpy value increases by a large amount, it suggests that there is a high degree of shape complementarity between ligand and receptor, which can be described with the induced fit.

The entropy **ΔS** represents the change in the disorder of the system. After binding, the ligand can lose conformational, translational, and rotational entropy due to the reduction of its degrees of freedom inside the binding pocket. During the entry of a hydrophobic ligand into the binding pocket or its binding to the target, it releases water molecules, causing an increase in disorder and solvent entropy. This gain in solvent entropy, defined as the hydrophobic effect, is a major driving force in biomolecular recognition.

An accurate determination of binding free energy is one of the most significant tasks in biomolecular studies; in fact, it drives all molecular processes, such as protein folding, molecular association, and chemical reactions [467].

From a biomolecular system, it is possible to measure the thermodynamic properties by several expensive and time-consuming experiments. So, the application of precise theoretical calculations of binding free energies is crucial in SBDD and LBDD [468-472], protein folding [473, 474], conformational changes of macromolecules [475], solubility of small molecules [476, 477], protein-protein interactions (PPIs), protein-ligand binding affinities and interactions [478-483], etc. In these areas, the binding free energy difference (ΔG) is calculated for an event or the relative free energy of two states, and it is usually used in drug design to describe the binding strength between a receptor and a drug molecule through numerous available theoretical methods, often with a trade-off between accuracy and efficiency [484-486].

3 main approaches used in drug discovery are: Alchemical free energy (AFE), [486-489] molecular mechanics Poisson−Boltzmann surface area (MM/PBSA), and molecular mechanics generalized Born surface area (MM/GBSA) [489-491].

AFE methods are theoretically rigorous and accurate, which necessitate the interconversion of the system from an initial state to a final state through infinitesimal alchemical changes of the energy function. Moreover, AFE methods have a slow convergence of the ΔGs, and they require a high computational cost [487]. The convergence of the ΔGs is difficult in systems characterized by slow conformational transitions or massive environmental reorganizations [487].

There are two techniques mainly used for AFE computations: Free energy perturbation (FEP) [492-496] and thermodynamic integration (TI) [497-500].

End-point free energy methods [501-505] are based on samplings of the final states of a system, computationally less expensive than AFE methods and more accurate than most docking scoring functions, and the most well-known endpoint-free energy methods are MM/PBSA and MM/GBSA [489-491].

MM/PBSA and MM/GBSA methods can quantitatively characterize the binding details of ligand−receptor systems and are mainly used for small-molecule drug design, SBVS, and lead compound optimization [506-512]. These Endpoint free energy methods are advantageous for the prediction of the interactions between macromolecules, such as protein-protein [513-521], protein−peptide [522-527], and protein−nucleic acid interactions [528-539].

Moreover, these methods can be used to re-score docking poses to identify the correct binding poses, since they are more precise, and rank the ligands under investigation by their binding affinities [540-549].

**Table 2**: Description of the different binding free energy methods, pros, and cons

| Method | Accuracy | Cost | Typical Use |
|---|---|---|---|
| FEP / TI | Very high | Very high | Lead optimization |
| MM/PBSA | Medium–High | Medium | Absolute ΔG |
| MM/GBSA | Medium | Low | Ranking / rescoring |

Both methods are mainly used to evaluate and/or identify [550, 551]:

A) Docking poses

B) Predicted binding affinities

C) Hotspots

D) Structural stability

E) Contributions from single residues

F) Energy terms by free energy decomposition analysis

G) Dominant interactions in the binding process.

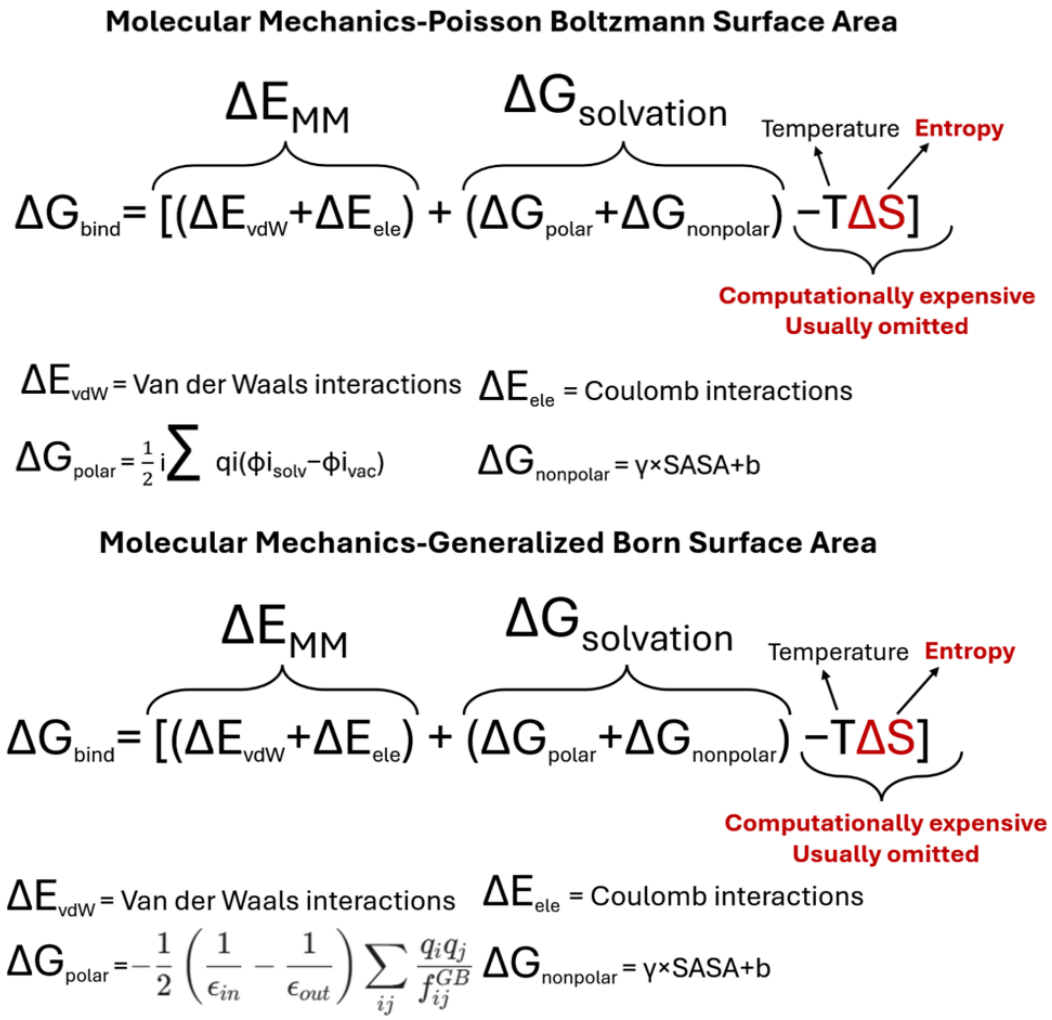



**Figure 6**: MM/PBSA and MM/GBSA formulas and components

MM/PBSA and MM/GBSA share most of the components, while only the calculation of the polar contribution is specific to the two methods. Meanwhile, nonpolar energy is estimated by evaluating the solvent-accessible surface area (SASA) [552]. Between them, the MM/GBSA method is much faster in resolving the polar solvation energy equation.

The conformational entropy ΔS calculation is performed on conformational snapshots gathered from MD simulations. However, this calculation has a high computational cost, and in fact, changes in ΔS are usually ignored when only the relative binding free energies of similar ligands are needed.

To obtain good results in the calculation of the ΔG, an explicit solvent model applied in MD simulations is vital [553]. Additionally, numerous published research papers demonstrate that the results of the MM/PB(GB)SA methods are dependent on the system and length of the MD simulations [554]. In fact, when the number of frames under analysis is low, it is hard to evaluate with MM/PB(GB)SA the incidence of numerous substates/rare events. In that case, the ΔG may have a large standard error/standard deviation [555, 556], and to avoid that, longer or several independent simulations need to be carried out to obtain better results [556]. An improvement was achieved in MM/PBSA predictions by discriminating MD snapshots/frames through a pre-evaluation of protein−ligand complexes using an ML-based approach (SVMSP) [557]. Moreover, MD sampling can help improve predictions; meanwhile, minimized conformations can yield optimal predictions that may even surpass those obtained by MD simulations [558, 559]. Nevertheless, the ΔG values can be calculated by MM/PB(GB)SA on a single minimized frame or on numerous MD frames. When a single minimized structure is gathered from an MD simulation, the ΔG calculation costs much less computational time, but it ignores the dynamic components, determining predictions that are mainly dependent on the starting structures [466].

With MM/PBSA and MM/GBSA methods, the ΔG prediction strongly depends on:

A) Force field [559]

B) Charge model [559]

C) Continuum solvation method [560]

D) Interior dielectric constant [560]

E) Sampling method [558]

F) Conformational entropy [561]

The MM/GBSA method is largely used because it is highly computationally efficient, and it can give better performance in ranking the binding affinities for systems without metals, but it is worse than MM/PBSA in calculating the absolute binding free energy [467]. Moreover, MM/GBSA and MM/PBSA methods have numerous weaknesses. The conformational ΔS is usually computed by normal-mode analysis (NMA), and it can determine mistakes in the overall ΔG calculation and is usually very time-consuming [562-573]. Because of that, when structurally similar molecules are under analysis, and their conformational ΔS values are similar, this ΔS is ignored, and only the relative ΔG values of those ligands are calculated, causing a probable overestimation of the ΔG itself [467].

Furthermore, the estimation of ΔG from highly polar or charged molecules by MM/GB(PB)SA methods could not be precise, because the uncertainty in the calculation of the solvation energy is proportional to the polarity of the considered molecules [467, 574]. Moreover, Machine-learning approaches represent a promising strategy to improve MM/PB(GB)SA-based binding free energy calculation and structure prediction in the coming years, specifically when a large amount of data concerning high-quality experimental structures and binding data becomes available [466].

## 2.5 Artificial Intelligence and Drug Discovery

Artificial intelligence (AI) is a machine that relies on computers to learn from a huge amount of data, defined as training datasets. AI has numerous real-world applications thanks to the wide availability of new computer hardware, such as graphical processing units (GPUs) [575]. The recent and gradual improvement of this technology has allowed the AI methods to predict biological activities and toxicities of drug molecules [576]. Moreover, AI can be applied in drug discovery (protein folding, protein-protein interaction, de novo drug design, VS, QSAR, evaluation of ADMET properties) [577]. In fact, this groundbreaking technology has invaded drug discovery in all aspects of this process [578–580]. The two most powerful methods that are mainly used in drug discovery are Machine Learning (ML) and Deep Learning (DL) [581]. Several ML algorithms are available nowadays, but the ones in drug discovery are Random Forest (RF) [582], Naive Bayesian (NB) [583], and support vector machine (SVM) [584].


**Figure 7**: Application of AI in drug discovery [585]

### 2.5.1 Machine Learning

ML methods use algorithms to gather and collect data, learn from it, and then make a prediction regarding any new data sets. To accomplish these predictions, the AI model is trained using large and high-quality data to improve their performance and overall accuracy. Thanks to this information, the AI model learns how to perform the task required by the operators [575].

ML algorithms are mainly applied to decode problems for which a large amount of data and countless variables are available [575], and their predictive power depends on these data. To obtain good results, the data needs to be complete, accurate, and tuned to amplify the predictability when used for the training segment [575]. Another method to improve the quality of the ML model is by performing feature selection before model building [586].

ML methods can be applied in two main techniques: supervised and unsupervised learning.

In the supervised learning methods, the goal is to predict future values for specific categories or continuous variables [575], and the model is trained with known input and output data that are correlated by tight relationships, making it possible to predict forthcoming outputs for new inputs. The unsupervised methods are mainly used for exploratory purposes to create models that can perform clustering of the data, and they can detect hidden patterns or fundamental structures inside the input data [575].

At the moment, AI is applied in all stages of drug discovery and development, comprising clinical trials, and the development and application of ML algorithms are aiming to identify novel targets [587], supply stronger marks for target-disease associations [380], enhance design and optimization of small-molecule ligands [588], improve the understading of the underlying mechanisms of a disease, increase knowdledge of phenotypes associated with disease and non-disease conditions [589], create new biomarkers for drug efficacy, prognosis and progression [590], expande the available analyses of biometric and supplemental data from patients observations and wearable devices, augment digital pathology imaging [591] to extract high-quality information from pictures [575]. One of the main issues associated with AI is model overfitting, which occurs when the model acquires the input training data plus uncommon features that will be included in the final model, causing a decrease in performance of the model on new data. To reduce this phenomenon, resampling methods or restraining sections of the training data to use as a validation set are worth considering.

Several regularization regression methods can impose penalties on parameters as the model intricacy increases, forcing the model to simplify the data and not overfit. Another available protocol to significantly decrease overfitting is the dropout method [592], which removes sections in the hidden layer in a random manner.

When a model can't be built with a specific training set or cannot generate new data, we are observing an underfitting problem. Several gold standard datasets are available, together with an independent dataset, which can be used to generate well-performing models [575].

Moreover, ML algorithms can be applied to predict druggable genes that are associated with:

1) Specific diseases [593]
2) Regulation/homeostasis [593]
3) Ageing [594]
4) Alternate splicing and disease-specific variants [595]
5) Cancer/drug response and drug effects from omics data [596]
6) Proteins that can bind small molecules [597]

### 2.5.2 Deep Learning

Deep Learning (DL) is a subset of AI that mimics a brain's neural networks to learn from large amounts of data, allowing machines to solve complex problems, and it is the most promising in big data. Several DL methods are available, such as deep neural network (DNN), restricted Boltzmann machine (RBN), recurrent neural network (RNN), and convolutional neural network (CNN). Moreover, multi-level deep neural networks (DNNs) can be developed to detect features from large datasets of unlabelled or labelled training data [598]. DL methods' complexity and scale are way above any other AI methods; In fact, in neural networks, input features are given to an input layer, and after numerous nonlinear transformations using hidden layers, the predictions are created by an output layer (**Figure 8**).

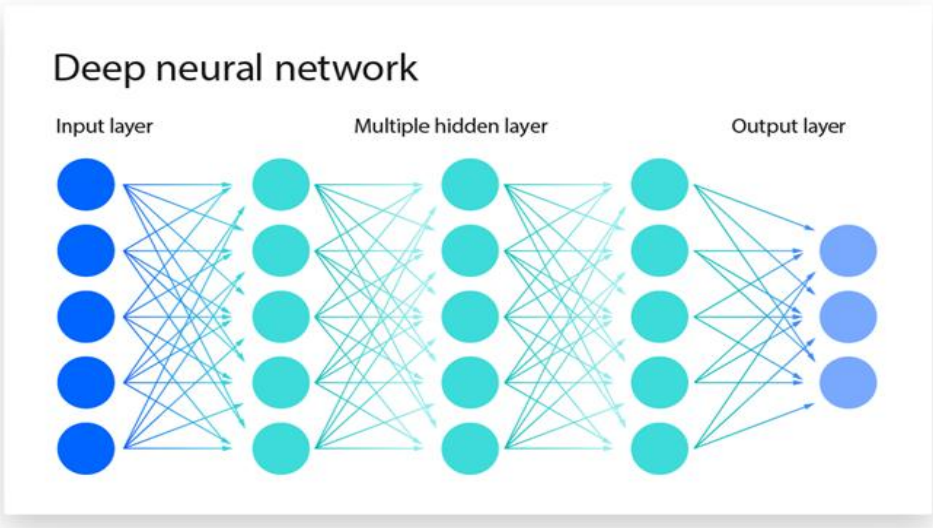



**Figure 8**: Scheme representation of a Deep Neural Network.

The backpropagation of errors allows for reducing the difference between the obtained and the expected values of the output, and this method is indispensable to optimizing DL algorithms. Each output node corresponds to a task to be predicted.

DNN algorithms can be applied in drug discovery, especially for bioactivity prediction [599], de novo molecular design, synthesis prediction, and biological image analysis [600].

As mentioned before, several different DL methods are available, and each of them is designed to perform a specific task.

Deep convolutional neural networks (CNNs) are mainly designed to work on spatial and image data and speech and image recognition. Structure-wise, they have some sections of the hidden layers locally connected to the next hidden layer. The main advantage of CNNs is that they can operate with high efficiency at capturing local patterns, such as edges or textures, while lowering the number of parameters under analysis. (**Figure 9**)

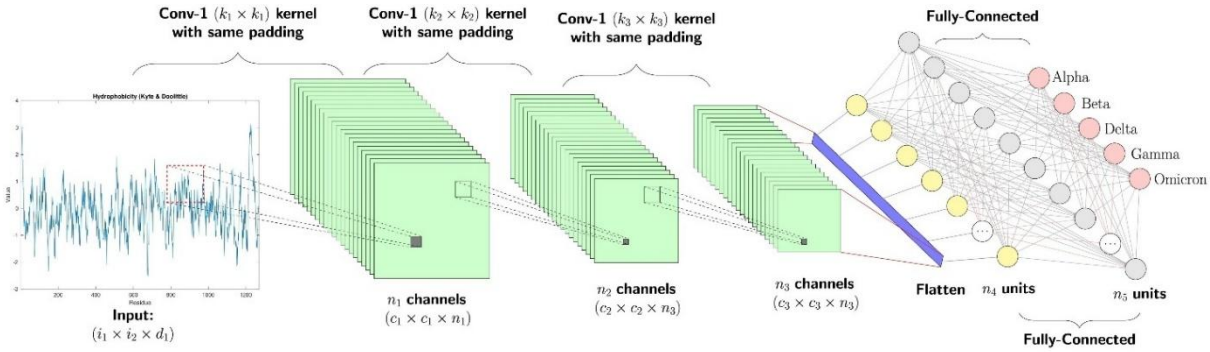



**Figure 9**: Representation in steps of the CNN methods

Recurrent neural network (RNN) algorithms are designed like a chain of multiple modules of NN that have connections between nodes, constituting a directed graph along a sequence. These methods possess “memory”, endowing them with the ability to process sequences of data, such as SMILES strings. They can be applied when it is important to predict the chemical properties of molecules based on the specific order of atoms/characters.

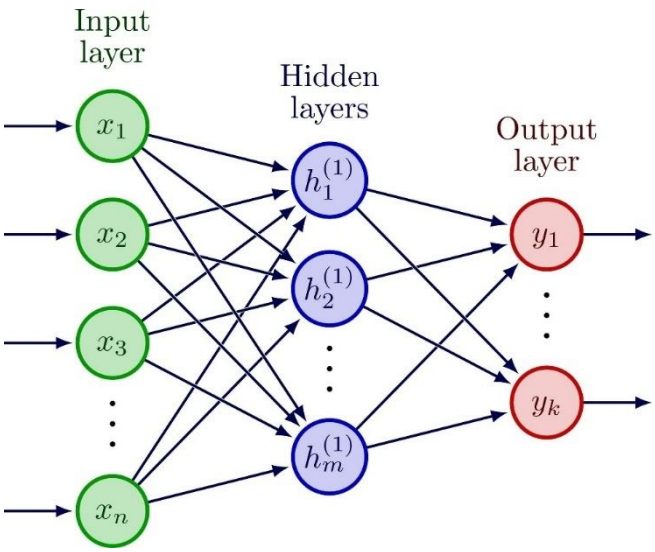



**Figure 10**: Representation of a simple RNN algorithm

Fully connected feedforward networks (FCFN) have a unique structure, where every input neuron is connected to every neuron in the next layer.

It's classified as a standard multi-layer architecture, and it can be used for basic classification and regression tasks in drug discovery. The key characteristic of this method is that data can flow only in one direction, without loops or cycles.

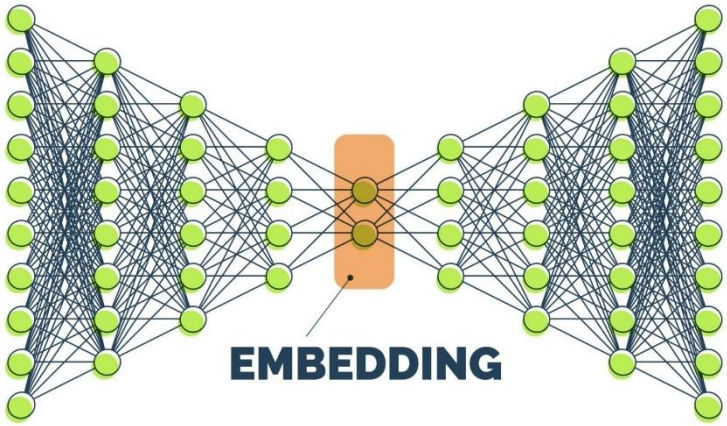



**Figure 11**: Representation of a classical FCFN algorithm

A deep autoencoder neural network (DAEN) is an unsupervised learning algorithm that aims to reduce the size of the system. To perform that, it uses backpropagation to project its input to its output [601], while maintaining the important random variables of the data, and removing the non-essential sections. This method is mainly applied to identify with good accuracy the most critical chemical features present in large datasets, performing high-quality, efficient screening processes.

Generative adversarial networks (GANs) contain two networks where one generates content and the other classifies that content. This system creates and discriminates the new content at the same time, making it possible to invent new molecular structures that could mimic the properties of known active ligands

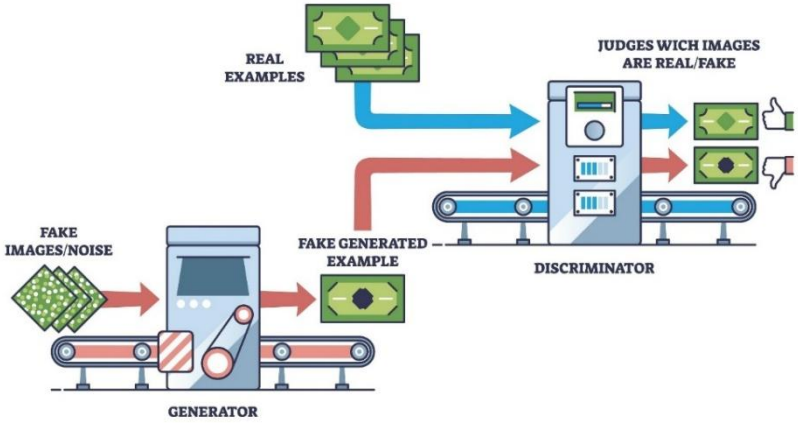



**Figure 12**: Illustration of a standard GAN algorithm

DNNs can be used on several occasions, such as drastically improving the predictive power when the aim is to understand the properties and activities of small molecules [602] and planning efficient routes of chemical synthesis. Another strong application of DL is molecular de novo design through reinforcement learning [603].

An example of the strong role of DL is well described by the researchers at AstraZeneca [603]. In fact, they used RNN methods to expand chemical space by fine-tuning a sequence-based generative model to design ligands with optimal solubility, bioactivity, pharmacokinetic properties, and other parameters.

Multi-task DNNs are optimal for predictions of lead identification/optimization, since they can synthesize data from numerous different biological sources [604] thanks to the presence of multiple nodes in the output layer. DL methods were applied for the first time in 2012 regarding the drug discovery process [605], attaining a higher accuracy prediction of ADMET properties compared to traditional ML methods [606].

Moreover, AI methods can also identify preclinical candidates at lower cost with high time-efficiency, while accurate predictions of binding affinity between a ligand and a target receptor remain challenging.

### 2.5.3 Limitations and Challenges

ML approaches are frequently used in several steps of the discovery and development pipeline by pharmaceutical companies. AI methods are data-driven, and their performance heavily relies on the amount and quality of the available data [575]. The requirements for the desired number and accuracy of the data are highly dependent on the complexity of the data themselves and the query to be resolved. Furthermore, the generation of these data sets could be expensive for pharmaceutical companies.

The large implementation of ML methods has numerous issues:

One of the main concerns is related to the lack of interpretability of deep-trained neural networks, making it difficult to obtain an appropriate explanation from the trained NN on how it arrives at the result. Another important issue for NNs is repeatability; in fact, ML outputs are extremely dependent on the order in which training examples are presented to the network and on the initial values/weights of the parameters (often chosen randomly).

This poor interpretability and repeatability reduce the possibility of using this powerful technique for biomarker identification, where different tools predict diverse prognosis biomarkers for breast cancer by using molecular expression signatures [606]. Since different ML methods can predict different results, it creates uncertainty about the adoption of these algorithms on a large scale. Data curation is key to providing reusable data, but it is an expensive process [607], and it requires advanced computational skills. Moreover, ML models can only predict within the known framework of the training data [575].

Soon, ML approaches will be applied to collect a large amount of data from several sources to drastically improve the predictive power of these methods, while helping medical decision-making concerning therapeutic benefits, clinical biomarkers, and side effects of therapies [575].

## 3. Design Of RNA-Based Drugs Influencing the Autophagy Process

Autophagy is an important, organized process that can selectively capture proteins and old or damaged organelles in eukaryotic cells and bacteria, using double-membrane vesicles called autophagosomes. When autophagosomes fuse with lysosomes, in the fusion stage, the captured components are degraded by the lytic enzymes and acidic environment that are available in the lysosome itself [608, 609].

The lysosomal breakdown releases several products, such as amino acids and fatty acids, which are recycled and transported into the cytoplasm, maintaining homeostasis [610].

When the organism is under metabolic stress or an insufficient nutrient supply, autophagy is rapidly upregulated to produce energy while providing building blocks for essential cellular functions.

When cells have low nutrient concentrations, such as glucose or amino acids, or are exposed to hypoxia, they depend on autophagy for their survival [611, 612]. Under physiological conditions, autophagy degrades harmful components such as damaged organelles, pathogens, and oxidized biomolecules (proteins, DNA, and lipids) in response to oxidative stress, preventing cell damage [613].

The autophagy regulators are involved in biochemical pathways that involve mTOR, class III PI3K (hVps34), Akt, V-ATPase, L-type $Ca^{++}$ channel, Calpain, proteasome, tyrosine kinases, histone deacetylase, and others [614–617].

To perform this process, several proteins are involved in the autophagy machinery called autophagy-related proteins (Atgs). The Atg8 family is responsible for autophagosome formation, and it facilitates cellular trafficking. In mammals, Atg8 proteins (mAtg8) are divided into two subfamilies: GABA-A receptor-associated protein (GABARAP) and microtubule-associated protein 1 light chain 3 (MAP1LC3), also known as LC3. The GABARAP subfamily consists of GABARAP, GARAPL1, and GABARAPL2, while the LC3 subfamily includes LC3A (LC3Aα and LC3Aβ as splicing variants), LC3B, LC3B2, and LC3C [611]. Each LC3/GABARAP family protein is characterized by 2 *N-terminal* α-helices and a *C-terminal* ubiquitin core [618].

The LC3 proteins have a primary role in cargo recruitment, while the GABARAP subfamily plays a crucial role in autophagosome closure and the recruitment of autophagy participants [609].

LC3 undergoes several post-translational modifications: it begins as pro-LC3 and is cleaved by the cysteine protease Atg4B to form its cytosolic isoform LC3-I.

During the initiation stage, LC3-I is then bound to phosphatidylethanolamine (PE) to attain LC3-II. This is finally positioned within the lipid membrane to constitute the autophagosomes [619]. The LC3-I and LC3-II levels in cells are the primary markers to evaluate autophagy activation.

LC3/GABARAP family has a pivotal role in the engulfment of different specific substrates through interaction with autophagy adaptor proteins such as sequestosome 1 (SQSTM1), neighbor of BRCA1 gene (NBR)-1, calcium-binding and coiled–coil domain 2 (NDP52/CALCOCO$_2$), BCL2/ adenovirus E1B 19-kDa interacting protein 3-like (NIX/ BNIP3L), optineurin (OPTN), and WD repeat and FYVE domain-containing (ALFY) [613]. These interactions are made possible by their LC3-interacting regions (LIRs) present on the ubiquitinated substrate/cargo and LC3/GABARAP family proteins [620].

The combination of LC3/GABARAP proteins and adaptor protein is therefore essential in cargo selection, targeting, and degradation [619]. RNA-binding Proteins (RBPs) can modulate autophagy by affecting the alternative splicing, stability, or translation of ATG mRNA [621-624]. When autophagy machinery is dysregulated, various diseases are associated, including neurodegenerative disorders [625], cardiomyopathies [626], infectious diseases [627], type II diabetes mellitus [628, 629], hepatic steatosis [630], and cancer [619, 631, 632]. Internal or external environmental factors can trigger the corruption of the autophagy process, causing the possible development of these pathologies.

Moreover, autophagy has a role in cancer progression and depends on the phase and context of disease progression. It can play a pro-survival role, reduce cell death, and promote resistance to cytotoxic therapies, or it can be associated with cell death [633].

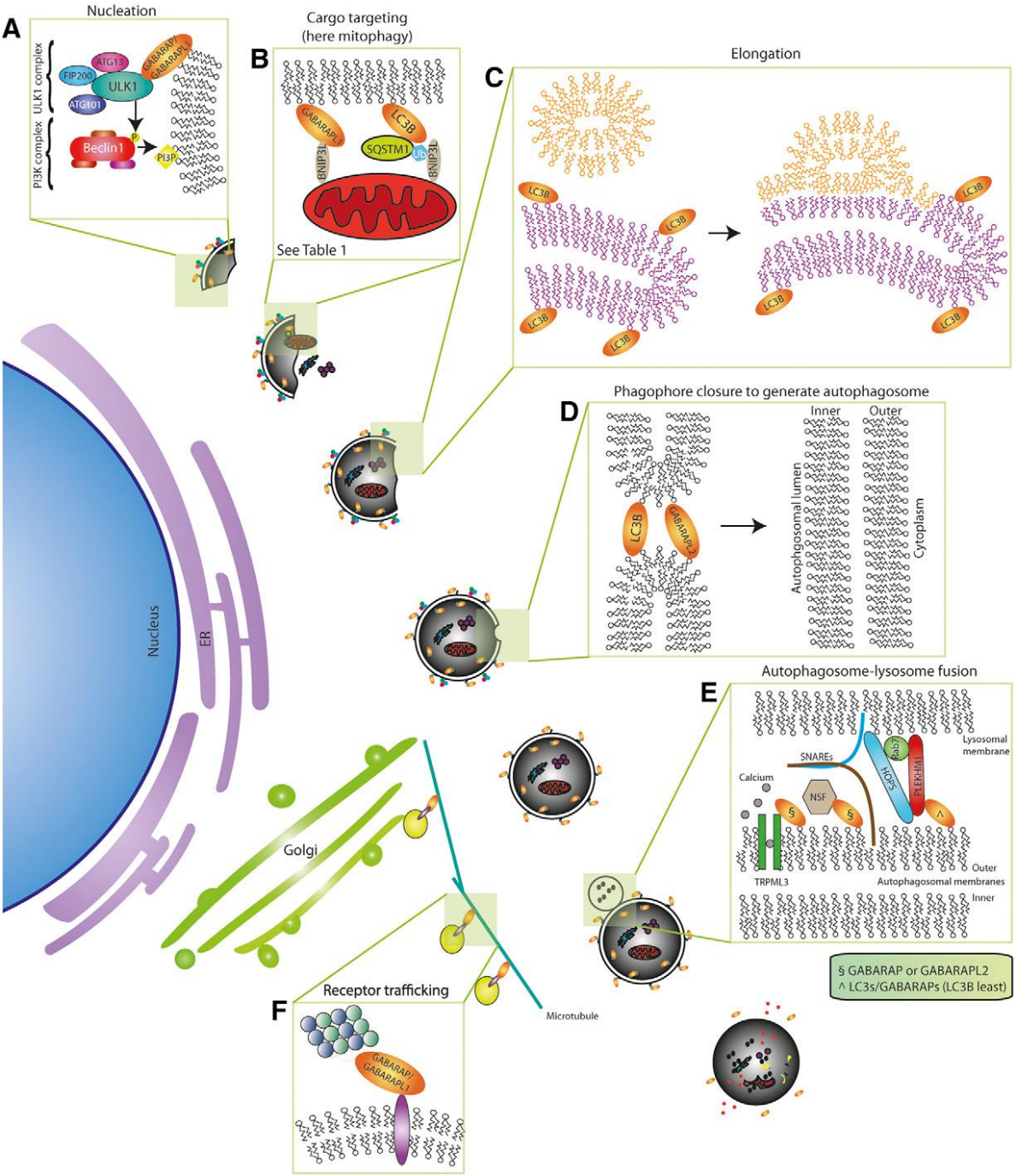



**Figure 13**. **Roles of LC3/GABARAP family proteins in autophagy-related processes.** A: nucleation phase is carried out by 2 protein complexes: the ULK1 and beclin-1 complex. The phagophore can engulf content either selectively or nonselectively. B: The selectivity in cargo tethering to the phagophore is established by the LC3/GABARAP family members. C: Elongation needs the delivery of membrane components (orange) to the growing phagophore (purple). D: Phagophore closure generates a closed double-membraned vesicle. E: Fusion of the autophagosome with a lysosome is necessary for cargo degradation. F: Role of GABARAP family proteins in receptor trafficking [619].

### 3.1 Autophagy/LC3B and Disease/Cancer Implications

The function of autophagy in cancer appears highly complex and may have opposite roles in different cancer cells, stages, and conditions. Autophagy plays a protective role by maintaining genome stability and limiting cellular alterations involved in malignant transformation. For this reason, a reduction in autophagy activity is observed in the early stages of tumorigenesis. Conversely, in the advanced and metastatic stages of cancer, autophagy increases, allowing tumor cells to survive and adapt to foreign sites. In addition, autophagy is activated in cancer cells exposed to various stresses such as anticancer treatments, leading to chemoresistance [634, 635]. LC3B is the most extensively studied Atg8 protein in humans. In fact, abnormal expression of LC3B is associated with autoimmune diseases, neurodegenerative disorders, cardiovascular diseases, and cancer [636–639]. The high level of LC3B expression is associated with key clinicopathological indicators of aggressive disease, highlighting its involvement in tumor progression and aggressive growth [620], and correlates with a poor prognosis in various solid tumors, including breast cancer, prostate carcinoma, and many others [640–649].

### 3.2 LC3B: Structure and Functions

LC3B localizes in the nucleus, cytoplasm, endomembrane system, lipid droplets, and autophagosomal membrane [650, 651] and participates in a wide range of physiological processes. LC3B contains an LC3-interacting region (LIR) docking site (LDS) that is associated with the LC3-interacting domain (LID) present in LIR-containing proteins, thereby enabling interactions with adaptor proteins, autophagy proteins, cargo receptors, and modulating the autophagic process [652, 653]. LC3B expression is regulated by several transcription factors, including CCAAT/enhancer binding protein [654], E2F1 [655], c-Jun [656, 657], and sterol regulatory element binding protein 2 (SREBP-2) [658]. In addition to transcriptional regulation, LC3B can undergo several post-translational modifications that are pivotal to modulate autophagy.

LC3B phosphorylation is a key trigger for autophagosome orientation towards the nucleus and cargo degradation [659]. STK4 phosphorylates LC3B during autophagy, promoting its interaction with other proteins. This process acts as a directional "switch" governing autophagy [660]. Several specific amino acids can be phosphorylated (Threonine and Serine residues), determining different impacts on the autophagy process [661-666].

During nutrient-depleted conditions (starvation mode), Protein kinase A (PKA) activity is reduced, causing the rapid activation of LC3B, promoting autophagy [667, 668].
LC3B's role in autophagy is strongly impacted by the acetylation of Lys residues [668].
When LC3B is acetylated, it accumulates in an inactive form in the nucleus, while deacetylated LC3B is moved to the cytoplasm to initiate autophagy. These post-translational modifications regulate LC3B availability and the rapid induction of autophagy [668].

### 3.3 LC3B and mRNA Regulation

The molecular characterization of LC3B as an RNA-binding protein (RNA-BP) was discovered by its ability to interact with mRNA through an arginine-rich motif identified by the three arginine residues (Arg68, Arg69, and Arg70), which is a feature shared with many other RBPs [669-674]. LC3B prefers mRNAs containing the AAUAAA consensus motif, a 3' poly-adenylation signal found in the untranslated region (UTR) of most eukaryotic mRNAs [675, 676]. When LC3B interacts with target mRNAs, it triggers their rapid degradation, and this process is defined as LC3B-mediated mRNA decay (LMD).

It has been recently discovered that the protein PRMT1 is one of the main LMD substrates. PRMT1 protein is a negative regulator of autophagy, and its efficient mRNA degradation via LMD facilitates autophagy. LC3B does not interact with all the mRNAs possessing the AAUAAA motif [677], and additional structural motifs and protein-mediated cues determine the substrate selectivity.

In this manner, LMD shapes the transcriptome during early autophagy, eliminating the mRNA/transcripts that could suppress the autophagy process while creating a positive intracellular environment associated with the cargo degradation [676].

### 3.4 Design of XNA Targeting RNA-BD of LC3B

The discovery of LC3B as an RBP and the identification of the arginine-enriched RNA-binding motif provide an ideal and potential target for the design of LC3B inhibitors. These compounds could selectively interfere with the LC3B-mRNA interactions, thereby avoiding the LMD in cancer cells and the degradation of its substrates (including PRMT1 mRNA), while reducing the impact and efficiency of autophagy on cell survival.

Based on these insights, I focused on the rational design of RNA-based therapeutics targeting the LC3B RNA-binding motif. The project was executed using a fully *in silico* protocol based on molecular docking, MD simulations, MM/GBSA calculations, and the development of an AI model to predict novel candidates.

The LC3B computational model used in this study was constructed from the 3D coordinates of chain A of the LC3B complex, retrieved from the Protein Data Bank (PDB accession code 1V49 [650], **Figure 14**). The original LC3B sequence, which consists of 125 amino acids, undergoes the removal of amino acids 121 to 125 from the *C-terminus*. Therefore, the computational model used in this project represents the pro-LC3B state of LC3B (amino acids 1 to 120), hereinafter identified as LC3B for simplicity.

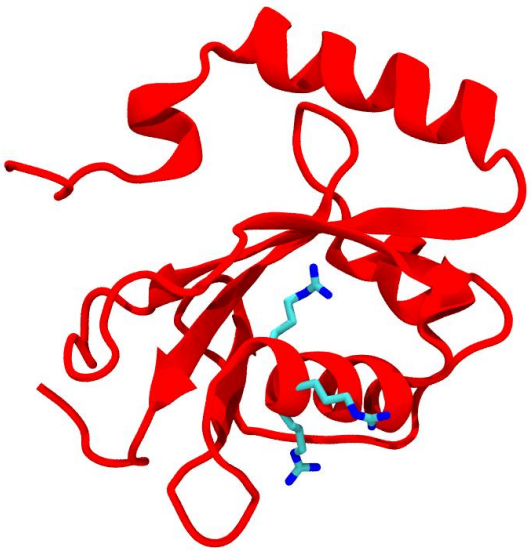


**Figure 14**. Ribbon representation of LC3B (PDB accession code 1V49 [650]). The protein structure is shown in red ribbons, while the three arginine residues (Arg68, Arg69, and Arg70), identified as the RNA-binding domain, are highlighted in cyan licorice representation.

In the first step, the sequence AAUAAA with a natural RNA backbone was designed by Maestro (Schrodinger Inc., USA, version 2021-2), using the *build tool*, after which a *protein preparation wizard* process was performed to optimize the structure retrieved from *PDB* and

the RNA sequence. This approach facilitated the initial steps of system setup, encompassing: (1) assessing the protonation states of residues at pH 7.4, (2) verifying residue completeness, (3) resolving atomic clashes, and (4) applying the OPLS4 force field.

Moreover, molecular docking calculations were performed using the HDOCK website [423] to dock the AAUAAA RNA sequence and LC3B to form the starting complex. The docking site was defined by the three arginine residues (Arg68, Arg69, and Arg70). After that, the RNA pose with the lowest binding free energy was selected (**Figure 15**).

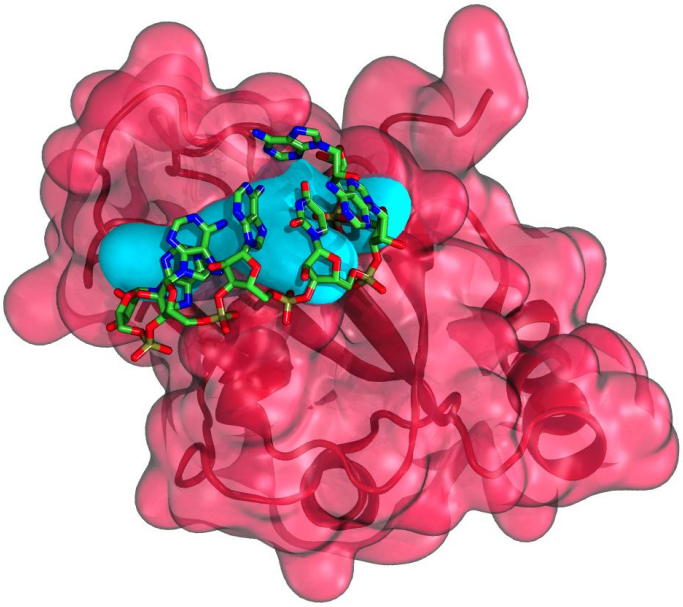


**Figure 15.** Predicted docking pose of LC3B in complex with the short RNA sequence AAUAAA, obtained using HDOCK. LC3B is shown as a red ribbon with a semi-transparent surface representation, while the RNA (AAUAAA) is displayed as a green stick representation. The RNA interacts with arginine residues Arg68, Arg69, and Arg70 (highlighted in cyan).

The resulting complex was solvated in a TIP3P water box [678], with 10 Å between the protein and the box edges. The system was then neutralized by adding a single $Cl^-$ counterion and, to avoid steric clashes, it underwent a multi-step energy minimization consisting of an initial steepest descent followed by a conjugate-gradient algorithm.

Additionally, the system was gradually heated from 50 K to 300 K with over 20 ps under constant volume conditions, followed by a 20 ps equilibration phase at constant pressure of 1 atm. This process was accomplished to adjust the overall system density.

Furthermore, two independent MD simulation replicas were carried out for each system.

The simulations were completed under constant temperature and pressure conditions, 300 K and 1 atm, respectively, using the Langevin thermostat [679] and Monte Carlo barostat [680]. The force fields used during these simulations were the ff19SB [681] for the protein, OL3 [682] for the RNA, and TIP3P [678] for the water molecules. Moreover, the MM/GBSA method was employed to estimate the ligand-binding free energy for each replica in each system. The analysis was conducted on 100 representative frames extracted from the stabilized portion of each MD trajectory, using a 10-frame stride over a total of 1000 frames.

Specifically, the frame selection was performed by visual inspection of the MD simulations themselves and by analyzing the ligand RMSD *vs* time plot. After MD simulations, the conformation representative of the most populated cluster, defined as the cluster centroid, was selected through the AMBER24/cpptraj cluster analysis tool [683].

Then, aiming at identifying new sequences capable of interacting more strongly with the LC3B-RNA-binding domain, a single-point mutation protocol was employed at position 3 of the starting RNA sequence (AAUAAA) using only natural nucleobases (Adenine A, Guanine G, Cytosine C, Thymine T, and Uracil U) (**Figure 16**). This approach was also useful for evaluating the sensitivity of our computational method in estimating the effects of a single RNA sequence mutation. In parallel, to address the RNA backbone pharmacokinetic issues, PMO and PNA analogs were designed and evaluated using the MD simulations/MM-GBSA protocol previously described (**Figure 16** and **Table 3**).

**Table 3:** Average binding free energies (ΔG, kcal/mol) and standard error of the mean (SEM) were calculated using the MM/GBSA method in AMBER24.

| Sequence | Avg ΔG ± SEM (kcal/mol) |
|---|---|
| **RNA** | |
| AAUAAA | −36.1 ± 0.8 |
| AACAAA | −38.5 ± 0.8 |
| AAGAAA | −33.2 ± 0.9 |
| AAAAAA | −48.7 ± 0.9 |
| **PMO** | |
| AAUAAA | −12.7 ± 0.9 |
| AACAAA | −19.1 ± 0.4 |
| AAGAAA | −18.9 ± 1.6 |
| AAAAAA | −25.2 ± 1.3 |
| AATAAA | −17.1 ± 1.4 |
| **PNA** | |
| AAUAAA | −21.9 ± 0.3 |
| AACAAA | −46.0 ± 0.3 |
| AAGAAA | −29.3 ± 0.2 |
| AAAAAA | −13.4 ± 0.3 |
| AATAAA | −58.0 ± 0.5 |


**Figure 16**. Protocol applied to investigate the functional role of the uracil (U) residue, and to explore the potential of alternative nucleobases at this position, by designing analogs in which U was replaced with other natural nucleobases (defined as 'X').

The results obtained highlighted that single-point mutations and backbone modifications can drastically change the calculated binding free energy of the selected XNA sequences. More in depth, as reported in **Table 3**, the sequences with the PMO backbone yielded unfavorable binding free energy. Conversely, the PNA analogs showed the best results, since the AATAAA sequence exhibited a binding free energy of −58.0 ± 0.5 kcal/mol (**Figure 17**), a value 22 kcal/mol lower than that of the starting RNA sequence AAUAAA (−36.1 ± 0.8 kcal/mol).

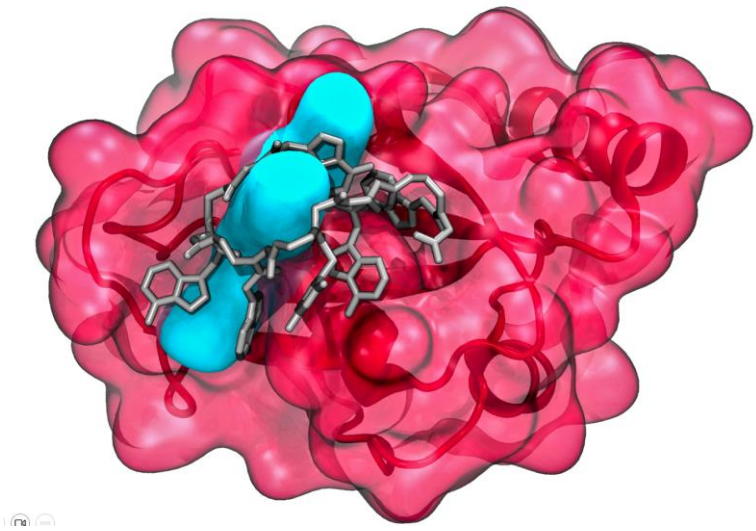


**Figure 17.** Structural representation of the PNA AATAAA (depicted in silver licorice representation)-LC3B complex (shown as a red ribbon with semi-transparent red surface representation). Highlighted in cyan are the arg68, arg69, and arg70 surfaces.

For this reason, the PNA backbone was then selected for further sequence optimization. In this study, the main goal was to identify PNA sequences that interact with the RNA-binding domain of LC3B with the highest complementarity and binding affinity. To this aim, the PNA AATAAA sequence was selected as a new template to create a new library generated by substituting each position with natural nucleobases, plus by implementing simultaneous modifications at all six positions (**Figure 18**).

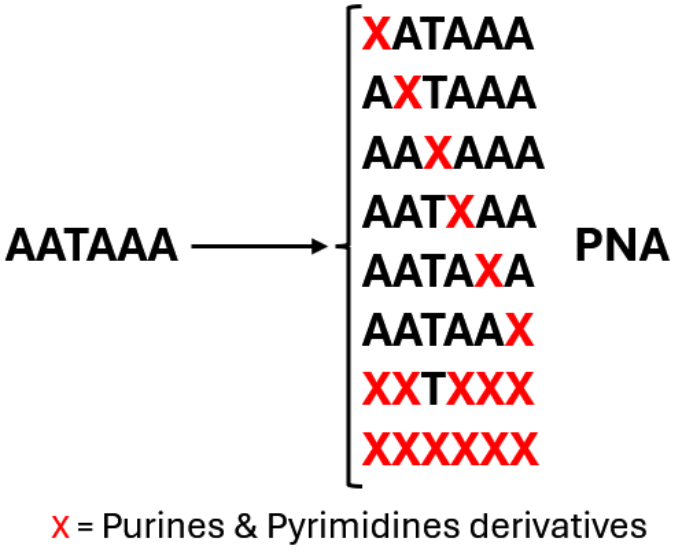



**Figure 18**: Systematic modification strategy of the AATAAA PNA template. A library of PNA sequences was generated by incorporating natural nucleobases, highlighted with the red X.

### 3.4.1 Design of PNAs and Development of an AI Model

The resulting library of PNA sequences was used as a training dataset to develop a new AI model to discover novel PNA sequences endowed with high binding affinities. This work was carried out in collaboration with Prof. F. Gentile and his lab members at the University of Ottawa, Canada. In this effort, the initial training set consisted of 42 PNA sequences, randomly selected from the residue mutations described in **Figure 18**.

Simulating them in complex with LC3B by MD simulations, the estimation of the ligand binding free energy suggested that a single point mutation can drastically change the binding mode and interaction profile, as highlighted by the ΔG values reported in **Table 4**.

**Table 4**. Average ΔG values for the PNA sequences training set, grouped by the most frequent nucleobase in each sequence (Thymine, Cytosine, Guanine, Uracil, Adenine), as well as several mixed-base combinations.

| Sequence | Avg ΔG ± SEM (kcal/mol) | Sequence | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|---|
| GATAAA | −21.2 ± 0.2 | CATAAA | −38.0 ± 0.5 |
| AGTAAA | −22.2 ± 0.4 | ACTAAA | −28.6 ± 0.3 |
| AAGAAA | −29.3 ± 0.2 | AACAAA | −46.0 ± 0.3 |
| AATGAA | −10.0 ± 0.3 | AATCAA | −28.7 ± 0.3 |
| AATAGA | −24.8 ± 0.3 | AATACA | −22.5 ± 0.2 |
| AATAAG | −29.1 ± 0.2 | AATAAC | −19.6 ± 0.2 |
| GGTGGG | −22.7 ± 0.3 | CCTCCC | −14.3 ± 0.2 |
| GGGGGG | −28.6 ± 0.2 | CCCCCC | −24.4 ± 0.3 |
| TATAAA | −27.9 ± 0.3 | UATAAA | −9.7 ± 0.2 |
| ATTAAA | −10.6 ± 0.3 | AUTAAA | −13.3 ± 0.2 |
| AATAAA | −58.0 ± 0.5 | AAUAAA | −21.9 ± 0.3 |
| AATTAA | −9.0 ± 0.3 | AATUAA | −13.3 ± 0.3 |
| AATATA | −18.0 ± 0.3 | AATAUA | −7.0 ± 0.2 |
| AATAAT | −22.8 ± 0.3 | AATAAU | −31.5 ± 0.3 |
| TTTTTT | −11.1 ± 0.3 | UUTUUU | −18.9 ± 0.2 |
| AAAAAA | −13.4 ± 0.3 | UUUUUU | −11.9 ± 0.4 |
| AATAUU | −14.4 ± 0.3 | AATTCC | −35.8 ± 0.4 |
| AATCGA | −34.2 ± 0.2 | AATTCG | −12.9 ± 0.3 |
| AATCGG | −40.4 ± 0.3 | AATUAT | −18.4 ± 0.4 |
| AATCTA | −21.5 ± 0.4 | AATUTU | −20.3 ± 0.3 |
| AATGCG | −17.6 ± 0.4 | AATUUU | −21.7 ± 0.3 |

Then, from each LC3B/PNA complex, a set of molecular descriptors and fingerprints was calculated for 1000 representative frames extracted from the MD simulations used for the MM/GBSA calculation. In this way, the 1D, 2D, 3D, and 4D descriptors were calculated only for the PNA sequences of each complex. Vega ZZ [684] supports the visualization and analysis of MD simulations, among other computational chemistry tasks, and CDK is a Java-based descriptor calculation tool. Below is a list of descriptors calculated by these tools.

- *Hybrid*: BCUT and WHIM
- *Constitutional*: AlogP, Acidic Group Count, Aromatic Atoms Count, Aromatic Bonds Count, Basic Group Count, Bond Count, Element Count, Largest Chain, Largest Pi Chain, Lipinski's Rule of Five, Longest Aliphatic Chain, Mannhold LogP, Molecular Weight, Rotatable Bonds Count, Xlog,
- *Topological*: Carbon Types, Chi Chain Indices, Chi Cluster Indices, Chi Path Indices, Chi Path-Cluster Indices, Eccentric Connectivity Index, FMF, Fragment Complexity, Hybridization Ratio, Kier & Hall SMARTS, Kier and Hall kappa molecular shape indices, Molecular Distance Edge, Moreau-Broto Autocorrelation (charge) descriptors, Moreau-Broto Autocorrelation (mass) descriptors, Moreau-Broto Autocorrelation (polarizability) descriptors, Petitjean Number, Topological Polar Surface Area, VABC Volume Descriptor, Vertex adjacency information magnitude, Weighted path descriptors, Wiener Numbers, Zagreb Index
- *Electronic*: Atomic Polarizabilities, Bond Polarizabilities, Charged Partial Surface Areas, Hydrogen Bond Acceptors, Hydrogen Bond Donors
- *Geometrical*: Gravitational Index, Length Over Breadth, Moments of Inertia, Petitjean Shape Indices

These data were used to develop the active learning (AL) AI model, an iterative machine learning approach that, unlike traditional supervised learning, trains a model on a static dataset before deployment. AL allows for the strategic selection of training samples in each iteration to optimize the model's performance. Moreover, AL is especially suitable for cases where we don't readily have access to the labels of our data or when labeling is time-consuming and expensive.

In AL, instead of labeling every sample at the beginning of training, specific samples are selected and labeled based on an acquisition function, and then the model is trained.

Acquisition functions determine which samples require labeling in each iteration of the active learning cycle.

In other words, the acquisition function defines the strategy that we use to determine which samples should be selected to be labeled and then added to the training set for the next cycle of training. The Input files for the AL are usually a labeled training set and a larger set of unlabeled samples. The outcomes after each iteration are sets of acquired samples from the unlabeled samples, which are sent to the domain expert to provide labels.

Generally, based on the goal of our study, different types of acquisition methods can be used. For instance, we use a very straightforward acquisition function called “maximum predicted probability” (MPP). This is an exploitative method, and as the name suggests, it selects the samples that the learning model is most confident about, i.e., from the unlabeled set, the samples with the highest predicted probability of belonging to a certain class are chosen for labeling. After each training round, information related to the MD simulations and molecular descriptors was calculated and used as input for the next step, improving, after each round, the AI model accuracy and precision. Specifically, we performed three rounds of iteration to improve the quality of the AI model, and all the sequences analyzed for each round are described in **Tables 5**, **6**, and **7**.

**Table 5**: Binding free energy values of the sequences predicted in the first round.

| Rank | Sequence | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|
| 1 | UAGGAG | −21.0 ± 0.3 |
| 2 | TGGGGG | −44.8 ± 0.3 |
| 3 | CAGGAG | −22.3 ± 0.3 |
| 4 | AAGGGG | −14.1 ± 0.3 |
| 5 | AAGGAG | −26.9 ± 0.3 |
| 6 | UGGGGG | −37.4 ± 0.3 |
| 7 | TGGAGG | −25.2 ± 0.2 |
| 8 | CGGAGG | −27.7 ± 0.2 |
| 9 | AGGGAG | −11.1 ± 0.2 |
| 10 | UAAAAG | −25.0 ± 0.3 |
| 11 | GGAAGG | −23.7 ± 0.2 |
| 12 | TAGAAA | −12.8 ± 0.3 |
| 13 | GAGGAG | −26.5 ± 0.2 |
| 14 | CAAAAG | −33.2 ± 0.2 |
| 15 | AGGGAA | −26.7 ± 0.3 |
| 16 | UAGAAG | −22.3 ± 0.3 |
| 17 | TGGAGA | −24.7 ± 0.3 |
| 18 | AAAAGA | −26.1 ± 0.2 |
| 19 | AAAAGC | −13.2 ± 0.3 |
| 20 | CAGAAG | −10.2 ± 0.3 |
| 21 | AGGAGG | −28.6 ± 0.2 |
| 22 | GAGGAA | −23.4 ± 0.4 |
| 23 | CGGGGG | −20.5 ± 0.2 |
| 24 | CAGAAA | −24.7 ± 0.3 |
| 25 | CCGAGG | −7.5 ± 0.2 |

**Table 6:** Binding free energy values of the sequences predicted in the second round

| Rank | Sequence | Avg ΔG ± SEM (kcal/mol) | Rank | Sequence | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|---|---|---|
| 1 | AGGGGG | −39.2 ± 0.2 | 26 | UCAGGG | −44.4 ± 0.3 |
| 2 | AGAAAG | −39.6 ±0.2 | 27 | CGACGG | −59.6 ± 0.3 |
| 3 | AAAAGG | −43.8 ± 0.3 | 28 | CGAAAA | −49.0 ± 0.3 |
| 4 | AGAAAA | −45.1 ± 0.4 | 29 | AGGGCG | −46.8 ± 0.3 |
| 5 | AAGGGA | −52.8 ± 0.4 | 30 | AGGGGA | −46.2 ± 0.3 |
| 6 | AAGGAA | −44.4 ± 0.3 | 31 | TGGGCG | −38.7 ± 0.3 |
| 7 | TAGGGA | −51.7 ± 0.3 | 32 | AAGAAG | −43.5 ± 0.3 |
| 8 | AGGAGA | −48.4 ± 0.3 | 33 | AGAGGG | −55.1 ± 0.3 |
| 9 | TAGAAG | −49.1 ± 0.4 | 34 | CAGCAG | −48.8 ± 0.3 |
| 10 | GGGAGG | −43.1 ± 0.3 | 35 | TCAGGG | −46.2 ± 0.3 |
| 11 | TAGGAG | −47.3 ± 0.3 | 36 | UAGAAC | −41.2 ± 0.3 |
| 12 | GGAGGG | −56.4 ± 0.3 | 37 | TAACAG | −51.9 ± 0.2 |
| 13 | UCAGAG | −46.5 ± 0.3 | 38 | TAAAAG | −43.8 ± 0.3 |
| 14 | AAAAAG | −51.7 ± 0.3 | 39 | AAGGCG | −51.8 ± 0.4 |
| 15 | CGGCGG | −47.3 ± 0.3 | 40 | GAGAAA | −41.7 ± 0.4 |
| 16 | AAAGGA | −44.7 ± 0.3 | 41 | CGGGAA | −39.8 ± 0.3 |
| 17 | CACAAG | −41.4 ± 0.3 | 42 | UCGAAG | −50.0 ± 0.4 |
| 18 | GAAGAG | −49.2 ± 0.3 | 43 | UCGAGG | −41.5 ± 0.4 |
| 19 | CGGAGA | −52.9 ± 0.3 | 44 | AACAAG | −49.2 ± 0.3 |
| 20 | UGGAGG | −44.3 ± 0.3 | 45 | GGGGGC | −45.0 ± 0.4 |
| 21 | AAGAGG | −43.1 ± 0.2 | 46 | CGGUGG | −43.2 ± 0.3 |
| 22 | CAGGAA | −44.0 ± 0.3 | 47 | UACGGG | −50.3 ± 0.3 |
| 23 | AAAGGG | −53.6 ± 0.3 | 48 | TGCAAG | −44.8 ± 0.5 |
| 24 | GAAAAA | −41.8 ± 0.3 | 49 | CGCGAA | −44.4 ± 0.4 |
| 25 | GAAAAG | −51.8 ± 0.3 | 50 | CGGGGU | −39.2 ± 0.3 |

**Table 7**: Binding free energy values of the sequences predicted in the third round

| Rank | Sequence | Avg ΔG ± SEM (kcal/mol) | Rank | Sequence | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|---|---|---|
| 1 | TGGCGG | −42.6 ± 0.3 | 26 | CGGAGC | −41.7 ± 0.3 |
| 2 | UGGCGG | −49.9 ± 0.4 | 27 | AGGGUG | −57.6 ± 0.4 |
| 3 | UGGAGA | −44.7 ± 0.3 | 28 | CACGGG | −53.0 ± 0.3 |
| 4 | CGGGCG | −46.8 ± 0.3 | 29 | UUGGGG | −46.7 ± 0.3 |
| 5 | UCGGGG | −40.1 ± 0.3 | 30 | TGGAGC | −49.7 ± 0.4 |
| 6 | CCAGGG | −51.4 ± 0.3 | 31 | TGGGGU | −51.3 ± 0.3 |
| 7 | UGGGCG | −57.9 ± 0.4 | 32 | TAGGCG | −39.9 ± 0.3 |
| 8 | UGGGAA | −48.6 ± 0.3 | 33 | UGGUGG | −42.7 ± 0.3 |
| 9 | AGGCGG | −59.6 ± 0.3 | 34 | CUGGGG | −42.4 ± 0.3 |
| 10 | TGGGGC | −51.1 ± 0.3 | 35 | CGAUGG | −41.4 ± 0.3 |
| 11 | CCGGGG | −40.0 ± 0.3 | 36 | UGGCGA | −60.3 ± 0.4 |
| 12 | TCGGAG | −37.8 ± 0.2 | 37 | TGGGUG | −48.1 ± 0.2 |
| **13** | **GGGGGA** | **−62.9 ± 0.4** | 38 | CGGGUG | −42.4 ± 0.3 |
| 14 | CGGGGC | −47.7 ± 0.3 | 39 | CCGGAG | −42.6 ± 0.2 |
| 15 | TCGAGG | −42.8 ± 0.3 | 40 | UGGAGC | −47.6 ± 0.4 |
| 16 | GGGGAG | −55.5 ± 0.3 | 41 | GGGAGA | −50.9 ± 0.3 |
| 17 | GAGGGG | −50.0 ± 0.2 | 42 | UAGGGC | −42.0 ± 0.3 |
| 18 | UGGGGC | −42.3 ± 0.2 | 43 | GAGGGA | −54.0 ± 0.4 |
| 19 | UGACGG | −57.3 ± 0.4 | 44 | CGGUGA | −36.5 ± 0.3 |
| 20 | TGACGG | −52.0 ± 0.3 | 45 | AGGGAC | −45.2 ± 0.3 |
| 21 | CAGGGA | −49.0 ± 0.4 | 46 | GGGGCG | −48.6 ± 0.4 |
| 22 | GCGAGG | −48.7 ± 0.3 | 47 | UAGGCG | −39.1 ± 0.3 |
| 23 | CGCGAG | −39.7 ± 0.3 | 48 | TACAGG | −51.1 ± 0.5 |
| 24 | UAGGGA | −49.4 ± 0.3 | 49 | GAGAAG | −44.8 ± 0.3 |
| 25 | AGCAGG | −42.4 ± 0.3 | 50 | AGGAAG | −47.2 ± 0.3 |

Finally, the sequence analysis of the top 25 sequences predicted by the AI model (**Figure 19**) shows that Guanine and Adenine are the most probable nucleobases at positions 2 to 6 for most rounds, while the first position shows high variety. This analysis established how structurally similar the best sequences were to each other, identifying a recurrent pattern.

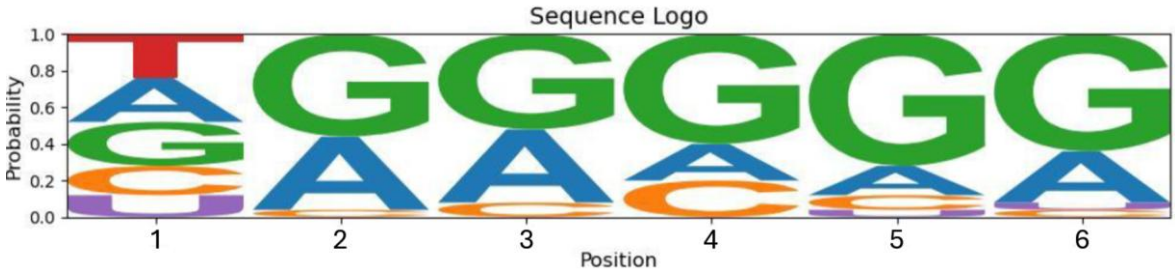



**Figure 19**: Sequence logo representing the positional nucleobase probability distribution calculated from the 25 sequences with the lowest binding free energy values.

Thanks to this approach, new sequences with lower binding free energy, compared to our starting template, were discovered, assessing how the AI model was progressively tuned to perform the task. The top 5 sequences found by this approach are depicted in **Table 8**.

**Table 8.** Average ΔG values of the top 5 sequences simulated in this project.

| **Sequence** | **Avg ΔG ± SEM (kcal/mol)** |
|---|---|
| GGGGGA | −62.9 ± 0.4 |
| UGGCGA | −60.3 ± 0.4 |
| AGGCGG | −59.6 ± 0.3 |
| CGACGG | −59.6 ± 0.3 |
| AATAAA | −58.0 ± 0.5 |

In conclusion, the PNA sequence identified by this approach was GGGGGA (−62.9 ± 0.4 kcal/mol) (**Figure 20**), which displayed an improvement of 5 kcal/mol in the calculated binding free energy value, compared to the starting AATAAA sequence (−58.0 ± 0.5 kcal/mol) (**Table 8**) and 26 kcal/mol, if compared to the RNA AAUAAA sequence (−36.1 ± 0.8 kcal/mol). These values look very promising, considering the high similarity among the natural nucleobase structures, and constitute a good starting point for the development of new PNA containing non-natural nucleobases.

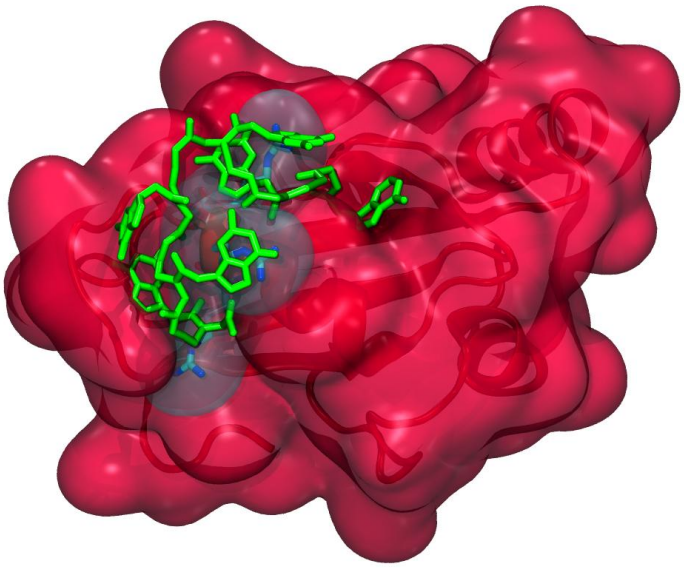


**Figure 20**: Structural representation of the PNA GGGGGA (depicted in green licorice representation)-LC3B complex (shown as a red ribbon with semi-transparent red surface representation). Highlighted in cyan are the arg68, arg69, and arg70 surfaces.

### 3.4.2 Synthesis, Biophysical, and Biological Assays

The synthesis of the PNA with sequence AATAAA was accomplished in Prof. Romanelli's lab, University of Milan. Once available, it was tested using Microscale Thermophoresis (MST), a versatile, robust, quantitative method that is largely used to quantify the interaction between interacting molecules, such as proteins and small molecules. This technique detects molecule interactions by quantifying the thermophoretic movement of fluorescent molecules in response to a temperature gradient. The effect of different concentrations of the nonfluorescent ligand on the movement of the fluorescent molecules is quantified to assess the interaction between molecules. If the fluorescent molecule interacts with the ligand, the molecular properties of the compounds, such as charge, size, and hydration shell, will influence the molecular motility [685].

In these experiments, the average $K_d$ value obtained testing the PNA with sequence AATAAA was 13.0 ± 7.8 nM, a very low value supporting the goodness of the computational design approach.

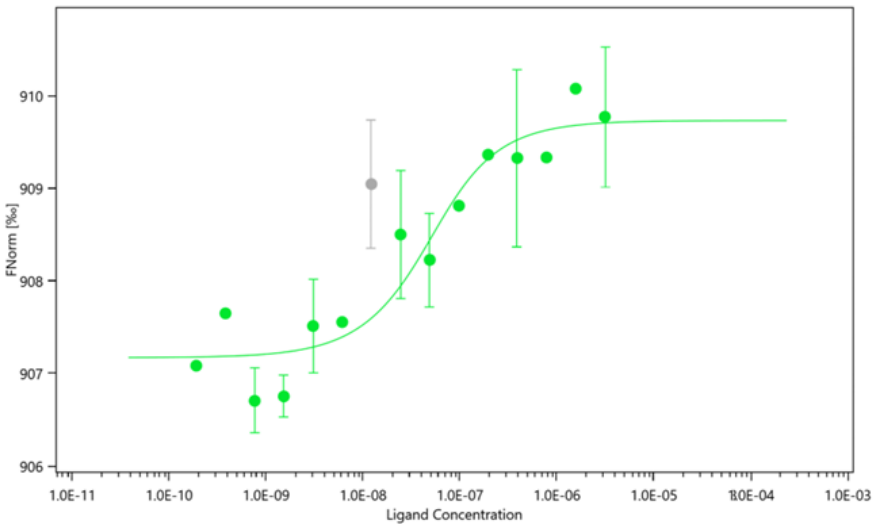



**Figure 21.** The figure illustrates a dose–response binding curve obtained from the MST experiment evaluating the interaction between PNA sequence AATAAA and LC3B. The x-axis reports the ligand concentration on a logarithmic scale (from ~$10^{-11}$ to $10^{-3}$ M), while the y-axis displays the normalized fluorescence (FNorm) in ‰ units. The green circles represent the individual measurements, measured FNorm values at each titration step, with vertical error bars showing the variability between replicates. A single outlier-like point is displayed in grey.

Furthermore, viability assays were performed on PNT2 cells by M. Garofalo's group, University of Padua, and PC3 cells by R. Moretti's Lab, University of Milan.

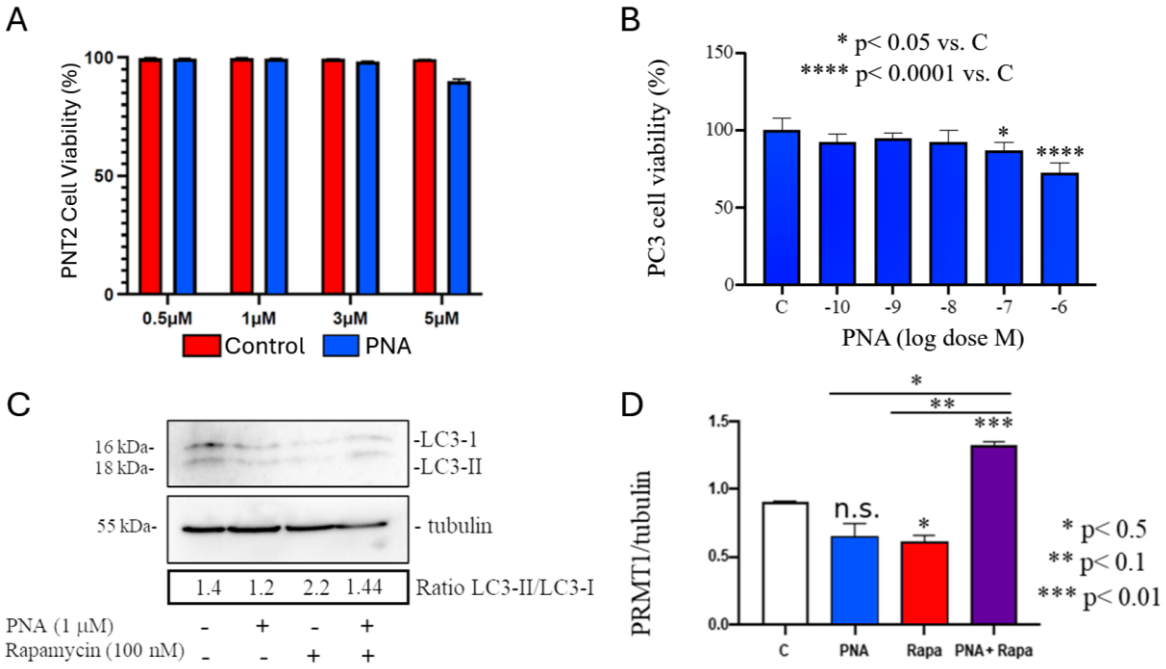



**Figure 22**: **A)** Cell viability of PNT2 cells treated with increasing concentrations (0.5–5 μM) of control or PNA AATAAA. B) Cell viability assay performed on PC3 cells treated with PNA AATAAA (0.1 nM–1 μM), with a considerable effect at 1 μM. C) Western blot analysis of LC3-I/LC3-II and LC3-II/LC3-I ratios after treatment with PNA AATAAA (1 μM) and/or rapamycin (100 nM); tubulin was used as a control. D) PRMT1/tubulin ratios after treatment with PNA (1 μM), rapamycin (100 nM), or both, showing raised PRMT1 levels with the combination of both compounds.

The attained results suggested that the prototypical PNA with sequence AATAAA displayed several significant features, starting with its safety on PNT2 cells (**Figure 22A**) and the cytotoxicity on PC3 cells (**Figure 22B**). Furthermore, the ligand showed the ability to interfere with the induced autophagy flux, promoted by rapamycin (100 nM), by strongly reducing the LC3-II/LC3-I ratio (**Figure 22C**). Moreover, the PNA AATAAA ability to reverse the decrease of PRMT1 expression induced by rapamycin confirmed that it can interact with the RNA-Binding domain of LC3B (**Figure 22D**).

To conclude, other biological assays are ongoing to investigate the PNA's ability to modulate the autophagy machinery. The synthesis of PNA sequences endowed with the best predicted affinity to LC3B, identified by the AI model, is ongoing in Prof. A. Romanelli's lab.

Overall, this project demonstrated that an *in silico* approach, integrated with experimental validation, can design useful ligands that target and modulate the autophagy machinery in prostate cancer cells, constituting a starting point for the design of new and innovative anticancer RNA-based drugs.

## 4. Design of RNA-Based Drugs as HMGB1 Inhibitors

High mobility group box 1 (HMGB1) is a highly conserved nucleoprotein belonging to the group of nonhistone chromatin-associated proteins. It was first extracted from calf-thymus chromatin in 1973 and named for its high mobility in gel electrophoresis [686]. HMGB1 is a multifunctional protein involved in a variety of cellular biological properties, depending on its subcellular localization, post-transcription modifications, and affects numerous cellular processes by modulating chromatin structure [687, 688]. Participating in the regulation of transcription, chromatin remodeling, recombination, and DNA repair [689-691], HMGB1 can bind several proteins, and these interactions are important for HMGB1's activity and function [692].

To play these functions, the human HMGB1, which comprises 215 amino acid residues, contains: 1) two DNA-binding domains: the HMG-A box domain (sequence 9–79 amino acids) and the HMG-B box domain (sequence 95–163 amino acids) [693], 2) a nine amino acid loop connecting the And B boxes, 3) a highly disordered negatively charged *C-terminal* tail (composed of continuous aspartate and glutamate residues) (sequence 186–215 amino acids) [694], and 4) a functionally important *N-terminal* region.

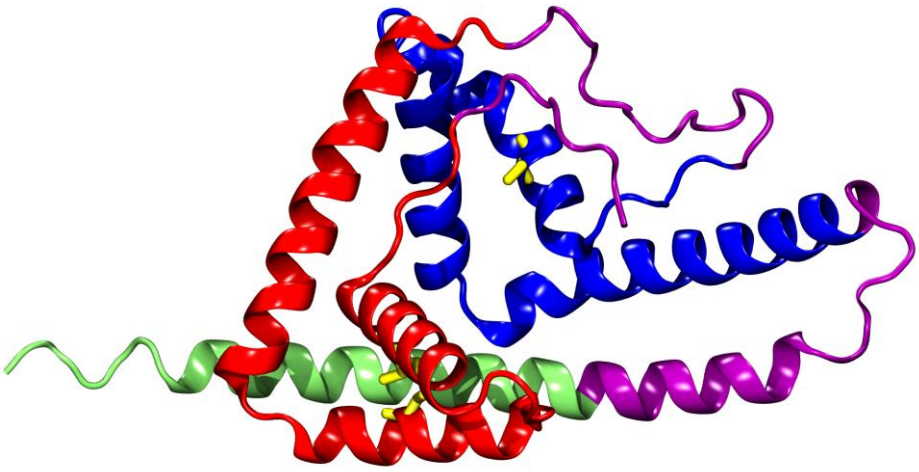


**Figure 23**: Ribbon representation of the human HMGB1 protein predicted by Alphafold 3. The structure comprises two DNA-binding domains, A Box (red) and B Box (blue), a long tail (green) connected by a flexible linker (purple). The cysteine residues are highlighted in yellow.

The two HMG boxes of HMGB1 are structurally similar to a characteristic DNA-binding domain consisting of 3 alpha helices (helix I, helix II, and helix III) and two loops (loop I and loop II), which then arrange in an "L" shape with an angle of 80° between the two arms [695-698].

Compared with the B box, the A box has a higher alpha-helix content and is more positively charged and straight than that found in helices I/II [699, 700]. The short arm contains helix I and helix II, whereas the long arm contains helix III and an *N-terminal* unstructured segment in parallel with the helix.

### 4.1 HMGB1 Localizations

HMGB1 can be found in four main compartments:

***Nucleus:*** Nuclear HMGB1 acts as a DNA chaperone with DNA-binding and bending activities, regulating many key DNA events and being involved in DNA repair, transcription, telomere maintenance, and genome stability [701, 702]. HMGB1 can translocate from the nucleus to the cytoplasm after post-translational modifications, including acetylation, phosphorylation, and methylation [687]. HMGB1 is normally located in the nucleus and translocates to the cytosol, including mitochondria and lysosomes, in response to various stressors (e.g., cytokines, chemokines, heat, hypoxia, $H_2O_2$, and oncogenes). The loss of HMGB1 or increased HMGB1 translocation from the nucleus to the cytoplasm could increase DNA damage, decrease DNA repair efficiency, and increase cell death in response to chemotherapy, irradiation, and oxidative stress [692].

***Cytoplasm:*** The main function of HMGB1 in the cytoplasm is to act as a positive regulator of autophagy, by binding with proteins involved in this pathway [703]. Moreover, it is involved in immune responses by increasing autophagy, inhibiting apoptosis, regulating mitochondrial function [704], and regulating autophagy in cancer cells [705].

***Membrane:*** HMGB1 has been reported to be present on cell surface membranes involved in neurite outgrowth and axonal sprouting [706], platelet activation [707, 708], cell differentiation [709], erythroid maturation [710], adhesion [711], cell migration [687], and innate immunity [712].

***Extracellular:*** Extracellular HMGB1 is involved in numerous processes, such as inflammation, immunity, immune cell migration, invasion, proliferation, differentiation, antimicrobial defense, and tissue regeneration [692].

HMGB1 is a danger signal in inflammatory conditions, including autoimmunity, cancer, cell growth, and cell proliferation/death [713, 714], and is massively released into the extracellular space by dead or dying cells.

During apoptosis, necrosis, cell scorching, and injury resulting from various stimuli, such as irradiation, hypoxia, and hyperthermia [715], soluble HMGB1 can be released into the extracellular space in large quantities by these dead or dying cells and elicit inflammatory responses in vitro and in vivo [716].

Several studies have demonstrated that endogenous stimuli, exogenous microbial products, and infections with various pathogens can induce active secretion of HMGB1 by immune cells, endothelial cells, epithelial cells, fibroblasts, or other cells [717-720].

HMGB1 induces the release of cytokines and chemokines from immune cells, mainly caused by ligand-receptor interactions. In monocytes and neutrophils, HMGB1 induces the secretion of several pro-inflammatory cytokines, such as IL1-β, TNF-α, and IL-6 [713].
In addition, extracellular HMGB1 can exert chemotactic effects and support the upregulation of vascular adhesion molecules, thereby impairing the epithelium barrier function [721]. Hence, extracellular HMGB1 has a powerful ability to coordinate different immune responses in the tumor microenvironment.

Another aspect is that the redox state of HMGB1 influences its functions [722, 723]. In fact, the nuclear and cytosolic environments are characterized by a negative redox potential that maintains HMGB1 in reduced form (fr-HMGB1). During an inflammatory process, the extracellular space, enriched in reactive oxygen species, drives a disulfide bond between Cys23 and Cys45 of A Box (ds-HMGB1) [724]. ds-HMGB1 activates TLR2 and TLR4, inducing the release of proinflammatory chemokines and cytokines, activating innate and adaptive immune responses. On the contrary, fr-HMGB1 binds to the receptor for advanced glycation end products (RAGE), modulating autophagy [725, 726].

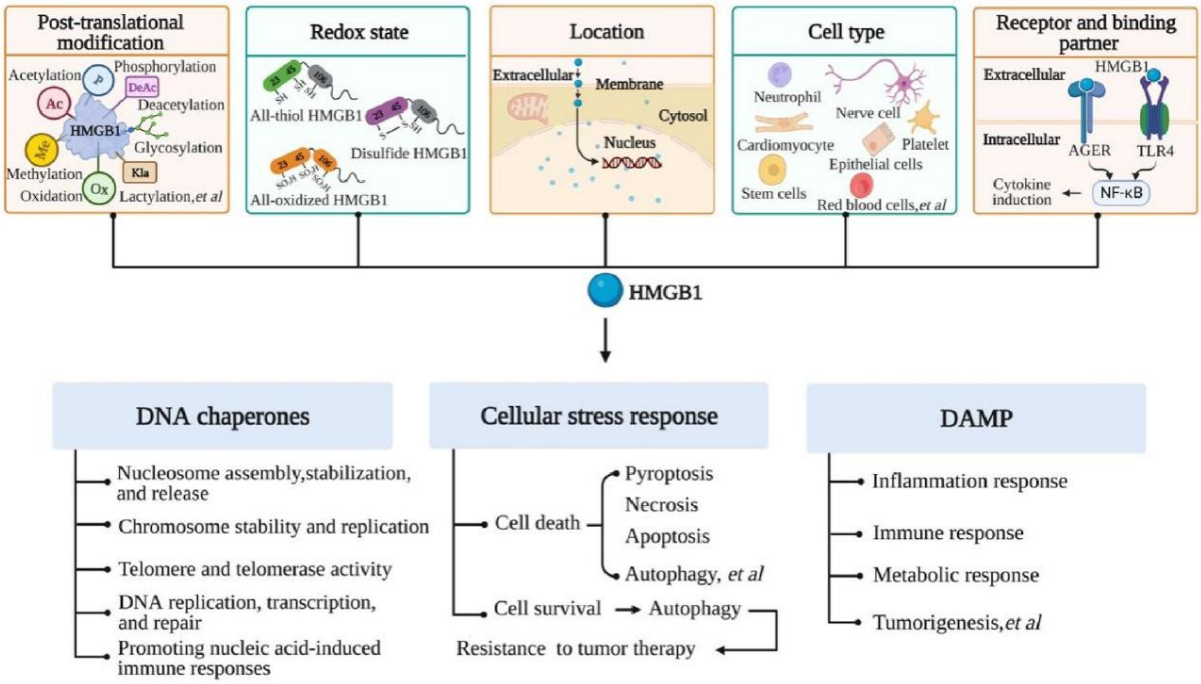



**Figure 24**: HMGB1 functions in the cells. serves as a DNA chaperone, a promoter of cellular stress responses, and a damage-associated molecular pattern (DAMP). These functions are influenced by various factors, including post-translational modifications, redox status, subcellular localization, cell types, as well as receptors and binding partners [727].

### 4.2 HMGB1 Inflammation and Diseases

HMGB1 dysregulation has been implicated in several pathologies, particularly autoimmune diseases and inflammation. Specifically, HMGB1 is found associated with hepatic infectious diseases [728], severe pulmonary inflammatory diseases, including COVID-19 [729]. Additionally, HMGB1 is reported to be an important factor in the development of Rheumatoid Arthritis (RA) [730, 731], and its suppression can inhibit the disease progression [732-737]. Also, HMGB1 is closely associated with Systemic lupus erythematosus (SLE), an autoimmune disease characterized by autoantibody production and systemic inflammation involving multiple organ systems [737]. Moreover, in type 1 diabetes mellitus (T1DM), HMGB1 can be passively released from damaged pancreatic cells and actively secreted by islet-infiltrating immune cells [738].

HMGB1 may also be involved in the pathogenesis of Autoimmune Thyroid Diseases (AITD), such as Hashimoto's thyroiditis (HT) and Graves' disease (GD), which are organ-specific autoimmune diseases characterized by lymphocytic infiltration of the thyroid gland [739-742].

### 4.3 HMGB1 in Cancer

HMGB1 is overexpressed in most tumors, including leukemia, hepatocellular carcinoma, and gastric and colorectal adenocarcinomas [743]. The aberrant expression of HMGB1 has been demonstrated and is closely associated with proliferation and metastasis in breast cancer cells [744, 745]. In HCC, HMGB1 mediates tumor growth by interacting with intracellular TLR9 under hypoxic conditions and attracts macrophages to the tumor site, leading to enhanced metastasis [746, 747].

HMGB1 plays a protective role in tumor suppression, tumor chemoradiotherapy, and immunotherapy. The loss of HMGB1 results in genome instability, autophagy deficiency, and increased apoptosis, leading to tumorigenesis [748].

In this scenario, HMGB1 appears to play paradoxical roles during the development and therapy of cancer. In fact, excessive HMGB1 production caused by chronic inflammatory response seems to be associated with tumorigenesis. In established cancers, HMGB1 produced by tumor cells may aggravate inflammation-related immunosuppression.

For instance, LPS induced the release of pro-inflammatory cytokines such as IL-1β, IL-6, and TNF-α in a HMGB1-dependent manner, improving colon cancer progression [749]. Furthermore, extracellular HMGB1 enhances chemotherapy efficacy by promoting tumor cells from apoptosis to senescence [750, 751], increasing the release of cytokines such as IL-6 and IL-8, which stimulate tumor cells' proliferation, angiogenesis, EMT, invasion, and metastasis. Moreover, nuclear and cytoplasmic HMGB1 promote autophagy and inhibit tumor cells' apoptosis, inducing chemotherapy resistance [750]. Additionally, it was discovered that the intracellular HMGB1 can function as a tumor suppressor by binding tumor suppressor proteins.

Moreover, HMGB1 can mediate immunogenic cell death during chemoradiotherapy, enhance anti-tumor immunity, recruit inflammatory cells, and mediate interactions with Natural Killer (NK) cells, dendritic cells (DCs), and macrophages. Activated NK cells provide an additional source of HMGB1, which is released into the immunological synapse between NK cells and immature DCs, promoting the maturation of DCs [752]. Moreover, several transcription factors, such as p53, retinoblastoma (RB) proteins, and NF-κB family members, can enhance their oncogenic activities through direct interaction with HMGB1 [753, 754]. Lastly, HMGB1 has emerged as a biomarker with prognostic and/or predictive value in various types of tumors.

Predictive biomarkers provide insights into treatment efficacy, aiding in treatment decision-making, while prognostic biomarkers offer information about patient outcomes [727].

HMGB1 is being used as a biomarker in several types of cancer, such as breast cancer, cutaneous T-cell lymphoma, ductal carcinoma, endometrial, esophageal, gastric, lung, prostate, ovarian, and kidney cancers. Serum and tissue HMGB1 levels are correlated with tumor immunoreactivity, size, differentiation, stage, grade, lymph node involvement, distant metastasis, and overall survival [755]. Therefore, modulating HMGB1 may provide a potential combination strategy for cancer chemoradiotherapy and immunotherapy, and targeting chromosomal architectural HMGB1 may provide a new perspective for cancer therapy [756].

### 4.4 Therapeutic Strategies Targeting HMGB1 in Cancers

Numerous therapeutic strategies have been proposed to target HMGB1 in cancer, aiming to directly or indirectly inhibit its expression, translocation, release, and activity. These strategies encompass a wide range of approaches, considering antibodies, peptides, proteins, RNAi, chemicals (including natural products and clinical drugs), the inhibition of HMGB1 receptors and signaling pathways, and cell-based therapies [727]. Hereinafter, a list of the available therapeutics described in the literature to target HMGB1 in a cancer environment:

The antibody Cetuximab [757] induces HMGB1 release, specifically described for head and neck cancer. PKHB1 is a Thrombospondin-1 peptide mimic designed to induce HMGB1's release, mainly for breast cancer [758]. Meanwhile, RAGE antagonist peptide (RAP) competes with HMGB1 for a site on the extracellular domain of RAGE, useful for breast cancer and non-small cell lung cancer [759].

Many other peptides were developed to target HMGB1 for several different cancer types [760-763]. Additionally, several RNA-based drugs are also in development to target HMGB1 expression [764-772]. miRNAs like miR22-3p, miR-107, and miR-410 are useful for inhibiting HMGB1 expression in several cancers, such as gastric, lung, and bladder cancer [773-792]. Moreover, several small molecules were developed to target HMGB1, such as apicidin, docetaxel, lisavambulin, docosahexaenoyl, genistin, imiquimod, sivelestat, abemaciclib, aspirin, carbenoxolone, chloroquine, enzalutamide, irinotecan, mifepristone, mitoxantrone, nafamostat, oxaliplatin, pemetrexed, rafoxanide, and many others [793-830] could be used to inhibit or induce HMGB1 release targeting melanoma, breast, lung, pancreatic cancers, and many more (**Figure 25**).

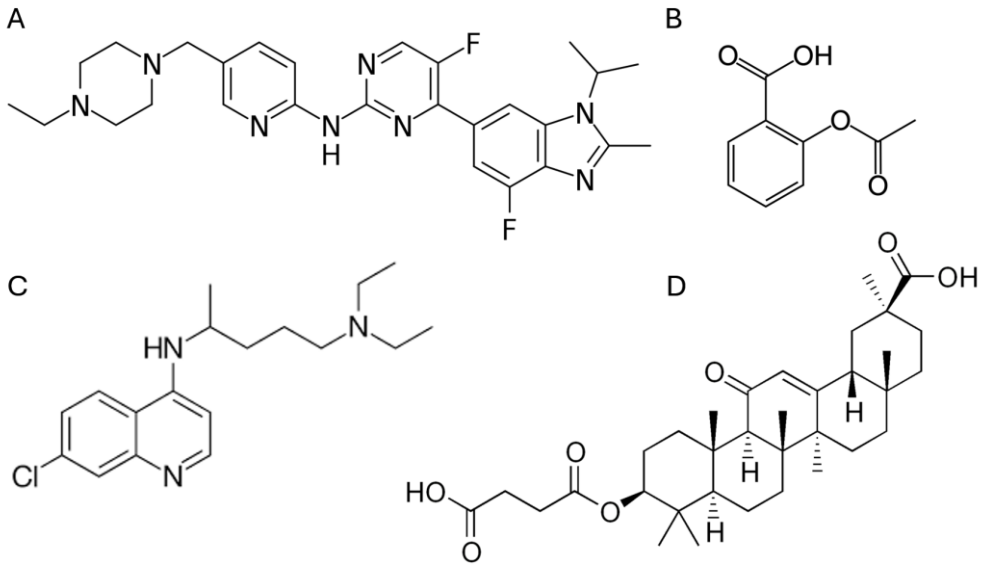



**Figure 25**: Examples of small molecule antagonists of HMGB1: A) Structure of Abemaciclib, B) Structure of Aspirin, C) Structure of Chloroquine, and D) Structure of Carbenoxolone

Furthermore, a list of natural chemical compounds, such as alantolactone, alternol, berberine, capsaicin, curcumin, glycyrrhizin, ginsenoside Rg3, oleandrin, P2Et, piceatannol, resveratrol, shikonin, triptolide, quercetin, diflunisal, and vitexin [831-848] could be useful to target HMGB1 in several types of cancer (breast, gastric cancer, etc.) (**Figure 26**).

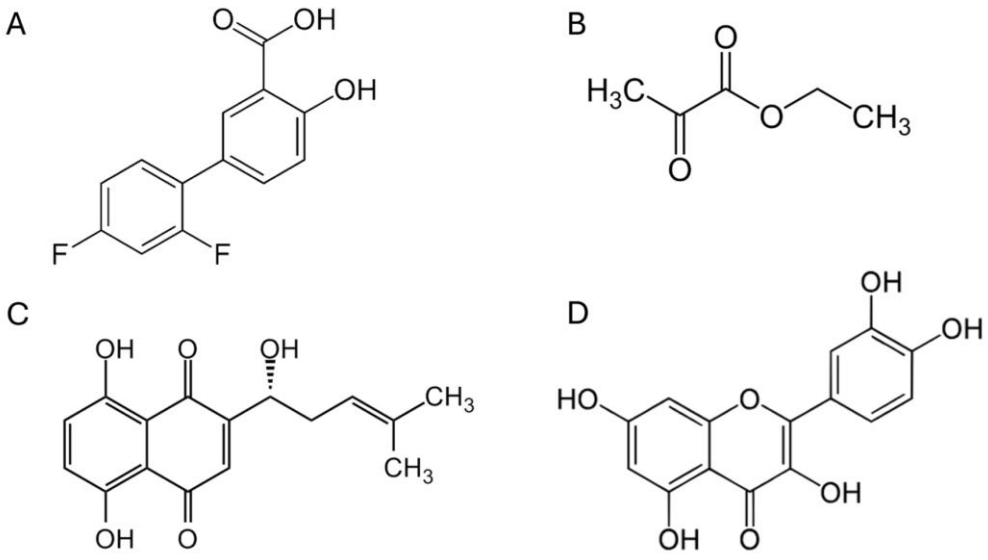



**Figure 26**: Examples of natural antagonists of HMGB1: A) Structure of Quercetin, B) Structure of Ethyl pyruvate, C) Structure of Shikonin, and D) Structure of Diflunisal

### 4.5 Design of XNAs as Potential Inhibitors of HMGB1

In this project, we aim to design new XNAs endowed with inhibitory activity on HMGB1. In this work, a set of *in silico* approaches, i. e., molecular docking, MD simulations, and MM/GBSA calculations, was applied to evaluate and rank new RNA analogs. The final goal was to inhibit the interaction between HMGB1 and DNA in cancer, to reduce its influence on cancer growth and inflammation.

The study began with structure identification; utilizing the work of Sanchez Giraldo et al. [849], we selected an NMR-derived structure (PDB: 4QR9; **Figure 27**) featuring two A Box domains interacting with a double-stranded DNA molecule (dsDNA). This HMGB1 structure presents a high homology to the human HMGB1 form and was obtained from *Rattus Norvegicus*.

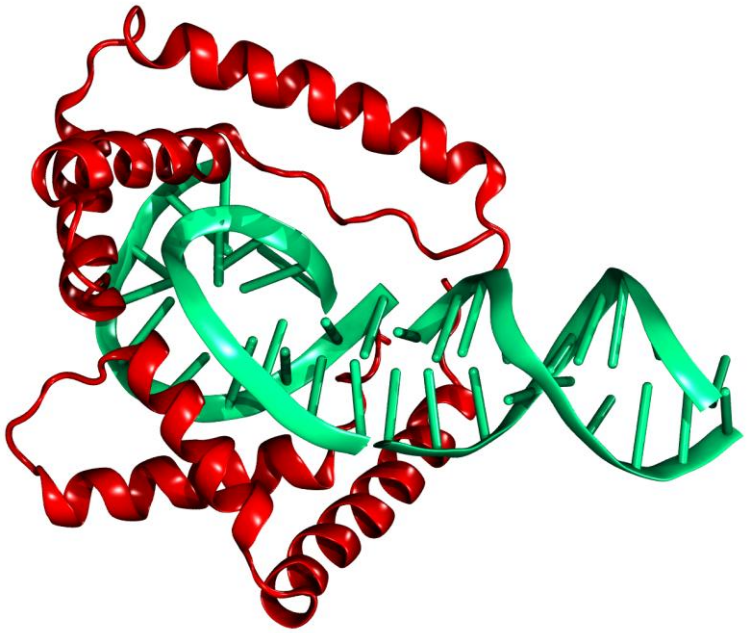


**Figure 27**: Structure of HMGB1-dsDNA complex (accession code: 4QR9). The model consists of two A boxes of HMGB1 (red), in complex with dsDNA filaments (green)

For that reason, the protein was converted to the human form thanks to the tools available in Maestro®. Furthermore, the part of the dsDNA that was not interacting with the protein was removed, while the 10-nucleotide palindrome sequence was maintained, obtaining the complex shown in **Figure 28**. The dsDNA sequence consisted of the following nucleotides: 5’ ATATCGATAT 3’

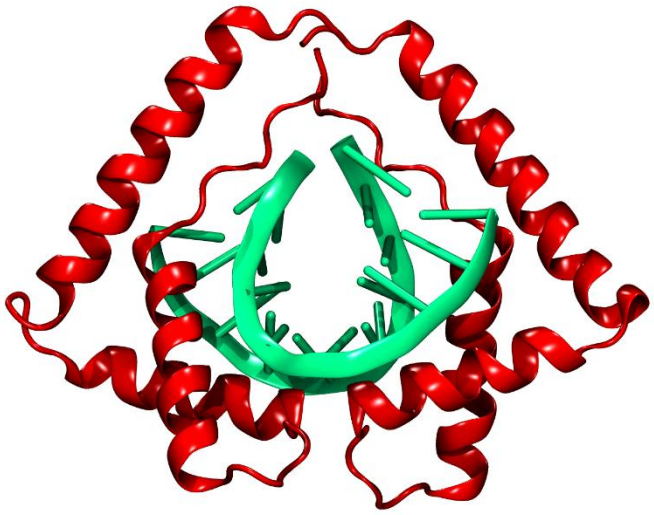


**Figure 28**: Starting complex consisting of two A Box (red) with the short dsDNA sequence (green).

After forming the starting complex, we investigated the shortest possible sequence capable of interacting with a single A Box. This study aimed to uncover which nucleobases are involved in the complex stabilization between the dsDNA and the HMGB1 DNA-binding domain. Consequently, the initial complex was built using the computational procedure previously described for LC3B, obtaining the system depicted in **Figure 29**.

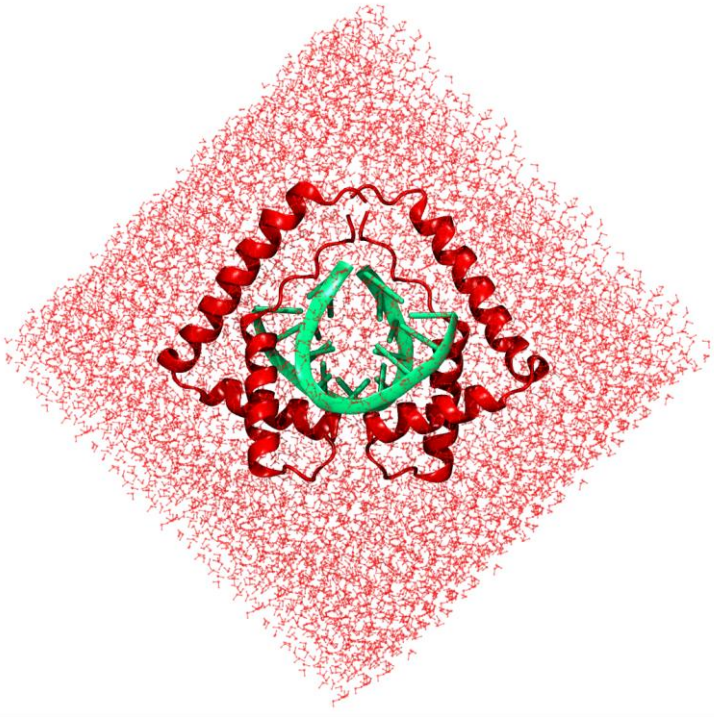


**Figure 29**: HMGB1 (red ribbons) in complex with dsDNA (green) inside the 10Å waterbox created by *System builder* (red).

Furthermore, two 250-ns MD simulation replicas were performed using the Desmond algorithm in Maestro to evaluate how the DNA sequence interacted with HMGB1. Finally, the *Prime* module [850] of Maestro was employed to calculate the ligand ΔG values. Here, the MM/GBSA single-trajectory approach was applied, neglecting the entropy contributions to the binding free energy. Specifically, 200 representative frames were selected for estimating the binding free energy. To shed light on the interaction between HMGB1 A Box and the DNA sequence, the two A boxes were separated while keeping the dsDNA bound to the protein, forming Complex A and Complex B (**Figure 30**). This approach enabled the evaluation of binding stability and interaction profiles at the DNA-binding domain for each complex, revealing the dynamic conformational properties.

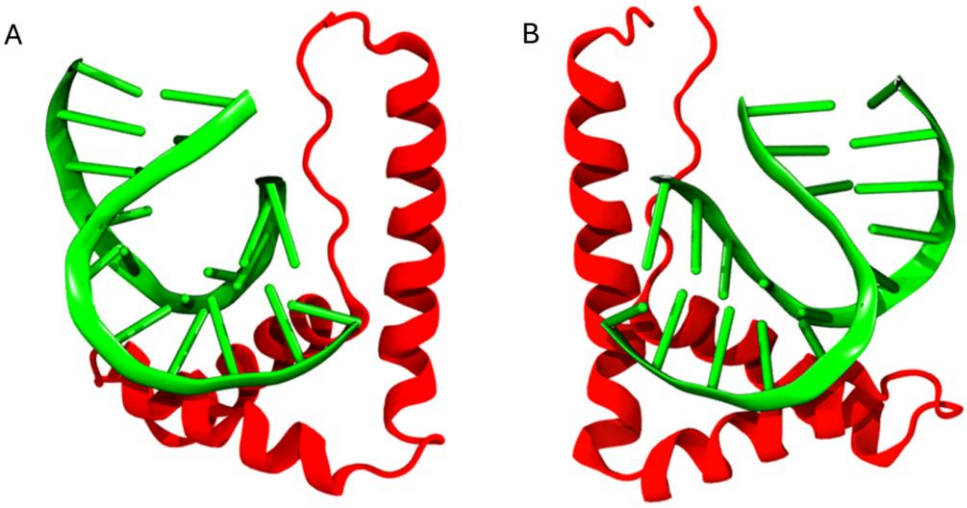



**Figure 30**: Isolated HMGB1-A Box-DNA complexes.
Complex A and B, with the A Box domain (red) and its bound dsDNA (green).

By this process, Complex B showed a lower average binding free energy than Complex A, with −55.0 ± 3.8 kcal/mol and −42.0 ± 4.2 kcal/mol, respectively, as described in **Table 9.**

**Table 9**: Binding free energy (ΔG) of isolated A Box-DNA complexes

| Complex | Backbone | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|
| A | DNA | −42.0 ± 4.2 |
| B | DNA | −55.0 ± 3.8 |

Considering this result, complex B was selected for further studies to identify a more compact active DNA sequence. From the Complex B MD simulations trajectory, the frame in which the dsDNA displayed the lowest binding free energy, detected by the MM/GBSA calculation, was selected and used to evaluate which strand of the DNA duplex (5‘or 3‘) had a better interaction profile in complex with HMGB1-A box. In this regard, two new A box/ssDNA complexes were created, each containing the HMGB1-A Box bound to a single 10-nucleotide strand (5' or 3').

**Table 10**: Binding free energy of A Box bound to single DNA strands

| Complex | Backbone | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|
| B + 5’ filament | DNA | −86.6 ± 5.2 |
| B + 3’ filament | DNA | −48.5 ± 4.3 |

As shown in **Table 10**, complex B, bearing the 5‘filament, acquired a ΔG (−86.6 ± 5.2 kcal/mol) value lower than the 3‘filament (−48.5 ± 4.3 kcal/mol). Consequently, it was further investigated to determine the minimal sequence endowed with the highest affinity on the A box. As shown in **Table 11**, at the end of this process, the 4-nucleotide sequence ATAG showed the best compromise between length and binding free energy (−75.3 ± 7.2 kcal/mol) (**Figure 31**).

**Table 11:** Effect of sequence truncation on binding free energy.

| Sequence | Backbone | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|
| TATAGCTATA | DNA | −86.6 ± 5.2 |
| TATAGCT | DNA | −84.7 ± 4.9 |
| TATAG | DNA | −82.3 ± 5.8 |
| ATAG | DNA | −75.3 ± 7.2 |
| TAG | DNA | Unbinding |

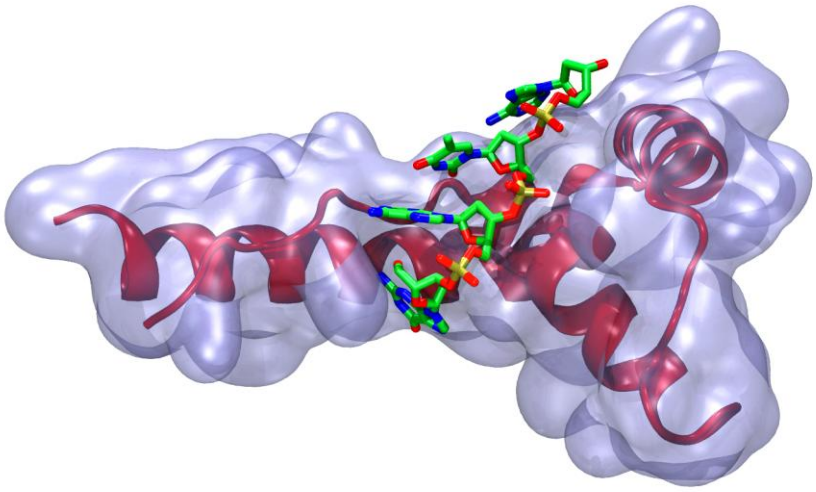


**Figure 31**: ATAG sequence (green licorice) in complex with A Box of HMGB1 (red ribbon and light purple for the surface)

### 4.5.1 Design of XNA Analogs of ATAG ssDNA

Here, the ATAG sequence was used as a template for a backbone and sequence optimization study, in which RNA and XNA backbones, such as PNAs, PMOs, Serinol nucleic acids (SNAs), and, lastly, Morpholino β amino acids, reported in literature by Contini et al [851] and herein defined as Mβaa (**Figure 32**), replaced the natural ssDNA.

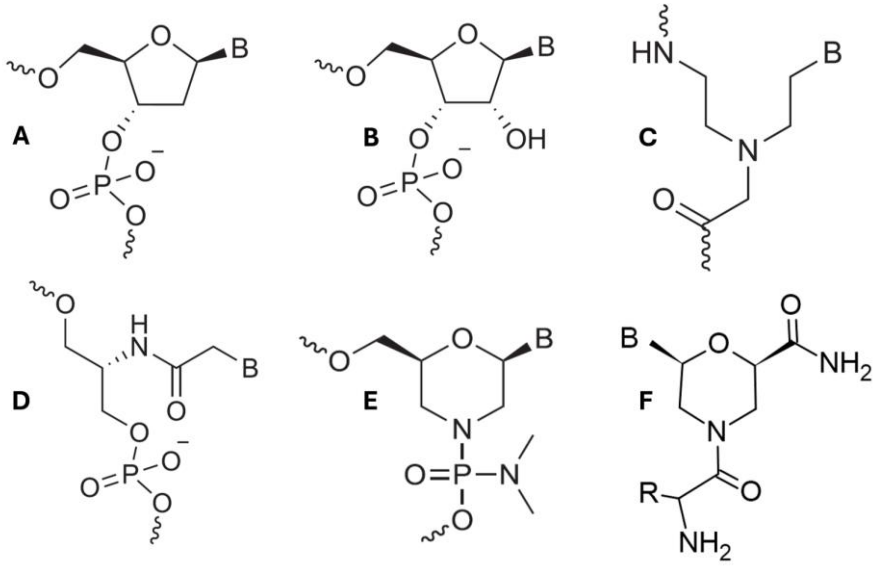



**Figure 32**. XNA backbones used in this study. A) DNA B) RNA C) PNA D) SNA E) PMO F) Mβaa

Using the appropriate Maestro tools and applying MD simulations and the MM/GBSA protocol, our calculations revealed that the Mβaa backbone performed slightly better than other backbones (average $\Delta G = -86.5 \pm 8.6$ kcal/mol).

**Table 12:** Comparative binding free energy of natural and XNAs backbones.

| Sequence | Backbone | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|
| ATAG | Mβaa | −86.5 ± 8.6 |
| ATAG | SNA | −84.1 ± 5.3 |
| ATAG | PNA | −76.4 ± 6.9 |
| ATAG | DNA | −75.3 ± 7.2 |
| ATAG | RNA | −68.7 ± 10.4 |
| ATAG | PMO | −64.1 ± 5.0 |

In fact, as can be observed in **Table 12**, the lowest ligand binding free energy was obtained with the Mβaa backbone, making it a new starting point for the sequence optimization process. To this aim, 32 randomly generated Mβaa sequences, containing the five natural nucleobases (A, C, G, T, and U), were designed and simulated in complex with the A box. At the end of these calculations, the sequence GAGG displayed the lowest binding free energy value (−127.5 ± 5.8 kcal/mol) (**Table 13**).

**Table 13**: Comparative binding free energy of Mβaa sequences

| Sequence | Avg ΔG ± SEM (kcal/mol) | Sequence | Avg ΔG ± SEM (kcal/mol) |
|---|---|---|---|
| GTAG | −110.5 ± 4.7 | GAAG | −110.5 ± 5.9 |
| TTAG | −76.3 ± 3.8 | GCAG | −122.1 ± 5.6 |
| CTAG | −96.6 ± 4.4 | GUAG | −113.6 ± 5.2 |
| UTAG | −99.0 ± 5.0 | GGAG | −120.0 ± 5.4 |
| AAAG | −81.1 ± 4.8 | **GAGG** | **−127.5 ± 5.8** |
| AGAG | −100.1 ± 5.1 | GACG | −101.0 ± 5.2 |
| ACAG | −84.7 ± 4.9 | GATG | Unbinding |
| AUAG | −96.2 ± 5.0 | GAUG | −103.9 ± 5.2 |
| ATCG | −74.8 ± 4.5 | GGGG | −110.9 ± 5.3 |
| ATGG | −85.6 ± 4.9 | GGCG | −117.5 ± 5.5 |
| ATTG | −78.8 ± 4.6 | GGTG | −119.0 ± 5.6 |
| ATUG | Unbinding | GGUG | −108.0 ± 5.1 |
| ATAA | −83.7 ± 4.8 | GAGA | Unbinding |
| ATAC | Unbinding | GAGT | Unbinding |
| ATAT | Unbinding | GAGC | −112.4 ± 5.1 |
| ATAU | Unbinding | GAGU | Unbinding |

Before proceeding with the synthesis, the Mβaa's *N*- and *C-termini* were substituted by introducing the acetyl group at the *N*-terminus and converting the C-terminus primary amide into a carboxylic acid (**Table 14**).

**Table 14**: Binding free energy differences determined by the *N*- and *C-terminus* modifications with Acetyl and Carboxylate groups on Mβaa backbone GAGG sequence.

| **Sequence** | **N-term** | **C-term** | **Avg ΔG ± SEM (kcal/mol)** |
|---|---|---|---|
| GAGG (ref) | $-NH_2$ | $-CONH_2$ | −127.5 ± 5.8 |
| GAGG | -NHAc | $-CONH_2$ | −76.3 ± 5.3 |
| GAGG | $-NH_2$ | -COOH | −129.8 ± 4.6 |
| GAGG (opt) | -NHAc | -COOH | −134.8 ± 5.4 |

The attained results (**Table 14**) suggested that both *N*- and *C-termini* can marginally affect GAGG's average binding free energy. In **Figure 33**, the pose of the GAGG *N*- and *C-termini* modification (GAGG-opt), which illustrates how the nucleobases interact with the surface of HMGB1, establishes a specific rearrangement of the DNA binding domain of the A box. In fact, the enriched purine scaffold sequence is perfectly adapted to the surface of HMGB1. Specifically, the 7 kcal/mol ΔG improvement was driven by a new interaction between the *C-terminus* and the guanidinium group of Arg5 residue of HMGB1, forming a new hydrogen bond-assisted salt bridge.

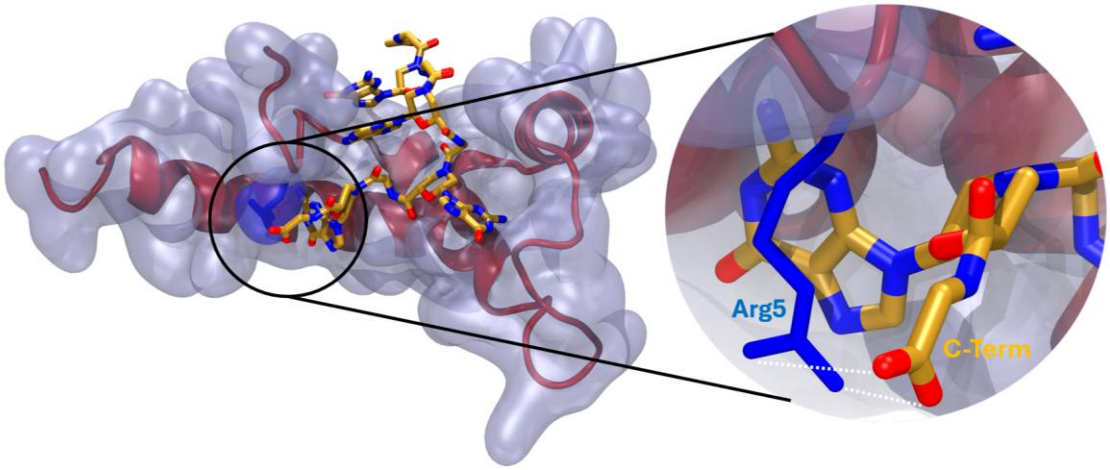



**Figure 33**: Structure of GAGG-opt ligand (orange licorice) bound to HMGB1 A box, shown (red ribbon) with light-purple molecular surface, with the molecular surface of Arg5 highlighted in blue. On the right, focus on the position of the carboxylic group and Arg5.

### 4.5.2 Mβaa Backbone Modification

In the last step of this study, we evaluated the effect of further backbone optimization of the Mβaa, which contained glycine residues (**Figure 34)**. Nevertheless, the synthetic procedure allows the use of multiple amino acids, obtaining Mβaas with heavy chains. The presence of amino acid side chains can influence the conformational mobility of the resulting Mβaa, offering also the possibility of adding functional groups that can create new interactions with the biological target.

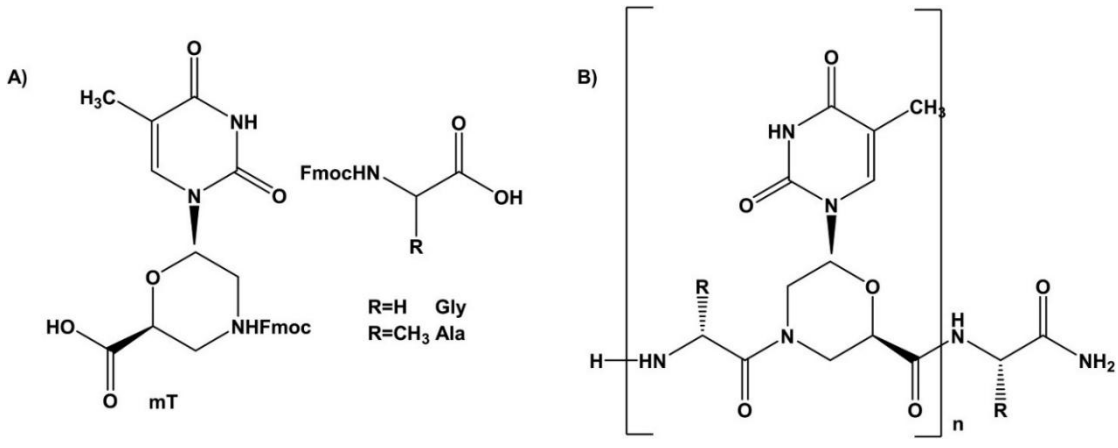



**Figure 34**: A) Building block of the Mβaa, constituted by a modified PMO backbone, in which glycine and alanine residues can be used to connect individual monomers. B) The Mβaa monomer unit [851]

The GAGG-opt Mβaa sequence contains three glycine residues, but as assessed by visual inspection of the MD simulation trajectories, only the glycine at position 2 was located on an HMGB1-A box surface with sufficient space to accommodate bulky groups (**Figure 35**). Consequently, the glycine at position 2 was substituted with eleven amino acids, representative of the main types of amino acids (considering polarity and steric hindrance of the side chain), and simulated in complex with HMGB1-A box, as depicted in **Table 15**.

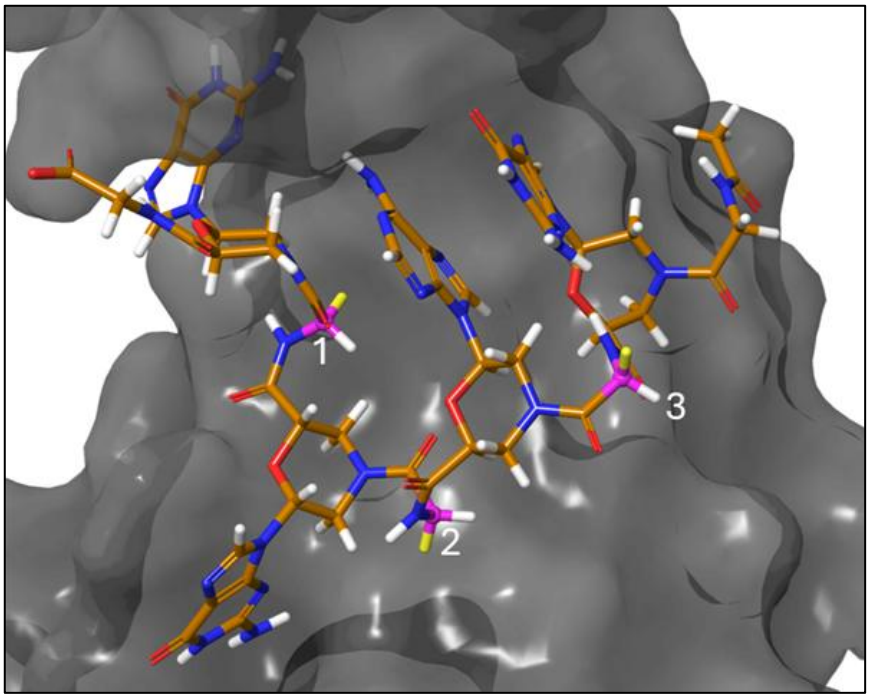



**Figure 35**: Structure of Mβaa GAGG opt (orange licorice) bound to HMGB1-A box shown as red ribbon with light-gray molecular surface.

**Table 15**: Effect of glycine 2 mutation on the predicted binding affinity of Mβaa GAGG sequence.

| Side chain | Avg ΔG ± SEM (kcal/mol) |
|---|---|
| Tyrosine | −149.3 ± 6.3 |
| Phenylalanine | −143.6 ± 7.1 |
| Methionine | −139.5 ± 5.6 |
| Tryptophan | −138.7 ± 5.6 |
| Alanine | −136.1 ± 6.5 |
| Lysine | −135.1 ± 7.5 |
| Aspartic acid | −134.1 ± 9.2 |
| Leucine | −131.8 ± 6.7 |
| Arginine | −125.8 ± 5.9 |
| Serine | −125.0 ± 6.8 |
| Histidine | −124.1 ± 5.9 |

Overall, the best-performing glycine mutations were associated with bulky and hydrophobic amino acids, specifically tyrosine, phenylalanine, and tryptophan (**Table 15**). Among them, the tyrosine analog displayed the lowest predicted average ΔG value (−149.3 ± 6.3) (**Figure 36)**.

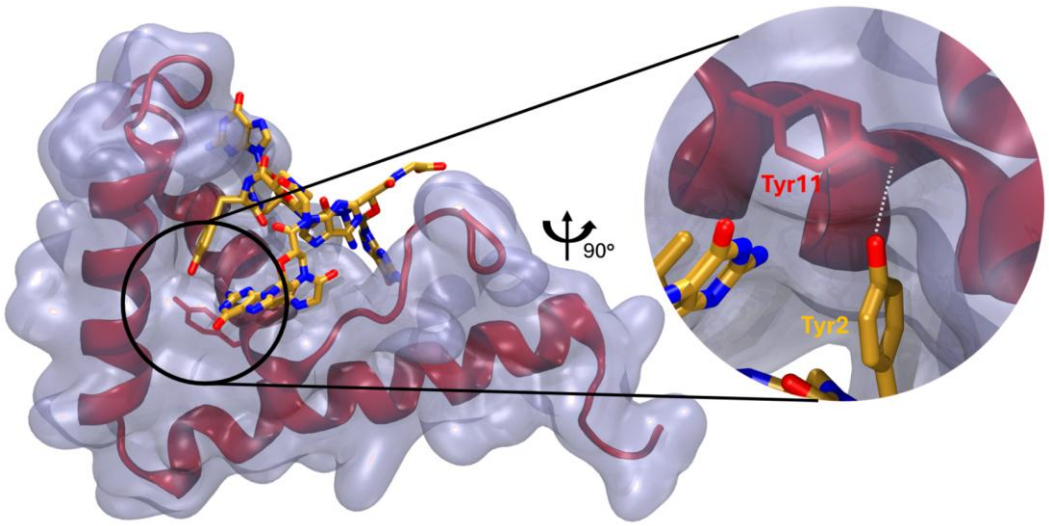



**Figure 36**: The Mβaa GAGG-opt Tyr2 (orange) is shown bound to the DNA-binding domain of HMGB1 (red ribbon, light purple transparent surface). On the right, focus on the interaction of Tyr2 and the new pocket on the HMGB1-A box, and with Tyr11 (red).

The structural basis for the 15 kcal/mol improvement in the estimated binding affinity of Mβaa GAGG-opt involved the Tyr2 side chain occupying a newly created pocket near the A box DNA-binding domain. This configuration assists the creation of a hydrogen bond with the A Box-Tyr11 OH group and is further reinforced by a π–π stacking interaction.

The effectiveness of the applied structural optimization was further confirmed by comparing the RMSD profiles of the two lead Mβaa candidates with the DNA ATAG sequence (Mβaa GAGG opt, and Mβaa GAGG opt Tyr2). The improvement observed in ligand stability, illustrated in **Figure 37**, demonstrates the robustness of the optimization process.

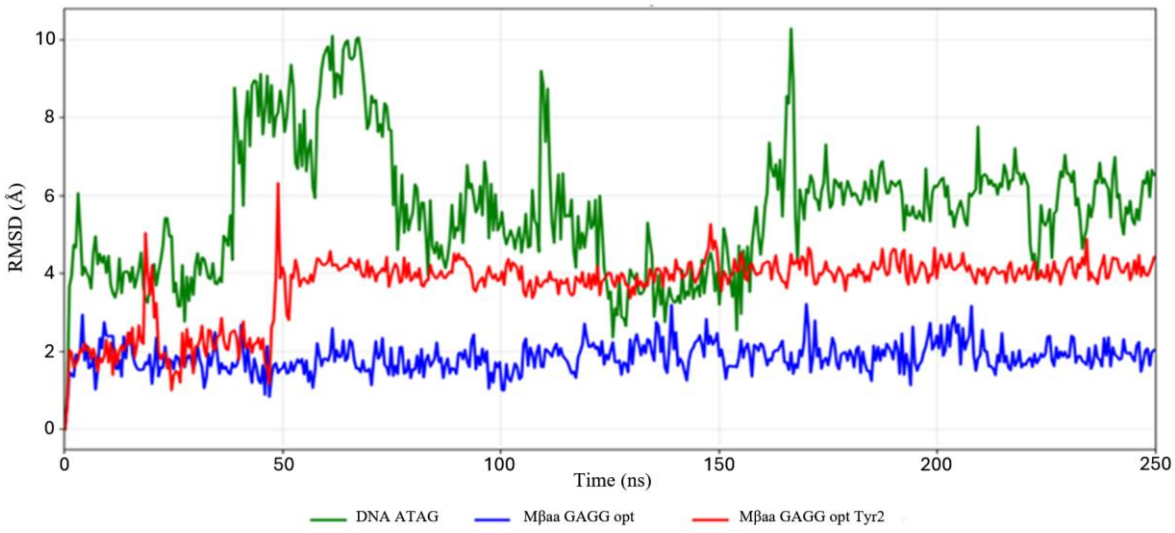



**Figure 37**: Root-mean-square deviation (RMSD) over MD simulation time of ligand-protein complexes.

In conclusion, several new interactions improved the stability and binding affinity of the tyrosine analog of the GAGG-opt Mβaa, for which a binding free energy value of −149.3 ± 6.3 kcal/mol was predicted. It is interesting to note that the sequence optimization protocol on ATAG (−75.3 ± 7.2 kcal/mol) led to a two-fold enhancement of the estimated binding affinity of the Mβaa considered in this study.

## 4.6 Biophysical Experiments

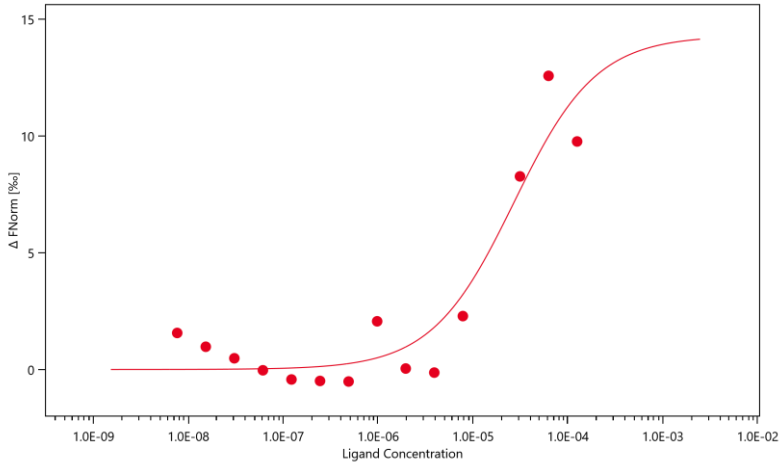



**Figure 38: Microscale Thermophoresis experiment:** The figure illustrates a dose–response binding curve obtained from the MST experiment evaluating the interaction between HMGB1 A box and Mβaa GAGG prot with Tyr in position 2. The x-axis reports the ligand concentration on a logarithmic scale (from ~$10^{-9}$ to $10^{-2}$ M), while the y-axis displays the normalized fluorescence (FNorm) in ‰ units. The red circles represent the individual measurements, FNorm values at each titration step, with vertical error bars showing variability between replicates.

The most promising Mβaa was synthesized in Prof. S. Pellegrino's Lab (University of Milan). Then, the compound was tested using MST experiments (**Figure 38**) to measure the binding affinity on the HMGB1-A box recombinant protein.

The obtained result showed a binding curve displaying a $K_d$ value of 26.9 ± 1.36 µM, confirming the theoretical data. Further biological assays are ongoing to investigate their *in vitro* biological activities in collaboration with Dr. Emilie Venereau (San Raffaele Hospital, Milano). The aim is to evaluate the effects of these Mβaa on muscle cell vitality. Moreover, the anti-inflammatory profile of these ligands will be investigated on suitable *in vitro* assays developed to evaluate the effects of the HMGB1 inhibition.

## 4.7 Conclusion

In this study, computational approaches have been applied to design novel XNAs specifically targeting the HMGB1 DNA-binding domain. The application of this *in silico* workflow facilitated the optimization of the initial sequence into a new XNA, a process supported by RMSD plots and ΔG values. Ultimately, this sequence optimization analysis highlights how specific nucleobases deeply impact ligand dynamics and their interaction profiles with HMGB1, providing a predictive framework for identifying new lead candidate molecules.

## 5. Design of Aptamers Interacting with HIV GP120

The human immunodeficiency virus type 1 (HIV-1) was first isolated in 1983 [852, 853] and associated with the acquired immunodeficiency syndrome (AIDS) in 1984 [854, 855]. Experimental evidence suggested that the first zoonotic transmission to humans occurred between 1920 and 1940 in Central Africa, when HIV originated from numerous cross-species transmission events from simian immunodeficiency viruses (SIVcpz; HIV-1) and West African sooty mangabey monkeys (SIVsmm; HIV-2) [856]. HIV is a member of the Lentivirus genus [857], a member of the Retroviridae family (retrovirus), that contains a genome that consists of two identical copies of single-stranded RNA (ssRNA) molecules [858]. HIV is divided into 2 main groups: HIV type 1 (HIV-1) and HIV type 2 (HIV-2). The HIV-1 variants are classified into four major groups: group M (main), group O (outlier), group N (non-M/non-O), and Group P. Among them, group M is responsible for most infections in the worldwide HIV-1 epidemic and is the main agent of AIDS [859, 860].

HIV-1 remains a global health crisis, having been responsible for 35.7 to 51.1 million deaths, with over 40.8 million people currently living with HIV-1 worldwide in 2024, and with approximately 1.3 million new infections annually, as reported by WHO [860]. The virus transmission occurs principally through sexual intercourse [861], by contact with infected blood, and by reusing or sharing contaminated syringes and needles [862]. Furthermore, vertical transmission is when HIV is transmitted from a pregnant mother to the fetus or newborn, either during pregnancy, during delivery, or by breastfeeding [863].

Multiple issues are related to HIV, such as its high genetic variability/mutation rate, with multiple quasispecies (genetically related variants), which allows the virus to overcome host immunity and the effects of drugs and prophylactic interventions [864]. Specifically, HIV's variability is determined by three factors:

1) The "error-prone" reverse transcriptase, which introduces, on average, one substitution per genome per replication round (RNA replication error rate of about 1 in $10^4$) [865, 866].

2) Rapid viral replication, which generates a high number of virions per day in the infected individual [867].

3) The ability of multiple diverse HIV strains to recombine genetic material in the host cell of the same infected individual, creating circulating recombinant forms (CRFs) [868, 869].

## 5.1 HIV Morphology

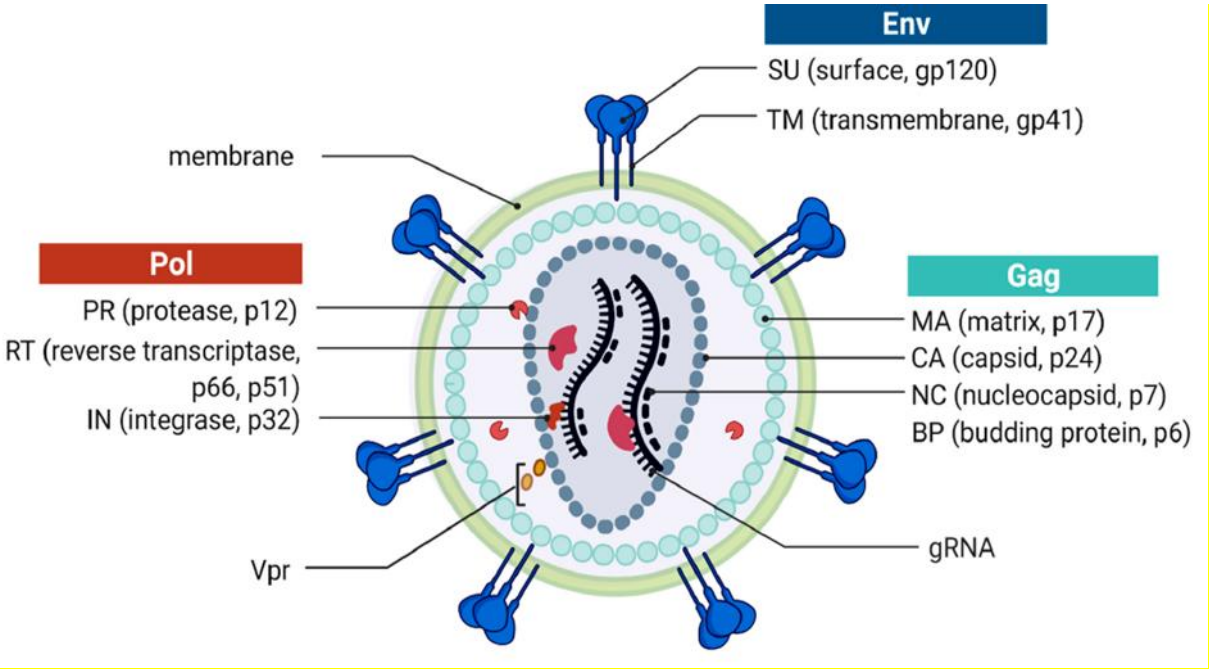



**Figure 39**: Representation of a mature HIV-1 virion, illustrating its lipid membrane, envelope glycoproteins, GAG and POL proteins, and the viral RNA genome (gRNA).

HIV virions have an average diameter of 120 nm, and their RNA genome consists of 9,750 nucleotides [869-871], containing all the information required to synthesize the 15 proteins necessary for the replication and assembly of new virions in infected host cells [872, 873]. The mature virions consist of a lipid bilayer envelope that contains several host cell proteins plus ~7–35 envelope trimers (ENV) composed of Glycoprotein 120 (GP120-SU) and Glycoprotein 41 (GP41-TM) [874-879]. The envelope covers the internal viral core, which is formed by the matrix protein (P17-MA), a highly myristoylated protein that mediates membrane association. Beneath the matrix lies the viral capsid, which is composed of 1000-1500 cone-shaped hexameric capsid proteins (P24-CA) [880]. The capsid protects two copies of positive-sense single-stranded genomic RNAs (gRNAs) bound to nucleocapsid proteins (P7-NC), preventing their degradation by nucleases. Furthermore, the mature virion contains essential viral enzymes, such as reverse transcriptase (P66-/P51-RT), integrase (P32-IN), protease (P10-PR), and accessory proteins VPR, VIF, and NEF [881].

## 5.2 Replication Cycle and Role of GP120

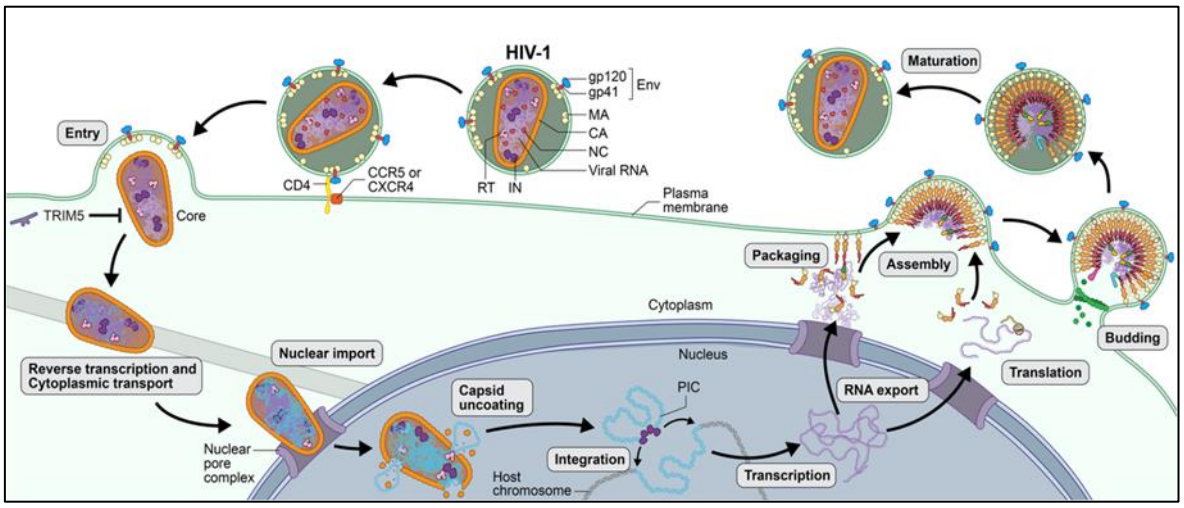



**Figure 40**: Schematic overview of the HIV-1 replication cycle [882].

The HIV replication can be summarised in six steps, as shown in **Figure 40**.

1. **Binding and entry**: The entry of HIV consists of three main events: virus binding to the cell, activation, and fusion with the host cell. The viral envelope glycoprotein ENV, composed of the heterodimer proteins GP120-SU and GP41-TM as aforementioned, targets cells by recognizing CD4, which is present on the cell surface of ≃60% of circulating T-lymphocytes, on T-cell precursors (present at the bone marrow and thymus), on monocytes/macrophages, eosinophils, dendritic cells, and microglial cells of the central nervous system (CNS). Right after the formation of the complex GP120-SU-CD4, the virus envelope complex rearranges its structure, exposing a specific domain that binds chemokine receptors on the host cell membrane. The most common chemokine coreceptors used by HIV are C-X-C motif chemokine receptor type 4 (CXCR4) and chemokine receptor type 5 (CCR5) [883]. The formation of the complex GP120-CD4, plus one chemokine receptor, allows for a stable association of the virus with the host cell. After that, the membranes fuse [884-886], followed by the delivery of the virion content into the host cell cytoplasm [887].
2. **Uncoating**: After the entry of the capsid into the cytoplasm, it moves toward the nucleus by microtubules and the intervention of microtubule motor proteins. Furthermore, the capsid binds the nuclear pore complex (NPC), contacting two nucleoporins. Simultaneously, after the nucleoplasm entry, the capsid disassembles, releasing CA [888, 889] and viral RNA.

3. **Reverse transcription**: The viral RNA needs to be converted into proviral DNA through the action of the RT in the cytoplasm. The reverse transcription begins by synthesizing ssDNA from the first strand transfer. The primer binding site is used to initiate the process. The ssDNA is hybridized to the 3'-end of the viral genome (RNA/DNA hybrid double helix), while negative-strand DNA synthesis persists, and the ribonuclease H site breaks down the RNA strand. The second strand transfer will lead to the transcription of the positive-strand DNA, terminating the dsDNA synthesis [890, 891].
4. **Integration and Transcription:** The newly synthesized dsDNA is integrated into the host cell genome by IN, which cleaves the nucleotides of each 3' end of the dsDNA, exposing a 3′-hydroxyl group and a 5′-overhang, forming two sticky ends. Subsequently, it transports the modified dsDNA into the host nucleus and assists and promotes its integration into the host genome. The cells infected after provirus integration are the important long-living reservoirs of HIV [892]. Notably, the integrated DNA, defined as provirus, can remain silent in the host genome to escape the immune system, establishing a long-lasting infection that characterizes the retroviruses [893].
5. **Protein Synthesis and Assembly**: The transcription of the provirus into mRNA determines the synthesis of regulatory HIV-1 proteins (TAT, REV, NEF, VIF, VPR, and VPU), and each of them has a specific role in the assembly and virion maturation. Protein expression is regulated by several factors, at the epigenetic, transcriptional, and post-transcriptional levels [894-896]. The env mRNA is translated into ENV precursor GP160 in the rough endoplasmic reticulum (ER), where GP160 assembles into trimers and travels to the Golgi apparatus, where it gets glycosylated and cleaved by furin-like proteases [897]. The structural proteins will form the nucleus of the new HIV particles, and the assembly of new virions begins when two viral RNA strands join with replication enzymes, while core proteins (MA) form the virus capsid.
6. **Budding**: The viral particles will obtain a new envelope by budding through the host cell membrane, which may incorporate diverse host cell proteins, phospholipids, and cholesterol. The viral particle matures and reorders its main components, such as structural proteins, gRNAs, and enzymes, to form the infectious virion. Overall, the structural changes are obligatory for viral infectivity, and after all these phases, the new HIV-1 virion is now ready for a new replication cycle [898].

### 5.3 Therapies

Several treatments have been developed during the last few decades. The following is a list of the most common treatments available on the market to treat HIV infection.

(I) Nucleoside reverse transcriptase inhibitors (NRTIs): inhibit provirus synthesis by causing chain termination during DNA strand elongation or by directly inhibiting RT activity [899].

(II) Non-nucleoside reverse transcriptase inhibitors (NNRTIs): they bind to a specific site of RT, causing a structural change that stops its functions [900].

(III) Integrase inhibitors (INSTIs): prevent the insertion and integration of the synthesized provirus into the host cell genome [901].

(IV) Protease inhibitors (PIs): block the cleavage of precursor proteins to interrupt the assembly and maturation phases [902].

(V) Fusion inhibitors (FIs): block the entry of HIV by gp41-TM fusion peptide binding [903].

(VI) CCR5 receptor antagonists: block the CCR5 co-receptor of CD4+ T cells, avoiding the initialization of the gp41-TM-mediated membrane fusion [904].

(VII) Attachment inhibitors (AIs): prevent retroviral entry by blocking HIV gp120-SU [905, 906].

(VIII) Post-attachment inhibitors (pAIs): prevent HIV entry by blocking HIV gp120-CD4 receptor

(IX) Capsid Inhibitory: target p24, they interfere with the uncoating phase, transport of the ssRNA into the nucleus, and the assembly and maturation of the new virions.

(X) Pharmacokinetic Enhancers/Boosters: inhibit the enzymes CYP3A4 responsible for the metabolism of other ART drugs, leading to their increased concentration for a longer period.

(XI) Fixed-Dose Combinations (FDCs): A Combination of two or more medications from diverse classes into one single pill, improving and simplifying therapy adherence

In **Table 16**, the full list of the FDA-approved drugs available for HIV.

**Table 16**: The FDA-approved drugs available for HIV and their classes [907]

| Drug Class | Generic Name (Abbreviations) | Brand Name | FDA Approval Date |
|---|---|---|---|
| **NRTIs** | abacavir (ABC) | Ziagen | Dec 17. 1998 |
| | emtricitabine (FTC) | Emtriva | July 2. 2003 |
| | lamivudine (3TC) | Epivir | nov 17. 1995 |
| | tenofovir disoproxil fumarate (TDF) | Viread | Oct 26. 2001 |
| | tenofovir alafenamide (TAF) | Vemlidy | nov 10. 2016 |
| | zidovudine (AZT. ZDV) | Retrovir | March 19. 1987 |
| **NNRTIs** | doravirine (DOR) | Pifeltro | Aug 30. 2018 |
| | efavirenz (EFV) | Sustiva (discontinued) | Sept 17. 1998 |
| | etravirine (ETR) | Intelence | Jan 18. 2008 |
| | nevirapine (NVP) | Viramune (discontinued) | June 21. 1996 |
| | nevirapine (NVP) | Viramune XR (discontinued) | March 25. 2011 |
| | rilpivirine (RPV) | Edurant | May 20. 2011 |
| | rilpivirine (RPV) | Edurant PED | March 15. 2024 |
| **INSTIs** | cabotegravir (CAB) | Apretude (injection) | Dec 20. 2021 |
| | cabotegravir (CAB) | Vocabria (tablet) | Jan 22. 2021 |
| | dolutegravir (DTG) | Tivicay | Aug 12. 2013 |
| | dolutegravir (DTG) | Tivicay PD | June 12. 2020 |
| | raltegravir (RAL) | Isentress | Oct 12. 2007 |
| | raltegravir (RAL) | Isentress HD | May 26. 2017 |
| **PIs** | atazanavir (ATV) | Reyataz | June 20. 2003 |
| | darunavir (DRV) | Prezista | June 23. 2006 |
| | fosamprenavir (FOS–APV) | Lexiva (discontinued) | Oct 20. 2003 |
| | ritonavir (RTV) | Norvir | March 1. 1996 |
| | tipranavir (TPV) | Aptivus | June 22. 2005 |
| **Fusion Inhibitors** | enfuvirtide (T–20) | Fuzeon (discontinued) | March 13. 2003 |
| **CCR5 Antagonists** | maraviroc (MVC) | Selzentry | Aug 6. 2007 |
| **Attachment Inhibitors** | fostemsavir (FTR) | Rukobia | July 2. 2020 |
| **Post–Attachment** | ibalizumab–uiyk (IBA) | Trogarzo | March 6. 2018 |
| **Capsid Inhibitors** | lenacapavir (LEN) | Sunlenca (treatment) | Dec 22. 2022 |
| | lenacapavir (LEN) | Yeztugo (prevention) | June 18. 2025 |
| **Pharmacokinetic** | cobicistat (COBI) | Tybost | Sept 24. 2014 |
| **Combinations** | abacavir / lamivudine (ABC/3TC) | Epzicom (discontinued) | Aug 2. 2004 |
| | ABC / DTG / 3TC | Triumeq | Aug 22. 2014 |
| | ABC / DTG / 3TC | Triumeq PD | March 30. 2022 |
| | ABC / 3TC / ZDV | Trizivir (discontinued) | nov 14. 2000 |
| | ATV / COBI | Evotaz | Jan 29. 2015 |
| | BIC / FTC / TAF | Biktarvy | feb 7. 2018 |
| | CAB / RPV | Cabenuva | Jan 22. 2021 |
| | DRV / COBI | Prezcobix | Jan 29. 2015 |
| | DRV / COBI / FTC / TAF | Symtuza | July 17. 2018 |
| | DTG / 3TC | Dovato | April 8. 2019 |
| | DTG / RPV | Juluca | nov 21. 2017 |
| | DOR / 3TC / TDF | Delstrigo | Aug 30. 2018 |
| | EFV / FTC / TDF | Atripla (discontinued) | July 12. 2006 |
| | EFV / 3TC / TDF | Symfi | March 22. 2018 |
| | EFV / 3TC / TDF | Symfi Lo (discontinued) | feb 5. 2018 |
| | EVG / COBI / FTC / TAF | Genvoya | nov 5. 2015 |
| | EVG / COBI / FTC / TDF | Stribild | Aug 27. 2012 |
| | FTC / RPV / TAF | Odefsey | March 1. 2016 |
| | FTC / RPV / TDF | Complera | Aug 10. 2011 |
| | FTC / TAF | Descovy | April 4. 2016 |
| | FTC / TDF | Truvada | Aug 2. 2004 |
| | 3TC / TDF | Cimduo | feb 28. 2018 |
| | 3TC / ZDV | Combivir (discontinued) | Sept 27. 1997 |
| | LPV / RTV | Kaletra | Sept 15. 2000 |

Furthermore, the most advanced therapy available on the market is ART, a combination of three/four antiviral molecules administered as a lifelong treatment program (comprising two NRTIs plus a third drug, which can be a boosted PIs or an NNRTI, or INSTIs**)** [908-910].

ART is not a cure, meaning that it cannot eliminate all the virions and proviruses in HIV-infected patients, but it permits control over HIV infection as a chronic disease by restraining virus replication and viral load to a steady low level, preventing CD4+ T cell decline [911, 912] and progression to AIDS. Furthermore, the almost undetectable plasma viremia reduces the risk of sexual transmission by 96%, preventing new infections [913, 914]. However, ART is a costly long-term treatment that can determine the development of multidrug-resistant viruses. Moreover, ART is associated with multiple adverse effects, such as anorexia, nausea, and vomiting [899]. The advent of drug Resistance to some drugs may cause complications for the selection of appropriate drug regimens [915]. To avoid drug resistance, ART is administered with different drug combinations [916-919].

During the ART regimen, the immune activation is significantly high, and it is associated with a higher risk of:

A) cardiovascular disease
B) renal disease
C) neurological disease
D) malignancy/non-AIDS-defining events [920-922]

Another valuable method to reduce the risk of HIV transmission and to prevent the infection, around 80%, is by post-exposure prophylaxis (PEP) treatment with multiple antiretrovirals (tenofovir, emtricitabine, and raltegravir). PEP can be started within the first 72 hours after direct contact with infected blood or blood-containing fluid (occupational contact) or after non-occupational [923, 924]. The challenges associated with this preventive method are the high costs and limited access [925]. Drammatic improvements have been achieved since the FDA approval of the first antiretroviral in 1987 [926, 927], and the overall health and life quality of the people infected by HIV have remarkably improved [910-913].

This phenomenon is associated with the increased potency, specificity, and half-life of the new generation of therapies available in the market [928, 929].

Novel approaches are compounds that can interact with TAT, REV, and P7-NC of GAG, including conserved mRNA secondary structures (hairpins, stem-loops, and bulges) present in TAR [930-933], RRE [934-936], and Ψ [937-939].

### 5.4 Vaccination and Prospects

The development of vaccines for HIV, followed by a global vaccination campaign, is considered the most effective strategy to eliminate the HIV pandemic. To date, however, the development of a potent vaccine has been unsuccessful due to HIV's complex biology [940-944].

Vaccination must yield immediate activity in containing and controlling viremia post-infection and reducing the chance of establishing viral reservoirs. The ideal HIV vaccine would consist of two components [945, 946].

1) Multiple broadly neutralizing antibodies (bNAbs) designed to target diverse sensitive epitopes of ENV strains/HIV variants to prevent the infection of new host cells [947-951]. These bNAbs can neutralize a wide range of HIV variants; in contrast, most induced neutralizing Igs are variant or strain-specific [952, 953].
2) A component that induces a quick and powerful T cell response to suppress initial viremia, subsequently preventing the formation of viral reservoirs.

However, it remains uncertain how the future HIV vaccine will prevent transmission, grant full immunity, and prevent the progression to AIDS [954]. Overall, the development of an efficient polyvalent vaccine requires the application of several advanced novel technologies, such as mRNA-, HIV-derived virus-like particle (VLP)-based, and viral vectored vaccines.

### 5.5 Design of XNA Aptamers Interacting with HIV GP120

As aforementioned, ENV is a crucial glycoprotein involved in the interaction between HIV and the host CD4 receptor, expressed on our cells. ENV initiates the entry of HIV into our cells, and the possibility of blocking this interaction is crucial. bNAbs and Fostemsavir can interact with the CD4 binding domain, but a cheaper and faster alternative to produce relies on aptamers. Different research groups have worked on the design and development of aptamers targeting and inhibiting HIV main components, such as RT [955-968], IN [969-974], and GP120 [975-986].

In this project, in collaboration with Prof. Amaro and her team, specifically PhD L. Casalino and PhD M. Shehata, I had the opportunity to perform simulations using the full-length and fully glycosylated 3D model of GP120, recently reported in literature by them [987] (**Figure 41**). This model was used to simulate the interaction between novels and known aptamers [986] targeting GP120 and its CD4-binding domain, to understand their mechanisms of action by a full *in silico* approach.

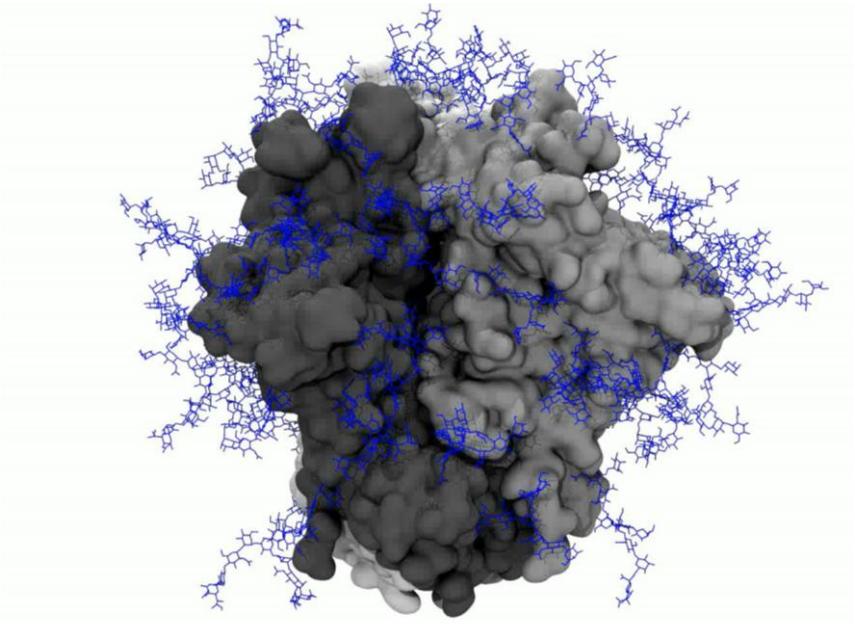


**Figure 41**: Full-length model of GP120. In shades of gray, the protein component (GP120-GP41); in blue, the glycan groups

Additionally, this project aimed to optimize the aptamer's backbone, using the ENV-CD4 complex as a reference.

To achieve these objectives, the following steps were taken (**Figure 42**):

A) Aptamer selection

B) Secondary structure prediction

C) 3D modelling prediction

D) Molecular Docking

E) MD simulations

F) Trajectory analyses and aptamers binding free energy estimation

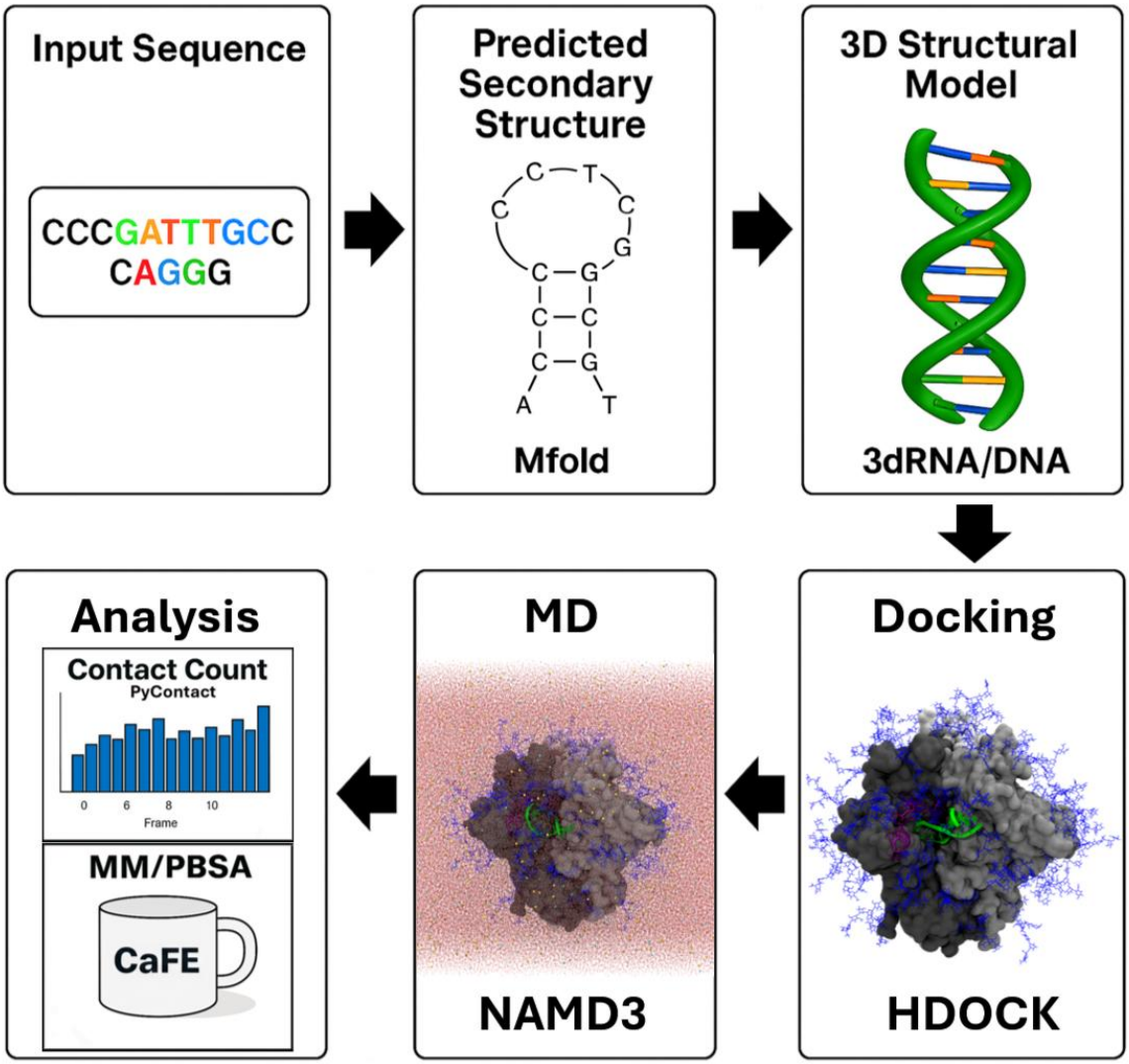



**Figure 42**: Workflow of the project and the list of several software used

### A) Aptamer selection

The aptamers utilized in this study, specifically HD2, HD3, HD4, and HD5 (**Table 17**), were formerly identified by Zarandi et al. [986]. These single-stranded DNA sequences exhibited high affinity for GP120, and biological assays confirmed their capacity to disrupt the interaction between HIV and its primary receptor, CD4 (**Figure 43**).

**Table 17:** Sequences of the aptamers described by Zarandi et al [986]

| Name | Backbone | Length | Sequence |
|---|---|---|---|
| HD2 | DNA | 34 | CGGCCTTAAACGCGCCTGAATAAATTCCTTGCTA |
| HD3 | DNA | 31 | AAGGGTTTTTCCCCGGCCTTAAATTTTCCAA |
| HD4 | DNA | 33 | GATTTGTTTTCCCGATTTTGCCCAGGGTTAATT |
| HD5 | DNA | 49 | CTTTAACGGCCACGCCCGATTTTACGCTTTCCCTACCGAAGGGCCGTAC |

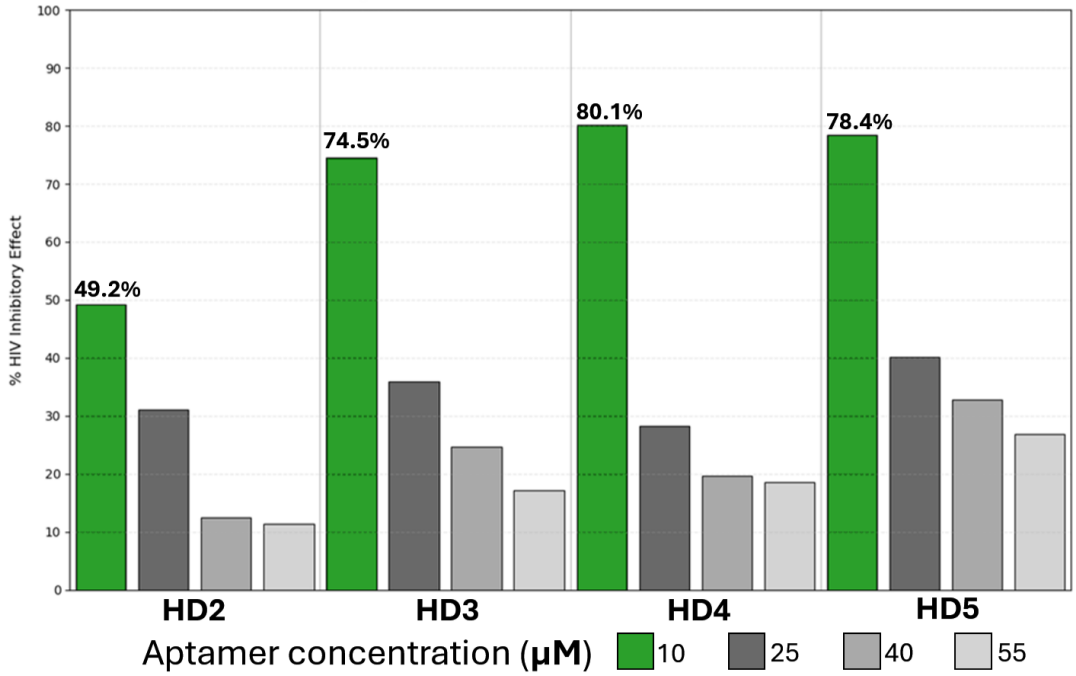



**Figure 43**: The graph illustrates the HIV inhibitory effects of four different aptamers (HD2–HD5) at concentrations between 10 and 55 μM. Remarkably, all tested aptamers displayed their best inhibitory effect at the lowest concentration of 10 μM, with HD4 (80.1%) and HD5 (78.4%) showing the highest inhibitory activity of the group [986].

In addition to these known aptamers, AptaSIM [988] was used to design new ssDNA aptamers against GP120 randomly. Initially, a primary library of 1 million random ssDNA sequences was built, containing a uniform A, C, T, and G nucleotide distribution (25% each). The protocol excluded 5′ and 3′ primer binding regions to emulate a PCR-free amplification model, and twelve iterative selection cycles were performed using the following parameters: 100 seeds per cycle, minimum seed affinity threshold 80%, maximum of 10 sequences selected per cycle, maximum sequence affinity cutoff of 25%, mutation rates: 26% for C–G base pairs; 24% for A–T pairs, 5% overall mutation probability per cycle, and 99% amplification efficiency.

At the end of each *in silico* SELEX cycle, AptaSIM quantifies how many times each sequence is amplified, generating variables such as amplification count for each sequence. In this project, the sequences with the highest amplification count were selected for further analysis. Finally, the aptamers' selection was achieved considering these factors (**Table 18**):

- Apt 0 was chosen to evaluate a long and bulky aptamer.
- Apt 1, 2, and 3 were selected to compare different sequences of the same length (40 nucleotides)
- Apt 4 was chosen to evaluate a short aptamer
- Apt 5 was selected to evaluate the shortest aptamer achievable

**Table 18**: Length and sequences of the aptamers randomly predicted by AptaSIM [988]

| Name | Backbone | Length | Sequence |
|---|---|---|---|
| Apt 0 | DNA | 72 | AAACCAGCTTATTCAATTAATATAGTAGTGAAAGAGCGGGGAGGGCGCAGGT GAGATAGTAAGTGGCAATCT |
| Apt 1 | DNA | 40 | CTAGCATGACCGCAGATCCTTTCGAGGAGAATGGTAGCTC |
| Apt 2 | DNA | 40 | TAGGGGGCAGCTACAAGTCCACGAGCCTGCCAGCCGAATA |
| Apt 3 | DNA | 40 | CAATTCACAAAGCATAAACAATAGGCCCGGTGGTGAAACT |
| Apt 4 | DNA | 22 | GGGTAGGACCCCTACTACCGAC |
| Apt 5 | DNA | 15 | GGCCTCCGGGGTGGA |

### B) Secondary structure prediction

In the second step, the sequences highlighted in **Tables 17** and **18** were used as input sequences for the secondary structure prediction performed by Mfold [989]. This algorithm, developed in the late 1980s [990], can predict the secondary structure of DNA and RNA, obtaining a dot-bracket representation and the minimum free energy structure (MFE) of each aptamer. The MFE values were used to rank all the different structures, and the one with the lowest value was selected for further studies. The secondary structure was predicted using default settings, 37°C, 1 bar, and an ionic condition of 1.0 M NaCl, without the inclusion of divalent ions, as suggested by Mfold developers.

### C) 3D model prediction

The secondary-structure information obtained by MFOLD was used, in combination with the primary sequence, to predict the tertiary structure. For this purpose, 3dRNA/DNA v0.7/0.8 [991] was used to obtain a representative 3D model for each sequence.

3dRNA/DNA is a template-based method that combines DNA and RNA 3D template libraries to predict DNA 3D structures [992-996]. This tool separates the target DNA into secondary structure elements (SSEs) to retrieve a 3D template for each from a DNA SSE 3D template library or assembles the 3D model using a bi-residue algorithm (**Figure 44**).

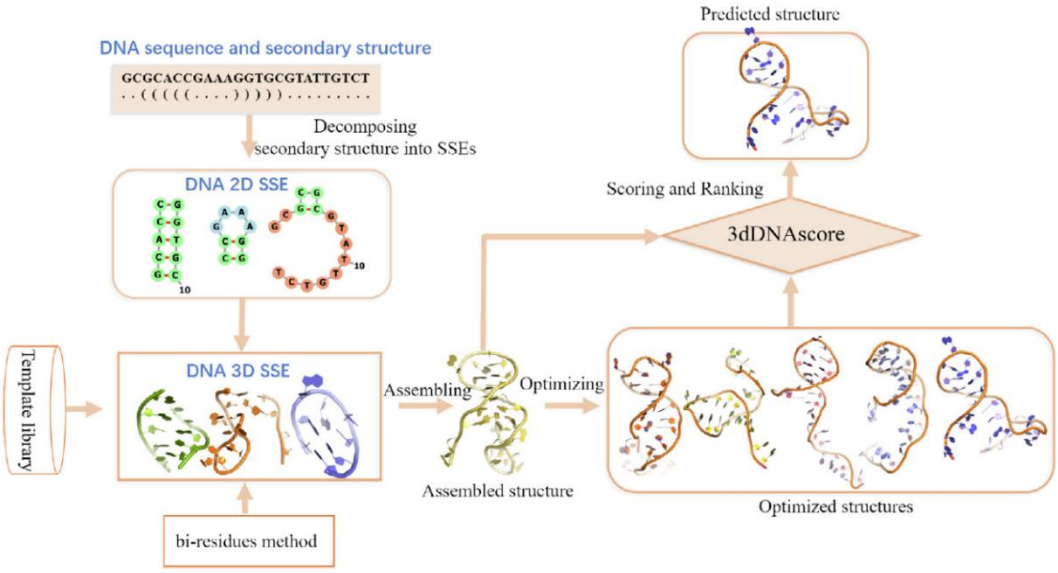



**Figure 44**: 3dRNA/DNA workflow [991]

For each aptamer, the 3D model with the lowest potential energy was selected for further investigations. No additional minimization or structural refinement methods were needed after this step (**Figures 45** and **46**).

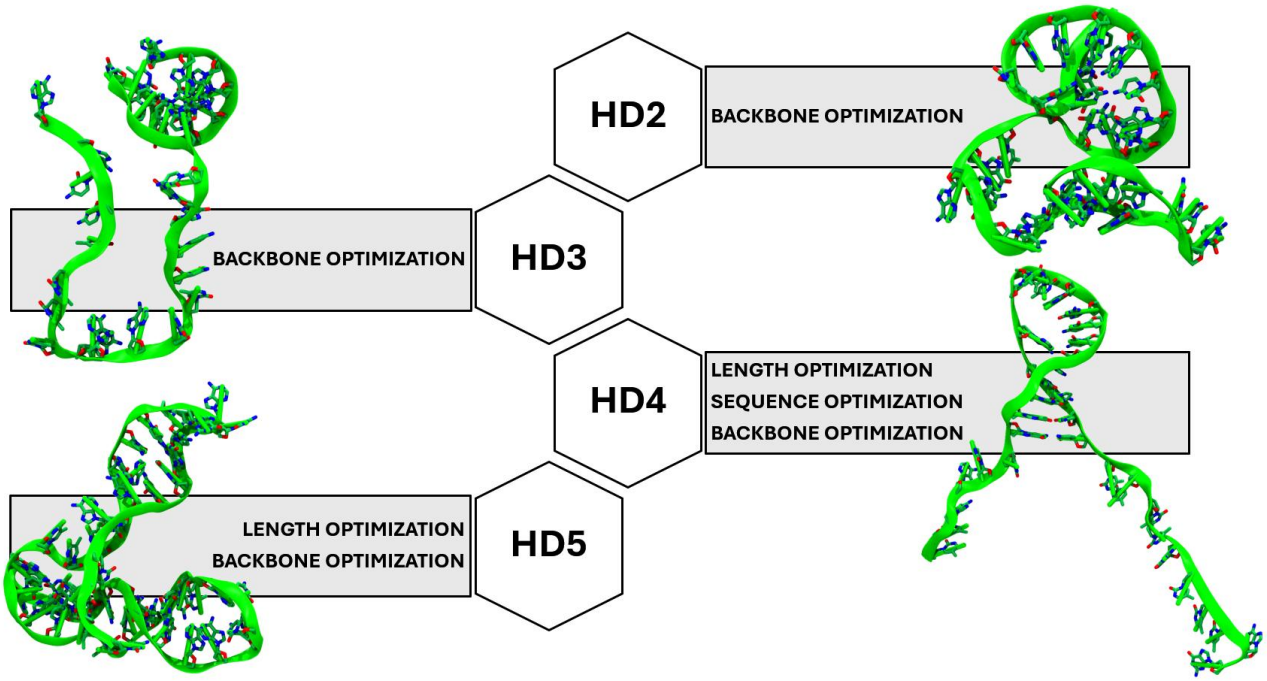



**Figure 45:** Aptamers gathered from Zarandi et al [986] and the aptamer optimization approaches applied for each of them

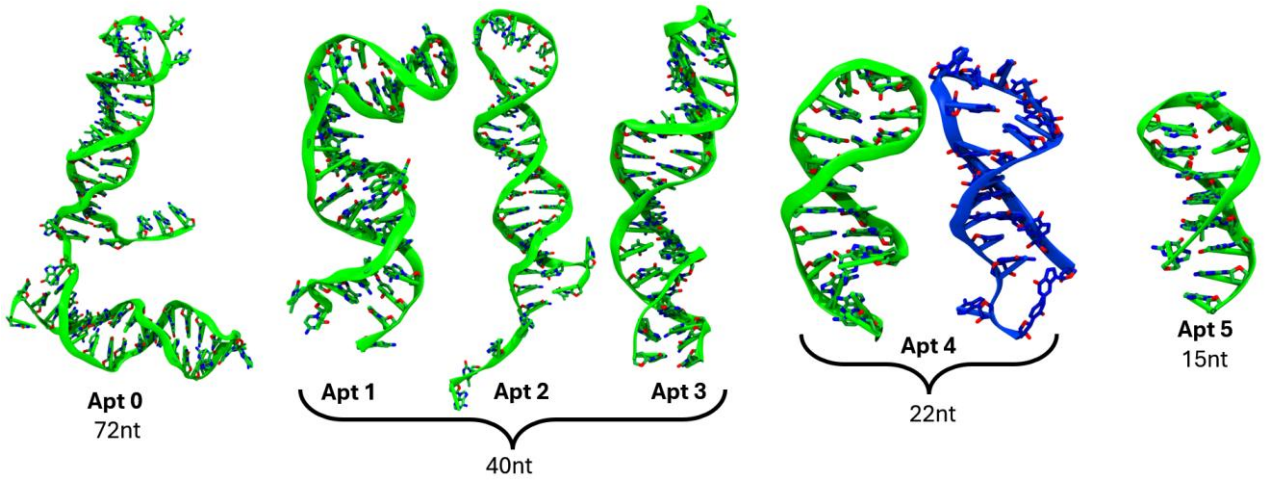



**Figure 46**: 3D model of the sequences predicted by AptaSIM [988]. In green, the DNA backbone, while in blue, the RNA backbone

### D) Aptamer optimization

Three approaches are available to ameliorate the pharmacokinetic properties of the aptamers:

1) Backbone optimization: improving biostability and nuclease-resistance through the introduction of natural chemical modifications or synthetic XNA backbones. Specifically, DNA, RNA, 2'F-RNA, and 2'OMe-RNA backbones were selected and applied to the sequences retrieved from Zarandi et al. [986].

2) Length optimization: aptamer length optimization to decrease the synthesis costs and to identify the key binding regions (applied to HD4 and HD5) (**Figure 47A/B**)

3) Sequence optimization: using site-directed mutagenesis to improve folding, binding free energy by incorporation of natural or unnatural nucleobases to form canonical or non-canonical base pairing (applied to HD4 by double point substitution on Thy28-29 by introducing two adenines, highlighted by the red box) (**Figure 47A**).

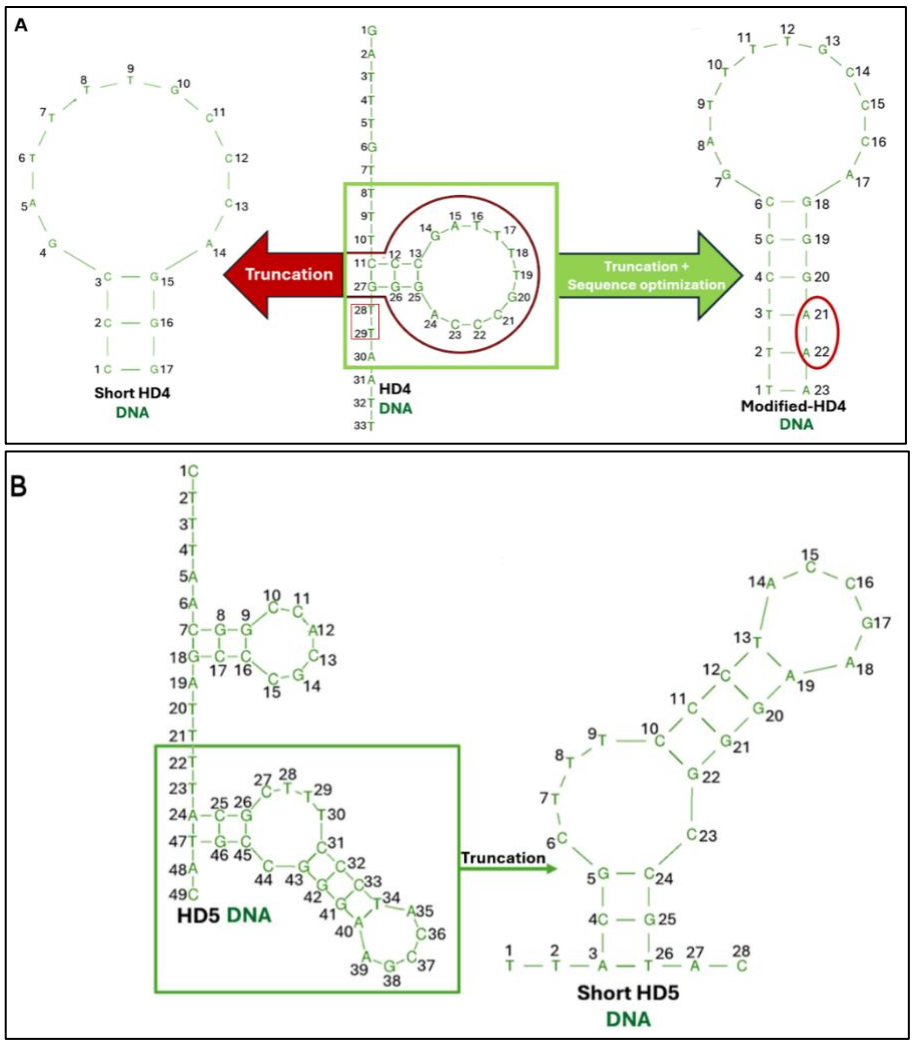



**Figure 47**: A) Aptamer optimization approaches applied on HD4, to form Short HD4 (S-HD4) and Modified HD4 (M-HD4) B) Length optimization applied on HD5 to form Short HD5

**Table 19**: Sequence specifications, lengths, and backbone composition of the HD4 and HD5 DNA aptamer variants.

| Name | Backbone | Length | Sequence |
|---|---|---|---|
| Short HD4 | DNA | 17 | CCCGATTTTGCCCAGGG |
| Mod HD4 | DNA | 23 | TTTCCCGATTTTGCCCAGGGAAA |
| Short HD5 | DNA | 28 | TTACGCTTTCCCTACCGAAGGGCCGTAC |

### E) Molecular Docking

The 3D model of each aptamer was docked into the three CD4-binding domains of the full glycosylated GP120 model. To this aim, HDOCK [423] was used. The docking sites were defined by conserved residues playing a critical role in CD4-GP120 interactions (**Figure 48A**) [886]. Docking poses were then evaluated using the aptamer binding energy scores and the structural complementarity with GP120. The visual inspection of the binding poses allowed for the avoidance of clashes and artefacts in the resulting poses. The pose of each aptamer with the lowest docking score was used to shape the final GP120/aptamer complex needed for the subsequent MD simulations.

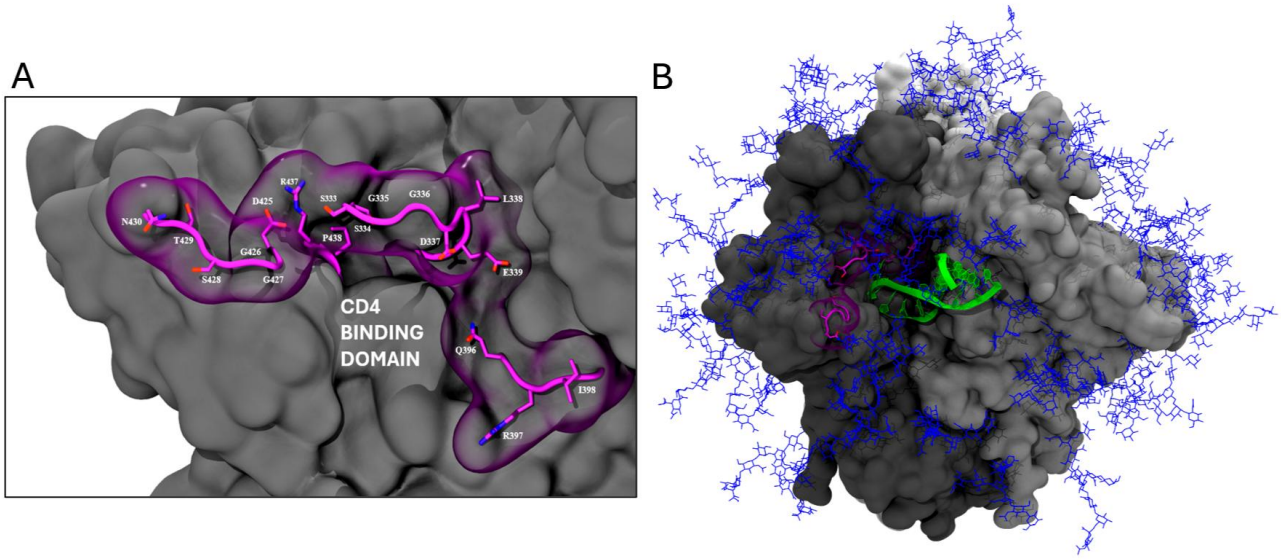



**Figure 48**: A) Surface representation of the GP120 glycoprotein (gray) with the CD4-binding site residues highlighted in purple. B) Representative binding pose of a gp120-aptamer complex predicted by HDOCK. The aptamer is shown in green, while interacting with the GP120 surface (gray), glycans (blue), and the CD4-binding site surface (purple)

### F) MD simulations

To conduct MD simulations, the force field for each XNA had to be available in the most common MD simulation software. However, while the parameters of DNA, RNA, and 2'OMe RNA are already incorporated into the CHARMM36m force field [443], for the 2'F RNA backbone, we developed a custom patch using the CHARMM General Force Field (CGenFF) website [997]. This approach was applied to the modified nucleotides: 2'F-uridine, 2'F-cytidine, 2'F-adenosine, and 2'F-guanosine.

The investigation and application of these newly developed force fields in the *in silico* protocol enabled an in-depth evaluation of the conformational dynamics and binding energetics of all aptamers, providing a foundation for the rational design of high-affinity, RNA-based aptamers.

The GP120-aptamer complexes resulting from docking calculations were then solvated by adding a cubic TIP3P water box, extending 15 Å beyond the solute in all directions, plus neutralizing counterions ($Na^+$ or $Cl^-$) to reach a physiological ionic strength of 0.15 M NaCl. MD simulations were carried out using explicit water molecules and ions to ensure the accuracy of the simulated biological system (**Figure 49**). The GP120-CD4 complex was also created and simulated to acquire reference values. The MD simulations protocol consisted of 6 steps, as described in **Figure 49** and **Table 20**. NAMD3 [998] was used to simulate these large complexes, each comprising over 500.000 atoms.

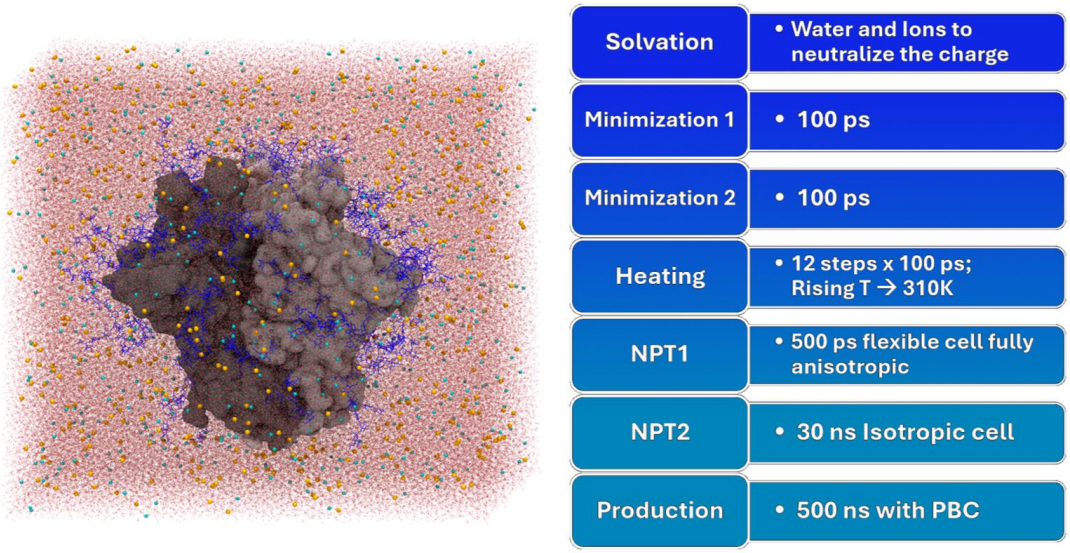



**Figure 49**: (Left) The solvated Env system, showing ENV surrounded by the TIP3P water box and neutralizing ions. (Right) The MD simulation protocol. It starts with energy minimization steps 1 and 2 (100 ps each) to reach a low-energy state. In the heating phase, the temperature gradually increased to reach 310 K. Equilibration was performed in two stages: NPT1 (500 ps) with a fully anisotropic flexible cell and NPT2 (30 ns) with an isotropic cell to stabilize density and pressure. Lastly, a production phase (500 ns) with periodic boundary conditions (PBC) was executed for each replica.

The MD simulations were carried out on the Triton shared computing cluster (TSCC) at the University of California, San Diego, USA. The total duration of all simulated processes exceeds 45 microseconds.

**Table 20**: MD simulation parameters applied to the Env-Aptamer/CD4 complexes

| Category | Parameter | Value/Description |
|---|---|---|
| **Integrator & Time** | Time Step | 2 fs |
| | Total Production Time | 1.5 μs (3 independent runs of 500 ns each) |
| | Bond Constraints | SHAKE algorithm (all bonds involving Hydrogen) |
| **Non—bonded Interactions** | 1–4 Scaling Factor | 1.0 |
| | Switching Function | 10–12 Å |
| | Cutoff Distance | 12 Å |
| | Pairlist Distance | 13.5 Å |
| **Electrostatics** | Long—range Method | Particle Mesh Ewald (PME) |
| | PME Grid Spacing | 2.0 Å |
| | Interpolation Order | 8 |
| | Update Frequency | Full electrostatics every 3 steps; Non—bonded every 1 step |
| **Ensemble & Boundary** | Ensemble | NPT (Constant Number. Pressure. Temperature) |
| | Boundary Conditions | Periodic Boundary Conditions (PBC) in 3D |
| | Atom Wrapping | All atoms wrapped to primary cell |
| **Thermostat** | Temperature | 310 K |
| | Method | Langevin Dynamics |
| | Damping Coefficient | 1 $ps^{-1}$ |
| **Barostat** | Pressure | 1.01325 bar |
| | Method | Langevin Piston |
| | Piston Period / Decay | 200 fs / 50 fs |
| | Piston Temperature | 310 K |
| | Cell Fluctuations | Isotropic |

### G) Post-processing analysis of MD trajectories

Several analyses were performed on the MD trajectories, such as-contact count analysis with PyContact, and binding free energy calculations through MM/PBSA calculations, thanks to the CaFE tool (Calculation of Free Energy) [999]. This VMD plugin can use CHARMM36m force fields, together with the custom patch developed with CGenFF for the analysis of 2'F RNA. To perform these calculations, CaFE executes the:

1) Conformational sampling from MD simulations using parameters retrieved from force fields
2) Post-processing of MD simulations by extracting the complex, receptor, and ligand conformations of interest
3) Calculations of the three energetic components by the MM/PBSA method (SASA term, Gas-phase, and PB energy) [1000].

CaFE ignores the entropic term due to the high computational cost and inaccuracy of current methods for entropy calculations.

Here, the aptamer binding free energy calculation was performed on 100 significant frames extracted from each MD replica, using these criteria: a) the first 1000 frames of the MD simulations were removed since they included the equilibration phase, and b) the last 4000 frames were considered for the analysis, using a stride value. To have a better understanding of the aptamer binding free energy profiles, the aptamer $\Delta G$ values were decomposed into the: i) Glycan contribution, ii) Protein contribution, and iii) CD4 binding domain residues contribution.

Moreover, to evaluate the weight of each nucleotide in the aptamers, the overall binding free energy was divided by the aptamer's length, yielding a normalized value. The software PyContact [1001] was used to identify and characterize the interactions between the aptamers and GP120. This analysis aimed to clarify which nucleotides, backbone chemistries, and key residues can provide the most effective aptamer/CD4 interaction profile with GP120. PyContact was run with a distance cutoff of 5.0 Å for general contacts, a donor–acceptor distance cutoff of 2.5 Å for hydrogen bonds, and an angular cutoff of 120° to ensure geometrically meaningful interactions. The same 100 representative frames used for the binding free energy calculation were also used for the interaction analysis.

### H) Outcomes

The results are divided into two main sections: 1) contact count analysis was carried out to elucidate the key residues and glycan groups involved in aptamer-GP120 interactions, and 2) binding free energy profiles were investigated to understand the relationships among aptamers and their backbones.

**Contacts count.** The contacts count analysis was used to identify the residues involved in the interactions between aptamers and GP120. The data attained suggest that, throughout the MD simulations, the natural binder CD4 exhibited the strongest and most persistent interactions with Gly335, Asp337, and Glu339 (CD4 BD residues), confirming its well-defined binding interface with the viral protein. In contrast, interactions with glycans were weak and transient. The aptamers showed a clear preference for Arg444, Arg397, and Arg153, all residues part of or located near the CD4 BD. Additionally, numerous asparagine residues and their linked glycans (Asn129, Asn197, and Asn234) play a pivotal role in stabilizing the aptamer–GP120 complexes. Notably, a distinct shift from protein-based to glycan-based interactions was observed for HD2, where the strongest contacts transitioned from Arg444 and Arg397 in the DNA form to the glycans bound to Asn197 and Asn234 in the 2′OMe RNA variant.

**Binding free energy.** As mentioned above, the reference CD4–GP120 complex was simulated to establish if the new aptamers can compete with the binding of the natural binders. In our simulations, CD4 exhibited an average ΔG value of −20.7 ± 2.5 kcal/mol (**Figure 50**). Interestingly, the CD4 binding profile yielded a perfect 50-50 ratio between glycan and protein contribution, describing how the natural binder interacts with GP120.

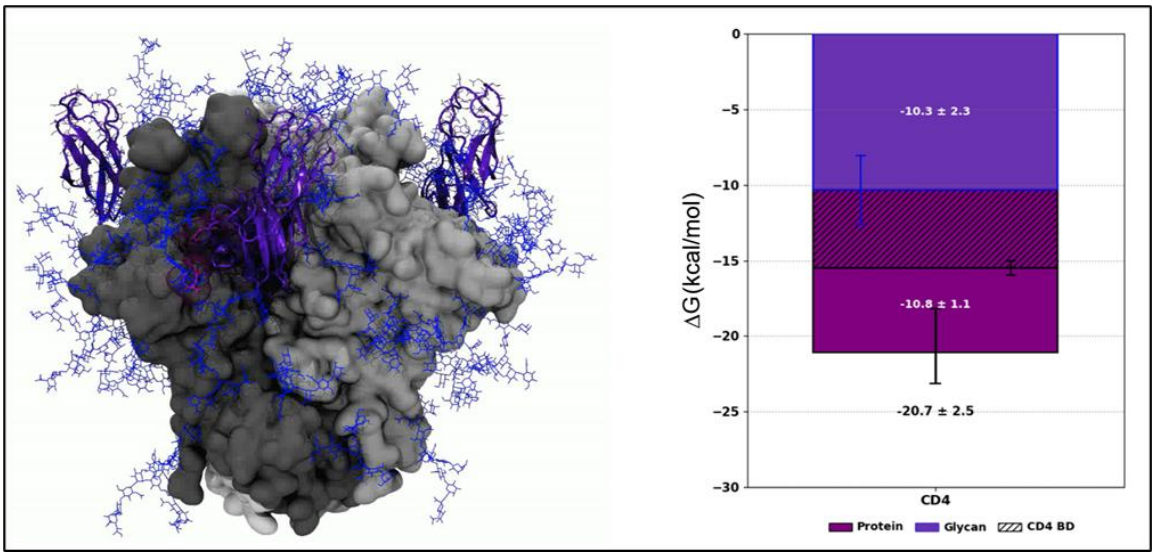



**Figure 50**: Average binding free energy evaluation of CD4 in complex with GP120. In the blue border, the glycan contribution (−10.3 ± 2.3 kcal/mol), in the purple stripes, the CD4 binding domain binding free energy value (comprises 25% of the overall ΔG), in dark purple with black border, the protein section that consists of the CD4 binding domain section + protein section (−10.8 ± 1.1 kcal/mol)

Considering the data collected for all the simulated aptamers (**Table 19, Figure 51**), among the aptamers bearing backbone modifications, the 2'OMe-RNA analog of HD2 displayed a ΔG value (−31.6 ± 2.2 kcal/mol) 27% lower than the DNA natural analog (−24.8 ± 3.1 kcal/mol). This data makes this aptamer the most stable candidate among the ones reported by Zarandi et al. (**Figure 51**). More in depth, the drastic improvement relies on the formation of new interactions displayed by the glycan moieties. On the other hand, the backbone optimization of HD4, from DNA to 2'OMe RNA, reported a drastic decrease in the predicted ΔG value, by approximately 30% (from −31.6 ± 2.9 to −21.9 ± 2.2 kcal/mol) (**Figure 51**). All these results show that the backbone modification does not universally improve the ΔG values but is highly dependent on the aptamer sequence and folding.

Regarding the aptamers randomly predicted by AptaSim (**Table 21**), Apt 0 yielded the most favorable binding free energy (−33.4 ± 3.3 kcal/mol).

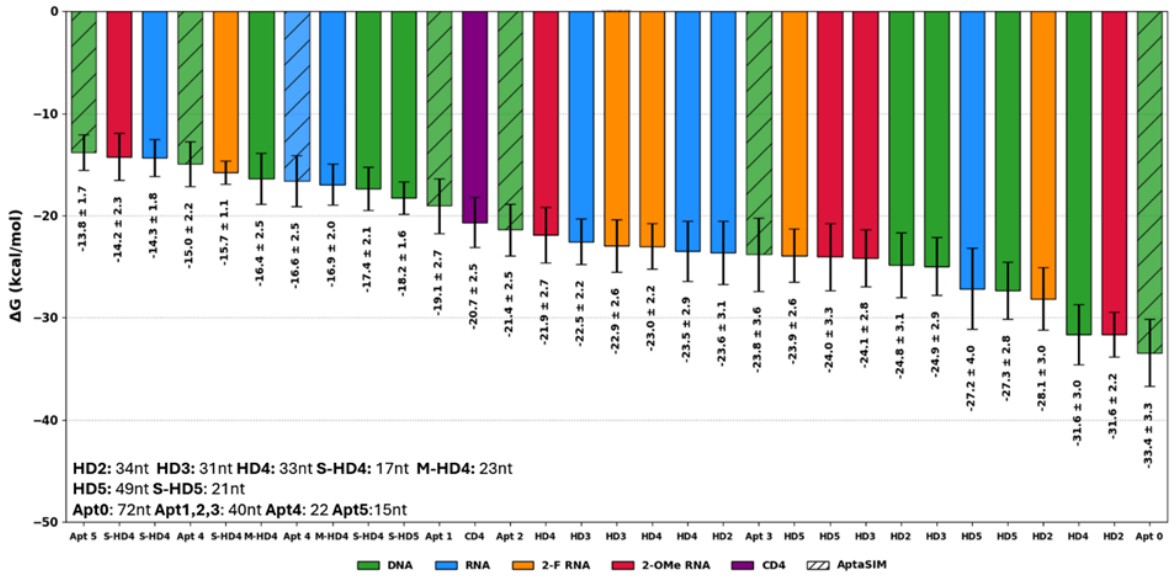



**Figure 51**: Absolute ΔG values (kcal/mol) of all aptamers analyzed in this study. Aptamers are ordered from weakest to strongest binding affinity (less negative to more negative ΔG). ΔG values are described with their related standard deviations (error bars). Aptamers are color-coded by backbone (DNA, RNA, 2'F RNA, 2'OMe RNA in green, blue, orange, and red, respectively), while hatched bars represent the aptamers predicted by AptaSIM.

Not surprisingly, the length optimization approach has yielded a significant reduction in the total binding free energy for all aptamers considered in this study. Specifically, when considering the DNA backbone, Short-HD4 (**Table 21**) exhibited a decrease of 45% relative to its original template HD4 (from −31.6 ± 2.9 to −17.4 ± 2.1 kcal/mol); in parallel, Mod-HD4 (**Table 21**) manifested a comparable loss of 49% compared to HD4 (from −31.6 ± 2.9 to −16.4 ± 2.5 kcal/mol). Meanwhile, Short-HD5 (**Table 21**) provided a 34% reduction in comparison with HD5 (−27.3 ± 2.8 to −18.2 ± 1.6 kcal/mol).

Most of the simulated aptamers can bind GP120 more strongly than CD4, underscoring their potential as competitive inhibitors of GP120. Moreover, short sequences such as Short-HD4 (17nt), Mod-HD4 (23nt), and Short-HD5 (28nt) displayed weaker ΔG values than CD4, highlighting that length, binding mode, interaction profile, and sequence are critical factors in optimizing nucleic acid binders. Specifically, Apt 5, the shortest aptamer of this series, established the highest average ΔG (−13.8 ± 1.7 kcal/mol) (**Figure 51).**

Apt 0, the longest aptamer simulated in this project (72nt), by yielding −33.4 ± 3.3 kcal/mol, is comparable to shorter aptamers, such as HD2 2'OMe RNA (34nt) and HD4 DNA (33nt), which achieved similar ΔG values (−31.6 ± 2.2 and −31.6 ± 2.9 kcal/mol, respectively).

These results indicate that aptamer length is a significant parameter in achieving high binding affinity, suggesting that the contributions of additional nucleotides deliver a small improvement in binding affinity while increasing structural complexity.

**Table 21**. The binding free energy, glycan, and protein percentages for all the aptamers and backbones analyzed throughout the project

| Aptamer | Backbone | Avg ΔG ± SEM (kcal/mol) | Avg Glycan % | Avg Protein % | Avg CD4 BD % | Normalized ΔG by length ± SEM (kcal/mol) |
|---|---|---|---|---|---|---|
| Apt 0 | DNA | −33.4 ± 3.3 | 70.3 | 29.7 | 5.9 | −0.4 ± 0.1 |
| Apt 1 | DNA | −19.0 ± 2.7 | 64.7 | 35.3 | 5.3 | −0.5 ± 0.1 |
| Apt 2 | DNA | −21.4 ± 2.5 | 62.5 | 37.5 | 10.6 | −0.5 ± 0.1 |
| Apt 3 | DNA | −23.8 ± 3.6 | 60.3 | 39.7 | 3.7 | −0.7 ± 0.1 |
| Apt 4 | DNA | −15.0 ± 2.2 | 55.0 | 45.0 | 15.0 | −0.7 ± 0.1 |
| Apt 4 | RNA | −16.6 ± 2.5 | 39.7 | 60.3 | 11.0 | −0.7 ± 0.1 |
| Apt 5 | DNA | −13.8 ± 1.7 | 57.6 | 42.4 | 14.5 | −0.9 ± 0.1 |
| HD5 | DNA | −27.3 ± 2.8 | 69.3 | 30.7 | 6.5 | −0.6 ± 0.1 |
| HD5 | RNA | −27.2 ± 4.0 | 61.9 | 38.1 | 8.4 | −0.5 ± 0.1 |
| HD5 | 2'F RNA | −23.9 ± 2.6 | 72.1 | 27.9 | 7.4 | −0.5 ± 0.1 |
| HD5 | 2'OMe RNA | −24.0 ± 3.3 | 61.0 | 39.0 | 7.4 | −0.4 ± 0.1 |
| S HD5 | DNA | −18.2 ± 1.6 | 55 | 45 | 7.3 | −0.6 ± 0.1 |
| HD4 | DNA | −31.6 ± 3.0 | 49.8 | 50.2 | 7.0 | −0.9 ± 0.1 |
| HD4 | RNA | −23.5 ± 2.9 | 55.9 | 44.1 | 8.8 | −0.7 ± 0.1 |
| HD4 | 2'F RNA | −23.0 ± 2.7 | 64.7 | 35.3 | 9.0 | −0.7 ± 0.1 |
| HD4 | 2'OMe RNA | −21.9 ± 2.2 | 54.3 | 45.7 | 11.9 | −0.6 ± 0.1 |
| S HD4 | DNA | −17.4 ± 2.1 | 44.7 | 55.3 | 16.8 | −1.0 ± 0.1 |
| S HD4 | RNA | −14.3 ± 1.8 | 50.0 | 50.0 | 20.9 | −0.8 ± 0.1 |
| S HD4 | 2'F RNA | −15.7 ± 1.4 | 61.0 | 39.0 | 12.2 | −0.9 ± 0.1 |
| S HD4 | 2'OMe RNA | −14.2 ± 2.3 | 48.7 | 51.3 | 15.6 | −0.8 ± 0.1 |
| M HD4 | DNA | −16.3 ± 2.5 | 57.6 | 42.4 | 16.4 | −0.7 ± 0.1 |
| M HD4 | RNA | −16.9 ± 2.0 | 53.4 | 46.6 | 13.9 | −0.7 ± 0.1 |
| HD3 | DNA | −24.9 ± 2.8 | 60.4 | 39.6 | 12.5 | −0.8 ± 0.1 |
| HD3 | RNA | −22.5 ± 2.2 | 56.5 | 43.5 | 11.6 | −0.7 ± 0.1 |
| HD3 | 2'F RNA | −22.9 ± 2.6 | 58.8 | 41.2 | 8.4 | −0.7 ± 0.1 |
| HD3 | 2'OMe RNA | −24.1 ± 2.8 | 62.2 | 37.8 | 5.9 | −0.8 ± 0.1 |
| HD2 | DNA | −24.8 ± 3.1 | 66.6 | 33.4 | 10.9 | −0.7 ± 0.1 |
| HD2 | RNA | −23.6 ± 3.1 | 61 | 39 | 10.1 | −0.7 ± 0.1 |
| HD2 | 2'F RNA | −28.1 ± 3.0 | 63.5 | 36.5 | 7.5 | −0.8 ± 0.1 |
| HD2 | 2'OMe RNA | −31.6 ± 2.2 | 66.6 | 33.3 | 7.6 | −0.9 ± 0.1 |
| CD4 | / | −20.7 ± 2.5 | 50.0 | 50.0 | 25 | / |

Two other key parameters were analyzed to evaluate the aptamers' behavior in MD-MM/PBSA calculations:

1. Efficiency: ΔG normalized per nucleotide (ΔG /length)
2. Specificity: the fraction of the total ΔG retrieved by considering only the CD4-Binding domain

Regarding the first parameter, Short-HD4 (−1.0 ± 0.1 kcal/mol/nt), HD4 DNA (−0.9 ± 0.1 kcal/mol/nt), and HD2 2'OMe RNA (−0.9 ± 0.1 kcal/mol/nt) emerged as the most promising sequences. In contrast, HD5 and Apt 0, with their bulky structures, showed poor efficiency (−0.5 ± 0.1 kcal/mol/nt for both) (**Figure 52**). These data suggest that the length optimization approach can increase efficiency at the expense of total binding affinity. In contrast, Apt5, which displayed the weakest predicted binding affinity, can obtain a high ΔG /length ratio (−0.9 ± 0.1 kcal/mol/nt), indicating that even with its short length, each nucleotide interacts more efficiently with GP120.

This comparison highlights a fundamental trade-off between binding efficiency and total free energy. Moreover, this trade-off illustrates that short aptamers (below 20 nucleotides) display a better contribution per nucleotide to the ΔG, but with an overall lower total binding free energy values through multiple non-specific interactions.

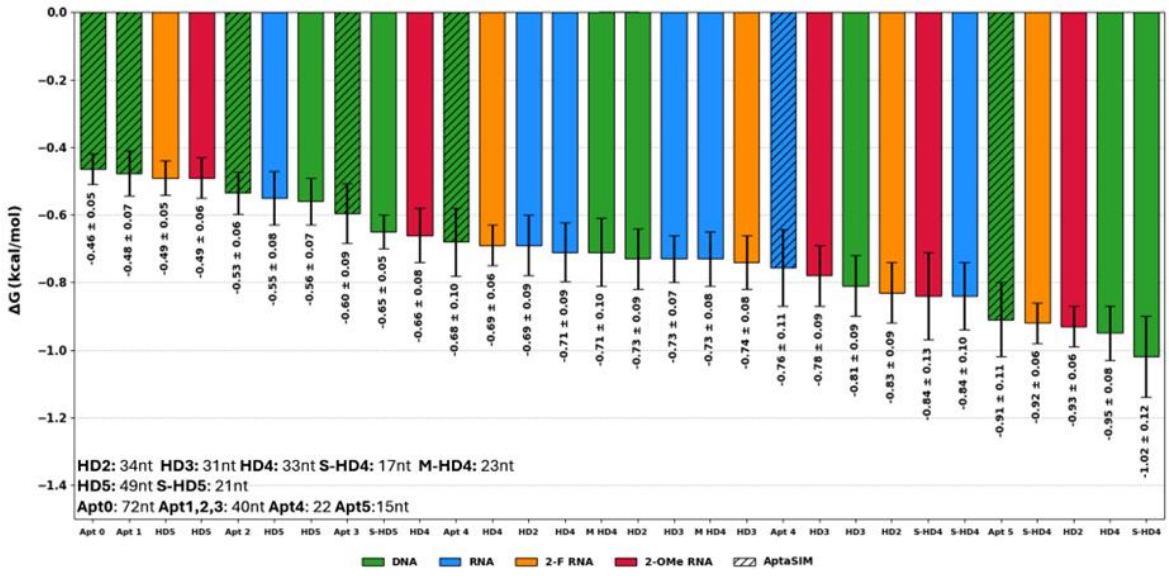



**Figure 52**: Normalized binding free energies (ΔG/length) for all aptamers analyzed in this study. The bars are ordered from weaker to stronger binding affinity (less negative to more negative ΔG). ΔG values are described with their related standard deviations (error bars). Aptamers are color-coded by backbone (DNA, RNA, 2'F RNA, 2'OMe RNA in green, blue, orange, and red, respectively), while hatched bars represent the aptamers predicted by AptaSIM.

When evaluating specificity, defined as the portion of the entire ΔG collected by considering only the CD4-BD, CD4 exhibited the highest value, accounting for 25% of the overall ΔG, reflecting its optimized binding to GP120, thanks to its conformation and the spread of its surface across the CD4-binding domain. On the other hand, aptamers under investigation showed a lower specificity, on average, between 7–10%, with small variation across the different backbones analyzed. Remarkably, Short-HD4 RNA achieved the highest specificity among aptamers (20.9%). This result can be explained by how the aptamers primarily interact with GP120, specifically due to a large exposition to glycans, with limited engagement of residues within the CD4 binding domain.

**Glycan and Protein contributions to aptamers' binding free energy.**

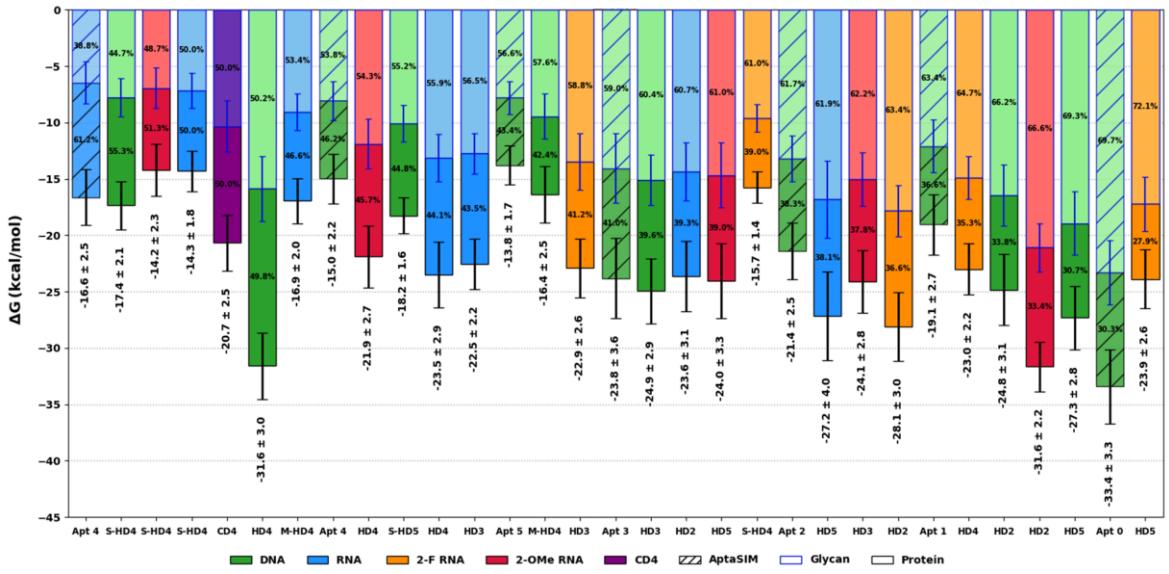



**Figure 53**: Aptamers ΔG binding values sorted by increasing glycan percentage. The chart discriminates energetic contributions of protein and glycans across various chemical modifications (DNA, RNA, 2-F RNA, and 2-OMe RNA).

To compare the CD4 binding profile, which described a 50/50 ratio regarding protein-glycan interactions, with the aptamers, an in-depth evaluation was performed. This evaluation was carried out to calculate the binding free energy of proteinaceous residues and the glycans. We found out that the glycan-aptamer interactions are the dominant source of binding energy, accounting for 44–72% of the total ΔG (**Figure 53).** Despite their similar binding efficiencies, the lead aptamers HD4/DNA and HD2/2'OMe-RNA utilize different binding patterns. HD4/DNA accurately replicates the CD4 interaction profile with a 50:50 glycan-protein ratio, whereas HD2/2'OMe-RNA favors glycan interactions at a 66:34 ratio.

The fact that these different profiles yield similar binding free energies emphasizes that high-quality aptamer design must account for a complex interplay of factors, including folding, sequence, and backbone structure. The advantage of targeting glycans permits targeting a broader spectrum of HIV variants, but it could establish non-specific interactions with other glycoproteins. However, due to a high protein-binding profile, aptamers could exert selective pressure on the virus, potentially leading to the mutation of key interacting residues, the formation of new strains, and the promotion of HIV resistance.

**I) Final remarks**

In conclusion, these computational insights suggest that aptamers can engage a broader range of residues than CD4, a capability primarily driven by their peculiar structural plasticity. Collectively, this research highlights the profound influence of backbone chemistry and primary sequence on binding affinity and efficiency, indicating that optimal design requires a strategic balance between molecular rigidity and adaptability.

More specifically, a rigid scaffold is required to ensure strong and stable interactions with the CD4-BD, while adaptability enables interactions with glycans, which display dynamic conformational behavior. Particularly, glycan-mediated binding may be less specific than CD4 BD-based recognition, but it offers a different therapeutic advantage by potentially reducing susceptibility to viral escape mutations within the ENV protein.

In conclusion, this study provides significant insights:

- Optimized binding affinities can be achieved without bulky and long aptamer structures, as highlighted by HD5 and Apt 0
- Short aptamers can reach high efficiency; however, since their total ΔG value is lower than that of CD4, this may suggest limited *in vivo* activity.
- Glycans are key components in the aptamers' binding. For this reason, they are required in the simulation systems to predict the aptamer binding stability and binding mode.
- Backbone modifications exert highly sequence-specific effects, rather than providing a universal enhancement across all candidates.

Considering all data produced in this study, the most promising aptamers, specifically HD2-2'OMe RNA, S-HD4, HD4, and Apt 0, could be purchased and investigated through biological and biophysical assays. An international collaboration between the University of Milan, the University of Verona, and the University of California could accomplish this task and validate these computational outcomes.

# 6. Appendix

**Appendix 1**

**Title**: PCSK9 inhibitors: a patent review 2018-2023

**Authors**: Fassi, E. M. A., Citarella, A., **Albani, M**., Milano, E. G., Legnani, L., Lammi, C., Silvani, A. and Grazioso, G.

**Status**: Published in *Expert Opinion on Therapeutic Patents (2024)*

**DOI**: 10.1080/13543776.2024.2340569

**Author Contribution**: Dr. M. Albani contributed to gathering the patent-related information and to preparing the original draft.

Expert Opinion on Therapeutic Patents

ISSN: 1354-3776 (Print) 1744-7674 (Online) Journal homepage: www.tandfonline.com/journals/ietp20

# PCSK9 inhibitors: a patent review 2018-2023

Enrico Mario Alessandro Fassi, Andrea Citarella, Marco Albani, Erica Ginevra Milano, Laura Legnani, Carmen Lammi, Alessandra Silvani & Giovanni Grazioso

**To cite this article:** Enrico Mario Alessandro Fassi, Andrea Citarella, Marco Albani, Erica Ginevra Milano, Laura Legnani, Carmen Lammi, Alessandra Silvani & Giovanni Grazioso (2024) PCSK9 inhibitors: a patent review 2018-2023, Expert Opinion on Therapeutic Patents, 34:4, 245-261, DOI: 10.1080/13543776.2024.2340569

**To link to this article:** https://doi.org/10.1080/13543776.2024.2340569

EXPERT OPINION ON THERAPEUTIC PATENTS
2024, VOL. 34, NO. 4, 245–261
https://doi.org/10.1080/13543776.2024.2340569

Taylor & Francis
Taylor & Francis Group

REVIEW



# PCSK9 inhibitors: a patent review 2018-2023

Enrico Mario Alessandro Fassi[a], Andrea Citarella[b], Marco Albani[a], Erica Ginevra Milano[a], Laura Legnani[c], Carmen Lammi[a], Alessandra Silvani[b] and Giovanni Grazioso[a]

[a]Dipartimento di Scienze Farmaceutiche, Università degli Studi di Milano, Milano, Italy; [b]Dipartimento di Chimica, Università degli Studi di Milano, Milano, Italy; [c]Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Milano, Italy

**ABSTRACT**

**Introduction:** Proprotein convertase subtilisin/kexin 9 (PCSK9) plays a crucial role in breaking down the hepatic low-density lipoprotein receptor (LDLR), thereby influencing the levels of circulating low-density lipoprotein cholesterol (LDL-C). Consequently, inhibiting PCSK9 through suitable ligands has been established as a validated therapeutic strategy for combating hypercholesterolemia and cardiovascular diseases.

**Area covered:** Patent literature claiming novel compounds inhibiting PCSK9 disclosed from 2018 to June 2023 available in the espacenet database, which contains more than 150 million patent documents from over 100 patent-granting authorities worldwide.

**Expert opinion:** The undisputable beneficial influence of PCSK9 as a pharmacological target has prompted numerous private and public institutions to patent chemical frameworks as inhibitors of PCSK9. While several compounds have advanced to clinical trials for treating hypercholesterolemia, they have not completed these trials yet. These compounds must contend in a complex market where new, costly, and advanced drugs, such as monoclonal antibodies and siRNA, are prescribed instead of inexpensive and less potent statins.

**ARTICLE HISTORY**
Received 30 November 2023
Accepted 28 March 2024

**KEYWORDS**
PCSK9; hypercholesterolemia; CVD; LDL-C; patents

## 1. Introduction

Hypercholesterolemia is the most widespread form of dyslipidemias, causing, over the past years, a significant increase in the global mortality. Hypercholesterolemia is caused by high levels of low-density lipoprotein cholesterol (LDL-C) in the bloodstream, a major contributing factor to the development of atherosclerotic cardiovascular disease (CVD) [1]. It has been reported that, over 5 years of trial, a low circulating level of LDL-C leads to: i) a decrease in myocardial infarction or coronary death, ii) a reduction in coronary revascularization treatments, iii) the decrease in the number of strokes and the major vascular events [2]. These authors, together with others reaching the same conclusions, led the Endocrine Society Guidelines to recommend an LDL-C circulating level lower than 55 mg/dL (equivalent to 1.4 mmol/L), especially for patients affected by CVDs [3]. Nevertheless, when this level is not respected by dietary habits, or for the presence of genetic disorders (such as familial hypercholesterolemia), individuals are recommended to use statins as a drug preventing fatal or non-fatal CV events. Unfortunately, not all patients are susceptible to the treatments with 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoAR) inhibitors such as statins, due to intolerance or statin-resistant hypercholesterolemia [4,5]. Moreover, long-term use of them resulted in tachyphylaxis, the decreasing response to therapeutic agents [6]. Specifically, a study on 254 patients found that LDL-C levels initially decreased during statins treatment but subsequently increased over time, even after reaching maximum LDL-C reduction. This was particularly noticeable in patients treated with atorvastatin at exposure doses of 10 or 20 mg/day [6]. From a molecular point of view, through the direct ability to inhibit the endogenous cholesterol synthesis via HMG-CoAR inhibition, statins activate the SREBP-2 pathway leading to the level increase of both LDL receptor (LDLR) and HMG-CoAR [7]. In addition, it has been observed that statins also raise levels of proprotein convertase subtilisin/kexin type 9 (PCSK9), a protein modulating the LDLR degradation. Increases in PCSK9 levels also reduce the LDL-C response to statin therapy [8]. Therefore, there has been a rapid and substantial expansion in the range of lipid-lowering medications, acting by novel mechanisms of action. These mechanisms include monoclonal antibodies (mAbs) acting on angiopoietin-like 3 (evinacumab, Ulterius Pharmaceuticals), microsomal triglyceride transfer protein inhibitors (lomitapide, Aegerion), ATP-citrate lyase inhibitors (bempedoic acid), and last but not least, PCSK9 inhibitors [9–12]. As stated before, PCSK9 is an enzyme circulating in the bloodstream and is responsible for regulating the population of LDLRs on the surface of liver cells. The PCSK9/LDLR interaction results in the internalization and degradation of LDLR, reducing the liver cells' ability to capture LDL-C from the bloodstream [13]. Consequently, inhibiting the interaction between PCSK9 and LDLR leads to an increased number of LDLRs on the cell membrane, enhancing the liver cells' capacity to uptake LDL-C [14]. In addition to its crucial role in LDL-C metabolism, PCSK9 has been linked to various processes relevant to cardiovascular health. Particularly, PCSK9 levels can predict the occurrence of recurrent cardiovascular events in individuals with coronary artery disease, even when their LDL-C levels are

CONTACT Enrico Mario Alessandro Fassi enrico.fassi@unimi.it; Giovanni Grazioso giovanni.grazioso@unimi.it Dipartimento di Scienze Farmaceutiche, Università degli Studi di Milano, Via L. Mangiagalli 25, Milano 20133, Italy

**Article highlights**

- PCSK9 inhibitors may become a new generation of *drugs* for the treatment of hypercholesterolemia and the prevention of fatal or non-fatal cardiovascular events.
- Several chemical scaffolds as PCSK9 inhibitors have been reported to date.
- This review reports on the PCSK9 inhibitors patented from 2018 to June 2023.
- Nineteen patents have been deposited by several companies and universities.
- One PCSK9 inhibitor (MK-0616) entered phase III clinical trial.

well-controlled. This underscores the growing recognition of PCSK9 s involvement in platelet reactivity and the formation of blood clots, implying the clinical importance of inhibiting PCSK9 through pharmacological methods [15]. In the past few years, both academia and pharmaceutical corporations have invested significant resources in discovering compounds with the ability to block PCSK9, both the biosynthesis and the interaction with LDL-R [16]. Since PCSK9 inhibitors exert a mechanism of action that is completely different to the statins one, this makes them eligible for the treatment of patients who cannot be treated by statins, due to harsh side effects or to therapeutic (total/partial) ineffectiveness in some, severe, forms of hypercholesterolemia (e.g. familial hypercholesterolemia) [17]. As a consequence, this led to the development of effective therapeutic strategies for addressing hypercholesterolemia resistant to the statin treatment, as demonstrated by the introduction of two monoclonal antibodies, alirocumab (Sanofi) and evolocumab (Amgen), onto the market [18,19]. Novartis pioneered the development of the first siRNA drug, known as Inclisiran, which can disrupt the liver's PCSK9 transcription. This disruption results in long-lasting reductions in cholesterol levels among treated patients [20]. However, these medications are costly and face challenges in terms of patients' compliance due to their subcutaneous administration. Consequently, pharmaceutical companies and academic institutions have a strong interest in advancing the clinical development of orally bioavailable small molecules. This interest is evidenced by the substantial number of patent applications in this field in the last decades [21,22]. Herein, among the 810 patents resulted from the worldwide patent search conducted on espacenet database (https://worldwide.espacenet.com), we summarized the disclosures regarding only the compounds which directly interact and inhibit PCSK9, obtaining a total of nineteen patents, deposited from 2018 to June 2023. In particular, here we report on the patents containing the word 'PCSK9' in the 'Title or abstract' considering each year of the time period examined in this review, making use of the 'advanced search' webtool of espacenet.

## 2. Patent review

### 2.1. Year 2018

#### 2.1.1. Cardio Therapeutics Pty Ltd

With this invention, Cardio Therapeutics Pty Ltd. patented a series of heterocyclic compounds as potential new PCSK9 small molecule inhibitors [23]. Based on the molecular modeling studies they have accomplished, it was concluded that these compounds can target PCSK9, thus preventing its interaction with the LDLR. This disclosure covers all the heterocyclic compounds with the general structure **1–4** (Figure 1). The embodiments of the patent cover any pharmaceutically acceptable salt, solvate, prodrug or polymorph of the compounds, in order to reduce LDL levels, through PCSK9 inhibition. The compounds of the invention may, moreover, be administered along with a pharmaceutical carrier, diluent, or excipient. Throughout the setting up and synthetic process of the molecules, the inventors selected some specific functional groups to put on the starting scaffold to enable these compounds to interact with some relevant amino acids in the protein sequence of PCSK9, such as: Ser221, Asp212, Lys223 and Lys258. The experimental approach used to establish these compounds' efficacy was, as a first thing, to measure their ability to disrupt the protein-protein interaction between LDLR and PCSK9 through a binding assay. Through this experiment, the inventors proved the mechanism of action of the compounds they proposed. The rate of inhibition was assessed as the absorbance of tested compounds at 450 nm and 540 nm, then plotted within $IC_{50}$ curves. Among all the 31 compounds reported as examples, 5 of them showed an $IC_{50}$ comprised between 0.1 and 1 µM (examples 5, 6, 7, 19, and 30), while only compound **1** (patented as example 27) exhibited an $IC_{50}$ lower than 0.1 µM, representing the most potent PCSK9 inhibitor reported in this patent (Figure 2).

Furthermore, they performed a cell-based LDL uptake assay. The assay was carried out on human liver HepG2 cells, considering the examples 6, 25, 27, and 30, tested at the concentration of 0.1 and 1 µM. AMB-657286, a small molecule inhibitor of the PCSK9-LDLR interaction developed by Shifa Biomedical Corporation [24], was employed as a positive control. In this assay, the dynamic range of measuring LDL uptake was improved through the addition of a gain-of-function mutant of PCSK9, that furtherly impairs LDL uptake (increased binding and degradation of LDLR) and when inhibited represents a sign of functionality against the target. The quantification of the uptake was determined by fluorescence, on a reader set at excitation/emission wavelengths of 485 and 530 nm, respectively. Also in this case, compound **1** exhibited the best results (Figure 2). More details are available in their research article published on the Bioorganic and Medicinal Chemistry where the compound **1** (example 27) can be found with the name of 3d [25].

#### 2.1.2. Merck sharp & Dohme corporation

In 2018, Merck Sharp & Dohme Corporation patented a series of novel PCSK9 allosteric modulators consisting of substituted 1-methyl-1,2,3,4-tetrahydroisoquinolines with general structure **5** (Figure 3), endowed with high affinity for a peculiar site of the protein, that is not involved in the PCSK9-LDLR interaction [26]. In particular, this unprecedented allosteric binding pocket was predicted *in silico* by using the Molecular Operating Environment Site Finder tool, and it is located between the catalytic and C-terminal domain of PCSK9 shaped

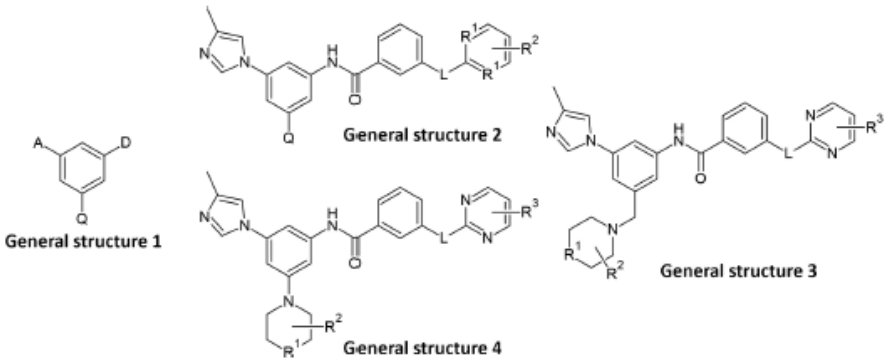



Figure 1. General structure 1–4 of the heterocyclic compounds patented by Cardio Therapeutics Pty Ltd as potential new PCSK9 small molecule inhibitors.

by Asp360, Arg357, Arg458 and Arg476 [27]. The authors also published the X-ray structures of several compounds, which are covered by the patent, in complex with the above-mentioned PCSK9 allosteric site (PDB accession codes: 6U26, 6U2N, 6U2P, 6U36, 6U38, 6U3X) [27]. Additionally, they seem to bind to such site independently from PCSK9's binding to LDLR. Besides inducing PCSK9 degradation through this unusual pathway, these compounds displayed the capacity of reducing intracellular levels of PCSK9 (both as pro-protein and mature form) and are potentially important in the treatment of clinical conditions affected by PCSK9 levels and functionality, such as atherosclerosis, hypercholesterolemia, coronary heart disease, metabolic syndrome, acute coronary syndrome, and related CV diseases. This invention covers any pharmaceutically acceptable salt or any pharmaceutical composition comprising one of said molecules or its salification product and any pharmaceutically accepted carrier. Talking about the combination, the other drug can be of various types, going from statins to Angiotensin – Converting Enzyme (ACE) inhibitors and sartans and many other classes. The embodiments cover all forms of the compounds presented, such as solvates, hydrates, stereoisomers, tautomers but also free acids and bases and alternative salts.

Several experiments were carried out to assess the compounds' capacity to bind the PCSK9 protein, such as Differential Scanning Calorimetry (DSC) and Fluorescence Polarization (FP), and their effectiveness in inhibiting its functionality performing a protein degradation assay. In the DSC assay, compounds designated as examples 1–92 and 109–112 were tested and most of them exhibited ΔTm values > 1°C. The higher the temperature shift, generally, the higher is the affinity. The stabilization of PCSK9 through these molecules' binding refers to an inactive form of the protein, that impairs its ability to bind to LDLR and, therefore, leads to an increased level of receptors on hepatocytes' surface and ameliorates the uptake of circulating LDL. The compound with the highest affinity in this test is compound **2** (example 85), exhibiting a ΔTm value of 9.623 (Figure 3). Also in the FP assay, the high affinity bond relates to the stabilization of an inactive conformation of PCSK9 and the consequent boost of

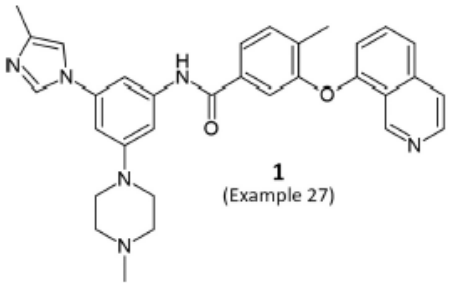



Figure 2. Chemical structure of the most potent PCSK9 inhibitor reported in the patent.

LDLR population. The lowest $IC_{50}$ value was observed in the case of compound **3** (example 14, 2.37 nM), represented in Figure 3. Finally, in the protein degradation assay, the reported $DC_{50}$ values represent the effective concentration of the tested molecule to lower PCSK9 protein levels, of both 'pro' and 'mature' forms, by 50%. The most effective compounds in lowering pro-PCSK9 are represented by compound **4** (example 94), examples 99 and 104, exhibiting $DC_{50}$ values of 1.36, 1.37, and 1.53 μM, respectively, while the significantly most active against the mature form is compound **5** (example 96), showing a $DC_{50}$ value of 1.05 μM (Figure 3).

### 2.1.3. Aarhus university

Liver heparan sulfate proteoglycans are PCSK9 receptors and essential for PCSK9-induced LDLR degradation. The heparan sulfate-binding site is located in the PCSK9 prodomain and formed by surface-exposed basic residues interacting with trisulfated heparan sulfate disaccharide repeats. Accordingly, heparan sulfate mimetics are potent PCSK9 inhibitors. It was proposed by Gustafsen et al., from Aarhus University of Denmark, that heparan sulfate proteoglycans lining the hepatocyte surface capture PCSK9 and facilitate subsequent PCSK9-LDLR complex formation [28].

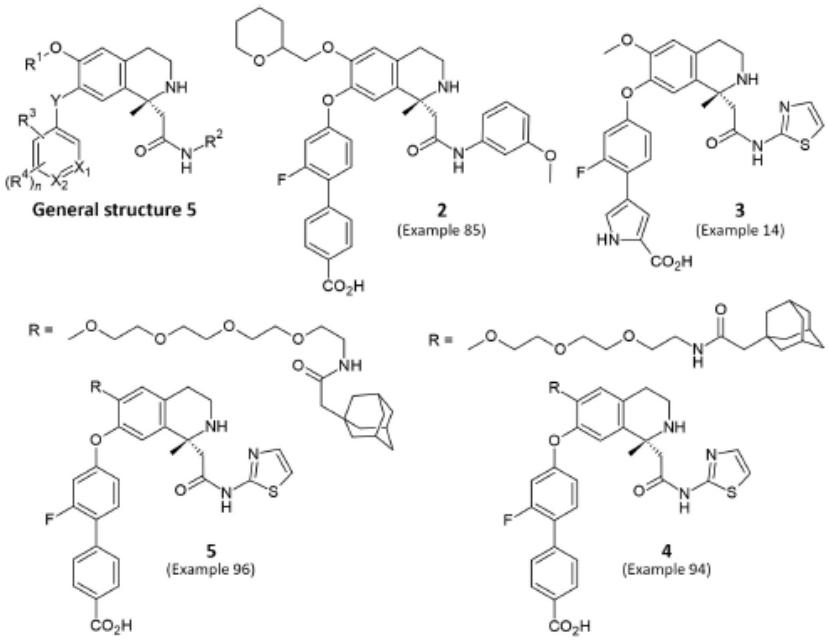



Figure 3. General structure and representative examples of PCSK9 allosteric binding tetrahydroisoquinolines patented by Merck Sharp & Dohme Corporation in 2018.

Notably, the patent entitled 'Compounds for treatment of lipoprotein metabolism disorders' by Gustafsen et al. relates to the use of heparin analogues as PCSK9 inhibitors for the treatment of lipid disorders [29]. In the first aspect, the invention relates to a composition comprising a compound with general structure **6**, or a pharmaceutically acceptable salt, solvate, polymorph or tautomer thereof, described in Figure 4.

Several heparin mimetic compounds were developed, and seven of them are currently being used in clinical settings for various purposes. These compounds are referred to as heparin/heparan sulfate mimetics and belong to different chemical classes, including oligosaccharides, oligonucleotides, and naphthalene derivatives [30]. Dedicated tests on a subset of these mimetics to evaluate their ability to bind PCSK9 and increase LDLR in HepG2 cells were carried out. The sulfated oligosaccharides dextran sulfate and pentosan sulfate both directly bound to PCSK9 with affinity constants of 180 and 381 µM, respectively, as estimated using Microscale Thermophoresis (MST). This binding resulted in a dose-dependent increase in cellular LDLR, reaching a plateau of around 400% compared to the control. It was observed that the interaction depended on the presence of sulfate groups, as non-sulfated dextran showed no affinity for PCSK9. Furthermore, the sulfated naphthalene derivative suramin, an antiparasitic drug used against African sleeping sickness, led to an up to 15-fold increase in LDLR and consequently ameliorated the HepG2 ability to uptake LDL. Phosphorothioate oligonucleotides are highly anionic and are known to interact with heparin-binding proteins [31,32]. Therefore, a modified 36–mer phosphorothioate oligodeoxycytidine (S-dC-36) was assessed demonstrating that it successfully binds to PCSK9 with a $K_d$ of 4.8 µM, with inhibitory activity within this range of concentration. Based on this evidence, it appears clear that several negatively charged compounds exhibit inhibitory activity toward PCSK9. To obtain a better understanding of the structure – activity relationships, crystallographic data on a PCSK9-heparin mimetic complex was acquired, suggesting that a dextran sulfate disaccharide interacting with Arg93, Arg97, Arg104, and His139 of PCSK9. Overall, these findings suggest that the number and placement of negatively charged

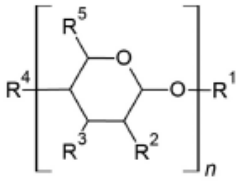


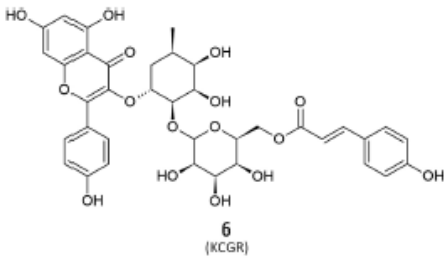

Figure 4. General structure of the heparin analogues patented by Aarhus University in 2018.

functional groups of sulfated sugars and mimetics are crucial determinants for their binding and activity toward PCSK9.

### 2.2. Year 2019

#### 2.2.1. China pharmaceutical university

The 'China Pharmaceutical University' of Nanjing patented the use of Kaempferol 3-O-2''-(6'''-*p*-coumaroyl)glucosyl rhamnoside (KCGR), a monomer component of a traditional Chinese medicine, for the treatment of diseases mediated by PCSK9 [33]. The chemical structure of KCGR (compound **6**) is made by a central rhamnose core which hemiacetal -OH group forms a glycosidic bond with a Kaempferol fragment while its hydroxymethyl functionality is esterified with a residue of *p*-coumaric acid (Figure 5). Study conducted by patent inventors demonstrated the efficacy of the inhibitor in providing a beneficial effect in the reduction of the level of lipids in blood. In particular, the authors discovered for the first time that KCGR possesses the specific activity of inhibiting both the combination of PCSK9 and LDLR. In fact, KCGR exhibited a reduction of the levels of TC (total cholesterol), TG (triglycerides) and LDL-C in the plasma of hyperlipidemic mice models, an increasement of the level of HDL-C (high-density lipoprotein cholesterol), and an improvement of the lipid metabolism. *In silico* studies showed the ability of KCGR to bind with high affinity to the active pocket region of PCSK9 crystal structure and establish stable H-bond interactions with multiple surrounding amino acid residues [34]. Moreover, a binding assay confirmed that KCGR and human recombinant PCSK9 protein had strong binding force ($K_d$ = 15 nM) [34]. In particular, the bio-layer interferometry method was employed to ascertain the binding kinetics between the ligand and PCSK9, utilizing the Octet RED96 system at 25°C. Taken all together, these findings represent an advancement in the treatment of cardiovascular diseases and certain forms of hyperlipidemia.




**6**
(KCGR)

Figure 5. Chemical structure of Kaempferol 3-O-2''-(6'''-p-coumaroyl)glucosyl rhamnoside (KCGR).

#### 2.2.2. Merck Sharp & Dohme corporation

In 2019, Merck Sharp & Dohme Corporation published two different patents covering novel macrocyclic inhibitors of PCSK9 with promising activity at nanomolar level. In the first patent, the chemical structures of the compound belong to a series of pseudopeptide macrocyclic derivatives bearing several modified aminoacidic residues, with general structure **7** (Figure 6) [35]. Selected compounds of the invention were examined by *in vitro* assays to determine their potency in blocking PCSK9 activity. Binding affinity was described as $K_i$ and the values were obtained through biological assessments using Alexa FRET Plus ($K_i$-plus) and Alexa FRET Ultra ($K_i$-ultra) TR-FRET kits and they are reported in nanomolar (nM). The eight most promising compounds showed a $K_i$-plus lower than 0.00558 nM and a $K_i$-ultra comprised between 0.00074 (exhibited only by compound **7**, patented as EX-31) and 0.00597 nM (Figure 6).

The second patent highlights the advantages of using novel macrocyclic compounds as potent PCSK9 inhibitors [36]. Chemical structure of the inhibitors resembles a pseudopeptide backbone, embodying a functionalized tryptophan, specifically bearing a fluorinated indole moiety and methylated tyrosine and proline frameworks, with general structure **8** (Figure 7). Selected compounds of the invention were examined by two *in vitro* assays to determine their potency in blocking PCSK9 activity. $K_i$ measures were obtained through biological assessments using Alexa FRET Standard ($K_i$ standard) and Alexa FRET Plus ($K_i$ plus) TR-FRET kits and the values are reported in nanomolar (nM). The most active compounds turned out to be compound **8** (patented as Ex-B03)

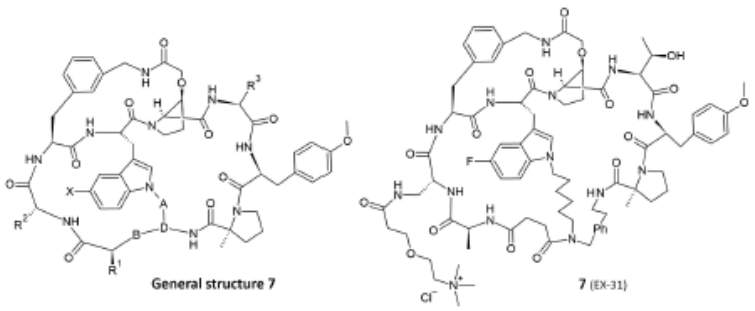



Figure 6. General structure of the compounds covered by Merck Sharp & Dohme Corporation, and the chemical structure of EX-31, which represent the most potent compound of the series.

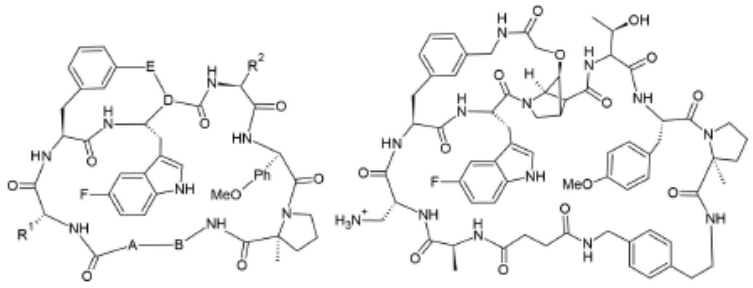



Figure 7. General structure of the compounds covered by the second patent of Merck Sharp & Dohme Corporation, and the chemical structure of Ex-B03, which represent the most potent compound of the series.

and Ex-B04, bearing an amide bridge between modified proline and phenylalanine residues, and derivatives EX-OT03–06, where the same amino acidic residues are connected via a triazole functionality to construct the tricyclic architecture. All these compounds showed a $K_i$ standard lower than 1.26 nM and $K_i$ plus comprised between 0.008 (exhibited only by compound **8**) and 0.32 nM (Figure 7).

Interestingly, this class of compounds was furtherly developed, finally attaining a promising drug candidate (see section 2.4.1 for details).

### 2.3. Year 2020

#### 2.3.1. Novartis AG

In 2020, Novartis AG has filed two different patents regarding tetrameric and pentameric cyclic polypeptide PCSK9 inhibitors, respectively, which are helpful in the treatment of cholesterol lipid metabolism and other metabolic disorders in which PCSK9 protein is involved.

The first patent covers all the compounds with general structure **9** (Figure 8) and comprises their pharmaceutically acceptable salts, hydrates, solvates, prodrugs, stereoisomers, N-oxides, or tautomers [37]. The PCSK9 binding capacity of the compounds was measured using a TR-FRET assay which was performed on EnVision or PheraStar instruments. These experiments measured the ability of the compounds to interfere with the binding of human PCSK9 protein to human LDLR, providing the measures of both potency ($IC_{50}$) and efficacy. A total of 178 compounds were reported as examples, and, among them, the $IC_{50}$ values ranged from 1.49 µM to 150 pM, however the $IC_{50}$ values of the majority of the compounds fall in the picomolar range. In fact, a total of 106 polypeptides out of 178 showed $IC_{50}$ values below 1 nM, and four of these showed values below 200 pM, as observed for compounds **9–12** (patented as example #10, #12, #138, and #139, respectively), representing the most potent PCSK9 inhibitors reported in the patent (Figure 8).

The second Novartis AG disclosure relates to PCSK9 inhibitors with general structure **10** (Figure 9) [38]. The company, already in the market with Inclisiran, in this patent covered cyclic pentameric *N*-substituted peptides in which X is C or N atoms and $R_n$ substituents can be often H, C1-C6 alkyl or haloalkyl groups (Figure 9). More than 300 examples of compounds were reported in the patent. For all of them the PCSK9 inhibiting activity was measured by TR-FRET assays using unlabeled human PCSK9 as positive control and buffer/DMSO as negative control. The highest inhibitory activity, associated with the lowest $IC_{50}$ value (0.9 nM), was displayed by compound **13** (patented as example 85) reported in Figure 9. Interestingly, 13 other compounds showed $IC_{50}$ values lower than 1 nM and a protein inhibition activity higher than 88%.

As usual in the patent descriptions, no structural or mechanistic information is provided and, to the best of our knowledge, the inventors did not publish any data on these compounds, only one of the inventors (Yang L.), together with other Novartis researchers, published an interesting paper in which macrocyclic natural peptides were discovered [39]. In this effort, the Novartis researchers, through an affinity-based screening process involving over a trillion in vitro-translated macrocyclic peptides, successfully pinpointed compounds that exhibit a strong binding affinity to PCSK9 at the low-nanomolar level. These compounds not only disrupted the function of PCSK9 but also made use of a distinctive, induced-fit pocket of PCSK9 [39]. It is possible that the compounds covered by this patent could share with the macrocyclic peptides reported in the published article the discovery process and the action mechanism.

#### 2.3.2. Dogma Therapeutics, Inc

The patent of Dogma Therapeutics, Inc. describes novel heteroaryl compounds of general structure **11** differently substituted (Figure 10), or their pharmaceutically acceptable salts, which are able to interact with a new discovered binding pocket defined by amino acid residues Val589 and Ser636 of human PCSK9 protein [40]. In September 2020, AstraZeneca AB acquired the oral PCSK9 inhibitor programme from Dogma Therapeutics, Inc. These compounds can inhibit the activity of PCSK9, thus resulting in an increase in LDLR levels on the liver surface. The affinity ($K_D$) of the compounds on human recombinant PCSK9 protein was measured through Surface Plasmon Resonance (SPR) experiments, in particular, the biophysical data were collected on a Biacore™ system (GE Healthcare).

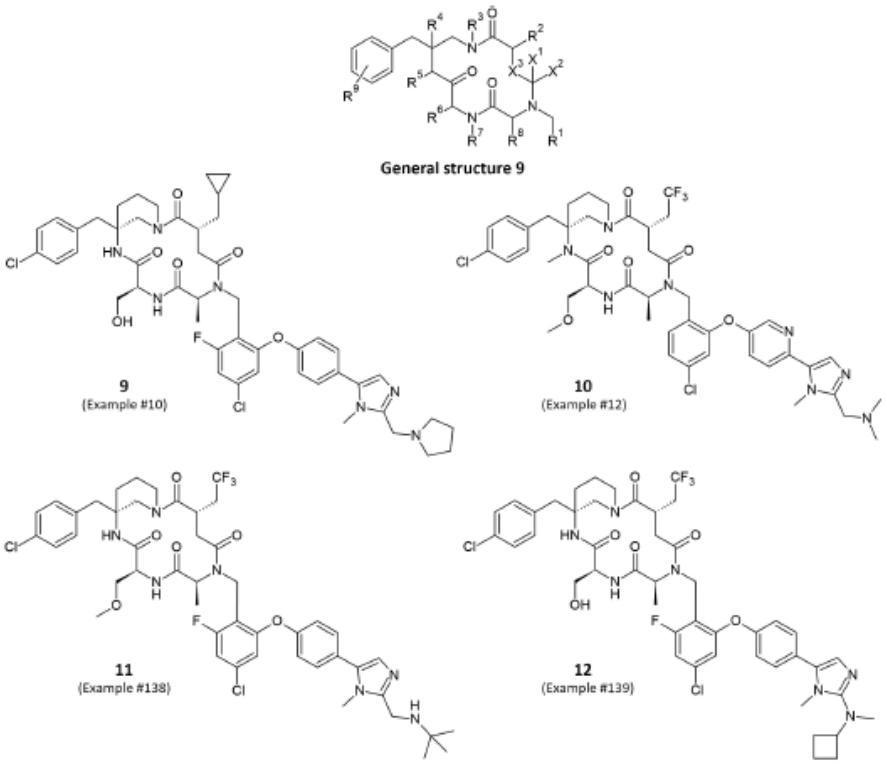



Figure 8. Chemical structure of the cyclic polypeptides which inhibit the PCSK9 protein, patented by Novartis AG, and chemical structures of the compounds 9–12, capable of inhibiting the PCSK9/LDLR interaction with an $IC_{50}$ of 180 pM, 152 pM, 152 pM, and 150 pM, respectively.

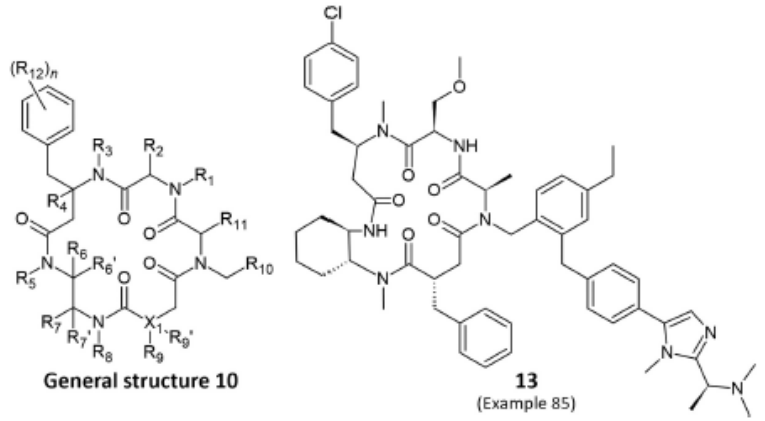



Figure 9. General structure of the compounds covered by Novartis AG patent, and chemical structure of compound 13 (example 85), which showed the highest inhibitory activity of the series.

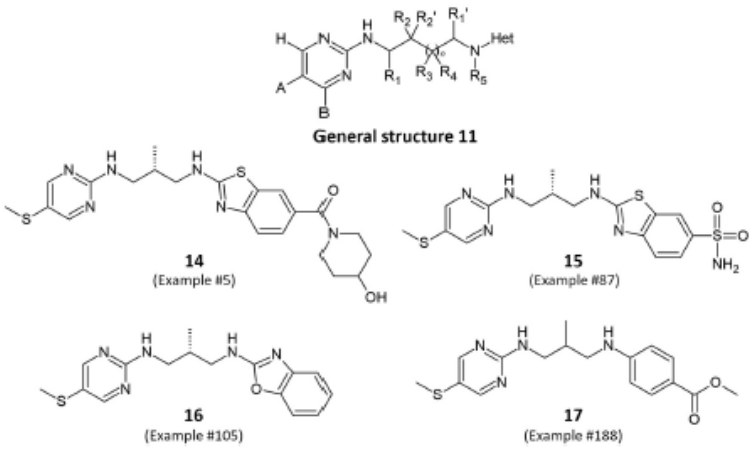



Figure 10. General structure of the novel heteroaryl allosteric inhibitors of PCSK9 protein, patented by Dogma Therapeutics Inc., and chemical structure of the representative compounds 14–17.

Out of 511 compounds tested, 261 showed a $K_D$ value lower than 200 nM.

The representative compounds with general structure **11**, for a total of eight compounds, were then selected for the *in vitro* cellular assay to measure the secreted PCSK9 levels, cellular LDLR levels and cell viability. An inhibitor of PCSK9 translation (SI – 1) was used as positive control and was synthesized according to the patent deposited by Pfizer in 2014 [41]. At the tested concentration of 75 µM, the compounds were able to increase LDLR levels up to 136% (in the case of example #89) and did not induce significant changes regarding cellular viability (values range between –14% and 12%). Surprisingly, the compounds inhibited also the PCSK9 secretion level up to 72% (in the case of example #175), but the mechanism of action was not highlighted in the patent description. Finally, the pharmacophore characterization was assessed by crystallographic experiments performed on compounds **14–17** (patented as examples #5, #87, #105, and #188, respectively), which are endowed with a $K_D$ value lower than 200 nM (Figure 10). Crystals were cryo-cooled for synchrotron data collection at ESRF beamline ID30A –1 or DLS beamline I04 on a Pilatus3 2 M or 6 M –F detector, respectively.

### 2.4. Year 2021

#### 2.4.1. Merck Sharp & Dohme Corporation

In 2021, Merck Sharp & Dohme Corporation patented macrocycles representing the final stage of the development of PCSK9 inhibitors obtained by mRNA display technology [42–44]. This technique is a robust *in vitro* affinity selection system, enabling the examination of extensive peptide libraries genetically encoded. This platform controls the transcription/translation apparatus and post-translation cyclization, all within cell-free settings, to assemble these libraries. A noteworthy aspect of this technique is the covalent linkage of mRNA to the peptides it encodes, allowing for the amplification of desired peptides exhibiting interactions with the target of interest. The use of this technique, together with structure-based studies, led to the development of highly potent macrocyclic peptides with general structure **12**, considering the patent (Figure 11) [45]. As can be noted, the chemical structure of macrocyclic peptides bears a tricyclic amino acid portion with naturally occurring and non-naturally occurring amino acids variously linked via side chains.

The patented structures resulted from studies on linear peptides, progressively cyclized creating bridges between different points of the loops, giving the structure considerable conformational rigidity, protecting the cleavage sites susceptible to gut enzymes like trypsin, chymotrypsin [44]. In the first patent examples, it is interesting to note that the $R^2$ substituents are long chains containing various PEG, ethanolamine, glutamic acid, glycolic acid, glutamate, and long-chain dicarboxylic acids (C16-C18). These groups, variably linked by ester, ether, or amide bonds, impart zwitterionic characteristics to the structures, maybe enabling the fine tuning of the chemical-physical properties of the resulting compounds. The published papers [43,44] do not report any of these long chains, neither clarify the importance of them. Conversely, in both papers the authors stated that this portion is responsible for the PK properties of the compounds, influencing also the OATP inhibition and the mast cell degranulation, particularly when a LYS residue composes the $R^2$ substituent of the macrocycles. Nevertheless, the use of a PEG linker, equipped with a trimethylammonium group substituting the side chain of LYS, solved this PK issue maintaining, at the same time, the low Ki value displayed by the compounds.

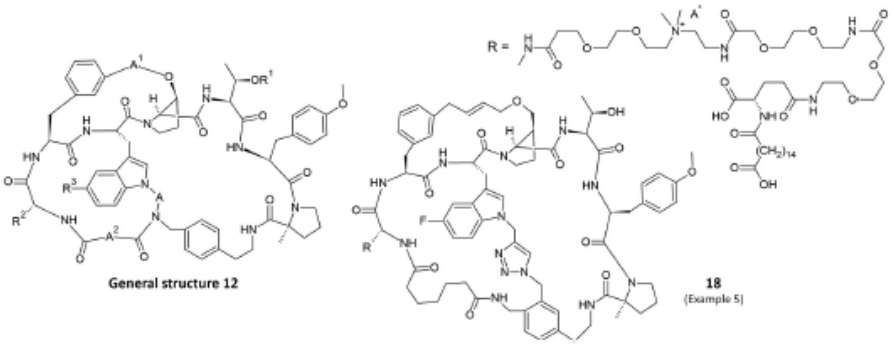



Figure 11. General structure of Merck Sharp & Dohme Corporation compounds, and chemical structure of compound 18 (Example 5), that represents the most potent compound. $A^-$ is any pharmaceutical acceptable counter ion.

Structural information is known about the PCSK9/peptide complex and the binding mode of the compounds on PCSK9. In fact, from the deposited X-ray structures (PDB accession codes 6XIB to -F [43], 7S5G and 7S5H [42]), the macrocycles interact with PCSK9, preventing PPI with LDL-R, in the area shaped by the S153 and the β-hairpin in which the disulfide bridge C375-C378 is located.

The biological activity of the peptides was measured by Alexa TR-FRET experiments. Particularly, biotinylated PCSK9 was used while the peptides were tagged with AlexaFluor647, in the presence of Lance streptavidin europium. Among the 13 compounds reported as examples (described in the Tables 1 and 2 of the patent original document), the most active was compound **18** (patented as example 5), reported in Figure 11, displaying a $K_i$ value of 0.13 nM.

In the same year, Merck Sharp & Dohme Corporation deposited another patent [46]. Comparing the structures covered by both patents, the general structure **13** (Figure 12) seems essentially the same but, in this patent, the $A^1$ and $A^2$ portions of the first patent (Figure 11) are better denoted. In particular, $A^1$ is -$CH_2$-NHCO-$CH_2$- while $A^2$ is -$CHR^1$-NHCO$(CH_2)_2$-. The previous patent already covered carbonyl derivatives as substituents at positions $A^1$ and $A^2$, but in this patent, 68 examples are given, displaying new compounds bearing new substituents A, $R^1$, and $R^2$ for which $K_i$ values spanned from 0.131 to 0.001 nM, as resulted by $K_i$ measurements accomplished by using the same TR-FRET method. The compound showing the lowest $K_i$ value (0.001 nM) is compound **19** (patented as example 18), reported in Figure 12.

As can be seen, compound **19** has a -$(CH_2)_6$- alkyl chain as A substituent, a methyl group as $R^1$, and an $R^2$ chain bearing the dicarboxylic acid of compound 5 of the first patent. The gain in activity could be due to the optimal $A^1$ and $A^2$ substituents of the generic formula covered by the first patent. Furthermore, in the 68 compounds shown as examples (Table 1 of the original patent), the A-spacer of the general formula is always an alkyl chain bearing 5 or 6 carbon atoms, sometimes also unsaturated (i.e. examples 25, 44, 46). Compounds exhibiting these variations continued to show $K_i$ <0.010 nM, an indication that these variations had little influence on the biological activity of the compounds tested. On the other hand, Merck Sharp & Dohme Corporation reported 68 examples bearing huge $R^2$ substituents, do not clearly displaying the structure of MK–0616, a very promising drug candidate entering the clinical phase III [47]. In the structure of MK-0616, the PEG linker as $R^2$ substituent, was replaced by a simpler -$(CH_2)_5$-$N^+(CH_3)_3$, do not present in any of the 68 reported example.

Despite the simplicity of the chain, MK-0616 retains a $K_i$ value of 5 pM, indicating that the $R^2$ is not really critical for the PCSK9 inhibiting activity of the compounds. MK–0616, not only exhibited a strong affinity for PCSK9 in laboratory settings, but it demonstrated poor renal and hepatic clearance, adequate safety, and oral bioavailability during preclinical

Table 1. Structure and inhibition rate values of the most potent compounds reported in the patent.

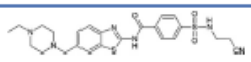
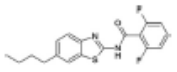
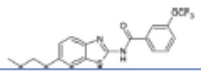
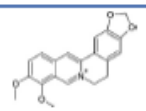
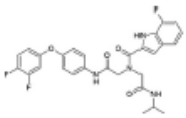
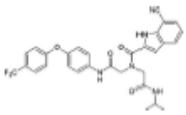
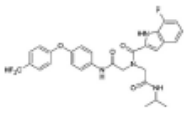
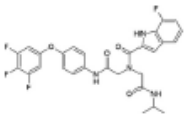
| Compound | Inhibition rate (%) | Structure | IUPAC |
|---|---|---|---|
| **24** (J14) | 90.2 |  | 4-[N(2-ethyl)sulfonyl]-N-[6(4–2-ethylpiperazinylmethyl)benzothiazol-2-yl]benzamide |
| **25** (J20) | 86.8 |  | *N*-(6-butylbenzothiazol-2-yl)-2–6-difluorobenzamide |
| **26** (J22) | 86.0 |  | *N*-(6-butylbenzothiazol-2-yl)-3-(trifluoromethoxy)benzamide |

Table 2. Biological activity and cytotoxicity of the positive control (Berberine) and compounds 28–31, ordered by PCSK9 inhibition. Notes: all values reported in the table below are intended to be percentage of activity at 10 μM of concentration. [1] the value refers to HepG2 ELISA test; [2] the value refers to a HepG2 RT-Q-PCR analysis PCSK9 mRNA analysis; [3] the value refers to cell viability measured in HepG2 cells.

| Name | Structure | PCSK9 Inhibition[1] | Reduction of the expression of PCSK9[2] | Cytotoxicity[3] |
|---|---|---|---|---|
| **Berberine** (positive control) |  | 52 | 35 | n/a |
| **28** (201 – 277) |  | ≃ 100 | 86 | 90 |
| **31** (201 – 284) |  | 94 | 90 | ≃ 100 |
| **30** (201 – 327) |  | 90 | 88 | 98 |
| **29** (201 – 331) |  | 87 | 92 | ≃ 100 |

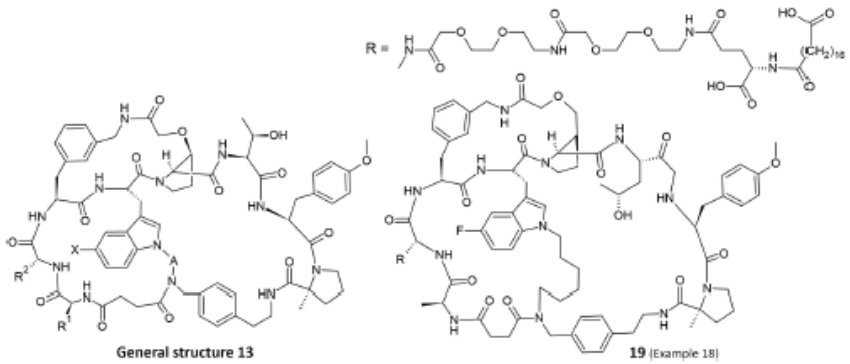



Figure 12. General structure of the second patent of Merck Sharp & Dohme Corporation compounds, deposited in 2021, and chemical structure of compound 19 (Example 18), that represents most potent compound described in the patent.

evaluations, facilitating its progression into clinical trials. MK-0616 suggests that, sometimes, the Lipinsky rule cannot be respected to attain drug candidates. In fact, it represents an example of high molecular weight drug deriving from the optimization of active peptides produced by mRNA display technologies. In Phase 1 clinical trials involving healthy adults, the administration of single oral doses of MK–0616 resulted in a substantial reduction of more than 93% in mean levels of free, unbound plasma PCSK9 (95% CI, 84–103). Among participants concurrently receiving statin therapy, the use of multiple oral doses of 20 mg MK-0616 once daily for 14 days led to a maximum reduction of 61% in mean levels of low-density

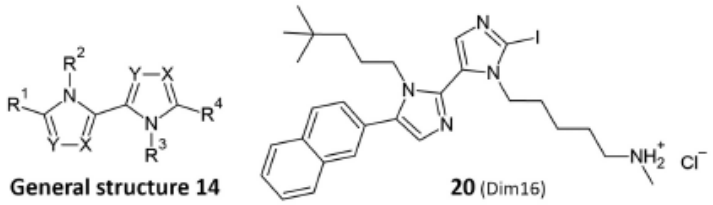



Figure 13. General structure of the diimidazole analogues and the most potent example of the Dim series (Dim16) patented by the University of Milan in 2021.

lipoprotein cholesterol from baseline (95% CI, 43–85) [48]. This work validates the approach of using peptide libraries from mRNA display technologies for the identification of novel oral therapeutic agents, exemplified by the invention of MK-0616.

#### 2.4.2. University of milan

In 2021, the University of Milan patented a disclosure relying on diimidazole analogues with general structure **14**, endowed with high PCSK9 inhibiting activity (Figure 13) [49]. As stated by the authors in the papers preceding the application for the patent [50–54], the source of inspiration of these compounds derived from the presence of a β-sheet shaping the interaction between the EGF-A domain of LDL-R and PCSK9, as can be clearly seen in the X-ray structure deposited by McNutt et al.. (PDB accession code 3GCX) [55]. Stucchi et al. demonstrated that a poli-imidazole chain well resembles the β-strand conformation of a poli-Ala chain [50]. Then, the synthesis and the preliminary biological investigation of a simple N-Methyltetraimidazole (MeIm) validated this hypothesis, since it displayed an $IC_{50}$ value in the micromolar range. The further optimization of the chemical structure of MeIm chain led to the Rim compounds [54] and, finally, to the synthesis of compound **20** (patented as Dim16) reported in Figure 13, displaying an $IC_{50}$ value lower than 1 nM [56]. Further biological investigations of the compounds demonstrated that they can decrease the HepG2 LDL-C uptake and increase the LDLR population on the cell surface. Remarkably, similarly to other natural peptides derived from lupin seeds hydrolysis, Dim series displayed an inhibitory activity on the LDL-C metabolic pathway. In fact, some of the compounds covered by the patent can inhibit HMG-CoAR, the enzyme targeted by the statins. Additionally, compound **20** exerted platelet antiaggregating activity, opening the way to the development of a new therapeutical application of PCSK9 inhibitors [56].

### 2.5. Year 2022

#### 2.5.1. Cardio Therapeutics Pty Ltd

Researchers from Cardio Therapeutics Pty Ltd. described the synthesis and biological evaluation of a series of small molecules, with general structure **15** (Figure 14) [57].

They are constituted by a central aromatic or heteroaromatic six-membered ring, which is decorated with a variety of substitution patterns. One hundred and fifty-four different compounds have been prepared by straightforward synthetic protocols and fully characterized by means of LC/MS(ESI) and $^1$H NMR experiments. All compounds were assayed for their ability to inhibit the binding between PCSK9 and the LDLR using a CircuLex PCSK9-LDLR *in vitro* binding assay kit, displaying $IC_{50}$ values in the high nanomolar range. All compounds were assessed for their oral suitability in Caco-2 assay. Caco-2 cells are a human colon epithelial cancer cell line used as a model of human intestinal absorption of drugs. The permeability coefficient (Papp) denotes the permeability of the drug through a monolayer of cells. *Ex vivo* LDLR surface expression on primary human lymphocytes was evaluated using most promising derivatives, namely example 3f ($IC_{50}$ 939 nM), **compound 21** (patented as example 29, $IC_{50}$ 320 nM, Figure 14), example 122 ($IC_{50}$ 279 nM) and example 131 ($IC_{50}$ 547 nM), in comparison with the positive control alirocumab, a fully humanized monoclonal antibody. These compounds, endowed with high bioavailability in the *in vitro* Caco2 study, were further assessed in a dedicated *in vivo* pharmacokinetic study. Within these tests, various additional

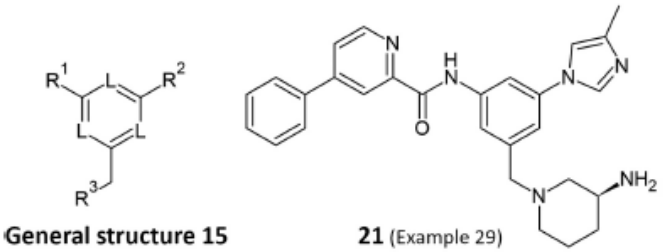
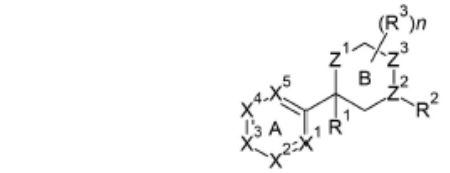
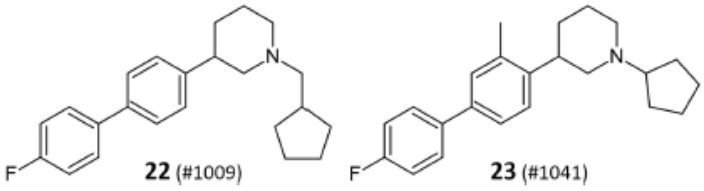



Figure 14. Representation of the general structure of the small molecules patented by Cardio Therapeutics Pty Ltd. and the most promising one, compound 21 (Example 29), showing potent PCSK9-LDLR inhibiting activity and tested in vivo.



General structure 16



22 (#1009) 23 (#1041)

Figure 15. (Top) General structure of the small molecules patented by SRX Cardio LLC in 2022. (Bottom) Structure of the most promising compound 22 (#1009) from biological assays, and structure of the representative compound 23 (#1041) selected to evaluate the oral bioavailability and efficacy in mice.

ADMET properties of compounds of the invention were assessed, such as HepG2 cytotoxicity (HepG2 is a human hepatoma commonly used in drug metabolism and hepatotoxicity studies), hERG assay (the hERG channel inhibition assay is a highly sensitive measurement which identifies compounds exhibiting cardiotoxicity), and CYP assay (the CYP enzyme inhibition assay predicts a drug candidate's predisposition to inhibit cytochrome P450 enzymes). Among them, compound **21** was selected for *in vivo* efficacy study, as monotherapy and in combination with atorvastatin. As monotherapy, oral dosing of compound **21** at 30 mg/kg and 50 mg/kg were able to reduce total plasma cholesterol levels by more than 30% and 50%, compared to the vehicle control, respectively, over a period of four weeks. The results also show an increase in plasma PCSK9 levels over the same time period. The increase in PCSK9 plasma levels indicates that compound **21** is inhibiting the interaction of the protein with LDLR *in vivo* and, as such, PCSK9 cannot be taken up by the liver cells from the blood. Finally, the results from experiments in combination with atorvastatin demonstrate at least an additive effect. All these results confirm the efficacy of compound **21** to modulate PCSK9 levels and to reduce the circulating LDL-C following the oral dosing of the compound.

More information is also available in the research article published on the Journal of Lipid Research,in which the compound **21** (example 29) is reported as NYX-PCSK9i [58].

### 2.5.2. SRX cardio LLC

The patent of SRX Cardio LLC describes a series of small molecules with general structure **16** (Figure 15). They are constituted by a six-membered aromatic or heteroaromatic ring A and a six-membered non-aromatic ring B, linked together and decorated with various substituents. They present a stereocenter and are described as single stereoisomers or mixtures of stereoisomers. Seventy-five molecules are listed in the examples section, but the synthesis and structural characterization, as racemic mixtures, are briefly reported only for the most promising molecules [59].

Biological assays data are reported for selected compounds. Cells, such as HepG2, HuH7, FL83B, or a cell line transfected with a short-hairpin PCSK9 knockdown sequence have been cultured and analyzed for both cell viability marker (dead cells) and LDLR levels in live cells, using a flow cytometer. Cells incubated with small molecule compounds that are inhibitors of PCSK9 are expected to show increased amounts of LDLR, relative to control (no compound) specimens. The percentage recovery in the LDLR assay at 10 µM concentration showed that only molecules #1003, #1006, compound **22** (patented as #1009), and molecule #1035, have demonstrated a LDLR percentage recovery higher than 80%. Fluorescent-LDL uptake analysis by flow cytometric analysis was also carried out. The LDL-uptake assay showed that only compound **22** (Figure 15), and molecules #1010 and #1011, have demonstrated an $EC_{50}$ at concentrations lower than 0.5 µM, representing the most promising compounds.

Finally, a representative compound **23** (patented as **#1041**) was tested for oral bioavailability and efficacy in mice (Figure 15). Following repeated treatment with the compound, a 72.1% reduction in AST (aspartate aminotransferase) levels and a 77.7% reduction in ALT (alanine aminotransferase) levels provide direct experimental evidence that the compound was reversing liver damage induced by a high fat diet and could be useful as a treatment for liver disease or liver dysfunction.

### 2.5.3. Hebei normal university & institute of medical biotechnology CAMS

Hebei Normal University and the Institute of Medical Biotechnology Chinese Academy of Medical Sciences patented a series of benzothiazole compounds with general structures **17 – 19**, or a pharmaceutically acceptable salt thereof, described in Figure 16 [60].

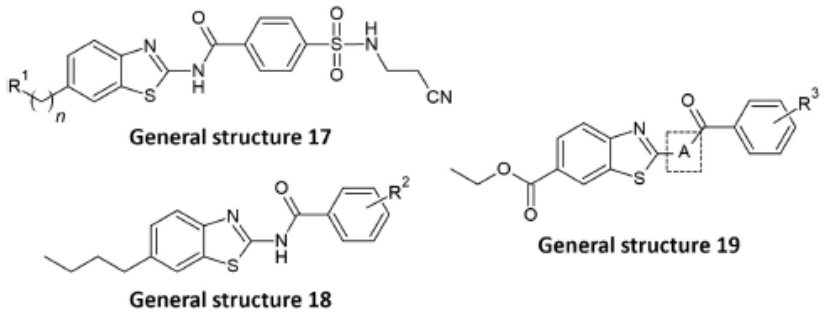



Figure 16. General structure of the benzothiazole patented by Hebei Normal University and Institute of Medical Biotechnology Chinese Academy of Medical Sciences in 2022.

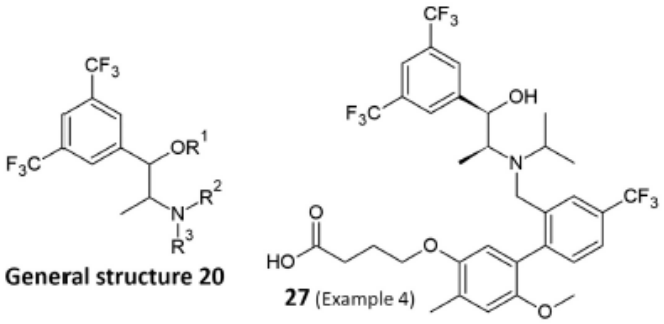



Figure 17. General structure of the new discovered amino alcohol derivatives and the most potent PCSK9-LDLR inhibitor reported in the patent of Daegu Gyeongbuk Medical Innovation Foundation and Kyungpook National University.

All the benzothiazole compounds mentioned above exhibit a significant inhibitory effect on PCSK9, making them suitable for the development of PCSK9 inhibitor drugs. In this experiment, HepG2 cells served as the experimental model and the inhibitory effect of the test compounds on PCSK9 activity was determined by chemiluminescent experiments.

In this patent, the inhibitory activity of 49 compounds against PCSK9 protein was tested. Only the examples J13–23, J32–36, J43–44, and J49, exhibited an inhibition rate equal to or greater than 50%. Among all, the compounds **24–26** (patented as J14, J20, and J22, respectively) demonstrated the most significant inhibitory effect (Table 1).

### *2.5.4. Daegu gyeongbuk medical innovation foundation & kyungpook national university*

The patent deposited by Daegu Gyeongbuk Medical Innovation Foundation and Kyungpook National University relates to the potential use of newly discovered amino alcohol derivatives with general structure **20** (Figure 17) [61]. This patent also comprises the pharmaceutically acceptable salts, hydrates, solvates, tautomeric forms, and stereoisomers of the compounds displayed. From the point of view of the formulation, the embodiments of the invention cover several administration methods: oral (sublingual, buccal), intrathoracic, subcutaneous, intramuscular, and parenteral.

A PCSK9-LDLR binding assay experiment was conducted on a culture of HepG2 cells to assess the inhibitory activity of the amino alcohol derivatives. As a positive control, the well-established PCSK9 inhibitor SBC – 115076 compound was used as described in the patent published by Shifa Biomedical Corporation in 2014 [24]. In particular, the binding inhibition rate of PCSK9 with LDLR was assessed by measuring the relative binding rate. The inhibitor's effect was evaluated by comparing the PCSK9-LDLR binding rate of the treated group to the control group (not treated with the inhibitor), set at 100. Results demonstrated that the reference compound SBC-115076, when treated at a single concentration of 10 μM, effectively inhibited the PCSK9-LDLR binding, showing a 47.39% inhibition rate.

Among all the compounds tested, 27 (patented as example 4), reported in Figure 17, demonstrated a superior PCSK9-LDLR binding inhibition (40.38% inhibition rate) compared to the positive control, indicating its potential as a PCSK9 inhibitor. These findings suggest that the compounds of the present invention have the potential to prevent, improve, or treat hypercholesterolemia by inhibiting the binding of PCSK9-LDLR.

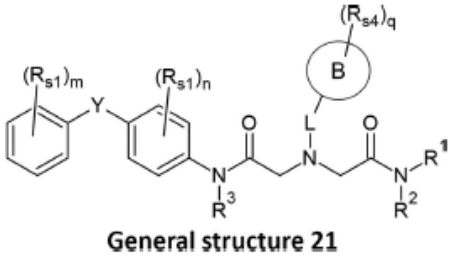



Figure 18. General structure of the novel compounds patented by Shengke Pharmaceuticals Jiangsu LTD in 2023.

## 2.6. Year 2023

### 2.6.1. Shengke Pharmaceuticals Jiangsu LTD

The patent of Shengke Pharmaceuticals Jiangsu LTD describes the synthesis of 28 novel compounds based on a central glycine core, modified with a para–amino biaryl structure and an indole moiety with general structure **21** (Figure 18) [62]. The invention showed promising results for most of the compounds and the most interesting examples are reported in Table 2. Fluorinated biaryl structures are endowed in the most active molecules, where the halogen directly functionalizes the phenyl ring (compounds **28** and **29**, patented as 201 – 277 and 201 – 331, respectively) or through difluoro- (compound **30**, patented as 201 – 327) or trifluoro- (compound **31**, patented as 201 – 284) groups.

The compounds turned out to be potent down-regulators of the activity of PCSK9, producing a blood cholesterol-reducing action. In particular, the biological activity of the compounds was measured through the treatment of HepG2 cells with the respective drug and a reduction of the level of PCSK9 was observed and described as percentage of inhibition (Table 2). Moreover, a RT-Q-PCR analysis was conducted to clarify the mechanism of action of the molecules, strengthening their role in the depletion of the expression of PCSK9 (Table 2). The compounds showed a good level of safety during conventional cell viability assays. In addition, the authors performed a metabolic stability test incubating the compounds with human or rat liver microsomes, indicating that compound **31** (201 – 284), bearing a -CN group on the indole moiety and a para-CF3 substituted phenyl ring, represented the most stable candidate. The isoquinoline alkaloid berberine was used as a positive control in the biological experiments performed (Table 2).

### 2.6.2. China pharmaceutical university

In March 2023, the public China Pharmaceutical University provided a class of structurally novel compounds endowed with PCSK9 inhibitory activity to be used alone or in combination for the prevention and/or treatment of hyperlipidemia and related metabolic diseases [63]. Compounds **32–35** (Figure 19) were discovered *via* virtual screening and their biological activity was verified through *in vitro* protein-protein interaction inhibition experiments, protein binding force test experiments and *in cell* test experiments.

Moreover, the compounds could also be subjected to pharmaceutical formulation such as capsules, powders, tablets, granules, pills, injections, syrups, oral liquids, inhalants, ointments, suppositories or patches and other conventional pharmaceutical preparation forms. Compounds **32–35** demonstrated to inhibit the binding of PCSK9 protein to LDLR on the surface of liver cells by directly binding to PCSK9 protein, thereby inhibiting the degradation of LDLR, increasing the level of LDLR on the surface of liver cells, promoting the uptake and processing of LDL from the blood, and reducing cholesterol levels.

A lipid–lowering effect was observed for compounds **32–35** and can be further used in the development of hypolipidemic drugs also in association with already marketed pharmaceuticals. The experimental results show that compounds **32** (AG-205/07687040, CAS: 339023-59-5), **33** (AN-153/13396332, CAS: 797808-17-4), **34** (AH-487/41802246, CAS: 664316-79-4) and **35** (AN-919/15529007, CAS: 2099167-44-7) can directly inhibit the protein–protein interaction between PCSK9 and LDLR, as shown PCSK9-LDLR binding inhibition in vitro. During the assay, the inhibitory rates of compounds with formula **32–35** at

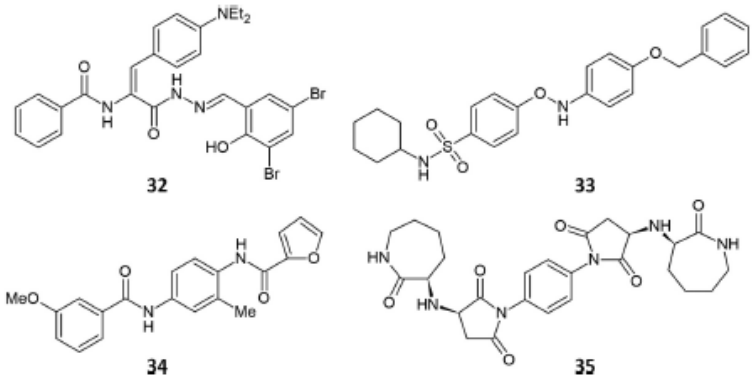



Figure 19. Chemical structures of compounds 32–35 patented by the public China Pharmaceutical University in 2023.

a concentration of 25 μM on PCSK9-LDLR protein-protein interaction were 42.0 ± 6.8%, 82.8 ± 1.7%, 45.9 ± 1.6% and 64.2 ± 0.7%, respectively. Further dilution experiments showed that the inhibitory activity of **35** on PCSK9-LDLR protein-protein interaction was dose–dependent ($IC_{50}$ = 7.57 ± 1.40 μM). A strong binding to PCSK9 was observed for compound **35** using a SPR method, and the stability of the binding was also confirmed by computational studies. Compound **35** was tested for its ability to restore LDL uptake showing optimal results and the treatment of cells with high concentration of compound **35** showed a good inhibition of the expression of LDLR. Such features, combined with low cytotoxicity, indicated that compound **35** possesses a potential cholesterol-lowering activity and lipid-lowering effect, and can be further used to prepare blood-lipid-lowering drugs.

### 2.7. Expert opinion

Genetic studies have revealed that variations in the PCSK9 gene, specifically the R46L polymorphism, present in 2–3% of the population, are linked to lower levels of LDL-C and a protective effect against cardiovascular issues. This finding prompted the development of strategies aimed at inhibiting or eliminating PCSK9 activity. The rationale behind this approach stems from the observation that PCSK9 appears to primarily regulate the population of LDL receptors on liver cells, thereby influencing circulating LDL-C, with no apparent physiological roles beyond this function. The significant therapeutic potential of the PCSK9 inhibition has led numerous entities worldwide, including companies and academic institutions, to patent diverse compounds such as small molecules, natural and non-natural peptides, siRNA, and monoclonal antibodies. Now, only monoclonal antibodies such as alirocumab (Sanofi) or evolocumab (Amgen) and siRNA (Inclisiran, Alnylam and Novartis) reached the market for the treatment of CVD. Nevertheless, the high cost makes these remedies unusable in the underdeveloped countries and for this reason orally bioavailable small molecules never lose interest. For this reason, to date, many companies and universities have continued to search for PCSK9 modulators, as evidenced by the 810 patents filed in the last 5 years.

It was interesting to note that some patented compounds do not act only inhibiting the PCSK9/LDLR interaction. In fact, the compounds patented by us (diimidazoles) are also active in inhibiting HMG-CoAR (as statins), improving their hypocholesterolemic effect. Remarkably, some of the patented structures have started the processes of clinical development, in particular MK-0616 (demonstrating PCSK9 affinity in the pico molar range) has successfully progressed to clinical application and has shown to be effective in reducing cholesterol levels. However, the intense competition in this therapeutic field is causing pharmaceutical companies, such as AstraZeneca or Novo Nordisk, to reconsider projects that seem competitive in cholesterol reduction. This is because these projects may face challenges when integrating into a market that already has a longer-acting alternative, like Novartis' inclisiran. For this reason, other targets implicated in dyslipidemia are being contemplated as potential alternatives. Among them, seems promising decreasing the level of lipoprotein(a) [Lp(a)], an LDL cholesterol variant that contains an apolipoprotein(a) [apo(a)]. Novartis is leading Phase III program for pelacarsen, an antisense oligonucleotide (ASO), interacting with hepatocyte apo(a) mRNA, forming an ASO/mRNA complex. This complex hinders the translation of apo(a), resulting in reduced apo(a) production and consequently, diminished levels of circulating Lp(a) [64]. In the same category of RNA-based drugs belong volanesorsen (Akcea Therapeutics Ireland Limited) and olezarsen (Ionis Pharmaceuticals), ASO directed to apoC-III mRNA, disrupting apoC-III translation. This disruption results in decreased apoC-III levels, leading to reduced chylomicron and triglyceride levels. Phase III clinical trial findings for volanesorsen, published in 2021, revealed a 71.8% reduction in triglyceride levels compared to the placebo group over a three-month period [65]. Unfortunately, volanesorsen creates reactions in the injection site, and thrombocytopenia has been observed. Fortunately, lower side effects were observed by the subcutaneous administration of olezarsen [65].

In conclusion, the identification of PCSK9's involvement in cholesterol homeostasis has illuminated a wonderful player in cholesterol metabolism and in the prevention of coronary heart disease. PCSK9 represents a promising target for reducing LDL-C levels, amplifying the efficacy of other lipid-lowering drugs, particularly statins, and advancing the prevention and treatment of hypercholesterolemia. However, a plethora of academia and pharmaceutical companies are still involved in the development of PCSK9 inhibitors. It is hoped that other compounds can proceed into clinical development soon, in order to obtain powerful and effective drugs, as siRNA and monoclonal antibodies are, but accessible to all patients at low prices.

## Funding

This paper was granted by "Linea2", grant from University of Milan.

## Declaration of interest

University of Milan, G Grazioso, A Silvani, and C Lammi are applicant and inventors, respectively, of patent entitled 'PCSK9 inhibiting compounds' and publication number WO2021234654, filing date 21 May 2021. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

## Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

## Author contribution statement

All authors contributed to gather the patent-related information and to the preparation of the original draft. EMA Fassi created the figures. EMA Fassi and G Grazioso contributed to the manuscript revisions. All authors read and approved the final version of the manuscript.

## References


**Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.**

1. Gu HM, Zhang DW. Hypercholesterolemia, low density lipoprotein receptor and proprotein convertase subtilisin/kexin-type 9. J Biomed Res. 2015;29(5):356–361. doi: 10.7555/JBR.29.20150067
2. Baigent C, Keech A, Kearney P, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. Lancet. 2005;366(9493):1267–1278.
3. Newman CB, Blaha MJ, Boord JB, et al. Lipid management in patients with endocrine disorders: an endocrine society clinical practice guideline. J Clin Endocrinol Metab. 2020;105(12):dgaa674. doi: 10.1210/clinem/dgaa674
4. Reiner Z. Resistance and intolerance to statins. Nutr Metab Cardiovasc Dis. 2014;24(10):1057–1066. doi: 10.1016/j.numecd.2014.05.009
5. Banach M, Rizzo M, Toth PP, et al. Statin intolerance - an attempt at a unified definition. Position paper from an international lipid expert panel. Arch Med Sci. 2015;11(1):1–23. doi: 10.5114/aoms.2015.49807
6. Cromwell WC, Ziajka PE. Development of tachyphylaxis among patients taking HMG CoA reductase inhibitors. Am J Cardiol. 2000;86(10):1123–1127. doi: 10.1016/S0002-9149(00)01171-1
7. Wong J, Quinn CM, Brown AJ. SREBP-2 positively regulates transcription of the cholesterol efflux gene, ABCA1, by generating oxysterol ligands for LXR. Biochem J. 2006;400(3):485–491. doi: 10.1042/BJ20060914
8. Taylor BA, Thompson PD. Statins and their effect on PCSK9–impact and clinical relevance. Curr Atheroscler Rep. 2016;18(8):46. doi: 10.1007/s11883-016-0604-3
9. Ahmad Z, Banerjee P, Hamon S, et al. Inhibition of angiopoietin-like protein 3 with a monoclonal antibody reduces triglycerides in hypertriglyceridemia. Circulation. 2019;140(6):470–486. doi: 10.1161/CIRCULATIONAHA.118.039107
10. Rizzo M. Lomitapide, a microsomal triglyceride transfer protein inhibitor for the treatment of hypercholesterolemia. IDrugs. 2010;13(2):103–111.
11. Brandts J, Ray KK. Bempedoic acid, an inhibitor of ATP citrate lyase for the treatment of hypercholesterolemia: early indications and potential. Expert Opin Investig Drugs. 2020;29(8):763–770. doi: 10.1080/13543784.2020.1778668
12. Mourikis P, Zako S, Dannenberg L, et al. Lipid lowering therapy in cardiovascular disease: from myth to molecular reality. Pharmacol Ther. 2020;213:107592. doi: 10.1016/j.pharmthera.2020.107592
13. Shapiro MD, Tavori H, Fazio S. PCSK9 from basic science discoveries to clinical trials. Circ Res. 2018;122(10):1420–1438. doi: 10.1161/CIRCRESAHA.118.311227
14. Horton JD, Cohen JC, Hobbs HH. Molecular biology of PCSK9: its role in LDL metabolism. Trends Biochem Sci. 2007;32(2):71–77. doi: 10.1016/j.tibs.2006.12.008
15. Peng J, Liu MM, Jin JL, et al. Association of circulating PCSK9 concentration with cardiovascular metabolic markers and outcomes in stable coronary artery disease patients with or without diabetes: a prospective, observational cohort study. Cardiovasc Diabetol. 2020;19(1):167. doi: 10.1186/s12933-020-01142-0
16. Seidah NG. The PCSK9 revolution and the potential of PCSK9-based therapies to reduce LDL-cholesterol. Glob Cardiol Sci Pract. 2017;2017(1):e201702.
    - **Thi s article highlights the great potential of targeting PCSK9 for the treatment of hypercholesterolemia.**
17. Chaplin S. PCSK9 inhibitors: when statins aren't enough. Prescriber. 2017;28(10):21–26. doi: 10.1002/psb.1616
18. Tavori H, Melone M, Rashid SA. Alirocumab: PCSK9 inhibitor for LDL cholesterol reduction. Expert Rev Cardiovasc Ther. 2014;12(10):1137–1144. doi: 10.1586/14779072.2014.954551
19. Sabatine MS, Giugliano RP, Keech AC, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. N Engl J Med. 2017;376(18):1713–1722. doi: 10.1056/NEJMoa1615664
20. German CA, Shapiro MD. Small interfering RNA therapeutic inclisiran: a new approach to targeting PCSK9. BioDrugs. 2020;34(1):1–9. doi: 10.1007/s40259-019-00399-6
21. Abifadel M, Pakradouni J, Collin M, et al. Strategies for proprotein convertase subtilisin kexin 9 modulation: a perspective on recent patents. Expert Opin Ther Pat. 2010;20(11):1547–1571. doi: 10.1517/13543776.2010.518615
    - **This review has to be considered to acquire the complete overview of the PCSK9 inhibitors until 2010.**
22. Elbitar S, El KP, Ghaleb Y, et al. Proprotein convertase subtilisin/kexin 9 (PCSK9) inhibitors and the future of dyslipidemia therapy: an updated patent review (2011-2015). Expert Opin Ther Pat. 2016;26(12):1377–1392. doi: 10.1080/13543776.2016.1206080
    - **This review has to be considered to acquire the complete overview of the PCSK9 inhibitors from 2011 to 2016.**
23. Cardio Therapeutics Pty Ltd. Heterocylic inhibitors of PCSK9. WO2018165718. 2018.
24. Shifa Biomedical Corporation. Anti-proprotein convertase subtilisin kexin type 9 (anti-PCSK9) compounds and methods of using the same in the treatment and/or prevention of cardiovascular diseases. WO2014150326. 2014.
25. Evison BJ, Palmer JT, Lambert G, et al. A small molecule inhibitor of PCSK9 that antagonizes LDL receptor binding via interaction with a cryptic PCSK9 binding groove. Bioorg Med Chem. 2020;28(6):115344. doi: 10.1016/j.bmc.2020.115344
26. Merck Sharp & Dohme Corporation. Substituted 1-methyl-tetrahydroisoquinoline molecules as PCSK9 allosteric binders. WO2018057409. 2018.
27. Petrilli WL, Adam GC, Erdmann RS, et al. From screening to targeted degradation: strategies for the discovery and optimization of small molecule ligands for PCSK9. Cell Chem Biol. 2020;27(1):32–40.e3. doi: 10.1016/j.chembiol.2019.10.002
28. Gustafsen C, Olsen D, Vilstrup J, et al. Heparan sulfate proteoglycans present PCSK9 to the LDL receptor. Nat Commun. 2017;8(1):503. doi: 10.1038/s41467-017-00568-7
29. Aarhus University. Compounds for treatment of lipoprotein metabolism disorders. WO2018054959. 2018.
30. Coombe DR, Kett WC. Heparin mimetics. Handb Exp Pharmacol. 2012;207:361–383.
31. Fennewald SM, Rando RF. Inhibition of high affinity basic fibroblast growth factor binding by oligonucleotides. J Biol Chem. 1995;270(37):21718–21721. doi: 10.1074/jbc.270.37.21718
32. Guvakova MA, Yakubov LA, Vlodavsky I, et al. Phosphorothioate oligodeoxynucleotides bind to basic fibroblast growth factor, inhibit its binding to cell surface receptors, and remove it from low affinity binding sites on extracellular matrix. J Biol Chem. 1995;270(6):2620–2627. doi: 10.1074/jbc.270.6.2620
33. China Pharmaceutical University. Application of kaempferol-3-O-rutinoside to preparation of medicine for treating PCSK9 (Proprotein convertase subtilisin/kexin type 9)-mediated disease. CN108567788. 2018.
34. Li L, Fan ML, Li YN, et al. A rapid strategy for screening high-efficiency PCSK9 inhibitors from ginkgo biloba leaves by ligand fishing, HPLC-Q-TOF-MS and interdisciplinary assay. J Food Drug Anal. 2020;28(2):273–282. doi: 10.38212/2224-6614.1061
35. Merck Sharp & Dohme Corporation. PCSK9 antagonist compounds. JOP20190150. 2019.
36. Merck Sharp & Dohme Corporation. PCSK9 antagonist bicyclo-compounds. WO2019246352. 2019.
37. Novartis AG Cyclic tetramer compounds as proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors for the treatment of metabolic disorders. US2020164024. 2020.
38. Novartis AG Cyclic pentamer compounds as proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors for the treatment of metabolic disorder. WO2020110008. 2020.
39. Brousseau ME, Clairmont KB, Spraggon G, et al. Identification of a PCSK9-LDLR disruptor peptide with in vivo function. Cell Chem Biol. 2022;29(2):249–258.e5. doi: 10.1016/j.chembiol.2021.08.012
40. Dogma Therapeutics Inc. PCSK9 inhibitors and methods of use thereof. WO2020150473. 2020.

41. Pfizer. N-piperidin-3-ylbenziamide derivatives for treating cardiovascular diseases. WO2014170786. 2014.
42. Tucker TJ, Embrey MW, Alleyne C, et al. A series of novel, highly potent, and orally bioavailable next-generation tricyclic peptide PCSK9 inhibitors. J Med Chem. 2021;64(22):16770–16800. doi: 10.1021/acs.jmedchem.1c01599
43. Alleyne C, Amin RP, Bhatt B, et al. Series of novel and highly potent cyclic peptide PCSK9 inhibitors derived from an mRNA display screen and optimized via structure-based design. J Med Chem. 2020;63(22):13796–13824. doi: 10.1021/acs.jmedchem.0c01084
44. Iskandar SE, Bowers AA. mRNA display reaches for the clinic with new PCSK9 inhibitor. ACS Med Chem Lett. 2022;13(9):1379–1383. doi: 10.1021/acsmedchemlett.2c00319
45. Merck Sharp & Dohme Corporation. PCSK9 antagonist compounds. WO2021127460. 2021.
46. Merck Sharp & Dohme Corporation. PCSK9 antagonist compounds. WO2021041770. 2021.
47. Burnett JR, Hooper AJ. MK-0616: an oral PCSK9 inhibitor for hypercholesterolemia treatment. Expert Opin Investig Drugs. 2023;32(10):873–878.
•• **An interesting review article highlighting the most recent results on the clinical development progress of MK-0616, which recently entered the Phase III.**
48. Johns DG, Campeau LC, Banka P, et al. Orally bioavailable macrocyclic peptide that inhibits binding of PCSK9 to the low density lipoprotein receptor. Circulation. 2023;148(2):144–158. doi: 10.1161/CIRCULATIONAHA.122.063372
49. Università degli studi di Milano. PCSK9-inhibiting compounds. WO2021234654. 2021.
50. Stucchi M, Grazioso G, Lammi C, et al. Disrupting the PCSK9/LDLR protein-protein interaction by an imidazole-based minimalist peptidomimetic. Org Biomol Chem. 2016;14(41):9736–9740. doi: 10.1039/C6OB01642A
51. Lammi C, Sgrignani J, Roda G, et al. Inhibition of PCSK9D374Y/LDLR protein-protein interaction by computationally designed T9 lupin peptide. ACS Med Chem Lett. 2019;10(4):425–430. doi: 10.1021/acsmedchemlett.8b00464
52. Lammi C, Sgrignani J, Arnoldi A, et al. Biological characterization of computationally designed analogs of peptide TVFTSWEEYLDWV (Pep2-8) with increased PCSK9 antagonistic activity. Sci Rep. 2019;9(1):2343. doi: 10.1038/s41598-018-35819-0
53. Lammi C, Zanoni C, Aiello G, et al. Lupin peptides modulate the protein-protein interaction of PCSK9 with the low density lipoprotein receptor in HepG2 cells. Sci Rep. 2016;6(1):29931. doi: 10.1038/srep29931
54. Lammi C, Sgrignani J, Arnoldi A, et al. Computationally driven structure optimization, synthesis, and biological evaluation of imidazole-based proprotein convertase subtilisin/kexin 9 (PCSK9) inhibitors. J Med Chem. 2019;62(13):6163–6174. doi: 10.1021/acs.jmedchem.9b00402
55. McNutt MC, Kwon HJ, Chen C, et al. Antagonism of secreted PCSK9 increases low density lipoprotein receptor expression in HepG2 cells. J Biol Chem. 2009;284(16):10561–10570. doi: 10.1074/jbc.M808802200
56. Lammi C, Fassi EMA, Manenti M, et al. Computational design, synthesis, and biological evaluation of diimidazole analogues endowed with dual PCSK9/HMG-CoAR-inhibiting activity. J Med Chem. 2023;66(12):7943–7958. doi: 10.1021/acs.jmedchem.3c00279
57. Cardio Therapeutics Pty Ltd. Heterocyclic inhibitors of PCSK9. WO2022133529. 2022.
58. Suchowerska AK, Stokman G, Palmer JT, et al. A novel, orally bioavailable, small-molecule inhibitor of PCSK9 with significant cholesterol-lowering properties in vivo. J Lipid Res. 2022;63(11):100293. doi: 10.1016/j.jlr.2022.100293
59. SRX Cardio LLC. Compounds for the modulation of proprotein convertase subtilisin/kexin type 9 (PCSK9). US2022193058. 2022.
60. Hebei Normal University & Institute of Medical Biotechnology CAMS. Benzothiazole compound as well as preparation method and application thereof. CN115109011. 2022.
61. Daegu Gyeongbuk Medical Innovation Foundation & Kyungpook National University. Aminoalcohol derivative as PCSK9 inhibitor, and pharmaceutical composition for preventing or treating hypercholesterolemia, containing same. WO2022075645. 2022.
62. Shengke Pharmaceuticals Jiangsu LTD. Novel compounds as inhibitors of PCSK9. WO2023280155. 2023.
63. China Pharmaceutical University. Medical application of compounds with PCSK9 inhibitory activity. CN115770247. 2023.
64. Yeang C, Karwatowska-Prokopczuk E, Su F, et al. Effect of pelacarsen on lipoprotein(a) cholesterol and corrected low-density lipoprotein cholesterol. J Am Coll Cardiol. 2022;79(11):1035–1046. doi: 10.1016/j.jacc.2021.12.032
65. Gouni-Berthold I, Alexander VJ, Yang Q, et al. Efficacy and safety of volanesorsen in patients with multifactorial chylomicronaemia (COMPASS): a multicentre, double-blind, randomised, placebo-controlled, phase 3 trial. Lancet Diabetes Endocrinol. 2021;9(5):264–275. doi: 10.1016/S2213-8587(21)00046-2

**Appendix 2**

**Title**: Computational Design of Novel Cyclic Peptides Endowed with Autophagy-Inhibiting Activity on Cancer Cell Lines

**Authors**: **Albani, M.**, Fassi, E. M. A., Moretti, R. M., Garofalo, M., Montagnani Marelli, M., Roda, G., Sgrignani, J., Cavalli, A., & Grazioso, G.

**Status**: Published in *International Journal of Molecular Sciences(2024)*

**DOI**: 10.3390/ijms25094622

**Author Contribution**: Dr. M. Albani performed molecular dynamics simulations and contributed to writing the original draft of the manuscript

International Journal of
*Molecular Sciences*

MDPI

*Article*

# Computational Design of Novel Cyclic Peptides Endowed with Autophagy-Inhibiting Activity on Cancer Cell Lines

**Marco Albani [1], Enrico Mario Alessandro Fassi [1,*], Roberta Manuela Moretti [2], Mariangela Garofalo [3], Marina Montagnani Marelli [2], Gabriella Roda [1], Jacopo Sgrignani [4], Andrea Cavalli [4,5] and Giovanni Grazioso [1,*]**

1 Department of Pharmaceutical Sciences, Università degli Studi di Milano, Via L. Mangiagalli 25, 20133 Milano, Italy; marco.albani@unimi.it (M.A.); gabriella.roda@unimi.it (G.R.)
2 Department of Pharmacological and Biomolecular Sciences, Università degli Studi di Milano, Via Balzaretti 9, 20133 Milano, Italy; roberta.moretti@unimi.it (R.M.M.); marina.marellimontagnani@unimi.it (M.M.M.)
3 Department of Pharmaceutical and Pharmacological Sciences, Università di Padova, Via F. Marzolo 5, 35131 Padova, Italy; mariangela.garofalo@unipd.it
4 Institute for Research in Biomedicine (IRB), Via Chiesa 5, 6500 Bellinzona, Switzerland; jacopo.sgrignani@irb.usi.ch (J.S.); andrea.cavalli@irb.usi.ch (A.C.)
5 Swiss Institute of Bioinformatics (SIB), University of Lausanne, Quartier UNIL-Sorge, Bâtiment Amphipôle, 1015 Lausanne, Switzerland
* Correspondence: enrico.fassi@unimi.it (E.M.A.F.); giovanni.grazioso@unimi.it (G.G.)

**Abstract:** (1) Autophagy plays a significant role in development and cell proliferation. This process is mainly accomplished by the LC3 protein, which, after maturation, builds the nascent autophagosomes. The inhibition of LC3 maturation results in the interference of autophagy activation. (2) In this study, starting from the structure of a known LC3B binder (LIR2-RavZ peptide), we identified new LC3B ligands by applying an in silico drug design strategy. The most promising peptides were synthesized, biophysically assayed, and biologically evaluated to ascertain their potential antiproliferative activity on five humans cell lines. (3) A cyclic peptide (named Pep6), endowed with high conformational stability (due to the presence of a disulfide bridge), displayed a $K_d$ value on LC3B in the nanomolar range. Assays accomplished on PC3, MCF-7, and A549 cancer cell lines proved that Pep6 exhibited cytotoxic effects comparable to those of the peptide LIR2-RavZ, a reference LC3B ligand. Furthermore, it was ineffective on both normal prostatic epithelium PNT2 and autophagy-defective prostate cancer DU145 cells. (4) Pep6 can be considered a new autophagy inhibitor that can be employed as a pharmacological tool or even as a template for the rational design of new small molecules endowed with autophagy inhibitory activity.

**Keywords:** peptide; LC3B; autophagy inhibitors; cancer; Atg8; LIR motif



**Citation:** Albani, M.; Fassi, E.M.A.; Moretti, R.M.; Garofalo, M.; Montagnani Marelli, M.; Roda, G.; Sgrignani, J.; Cavalli, A.; Grazioso, G. Computational Design of Novel Cyclic Peptides Endowed with Autophagy-Inhibiting Activity on Cancer Cell Lines. *Int. J. Mol. Sci.* **2024**, *25*, 4622. https://doi.org/10.3390/ijms25094622

Academic Editors: Artur M. S. Silva and Joana L. C. Sousa

Received: 26 March 2024
Revised: 17 April 2024
Accepted: 21 April 2024
Published: 24 April 2024



## 1. Introduction

In living organisms, autophagy is a highly organized process that selectively captures proteins and old or damaged organelles using double-membrane vesicles called autophagosomes. When autophagosomes fuse with lysosomes, the contents within them are degraded by the acidic environment and lytic enzymes present in the lysosome [1,2]. The recycling ability of autophagy machinery is present in bacteria as well as in eukaryotic cells, allowing for the conservation of physiological conditions [2,3]. The autophagy machinery involves more than 50 proteins known as Atgs, but the ones responsible for forming the autophagosome and facilitating cellular trafficking are members of the Atg8 family. In mammals, Atg8 proteins (mAtg8) are categorized into two subfamilies: GABA-A receptor-associated protein (GABARAP) and microtubule-associated protein 1 light chain 3 (MAP1LC3), also known as LC3. The GABARAP subfamily consists of GABARAP, GARAPL1, and GABARAPL2, while the LC3 subfamily includes LC3A (with its two splicing variants LC3Aα and LC3Aβ, LC3B, LC3B2, and LC3C) [4]. Proteins within the same

subfamily exhibit significant sequence similarities and perform similar functions within the cell. The GABARAP subfamily plays a crucial role in autophagosome closure and the recruitment of autophagy participants, while LC3 proteins primarily participate in the process of cargo recruitment [2]. Interference with autophagy by compounds capable of interfering with Atg8 proteins has still not been completely clarified, although the use of peptides or peptidomimetics capable of inhibiting Atg3–Atg8 interaction in Plasmodium falciparum has the potential to fight malaria [5,6]. In physiological conditions, autophagy degrades unfavorable components such as damaged organelles, pathogens, and oxidized biomolecules (proteins, DNA, and lipids) in response to oxidative stress, preventing cell damage. Conversely, in mammals, it has been demonstrated that dysregulations in the complex autophagy machinery are associated with various diseases, including neurodegenerative disorders [7], cardiomyopathies [8], infectious diseases [6,9], type II diabetes mellitus [10,11], hepatic steatosis [12], and cancer [13–15]. These pathologies are due to the autophagy corruption triggered by different stimuli originating from internal or external environmental factors. The role of autophagy in cancer is complex and depends on the phase and context of disease progression. It can play a pro-survival role, reducing cell death and promoting resistance to cytotoxic therapies, or it can be associated with cell death [16]. Numerous compounds have been identified to either stimulate or suppress autophagy, with the aim of obtaining therapeutic effects. These regulators serve as valuable research tools for investigating the intricate machinery of autophagy at a molecular level. Additionally, there is potential for their future development into promising drug candidates, with the aim of addressing cancer and other associated medical conditions [17–20]. The known autophagy regulators are active on biochemical pathways in which mTOR, class III PI3K (hVps34), Akt, V-ATPase, L-type $Ca^{++}$ channel, Calpain, proteasome, tyrosine kinases, histone deacetylase, and some others are involved [17–20]. In our previous paper [18], we computationally designed two peptides (WC8 and WC10) endowed with high predicted and measured affinity on GABARAP, one of the subfamilies of the mAtg8 proteins. Interestingly, the treatments of human metastatic castration-resistant prostate cancer (CRPC) cells PC3 with WC8 and WC10 (from 1 to 10 μM) proved the high therapeutic potential of autophagy inhibition since the peptides were more active than paclitaxel, a common anticancer drug [21]. Nevertheless, LC3B is the most extensively studied Atg8 protein in humans due to its clear associations with cancer. In fact, a correlation between LC3B expression and higher tumor grade in clear cell renal carcinoma and other cancers has been established [22–25]. Moreover, studies on ovarian cancer cell lines revealed that the direct targeting of LC3B, to inhibit autophagy and promote apoptosis, enhances the sensitivity of cancer cells to chemotherapy [26]. Similarly, Quan et al. demonstrated that combination therapy with autophagy inhibitors and enzalutamide (a known antiandrogen for the treatment of prostate cancer) effectively induced bladder cancer apoptosis in vitro and in vivo [27]. Intriguingly, LC3 is initially expressed as pro-LC3 and undergoes cleavage by the cysteine protease Atg4B to form its cytosolic isoform LC3-I. Upon initiation of autophagy, LC3-I is bound to phosphatidylethanolamine (PE), becoming LC3-II, and localizes within the lipid membrane of developing autophagosomes [15]. Atg4B, together with other proteins involved in the autophagy machinery [28], and capable of interacting with LC3, possess a specific amino acidic sequence known as the "LC3 interacting region" (LIR), a small protein sequence containing four conserved residues. These 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 acid (often acidic or hydrophobic residues), and $X_3$ is a large hydrophobic residue like Leu, Val, or Ile [29]. For this reason, the LIR domain represents a wonderful starting point for designing ligands capable of interacting with the LC3 subfamily by applying the structure-based drug design approach. Among the proteins bearing the LIR domain, herein, we focus our attention on the proteome of Legionella pneumophila, an intracellular pathogen that produces a protein called RavZ. This has the peculiarity of triggering a decrease in the autophagy level of cells infected by *L. pneumophila*, limiting their ability to fight the infection through bacterial internalization in autophagosomes [30]. In this paper,

we adopted a computational protocol that was successfully applied to select biologically active peptides [21,31,32]; we designed new peptides capable of binding LC3B, using as a template the structure of the bacterial protein RavZ. The newly designed RavZ analogs, also containing non-natural amino acids as well as conformational rigidification, were simulated through molecular dynamics (MD) simulations. Finally, the most promising peptides were synthesized, tested through biophysical experiments on recombinant LC3B protein, and in vitro assayed to demonstrate their biological effects on different cancer cell lines.

## 2. Results and Discussion

*Computational design of LIR2-RavZ analogs.* RavZ is an *L. pneumophila* protein capable of deconjugating LC3 proteins coupled to PE on autophagosomal membranes [33]. This is thanks to the presence of three LIR domains in the sequence. Among them, LIR2 (residues 27–32) seems to be the one responsible for the initial recognition of LC3, with it playing a crucial role in RavZ activity [33]. In fact, testing the binding of the LIR2-RavZ peptide (sequence DIDEFDLLEGDE, Table 1) on LC3B using isothermal titration calorimetry (ITC) and fluorescence polarization (FP), the measured dissociation constants ($K_d$) were 360 nM and 550 nM, respectively, values comparable to that of the full RavZ in complex with LC3B ($K_d$ = 260 nM) [33]. These data unambiguously demonstrate that the interaction between RavZ and LC3B can be mainly attributed to the LIR2 domain of the RavZ peptide (LIR2-RavZ, Table 1) [33]. Since structural data are not available in the protein data bank, in this study, we predicted the binding mode of LIR2-RavZ on LC3B by performing docking calculations and MD simulations on the resulting complex (see the Materials and Methods section for details). The visual inspection of the attained MD trajectory frames, together with the analysis of the ligand atoms' root mean square fluctuation (RMSF) plot (see Figure 1A), suggested the high conformational mobility of the peptide *N*- and *C*-terms (mean RMSF value of 2.7 Å). Conversely, the residues located in the core of the peptide (residues 3–10) were almost stable (Figure 1A). The low fluctuation in the core was due to the presence of two salt bridges shaped by the side chains of LIR2-RavZ-Glu4 and -Asp6 with the ones of LC3B-Lys65 and -Arg69, respectively. Moreover, the hydrophobic tails of LIR2-RavZ-Phe5 and -Leu7 were inserted into a lipophilic pocket shaped by LC3B-Phe52, -Val54, -Pro55, -Val58, -Leu63 and -Ile66. Furthermore, the side chains of LIR2-RavZ-Glu9 and -Asp11 could create a H-bond network with LC3B-Lys30 and LC3B-Lys49, -Thr50, -Lys51 and -Thr50, respectively (Figure 1B).

**Table 1.** Primary structure and estimated binding free energy values (ΔG*) of the reference peptide LIR2-RavZ and its structural analogs.

| Peptide | Sequence | ΔG* [kcal/mol] | SD |
|---|---|---|---|
| LIR2-RavZ | DIDEFDLLEGDE | −86.7 | 13.1 |
| Pep1 | DIDEFDLLEC$\overline{DC}$ | −87.5 | 8.7 |
| Pep2 | Ac-DIDEFDLLEC$\overline{DC}$-$NH_2$ | −109.1 | 7.2 |

Considering these data, we tried to design new LIR2-RavZ analogs with the aim of selecting new peptides endowed with increased affinity on LC3B, possibly by using unnatural amino acids to improve the metabolic resistance of the resulting peptides. In this process, three different strategies were applied:

1. Rigidification: Designing cyclic peptides maintaining the original length of the peptide (12 residues) and inserting cysteines to create the disulfide bridge (replacing two residues, identified as having minor significance in the interaction with LC3B through alanine scanning);
2. Terms protection through the amidation and acetylation of the *C*- and *N*- terms, respectively, to prevent peptide self-cyclization;

3. Affinity maturation: Optimization of the sequence to attain new peptides with improved affinity on LC3B.

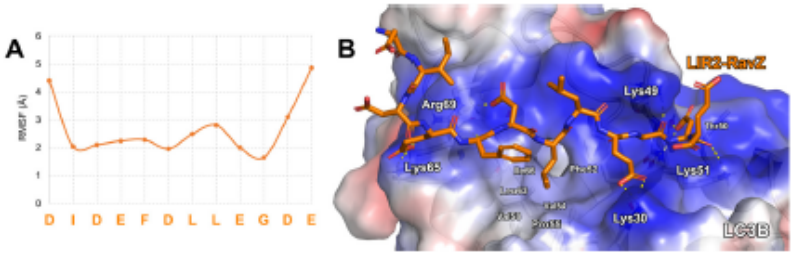



**Figure 1.** (**A**) LIR2-RavZ Cα atom RMSF plot (orange line). (**B**) Predicted binding mode of LIR2-RavZ (orange sticks) in complex with LC3B resulting at the end of MD simulations. The protein surface is colored depending on the atomic partial charges of the protein residues: blue for positive and red for negative charges, respectively. The H-bonds are represented as yellow dotted lines.

By applying these strategies, according to the "Materials and Methods" section, we initially calculated the binding free energy value (ΔG*) of the reference peptide LIR2-RavZ, attaining a value of −86.7 kcal/mol (SD = 13.1) (Table 1).

Then, with the aim of establishing which LIR2-RavZ residues were mainly involved in the interaction with LC3B, and recognizing the hot and non-hot spots, computational alanine scanning was carried out (Supplementary Materials, Table S1). From these calculations, it could be seen that all residues seemed critical for interaction with the target since their substitution by alanine led to an increase in the estimated ΔG* values. Thus, in the first attempt to design new potent LIR2-RavZ analogs, we tried to limit the conformational flexibility of the peptide by creating a disulfide bridge between the residues in positions 10 and 12. Position 12 was chosen because the Ala substitution of the residues in position 12 led to peptides with estimated ΔAffinity not greatly diverse from the parent peptide, while a residue without a side chain (a Gly) is naturally present in the sequence. The resulting peptide, Pep1, was simulated in complex with LC3B through MD simulations and the further application of the molecular mechanics-generalized Born surface area (MM-GBSA) approach suggested that it possessed a ΔG* value of −87.5 kcal/mol (SD = 8.7) (Table 1). This value, effectively equal to the one calculated for the natural peptide LIR2-RavZ, suggested that the presence of conformational rigidification at the C-term did not elicit any change in the LC3B/peptide interaction strength. With variance, the Cα atoms' RMSF, evaluated by MD simulations, demonstrated that the cyclic peptide Pep1 fluctuated around a value lower than the one displayed in the MD simulations on the parent peptide, suggesting that the conformational rigidification stabilizes the new peptide on the LC3B surface (Supplementary Materials, Figure S1). In the second attempt, the peptide termini were protected by amidation and acetylation, with the aim of reducing the *N*-/*C*- reciprocal interaction and the consequent creation of a cyclic peptide in solution. Thus, performing MD simulations and MM-GBSA calculations again on the termini-protected peptide (Pep2), the predicted ΔG* value of Pep2 decreased to the value of −109.1 kcal/mol (SD = 7.2), a value almost 21 kcal/mol lower than the unprotected peptide (Table 1). Subsequently, the "affinity maturation" procedure was applied to Pep2, with the aim of designing new Pep2 analogs with primary structure endowed with improved complementarity with LC3B. This approach allowed for the identification of new peptides through the substitution of the non-hot spot residues with new ones displaying higher affinity to the LC3B binding site. Specifically, to identify the non-hot spots, a new alanine scanning calculation was accomplished on Pep2, and, at the end of calculations, the attained results suggested that the non-hot spot residues were the ones in positions 1, 2, 3, and 6 (Supplementary Materials, Table S2). Consequently, in the next affinity maturation calculation, Pep2 was randomly

mutated by new amino acids, creating a final library of 100,000 peptides, containing L- and D-, side chain-protected, phosphorylated, and unnatural amino acids. The newly designed peptides were ranked considering the predicted affinity on the target and the resulting three top-ranked peptides (Pep3–5) were simulated in complex with LC3B through MD simulations, with their ΔG* values finally being calculated via the MM-GBSA approach as well. Interestingly, the estimated ΔG* values of these peptides were almost 30 kcal/mol lower compared to the parent peptide Pep2, confirming the strength of the applied affinity maturation procedure (Table 2).

**Table 2.** Primary structure and estimated binding free energy values (ΔG*) of the Pep2 analogs at the end of the affinity maturation procedure.

| Peptide | Sequence | ΔG* [kcal/mol] | SD |
|---|---|---|---|
| Pep3 | Ac-DE(Bz)T(Bz)EFT(P)Y(P)LEC$\overline{\text{DC}}$-NH$_2$ | −149.8 | 12.1 |
| Pep4 | Ac-H(Bz)E(Bz)T(Bz)EFT(P)LLEC$\overline{\text{DC}}$-NH$_2$ | −146.5 | 9.5 |
| Pep5 | Ac-H(Bz)E(Bz)E(Bz)EFE(Bz)LLEC$\overline{\text{DC}}$-NH$_2$ | −138.2 | 11.6 |
| Pep6 | Ac-DE(Bz)t(Bz)Eft(P)Y(P)LEc$\overline{\text{Dc}}$-NH$_2$ | −127.1 | 10.8 |

Bn = benzyl group; P = phosphorylation on the side chain. Lowercase letters are D—amino acids.

Among them, Pep3 was the most promising one since it showed a ΔG* value of −149.8 kcal/mol. The sequence of this peptide contains two phosphorylated residues, in positions 6 and 7, together with two benzyl (Bz)-protecting groups on the side chains of Glu2 and Thr3. The Cα RMSF plot retrieved from MD simulations suggested that Pep3 was firmly bound on the LC3B surface, displaying a mean RMSF value of 0.9 Å (Figure 2A). Interestingly, the *N*- and *C*-terms displayed low fluctuation, suggesting that they could create anchoring contacts with LC3B. Indeed, the acetyl group protecting the Pep3 N-terminus formed a water-mediated hydrogen bond with LC3B-Asn59, while the side chain of the same residue established a salt bridge with the one of LC3B-Lys65. The Bz group attached to the side chain of Pep3-Glu2 reached LC3B-His57, resulting in a π-π interaction. Pep3-Glu4′s side chain formed a salt bridge with LC3B-Arg69, while Pep3-Phe5′s side chain was inserted into the hydrophobic pocket defined by LC3B-Ile35, -Val54, -Pro55, -Leu63, and -Ile66. Thr(P) in position 5 of Pep3 was positioned in proximity to the positively charged area formed by the side chain of LC3B-Lys30, whereas the Tyr(P) in position 7 interacted with the alkaline moieties of LC3B-Lys49, -Thr50, and -Lys51. The latter LC3B residues established a salt bridge with the side chain of Pep3-Glu9 while the disulfide bridge (Cys10-Cys12) of Pep3 was inserted into the crevice sized by LC3B-Leu53, -Phe108, -Ile34, and -Pro32. The amidated C-terminus of Pep3-Cys12 created an H-bond network with LC3B-Asp19 and -Lys51. All of these contacts (Figure 2B) strongly stabilize Pep3 on the LC3B LIR binding domain, as displayed by the RMSF plot (Figure 2A).

*Biological activity of Pep3.* Pep3, which proved to be the most promising peptide among the ones investigated, was acquired by Pepmic (Pepmic Co., Ltd., Suzhou, Jiangsu, RPC), and it was assayed using preliminary cell viability assays on the PC3 prostatic cancer cell line (Figure 3). PC3 cells were treated with a range of doses from 0.0025 μM to 5 μM for 24, 48, and 72 h, and viability assays were performed. The results obtained highlight that Pep3 exerted a significant cytotoxic effect at both 2.5 and 5 μM doses after 24 h of treatment, while at 48 h, only 5 μM was effective. Furthermore, the treatment for 72 h did not modify cell viability compared to the control cells, resulting in a lack of efficacy (Figure 3A). We then analyzed the effect of Pep3 on LC3 expression. The analysis of the ratio of LC3-II/LC3-I was not different at any time considered, but in the treated samples, there was a reduction in both LC3-I and LC3-II expression compared to the control sample after 24 h and to a lesser extent after 48 h. It is possible to speculate that in the presence of basal physiological autophagy activation, Pep3 binds on the LC3 precursor, preventing the formation of both LC3-I and lipidated LC3-II (Figure 3B). The same effect on LC3-I

and LC3-II expression was obtained with the treatment of PC3 cells with the autophagy inhibitor 3-metyladenine (3-MA, 1 mM) (Sigma-Aldrich), which blocks the formation of phagophore, inhibiting the phosphoinositide 3-kinases (PI3K) [34]. It is possible that the effectiveness of Pep3 decreased in the considered time from 24 h to 72 h, possibly due to the peptide's low metabolic stability.

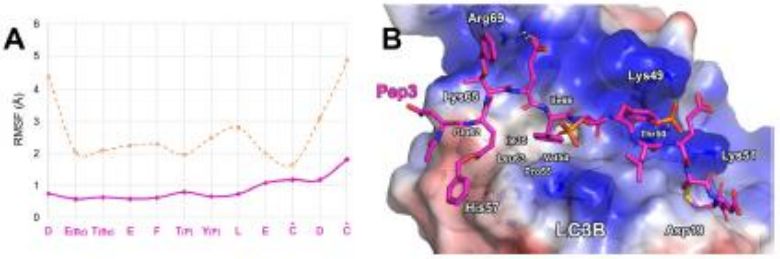



**Figure 2.** (**A**) Pep3 Cα atom RMSF plot (pink line) compared to LIR2-RavZ (orange line). Asterisks indicate the residues involved in the disulfide bond. (**B**) Predicted binding mode of Pep3 (pink sticks) in complex with LC3B resulting at the end of MD simulations. The protein surface is colored depending on the atomic partial charges of the protein residues: blue for positive and red for negative charges, respectively. The H-bonds are represented as yellow dotted lines.

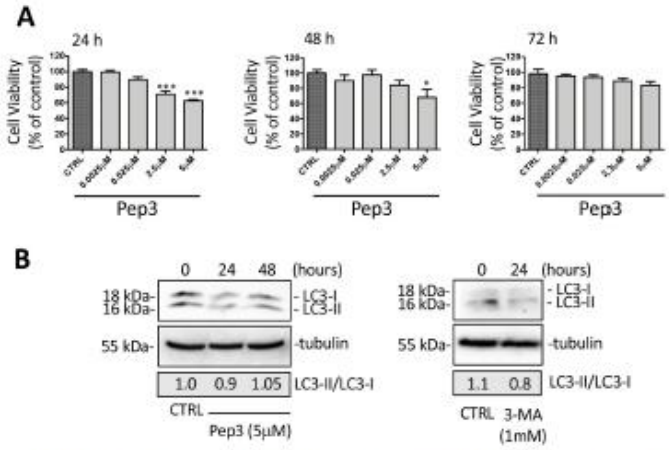



**Figure 3.** (**A**) Effect of Pep3 on PC3 cell viability. Cell viability was determined using MTT assay after 24 h, 48 h, and 72 h. Six independent biological samples for each condition were analyzed ($n$ = 6). Statistical analysis was performed using one-way ANOVA followed by Dunnett's test (* $p < 0.05$ vs. CTRL; *** $p < 0.001$ vs. CTRL). (**B**) Western blot analysis of the LC3-II/LC3-I ratio in the PC3 cells treated with Pep3 (5 μM) or 3-methyladenine (3-MA) (1 mM).

Consequently, we surmised that a new peptide containing D-amino acids in its sequence (Pep6, Table 2) could be more resistant to peptidase activity. Therefore, we used the peptidecutter web service "https://web.expasy.org/peptide_cutter (accessed on 8 January 2024)" to predict the Pep3 cleavage sites. Here, by inserting the Pep3 sequence (without any side chain derivatization) and selecting the prediction of cleavage sites cleaved using proteases (excluding the bacterial ones and chemical reagents), it appeared that the residues in the middle positions could be the most susceptible. Consequently, we designed a Pep3 analog (Pep6) in which the residues at positions 3, 5, 6, 10, and 11 were

mutated by D-amino acids. The predicted ΔG* value of the resulting peptide (Pep6) was −127.1 kcal/mol (SD = 10.8), a value about 23 kcal/mol higher than the one of Pep3 but still about 40 kcal/mol lower than the ΔG* value calculated for the template peptide LIR2-RavZ. The MD simulations on the LC3B/Pep6 complex suggested that the overall conformational stability of the peptide on the LC3B remained essentially like that of Pep3 (0.9 Å, Figure 2A), since the Pep6 mean Cα RMSF value was 0.7 Å (Figure 4A). Moreover, the "simulation interaction analysis" tool, together with the visual inspection of the MD trajectory frames, suggested that the alkaline side chains of LC3B-Lys65, -Arg69, -Arg70, and -Lys49 could create H-bond assisted salt bridges with Pep6-Asp1, -Glu4, -Tyr(P)7, -Glu9, and -Asp11 (Figure 4B). Moreover, Pep6-Asp1 established H-bond interactions with LC3B-Ser61 and -Glu62. Both the amide and carbonyl backbone atoms of Pep6-Leu8 create an H-bond network with the backbone atoms of LC3B-Leu53, while its lipophilic chain was in contact with an LC3B pocket shaped by His27 and Leu53. The disulfide bridge remained in contact with LC3B-Leu53, -Phe108, -Ile34, and -Pro32. The side chain of Pep6-D-Tyr(P) created a salt bridge with LC3B-Lys30, and the phenyl ring of Pep6-D-Phe5 was in contact with the imidazole ring of LC3B-His57, creating a π-π interaction (Figure 4B). Finally, the amidated C-terminus of Pep3-D-Cys12 created an H-bond network with LC3B-Asp19 and -Gln26. For a comparison with a known LC3 binder, in 2021, Fan et al. reported on covalent ligands capable of reducing autophagic structure formation and subsequent substrate degradation [35]. Superimposing our LC3B/Pep6 complex with LC3B covalently bound with compound a4 at Lys49 (PDB accession code 7ELG), we noted that the small molecule is aligned on the Pep6 residues in position 7–9, amino acids T(P)Y(P)L (Supplementary Materials, Figure S2). Interestingly, both Lys49 and Lys51 are located in the LIR docking site, and their occupation by the presence of any ligand impairs the protein–protein interaction between LC3B and its biological partners involved in autophagy. Thus, we can speculate that the presence of our peptide in the LIR docking site could avoid pro-LC3B maturation through prodomain removal catalyzed by Atg4. Additionally, the presence of the peptide in proximity to Lys49 and Lys51 could prevent endogenous acetylation and deacetylation modifications in cells [36]. In fact, it has been reported that the acetylation of these residues blocks the interactions of LC3B with p62 and Atg7 proteins, leading to the accumulation of LC3B-I and autophagy inhibition [37,38]. Nevertheless, additional biochemical experiments are needed to better understand the action mechanism of our peptides.

Pep6 was then acquired by Pepmic (Pepmic Co., Ltd., Suzhou, China), and it was tested using cell viability assays on a non-cancerous cell line (PNT2, to evaluate the cytotoxicity) and PC3, DU145 CRPC cells, A549, and MCF-7 cancer cell lines. These cell lines were chosen since it has been demonstrated that their growth is influenced by autophagy inhibition [23–25].

*Biological activity of LIR2-RavZ and Pep6 on cancer cell lines.* The biological activity of LIR2-RavZ and Pep6 (from 0.0025 to 5 μM) was evaluated with MTT cell viability assay on PC3 cells (Figure 5A). Treatment for 72 h with both compounds determines a reduction in cell viability with greater efficacy of Pep6 compared to LIR2-RavZ at all doses used (Figure 5A). The effects of both compounds were then evaluated on the CRPC cell line DU145, which is defective in autophagy due to its lack of functional ATG5 [39]. Treatment with LIR2-RavZ and Pep6 (5 μM for 72 h) in DU145 cells was ineffective, confirming that the antitumoral activity of the peptides was attributable to autophagy activation (Figure 5B). We then analyzed the expression of LC3 and SQSTM1/p62 (p62) in PC3 cells treated with both compounds for 48 h. In analogy to results obtained with Pep3, both LIR2-RavZ and Pep6 significantly reduced the expression of LC3-I and LC3-II without changing the LC3-II/LC-I ratio. Furthermore, p62 expression was significantly increased after treatment with both compounds (Figure 5C). We also evaluated the ability of Pep6 to interfere with trehalose-dependent autophagy activation. PC3 cells were co-treated with Pep6 (5 μM dose for 48 h) and trehalose (100 mM for 48 h) (Sigma-Aldrich). The Western blot (WB) shown in Figure 5D highlights that Pep6 in association with trehalose was able to decrease the LC3-II/LC3- ratio compared to trehalose alone. These data suggest that Pep6 inhibits trehalose-induced autophagy, causing less autophagosome formation. We can conclude

that the antitumoral effect of Pep6 in PC3 cells is more stable over time than Pep3, and this antitumoral activity is related to autophagy inhibition.

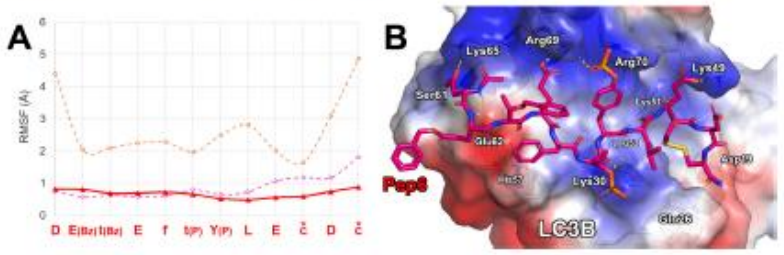



**Figure 4.** (**A**) Pep6 Cα atom RMSF plot (red line) compared to LIR2-RavZ (orange line) and Pep3 (pink line). Asterisks indicate the residues involved in the disulfide bond. The D-amino acids of the Pep6 sequence are reported as lowercase letters. (**B**) Predicted binding mode of Pep6 (magenta sticks) in complex with LC3B resulting at the end of MD simulations. The protein surface is colored depending on the atomic partial charges of the protein residues: blue for positive and red for negative charges, respectively. The H-bonds are represented as yellow dotted lines.

The biological activity of different concentrations of LIR2-RavZ and Pep6 (from 0.0025 to 5 µM) was evaluated with an MTS cell viability assay on non-cancerous PNT2 prostate cells and different cancer cell lines such as A549 and MCF-7 (Figure 6). The results reported in Figure 6A show that 96 h post-treatment, none of the tested samples displayed significant cytotoxicity at the concentration of 5 mM (cell availability > 90%), confirming the excellent biocompatibility and potential pharmacological selectivity for tumor cells. Indeed, as shown in Figure 6B,C, a reduction in cell viability (expressed as percent-age % of viable cells) was observed in A549 and MCF-7 cells treated with LIR2-RavZ and Pep6 compared to the untreated control cells. These in vitro data demonstrate that the compounds exhibited slight anticancer activity, with a reduction in cell viability especially at the highest tested concentration (Figures 5A and 6B,C). More in depth, the obtained results indicated that A549 cells are more responsive than MCF-7 cells. In fact, the cell viability for A549 was 81.1% and 76.7% for LIR2-RavZ and Pep6, respectively, whereas for MCF-7 cells, the viability was 85.4% and 80.7% for LIR2-RavZ and Pep6, respectively. Finally, it is possible to speculate that, since these peptides did not elicit any toxicity on PNT2 cells, they could have an interesting application for the inhibition of pro-survival autophagy induced by chemotherapy or anti-tumor drugs, reducing the onset of drug resistance [26].

*Biophysical assays*. To finally confirm that Pep6 can really affect autophagy machinery by binding on LC3B, MST experiments were performed on human recombinant His-tagged LC3B protein [21]. This biophysical technique quantifies the interactions between two entities in contact in the liquid phase, avoiding any sample immobilization, as needed in other approaches, such as in the case of surface plasmon resonance (SPR) [40]. Initially, the dissociation constant ($K_d$) value of the LIR2-RavZ peptide (as a positive control) was estimated through MST experiments in order to validate the applied biophysical method, leading to a $K_d$ value of 428 ± 162 nM (Figure 7A). This value resulted in being comparable with the one reported by Yang A. and co-workers, where LIR2-RavZ was tested through ITC and FP, displaying dissociation constants of 360 nM and 550 nM, respectively [33]. Finally, we evaluated the binding of Pep6 to the LC3B protein using MST, obtaining a $K_d$ value of 159 ± 56 nM (Figure 7B), which is a value about three times lower compared to the one acquired for LIR2-RavZ. The experimental conditions of the MST experiments carried out are detailed in the Material and Methods section.

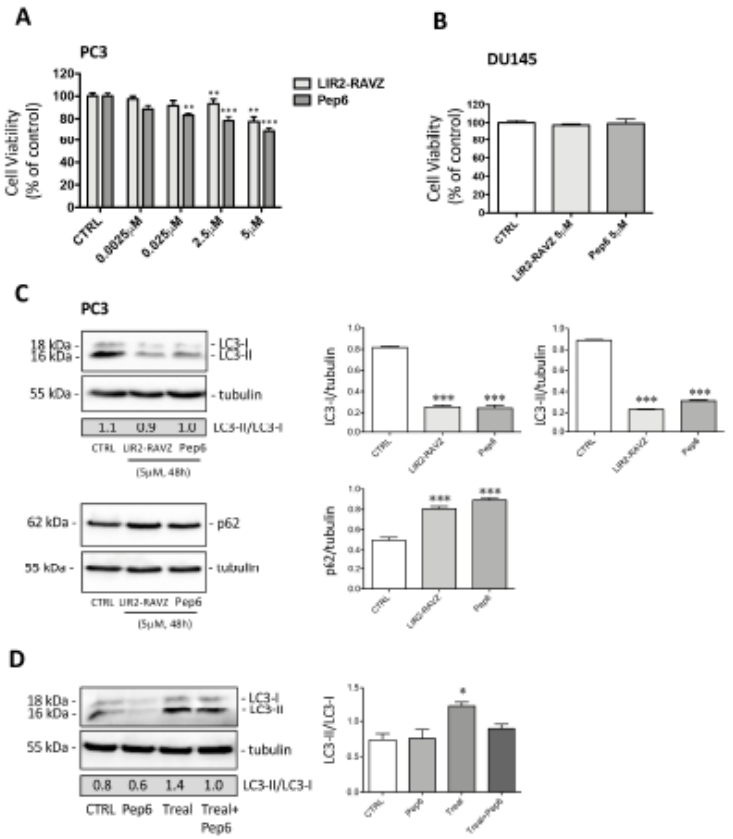



**Figure 5.** (**A**) Effect of LIR2-RavZ and Pep6 on PC3 cell viability. Cell viability was determined using MTT assay after 72 h. Six independent biological samples for each condition were analyzed ($n = 6$). Statistical analysis was performed using one-way ANOVA followed by Dunnett's test (** $p < 0.01$ vs. CTRL; *** $p < 0.001$ vs. CTRL). (**B**) Effect of LIR2-RavZ and Pep6 on DU145 cell viability. Cell viability was determined using MTT assay after 72 h. Six independent biological samples for each condition were analyzed ($n = 6$). Statistical analysis was performed using one-way ANOVA followed by Dunnett's test. (**C**) Western blot analysis of LC3-II/LC3-I ratio and p62 in PC3 cells treated with LIR2-RavZ and Pep6. The relative optical density of LC3-I/tubulin, LC3-II/tubulin, and p62/tubulin was quantified using ImageJ software. The bar graph represents the mean ± SD calculated from three independent experiments. Statistical analysis was performed using one-way ANOVA followed by Dunnett's post-test (*** $p < 0.001$ vs. CTRL). (**D**) Western blot analysis of LC3-II/LC3-I ratio in PC3 cells treated with Pep6 (5 µM) and trehalose (100 mM) for 48 h. The relative optical density of LC3-II/LC3-I was quantified using ImageJ software (version 1.50i). The bar graph represents the mean ± SD calculated from three independent experiments. Statistical analysis was performed using one-way ANOVA followed by Bonferroni's post test (* $p < 0.05$ vs. CTRL).

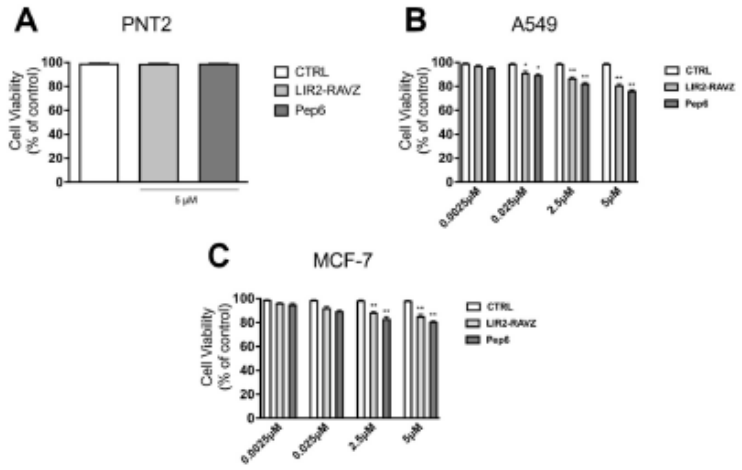



**Figure 6.** Effect of LIR2-RavZ and Pep6 on cell viability. Cell viability was determined using MTS assay on PNT2 (**A**), A549 (**B**), and MCF-7 (**C**) 96 h post-treatment. Absorbance was measured with a 96-well plate spectrophotometer (Varioskan Flash Multimode Reader) at 490 nm (* $p < 0.05$ vs. CTRL; ** $p < 0.01$ vs. CTRL).

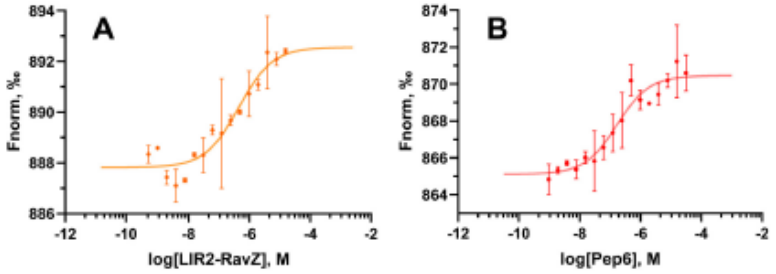



**Figure 7.** MST curves acquired using human recombinant His-tagged LC3B protein incubated with different concentrations of the control peptide LIR2-RavZ (**A**) and Pep6 (**B**), using the Monolith NT.115$^{Pico}$ instrument. Two independent experiments were performed to compute the $K_d$ curve.

## 3. Materials and Methods

*Computational studies.* The LC3B computational model utilized in this study was constructed using the 3D coordinates of the chain A of the LC3B/FYCO1-LIR complex, retrieved from the Protein Data Bank (PDB accession code 5D94) [41]. The structure available in this databank represents the sequence from amino acids 5 to 123. During the maturation process, the signal peptide spanning amino acids 121 to 125 is cleaved from the C-terminal end. Therefore, the X-ray, as well as the computational models built using it, represents the pro-LC3B state of the LC3 maturation process. In this paper, we named it "LC3B" for simplicity. The LC3B model underwent optimization using the "Protein Preparation Wizard" tool within Maestro Software (release 2021-2, Schrödinger, LLC, New York, NY, USA). This tool facilitated the initial steps of system setup, encompassing: (1) assessing the protonation states of residues at pH 7.4, (2) verifying residue completeness, (3) resolving atomic clashes, and (4) applying the OPLS4 force field [42]. The docking of LIR2-RavZ was executed using the "Peptide Docking" tool in Maestro software (release 2021-2, Schrödinger, LLC, New York, NY, USA). This process involved defining a grid,

creating a centroid through the "Centroid of Selected Residues" option, and selecting the residues in complex with the LIR domain in the X-ray structure, such as Arg10, Lys30, Tyr50, Lys51, Leu53, Arg69, and Phe108. The grid dimensions accommodated a linear 12-residue peptide. The sequence of LIR2-RavZ (Table 1) was introduced, cis amide bonds were deactivated, 150 poses were generated, and the Glide score [43] served as the scoring function, resulting in the LC3B/LIR2-RavZ starting complex. After this, alanine scanning was accomplished using the "Residue Scanning" tool available in Maestro (release 2021-2, Schrödinger, LLC, New York, NY, USA). Each peptide residue was individually mutated to alanine to pinpoint crucial hot/non-hot spots for LC3B and LIR2-RavZ interaction. The same tool was employed for the affinity maturation process. All residues, except the two cysteines, went through simultaneous mutations including natural and unnatural residues, selected from the library available in the tool. Affinity maturation utilized Monte Carlo optimization with a maximum of 100,000 steps, focusing on optimizing affinity and generating a maximum of 150 structures. Subsequently, a cubic box of water molecules, represented by the TIP3P model, was built around the protein–ligand complex, and subsequent system energy minimization was followed by 250 ns long MD simulations, using the Desmond algorithm in Maestro (release 2021-2, Schrödinger, LLC, New York, NY, USA). The "Simulation Interactions Diagram" tool evaluated peptide and ligand stability (see Supplementary Materials, Figure S3, for the Cα atom RMSD plots of the main simulated systems). Finally, ligand binding free energy (ΔG) calculations were conducted using the Prime algorithm within Maestro software (release 2021-2, Schrödinger, LLC, New York, NY, USA), employing the MM-GBSA algorithm. The single-trajectory approach was adopted, neglecting entropy contributions to the binding free energy. The resulting binding free-energy values were denoted as ΔG* by our group [21,44,45]. This protocol was applied to all peptides under investigation (Tables 1 and 2).

*Cell lines.* The PNT2 (European Collection of Authenticated Cell Cultures, ECACC, UK) human prostate cell line was purchased from Sigma-Aldrich (St. Louis, MO, USA). PC3 and DU145 human CRPC cell lines, A549 human lung cancer cell lines, and MCF-7 human breast cancer cell lines were purchased from the American Type Culture Collection (ATCC, USA). PNT2 cells were cultured at 37 °C and 5% $CO_2$ in RPMI 1640 (Gibco Laboratories) supplemented with 10% FBS (Gibco Laboratories), 1% 100 µ/mL penicillin/streptomycin (Gibco Laboratories), and 2% L-glutamine (Gibco Laboratories). PC3 and DU145 cells were maintained in RPMI 1640 medium (EuroClone, Milano, Italy) supplemented with 7.5% (PC3) and 5% (DU145) FBS (Gibco Laboratories), 1% L-glutamine, and antibiotics (100 IU/mL penicillin G). A549 and MCF-7 cells were cultured in Dulbecco's modified Eagle medium (DMEM, Lonza, Switzerland) supplemented with 10% fetal bovine serum (FBS, Gibco Laboratories, USA), 1% 100 u/mL penicillin/streptomycin (Gibco Laboratories), and 1% L-glutamine (Gibco Laboratories).

*Cytotoxicity studies.* PNT2, A549, and MCF-7 cells were seeded at a density of $1 \times 10^4$ cells/well in 96-well plates and maintained under standard growth conditions. After 24 h, the cells were treated for 1h at 0.0025 µM, 0.025 µM, 2.5 µM, and 5 µM to a final volume of 100 µL. After 96 h, cell viability was assessed using MTS assay, according to the manufacturer's protocol (Cell Titer 96 Aqueous One Solution Cell Proliferation Assay; Promega, Nacka, Sweden) using a 96-well-plate spectrophotometer (Varioskan Flash Multimode Reader; ThermoFisher Scientific, Waltham, MA, USA) set at λ = 490 nm. The absorbance value of untreated cells was set at 100% (control), and the viability of treated cells was expressed as a percentage of the control. PC3 and DU145 cells were plated at a density of $3 \times 10^4$ cells/well in 24-well plates. After 48 h, the cells were treated with compounds at 0.0025 µM, 0.025 µM, 2.5 µM, and 5 µM doses. After 24 h, 48 h, or 72 h, cell viability was analyzed using 3-(4,5-dimethylthiazole-2-yl)-2,5-diphenyltetrazolium bromide (MTT) (Sigma-Aldrich, St. Louis, MO, USA) assay. At the end of the treatment, the medium was replaced with MTT solution (0.5 mg/mL) in RPMI without phenol red and FBS. After 30–45 min at 37 °C, the precipitate of formazan was dissolved with isopropanol. Absorbance (λ = 550 nm) was measured through the use of an EnSpire Multimode Plate

reader (Perkin Elmer, Milano, Italy). Three independent experiments were performed for each condition.

*Western blot (WB) assay.* To investigate the effects of compounds on autophagy modulation, PC3 cells were plated at $2 \times 10^5$ cells/dish in 6 cm dishes for 24 and 48 h. The cells were then treated and adherent, and floating cells were harvested and lysed in RIPA buffer. Protein extracts (10–20 µg) were resuspended in reducing sample buffer (Bio-Rad Laboratories, Segrate, Milano, Italy) and heated at 95 °C for 5 min. Following electrophoretic separation via SDS-PAGE, the proteins were transferred onto PVDF or nitrocellulose membranes. After blocking, the membranes were incubated with anti-LC3 (L8918) (Sigma-Aldrich, St. Louis, MO, USA)) and anti-SQSTM1/p62 (PA5-20839) (Thermo Fisher Scientific, Watham, MA, USA) primary antibody. Peroxidase-conjugated secondary anti-rabbit or anti-mouse antibodies were used for 1 h at room temperature, and membranes were processed using the enhanced chemiluminescence kit Cyanagen Ultra (Cyanagen, Bologna, Italy). In each WB experiment, alpha-tubulin expression (T6199) (Sigma-Aldrich) was evaluated as a loading control.

*MST Experiments.* The binding of the peptides on LC3B protein was assessed by employing the Monolith NT.115$^{Pico}$ instrument (NanoTemper Technologies GmbH, München, Germany), which allows for estimation of the dissociation constant ($K_d$) comprising the concentration range from 1 pM to mM. Specifically, His-tagged human recombinant LC3B (catalog number 14555-H07E, Sino Biological, Beijing, China) was red-labeled using the dedicated His-Tag Labeling Kit RED-tris-NTA 2nd Generation from NanoTemper Technologies (Product No. MO-L018) for 30 min at room temperature. The "Binding Affinity" mode of the dedicated software MO.Control v1.6 (NanoTemper Technologies GmbH, München, Germany) was used to perform the MST experiments. In detail, a fixed concentration of red-labeled LC3B (10 nM) was mixed with sixteen 1:1 serial dilutions of LIR2-RavZ peptide (concentration range: 15.6 µM–0.477 nM), which was used as a positive control, and Pep6 (concentration range: 31.3 µM–0.954 nM), respectively. Both interacting species were dissolved in PBS-T buffer (phosphate-buffered saline + 0.05% Tween™ 20) from NanoTemper Technologies with 2.5% dimethyl sulfoxide (DMSO) for molecular biology (Product No. D8418; Sigma-Aldrich, Saint Louis, USA). The samples were incubated for 15 min at room temperature, then filled into standard capillaries (Product No. MO-K022; NanoTemper Technologies GmbH, München, Germany), and, finally, measured through the employment of an excitation power of 20% and the medium MST power (40%) fixing the temperature at 25 °C. The auto-fluorescence of each peptide was evaluated before proceeding to the determination of the $K_d$. The experimental data were processed by employing dedicated MO.Affinity Analysis software v2.3 (NanoTemper Technologies GmbH, München, Germany) and the $K_d$ model for fitting the binding curve, while the figures were generated using GraphPad Prism software v8.0.2 (GraphPad, Boston, MA, USA). The MST analysis report can be consulted in the Supplementary Materials, namely Table S3 (dataset overview) and Figure S4 (MST traces and capillary scan graphs).

## 4. Conclusions

Our computational approach aimed at designing new peptides endowed with high affinity on a specific target led to the desired goal. In fact, in this study, starting from the sequence of a known peptide demonstrating affinity on LC3B, we applied an approach combining MD simulations and MM-GBSA calculations to computationally predict the binding affinity of new peptides obtained by mutating the sequence of a known peptide (affinity maturation process). As can be seen in Figures 1, 2 and 4 and Tables 1 and 2, using our approach, both the stability and the affinity of the new peptides were significantly improved. In fact, considering the predicted binding modes of LIR2-RavZ and the ones of Pep3 and Pep6, the new peptides can create numerous electrostatic interactions with the LC3B positive charged area shaped by Lys30, Lys65, Arg69, and Arg70 residues, explaining the low predicted $\Delta G^*$ values. Moreover, the backbone rigidification conferred by the presence of the disulfide bridge in the sequences led to a low fluctuation in the peptides on

the LIR binding site of LC3B, together with better filling of the LC3B basin shaped by Ile23, Leu58, and Phe108 residues (close to the W-site). Finally, it has to be stressed that Pep6 contains D-amino acids and has side chain modification to improve metabolic resistance. We are aware that cellular phosphatases could remove the phosphate groups on serine, threonine, and tyrosine. Therefore, we understand that the phospho-peptides Pep3 and Pep6 may be subject to hydrolysis in the cellular environment; in particular, the L-Tyr(P) at position 7 may be the most susceptible. Nevertheless, we anticipate that this susceptibility to cellular phosphatases could be reduced by incorporating the disulfide bridge, the Bz groups, and the D-amino acids in the sequence.

Thus, based on the biological assessments mentioned earlier, it is reasonable to hypothesize that, if adequately delivered into the intracellular compartment via nanocarriers or liposomes, Pep6 could serve as a novel modulator of autophagy, potentially beneficial for cancer treatment.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms25094622/s1.

**Author Contributions:** Conceptualization, E.M.A.F. and G.G.; methodology, E.M.A.F., J.S., R.M.M. and M.G.; investigation, M.A., E.M.A.F., M.M.M., J.S. and M.G.; resources, G.G., G.R., A.C. and R.M.M.; data curation, E.M.A.F. and G.G.; writing—original draft preparation, all authors; writing—review and editing, E.M.A.F. and G.G.; funding acquisition, G.G., G.R., R.M.M., M.G. and A.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article and Supplementary Materials.

**Acknowledgments:** G.G. would like to thank INDACO for providing high-performance computing resources and support.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Behrends, C.; Sowa, M.E.; Gygi, S.P.; Harper, J.W. Network Organization of the Human Autophagy System. *Nature* **2010**, *466*, 68–76. [CrossRef] [PubMed]
2. Xie, Z.; Klionsky, D.J. Autophagosome Formation: Core Machinery and Adaptations. *Nat. Cell Biol.* **2007**, *9*, 1102–1109. [CrossRef] [PubMed]
3. Huang, J.; Brumell, J.H. Bacteria-Autophagy Interplay: A Battle for Survival. *Nat. Rev. Microbiol.* **2014**, *12*, 101–114. [CrossRef] [PubMed]
4. Jacquet, M.; Guittaut, M.; Fraichard, A.; Despouy, G. The Functions of Atg8-Family Proteins in Autophagy and Cancer: Linked or Unrelated? *Autophagy* **2021**, *17*, 599–611. [CrossRef]
5. Hain, A.U.P.; Weltzer, R.R.; Hammond, H.; Jayabalasingham, B.; Dinglasan, R.R.; Graham, D.R.M.; Colquhoun, D.R.; Coppens, I.; Bosch, J. Structural Characterization and Inhibition of the Plasmodium Atg8–Atg3 Interaction. *J. Struct. Biol.* **2012**, *180*, 551–562. [CrossRef]
6. Villa, S.; Legnani, L.; Colombo, D.; Gelain, A.; Lammi, C.; Bongiorno, D.; Ilboudo, D.P.; Mcgee, K.E.; Bosch, J.; Grazioso, G. Structure-Based Drug Design, Synthesis and Biological Assays of *P. Falciparum* Atg3–Atg8 Protein-Protein Interaction Inhibitors. *J. Comput. Aided Mol. Des.* **2018**, *32*, 473–486. [CrossRef]
7. Wong, Y.C.; Holzbaur, E.L.F. Optineurin is an Autophagy Receptor for Damaged Mitochondria in Parkin-Mediated Mitophagy That is Disrupted by an ALS-Linked Mutation. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E4439–E4448. [CrossRef]
8. McLendon, P.M.; Ferguson, B.S.; Osinska, H.; Shenuarin Bhuiyan, M.; James, J.; McKinsey, T.A.; Robbins, J. Tubulin Hyperacetylation is Adaptive in Cardiac Proteotoxicity by Promoting Autophagy. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E5178–E5186. [CrossRef]
9. Chen, M.; Hong, M.J.; Sun, H.; Wang, L.; Shi, X.; Gilbert, B.E.; Corry, D.B.; Kheradmand, F.; Wang, J. Essential Role for Autophagy in the Maintenance of Immunological Memory against Influenza Infection. *Nat. Med.* **2014**, *20*, 503–510. [CrossRef]
10. Quan, W.; Lim, Y.M.; Lee, M.S. Role of Autophagy in Diabetes and Endoplasmic Reticulum Stress of Pancreatic β-Cells. *Exp. Mol. Med.* **2012**, *44*, 81–88. [CrossRef]

11. Sarparanta, J.; García-Macia, M.; Singh, R. Autophagy and Mitochondria in Obesity and Type 2 Diabetes. *Curr. Diabetes Rev.* **2016**, *13*, 352–369. [CrossRef]
12. Mao, Y.; Yu, F.; Wang, J.; Guo, C.; Fan, X. Autophagy: A New Target for Nonalcoholic Fatty Liver Disease Therapy. *Hepat. Med.* **2016**, *8*, 27–37. [CrossRef] [PubMed]
13. Giannopoulos, S.; Bolkus, C.C.; Zografos, E.; Athanasiou, A.; Bongiovanni, A.M.; Dolliver's, G.; Bakoyiannis, C.N.; Theodoropoulos, G.E.; Zografos, G.C.; Witkin, S.S.; et al. Targeting Both Autophagy and Immunotherapy in Breast Cancer Treatment. *Metabolites* **2022**, *12*, 966. [CrossRef] [PubMed]
14. Pérez-Hernández, M.; Arias, A.; Martínez-García, D.; Pérez-Tomás, R.; Quesada, R.; Soto-Cerrato, V. Targeting Autophagy for Cancer Treatment and Tumor Chemosensitization. *Cancers* **2019**, *11*, 1599. [CrossRef]
15. Amaravadi, R.K.; Lippincott-Schwartz, J.; Yin, X.M.; Weiss, W.A.; Takebe, N.; Timmer, W.; DiPaola, R.S.; Lotze, M.T.; White, E. Principles and Current Strategies for Targeting Autophagy for Cancer Treatment. *Clin. Cancer Res.* **2011**, *17*, 654–666. [CrossRef]
16. Yu, G.; Klionsky, D.J. Life and Death Decisions—The Many Faces of Autophagy in Cell Survival and Cell Death. *Biomolecules* **2022**, *12*, 866. [CrossRef] [PubMed]
17. Zhou, M.; Wang, R. Small-Molecule Regulators of Autophagy and Their Potential Therapeutic Applications. *ChemMedChem* **2013**, *8*, 694–707. [CrossRef] [PubMed]
18. Akin, D.; Wang, S.K.; Habibzadegah-Tari, P.; Law, B.; Ostrov, D.; Li, M.; Yin, X.M.; Kim, J.S.; Horenstein, N.; Dunn, W.A. A Novel ATG4B Antagonist Inhibits Autophagy and Has a Negative Impact on Osteosarcoma Tumors. *Autophagy* **2014**, *10*, 2021–2035. [CrossRef]
19. Fu, Y.; Hong, L.; Xu, J.; Zhong, G.; Gu, Q.; Gu, Q.; Guan, Y.; Zheng, X.; Dai, Q.; Luo, X.; et al. Discovery of a Small Molecule Targeting Autophagy via ATG4B Inhibition and Cell Death of Colorectal Cancer Cells in Vitro and in Vivo. *Autophagy* **2019**, *15*, 295–311. [CrossRef] [PubMed]
20. Bosc, D.; Vezenkov, L.; Bortnik, S.; An, J.; Xu, J.; Choutka, C.; Hannigan, A.M.; Kovacic, S.; Loo, S.; Clark, P.G.K.; et al. A New Quinoline-Based Chemical Probe Inhibits the Autophagy-Related Cysteine Protease ATG4B. *Sci. Rep.* **2018**, *8*, 11653. [CrossRef]
21. Fassi, E.M.A.; 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**, *23*, 5070. [CrossRef] [PubMed]
22. Mikhaylova, O.; Stratton, Y.; Hall, D.; Kellner, E.; Ehmer, B.; Drew, A.F.; Gallo, C.A.; Plas, D.R.; Biesiada, J.; Meller, J.; et al. VHL-Regulated MiR-204 Suppresses Tumor Growth through Inhibition of LC3B-Mediated Autophagy in Renal Clear Cell Carcinoma. *Cancer Cell* **2012**, *21*, 532–546. [CrossRef] [PubMed]
23. Shin, S.W.; Kim, S.Y.; Park, J.W. Autophagy Inhibition Enhances Ursolic Acid-Induced Apoptosis in PC3 Cells. *Biochim. Biophys. Acta Mol. Cell Res.* **2012**, *1823*, 451–457. [CrossRef] [PubMed]
24. Xie, M.; Liu, J.; Wang, Z.; Sun, B.; Wang, J. Inhibitory Effects of 5-Heptadecylresorcinol on the Proliferation of Human MCF-7 Breast Cancer Cells through Modulating PI3K/Akt/MTOR Pathway. *J. Funct. Foods* **2020**, *69*, 103946. [CrossRef]
25. Fu, R.; Deng, Q.; Zhang, H.; Hu, X.; Li, Y.; Liu, Y.; Hu, J.; Luo, Q.; Zhang, Y.; Jiang, X.; et al. A Novel Autophagy Inhibitor Berbamine Blocks SNARE-Mediated Autophagosome-Lysosome Fusion through Upregulation of BNIP3. *Cell Death Dis.* **2018**, *9*, 243. [CrossRef] [PubMed]
26. Tang, J.; Zhu, J.; Ye, Y.; Liu, Y.; He, Y.; Zhang, L.; Tang, D.; Qiao, C.; Feng, X.; Li, J.; et al. Inhibition LC3B Can Increase Chemosensitivity of Ovarian Cancer Cells. *Cancer Cell Int.* **2019**, *19*, 199. [CrossRef] [PubMed]
27. Quan, Y.; Lei, H.; Wahafu, W.; Liu, Y.; Ping, H.; Zhang, X. Inhibition of Autophagy Enhances the Anticancer Effect of Enzalutamide on Bladder Cancer. *Biomed. Pharmacother.* **2019**, *120*, 109490. [CrossRef] [PubMed]
28. Lazova, R.; Camp, R.L.; Klump, V.; Siddiqui, S.F.; Amaravadi, R.K.; Pawelek, J.M. Punctate LC3B Expression is a Common Feature of Solid Tumors and Associated with Proliferation, Metastasis, and Poor Outcome. *Clin. Cancer Res.* **2012**, *18*, 370–379. [CrossRef] [PubMed]
29. Johansen, T.; Lamark, T. Selective Autophagy: ATG8 Family Proteins, LIR Motifs and Cargo Receptors. *J. Mol. Biol.* **2020**, *432*, 80–103. [CrossRef]
30. Choy, A.; Dancourt, J.; Mugo, B.; O'Connor, T.J.; Isberg, R.R.; Melia, T.J.; Roy, C.R. The Legionella Effector RavZ Inhibits Host Autophagy through Irreversible Atg8 Deconjugation. *Science* **2012**, *338*, 1072–1076. [CrossRef]
31. 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**, *9*, 2343. [CrossRef] [PubMed]
32. Sgrignani, J.; Arnoldi, A.; Lammi, C.; Grazioso, G.; Roda, G. Inhibition of PCSK9 D374Y/LDLR Protein–Protein Interaction by Computationally Designed T9 Lupin Peptide. *ACS Med. Chem. Lett.* **2018**, *10*, 425–430.
33. Yang, A.; Pantoom, S.; Wu, Y.W. Elucidation of the Anti-Autophagy Mechanism of the Legionella Effector Ravz Using Semisynthetic LC3 Proteins. *Elife* **2017**, *6*, e23905. [CrossRef] [PubMed]
34. Cristofani, R.; Montagnani Marelli, M.; Cicardi, M.E.; Fontana, F.; Marzagalli, M.; Limonta, P.; Poletti, A.; Moretti, R.M. Dual Role of Autophagy on Docetaxel-Sensitivity in Prostate Cancer Cells. *Cell Death Dis.* **2018**, *9*, 889. [CrossRef] [PubMed]
35. Fan, S.; Yue, L.; Wan, W.; Zhang, Y.; Zhang, B.; Otomo, C.; Li, Q.; Lin, T.; Hu, J.; Xu, P.; et al. Inhibition of Autophagy by a Small Molecule through Covalent Modification of the LC3 Protein. *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 26105–26114. [CrossRef] [PubMed]
36. Hamaï, A.; Codogno, P. New Targets for Acetylation in Autophagy. *Sci. Signal.* **2012**, *5*, pe29. [CrossRef] [PubMed]

37. Huang, R.; Xu, Y.; Wan, W.; Shou, X.; Qian, J.; You, Z.; Liu, B.; Chang, C.; Zhou, T.; Lippincott-Schwartz, J.; et al. Deacetylation of Nuclear LC3 Drives Autophagy Initiation under Starvation. *Mol. Cell* **2015**, *57*, 456–466. [CrossRef] [PubMed]
38. Song, T.; Su, H.; Yin, W.; Wang, L.; Huang, R. Acetylation Modulates LC3 Stability and Cargo Recognition. *FEBS Lett.* **2019**, *593*, 414–422. [CrossRef] [PubMed]
39. Ouyang, D.Y.; Xu, L.H.; He, X.H.; Zhang, Y.T.; Zeng, L.H.; Cai, J.Y.; Ren, S. Autophagy Is Differentially Induced in Prostate Cancer LNCaP, DU145 and PC-3 Cells via Distinct Splicing Profiles of ATG5. *Autophagy* **2013**, *9*, 20–32. [CrossRef]
40. Seidel, S.A.I.; Dijkman, P.M.; Lea, W.A.; van den Bogaart, G.; Jerabek-Willemsen, M.; Lazic, A.; Joseph, J.S.; Srinivasan, P.; Baaske, P.; Simeonov, A.; et al. Microscale Thermophoresis Quantifies Biomolecular Interactions under Previously Challenging Conditions. *Methods* **2013**, *59*, 301–315. [CrossRef]
41. Olsvik, H.L.; Lamark, T.; Takagi, K.; Larsen, K.B.; Evjen, G.; Øvervatn, A.; Mizushima, T.; Johansen, X.T. FYCO1 Contains a C-Terminally Extended, LC3A/B-Preferring LC3-Interacting Region (LIR) Motif Required for Efficient Maturation of Autophagosomes during Basal Autophagy. *J. Biol. Chem.* **2015**, *290*, 29361–29374. [CrossRef] [PubMed]
42. Lu, C.; Wu, C.; Ghoreishi, D.; Chen, W.; Wang, L.; Damm, W.; Ross, G.A.; Dahlgren, M.K.; Russell, E.; Von Bargen, C.D.; et al. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *J. Chem. Theory Comput.* **2021**, *17*, 4291–4300. [CrossRef] [PubMed]
43. Friesner, R.A.; Murphy, R.B.; Repasky, M.P.; Frye, L.L.; Greenwood, J.R.; Halgren, T.A.; Sanschagrin, P.C.; Mainz, D.T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196. [CrossRef] [PubMed]
44. Fassi, E.M.A.; Manenti, M.; Citarella, A.; Dei Cas, M.; Casati, S.; Micale, N.; Schirmeister, T.; Roda, G.; Silvani, A.; Grazioso, G. Computational Design, Synthesis and Biophysical Evaluation of β-Amido Boronic Acids as Potent SARS-CoV-2 $M^{pro}$ Inhibitors. *Molecules* **2023**, *28*, 2356. [CrossRef]
45. Lammi, C.; Fassi, E.M.A.; Manenti, M.; Brambilla, M.; Conti, M.; Li, J.; Roda, G.; Camera, M.; Silvani, A.; Grazioso, G. Computational Design, Synthesis, and Biological Evaluation of Diimidazole Analogues Endowed with Dual PCSK9/HMG-CoAR-Inhibiting Activity. *J. Med. Chem.* **2023**, *66*, 7943–7958. [CrossRef]

**Appendix 3**

**Title**: FYCO1 Peptide Analogs: Design and Characterization of Autophagy Inhibitors as Co-Adjuvants in Taxane Chemotherapy of Prostate Cancer

**Authors**: Fassi, E. M. A., Moretti, R. M., Montagnani Marelli, M., Garofalo, M., Gori, A., Pesce, C., **Albani, M**., Milano, E. G., Sgrignani, J., Cavalli, A., and Grazioso, G.

**Status**: Published in *International Journal of Molecular Sciences (2025)*

**DOI**: 10.3390/ijms26115365

**Author Contribution**: Dr. M. Albani performed microscale thermophoresis experiments and contributed to manuscript drafting.

International Journal of
*Molecular Sciences*

MDPI

*Article*

# FYCO1 Peptide Analogs: Design and Characterization of Autophagy Inhibitors as Co-Adjuvants in Taxane Chemotherapy of Prostate Cancer

**Enrico Mario Alessandro Fassi [1,*], Roberta Manuela Moretti [2], Marina Montagnani Marelli [2], Mariangela Garofalo [3], Alessandro Gori [4], Cristiano Pesce [3], Marco Albani [1], Erica Ginevra Milano [1], Jacopo Sgrignani [5], Andrea Cavalli [5] and Giovanni Grazioso [1]**

[1] Department of Pharmaceutical Sciences, Università degli Studi di Milano, Via L. Mangiagalli 25, 20133 Milano, Italy; marco.albani@unimi.it (M.A.); ericaginevra.milano@unimi.it (E.G.M.); giovanni.grazioso@unimi.it (G.G.)
[2] Department of Pharmacological and Biomolecular Sciences, Università degli Studi di Milano, Via G. Balzaretti 9, 20133 Milano, Italy; roberta.moretti@unimi.it (R.M.M.); marina.marellimontagnani@unimi.it (M.M.M.)
[3] Department of Pharmaceutical and Pharmacological Sciences, Università di Padova, Via F. Marzolo 5, 35131 Padova, Italy; mariangela.garofalo@unimi.it (M.G.); cristiano.pesce@phd.unipd.it (C.P.)
[4] National Research Council of Italy, Istituto di Scienze e Tecnologie Chimiche (SCITEC-CNR), Via M. Bianco 9, 20131 Milano, Italy; alessandro.gori@icrm.cnr.it
[5] Institute for Research in Biomedicine (IRB), Via Chiesa 5, 6500 Bellinzona, Switzerland; jacopo.sgrignani@irb.usi.ch (J.S.); andrea.cavalli@irb.usi.ch (A.C.)
* Correspondence: enrico.fassi@unimi.it

**Abstract:** Autophagy plays a central role in cellular degradation and recycling pathways involving the formation of autophagosomes from cellular components. The Atg8 protein family, particularly LC3, is essential to this process, and dysregulation has been implicated in many diseases (including cancer). Furthermore, therapeutic strategies targeting Atg8 proteins like LC3 can be advanced by exploiting the expanding knowledge of the "LC3 interacting region" (LIR) domain to develop inhibitory ligands. Here, we report a computational approach to design novel peptides that inhibit LC3B. The LIR domain of a known LC3B binder (the FYCO1 peptide) was used as a starting point to design new peptides with unnatural amino acids and conformational restraints. Accomplishing molecular dynamics simulations and binding free energy calculations on the complex of peptide–LC3B, new promising FYCO1 analogs were selected. These peptides were synthesized and investigated by biophysical and biological experiments. Their ability to affect cellular viability was determined in different cancer cell lines (prostate cancer, breast cancer, lung cancer, and melanoma). In addition, the ability to inhibit autophagy and enhance the apoptotic activity of Docetaxel was evaluated in PC-3 prostate cancer cells. In conclusion, this research presents a rational approach to designing and developing LC3B inhibitors based on the FYCO1-LIR domain. The designed peptides hold promise as potential therapeutic agents for cancer and as tools for further elucidating the role of LC3B in autophagy.

**Keywords:** peptide; LC3B binders; autophagy; cancer; Atg8; LIR motif; FYCO1



Academic Editor: Giuseppe Lucarelli

Received: 7 May 2025
Revised: 30 May 2025
Accepted: 1 June 2025
Published: 3 June 2025

**Citation:** Fassi, E.M.A.; Moretti, R.M.; Montagnani Marelli, M.; Garofalo, M.; Gori, A.; Pesce, C.; Albani, M.; Milano, E.G.; Sgrignani, J.; Cavalli, A.; et al. FYCO1 Peptide Analogs: Design and Characterization of Autophagy Inhibitors as Co-Adjuvants in Taxane Chemotherapy of Prostate Cancer. *Int. J. Mol. Sci.* **2025**, *26*, 5365. https://doi.org/10.3390/ijms26115365



## 1. Introduction

Within living organisms, autophagy operates as a meticulously orchestrated mechanism, targeting specific proteins and aged or impaired organelles through the use of double-membrane vesicles termed autophagosomes. Upon the fusion of autophagosomes

with lysosomes, the enclosed contents undergo degradation facilitated by the acidic milieu and lytic enzymes contained within the lysosomes [1,2]. The capacity for recycling within the autophagy machinery is found not only in eukaryotic cells but also in bacteria, enabling the preservation of physiological conditions [2,3]. The autophagy machinery comprises over 50 proteins named Atgs, with those belonging to the Atg8 family primarily responsible for autophagosome formation and cellular trafficking. In mammals, Atg8 proteins (mAtg8) consist of two subfamilies: GABARAP (GABA-A receptor-associated protein) and MAP1LC3 (microtubule-associated protein 1 light chain 3), also known as LC3. The GABARAP subfamily encompasses GABARAP, GARAPL1, and GABARAPL2, whereas the LC3 subfamily includes LC3A (comprising two splicing variants, LC3Aα and LC3Aβ), LC3B, LC3B2, and LC3C [4]. Proteins belonging to the same subfamily demonstrate notable sequence resemblances and fulfill analogous roles within the cellular environment. The GABARAP subfamily is pivotal in autophagosome closure and the enlistment of autophagic components, whereas LC3 proteins primarily engage in cargo recruitment during the process [2]. The exact impact of compounds that disrupt Atg8 proteins on autophagy remains incompletely understood. However, utilizing peptides or peptidomimetics to inhibit the Atg3–Atg8 interaction in *Plasmodium falciparum* holds promise as a strategy to combat malaria [5,6]. Conversely, it has been confirmed that dysregulations of the complex autophagy machinery are associated with diseases like neurodegenerative disorders [7], cardiomyopathies [8], infectious diseases [6,9], type II diabetes mellitus [10,11], hepatic steatosis [12], and cancer [13–15]. These conditions arise from the dysregulation of autophagy induced by various stimuli originating from internal or external environmental factors.

The function of autophagy in cancer appears to be highly complex and may have opposite roles in different cancer cells, stages, and conditions. In normal cells, autophagy removes altered molecules or dysfunctional organelles, maintaining cellular health. Furthermore, autophagy plays a protective role by maintaining genome stability and reducing cellular alterations involved in cellular transformation. For this reason, a reduction in autophagy activity is observed in the early stages of tumorigenesis. Conversely, in the advanced and metastatic stages of cancer, autophagy increases, allowing tumor cells to survive and adapt to foreign sites. In addition, autophagy is activated in cancer cells exposed to various stresses such as anticancer treatments, leading to chemoresistance [16,17].

Specifically, LC3B is significantly upregulated in prostate cancer (PC) tissues, especially in metastatic castration-resistant PC (mCRPC), as compared to benign prostate tissues [18,19]. In PC tissue, high-level LC3B expression is associated with key clinicopathological indicators of aggressive disease, including high Gleason scores and advanced tumor grades, highlighting its involvement in tumor progression and aggressive growth [19]. In addition, a lack of immunoreactivity for LC3B is an independent predictor of PC specific mortality, indicating that autophagy is complex and context-dependent in PC evolution [18]. Moreover, because taxane chemotherapy induces cellular stress that often triggers autophagy as a protective mechanism in tumor cells, targeting LC3B-mediated autophagy represents a promising adjuvant strategy to sensitize PC cells to taxanes, potentially enhancing treatment efficacy and reducing recurrence rates [20].

In this complex context, numerous studies have been conducted in PC evaluating the effects of autophagy activators, such as Rapamycin, Everolimus, and Temsirolimus, or inhibitors, such as Chloroquine (CQ) and hydroxychloroquine (HCQ), alone or in association with conventional therapies [21,22] to better understand the impact of autophagy on cancer progression.

These agents serve as invaluable tools for delving into the complex mechanisms of autophagy on a molecular scale. Furthermore, there exists potential for their further ad-

vancement into promising drug candidates, with the aim of dealing with pathological conditions such as cancer and other related medical ailments [23]. In this field, in our previous paper [24], we utilized computational methods to design peptides called WC8 and WC10, which demonstrated both a high calculated and measured affinity for GABARAP. Intriguingly, when prostate cancer cells (PC-3) were treated with WC8 and WC10 at concentrations ranging from 1 to 10 μM, the attained results highlighted the significant therapeutic potential of this approach. Notably, the peptides exhibited greater activity compared to Paclitaxel, a widely used anticancer drug [24]. Nevertheless, LC3B remains the most extensively studied Atg8 protein in humans since it is clearly associated with cancer.

Interestingly, the proteins involved in the autophagy process and capable of interacting with LC3 feature a distinct amino acid sequence referred to as the "LC3 interacting region" (LIR). This small protein sequence consists of four conserved residues that can be succinctly represented as a sequence of "$X_0$–$X_1$–$X_2$–$X_3$", where $X_0$ represents an aromatic residue (Trp/Phe/Tyr), $X_1$ and $X_2$ can denote any amino acids (often acidic or hydrophobic residues), and $X_3$ signifies a large hydrophobic residue such as Leu, Val, or Ile [25]. Consequently, the LIR domain can be considered a promising starting point to design a ligand capable of interacting with the LC3 subfamily. For this reason, among the proteins bearing the LIR domain, we focused our attention on the "FYVE and coiled-coil protein 1" (FYCO1), a protein involved in the transport of autophagosomes along microtubules in the plus-end direction [26,27]. In more depth, FYCO1 is an adaptor that connects LC3B on autophagosomal membranes to Rab7 and phosphatidylinositol-3-phosphate (PI3P), allowing for coordinated movement along the cytoskeleton and facilitating the microtubule plus-end-directed transport of autophagic vesicles. Rab7 binding facilitates vesicle docking and fusion with lysosomes, while its interaction with LC3B through the conserved LIR motif guarantees selective recruitment to autophagosomal membranes. Through PI3P interaction, the FYVE domain of FYCO1 maintains membrane association, and its coiled-coil region facilitates microtubule motor recruitment and dimerization, both of which are necessary for directional transport. FYCO1 behaves as a crucial regulator of autophagosomal maturation and intracellular trafficking due to its dual binding to LC3B and Rab7, which guarantees the effective lysosomal degradation of cargo [27]. Effective cargo degradation is made possible by LC3B binding, which stabilizes FYCO1's association with autophagosomal membranes and guarantees appropriate vesicle docking and fusion with lysosomes. This interaction's specificity keeps autophagosomes from mislocalizing, preserving the integrity of intracellular trafficking during autophagy. Moreover, FYCO1 plays a crucial role in LC3-associated phagocytosis by being recruited to Dectin-1 phagosomes and aiding in their maturation, transitioning them from early p40phox-containing phagosomes to late LAMP1-positive phagosomes [28]. In addition, FYCO1 and protrudin collaborate to promote the microtubule-mediated transport of late endosomes through endoplasmic reticulum–endosome contact sites [29]. Cerulli et al. [30] employed the FYCO1-LIR peptide to conduct structure–activity relationship (SAR) studies, systematically investigating how to enhance its preferential binding affinity and selectivity for LC3B over GABARAP. In particular, they systematically deleted some residues and identified the crucial determinants of the binding, e.g., the *N*-terminal region, E1287, critical hydrophobic interactions. Artificial amino acids (e.g., 1-naphthylalanine at position F1280 and tert-butylalanine at L1288) further enhanced binding to hydrophobic pockets. This improved affinity was also due to the introduction of an *N*-terminal arginine. Their effort resulted in a peptide (Comb1) with a 2.4-fold increase in binding affinity vs. the original FYCO1 peptide, as well. Indeed, the researchers also used diversity-oriented stapling to improve both the stability and efficacy of their peptides, culminating in novel inhibitors of LC3B [30].

Here, adopting the affordable computational protocol reported in our previous papers [24,31], we have designed new 12-residue-long FYCO1 peptide analogs capable of inhibiting LC3B. The newly designed FYCO1 analogs are shorter than the ones reported by Cerulli et al. [30] but, similarly, contain unnatural amino acids as well as conformational rigidification. All peptides were simulated in complex with LC3B protein by molecular dynamics (MD) simulations and the peptide binding free energy values were predicted by the Molecular Mechanics–Generalized Born surface area (MM-GBSA) approach, aiming to select the most promising FYCO1 analogs to be subsequently synthesized and tested by biophysical and biological in vitro assays on cancer cell lines.

## 2. Results and Discussion

### 2.1. Computational Design of FYCO1-LIR Analogs

FYCO1, a protein composed of 1478 amino acids, binds the LC3A and LC3B protein, principally by means of the amino acids belonging to the LIR domain, especially by the sequence FDIITDEE (1280–1288 region) [26], as demonstrated in the X-ray structure of the human LC3B in complex with the LIR domain of FYCO1 (DAVFDIITDEEL, PDB accession code 5D94 [26]). Here, we used this domain as the starting point to design new FYCO1 analogs, aiming at computationally designing new peptides endowed with improved affinity for LC3B. These new peptides, could weaken the maturation process of LC3B to LC3-II, finally shaping the autophagosome.

In our approach, we optimized the FYCO1 sequence by the rigidification of the peptide backbone and incorporating unnatural amino acids to limit the protease liability of the new peptides. In particular, the FYCO1-LIR domain (DAVFDIITDEEL) was initially simulated in complex with LC3B by accomplishing energy minimization and MD simulations. When the peptide reached geometrical stability in complex with LC3B, the peptide binding free energy value (ΔG*) was calculated by the MM-GBSA approach, attaining a value of −110.6 kcal/mol (Table 1). This value served as a reference for the subsequent design steps.

**Table 1.** Primary structure and estimated binding free energy values (ΔG*) of the reference peptide FYCO1-LIR and its analogs. Disulfide bonds are represented as square brackets (⌐¬). The curly bracket (⌐┴¬) of **AM7** joins two cysteines bis-alkylated by a para-dibromomethylbenzene (see the Materials and Methods section for details).

| Peptide | Sequence | ΔG* (kcal/mol) | SD (kcal/mol) | Avg. Cα RMSF (Å) |
|---|---|---|---|---|
| **FYCO1-LIR** | DAVFDIITDEEL | −110.6 | 7.0 | 1.49 |
| | **Sequence mutation** | | | |
| **AM1** | DAVFDI**M**TDEEL | −113.0 | 7.3 | 1.52 |
| **AM2** | DAV**F$_I$**DIITDEEL | −119.2 | 10.1 | 1.29 |
| **AM3** | DAV**F$_{Br}$**DIITDEEL | −111.6 | 7.7 | 1.44 |
| **AM4** | DAV**F$_I$**DI**M**TDEEL | −123.8 | 5.8 | 1.13 |
| | **Backbone rigidification** | | | |
| **AM5** | ⌐DAVFDIITCEEC¬ | −109.5 | 6.2 | 1.26 |
| **AM6** | ⌐DAVFDI**M**TCEEC¬ | −126.6 | 8.4 | 1.38 |

Table 1. *Cont.*

| Peptide | Sequence | ΔG* (kcal/mol) | SD (kcal/mol) | Avg. Cα RMSF (Å) |
|---|---|---|---|---|
| **AM7** | DAVFDIITCEEC | −118.5 | 8.3 | 1.00 |
| **AM8** | DAV$F_I$DIITCEEC | −109.2 | 5.3 | 1.03 |
| **AM9** | DAV$F_{Br}$DIITCEEC | −110.4 | 7.3 | 1.17 |
| **AM10** | DAV$F_I$DIMTCEEC | −137.1 | 6.1 | 0.60 |

### 2.2. Design of FYCO1 Analogs

Analysis of the final frame of the LC3B/FYCO1-LIR complex MD trajectory, through visual inspection, highlighted several key interactions (Figure 1A):

- The D1 of FYCO1-LIR establishes contact with K51 on LC3B.
- FYCO1-LIR's F4 residue engages in cation-π stacking with the side chain of LC3B-K51. Its phenyl ring is also positioned within the hydrophobic pocket 1 (HP1) on LC3B, formed by the residues P7, I23, P32, I34, L53, and F108.
- A hydrogen bond is formed between the NH group of D5 in FYCO1-LIR and the carbonyl group of LC3B-K51. Additionally, the acidic tail of D5 creates a salt bridge with LC3B-K49.
- The side chain of I6 in FYCO1-LIR is solvent-exposed, while the side chain of I7 projects inward, interacting with the hydrophobic pocket 2 (HP2) on LC3B, which is composed of the amino acids I35, F52, V54, L63, I66, and I67.
- The residue E10 of FYCO1-LIR interacts with the side chains of R69 and R70 on LC3B.

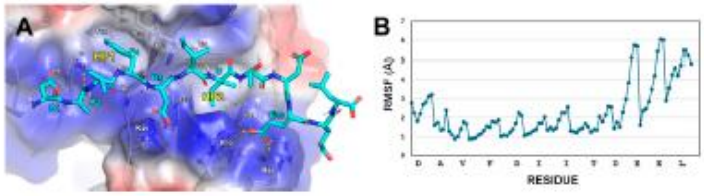



**Figure 1.** (**A**) Predicted binding mode of FYCO1-LIR (cyan sticks) in complex with LC3B resulting at the end of the MD simulation. The protein surface is colored depending on the atomic partial charges of the protein residues: blue for positive- and red for negative-charge areas, respectively. The H-bonds are represented as yellow dotted lines. (**B**) RMSF plot of FYCO1-LIR no-H heavy atoms (highlighted as dots).

The root mean square fluctuation (RMSF) plot of the FYCO1-LIR heavy atoms (Figure 1B) indicated that the C-terminal region of the reference peptide does not strongly bind to the LC3B surface but, interestingly, the LC3B-HP2 pocket surrounding FYCO1-LIR I7 appeared to have some extra space, and this led us to consider replacing I7 with a bulkier amino acid like methionine (**AM1** peptide, Table 1), potentially leading to a more tightly packed structure.

To validate this hypothesis, we simulated the **AM1** peptide in complex with LC3B, observing a predicted ΔG* value almost 3 kcal/mol lower than that of the parent peptide, though the RMSF value was comparable to that of FYCO1-LIR (Table 1). Building upon this, we reasoned that the LC3B-HP1 pocket, which binds F4 of FYCO1-LIR, could potentially

accommodate a larger hydrophobic side chain. To test this, we simulated two unnatural peptides, **AM2** and **AM3**, incorporating bromo- and iodo-phenylalanine at position 4. The growing applications of bromine and iodine substituents in anticancer treatment lend support to their selection. By inducing mitochondrial apoptosis and activating MAPK pathways, bromine-containing compounds, like bromamine T [32], have shown strong cytotoxic effects against breast and colon cancer cells, exhibiting superior anticancer activity when compared to non-halogenated analogs. Furthermore, because bromine and iodine isotopes can produce deadly Auger electrons that destabilize cancer cells at a very short range, they are being investigated in radiopharmaceuticals for targeted cancer therapy [33]. Notably, only **AM2** exhibited a lower predicted peptide binding free energy value and a significant reduction in the overall peptide conformational fluctuation (see RMSF values, Table 1). Integrating all these findings, we designed and simulated a novel peptide, **AM4**, incorporating both bulkier residues targeting LC3B-HP1 and HP2 pockets. Interestingly, the predicted $\Delta G^*$ and RMSF values for **AM4** were approximately 13 kcal/mol and 0.36 Å lower, respectively, than those of FYCO1-LIR (Table 1). These results suggested that the concurrent optimization of bulky side chains at positions 4 and 7, to enhance interactions with HP1 and HP2, yielded the most promising peptide among those designed up to that point.

### 2.3. Design of Stapled Peptides

We also sought to enhance peptide structural stability and reduce C-terminal RMSF by introducing covalent crosslinks. Disulfide bonds, a common strategy for restricting conformational freedom and stabilizing secondary structure in peptides, often improve drug-like properties and metabolic resistance. This approach aimed to enhance binding affinity, selectivity, cell permeability, and proteolytic degradation resistance [34–36]. Given the spatial proximity of D9 and L12 in FYCO1-LIR during MD simulations of the LC3B/**AM4** complex, we substituted these residues with cysteines to create a disulfide bridge, resulting in the **AM5** peptide. Subsequent simulation of **AM5** in complex with LC3B revealed a comparable predicted $\Delta G^*$ value, but a markedly lower average Cα atom RMSF (Table 1).

Building upon our previous designs, the peptide **AM6** incorporated both the productive I7M mutation from FYCO1-LIR and the conformational rigidification provided by the disulfide bonds in **AM5**. Remarkably, the simulation of **AM6** in complex with LC3B yielded a new low predicted $\Delta G^*$ value (Table 1). To explore the impact of linker length, we designed **AM7**, a stapled peptide with a -$CH_2$-Ph-$CH_2$- spacer connecting the sulfur atoms of C9 and C12. MD simulations indicated that this modification improved the predicted $\Delta G^*$ value by approximately 8 kcal/mol compared to FYCO1-LIR, while concurrently enhancing the structural stability of the peptide (Table 1). Taking into account all these data, we extended our simulations to analogs incorporating iodo- and bromo-phenylalanine at position 4 (**AM8** and **AM9**, respectively), while maintaining the **AM5** scaffold. This approach allowed us to assess the impact of halogen substitution on the peptide's conformational behavior and the interaction network. The attained results suggested that this modification significantly increased the structural stability of the peptides, as indicated by the reduced RMSF values (≈1 Å, Table 1). However, in both cases, the binding affinity toward LC3B remained largely unchanged compared to the native FYCO1-LIR peptide (Table 1).

In a final design iteration, we combined the most advantageous individual substitutions, namely, the I7M mutation (as in **AM6**) and the incorporation of iodo-phenylalanine at position 4 (as in **AM8**), within the **AM5** scaffold, yielding the **AM10** peptide. This peptide exhibited the lowest predicted $\Delta G^*$ (−137.1 kcal/mol) and Cα RMSF (0.60 Å) values among all designed peptides, representing the most promising candidate. All these findings

indicate that individual modifications are insufficient to substantially enhance peptide performance. However, the combination of all modifications exerts a strong synergistic effect, resulting in a peptide in which all residues are stably anchored on the LC3B protein surface. In fact, the calculated RMSF, with all values significantly below 1 Å (Figure 2), indicates minimal fluctuations. This stability is further corroborated by the RMSD analysis, which reveals that this peptide maintained a remarkably lower average deviation compared to all other designed peptides throughout the simulation (Supplementary Materials, Figure S1).

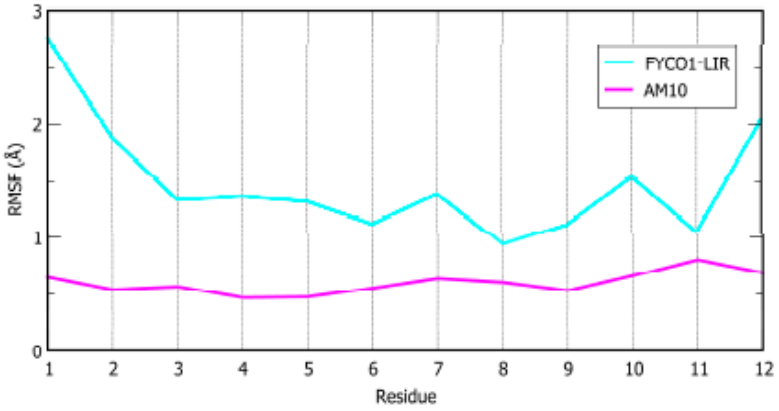



**Figure 2.** Cα atoms RMSF plot of FYCO1-LIR (cyan line) and **AM10** peptides (magenta line).

To compare our findings with Cerulli et al.'s 2020 study, we conducted docking and MD simulations on Comb1 (sequence RDDAV2DIITDEEαCQIQEW, in which "2" denotes 2-naphthylalanine and "α" is a tert-butylalanine), one of most potent and LC3B-selective peptides [26]. Structural alignment of the MD-stabilized LC3B-Comb1 complex with the LC3B/**AM10** complex revealed key interaction similarities (Figure 3). Specifically, the *N*-terminal residues of both peptides were located in an LC3B region enriched with positively charged residues, facilitating the formation of electrostatic and H-bond interactions. Of particular note is that the interaction sites involving LC3B-K49 and K51 are common to both peptides. However, only Comb1 engages H-bonds with R3 and R10, whereas **AM10** uniquely forms two H-bonds with T50. The 2-naphthylalanine (residue "2" in the sequence) of Comb1 at position 6, projected into the LC3B-HP1 hydrophobic pocket, was structurally mimicked by the Iodo-F4 of **AM10**. Furthermore, Comb1 exhibited similar interactions to **AM10** at several key sites: D7 and E13 of Comb1 mirrored D5 and E10 of **AM10** in their interactions (H-bonds/salt bridges) with LC3B-K49, -R69, and -R70. Comb1-E19 interacted with LC3B-K65 similarly to the C-terminus of **AM10**-C12. Additionally, I9 of Comb1 occupied the LC3B-HP2 region analogously to **AM10**-M7. However, a key difference was the deeper penetration of **AM10** into the hydrophobic pocket, allowing for additional hydrophobic contacts with LC3B-I35 and -I67.

In conclusion, computational analyses suggested that the shorter decapeptide **AM10** can emulate the interaction patterns of the longer Comb1 eicosapeptide. In addition, **AM10**'s cyclic conformation imposes geometric rigidity, enabling side chains to adopt conformations that maximize binding interactions through spatially matched residue complementarity.

In summary, our computational investigations have identified novel FYCO1 analogs possessing improved predicted binding affinity (lower $\Delta G^*$ values), enhanced stability on the LC3B surface, and increased metabolic stability through the incorporation of disulfide bonds and unnatural amino acids.

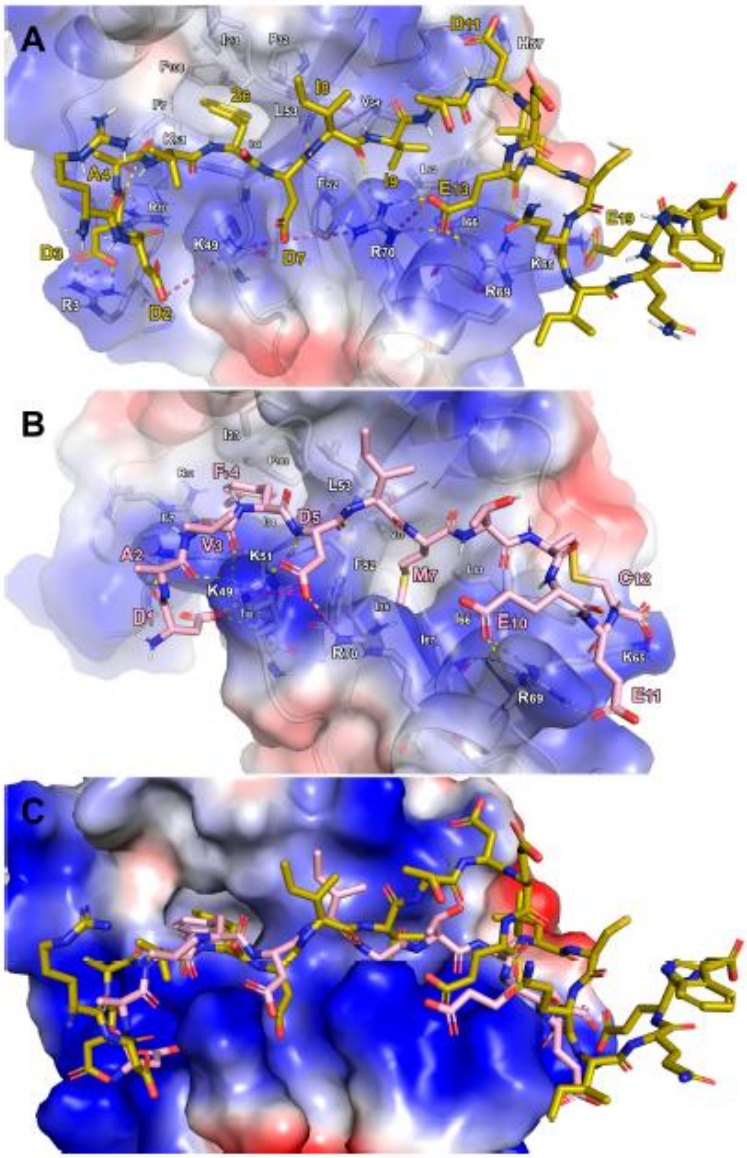



**Figure 3.** Predicted binding mode of (**A**) Comb1 (sand sticks) and (**B**) **AM10** (pink sticks) in complex with LC3B resulting in the most stable RMSD conformation during the MD simulation. (**C**) Superposition of the Comb1 and **AM10** binding mode on the LC3B protein surface. The protein surface is colored depending on the atomic partial charges of the protein residues: blue for positive- and red for negative-charge areas, respectively. The H-bonds and salt bridges are represented as yellow and purple dotted lines, respectively. The label sizes are proportional to the distance from the viewpoint.

### 2.4. Synthesis of Peptides

Based on the computational outcomes, the **AM2**, **AM6**, **AM7**, and **AM10** peptides were selected for synthesis and biophysical analysis (Table 1), while FYCO1-LIR was chosen as a control. In addition, these peptides incorporate key mutations that provide valuable insights into structure–activity relationship (SAR) studies. They were assembled by conventional solid-phase peptide synthesis (Fmoc-) and HPLC-purified before further processing. Disulfide cyclic analogs were obtained by $H_2O_2$ oxidation in mild conditions, whereas covalent cysteine crosslinking was induced by mixing the linear peptide form with the corresponding di-bromol linker (**AM7** peptide) in aqueous $NaHCO_3$/acetonitrile buffer (pH 8.0). Upon full conversion into their macrocyclic forms (<1 h, HPLC monitoring), the resulting compounds were HPLC-purified (see the Materials and Methods for details).

### 2.5. Biophysical Assays

To measure the dissociation constants ($K_d$) of the peptides on human recombinant His-tagged LC3B protein, MST experiments were carried out on a Monolith NT.115$^{Pico}$ instrument (see the Materials and Methods section for details). To validate the applied biophysical method, the $K_d$ value of the FYCO1-LIR peptide (DAVFDIITDEEL) was used as a positive control, yielding a $K_d$ value of 3.6 ± 1.2 µM (Figure 4A). This result closely aligns with the values reported by Cerulli et al. [30], who measured the $K_d$ of FYCO1 (DDAVFDIITDEELW) using Biolayer Interferometry (BLI) assays, obtaining a value of 3.1 ± 0.6 µM [30]. To note, the slightly lower $K_d$ value observed by Cerulli and coworkers could be explained by the presence of additional D and W residues in the amino- and carboxy-terminal groups, respectively, in the FYCO1 sequence tested by them. Next, the binding affinity of the peptides **AM2**, **AM6**, **AM7**, and **AM10** to the human LC3B protein was assessed. Interestingly, in the case of **AM2**, the addition of an iodine atom to the benzyl group of F4, which is projected into the HP1 pocket, nearly halved the $K_d$ value (2.2 ± 0.5 µM, Figure 4B) compared to FYCO1-LIR (Figure 4A). In addition, from the MST experiments, it is clearly observable that the iodine group of **AM2** stabilizes the complex; in fact, the points of the $K_d$ curve fit much better, and this is reflected in a high signal-to-noise ratio (SNR) value (14.4, Supplementary Materials, Table S1). Interestingly, these results reflected both the predicted computational ΔG* and RMSF values for **AM2** (Table 1).

The binding affinity to the LC3B protein significantly improved for the peptides with backbone rigidification and I7M mutation (i.e., **AM6** and **AM7**), as shown by their $K_d$ values of 0.6 ± 0.2 µM and 0.9 ± 0.4 µM, respectively (Figure 4C,D). Interestingly, in this case it is also observable that there is a good correlation between the computationally predicted ΔG* and experimental $K_d$ values (Table 1). The MST experiments reveal that a disulfide bond (**AM6**) facilitates binding interactions more effectively than a bulkier linker bridging the sulfur atoms (**AM7**). Furthermore, the -$CH_2$-Ph-$CH_2$- linker displays aggregation at concentrations of 31.25 µM or higher (Figure 4D), interfering with the MST analysis. The reliability of computational studies is further confirmed by MST analysis of the stapled peptide with the lowest predicted ΔG* and RMSF values (**AM10**, Table 1). Notably, this peptide exhibited the highest binding affinity for human LC3B protein ($K_d$ = 0.04 ± 0.01 µM). This finding highlights the crucial role of the iodine group in F4 and demonstrates that its combination with other mutations, such as M7 and C9–C12 (involved in a disulfide bond), leads to a dramatic improvement in $K_d$. In fact, the $K_d$ value is about 90-fold lower than that of the parent peptide FYCO1-LIR.

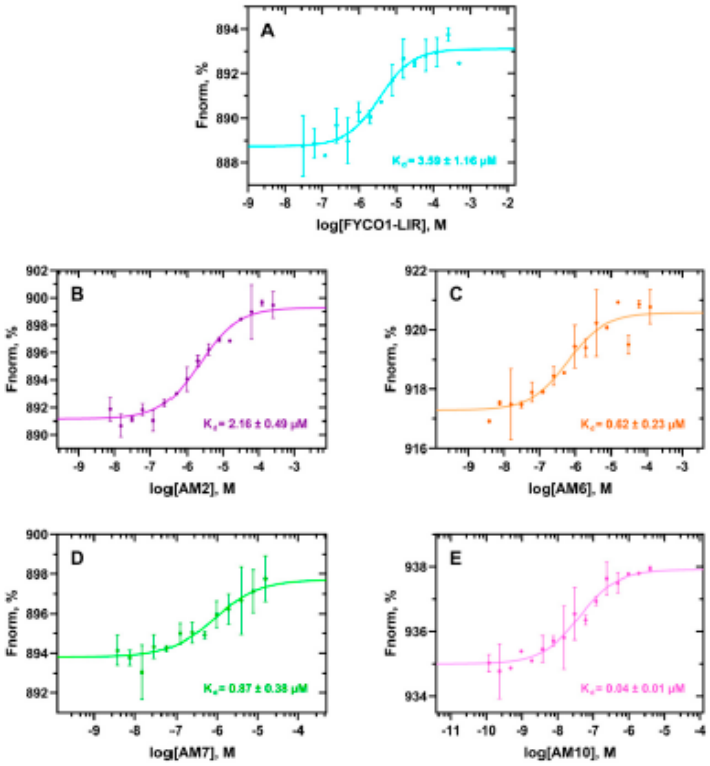



**Figure 4.** MST curves obtained by incubating human recombinant His-tagged LC3B protein with different concentrations of the control peptide FYCO1-LIR (**A**) and **AM2** (**B**), **AM6** (**C**), **AM7** (**D**), and **AM10** (**E**), using the Monolith NT.115$^{Pico}$ instrument. In the case of **AM7**, the three highest concentration points (125, 62.5, and 31.25 µM) were discarded due to aggregation. ($n$ = 2 independent measurements; error bars represent the standard deviation).

### 2.6. Biological Assays on PC-3 Cells

Based on the biophysical results, the biological activity of the peptides FYCO1-LIR (used as a reference), **AM6**, and **AM10** was evaluated on the viability of two CRPC cell lines which differ in their ability to activate the autophagic process: PC-3 cells, which show measurable endogenous autophagic activity; and DU145 cells which, due to lacking the ATG5 protein, are unable to form autophagosomes and activate the autophagic process [37]. PC-3 cells were treated for 72 h with increasing doses of FYCO1-LIR, **AM6**, and **AM10**, and at the end of the treatment, an MTT assay was conducted. Figure 5A–C shows that all peptides reduce cell viability in a significant dose-dependent manner starting from the 0.025 mM dose up to the 5 mM dose. Treatment of DU145 cells with the same compounds for 72 h at a dose of 5 mM showed no effect on cell viability (Figure 5D). This result highlights that the compounds are specific and selective; in fact, they are ineffective in DU145 cells that are ATG5-deficient, as shown in Figure 5E.

In order to evaluate in more depth the ability of the compounds to act as autophagy inhibitors in PC-3 cells, the expression of the LC3 and SQSTM1 (sequestosome1, p62) proteins was analyzed. Treatment with FYCO1-LIR, **AM6**, and **AM10** for 48 h at a dose of 5 µM significantly reduced the level of LC3-I and LC3-II, without modifying the LC3-

II/LC3-I ratio. It is plausible that all compounds bind to the LC3 precursor and prevent the efficient processing of LC3 by the cysteine protease Atg4, reducing the formation of both LC3-I and LC3-II without changing the ratio of LC3-II/LC3-I. This action inhibits the formation of autophagosomes and then the basal autophagy process. The analysis of LC3 expression after 72 h of treatment shows that FYCO1-LIR and **AM6** lose their efficacy, while **AM10** retains the ability to inhibit both LC3-I and LC3-II (Figure 6A). In addition, we analyzed the expression of p62, another protein involved in the autophagic process. This protein is recruited into autophagosomes linked to the material to be addressed for degradation in lysosomes. The p62 protein represents a marker of the autophagic flux; in fact, when autophagosomes fuse with lysosomes, the materials in autophagolysosomes were degraded including p62. The results obtained show that after treatment with the compounds for 72 h and 96 h, the expression of p62 increases significantly, demonstrating an impairment of autophagic flux. This result highlights that the compounds, mainly **AM10**, could determine an accumulation of materials that are not correctly degraded at the lysosomal level (Figure 6B). The analysis of the molecular mechanism suggests that the compounds interfere with the endogenous autophagy that preserves the tumor cells' survival; furthermore, it is presumable that the inhibition of a pro-survival basal autophagy determines a stressful condition which leads to a reduction in cell viability as demonstrated by the MTT assays.

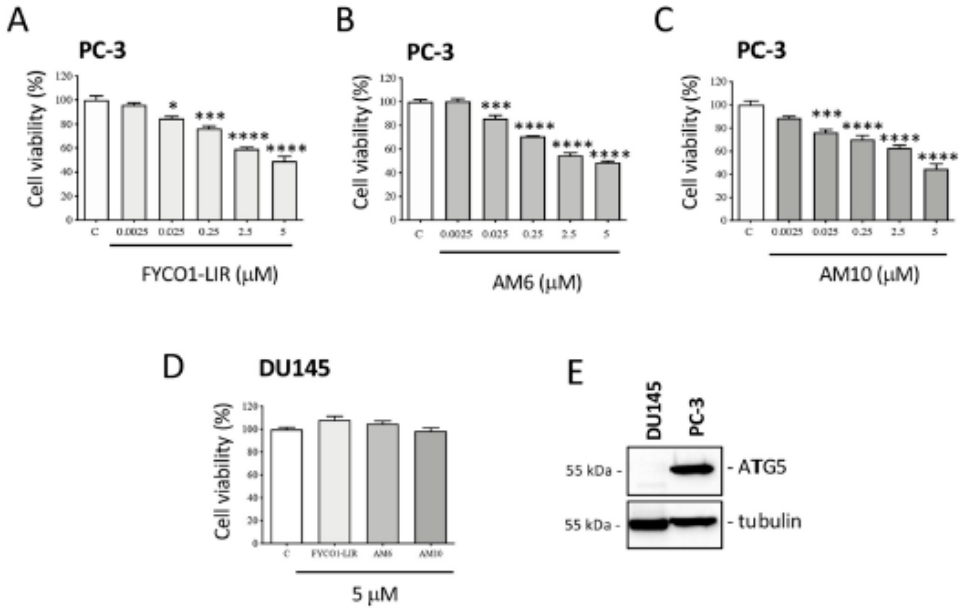



**Figure 5.** Effect of FYCO1-LIR analogs on CRPC cell viability. (**A**–**C**) PC-3 cells were treated with FYCO1-LIR, **AM6**, and **AM10** for 72 h. Cell viability was evaluated by MTT assay. (**D**) DU145 cells were treated with FYCO1-LIR, **AM6**, and **AM10** for 72 h. Cell viability was evaluated by MTT assay. Data represent the mean values ± SEM of six biological samples (n = 6) and were analyzed by one-way analysis of variance ANOVA followed by Dunnett's post hoc test (* $p < 0.05$ vs. C; *** $p < 0.001$ vs. C; **** $p < 0.0001$ vs. C). (**E**) Analysis of ATG5 expression in DU145 and PC-3 cell lines. Tubulin expression was evaluated as a protein control.

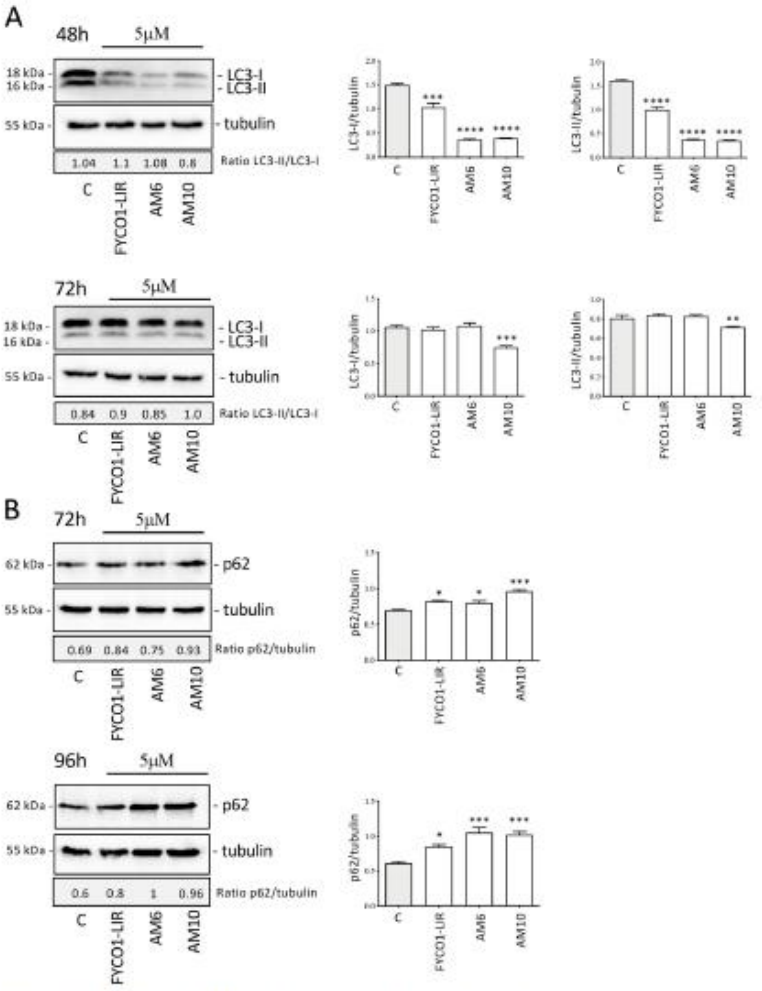



**Figure 6.** Effect of FYCO1-LIR analogs on autophagy in PC-3 cells. (**A**) Analysis of LC3 expression after treatment with FYCO1-LIR, **AM6**, and **AM10** (48 h and 72 h). (**B**) Analysis of p62 expression after treatment with FYCO1-LIR, **AM6**, and **AM10** (72 h and 96 h). Tubulin expression was evaluated as a protein control. Relative optical density was quantified by ImageJ software (version 1.54g). WB was performed independently three times, and a representative blot is presented. Data represent the mean values ± SEM and were analyzed by one-way analysis of variance ANOVA followed by Dunnett's post hoc test (* $p < 0.05$ vs. C; ** $p < 0.01$ vs. C; *** $p < 0.001$ vs. C; **** $p < 0.0001$ vs. C).

An important problem of tumor is represented by therapy resistance [38]. In fact, in prostate cancer, numerous studies have investigated the implications of autophagy in resistance to hormonal therapies or chemotherapeutic agents. Abiraterone and Enzalutamide, currently employed in CRPC therapy, activate an autophagic response that reduces their effectiveness [39]. Docetaxel (**Doc**), a chemotherapeutic drug which inhibits microtubule depolymerization, represents the first-line treatment for metastatic CRPC. The treatment is effective in the early stages, but over time its efficacy is drastically reduced [40,41]. For this

reason, a lot of research was focused on the molecular mechanisms involved in chemotherapy resistance including autophagy [21,22,42]. The action of **AM10** in combination with **Doc** was then analyzed in PC-3 cells to evaluate its ability to modulate the cytotoxic action of **Doc**. An MTT assay was performed to determine the dose of **Doc** capable of significantly reducing cell viability. PC-3 cells were treated with **Doc** at concentrations of 1 nM, 10 nM, 20 nM, 50 nM, and 100 nM, for an incubation time of 48 h. The results showed a significant reduction in cell viability at concentrations of 10 nM, 25 nM, 50 nM, and 100 nM, with a dose-dependent effect. The effect of simultaneous treatment with **Doc** (25 nM) and **AM10** (5 µM) on cell growth was subsequently evaluated by cell count. The results obtained showed that treatment with **AM10** increases the antitumoral activity of **Doc** in a significant manner (Figure 7A). The impact of **AM10** on **Doc**-induced autophagy activation was then explored by analysis of LC3 expression. A Western blot of LC3 shows that **AM10** treatment does not determine a change in the LC3-II/LC3-I ratio but reduces the expression of both LC3-I and LC3-II. On the contrary, **Doc** treatment determines a significant increase in LC3-II expression, indicative of autophagy activation in response to the stress induced by the compound. Combination treatment conducted simultaneously for 48 h with **Doc** and **AM10** determines a reduction in LC3-II expression compared to **Doc** alone, suggesting that **AM10** can inhibit **Doc**-induced autophagy (Figure 7B). Then, **AM10** in combination with **Doc** can counteract the autophagic activation induced by **Doc**, potentially influencing the sensitivity of PC-3 cells to chemotherapy.

It is known that autophagy and apoptosis are interconnected phenomena, and several studies have examined how autophagy may influence the ability of **Doc** to trigger apoptotic cell death on PC cells. The observations obtained are still discordant, and to date it is impossible to draw conclusions on the role of autophagy in the regulation of **Doc** cytotoxicity and in the development of resistance in CRPC cells [43–46]. For this reason, we evaluated the activation of apoptosis by analysis of caspase-3 cleavage after simultaneous treatment with **Doc** and **AM10**. Figure 7C highlights that **AM10** does not induce caspase-3 activation, unlike **Doc** which is known to trigger apoptotic cell death by activating the executor caspase-3. The combined treatment enhances the expression of cleaved caspase-3, demonstrating that the inhibition of **Doc**-induced autophagy by **AM10** enhances the apoptotic cell response. Therefore, **AM10** acts as sensitizing chemotherapy rather than directly promoting apoptosis (Figure 7C). Consequently, we can affirm that **AM10**, by inhibiting the pro-survival autophagy induced by **Doc**, could represent a therapeutic opportunity to enhance the efficacy of **Doc** and decrease the dose of taxanes that are responsible for numerous side effects and reduce the resistance to this chemotherapy in CRPC.

Many studies have analyzed the nature of autophagy activated by **Doc** in PC and the impact of autophagy inhibitors on cell proliferation. The results obtained using 3-Methyladenine (3-MA) in association with **Doc** were contradictory. Hu and collaborators reported how 3-MA enhanced the cytotoxic action of **Doc** [46], while other studies showed that 3-MA decreased the chemotherapy efficacy [43,47,48]. Finally, our study showed that 3-MA does not modify the cytotoxicity of **Doc** in PC-3 cells [45]. A more convincing result was obtained using the autophagy inhibitor CQ which enhanced the action of **Doc** by reducing resistance to chemotherapy [49].

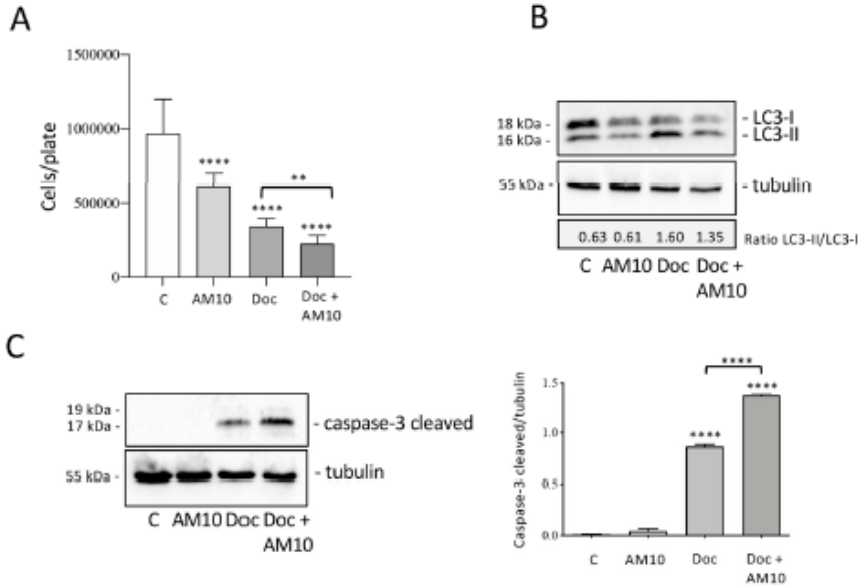



**Figure 7.** Effect of Docetaxel (**Doc**) in combination with AM10 on PC-3 cell proliferation and death. (**A**) PC-3 cells were simultaneously treated with **Doc** (25 nM) and **AM10** (5 mM) for 48 h. Cell growth was evaluated by cell count. Data represent the mean values ± SEM of four biological samples (n = 4) and were analyzed by one-way analysis of variance ANOVA followed by Tukey's post hoc test (**** $p < 0.0001$ vs. C; ** $p < 0.01$). (**B**) Analysis of LC3 expression after treatment with **Doc** (25 nM) and **AM10** (5 mM) for 48 h. Tubulin expression was evaluated as protein control. (**C**) Analysis of caspase-3 cleaved after treatment with **Doc** (25 nM) and **AM10** (5 mM) for 48 h. The tubulin expression was evaluated as a protein control. WB was performed independently three times, and a representative blot is presented. Relative optical density was quantified by ImageJ software (version 1.54g). Data represent the mean values ± SEM and were analyzed by one-way analysis of variance ANOVA followed by Tukey's post hoc test (**** $p < 0.0001$ vs. C).

***Biological assays on different cancer cell lines.*** The cell viability of different concentrations of FYCO1-LIR, **AM6**, and **AM10** (from 0.0025 to 5 µM) peptides was evaluated with an MTS assay on the MCF-7, A549, and A375 cancer cell lines. Interestingly, as shown in Figure 8, a concentration-dependent reduction in cell viability (expressed as percentage % of viable cells) was observed in all tested cell lines as compared to the untreated control cells. In more depth, the results showed that MCF-7 cells are more responsive than the other tested cell lines. Overall, these findings support the effectiveness of these compounds in tumor cell lines with various origins and characteristics (Figure 8). Additionally, considering these results suggests that the AM10 peptide could have interesting application in cancer therapy; nevertheless, further studies should be conducted to better understand the translational aspects of the tested therapy.

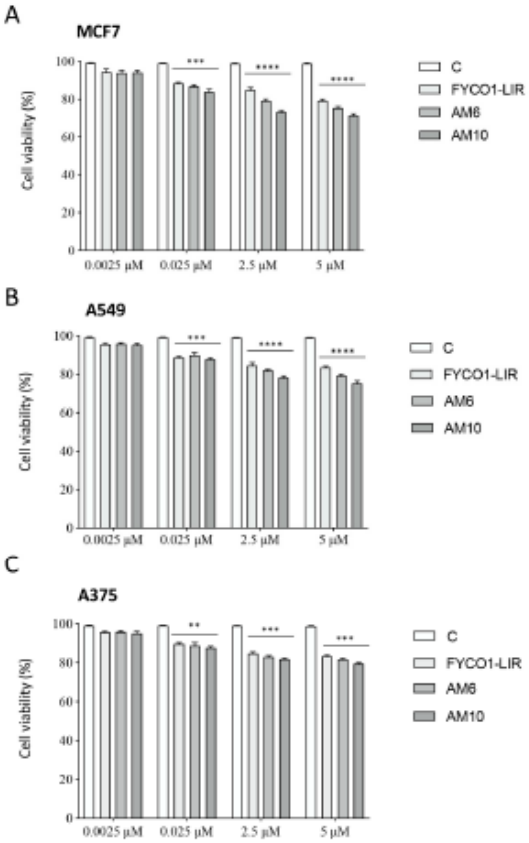



**Figure 8.** Effect of FYCO1-LIR analogs on different cancer cell lines. MCF-7 (**A**), A549 (**B**), and A375 (**C**) cells were treated with FYCO1-LIR, **AM6**, and **AM10** for 72 h. Cell viability was evaluated by MTS assay. Data the represent the mean values ± SEM of four biological samples (n = 4) and were analyzed by one-way analysis of variance ANOVA followed by Tukey's post hoc test (** $p < 0.01$ vs. C; *** $p < 0.001$ vs. C; **** $p < 0.0001$ vs. C).

## 3. Materials and Methods

### *3.1. Computational Design of FYCO1 Analogs*

The starting computational model of LC3B was generated using the 3D coordinates from the LC3B/FYCO1-LIR complex (PDB accession code 5D94 [26]). The complex model was optimized using the Protein Preparation Wizard in Maestro (release 2021–2, Schrödinger, LLC, New York, NY, USA), which included the following: residue protonation state assignment at pH 7.4, residue verification, clash resolution, and application of the OPLS4 force field. Then, the protein-ligand complex was solvated in a cubic box of TIP3P water molecules and subjected to energy minimization, followed by 250 ns MD simulations using the Desmond algorithm of Maestro (release 2021–2, Schrödinger, LLC, New York, NY, USA) [50]. The peptide stability in complex with LC3B was assessed using the "Simulation Interactions Diagram" tool. The Cα atoms RMSD and RMSF graphs of all the simulated systems are available in the Supplementary Materials, Figure S1 and Figure S2, respectively. The peptide binding free energy value was calculated using the Prime MM-GBSA algo-

rithm in Maestro (release 2021–2, Schrödinger, LLC, New York, NY, USA) [51], employing the single-trajectory approach. The resulting binding free-energy value was designated as ΔG* [24,31] and calculated for all peptides (Table 1). The peptides of the AM series investigated in this paper (Table 1) were manually built starting from the LC3B/FYCO1-LIR complex and using the "mutate residue" and drawing tools available in Maestro (release 2021–2, Schrödinger, LLC, New York, NY, USA). The ΔG* values of the FYCO1-LIR analogs were calculated using the computational protocol adopted for the reference FYCO1-LIR peptide. The LC3B/Comb1 complex was generated through a three-step protocol:

1. Structural Alignment: The LC3B/FYCO1-LIR (PDB: 5D94 [26]) and LC3A/FYCO1 (PDB: 5CX3 [52]) X-ray structures were superimposed to establish a common spatial framework.
2. Peptide Transfer and Adaptation: The DDAVFDIITDEELCQIQESG peptide from the LC3A/FYCO1 complex was transferred onto the LC3B surface. This transfer was feasible because the FYCO1 segment (DAVFDIITDEEL) in the LC3B complex aligned perfectly with its counterpart in the LC3A structure. The sequence was then manually adjusted to match Comb1.
3. Refinement: The resulting LC3B/Comb1 model was optimized using energy minimization and MD simulations, following the protocol previously described.

### 3.2. Synthesis of Peptides

All material and reagents were purchased by Sigma Aldrich, if not otherwise stated. Peptides were assembled on a 2-CTC resin by stepwise solid phase Fmoc-chemistry in a 0.15 mmol scale. Concentrations of 0.5 M Oxyma and 0.5 M DIC were used as activators, while a 20% piperidine solution in DMF was used for Fmoc-removal. Upon iterative chain assembly, peptides were cleaved off the resin by treatment with a TFA-based mixture (92.5% TFA, 2.5% TIS, 2.5% thioanisole, 2.5% water), and precipitated in cold ether. Crude peptides were recovered by centrifugation, and HPLC-purified (C-18 column, Phenomenex). For cyclization purposes, peptides were oxidized by dissolution at 100 mM in a phosphate buffer (pH = 7.8), where 1.1 eq. of 10 mM $H_2O_2$ were added. The reaction was HPLC-monitored until completion and the resulting product HPLC-purified. For linker-based cyclization, to peptide solutions in 50:50 $NaHCO_3$ aq./acetonitrile (pH = 8.0), 1.5 eq. of bis-alkylating reagent was added. The reaction was HPLC-monitored until completion and the resulting product HPLC-purified. The mass spectra and HPLC graphs of the FYCO1-LIR, **AM2**, **AM6**, **AM7**, and **AM10** peptides are available in the Supplementary Materials, Figure S3 and Figure S4, respectively.

### 3.3. Biophysical Assays

The interaction between peptides (FYCO1-LIR, **AM2**, **AM6**, **AM7**, **AM10**) and the LC3B protein was evaluated using the Monolith NT.115$^{Pico}$ instrument (NanoTemper Technologies GmbH, München, Germany). This technique enables the determination of the dissociation constant ($K_d$) across a concentration range spanning from 1 picomolar (pM) to the millimolar (mM) level. The experimental conditions mirrored those in our previously published work [31]. In summary, His-tagged human recombinant LC3B (Catalog No. 14555-H07E, Sino Biological, Beijing, China) was labelled with the His-Tag Labelling Kit RED-tris-NTA 2nd Generation (Product No. MO-L018, NanoTemper Technologies GmbH, München, Germany) for 30 min at room temperature. A constant concentration of red-labelled LC3B (10 nM) was incubated with sixteen 1:1 serial dilutions of peptides (concentration details are consultable in the Supplementary Materials, Table S1). PBS-T (phosphate-buffered saline + 0.05% Tween™ 20), from NanoTemper Technologies, with 2.5% dimethyl sulfoxide (DMSO) (Product No. D8418; Sigma-Aldrich, Saint Louis, MO,

USA) was used as buffer. After 15 min of incubation at room temperature, samples were loaded into standard capillaries (Product No. MO-K022; NanoTemper Technologies) and analyzed at 25 °C with 20% excitation power and medium MST power (40%) using the "Binding Affinity" mode available in the MO.Control v1.6 software (NanoTemper Technologies GmbH, München, Germany). Prior to $K_d$ determination, peptide auto-fluorescence was assessed. Data analysis was performed using MO.Affinity Analysis v2.3 software (NanoTemper Technologies GmbH, München, Germany) applying the $K_d$ model for fitting the binding curve, while the figures were generated using GraphPad Prism v8.0.2 software (GraphPad, Boston, MA, USA). In the case of **AM7**, the three highest concentrations (125, 62.5, and 31.25 µM) were excluded from analysis due to aggregation phenomena, as evidenced in the Supplementary Materials, Figure S5D. The "Capillary Scan" graphs, showing the fluorescence homogeneity in the capillaries of all biophysically tested peptides, are provided in the Supplementary Materials, Figure S5.

### 3.4. Cell Lines

The study was conducted on several cancer cells. Human CRPC cell lines (PC-3 and DU145) were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA) and cultured at 37 °C and 5% $CO_2$ in RPMI 1640 (EuroClone, Milano, Italy) supplemented, respectively, with 7.5% (PC3) and 5% (DU145) FBS (Gibco, ThermoFisher Scientific, Waltham, MA, USA), 1% L-glutamine, and antibiotics (100 IU/mL penicillin G). Luminal A MCF-7 breast cancer cells (ATCC, Manassas, VA, USA) were cultured in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM-F12, Gibco Laboratories, Waltham, MA, USA) supplemented with 10% fetal bovine serum (FBS, Gibco Laboratories, Waltham, MA, USA), 1% penicillin/streptomycin (Gibco Laboratories), and 1% L-glutamine (Gibco Laboratories). A549 human lung cancer cell line (ATCC, Manassas, VA, USA) were cultured in Dulbecco's modified eagle medium (DMEM, Lonza, Switzerland) supplemented with 10% FBS (Gibco Laboratories, Waltham, MA, USA), 1% of 100 u/mL penicillin/streptomycin (Gibco Laboratories), and 1% L-glutamine (Gibco Laboratories). The A375 melanoma cell line (ATCC, Manassa, VA, USA), derived from skin lesions was cultured in RPMI 1640 media (Gibco Laboratories, Waltham, MA, USA) supplemented with 1% of penicillin/streptomycin (Gibco Laboratories, Waltham, MA, USA), 1% L-glutamine (Gibco Laboratories, Waltham, MA, USA), and 10% FBS (Gibco Laboratories).

### 3.5. Cell Viability Studies

For viability studies, MTS (Cell Titer 96 Aqueous One Solution Cell Proliferation Assay) (Promega, Nacka, Sweden) and 3-(4,5-dimethylthiazole-2-yl)-2,5-diphenyltetrazolium bromide (MTT) (Sigma-Aldrich, St. Louis, MO, USA) assays were conducted. PC-3 and DU145 cells were plated at the density of $3 \times 10^4$ cells/well in 24-well plates. After 48 h, cells were treated with FYCO1-LIR, **AM6**, and **AM10** at 0.0025 µM, 0.025 µM, 0.25 µM, 2.5 µM, and 5 µM doses for 72 h. At the end of the treatments the medium was replaced with MTT solution (0.5 mg/mL) in RPMI without phenol red and FBS. Following 30–45 min of incubation at 37 °C, the precipitate was dissolved with isopropanol. The absorbance ($\lambda$ = 550 nm) was measured by an EnSpire Multimode Plate reader (Perkin Elmer, Milano, Italy). The absorbance value of untreated cells was set at 100% (control), and the viability of treated cells was expressed as a percentage of the control. Three independent experiments were performed for each condition. MCF-7, A549, and A375 cells were seeded at a density of $1 \times 10^4$ cells/well in 96-well plates and maintained under standard growth conditions. After 24 h, the cells were treated with FYCO1-LIR, **AM6** and **AM10** at 0.0025 µM, 0.025 µM, 2.5 µM, and 5 µM doses. After 72 h, cell viability was assessed using an MTS assay according to the manufacturer's protocol using a 96-well-plate spectrophotometer (Varioskan

Flash Multimode Reader; ThermoFisher Scientific, Waltham, MA, USA) set at $\lambda = 490$ nm. The absorbance value of untreated cells was set at 100% (control), and the viability of treated cells was expressed as a percentage of the control. Three independent experiments were performed for each condition.

### 3.6. Western Blot (WB) Assay

PC-3 cells were plated at $2 \times 10^5$ cells/dish in 6 cm dishes and treated with different compounds. At the end of the treatments the cells were lysed in RIPA buffer. Protein extracts (15–35 µg) were resuspended in presence of reducing Sample buffer (Bio-Rad Laboratories, Segrate, Milano, Italy), heated at 95 °C for 5 min and separated by SDS-PAGE WB. Proteins were transferred onto nitrocellulose or PVDF membranes. After blocking with nonfat dried milk, membranes were incubated with anti-ATG5 (1:1000) (#12994) Cell Signaling Technology Inc. (Boston, MA, USA), anti-LC3 (1:1000) (L8918) (Sigma-Aldrich), anti-SQSTM1/p62 (1:2000) (PA5-20839) (Thermo Fisher Scientific, Waltham, MA, USA), anti-cleaved-caspase-3 (1:500) (#9664) Cell Signaling Technology Inc. primary antibodies overnight at 4 °C. Horseradish peroxidase (HRP) conjugated secondary anti-rabbit or anti-mouse antibodies were used for 1 h at room temperature and the membranes were processed using chemiluminescence kit Cyanagen Ultra (Cyanagen, Bologna, Italy). In each WB experiment alpha-tubulin expression (T6199) (Sigma-Aldrich, St. Louis, MO, USA) was evaluated as a protein control. Relative optical density of the bands was assessed by ImageJ software (version 1.54g). Uncropped WB images are available in the Supporting Information.

### 3.7. Cell Proliferation Studies

PC-3 cells were plated at $2 \times 10^5$ cells/dish in 6 cm dishes and treated simultaneously with **AM10** (5 mM) in combination with **Doc** (25 nM) for 48 h. The cells were then collected and counted using a hemocytometer.

### 3.8. Statistical Analysis

Statistical analysis of the results was performed by one way performed of variance (ANOVA) followed by Dunnett's test or Tukey's multiple comparison post-test. Prism software was used for the analyses (Prism 8 for Mac OS version 8.2.1, GraphPad Software, San Diego, CA, USA).

## 4. Conclusions

This work began by studying the sequence of FYCO1-LIR, a peptide known to bind LC3B, and applied a computational approach combining MD simulations and MM-GBSA calculations to predict the binding affinity of novel peptides generated through sequence mutations. Our methodology significantly enhanced both the stability and affinity of the redesigned peptides (Table 1, Figures 1 and 2). The electrostatic interactions between the new peptides and LC3B's positively charged surface—defined by the residues K49, K51, R69, and R70—correlated with low predicted $\Delta G^*$ values, explaining their improved binding efficiency. Structural rigidity conferred by disulfide bridges in the peptide sequences reduced fluctuations at LC3B's LIR binding site while enhancing the occupancy of the HP1 hydrophobic basin (P7, I23, P32, I34, L53, and F108). Notably, the top analog, **AM10** (a decamer), demonstrated a binding mode comparable to Comb1 (eicosapeptide, $K_d = 0.12$ µM), despite its shorter length [30]. Experimental validation via MST confirmed **AM10**'s superior affinity ($K_d = 0.04$ µM), representing an approximately 80-fold improvement over the parent FYCO1-LIR peptide ($K_d = 3.1$ µM). Biological assays in PC-3 cells revealed that **AM6** and **AM10** diminished cell viability and inhibited autophagosome formation and autophagic flux, as evidenced by the expression of LC3-I, LC3-II, and p62.

Furthermore, co-treatment with **Doc** and **AM10** determined an enhancement of apoptosis, demonstrating that **AM10** is able to counteract **Doc**-induced autophagy which causes a reduction in chemotherapy-induced apoptotic cell death. Overall, the results obtained may suggest the use of **AM10** to reduce resistance to pharmacological therapies that activate a cellular protective pro-survival autophagic response.

The efficacy of **AM10** against prostate cancer and other different tumoral cells suggests its use as a novel autophagy modulator for cancer treatment, particularly if intracellular delivery via nanocarriers or liposomes is optimized. Moreover, since **AM10** can bind to the LC3B area close to the arginine-rich motif (residues 68 to 70), the one regulating the mRNA degradation during autophagy [53], our results open the way to design new peptide nucleic acids (PNAs) for RNA-based therapeutics, which could represent a new class of autophagy modulators.

**Supplementary Materials:** The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/ijms26115365/s1.

**Author Contributions:** Conceptualization, E.M.A.F. and G.G.; methodology, E.M.A.F., J.S., R.M.M. and M.G.; investigation, M.A., E.M.A.F., E.G.M., M.M.M., C.P., A.G., J.S. and M.G.; resources, G.G., A.G. and R.M.M.; data curation, E.M.A.F. and G.G.; writing—original draft preparation, all authors; writing—review and editing, E.M.A.F. and G.G.; funding acquisition, G.G., R.M.M., M.G. and A.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by National Center for Gene Therapy and Drugs Based on RNA Technology—MUR (Project no. CN_00000041), funded by the NextGeneration EU program.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article and Supplementary Materials.

**Acknowledgments:** We'd like to extend our special thanks to Arianna Maggi for her support with this work. Moreover, G.G. and E.M.A.F. would like to thank INDACO for providing high-performance computing resources and support.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Behrends, C.; Sowa, M.E.; Gygi, S.P.; Harper, J.W. Network Organization of the Human Autophagy System. *Nature* **2010**, *466*, 68–76. [CrossRef] [PubMed]
2. Xie, Z.; Klionsky, D.J. Autophagosome Formation: Core Machinery and Adaptations. *Nat. Cell Biol.* **2007**, *9*, 1102–1109. [CrossRef] [PubMed]
3. Huang, J.; Brumell, J.H. Bacteria-Autophagy Interplay: A Battle for Survival. *Nat. Rev. Microbiol.* **2014**, *12*, 101–114. [CrossRef]
4. Jacquet, M.; Guittaut, M.; Fraichard, A.; Despouy, G. The Functions of Atg8-Family Proteins in Autophagy and Cancer: Linked or Unrelated? *Autophagy* **2021**, *17*, 599–611. [CrossRef] [PubMed]
5. Hain, A.U.; Miller, A.S.; Levitskaya, J.; Bosch, J. Virtual Screening and Experimental Validation Identify Novel Inhibitors of the Plasmodium Falciparum Atg8-Atg3 Protein-Protein Interaction. *ChemMedChem* **2016**, *11*, 900–910. [CrossRef]
6. Villa, S.; Legnani, L.; Colombo, D.; Gelain, A.; Lammi, C.; Bongiorno, D.; Ilboudo, D.P.; Mcgee, K.E.; Bosch, J.; Grazioso, G. Structure-Based Drug Design, Synthesis and Biological Assays of P. Falciparum Atg3-Atg8 Protein-Protein Interaction Inhibitors. *J. Comput. Aided Mol. Des.* **2018**, *32*, 473–486. [CrossRef]
7. Wong, Y.C.; Holzbaur, E.L.F. Optineurin Is an Autophagy Receptor for Damaged Mitochondria in Parkin-Mediated Mitophagy That Is Disrupted by an ALS-Linked Mutation. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E4439–E4448. [CrossRef]
8. McLendon, P.M.; Ferguson, B.S.; Osinska, H.; Shenuarin Bhuiyan, M.; James, J.; McKinsey, T.A.; Robbins, J. Tubulin Hyperacetylation Is Adaptive in Cardiac Proteotoxicity by Promoting Autophagy. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E5178–E5186. [CrossRef]
9. Chen, M.; Hong, M.J.; Sun, H.; Wang, L.; Shi, X.; Gilbert, B.E.; Corry, D.B.; Kheradmand, F.; Wang, J. Essential Role for Autophagy in the Maintenance of Immunological Memory against Influenza Infection. *Nat. Med.* **2014**, *20*, 503–510. [CrossRef]

10. Quan, W.; Lim, Y.M.; Lee, M.S. Role of Autophagy in Diabetes and Endoplasmic Reticulum Stress of Pancreatic β-Cells. *Exp. Mol. Med.* **2012**, *44*, 81–88. [CrossRef]
11. Sarparanta, J.; García-Macia, M.; Singh, R. Autophagy and Mitochondria in Obesity and Type 2 Diabetes. *Curr. Diabetes Rev.* **2016**, *13*, 352–369. [CrossRef] [PubMed]
12. Mao, Y.; Yu, F.; Wang, J.; Guo, C.; Fan, X. Autophagy: A New Target for Nonalcoholic Fatty Liver Disease Therapy. *Hepat. Med.* **2016**, *ume 8*, 27–37. [CrossRef]
13. Amaravadi, R.K.; Lippincott-Schwartz, J.; Yin, X.M.; Weiss, W.A.; Takebe, N.; Timmer, W.; DiPaola, R.S.; Lotze, M.T.; White, E. Principles and Current Strategies for Targeting Autophagy for Cancer Treatment. *Clin. Cancer Res.* **2011**, *17*, 654–666. [CrossRef] [PubMed]
14. Pérez-Hernández, M.; Arias, A.; Martínez-García, D.; Pérez-Tomás, R.; Quesada, R.; Soto-Cerrato, V. Targeting Autophagy for Cancer Treatment and Tumor Chemosensitization. *Cancers* **2019**, *11*, 1599. [CrossRef] [PubMed]
15. Giannopoulos, S.; Bozkus, C.C.; Zografos, E.; Athanasiou, A.; Bongiovanni, A.M.; Doulaveris, G.; Bakoyiannis, C.N.; Theodoropoulos, G.E.; Zografos, G.C.; Witkin, S.S.; et al. Targeting Both Autophagy and Immunotherapy in Breast Cancer Treatment. *Metabolites* **2022**, *12*, 966. [CrossRef]
16. Pimentel, J.M.; Zhou, J.Y.; Wu, G.S. Autophagy and Cancer Therapy. *Cancer Lett.* **2024**, *605*, 217285. [CrossRef]
17. Niu, X.; You, Q.; Hou, K.; Tian, Y.; Wei, P.; Zhu, Y.; Gao, B.; Ashrafizadeh, M.; Aref, A.R.; Kalbasi, A.; et al. Autophagy in Cancer Development, Immune Evasion, and Drug Resistance. *Drug Resist. Updates* **2025**, *78*, 101170. [CrossRef]
18. Mortezavi, A.; Salemi, S.; Rupp, N.J.; Rüschoff, J.H.; Hermanns, T.; Poyet, C.; Randazzo, M.; Simon, H.U.; Moch, H.; Sulser, T.; et al. Negative LC3b Immunoreactivity in Cancer Cells Is an Independent Prognostic Predictor of Prostate Cancer Specific Death. *Oncotarget* **2017**, *8*, 31765–31774. [CrossRef]
19. Holah, N.S.; El-Dien, M.M.S.; Mahmoud, S.F. Expression of Autophagy Markers Beclin1 and LC3B in Prostatic Carcinoma: An Immunohistochemical Case-Control Study. *Iran. J. Pathol.* **2022**, *17*, 75–84. [CrossRef]
20. Lamprou, I.; Tsolou, A.; Kakouratos, C.; Mitrakas, A.G.; Xanthopoulou, E.T.; Kassela, K.; Karakasiliotis, I.; Zois, C.E.; Giatromanolaki, A.; Koukourakis, M.I. Suppressed PLIN3 Frequently Occurs in Prostate Cancer, Promoting Docetaxel Resistance via Intensified Autophagy, an Event Reversed by Chloroquine. *Med. Oncol.* **2021**, *38*, 116. [CrossRef]
21. Ashrafizadeh, M.; Paskeh, M.D.A.; Mirzaei, S.; Gholami, M.H.; Zarrabi, A.; Hashemi, F.; Hushmandi, K.; Hashemi, M.; Nabavi, N.; Crea, F.; et al. Targeting Autophagy in Prostate Cancer: Preclinical and Clinical Evidence for Therapeutic Response. *J. Exp. Clin. Cancer Res.* **2022**, *41*, 105. [CrossRef] [PubMed]
22. Lemos, G.; Fernandes, C.M.A.d.S.; Silva, F.H.; Calmasini, F.B. The Role of Autophagy in Prostate Cancer and Prostatic Diseases: A New Therapeutic Strategy. *Prostate Cancer Prostatic Dis.* **2024**, *27*, 230–238. [CrossRef] [PubMed]
23. Zhou, M.; Wang, R. Small-Molecule Regulators of Autophagy and Their Potential Therapeutic Applications. *ChemMedChem* **2013**, *8*, 694–707. [CrossRef] [PubMed]
24. Fassi, E.M.A.; 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**, *23*, 5070. [CrossRef]
25. Johansen, T.; Lamark, T. Selective Autophagy: ATG8 Family Proteins, LIR Motifs and Cargo Receptors. *J. Mol. Biol.* **2020**, *432*, 80–103. [CrossRef]
26. Olsvik, H.L.; Lamark, T.; Takagi, K.; Larsen, K.B.; Evjen, G.; Øvervatn, A.; Mizushima, T.; Johansen, X.T. FYCO1 Contains a C-Terminally Extended, LC3A/B-Preferring LC3-Interacting Region (LIR) Motif Required for Efficient Maturation of Autophagosomes during Basal Autophagy. *J. Biol. Chem.* **2015**, *290*, 29361–29374. [CrossRef]
27. Pankiv, S.; Alemu, E.A.; Brech, A.; Bruun, J.-A.; Lamark, T.; Øvervatn, A.; Bjørkøy, G.; Johansen, T. FYCO1 Is a Rab7 Effector That Binds to LC3 and PI3P to Mediate Microtubule plus End–Directed Vesicle Transport. *J. Cell Biol.* **2010**, *188*, 253–269. [CrossRef]
28. Ma, J.; Becker, C.; Reyes, C.; Underhill, D.M. Cutting Edge: FYCO1 Recruitment to Dectin-1 Phagosomes Is Accelerated by Light Chain 3 Protein and Regulates Phagosome Maturation and Reactive Oxygen Production. *J. Immunol.* **2014**, *192*, 1356–1360. [CrossRef]
29. Raiborg, C.; Wenzel, E.M.; Pedersen, N.M.; Olsvik, H.; Schink, K.O.; Schultz, S.W.; Vietri, M.; Nisi, V.; Bucci, C.; Brech, A.; et al. Repeated ER–Endosome Contacts Promote Endosome Translocation and Neurite Outgrowth. *Nature* **2015**, *520*, 234–238. [CrossRef]
30. Cerulli, R.A.; Shehaj, L.; Brown, H.; Pace, J.; Mei, Y.; Kritzer, J.A. Stapled Peptide Inhibitors of Autophagy Adapter LC3B. *ChemBioChem* **2020**, *21*, 2777–2785. [CrossRef]
31. Albani, M.; Fassi, E.M.A.; Moretti, R.M.; Garofalo, M.; Montagnani Marelli, M.; Roda, G.; Sgrignani, J.; Cavalli, A.; Grazioso, G. Computational Design of Novel Cyclic Peptides Endowed with Autophagy-Inhibiting Activity on Cancer Cell Lines. *Int. J. Mol. Sci.* **2024**, *25*, 4622. [CrossRef] [PubMed]
32. Baliou, S.; Goulielmaki, M.; Ioannou, P.; Cheimonidi, C.; Trougakos, I.P.; Nagl, M.; Kyriakopoulos, A.M.; Zoumpourlis, V. Article Bromamine T (BAT) Exerts Stronger Anti-Cancer Properties than Taurine (Tau). *Cancers* **2021**, *13*, 182. [CrossRef]

33. Potapskyi, E.; Kustrzyńska, K.; Łażewski, D.; Skupin-Mrugalska, P.; Lesyk, R.; Wierzchowski, M. Introducing Bromine to the Molecular Structure as a Strategy for Drug Design. *J. Med. Sci.* **2024**, *93*, e1128. [CrossRef]
34. Xu, L.; Fan, X.; He, Y.; Xia, X.; Zhang, J. Design, Synthesis, and Biological Evaluation of Lysine-Stapled Peptide Inhibitors of P53-MDM2/MDMX Interactions with Potent Antitumor Activity In Vivo. *J. Med. Chem.* **2024**, *67*, 17893–17904. [CrossRef]
35. Walensky, L.D.; Bird, G.H. Hydrocarbon-Stapled Peptides: Principles, Practice, and Progress. *J. Med. Chem.* **2014**, *57*, 6275–6288. [CrossRef]
36. Tombling, B.; Lammi, C.; Lawrence, N.; Gilding, E.; Grazioso, G.; Craik, D.; Wang, C. Bioactive Cyclization Optimizes the Affinity of a Proprotein Convertase Subtilisin/Kexin Type 9 (PCSK9) Peptide Inhibitor. *J. Med. Chem.* **2021**, *64*, 2523–2533. [CrossRef]
37. Ouyang, D.Y.; Xu, L.H.; He, X.H.; Zhang, Y.T.; Zeng, L.H.; Cai, J.Y.; Ren, S. Autophagy Is Differentially Induced in Prostate Cancer LNCaP, DU145 and PC-3 Cells via Distinct Splicing Profiles of ATG5. *Autophagy* **2013**, *9*, 20–32. [CrossRef] [PubMed]
38. Wade, C.A.; Kyprianou, N. Profiling Prostate Cancer Therapeutic Resistance. *Int. J. Mol. Sci.* **2018**, *19*, 904. [CrossRef]
39. Loizzo, D.; Pandolfo, S.D.; Rogers, D.; Cerrato, C.; Di Meo, N.A.; Autorino, R.; Mirone, V.; Ferro, M.; Porta, C.; Stella, A.; et al. Novel Insights into Autophagy and Prostate Cancer: A Comprehensive Review. *Int. J. Mol. Sci.* **2022**, *23*, 3826. [CrossRef]
40. Bumbaca, B.; Li, W. Taxane Resistance in Castration-Resistant Prostate Cancer: Mechanisms and Therapeutic Strategies. *Acta Pharm. Sin. B* **2018**, *8*, 518–529. [CrossRef]
41. Giacinti, S.; Poti, G.; Roberto, M.; Macrini, S.; Bassanelli, M.; Di Pietro, F.; Aschelter, A.M.; Ceribelli, A.; Ruggeri, E.M.; Marchetti, P. Molecular Basis of Drug Resistance and Insights for New Treatment Approaches in MCRPC. *Anticancer. Res.* **2018**, *38*, 6029–6039. [CrossRef] [PubMed]
42. Kurganovs, N.J.; Engedal, N. To Eat or Not to Eat: A Critical Review on the Role of Autophagy in Prostate Carcinogenesis and Prostate Cancer Therapeutics. *Front. Pharmacol.* **2024**, *15*, 1419806. [CrossRef] [PubMed]
43. Pickard, R.D.; Spencer, B.H.; McFarland, A.J.; Bernaitis, N.; Davey, A.K.; Perkins, A.V.; Chess-Williams, R.; McDermott, C.M.; Forbes, A.; Christie, D.; et al. Paradoxical Effects of the Autophagy Inhibitor 3-Methyladenine on Docetaxel-Induced Toxicity in PC-3 and LNCaP Prostate Cancer Cells. *Naunyn Schmiedebergs Arch. Pharmacol.* **2015**, *388*, 793–799. [CrossRef] [PubMed]
44. Wang, Q.; He, W.Y.; Zeng, Y.Z.; Hossain, A.; Gou, X. Inhibiting Autophagy Overcomes Docetaxel Resistance in Castration-Resistant Prostate Cancer Cells. *Int. Urol. Nephrol.* **2018**, *50*, 675–686. [CrossRef]
45. Cristofani, R.; Montagnani Marelli, M.; Cicardi, M.E.; Fontana, F.; Marzagalli, M.; Limonta, P.; Poletti, A.; Moretti, R.M. Dual Role of Autophagy on Docetaxel-Sensitivity in Prostate Cancer Cells. *Cell Death Dis.* **2018**, *9*, 889. [CrossRef]
46. Hu, F.; Zhao, Y.; Yu, Y.; Fang, J.M.; Cui, R.; Liu, Z.Q.; Guo, X.L.; Xu, Q. Docetaxel-Mediated Autophagy Promotes Chemoresistance in Castration-Resistant Prostate Cancer Cells by Inhibiting STAT3. *Cancer Lett.* **2018**, *416*, 24–30. [CrossRef]
47. Zeng, J.; Liu, W.; Fan, Y.Z.; He, D.L.; Li, L. Prl-Z Increases Prostate Cancer Docetaxel Resistance by Inhibiting LKB1/AMPK-Mediated Autophagy. *Theranostics* **2018**, *8*, 109–123. [CrossRef]
48. Peng, K.; Sun, A.; Zhu, J.; Gao, J.; Li, Y.; Shao, G.; Yang, W.; Lin, Q. Restoration of the ATG5-Dependent Autophagy Sensitizes DU145 Prostate Cancer Cells to Chemotherapeutic Drugs. *Oncol. Lett.* **2021**, *22*, 638. [CrossRef]
49. Liu, B.; Miyake, H.; Nishikawa, M.; Tei, H.; Fujisawa, M. Expression Profile of Autophagy-Related Markers in Localized Prostate Cancer: Correlation with Biochemical Recurrence after Radical Prostatectomy. *Urology* **2015**, *85*, 1424–1430. [CrossRef]
50. Bowers, K.J.; Chow, E.; Xu, H.; Dror, R.O.; Eastwood, M.P.; Gregersen, B.A.; Klepeis, J.L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti, F.D.; et al. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In Proceedings of the ACM/IEEE Conference on Supercomputing (SC06), Tampa, FL, USA, 11–17 November 2006. [CrossRef]
51. Jacobson, M.P.; Friesner, R.A.; Xiang, Z.; Honig, B. On the Role of the Crystal Environment in Determining Protein Side-Chain Conformations. *J. Mol. Biol.* **2002**, *320*, 597–608. [CrossRef]
52. Cheng, X.; Wang, Y.; Gong, Y.; Li, F.; Guo, Y.; Hu, S.; Liu, J.; Pan, L. Structural Basis of FYCO1 and MAP1LC3A Interaction Reveals a Novel Binding Mode for Atg8-Family Proteins. *Autophagy* **2016**, *12*, 1330–1339. [CrossRef] [PubMed]
53. Hwang, H.J.; Ha, H.; Lee, B.S.; Kim, B.H.; Song, H.K.; Kim, Y.K. LC3B Is an RNA-Binding Protein to Trigger Rapid MRNA Degradation during Autophagy. *Nat. Commun.* **2022**, *13*, 1436. [CrossRef] [PubMed]

**Appendix 4**

**Title**: Computational characterization of peptide binding stability to HLA-C allotypes and its association with HIV-1 infection progression and HIV-1 related neurocognitive impairment

**Authors**: Voi M., Sangalli A., Milano E.G., De Martinis C., Orlandi E., Tamburin S., Mantovani E., Federico A., Lanzafame M., Lattuada E., Argañaraz G.A., Da Silva B.C.M., Da Silva Duarte A.J., Casseb J., Argañaraz E.R., Malena M., **Albani M.**, Ruggiero A., Romanelli M.G., Valenti M.T., Grazioso G. and Zipeto D.

**Status**: Published in Frontiers in Immunology (2025)

**DOI**: 10.3389/fimmu.2025.1703026

**Author Contribution**: Dr. M. Albani led the study design, performed molecular dynamics simulations, analyzed the data, and wrote the manuscript.

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 11 December 2025
DOI 10.3389/fimmu.2025.1703026



OPEN ACCESS

EDITED BY
Simone Agostini,
Fondazione Don Carlo Gnocchi Onlus (IRCCS), Italy

REVIEWED BY
Nobubelo Ngandu,
South African Medical Research Council, South Africa
Pengcheng Wei,
Guangxi University, China

*CORRESPONDENCE
Donato Zipeto
donato.zipeto@univr.it

RECEIVED 10 September 2025
REVISED 25 November 2025
ACCEPTED 26 November 2025
PUBLISHED 11 December 2025

CITATION
Voi M, Sangalli A, Milano EG, De Martinis C, Orlandi E, Tamburin S, Mantovani E, Federico A, Lanzafame M, Lattuada E, Argañaraz GA, Da Silva BCM, Da Silva Duarte AJ, Casseb J, Argañaraz ER, Malena M, Albani M, Ruggiero A, Romanelli MG, Valenti MT, Grazioso G and Zipeto D (2025) Computational characterization of peptide binding stability to HLA-C allotypes and its association with HIV-1 infection progression and HIV-1 related neurocognitive impairment. *Front. Immunol.* 16:1703026. doi: 10.3389/fimmu.2025.1703026



# Computational characterization of peptide binding stability to HLA-C allotypes and its association with HIV-1 infection progression and HIV-1 related neurocognitive impairment

Mauro Voi[1], Antonella Sangalli[1], Erica Ginevra Milano[2], Carola De Martinis[1], Elisa Orlandi[1], Stefano Tamburin[1], Elisa Mantovani[1], Angela Federico[1], Massimiliano Lanzafame[3,4], Emanuela Lattuada[3], Gustavo Adolfo Argañaraz[5], Bosco Christiano Maciel Da Silva[6], Alberto Jose Da Silva Duarte[6], Jorge Casseb[7], Enrique Roberto Argañaraz[5], Marina Malena[8], Marco Albani[2], Alessandra Ruggiero[1], Maria Grazia Romanelli[1], Maria Teresa Valenti[1], Giovanni Grazioso[2] and Donato Zipeto[1*]

[1]Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy, [2]Department of Pharmaceutical Sciences, University of Milan, Milan, Italy, [3]Unit of Infectious Diseases, Santa Chiara Hospital, Azienda Provinciale per i Servizi Sanitari, Trento, Italy, [4]Centre for Medical Sciences (CISMED), University of Trento, Trento, Italy, [5]Laboratory of Molecular Neurovirology, Department of Pharmacy, Faculty of Health Science, University of Brasilia, Brasilia, Brazil, [6]Medical Investigation Laboratory Unit 56 (LIM/56), Faculdade de Medicina FMUSP, University of São Paulo, São Paulo, Brazil, [7]Faculty of Medicine, Institute of Tropical Medicine, University of São Paulo, São Paulo, Brazil, [8]U.O.S. Infectious Diseases, Santa Maria della Misericordia Hospital AULSS5, Rovigo, Italy

**Background:** HLA-C molecules play a critical role in the immune response, particularly in antigen presentation and immune modulation.
**Methods:** To investigate the effect of the most common HLA-C allotypes on the stability of the HLA-C-β-2 microglobulin-peptide complex, we used the NetMHCpan-4.2 bioinformatic tool that predicts peptide binding to MHC class I molecules. This allowed us to predict the probability of a broad set of peptides to be naturally processed, presented on each HLA-C allotype, and ultimately recognised by the immune system, measured by EL-score. By plotting the EL-score against the percentile of the peptide's stability rank position, curves were drawn to illustrate the relative stability of the binding interaction of each HLA-C allotype tested, and the area under the curve was calculated to determine a stability score for each HLA-C variant.
**Results:** This approach permits us to greatly improve the classification of HLA-C allotypes according to their stability, overcoming the previous coarse stable and unstable binary classification. Analysis of two well-characterised HIV-1 patient cohorts, one focused on disease progression and the other on neurocognitive

10.3389/fimmu.2025.1703026

impairment, demonstrated a significant association between unstable HLA-C alleles, faster disease progression, and worse HIV-associated neurocognitive outcomes.
**Conclusions:** These findings underscore the role of HLA-C stability in AIDS progression, suggesting that profiling HLA-C stability may serve as a predictive tool for HIV-1 disease management and assessing neurocognitive risk, with potential implications in personalised medicine.

KEYWORDS

HLA-C, HIV-1, antigen presentation, immune response, neurocognitive disorders, HAND, peptide-binding stability, personalized medicine

## Introduction

The human leukocyte antigen (HLA) system is a key component of the innate and adaptive immune responses, playing a critical role in antigen presentation (1, 2). HLA-C is a class I major histocompatibility complex (MHC-I) molecule involved in the immune surveillance against infections and malignancies. Unlike HLA-A and HLA-B, HLA-C exhibits lower surface expression and distinct peptide-binding properties due to less efficient assembly and cell membrane expression (3–5), making its functional stability a critical aspect of immune response regulation (6). Despite its historically perceived lower relevance, emerging evidence has demonstrated that HLA-C plays a significant role in modulating immune responses, particularly in HIV-1 infection, influencing viral replication, immune escape mechanisms, and disease progression (7).

HLA-C is central to antigen presentation to CD8+ T cells and immune surveillance. Its genetic variability influences disease susceptibility and transplant compatibility (1, 2). Furthermore, the stability of the HLA-C-β-2 microglobulin ($\beta_2$m)-peptide complex determines antigen presentation efficiency, with less stable variants linked to impaired immune responses, increased infection risk, and reduced immunotherapy efficacy (7–9).

HLA-C expression varies significantly among individuals (10) and is modulated by genetic factors, including promoter polymorphisms, microRNA interactions, and alternative splicing (9, 11–13). Single-nucleotide polymorphisms (SNPs) in the 3' untranslated region (3'UTR) of the HLA-C gene affect its expression levels, particularly through regulation by miR-148a (14). Furthermore, HLA-C upregulation has been linked to greater immune pressure on HIV-1, driving viral evolution and immune escape mutations (10).

The efficiency of antigen presentation is directly linked to the stability of the HLA-C/peptide complex, which varies across alleles and can impact immune recognition and viral control. Unlike HLA-A and HLA-B, HLA-C displays more selective peptide-binding preferences, which impact its effectiveness in presenting viral antigens. HLA-C allotypes display considerable variation in their peptide-binding clefts, which directly influences peptide binding stability (9). These differences affect the capacity of HLA-C molecules to stabilize peptide-MHC complexes and maintain antigen presentation on the cell surface. While allotypes like HLA-C*05 and HLA-C*08 are characterized by higher binding stability, others, such as HLA-C*07, are more prone to peptide dissociation and degradation (6). Sibilio and colleagues evaluated the post-assembly stability of HLA-C variants using pulse-labelling assays in homozygous cell lines expressing eight serologically defined HLA-C alleles. Their findings revealed that some HLA-C variants, such as HLA-C*05, C*06, C*08, exhibit stronger binding to $\beta_2$m, whereas others, including HLA-C*04 and C*07, display weaker $\beta_2$m association and reduced complex stability (15).

The extensive polymorphism of HLA alleles presents challenges for experimentally characterizing HLA/peptide stability across variants *in vitro*. To overcome this limitation, several computational tools have been developed to predict the interaction between HLA and peptides (16), among which NetMHCpan stands out for its high predictive accuracy (17). This model incorporates multiple parameters, including binding affinity, stability, length, processing and presentation pathways, to estimate peptide presentation and immunogenicity (17). NetMHCpan uses artificial neural networks trained on experimental binding affinity and mass spectrometry-eluted ligand (EL) data to significantly improve the accuracy of peptide-binding predictions across different HLA alleles. Evolving through versions 4.0, 4.1 and most recently 4.2, it incorporates enhanced machine learning techniques and new training datasets to improve performance in identifying immunogenic peptides (17–19). A key innovation in recent versions is the integration of structural features to enhance the predictive power of the algorithm (18), as well as the EL-score (19), which incorporates data on naturally processed and presented peptides obtained via mass spectrometry-based ligand elution assays (20). This overcomes the limitations of the traditional IC50-based calculation of binding affinity. Therefore, the EL-score represents a composite measure that integrates binding affinity (BA), the traditional indicator of peptide-MHC interaction strength, with eluted ligand data from naturally presented peptides, enhancing biological relevance. It also incorporates peptide stability and processing parameters, reflecting the likelihood that a bound

peptide remains stably displayed on the cell surface. The EL-score aims to predict the likelihood that a peptide will be naturally processed, presented on an MHC class I molecule, and ultimately trigger immunity. Studies have shown that the EL-score provides a better correlation with immunogenicity than BA alone. For instance, Harndahl et al. demonstrated that peptide-MHC stability is a stronger predictor of T cell activation than affinity, supporting the inclusion of EL data in prediction models (8). Similarly, Rasmussen et al. reported that combining affinity and stability scores improves the identification of CTL epitopes, reinforcing NetMHCpan's value in epitope discovery for vaccine development (21). Recent improvements using deep learning frameworks and large-scale mass spectrometry have further refined EL-score predictions, supporting it as a robust metric linking computational models to experimental outcomes.

HLA-C expression levels and the stability of its association with $\beta_2$m and peptide are key factors influencing antiviral immunity. Certain HLA-C alleles are linked to better control of HIV-1 replication (10). Variations in viral load associated with different HLA-C alleles are thought to result from differences in their capacity to present HIV-derived peptides to CD8+ T cells and NK cells—key players in targeting infected cells (22). Moreover, research has shown that HLA-C expression levels and peptide-binding stability are critical determinants of immune control (23).Unstable HLA-C variants have also been linked to increased HIV-1 infectivity. Previous findings from our group indicate that these variants may dissociate from $\beta_2$m, forming free heavy chains that interact with the HIV-1 envelope glycoprotein (Env), enhancing viral infectivity (24). In contrast, stable HLA-C variants remain bound to $\beta_2$m and retain strong antigen-presenting functions, which are associated with better immune control. Thus, binding stability to peptides confers to HLA-C the ability to act as a conventional molecule involved in cellular immunity, or as an accessory factor modulating HIV-1 infectivity (25).

HIV-associated neurocognitive disorders (HAND) encompass a spectrum of subjective and objective cognitive impairments, from mild to severe, that occur in approximately 50% of HIV-infected individuals and persist despite effective antiretroviral therapy, suggesting that host genetic factors influence their onset and progression (26). HLA-C polymorphism also plays a role in HAND (27, 28). Unstable HLA-C variants may contribute to higher levels of free $\beta_2$m in cerebrospinal fluid, leading to chronic neuroinflammation and neuronal damage (29). This mechanism is supported by associations between high $\beta_2$m levels and neurodegenerative conditions, including Alzheimer's disease (30). Additionally, HLA-C*07 has been specifically linked to a higher incidence of HAND in HIV-positive individuals, further supporting the hypothesis that HLA-C variation plays a role in neurocognitive impairment (31).

In this study, we sought to thoroughly investigate HLA-C stability to evaluate its relevance to HIV-1 clinical outcomes. To these aims, we integrated bioinformatic predictions with clinical progression and neurocognitive outcomes.

# Materials and methods

## Molecular dynamics analysis

Starting from the crystal structures of five human allelic variants of HLA-C currently available on Protein Data Bank (specifically: HLA-C*03:04 - PDB ID: 1EFX; HLA-C*04:01 – PDB ID: 1QQD; HLA-C*05:01 – PDB ID: 5VGD; HLA-C*06:02 – PDB ID: 5W6A; HLA-C*07:02 – PDB ID: 5VGE), single point mutations were introduced on the co-crystallized nonapeptides using Maestro's Workspace tool 'Mutate Residue' (release 2021–2, Schrödinger, LLC, New York, NY, USA). Subsequently, ff19SB force field (32) were applied on the systems using the "Protein Preparation Wizard" available in Maestro (release 2021–2, Schrödinger, LLC, New York, NY, USA), then the complexes were immersed in TIP3P (33) water cubic boxes and their geometry was energy minimized, allowing the remainder of the system (HLA-C and $\beta_2$m) to adapt to the newly introduced oligopeptide sequences. Then, Molecular Dynamics (MD) simulations were accomplished on these systems, using the Amber24. The parameters for these MD simulations were configured as follows: 300 ns, 300 K, and 1 atm. For each system, we carried out three independent replicas. The attained trajectories were clustered, by means of Amber24's cpptraj tool (34), to identify the different families of complex conformations and to identify the most populated ones. On these, the Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) approach was applied to decompose the pairwise binding energetic contributions within the HLA/$\beta_2$m interactions.

## Peptide and HLA-C stability binding analysis

Peptides binding to the 21 most frequent HLA-C allotypes in the human population (C*01:02, C*02:02, C*03:02, C*03:03, C*03:04, C*04:01, C*04:03, C*05:01, C*06:02, C*07:01, C*07:02, C*07:04, C*08:01, C*08:02, C*12:02, C*12:03, C*14:02, C*14:03, C* 15:02, C* 16:01, C* 17:01) were experimentally validated by Sarkizova et al. (35). The final number of selected peptides was 36070, with a median of 1403 (range 730-3311) (Table 1). Their database cumulatively covers 95.8% of individuals worldwide based on allele frequencies. For our analysis, we focused on peptides of 8–12 amino acids, as they comprised most of the initial list and were the ones preferentially bound by MHC-I complexes (36). Allele frequencies were reported according to Sarkizova et al. (35). The Eluted Ligand (EL) score for each peptide binding to its specific HLA-C allotype was determined using NetMHCpan-4.2 (https://services.healthtech.dtu.dk/services/NetMHCpan-4.2/) with default parameters. Peptides were then ranked by NetMHCpan-4.2 EL-score for each HLA-C allele. The ranking was expressed in percentiles (% Ranking) to account for different peptide pool sizes. The resulting EL-score versus % Ranking curves were used to calculate the area under the curve (AUC) to determine a stability score for each HLA-C allotype considered (Figure 1).

10.3389/fimmu.2025.1703026

TABLE 1 The most frequent human HLA-C allotypes and their calculated stability score.

| Allotype | Peptides number | Frequency | Stability score (AUC) | Weighted average |
|---|---|---|---|---|
| C*01:02 | 1357 | 0.085 | 63.87 | – |
| C*02:02 | 930 | 0.028 | 56.76 | – |
| C*03:02 | 1194 | 0.025 | 65.09 | *65.52* |
| C*03:03 | 2123 | 0.056 | 63.17 | |
| C*03:04 | 2356 | 0.091 | 67.08 | |
| C*04:01 | 1854 | 0.112 | 67.28 | *68.06* |
| C*04:03 | 1038 | 0.019 | 72.67 | |
| C*05:01 | 1442 | 0.026 | 84.84 | - |
| C*06:02 | 1324 | 0.062 | 56.13 | - |
| C*07:01 | 808 | 0.069 | 29.74 | *40.96* |
| C*07:02 | 1072 | 0.131 | 50.45 | |
| C*07:04 | 730 | 0.015 | 9.70 | |
| C*08:01 | 1802 | 0.045 | 35.23 | - |
| C*08:02 | 3148 | 0.020 | 78.34 | - |
| C*12:02 | 1403 | 0.032 | 52.20 | - |
| C*12:03 | 2175 | 0.020 | 50.34 | - |
| C*14:02 | 1371 | 0.025 | 77.80 | *76.05* |
| C*14:03 | 2784 | 0.015 | 73.12 | |
| C*15:02 | 3311 | 0.034 | 57.39 | – |
| C*16:01 | 2883 | 0.024 | 51.12 | – |
| C*17:01 | 965 | 0.019 | 55.18 | – |

For the 21 most frequent HLA-C allotypes (first column), the number of specific binding peptides (8–12 mer) identified by Sarkizova et al. (35) is reported (second column) alongside their frequency in the human population (third column). The area under the curve (AUC) calculation (fourth column) provides a stability score for each allotype. When second-digit typing was not possible, an average stability score weighted by frequencies was calculated (fifth column).

## Patient cohorts

Two groups of HIV-1 infected individuals were enrolled in this study and categorized based on disease progression and neurocognitive status.

The first cohort included 47 patients classified as progressors (P; >10,000 copies HIV genome/mm$^3$, and ≤200 CD4+ T lymphocytes/mm$^3$) and 37 patients classified as long-term non-progressors (LTNPs; <10,000 copies HIV genome/mm$^3$, and ≥400 CD4+ T lymphocytes/mm$^3$). HIV viral-load measurements were obtained periodically throughout each patient's follow-up period (defined as the interval between initial infection detection and the end of monitoring). LTNPs were followed for a minimum of seven years. Progressors were observed until their CD4$^+$ T-cell count declined to ≤200 cells/mm$^3$, at which point antiretroviral treatment was initiated and sample collection was discontinued. All subjects gave informed consent, and the research protocols were approved by the relevant institutional review boards and research ethics committees (Ethics Committee of the Health Department of the Federal District (#066/07); and the Ethics Committee for Analysis of the public network of the Federal District, Brazil; and the Ethics Committee for Analysis of Research Projects (CAPPesq) of Hospital das Clinicas HCFMUSP (#CAPPesq #0306/10, Online registration #5867), from the Faculty of Medicine at the University of São Paulo, Brazil. The cohort is unique since the study participants had not received any treatment at the time of sampling. Further details on the patient cohort are reported in our previous study (37).

The second cohort included 57 patients referred by the Infectious Diseases Outpatient Clinics of the Verona University Hospital a) aged < 67 years, b) on dual/triple antiretroviral therapy regimens, and with c) stable suppressed plasma viremia. Participants with a history of recurrent drug abuse, current drug addiction, or other neurological diseases (e.g., history of cerebrovascular events) that could cause cognitive decline were excluded. Each patient underwent a comprehensive neuropsychological test battery designed to measure the cognitive domains recommended by the three main diagnostic guidelines (i.e., attention, executive function, learning and memory, language, speed of processing, complex motor skills) (26). The impact of cognitive difficulties on the ability to perform everyday activities was also assessed. As neuropsychiatric symptoms frequently occur alongside HIV infection, anxiety and depression were evaluated.

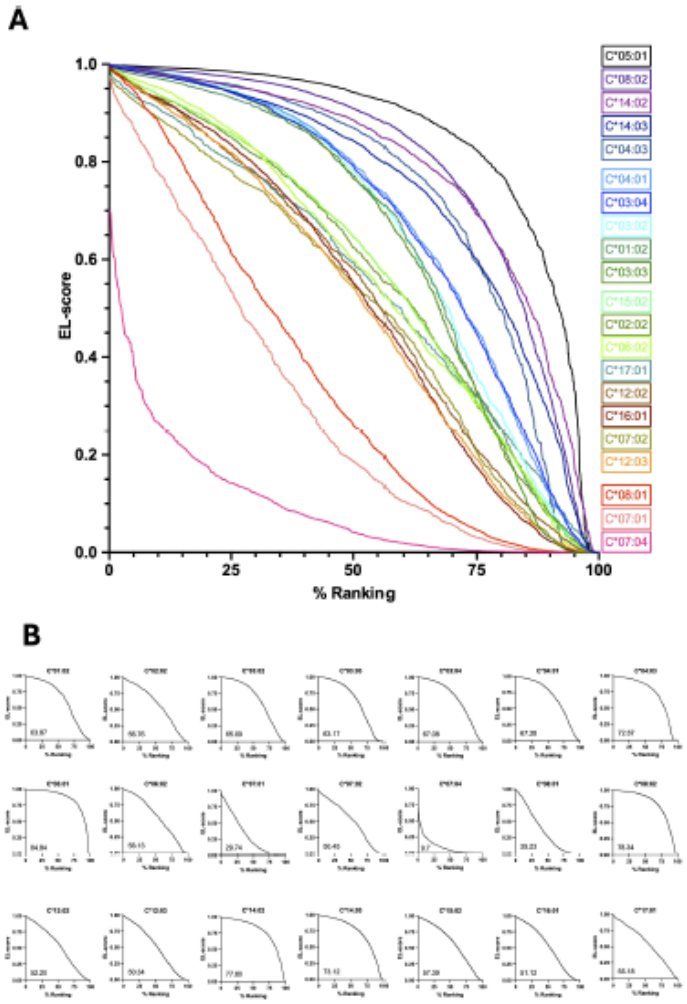



FIGURE 1
EL-score vs percentile ranking curves for each HLA-C/peptide pool. Individual EL-score values of each specific HLA-C allotype and its corresponding peptides, as predicted by NetMHCpan-4.2, were ranked and plotted against the percentile of the peptides' ranking position, resulting in distinct stability distribution profiles for each HLA-C allele. HLA-C *05:01, C*08:02, C*14:02, C*14:03 and C*04:03 clearly show a strong binding trend with most peptides, whereas HLA-C*07:04, C*07:01 and C*08:01 exhibit the lowest EL-score values for most peptides. The curves are shown overlapped in (**A**), where each curve is color-coded according to the corresponding HLA-C allotype, and separately in (**B**), where the area under the curve for each HLA-C allotype is reported in the lower left corner of each plot.

The Frascati criteria (38) were used as the gold standard for diagnosing HIV-associated neurocognitive disorders (HAND) and for dividing patients into HAND-positive (n=16) and HAND-negative (n=41) groups. Patients reporting subjective cognitive complaints were added to the first group, in line with the most recent recommendations provided by the International HIV-Cognition Working Group, which highlight the importance of changes in cognition that have been noticed by the individual or an observer, even in the absence of impact on daily functioning (39). All patients signed an informed consent for the study that was

conducted according to the Declaration of Helsinki and approved by the Institutional Ethics Committee of Verona and Rovigo (Italy) (#2459 CESC).

## HLA-C genotyping and sequencing

DNA samples from all patients were extracted from peripheral blood lymphocytes and subjected to allele-specific polymerase chain reaction (AS-PCR) and Sanger sequencing to determine HLA-C allotypes as described in our previous study (37). HLA-C allotypes C*01, C*02, C*05, C*06, C*12, C*14, C*15, C*16 and C*17 were typed at one digit resolution, since less common allotypes were mostly clustering in the same subgroup according to Shen et al. (40). Allele-specific PCR was used to type C*12:03 and C*16:01. Sequencing analysis of the HLA-C region between exons 2 and 3 was utilized to further characterize HLA-C*03:02, C*03:03, C*03:04, C*04:01, C*04:03, C*07:01, C*07:02, C*07:04, C*08:01 and C*08:02 variants at second digit resolution, by performing sequence alignments at the Immuno Polymorphism Database (IPD https://www.ebi.ac.uk/ipd/index.html).

## Determination of patient-specific HLA-C stability score

A stability score was determined for each patient based on their HLA-C genotype by multiplying the allotype stability scores (determined by the AUC of the corresponding allotypes). For HLA-C alleles typed at the first-digit resolution, the stability score of the most frequent allotype was used.

For HLA-C*14, the two most common allotypes (C*14:02 and C*14:03) have similar frequencies (35) and are clustered within the same subgroup, according to Shen et al. (40). In this case, the allotype stability score was determined by calculating the frequency-weighted mean. Similarly, when a patient's DNA was insufficient for second-digit genotyping, the frequency-weighted mean was calculated.

Finally, since allele-specific PCR for HLA-C*12 only identifies the common subtype C*12:03 at the second-digit, the stability score of the other most frequent C*12 allotype (C*12:02) was assigned when C*12 typing did not match C*12:03.

## Statistical analysis

Statistical analyses were performed using GraphPad Prism (version 10). Comparisons between independent patient groups (LTNP vs P; HAND- vs HAND+) were performed using the two-tailed Mann–Whitney U test. A p-value $< 0.05$ was considered statistically significant.

# Results

## HLA-C allotypes exhibit different binding stability with their specific peptides

The analysis of the EL-score distribution revealed significant variations in HLA-C binding stability. Indeed, allele C*05:01 showed a strong binding for most peptides, while peptides specifically binding allele C*07:04 presented much lower binding prediction values. The determination of a stability score by calculating the AUC allowed a clear identification of different stability values describing the binding of each allele to its own peptide pool. Accordingly, some alleles, such as HLA-C*07:04, C*07:01 and C*08:01, displayed weak binding interactions, while others, such as HLA-C*05:01, C*08:02 and C*14, exhibited strong binding stability (Figure 1). Notably, we identified discrepancies within allele subtypes that challenged previous classifications of stability. For example, HLA-C*07, previously considered an unstable allele (37), showed subtype-dependent variability, with C*07:02 demonstrating a greater stability score than C*07:01 and C*07:04. Similarly, HLA-C*08, considered a stable allele, presented a similar subtype variability, with C*08:02 among the variants with the highest stability, but C*08:01 among those with the lowest. The considered HLA-C allotypes and the calculated stability scores are reported in Table 1.

## The HLA-C complex stability is determined by the interactions with the peptide

Computational studies were performed to acquire atomistic details on the interaction between HLA-C/peptide and $\beta_2$m. Analyses were performed after selecting a group of HLA-C variants for which the crystal structures were available in the Protein Data Bank (www.rcsb.org). We randomly selected two 9-mer peptides specific for each allotype tested. After accomplishing MD simulations of the selected peptides in complex with the HLA-C/$\beta_2$m heterodimer, the analysis of the MD trajectories and the data retrieved from the pairwise energy decomposition revealed that there was a clear recurring pattern in $\beta_2$m's "hot spot" residues involved in the interaction with HLA-C. In fact, the trend of the residues mainly contributing and their energetic contributions' values were superimposable within the different allelic variants, but also amongst the different peptides (Table 2). On average (Figure 2 and Table 2), our analysis revealed that a) $\beta_2$m-Trp80 and -Trp61 give the main energetic contribution in each considered complex, since their energetic contribution's values are always among -9 and -12 kcal/mol; b) $\beta_2$m-Phe76 always stands in second position, with values around -5/-6 kcal/mol; c) The $\beta_2$m-Tyr30 residue is another key contributor to the interaction, providing a consistent energetic contribution of around -4 to -5 kcal/mol across all complexes.

10.3389/fimmu.2025.1703026

TABLE 2 Energetic contributions of $\beta_2$m residues to HLA-C complex stability.

| Allotype | Peptide | EL-score | % Ranking | β2m residue | ΔG [kcal/mol] |
|---|---|---|---|---|---|
| C*03:04 | QATMPHLSM | 0.610 | 64.941 | Trp60 | -10.1 |
| | | | | Phe76 | -5.8 |
| | TITDIISAL | 0.712 | 57.598 | Trp60 | -10.5 |
| | | | | Phe76 | -5.7 |
| C*04:01 | YHDKNIVLL | 0.813 | 48.220 | Trp60 | -10.7 |
| | | | | Phe76 | -4.7 |
| | TFESLVAKL | 0.533 | 70.712 | Trp61 | -9.4 |
| | | | | Phe76 | -5.2 |
| C*05:01 | NLDQPPAFF | 0.924 | 56.380 | Trp60 | -9.3 |
| | | | | Phe76 | -5.6 |
| | NAEAKITKL | 0.621 | 88.141 | Trp 61 | -10.3 |
| | | | | Phe76 | -5.5 |
| C*06:02 | FKMTIPLLV | 0.490 | 59.063 | Trp60 | -11.7 |
| | | | | Phe76 | -6.1 |
| | VYYLKNREV | 0.563 | 53.852 | Trp60 | -9.8 |
| | | | | Phe76 | -6.5 |
| C*07:02 | LRHPVCVEL | 0.439 | 61.007 | Trp60 | -10.4 |
| | | | | Phe76 | -5.8 |
| | FYRVTTEQY | 0.499 | 55.877 | Trp60 | -10.7 |
| | | | | Phe76 | -5.6 |

Binding free energy values (ΔG) of $\beta_2$m residues that are functionally important for interaction were analyzed for five different HLA-C allotypes for which crystallographic structures are available in the Protein data bank (www.rcsb.org). The analysis was performed on the trimeric complex consisting of the HLA-C heavy chain, $\beta_2$m, and peptide. Two randomly chosen peptides specific to each allotype were analyzed. The $\beta_2$m residues primarily involved in interacting with HLA-C were Trp60 and Phe76. These residues formed comparable contacts and exhibited similar ΔG values across different HLA-C allotypes.

These data clearly indicate that $\beta_2$m maintains an invariant binding pattern with HLA-C heavy chain molecules, with consistent binding affinity and interacting residues, regardless of the bound peptide or HLA-C subtype.

## Lower HLA-C stability scores are associated with HIV-1 progression and HIV-1-related neurocognitive impairment

Peptide specificity is determined by interactions of peptide side chains with six binding pockets in the HLA-C peptide-binding site (41). Kangueane et al. (42) assigned common HLA-C alleles to 18 different sub-supertypes, with variants within the same sub-supertype generally binding a highly shared set of peptides. It is thus possible to predict peptide binding of other members of a supertype using experimental results based on just one member of the type (42). A more recent classification by Shen et al. (40) proposed three main subtypes for HLA-C (C7 includes various C*07 alleles; C1 includes C*05, C*17 and most of C*01 and C*15; C2 includes C*02, C*06, C*14, C*16). Because structural clustering of the peptide-binding region correlates with binding specificity (40), according to the supertype classification by Shen et al., we adopted typing at the second-digit level for structurally divergent alleles belonging to different subtypes (for instance C*03:02 to C2, while C*03:03 and C*03:04 to C1) or to different subgroups within C1 (C*04:03 and C*04:01, or C*08:01 and C*08:02), C7 (C*07:01, C*07:02 and C*07:04) and C2 (C*12:02 and C*12:03) subtypes, in order to obtain a more precise identification of the specific allotype. Based on these considerations, HLA-C typing made it possible to calculate a stability score related to each patient's genotype and to correlate this value to different outcomes of HIV-1 infection. The stability scores calculated based on the HLA-C genotype of each HIV-1 patient considered are shown in Table 3 and Table 4.

We found that Progressors (P) exhibited significantly lower stability scores compared to Long Term Non Progressors (LTNP) ($p = 0.0143$), supporting the hypothesis that unstable HLA-C alleles are associated with more severe disease outcomes (Figure 3A).

The examination of neurocognitive outcomes showed that HAND-positive patients had a higher prevalence of unstable alleles, with a statistically significant difference in stability scores compared to HAND-negative patients ($p = 0.0221$) (Figure 3B).

10.3389/fimmu.2025.1703026

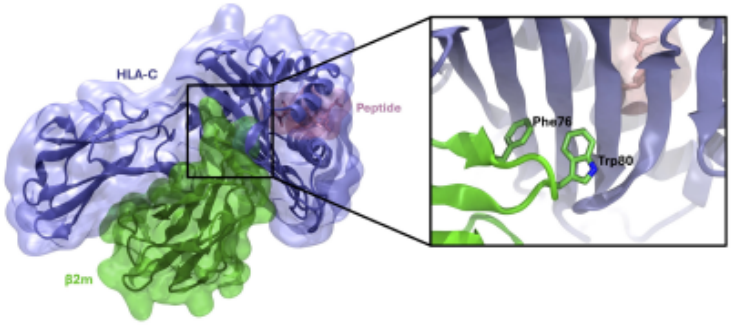



FIGURE 2
HLA-C/$\beta_2$m/peptide complex. A 3D representation of the HLA-C/$\beta_2$m/peptide complex, as obtained at the end of MD simulations, is shown on the left. On the right, a detailed visualization of the HLA-C/$\beta_2$m interface, highlighting residues critical for mediating protein-protein interactions, is shown.

## Discussion

We explored HLA-C stability by means of advanced computational techniques and found that HLA-C stability influenced HIV-1 progression and cognitive outcomes. The identification of allele subtypes with different stability levels challenges traditional classifications of HLA-C as simply "stable" or "unstable," highlighting the advantages of a more precise and individualized approach in assessing HLA-C function. Moreover, our findings provide compelling evidence that HLA-C stability may contribute to shaping both immune responses and neurological outcomes in HIV-1 infected individuals.

### HLA-C function is related to its stability

HLA-C stability is significantly influenced by genetic variations, which can either enhance or reduce its functionality within the immune response system (43, 44). HLA-C molecules exhibit differential stability based on genetic variations in their promoter regions, affecting their expression levels and antigen presentation efficiency. Some alleles, e.g., HLA-C*05, exhibit high stability and remain on the cell surface for prolonged surface expression, whereas others, e.g., HLA-C*07, display lower stability, impairing peptide loading and presentation (9). These observations are consistent with previous findings showing that inefficient peptide binding contributes to lower surface expression of HLA-C compared to HLA-A and HLA-B (6), likely due to its greater selectivity for peptide binding (3).

Several experimental and computational methodologies have been employed to assess HLA-C peptide binding stability. Mass spectrometry-based ligand elution assays have been used to profile naturally processed peptides, revealing a preference for 9-mer peptides in HLA-C binding (35, 36). Computational tools such as NetMHCpan and MHCflurry have demonstrated increased predictive accuracy in assessing binding affinities and stability of peptide-HLA interactions (19). Notably, peptide-MHC stability, rather than binding affinity alone, has been identified as a better predictor of T-cell immunogenicity (8). Additionally, computational tools like NetMHCpan and mass spectrometry-based ligand profiling have significantly improved the prediction and validation of HLA-C peptide-binding stability (19). These methodologies may provide insights into how stability affects immune function, disease progression, and therapeutic outcomes.

Recent studies have shown that HLA-C expression varies widely in an allele-specific manner, with higher expression levels exerting greater selection pressure on HIV-1, leading to virus-mediated downregulation of HLA-C (9). Certain single-nucleotide polymorphisms (SNPs) in the promoter and 3' untranslated region (UTR) of HLA-C, such as rs9264942 (located in the 5' UTR of the HLA-C gene), have been correlated with HIV viral load and disease progression (45). Additionally, the 3'UTR of HLA-C is a target for microRNA regulation (miR-148a), which influences HLA-C surface expression (14). This regulatory mechanism plays a crucial role in determining antigen presentation efficiency and immune escape strategies employed by HIV-1.

Our computational analysis supports evidence that HLA-C allotypes differ in their ability to bind and stabilize peptides, influencing their expression levels on the cell surface. Structural studies have demonstrated that HLA-C*07 has a deeper and narrower antigen-binding cleft, while HLA-C*05 has a relatively flat peptide-binding groove, allowing it to bind a broader range of peptides and remain more stably expressed on the cell membrane (9). These findings align with our stability analysis, which demonstrated that HLA-C*07:04 is the most unstable allele, while HLA-C*05:01 exhibits the highest stability. Importantly, the inefficient association of certain HLA-C variants with $\beta_2$m leads to the accumulation of misfolded HLA-C heavy chains, further affecting antigen presentation efficiency (6, 15).

10.3389/fimmu.2025.1703026

TABLE 3 HLA-C stability scores in HIV-1 patients according to AIDS progression.

| Sample code | HIV-1 progression | Sex | Age | HLA-C genotype and stability scores | | | | Stability score |
|---|---|---|---|---|---|---|---|---|
| | | | | 1st allele | Score (1st) | 2nd allele | Score (2nd) | |
| PR1 | LTNP | M | 42 | *03 | 65.52 | *08:02 | 78.34 | 5132.84 |
| PR2 | LTNP | M | 29 | *03 | 65.52 | *03 | 65.52 | 4292.87 |
| PR3 | LTNP | F | 36 | *06 | 56.13 | *07:02 | 50.45 | 2831.76 |
| PR4 | LTNP | M | 54 | *02 | 56.76 | *12:03 | 50.34 | 2857.30 |
| PR5 | LTNP | F | 47 | *07:02 | 50.45 | *12 | 52.20 | 2633.49 |
| PR6 | LTNP | M | 23 | *05 | 84.84 | *05 | 84.84 | 7197.83 |
| PR7 | LTNP | M | 55 | *05 | 84.84 | *06 | 56.13 | 4762.07 |
| PR8 | LTNP | M | 35 | *02 | 56.76 | *06 | 56.13 | 3185.94 |
| PR9 | LTNP | M | 58 | *05 | 84.84 | *08:02 | 78.34 | 6646.37 |
| PR10 | LTNP | M | 46 | *07:02 | 50.45 | *14 | *76.05* | 3836.72 |
| PR11 | LTNP | M | 32 | *07:02 | 50.45 | *12:03 | 50.34 | 2539.65 |
| PR12 | LTNP | M | 26 | *05 | 84.84 | *12:03 | 50.34 | 4270.85 |
| PR13 | LTNP | M | 33 | *02 | 56.76 | *15 | 57.39 | 3257.46 |
| PR14 | LTNP | M | 52 | *07:01 | 29.74 | *08:02 | 78.34 | 2329.83 |
| PR15 | LTNP | M | 42 | *04:01 | 67.28 | *07:01 | 29.74 | 2000.91 |
| PR16 | LTNP | M | 34 | *04:01 | 67.28 | *07:02 | 50.45 | 3394.28 |
| PR17 | LTNP | M | 44 | *04:01 | 67.28 | *08:02 | 78.34 | 5270.72 |
| PR18 | LTNP | M | 55 | *01 | 63.87 | *12 | 52.20 | 3334.01 |
| PR19 | LTNP | M | 32 | *14 | 76.05 | *08:02 | 78.34 | 5957.76 |
| PR20 | LTNP | M | 49 | *04:01 | 67.28 | *08:02 | 78.34 | 5270.72 |
| PR21 | LTNP | F | 42 | *04:01 | 67.28 | *16:01 | 51.12 | 3439.35 |
| PR22 | LTNP | F | 54 | *07:01 | 29.74 | *12 | 52.20 | 1552.43 |
| PR23 | LTNP | M | 51 | *04:01 | 67.28 | *08:02 | 78.34 | 5270.72 |
| PR24 | LTNP | F | 58 | *01 | 63.87 | *07:02 | 50.45 | 3222.24 |
| PR25 | LTNP | F | 77 | *01 | 63.87 | *04:01 | 67.28 | 4297.17 |
| PR26 | LTNP | M | 39 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| PR27 | LTNP | M | 40 | *04:01 | 67.28 | *12 | 52.20 | 3512.02 |
| PR28 | LTNP | F | 43 | *05 | 84.84 | *02 | 56.76 | 4815.52 |
| PR29 | LTNP | F | 47 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| PR30 | LTNP | M | 28 | *07:01 | 29.74 | *05 | 84.84 | 2523.14 |
| PR31 | LTNP | F | 49 | *03 | 65.52 | *02 | 56.76 | 3718.92 |
| PR32 | LTNP | M | 53 | *07 | 40.96 | *16:01 | 51.12 | 2093.88 |
| PR33 | LTNP | M | 30 | *03:02 | 65.09 | *06 | 56.13 | 3653.50 |
| PR34 | LTNP | M | 49 | *05 | 84.84 | *06 | 56.13 | 4762.07 |
| PR35 | LTNP | F | 43 | *01 | 63.87 | *06 | 56.13 | 3585.02 |
| PR36 | LTNP | M | 27 | *05 | 84.84 | *12:03 | 50.34 | 4270.85 |
| PR37 | LTNP | M | 57 | *06 | 56.13 | *12:03 | 50.34 | 2825.58 |

*(Continued)*

10.3389/fimmu.2025.1703026

TABLE 3 Continued

| Sample code | HIV-1 progression | Sex | Age | HLA-C genotype and stability scores | | | | Stability score |
|---|---|---|---|---|---|---|---|---|
| | | | | 1st allele | Score (1st) | 2nd allele | Score (2nd) | |
| PR38 | P | M | 45 | *04:01 | 67.28 | *07:02 | 50.45 | 3394.28 |
| PR39 | P | M | 57 | *03 | 65.52 | *07:01 | 29.74 | 1948.56 |
| PR40 | P | M | 28 | *06 | 56.13 | *07:01 | 29.74 | 1669.31 |
| PR41 | P | M | 41 | *07:02 | 50.45 | *12 | 52.20 | 2633.49 |
| PR42 | P | M | 40 | *02 | 56.76 | *04:01 | 67.28 | 3818.81 |
| PR43 | P | M | 42 | *04:01 | 67.28 | *07:01 | 29.74 | 2000.91 |
| PR44 | P | M | 25 | *04:01 | 67.28 | *08:02 | 78.34 | 5270.72 |
| PR45 | P | M | 23 | *05 | 84.84 | *07:01 | 29.74 | 2523.14 |
| PR46 | P | M | 30 | *03:02 | 65.09 | *04 | 68.06 | 4430.03 |
| PR47 | P | M | 30 | *05 | 84.84 | *07:01 | 29.74 | 2523.14 |
| PR48 | P | M | 28 | *07:02 | 50.45 | *12:03 | 50.34 | 2539.65 |
| PR49 | P | M | 31 | *03 | 65.52 | *07:01 | 29.74 | 1948.56 |
| PR50 | P | M | 39 | *02 | 56.76 | *04:01 | 67.28 | 3818.81 |
| PR51 | P | M | 37 | *04:01 | 67.28 | *07:02 | 50.45 | 3394.28 |
| PR52 | P | M | 26 | *03 | 65.52 | *02 | 56.76 | 3718.92 |
| PR53 | P | M | 36 | *04:01 | 67.28 | *04:01 | 67.28 | 4526.60 |
| PR54 | P | M | 51 | *07:04 | 9.7 | *15 | 57.39 | 556.68 |
| PR55 | P | F | 47 | *04 | 68.06 | *16 | 51.12 | 3479.23 |
| PR56 | P | M | 28 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| PR57 | P | M | 65 | *07:04 | 9.7 | *15 | 57.39 | 556.68 |
| PR58 | P | M | 61 | *07 | 40.96 | *07 | 40.96 | 1677.72 |
| PR59 | P | F | 30 | *07:01 | 29.74 | *04:01 | 67.28 | 2000.91 |
| PR60 | P | M | 57 | *04:01 | 67.28 | *12:03 | 50.34 | 3386.88 |
| PR61 | P | M | 41 | *03:02 | 65.09 | *07:02 | 50.45 | 3283.79 |
| PR62 | P | M | 54 | *07:02 | 50.45 | *07:02 | 50.45 | 2545.20 |
| PR63 | P | M | 34 | *04 | 68.06 | *05 | 84.84 | 5774.21 |
| PR64 | P | M | 31 | *03:02 | 65.09 | *06 | 56.13 | 3653.50 |
| PR65 | P | M | 30 | *07:02 | 50.45 | *08:02 | 78.34 | 3952.25 |
| PR66 | P | M | 64 | *01 | 63.87 | *02 | 56.76 | 3625.26 |
| PR67 | P | M | 48 | *04:01 | 67.28 | *15 | 57.39 | 3861.20 |
| PR68 | P | F | 55 | *07 | 40.96 | *05 | 84.84 | 3475.05 |
| PR69 | P | F | 41 | *06 | 56.13 | *16 | 51.12 | 2869.37 |
| PR70 | P | F | 50 | *07:02 | 50.45 | *16:01 | 51.12 | 2579.00 |
| PR71 | P | M | 30 | *04:01 | 67.28 | *08:02 | 78.34 | 5270.72 |
| PR72 | P | M | 49 | *03 | 65.52 | *06 | 56.13 | 3677.64 |
| PR73 | P | F | 32 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| PR74 | P | M | 61 | *03 | 65.52 | *07:01 | 29.74 | 1948.56 |

(Continued)

10.3389/fimmu.2025.1703026

TABLE 3 Continued

| Sample code | HIV-1 progression | Sex | Age | HLA-C genotype and stability scores | | | | Stability score |
|---|---|---|---|---|---|---|---|---|
| | | | | 1st allele | Score (1st) | 2nd allele | Score (2nd) | |
| PR75 | P | M | 29 | *03:02 | 65.09 | *07:02 | 50.45 | 3283.79 |
| PR76 | P | F | 53 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| PR77 | P | M | 57 | *04 | 68.06 | *12 | 52.20 | 3552.73 |
| PR78 | P | F | 57 | *04 | 68.06 | *07:04 | 9.70 | 660.18 |
| PR79 | P | M | 33 | *07:01 | 29.74 | *07:02 | 50.45 | 1500.38 |
| PR80 | P | M | 32 | *03 | 65.52 | *15 | 57.39 | 3760.19 |
| PR81 | P | F | 39 | *04 | 68.06 | *12:03 | 50.34 | 3426.14 |
| PR82 | P | F | 32 | *07 | 40.96 | *16:01 | 51.12 | 2093.88 |
| PR83 | P | M | 63 | *06 | 56.13 | *07:02 | 50.45 | 2831.76 |
| PR84 | P | F | 50 | *07 | 40.96 | *17 | 55.18 | 2260.17 |

HIV-1 infected individuals who showed either slow (Long Term Non Progressors, LTNP) or rapid progression (P) to AIDS were genotyped for HLA-C alleles. The stability values corresponding to each allotype were multiplied together to give a final HLA-C stability score for each patient. Allele scores calculated using a frequency-weighted average are reported in italics. Patient code, sex (M, male; F, female), HLA-C alleles and their corresponding stability scores, and the final calculated stability scores are reported.

In addition, we noted that some very unstable variants, such as C*07:01 or C*07:04, are also the lowest expressed variants, whereas some more stable variants, such as C*05:01 or C*14:02, are among the most highly expressed alleles. While numerous factors contribute to the regulation of HLA-C expression levels, the peptide-binding capacity of different HLA-C variants is also recognized as a key determinant influencing their surface expression (46). Therefore, it is plausible that inefficient peptide binding may contribute to the lower expression of certain HLA-C variants on the cell membrane (9), suggesting a potential correlation between stability and expression levels.

Our analysis improved the definition of the stability of HLA-C allotypes, overcoming the previous binary classification, which was too simplistic and inaccurate. The in-depth characterization of binding stability to peptides specific to the most common variants in the human population was carried out using pools of experimentally validated peptides known to bind to the main allotypes (35). This approach enabled obtaining a "stability coefficient" for each of these HLA-C variants and thereby quantified the overall stability profile of each HLA-typed individual based on the combination of their specific allotype stability coefficients.

## HLA-C stability influences HIV-1 infection progression

HLA-C stability plays a critical role in modulating defense mechanisms against HIV-1. Unstable alleles may increase HIV-1 progression, suggesting that by influencing antigen presentation efficiency, they may alter the ability of CD8+ T cells to recognize and eliminate infected cells, in keeping with previous findings indicating that HLA-C expression levels directly affect HIV-1 immune control (10). Additionally, in our previous studies (24, 25) we reported that stable HLA-C alleles are associated with lower viral loads and more effective immune responses, supporting our finding that unstable alleles contribute to a faster progression to AIDS.

To assess the correlation between stability metrics and progression of HIV-1 infection, we re-analyzed a population of HIV-1 positive, treatment-naïve subjects described and characterized in our previous study (37). In the original analysis, the different allotypes were binary divided into stable and unstable, revealing a significant correlation between rapid progression to AIDS and the presence of unstable HLA-C variants. In the present study, we refined the analysis by performing high-resolution typing of the second digit for divergent subtypes and assessing a quantitative stability coefficient based on their HLA-C allotype combination. This enhanced approach confirmed and strengthened our earlier findings, showing a robust and highly significant correlation between accelerated disease progression and a higher burden of unstable HLA-C allotypes.

## HLA-C stability impact on HIV-associated neurocognitive outcomes

The association between unstable HLA-C alleles and HAND is particularly noteworthy. Previous studies have suggested that unstable HLA-C variants may lead to increased levels of free $\beta_2$m, contributing to neuroinflammation and neuronal damage (29). Additionally, a recent study has specifically linked HLA-C*07 to HAND in HIV patients (31).

To test the clinical relevance of the specific stability coefficient for each major HLA-C allotype on neurological outcomes in HIV-1 infection, we analyzed a population of HIV-positive subjects with subjective report or objective evidence of cognitive impairment and compared them to cognitively unimpaired ones. We observed a

10.3389/fimmu.2025.1703026

TABLE 4 HLA-C stability scores in HIV-1–infected patients with and without neurocognitive impairment.

| Sample code | Neurocognitive status | Sex | Age | HLA-C genotype and stability scores | | | | Stability score |
|---|---|---|---|---|---|---|---|---|
| | | | | 1st allele | Score (1st) | 2nd allele | Score (2nd) | |
| H1 | HAND + | M | 42 | *01 | 63.87 | *06 | 56.13 | 3585.02 |
| H2 | HAND + | F | 46 | *03:02 | 65.09 | *07:02 | 50.45 | 3283.79 |
| H3 | HAND + | M | 60 | *02 | 56.76 | *16:01 | 51.12 | 2901.57 |
| H4 | HAND + | M | 59 | *03:02 | 65.09 | *04:01 | 67.28 | 4379.26 |
| H5 | HAND + | F | 52 | *07:02 | 50.45 | *07:04 | 9.70 | 489.37 |
| H6 | HAND + | M | 66 | *03:02 | 65.09 | *16:01 | 51.12 | 3327.40 |
| H7 | HAND + | M | 40 | *06 | 56.13 | *07:01 | 29.74 | 1669.31 |
| H8 | HAND + | M | 63 | *01 | 63.87 | *07:01 | 29.74 | 1899.49 |
| H9 | HAND + | M | 56 | *04:01 | 67.28 | *07:01 | 29.74 | 2000.91 |
| H10 | HAND + | F | 53 | *06 | 56.13 | *12:03 | 50.34 | 2825.58 |
| H11 | HAND + | M | 52 | *05 | 84.84 | *07:01 | 29.74 | 2523.14 |
| H12 | HAND + | M | 56 | *04:01 | 67.28 | *12:03 | 50.34 | 3386.88 |
| H13 | HAND + | M | 51 | *07:01 | 29.74 | *07:01 | 29.74 | 884.47 |
| H14 | HAND + | F | 65 | *02 | 56.76 | *07:01 | 29.74 | 1688.04 |
| H15 | HAND + | M | 58 | *07:01 | 29.74 | *12:03 | 50.34 | 1497.11 |
| H16 | HAND + | M | 64 | *07:01 | 29.74 | *08:02 | 78.34 | 2329.83 |
| H17 | HAND - | M | 57 | *07:01 | 29.74 | *07:02 | 50.45 | 1500.38 |
| H18 | HAND - | M | 58 | *07:01 | 29.74 | *15 | 57.39 | 1706.78 |
| H19 | HAND - | F | 54 | *07:01 | 29.74 | *14 | 76.05 | 2261.73 |
| H20 | HAND - | M | 45 | *04:01 | 67.28 | *15 | 57.39 | 3861.20 |
| H21 | HAND - | M | 31 | *07:01 | 29.74 | *12:03 | 50.34 | 1497.11 |
| H22 | HAND - | F | 53 | *02 | 56.76 | *12:03 | 50.34 | 2857.30 |
| H23 | HAND - | M | 39 | *03:02 | 65.09 | *12:03 | 50.34 | 3276.63 |
| H24 | HAND - | F | 56 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| H25 | HAND - | M | 41 | *04:01 | 67.28 | *12:03 | 50.34 | 3386.88 |
| H26 | HAND - | M | 55 | *04:01 | 67.28 | *04:01 | 67.28 | 4526.60 |
| H27 | HAND - | F | 56 | *05 | 84.84 | *12:03 | 50.34 | 4270.85 |
| H28 | HAND - | NA | NA | *04:01 | 67.28 | *07:01 | 29.74 | 2000.91 |
| H29 | HAND - | M | 62 | *05 | 84.84 | *07:01 | 29.74 | 2523.14 |
| H30 | HAND - | F | 57 | *07:02 | 50.45 | *12:03 | 50.34 | 2539.65 |
| H31 | HAND - | M | 47 | *07:02 | 50.45 | *15 | 57.39 | 2895.33 |
| H32 | HAND - | M | 60 | *02 | 56.76 | *15 | 57.39 | 3257.46 |
| H33 | HAND - | M | 60 | *06 | 56.13 | *07:01 | 29.74 | 1669.31 |
| H34 | HAND - | M | 62 | *03:02 | 65.09 | *07:02 | 50.45 | 3283.79 |
| H35 | HAND - | M | 59 | *08:02 | 78.34 | *17 | 55.18 | 4322.80 |
| H36 | HAND - | M | 65 | *01 | 63.87 | *14 | 76.05 | 4857.31 |
| H37 | HAND - | M | 62 | *06 | 56.13 | *12:03 | 50.34 | 2825.58 |

(Continued)

10.3389/fimmu.2025.1703026

TABLE 4 Continued

| Sample code | Neurocognitive status | Sex | Age | HLA-C genotype and stability scores | | | | Stability score |
|---|---|---|---|---|---|---|---|---|
| | | | | 1st allele | Score (1st) | 2nd allele | Score (2nd) | |
| H38 | HAND - | M | 37 | *01 | 63.87 | *07:01 | 29.74 | 1899.49 |
| H39 | HAND - | M | 54 | *02 | 56.76 | *08:02 | 78.34 | 4446.58 |
| H40 | HAND - | F | 64 | *04:01 | 67.28 | *07:02 | 50.45 | 3394.28 |
| H41 | HAND - | M | 59 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| H42 | HAND - | M | 46 | *07:02 | 50.45 | *14 | 76.05 | 3836.72 |
| H43 | HAND - | M | 64 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| H44 | HAND - | F | 54 | *05 | 84.84 | *16:01 | 51.12 | 4337.02 |
| H45 | HAND - | M | 58 | *05 | 84.84 | *07:01 | 29.74 | 2523.14 |
| H46 | HAND - | M | 54 | *07:01 | 29.74 | *07:02 | 50.45 | 1500.38 |
| H47 | HAND - | M | 55 | *03:02 | 65.09 | *14 | 76.05 | 4950.09 |
| H48 | HAND - | M | 54 | *06 | 56.13 | *06 | 56.13 | 3150.58 |
| H49 | HAND - | M | 47 | *04:01 | 67.28 | *07:02 | 50.45 | 3394.28 |
| H50 | HAND - | M | 60 | *04:01 | 67.28 | *06 | 56.13 | 3776.43 |
| H51 | HAND - | M | 51 | *03:02 | 65.09 | *12:03 | 50.34 | 3276.63 |
| H52 | HAND - | M | 64 | *04:01 | 67.28 | *04:01 | 67.28 | 4526.60 |
| H53 | HAND - | M | 65 | *03:02 | 65.09 | *12:03 | 50.34 | 3276.63 |
| H54 | HAND - | M | 47 | *04:01 | 67.28 | *05 | 84.84 | 5708.04 |
| H55 | HAND - | M | 48 | *04:01 | 67.28 | *16:01 | 51.12 | 3439.35 |
| H56 | HAND - | M | 45 | *07:01 | 29.74 | *15 | 57.39 | 1706.78 |
| H57 | HAND - | M | 53 | *04:01 | 67.28 | *07:01 | 29.74 | 2000.91 |

HIV-infected individuals presenting (HAND-positive, HAND+) or not presenting (HAND-negative, HAND-) neurocognitive impairment were genotyped for HLA-C alleles. For each patient, the stability score of each allele was multiplied to obtain the final HLA-C stability score. Allele scores calculated using a weighted average are reported in italics. Patient code, sex (M, male; F, female; NA, not available), HLA-C alleles and their respective stability scores, and the final calculated stability scores are reported.

significantly higher presence of unstable HLA-C variants in HAND-positive subjects than in HAND-negative ones, further confirming our preliminary observation on a small case series (29). Our findings confirm this association, suggesting that HLA-C instability may exacerbate neurocognitive decline by promoting chronic immune activation and neuroinflammatory processes. Extensive scientific evidence supports the notion that genetic determinants of immune function are critical in shaping disease outcomes in HIV-1. In particular, previous studies have emphasized the role of HLA alleles in HIV-1 replication and progression (7, 10, 24, 25). Our findings extend this knowledge by highlighting the relevance of HLA-C allele stability at a more granular level. The association between HLA-C stability and neurocognitive disorders also aligns with research on $\beta_2$m in neurodegeneration (29), providing a potential molecular framework for future investigations into HIV-associated neurocognitive impairment.

## HLA-C stability: towards a precision medicine approach

A major advantage of this refined stability analysis is the ability to calculate a personalized HLA-C stability score for each patient. By integrating computational stability assessments, such as EL-scores and AUC-derived stability coefficients, with patient-specific genetic data, it is possible to develop a predictive model that can help stratify patients based on their risk of rapid disease progression or neurocognitive complications. This approach aligns with the broader movement toward personalized medicine, where treatments and monitoring strategies are tailored to an individual's genetic profile. From a clinical perspective, the ability to predict disease progression based on HLA-C stability could lead to more targeted interventions. For instance, patients identified as having unstable HLA-C variants could be prioritized for early

10.3389/fimmu.2025.1703026

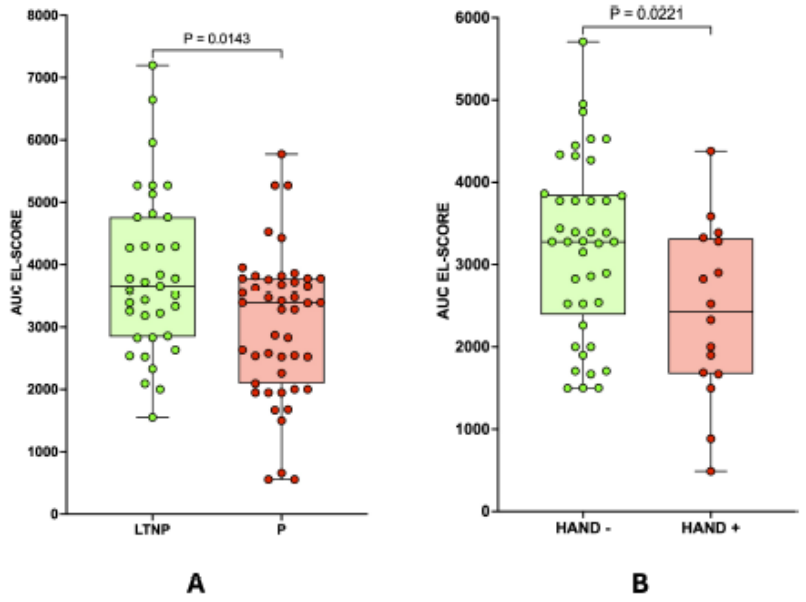



FIGURE 3
Association of patient HLA-C stability scores with HIV-1 progression and HIV-1-related neurocognitive impairment. (**A**) HIV-1 patients who experience rapid disease progression (Progressors, P, n=47) exhibit lower stability scores than those with slower progression (Long Term Non Progressors, LTNP, n=37). (**B**) HIV-1 patients presenting neurocognitive impairment (HAND+, n=16) have lower stability scores than cognitively normal patients (HAND -, n= 41). Statistical analyses were performed using the two-tailed Mann-Whitney U test; $p < 0.05$ was considered statistically significant.

intervention strategies, including more intensive monitoring, earlier initiation of antiretroviral therapy, or adjunctive therapies targeting immune modulation. Furthermore, understanding the link between HLA-C instability and HAND could open new therapeutic avenues, such as early interventions targeting neuroinflammatory pathways. The use of HLA-C genotyping could improve personalized treatment strategies, identifying individuals at higher risk of rapid disease progression or HAND.

The study of HLA-C peptide binding stability is critical for understanding immune regulation and its implications in infectious diseases, cancer, and autoimmunity. Advances in computational modelling and experimental methodologies provide valuable insights into the structural and functional aspects of HLA-C stability. Future research should focus on refining predictive models, defining clinically relevant stability thresholds and exploring therapeutic interventions aimed at enhancing HLA-C-mediated immune responses.

The main limitation of the study was the inability to perform more precise typing at the second digit for several biological samples due to insufficient DNA for testing. This reduced the accuracy of the data obtained, necessitating the use of a frequency-weighted average of stability scores in some cases.

## Conclusions

HLA-C plays a crucial and multifaceted role in HIV-1 infection, influencing immune recognition, disease progression, and neurocognitive outcomes. Despite its lower surface expression compared to HLA-A and HLA-B, HLA-C contributes significantly to both adaptive and innate immune responses. Advances in genomic and immunological research continue to reveal the complex interactions between HLA-C and HIV-1, providing valuable insights into potential therapeutic and vaccine strategies. By using an innovative cutting-edge bioinformatic pipeline, we demonstrated that reduced HLA-C stability is associated with faster HIV-1 disease progression and a higher prevalence of HIV-associated neurocognitive disorders. Overall, the findings from this study emphasize that HLA-C stability analysis should become an integral part of HIV-1 disease management and research. Future studies should focus on refining predictive models for personalized stability scoring, validating these findings in larger, independent cohorts, to early identify those at greater risk of progression or developing neurocognitive symptoms and intervene early with the most appropriate treatment approaches to improve patient outcomes.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving humans were approved by Ethics Committee of the Health Department of the Federal District (#066/07) from hospitals of the public network of the Federal District. Brazil; Ethics Committee for Analysis of Research Projects (CAPPesq) of Hospital das Clinicas HCFMUSP (#CAPPesq #0306/10, online registration #5867) Faculdade de Medicina da Universidade de Sao Paulo, Brazil Institutional Ethics Committee of Verona and Rovigo (Italy) (#2459 CESC). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

MV: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. AS: Investigation, Methodology, Validation, Writing – review & editing. EGM: Investigation, Methodology, Software, Writing – review & editing. CDM: Investigation, Methodology, Writing – review & editing. EO: Investigation, Methodology, Writing – review & editing. ST: Conceptualization, Data curation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. EM: Data curation, Investigation, Validation, Writing – review & editing. AF: Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing. ML: Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – review & editing. EL: Data curation, Methodology, Validation, Writing – review & editing. GAA: Data curation, Methodology, Validation, Writing – review & editing. BCMDS: Data curation, Methodology, Validation, Writing – review & editing. AJDSD: Data curation, Methodology, Validation, Writing – review & editing. JC: Data curation, Methodology, Validation, Writing – review & editing. ERA: Data curation, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. MM: Data curation, Methodology, Writing – review & editing. MA: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – review & editing. AR: Conceptualization, Supervision, Visualization, Writing – review & editing. MGR: Conceptualization, Supervision, Visualization, Writing – review & editing. MTV: Conceptualization, Supervision, Visualization, Writing – review & editing. GG: Conceptualization, Investigation, Methodology, Software, Supervision, Validation, Writing – original draft, Writing – review & editing. DZ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Software, Supervision, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was financially supported by the Gilead Fellowship Program, 2019 Edition, by the Federal District Research Support Foundation (FAPDF), Bench Amendment No. 7108001, and by the Italian Ministry of Research (MUR) as part of the Excellence Project 2023-2027 of the Department of Neuroscience, Biomedicine and Movement Sciences of the University of Verona.

## Acknowledgments

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Health Department of the Federal District (#066/07) from hospitals of the public network of the Federal District. Brazil; Ethics Committee for Analysis of Research Projects (CAPPesq) of Hospital das Clinicas HCFMUSP (#CAPPesq #0306/10, online registration #5867) Faculdade de Medicina da Universidade de Sao Paulo, Brazil and the Institutional Ethics Committee of Verona and Rovigo (Italy) (#2459 CESC). We acknowledge the contribution of Brazilian, Canadian, and US patients who agreed to participate in this study. We thank INDACO for providing high-performance computing resources and support.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

## References


1. Robinson J, Guethlein LA, Cereb N, Yang SY, Norman PJ, Marsh SGE, et al. Distinguishing functional polymorphism from random variation in the sequences of >10,000 HLA-A, -B and -C alleles. *PLoS Genet.* (2017) 13:e1006862. doi: 10.1371/journal.pgen.1006862

2. Parham P. MHC class I molecules and kirs in human history, health and survival. *Nat Rev Immunol.* (2005) 5:201–14. doi: 10.1038/nri1570

3. Neisig A, Melief CJ, Neefjes J. Reduced cell surface expression of HLA-C molecules correlates with restricted peptide binding and stable TAP interaction. *J Immunol.* (1998) 160:171–9. doi: 10.4049/jimmunol.160.1.171

4. Zemmour J, Parham P. Distinctive polymorphism at the HLA-C locus: implications for the expression of HLA-C. *J Exp Med.* (1992) 176:937–50. doi: 10.1084/jem.176.4.937

5. McCutcheon JA, Gumperz J, Smith KD, Lutz CT, Parham P. Low HLA-C expression at cell surfaces correlates with increased turnover of heavy chain mRNA. *J Exp Med.* (1995) 181:2085–95. doi: 10.1084/jem.181.6.2085

6. Blais ME, Dong T, Rowland-Jones S. HLA-C as a mediator of natural killer and T-cell activation: spectator or key player? *Immunology.* (2011) 133:1–7. doi: 10.1111/j.1365-2567.2011.03422.x

7. Zipeto D, Beretta A. HLA-C and HIV-1: friends or foes? *Retrovirology.* (2012) 9:39. doi: 10.1186/1742-4690-9-39

8. Harndahl M, Rasmussen M, Roder G, Dalgaard Pedersen I, Sørensen M, Nielsen M, et al. Peptide-MHC class I stability is a better predictor than peptide affinity of CTL immunogenicity. *Eur J Immunol.* (2012) 42:1405–16. doi: 10.1002/eji.201141774

9. Kaur G, Gras S, Mobbs JI, Vivian JP, Cortes A, Barber T, et al. Structural and regulatory diversity shape HLA-C protein expression levels. *Nat Commun.* (2017) 8:15924. doi: 10.1038/ncomms15924

10. Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, et al. Influence of HLA-C expression level on HIV control. *Science.* (2013) 340:87–91. doi: 10.1126/science.1232685

11. Kulkarni S, Qi Y, O'HUigin C, Pereyra F, Ramsuran V, McLaren P, et al. Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. *Proc Natl Acad Sci U.S.A.* (2013) 110:20705–10. doi: 10.1073/pnas.1312237110

12. Vince N, Li H, Ramsuran V, Naranbhai V, Duh FM, Fairfax BP, et al. HLA-C level is regulated by a polymorphic octl binding site in the HLA-C promoter region. *Am J Hum Genet.* (2016) 99:1353–8. doi: 10.1016/j.ajhg.2016.09.023

13. Goodson-Gregg FJ, Rothbard B, Zhang A, Wright PW, Li H, Walker-Sperling VE, et al. Tuning of NK-specific HLA-C expression by alternative mRNA splicing. *Front Immunol.* (2019) 10:3034. doi: 10.3389/fimmu.2019.03034

14. O'Huigin C, Kulkarni S, Xu Y, Deng Z, Kidd J, Kidd K, et al. The molecular origin and consequences of escape from miRNA regulation by HLA-C alleles. *Am J Hum Genet.* (2011) 89:424–31. doi: 10.1016/j.ajhg.2011.07.024

15. Sibilio L, Martayan A, Setini A, Monaco EL, Tremante E, Butler RH, et al. A single bottleneck in HLA-C assembly. *J Biol Chem.* (2008) 283:1267–74. doi: 10.1074/jbc.M708068200

16. Wang M, Kurgan L, Li M. A comprehensive assessment and comparison of tools for HLA class I peptide-binding prediction. *Briefings Bioinf.* (2023) 24(3):1–9. doi: 10.1093/bib/bbad150

17. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol.* (2017) 199:3360–8. doi: 10.4049/jimmunol.1700893

18. Nilsson JB, Greenbaum J, Peters B, Nielsen M. NetMHCpan-4.2: improved prediction of CD8+ epitopes by use of transfer learning and structural features. *Front Immunol.* (2025) 16:1616113. doi: 10.3389/fimmu.2025.1616113

19. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* (2020) 48:W449–54. doi: 10.1093/nar/gkaa379

20. Oip MD, Laufer VA, Valesano AL, Zimmerman A, Woodside KJ, Lu Y, et al. HLA-C peptide repertoires as predictors of clinical response during early SARS-coV-2 infection. *Life.* (2024) 14:1181. doi: 10.3390/life14091181

21. Rasmussen M, Renoy E, Harndahl M, Kristensen AB, Nielsen IK, Nielsen M, et al. Pan-specific prediction of peptide–MHC class I complex stability, a correlate of T cell immunogenicity. *J Immunol.* (2016) 197:1517–24. doi: 10.4049/jimmunol.1600582

22. Malnati MS, Ugolotti E, Monti MC, Battista D, Vanni I, Bordo D, et al. Activating Killer Immunoglobulin Receptors and HLA-C: a successful combination providing HIV-1 control. *Sci Rep.* (2017) 7:42470. doi: 10.1038/srep42470

23. The International HIV Controllers Study. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science.* (2010) 330:1551–7. doi: 10.1126/science.1195271

24. Serena M, Parolini F, Biswas P, Sironi F, Blanco Miranda A, Zoratti E, et al. HIV-1 Env associates with HLA-C free-chains at the cell membrane modulating viral infectivity. *Sci Rep.* (2017) 7:40037. doi: 10.1038/srep40037

25. Parolini F, Biswas P, Serena M, Sironi F, Muraro V, Guizzardi E, et al. Stability and expression levels of HLA-C on the cell membrane modulate HIV-1 infectivity. *J Virol.* (2018) 92. doi: 10.1128/JVI.01711-17

26. Tierney SM, Sheppard DP, Kordowski VM, Faytell MP, Avci G, Woods SP. A comparison of the sensitivity, stability, and reliability of three diagnostic schemes for HIV-associated neurocognitive disorders. *J Neurovirol.* (2017) 23:404–21. doi: 10.1007/s13365-016-0510-z

27. Kallianpur AR, Levine AJ. Host genetic factors predisposing to HIV-associated neurocognitive disorder. *Curr HIV/AIDS Rep.* (2014) 11:336–52. doi: 10.1007/s11904-014-0222-z

28. Leslie S, Donnelly P, McVean G. A statistical method for predicting classical HLA alleles from SNP data. *Am J Hum Genet.* (2008) 82:48–56. doi: 10.1016/j.ajhg.2007.09.001

29. Zipeto D, Serena M, Mutascio S, Parolini F, Diani E, Guizzardi E, et al. HIV-1-associated neurocognitive disorders: is HLA-C binding stability to beta(2)-microglobulin a missing piece of the pathogenetic puzzle? *Front Neurol.* (2018) 9:791. doi: 10.3389/fneur.2018.00791

30. McArthur JC, Nance-Sproson TE, Griffin DE, Hoover D, Selnes OA, Miller EN, et al. The diagnostic utility of elevation in cerebrospinal fluid beta 2-microglobulin in HIV-1 dementia. Multicenter AIDS Cohort Study. *Neurology.* (1992) 42:1707–12. doi: 10.1212/wnl.42.9.1707

31. Pons-Fuster E, Bernal E, Guillamón CF, Gimeno L, Martínez-Sánchez MV, Ruiz-Lorente I, et al. HLA-C*07 is associated with symptomatic HIV-1-associated neurocognitive disorders (HAND) and immune dysregulation. *Infect Dis Nor.* (2024) 56:818–29. doi: 10.1080/23744235.2024.2351047

32. Tian C, Kasavajhala K, Belfon KAA, Raguette L, Huang H, Migues AN, et al. ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J Chem Theory Comput.* (2020) 16:528–52. doi: 10.1021/acs.jctc.9b00591

33. Mark P, Nilsson L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J Phys Chem A.* (2001) 105:9954–60. doi: 10.1021/jp003020w

34. Roe DR, Cheatham TE. III, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput.* (2013) 9:3084–95. doi: 10.1021/ct400341p

35. Saitkinova S, Klaeger S, Le PM, Li LW, Oliveira G, Keshishian H, et al. A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat Biotechnol.* (2019) 38:199–209. doi: 10.1038/s41587-019-0322-9

36. Karnaukhov V, Paes W, Woodhouse IB, Partridge T, Nicastri A, Brackenridge S, et al. HLA variants have different preferences to present proteins with specific molecular functions which are complemented in frequent haplotypes. *Front Immunol.* (2022) 13:1067463. doi: 10.3389/fimmu.2022.1067463

37. Stefani C, Sangalli A, Locatelli E, Federico T, Malerba G, Romanelli MG, et al. Increased prevalence of unstable HLA-C variants in HIV-1 rapid-progressor patients. *Int J Mol Sci.* (2022) 23:14852. doi: 10.3390/ijms232314852

38. Antinori A, Arendt G, Becker JT, Brew BJ, Byrd DA, Cherner M, et al. Updated research nosology for HIV-associated neurocognitive disorders. *Neurology.* (2007) 69:1789–99. doi: 10.1212/01.WNL.0000287431.88658.8b

39. Nightingale S, Ances B, Cinque P, Dravid A, Dreyer AJ, Gisslen M, et al. Cognitive impairment in people living with HIV: consensus recommendations for a new approach. *Nat Rev Neurol.* (2023) 19:424–33. doi: 10.1038/s41582-023-00813-2

40. Shen Y, Parks JM, Smith JC. HLA class I supertype classification based on structural similarity. *J Immunol.* (2023) 210:103–14. doi: 10.4049/jimmunol.2200685

41. Nguyen AT, Szeto C, Gras S. The pockets guide to HLA class I molecules. *Biochem Soc Trans.* (2021) 49:2319–31. doi: 10.1042/bst20210410

42. Kangueane P, Sakharkar MK, Rajarager G, Bolisetty S, Sivasekari B, Zhao B, et al. A framework to sub-type HLA supertypes. *Front Biosci.* (2005) 10:879–86. doi: 10.2741/1582

43. Liu Q, Yang M, Zhong P, Wei Q, Jiao H, Meng J, et al. Micropolymorphism outside the peptide-binding groove of human leukocyte antigen (HLA)-C*14 modulates structural stability and shapes immune responses. *Int J Biol Macromol.* (2025) 309:142772. doi: 10.1016/j.ijbiomac.2025.142772

44. Yang M, Zhong P, Liu Q, Jiao H, Lei J, Wei P. Biochemical and structural insights into position 97 micropolymorphisms in human leukocyte antigen (HLA)-C*12 allotypes and their differential disease associations. *Int J Biol Macromol.* (2025) 306:141681. doi: 10.1016/j.ijbiomac.2025.141681

45. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, et al. A whole-genome association study of major determinants for host control of HIV-1. *Science.* (2007) 317:944–7. doi: 10.1126/science.1143767

46. Dendrou CA, Petersen J, Rossjohn J, Fugger L. HLA variation and disease. *Nat Rev Immunol.* (2018) 18:325–39. doi: 10.1038/nri.2017.143

**Appendix 5**

**Title**: Support Vector Machine Identification of Small Molecule Binders to an Understudied Allosteric Site of SARS-CoV-2 Mpro for Next-Generation PROTAC-Based Therapeutics

**Authors**: Fassi, E. M. A., Mekni, N., **Albani, M**., Maehrlein, S., Weldert, A. C., Schirmeister, T., Langer, T., and Grazioso, G.

**Status**: Published in *Archiv der Pharmazie* (2025)

**DOI**: 10.1002/ardp.70169

**Author Contribution**: Dr. M. Albani performed microscale thermophoresis experiments and contributed to manuscript drafting.

*Archiv der Pharmazie*

DPhG

FULL PAPER OPEN ACCESS

# Support Vector Machine Identification of Small Molecule Binders to an Understudied Allosteric Site of SARS-CoV-2 Mpro for Next-Generation PROTAC-Based Therapeutics

Enrico Mario Alessandro Fassi[1] | Nedra Mekni[2] | Marco Albani[1] | Sabine Maehrlein[3] | Annabelle Carolin Weldert[3] | Tanja Schirmeister[3] | Thierry Langer[2] | Giovannf razioso[1]

[1]Department of Pharmaceutical Sciences, Università degli Studi di Milano, Milano, Italy | [2]Department of Pharmaceutical Sciences, Division of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria | [3]Department of Medicinal Chemistry, Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University, Mainz, Germany

**Correspondence:** Nedra Mekni (nedram19@univie.ac.at)

**Received:** 31 July 2025 | **Revised:** 22 October 2025 | **Accepted:** 3 December 2025

**Keywords:** 3CL protease | covid-19 | drug design | machine learning | Mpro | PROTAC

## ABSTRACT

The emergence of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has underscored the urgent need for novel antiviral strategies. One of the primary targets of interest is the SARS-CoV-2 main protease (Mpro), which plays a crucial role in viral replication. Building on our prior work involving machine learning (ML)-based virtual screening for potential Mpro inhibitors, we sought to experimentally validate top-ranked candidates. Microscale thermophoresis (MST) was used to assess the binding affinity, leading to the identification of three promising hits from a library of 180 compounds. Notably, one compound demonstrated high-affinity binding to SARS-CoV-2 Mpro ($K_d = 2.8 \pm 0.9\ \mu M$). However, enzymatic assays revealed that none of the hit compounds inhibited the activity of the protease, suggesting a non-competitive binding. Docking and molecular dynamics (MD) simulations allowed to identify an accessory site in which the compounds exhibited stable interactions. These findings suggest that the identified compounds may serve as a starting point for the rational design of degradation-inducing strategies, such as proteolysis-targeting chimeras (PROTACs), targeting SARS-CoV-2 Mpro, and highlight the value of integrating ML-driven discovery with biophysical and computational validation in antiviral drug development.

## 1 | Introduction

The global coronavirus disease 2019 (COVID-19) outbreak has intensified the demand for potent antiviral agents [1]. A key enzyme in the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) replication process is represented by main protease (Mpro), also known as 3CL protease, which gained considerable attention as a drug target [2]. Its therapeutic relevance is reinforced by its high sequence conservation across β-coronaviruses and its lack of close homologs in the human proteome, minimizing the potential off-target side effects [3]. Additionally, given its essential role in the viral life cycle, Mpro has become a central focus for the development of selective inhibitors with clinical potential [4]. In 2022, the FDA approved the combination of ritonavir (an anti-HIV drug) and nirmatrelvir (also known as PF-07321332) with the commercial name Paxlovid (Pfizer Inc., New York, NY, USA), for treating severe COVID-19, underscoring the importance of incorporating Mpro inhibitors into the therapeutic arsenal against the virus [5, 6].

SARS-CoV-2 Mpro is a homodimeric cysteine protease resembling chymotrypsin, featuring a non-standard catalytic dyad comprising Cys145 and His41 [2]. This dyad forms crucial hydrogen bonds with a water molecule necessary for hydrolyzing the amide bond of substrates [7]. SARS-CoV-2 Mpro, along with the SARS-CoV-2 papain-like protease (PLpro), plays

a pivotal role in processing viral polyproteins, acting during the initial phases of viral replication within host cells [8]. Inhibiting one or both enzymes represents a promising therapeutic approach to impede the proliferation of SARS-CoV-2 within infected cells. The literature documents a variety of compounds containing diverse warheads, demonstrating either non-covalent or covalent activity against SARS-CoV-2 Mpro (Figure 1) [12].

Covalent inhibitors exert their inhibitory effects on SARS-CoV-2 Mpro by establishing covalent bonds with specific nucleophilic residues in the catalytic site, notably the cysteine residue at position 145 (Cys145) [13]. On the other hand, non-covalent inhibitors competitively bind to the catalytic site through classical hydrogen bonds, electrostatic, and non-polar interactions. To date, nirmatrelvir remains the only approved peptide-based therapeutic for COVID-19 that exerts its antiviral effect through covalent inhibition of the SARS-CoV-2 Mpro [14]. Conversely, most non-covalent Mpro inhibitors are still in the early stages of development [15–17]. For example, compound **47** (CCF981, Figure 1B) exhibited an $IC_{50}$ of 68 nM against SARS-CoV-2 Mpro and demonstrated $EC_{50}$ values of 0.50 and 0.56 μM in virus-infected VeroE6 cells, measured by cytopathic effect (CPE) inhibition and plaque reduction assays, respectively. Similarly, compound **51** (GC-14, Figure 1B), a 1,2,4-trisubstituted piperazine optimized through structure-based design, showed a significant antiviral activity in vitro, with an $EC_{50}$ of about 1.1 μM in assays using infected cells and an $IC_{50}$ of about 0.40 μM against Mpro [9–11]. Therefore, there is a pressing need to intensify research efforts toward the development of non-covalent inhibitors. These compounds show potential for addressing the challenges linked to peptide drugs, providing better pharmacological features like lower immunogenicity and increased stability [18].

Considering a previous study in which a support vector machine (SVM)-based machine learning (ML) model was developed to identify potential SARS-CoV-2 Mpro inhibitors [19], this work presents the next phase of our drug discovery pipeline, focusing on biophysical validation of the computational predictions. The SVM classifier was trained to distinguish active from inactive compounds based on physicochemical properties rather than structural similarity, offering an advantage in

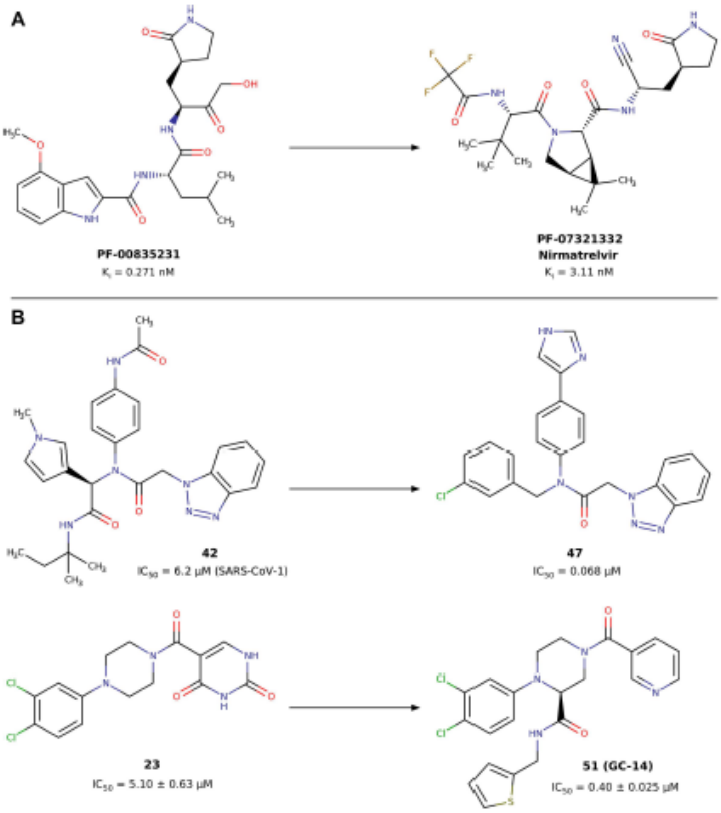



**FIGURE 1** | Covalent and non-covalent inhibitors of the SARS-CoV-2 Mpro. (A) The covalent inhibitor nirmatrelvir (also known as PF-07321332). (B) Non-covalent inhibitors: compound **47** (CCF981) derived from optimization of lead **42** (ML300), and hit compound **23** (MCULE-5948770040), selected as lead compounds in the rational design of highly potent and selective SARS-CoV-2 Mpro inhibitors **51** (**GC-14**) [9–11].

scenarios where mechanistic understanding is limited [19]. In the field of drug design, SVM has proven effective in various drug design applications, including predicting drug metabolism, identifying P-glycoprotein substrates, and evaluating blood–brain barrier permeability [20]. After extensive model optimization and cross-validation, 180 top-ranked commercially available compounds were selected and experimentally tested through microscale thermophoresis (MST). Despite exhibiting promising binding profiles in those biophysical assays, the most promising compounds failed to inhibit enzymatic activity, leading to the hypothesis that their binding does not occur in the active site of SARS-CoV-2 Mpro.

Consequently, in this study, we report on the computational studies aimed at the identification of the accessory site of SARS-CoV-2 Mpro in which the selected compounds exhibit stable interactions. Specifically, using a combination of pocket detection algorithms, molecular docking studies, and molecular dynamics (MD) simulations, we identified a pocket that has received limited attention in the literature, further suggesting its potential as a target for antiviral drug development. In fact, while most therapeutic efforts have focused on targeting the catalytic site of SARS-CoV-2 Mpro, accessory sites represent promising targets for designing alternative antiviral drugs with potentially lower resistance risk [21]. Depending on their position, the external sites may influence the catalytic activity of the enzyme, in which case they are referred to as "allosteric sites." In contrast, when the accessory sites do not alter enzymatic activity, they can serve as strategic anchoring points for designing molecular glues such as PROteolysis-TArgeting Chimeras (PROTACs). PROTACs represent a novel class of promising drugs inducing the selective degradation of disease-associated proteins, with significant applications in different fields. PROTACs are heterobifunctional molecules that simultaneously bind a target protein and an E3 ubiquitin ligase, facilitating the formation of a ternary complex [22, 23]. The E3 ligase itself is in complex with an activated ubiquitin-loaded E2 enzyme, and this close association facilitates the (poly)-ubiquitination of the target protein, marking it for degradation by the 26S proteasome [24]. In particular, they consist of a "warhead" group that binds to a target protein and an "anchor" group that binds to the substrate-binding domain of an E3 ubiquitin ligase, connected by a more or less flexible linker [23, 25].

Several studies have reported the development of PROTACs targeting the SARS-CoV-2 Mpro. Notably, Alugubelli and coworkers designed PROTAC degraders derived from previously characterized reversible covalent SARS-CoV-2 Mpro inhibitors (MPI8 and MPI29), conjugated to a CRBN E3 ligase anchor, which reported to reduce Mpro levels in human cells through a CRBN-dependent, proteasome-mediated mechanism [26–28]. Among these, MPD2 demonstrated potent degradation of Mpro in both 293T cells and SARS-CoV-2-infected A549-ACE2 cells, exhibiting antiviral activity across multiple viral strains, including nirmatrelvir-resistant variants, with an $EC_{50}$ of approximately 492 nM [27]. In addition, a series of indomethacin-based PROTACs recruiting either VHL or CRBN E3 ligases were developed, showing broad-spectrum antiviral effects, with enhanced antiviral activity compared with indomethacin alone, achieving low micromolar to nanomolar $EC_{50}$ values [29].

Collectively, these findings establish that PROTACs can be rationally designed to selectively degrade SARS-CoV-2 Mpro, offering a promising therapeutic strategy that may overcome challenges such as drug resistance and toxicity associated with conventional inhibitors. In fact, unlike conventional orthosteric ligands that function by occupying the active site to block activity, PROTACs act through a catalytic mechanism, eliminating the protein entirely. This catalytic mechanism allows PROTACs to exert their effects at lower concentrations compared with traditional inhibitors, which often require higher systemic exposure for therapeutic efficacy [30, 31].

## 2 | Results and Discussion

### 2.1 | Selection of the SARS-CoV-2 Mpro Inhibitors

In a previous study, published by some of us, an ML-based virtual screening model was developed to identify novel SARS-CoV-2 Mpro inhibitors [19]. The model architecture employed a supervised classification framework using an SVM with a radial basis function kernel [32] and was trained on the PostEra COVID-19 Moonshot dataset, an open-access collection of small molecules annotated with experimentally measured $IC_{50}$ values [33]. In this work, the trained model (see Section 4 for details) was employed to perform a virtual screening of a commercial compound library comprising approximately two million small molecules obtained from different vendors such as MolPort, Asinex, and ChEMBL. Out of 200 compounds predicted as potential Mpro inhibitors, 180 were selected for experimental validation using MST due to their commercial availability.

### 2.2 | Biophysical Experiments

Following an experimental protocol previously reported by us [34, 35], MST experiments were conducted to estimate the $K_d$ values of the 180 compounds selected from the SVM-based virtual screening model against the SARS-CoV-2 Mpro, quantifying interactions between the designed ligands and the protein binding site [19]. This biophysical technique allows for the analysis and measurement of molecular interactions between two entities under solution equilibrium conditions, avoiding the need for the sample immobilization which could potentially interfere with binding, as in the case of surface plasmon resonance method [36]. Initially, binding check experiments were accomplished using a ligand fixed concentration of 50 μM (see Section 4 for details). This initial step was designed to determine which small molecules are capable of binding to the SARS-CoV-2 Mpro protein at the tested concentration. Out of the 180 compounds screened, 17 showed positive results in the binding assay and were chosen for further analysis of their binding affinities. The chemical structures of the 17 compounds are provided in the Supporting Information S1: Table S1. To obtain a complete $K_d$ curve, a fixed concentration of labeled SARS-CoV-2 Mpro enzyme was mixed with 16 1:1 serial dilutions of the compounds, ranging in concentration from 250 μM to 7.6 nM (see Section 4 for details). Three out of the 17 tested compounds (**7**, **8**, and **9**) displayed a clear $K_d$ curve in two independent replicas. In the remaining cases, a significant number of outlier points prevented the computation of the complete affinity curve making impossible to estimate the $K_d$ value for those compounds. Among the compounds showing

clear $K_d$ curves, the compound **7** showed the highest affinity for SARS-CoV-2 Mpro enzyme displaying a $K_d$ of $2.8 \pm 0.9\,\mu M$, followed by **8** ($K_d = 23.9 \pm 7.4\,\mu M$), and **9** ($K_d = 39.0 \pm 23.4\,\mu M$) as detailed in Table 1 and shown in Figure 2. The summary table of the MST assays experimental conditions performed, and the detailed MST $K_d$ curve of each compound are available in the Supporting Information S1: Table S2 and Figure S1, respectively. In the case of compound **7**, the first three concentration points (corresponding to 250, 125, and 62.5 µM, respectively) were excluded by fitting, since they seem to be part of another binding curve observed in the high micromolar range, probably due to its nonspecific binding on the SARS-CoV-2 Mpro protein (Supporting Information S1: Figure S2).

### 2.3 | Enzymatic Assays

The three compounds identified as binders in the MST assay were tested for their inhibitory properties against SARS-CoV-2 Mpro at a 20 µM concentration (see Section 4 for details). However, none of the compounds caused a significant reduction in the enzyme's activity (Table 2 and Supporting Information S1: Figure S3).

These results suggest that, at the tested concentration, the compounds likely do not bind within the active site as orthosteric or allosteric ligands. Consequently, to investigate alternative binding sites, we conducted further computational studies, including the identification of protein accessory sites, molecular docking calculations, and MD simulations.

### 2.4 | Accessory Sites Identification

The SiteMap tool, available in Maestro software (Schrödinger LLC, New York, NY, USA), was used to identify potential binding sites on the crystal structure of the target enzyme (PDB accession code: 6W63), which were ranked based on their "SiteScore" values (Table 3). Notably, the SARS-CoV-2 Mpro structure comprises three domains: Domains I (residues 8-101) and II (residues 102–184) consist of antiparallel β-barrel structures and serve as catalytic domains, while Domain III (residues 201–303) comprises five α-helices and is responsible for enzyme dimerization (Figure 3A) [3].

SiteMap identified a total of five potential binding sites and correctly recognized the active site as a promising binding region. However, this was not ranked as the most favorable site (Table 3). On the other hand, two sites (Site1 and Site2) were found to be adjacent to each other and located at the interface between Domain II and Domain III, on the opposite side of the active site (Figure 3B). These two sites showed SiteScore values slightly

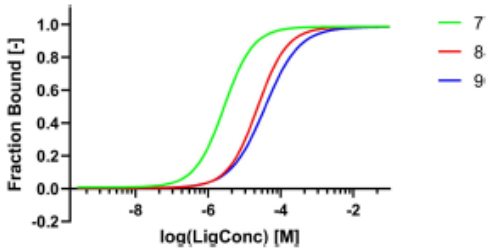
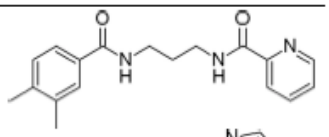
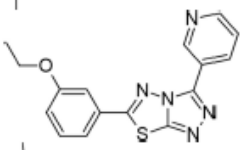
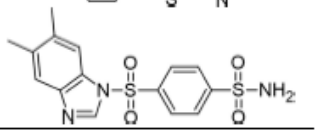



**FIGURE 2** | MST curves acquired by SARS-CoV-2 Mpro incubated with scaling concentrations of compounds **7** (green), **8** (red), and **9** (blue) in comparison with each other, using the Monolith NT.115 instrument. The curves were normalized by fraction bound. Two independent experiments were performed to compute the $K_d$ curve.

**TABLE 2** | Residual activity of compounds screened against SARS-CoV-2 Mpro at 20 µM in enzymatic assay.

| Compound ID | Residual activity (%) | Error |
|---|---|---|
| **7** | 104.66 | 3.03 |
| **8** | 100.80 | 2.94 |
| **9** | 106.45 | 4.02 |

**TABLE 1** | ID number, $K_d$ values, and chemical structure of compounds that displayed a clear $K_d$ curve in MST.

| Compound ID | $K_d \pm K_d$ confidence (µM) | Structure |
|---|---|---|
| **7** | 2.8 ± 0.9 |  |
| **8** | 23.9 ± 7.4 |  |
| **9** | 39.0 ± 23.4 |  |

higher (Site1) or comparable (Site2) to that of the active site (Table 3). Site3 and Site4 (Table 3), located within Domain II and Domain III, respectively, showed considerably low SiteScore values and were found to be very small and shallow, lacking well-defined pockets (Supporting Information S1: Figure S4). Thus, they were excluded from the following docking phase. Consequently, Site1 and Site2 were selected for molecular docking calculations.

It was interesting to note that these sites were also investigated by Günther and co-workers, whose employed X-ray crystallography to screen over 5000 compounds, either approved drugs or those in clinical trials, identifying 37 compounds targeting SARS-CoV-2 Mpro, primarily at the active site, but also revealing two additional binding sites [37]. Specifically, Site1 in our study was targeted by only a single compound, AT7519 (a CDK inhibitor) [38], for which the crystal structure has also been published (PDB accession code 7AGA) [37]. However, AT7519 displayed weak antiviral activity ($EC_{50} = 25.2\,\mu M$), and its inhibitory effect on SARS-CoV-2 Mpro enzymatic activity was not experimentally evaluated [39]. The absence of these data prevents determining whether this site can be classified as an allosteric pocket. Indeed, Samrat and colleagues reported on the identification of three niclosamide derivatives as potent non-competitive inhibitors of SARS-CoV-2 Mpro, with $IC_{50}$ values comparable to that of the known covalent inhibitor boceprevir, thereby suggesting a potential allosteric mechanism of inhibition [40, 41]. Through computational studies, including molecular docking and MD simulations, the authors provided an evidence that the binding site of these compounds could be the same site targeted by the compound AT7519 [37, 41].

**TABLE 3** | Potential binding sites identified by SiteMap, along with their SiteScore, size, and volume values.

| Binding site | SiteScore | Size (Å$^2$) | Volume (Å$^3$) |
|---|---|---|---|
| Site1 | 0.83 | 74 | 183 |
| Active Site | 0.81 | 46 | 129 |
| Site2 | 0.80 | 51 | 159 |
| Site3 | 0.56 | 32 | 61 |
| Site4 | 0.51 | 16 | 66 |

## 2.5 | Docking and MD Simulations

The residues defining the Site1 and Site2 pockets (see Section 4 for details) were selected to construct the receptor grid, ensuring that the bounding box fully encompassed the spatial volume of both binding sites, thereby capturing all relevant interactions for subsequent docking simulations. Subsequently, molecular docking calculations were conducted for the three ligands targeting the SARS-CoV-2 Mpro using the Extra Precision (XP) mode of the Glide module implemented in the Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA) [42]. Interestingly, the obtained docking scores are in agreement with the experimental binding affinities determined by MST, providing a supporting evidence that the identified allosteric pocket may represent the actual binding site of these compounds. Specifically, compound **7** exhibited the most favorable docking score of −3.440 kcal/mol, consistent with its highest binding affinity, while compounds **8** and **9** showed lower scores of −3.279 and −2.026 kcal/mol, respectively. For each compound, the top-ranked binding pose was subjected to three independent 500 ns MD simulations using the Desmond module integrated within the Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA), with the goal of evaluating the stability and binding mode of the ligands over the simulation time within the identified allosteric site (see Section 4 for details).

The MD simulations revealed that all three ligands remained stably bound within the identified allosteric site of SARS-CoV-2 Mpro throughout the entire simulation period (Supporting Information S1: Figure S5). Notably, compound **7** exhibited the most significant rearrangement from its initial docking pose, as indicated by the highest RMSD peaks before stabilizing into a conformation that maintained persistent interactions with key residues of the allosteric site (Supporting Information S1: Figure S5).

To derive a reliable computational model of the ligand/Mpro complexes, we conducted a cluster analysis of the 1.5 µs MD simulations. The representative structures from the most populated clusters for compounds **7**, **8**, and **9** are shown in Figure 4. For compound **7**, which represents about the 49% of the conformational ensemble explored, the interaction network with SARS-CoV-2 Mpro is predominantly characterized by a π–π stacking interaction with Phe294, along with several hydrophobic contacts

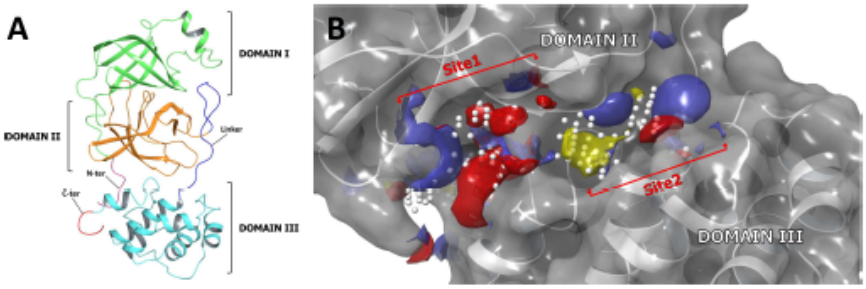



**FIGURE 3** | General overview of Mpro structure. Representation (A) of the three structural domains of SARS-CoV-2 Mpro (Domain I: green, Domain II: orange, Domain III: cyan) and of (B) the two allosteric sites (Site1 and Site2), located at the interface between Domain II and Domain III, identified with SiteMap.

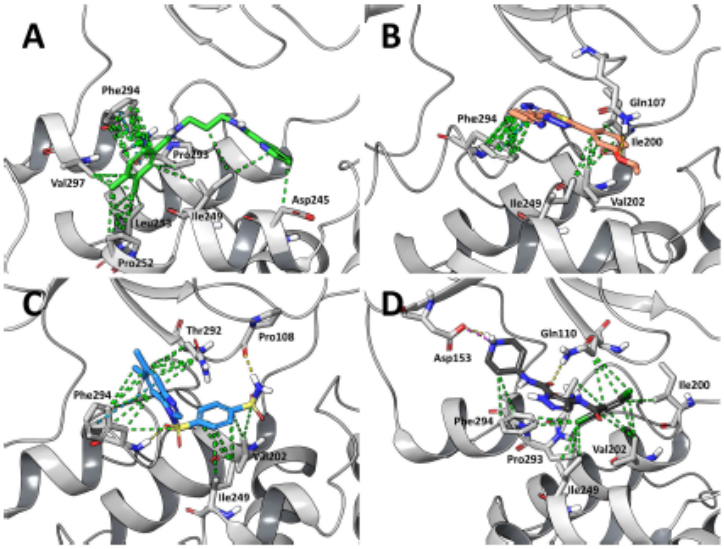



**FIGURE 4** | Binding modes of selected compounds. Representative structure from the most populated cluster of (A) compound **7** (green sticks), (B) compound **8** (light red sticks), and (C) compound **9** (light blue sticks) in complex with SARS-CoV-2 Mpro (white cartoon) after 500 ns of MD simulation. Hydrophobic, H-bond, and π–π interactions are represented as green, yellow, and cyan dashed lines, respectively. (D) Crystal structure of AT7519 (dark gray sticks) in complex with SARS-CoV-2 Mpro (white cartoon), with PDB accession code 7AGA.

involving surrounding nonpolar residues that form a hydrophobic pocket (Ile249, Pro252, Leu253, Pro293, Phe294, and Val297), which collectively contribute to the stabilization of the entire ligand within the allosteric binding pocket (Figure 4A). The representative structure of compound **8**, which constitutes the 64.8% of the sampled conformational ensemble, predominantly exhibits hydrophobic interactions, although fewer than those observed for compound **7**. Notably, it forms interactions with Phe294 (including a π–π stacking interaction), Ile200, Val202, and Ile249. Additionally, a hydrogen bond with Gln107 is observed (Figure 4B). This limited interaction profile may account for the lower binding affinity experimentally determined by MST (Figure 2). Compound **9**, representing about the 86% of the sampled conformational ensemble, forms two hydrogen bonds with the backbone amine of Phe294 and the carboxyl group of Pro108 and an important hydrophobic interaction network, including a π–π stacking interaction with Phe294 (Figure 4D). Unexpectedly, despite these interactions, it shows the lowest binding affinity among the three compounds based on MST data (Table 1).

The binding modes of the compounds were compared with the crystallographic pose of AT7519 (Figure 4D). Superimposition of the complexes (Supporting Information S1: Figure S6) revealed that none of our compounds induce a significant conformational change in the apo protein via interaction with Asp153, unlike what is observed for AT7519 [37]. This interaction is critical since the conformational change induced by AT7519 facilitates the formation of strong interactions between Asp153 and Arg298, including H-bonds and salt bridges and it is important to note that Arg298 is essential for dimerization [37, 43]. In particular, its replacement with Ala residue shifts the spatial arrangement between the dimerization and catalytic domains, thereby altering the oxyanion hole and weakening the S1 pocket through perturbations of the N-terminal structure [43]. This difference may explain why compounds **7**, **8**, and **9** may have no effect on the catalytic activity of SARS-CoV-2 Mpro at the tested concentration of 20 μM (Table 2). However, it is also noteworthy that AT7519 exhibits only low antiviral activity against SARS-CoV-2 replication in Vero E6 cells, with an $EC_{50}$ of 25.16 μM, suggesting that the binding site for AT7519 (and likely our compounds as well) may have limited influence on the enzyme's physiological function.

The generation of these computational binding models enabled a deeper understanding of the molecular regions of the compounds most involved in the interactions with SARS-CoV-2 Mpro, as well as those predominantly solvent-exposed. This insight is crucial for guiding the strategic addition of functional moieties needed for the design of potential PROTACs.

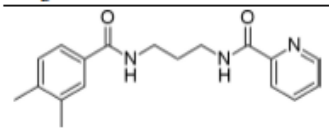
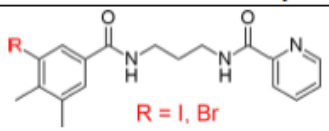
### 2.6 | Rational Design of Protacs

Here, we propose the design of potential warheads for PROTAC development targeting SARS-CoV-2 Mpro, based on the compounds we have identified that exhibit a binding affinity to the target protein in the micromolar range, as shown by MST experiments (Table 1). The proposed warheads were rationalized based on their favorable binding modes observed after three independent replicas of 500 ns-long MD simulations (Figure 4), and considering the synthetic feasibility. This latter point was verified through the retrosynthesis tool implemented into CAS SciFinder (Columbus, OH, USA). Based on the chemical structures of the compounds, classical click-type reactions to generate

the corresponding PROTACs can be accomplished, primarily via the Sonogashira cross-coupling and, in the case of compound **8**, also through the Huisgen cycloaddition. Both the Sonogashira cross-coupling of terminal alkynes with aryl iodides, bromides, or chlorides (not preferable, as it operates under more forcing conditions), and the Huisgen cycloaddition between an azide and an alkyne are widely used for the introduction of alkynyl or 1,2,3-triazole linkers, respectively, in the assembly of PROTACs [44, 45]. Given the significantly higher affinity observed for compound **7**, as well as its better synthetic feasibility, this molecule is undoubtedly the most promising among the three to serve as a starting point for PROTAC design (Table 4). Therefore, compound **7** is described in detail here (Figure 5), while the characterization of the other compounds are described in a dedicated section of the Supporting Information, and represented in Supporting Information S1: Figures S7–S10.

The analysis of the compound **7** binding mode within the accessory pocket reveals that the free meta-position of the dimethyl-benzene group is the most favorably oriented for linker attachment (Figure 4A), as it is predominantly solvent-exposed. Hence, we recommend its substitution with iodides or bromides in the rational design of a PROTAC applying the Sonogashira cross-coupling reaction (Table 4). The retrosynthetic approach suggested that bromine is the preferred halogen substituent, with an estimated yield of 49% and an estimated cost of $44 per 100 g, versus a 29% estimated yield and a cost of $73 per 100 g for iodine. In particular, the commercially available 3,4-dimethylbenzoic acid was first brominated to yield 3-bromo-4,5-dimethylbenzoic acid. This intermediate was then amidated with commercially available 1,3-propanediamine to produce N-(3-aminopropyl)-3-bromo-4,5-dimethylbenzamide. Subsequently, a second amidation with commercially available picolinic acid afforded N-(3-(3-bromo-4,5-dimethylbenzamido)propyl)picolinamide (Table 4), which represents the final compound to be used as substrate for the subsequent Sonogashira cross-coupling reaction to synthesize the PROTAC. The retrosynthesis scheme is reported in Figure 5.

**TABLE 4** | Proposed optimal linker position for rational design of potential PROTACs targeting SARS-CoV-2 Mpro based on the compound **7**, indicated by red "R" labels.

| **Compound 7** | |
|---|---|
| **Original structure** | **Intermediate relevant to PROTAC synthesis** |
|  |   |

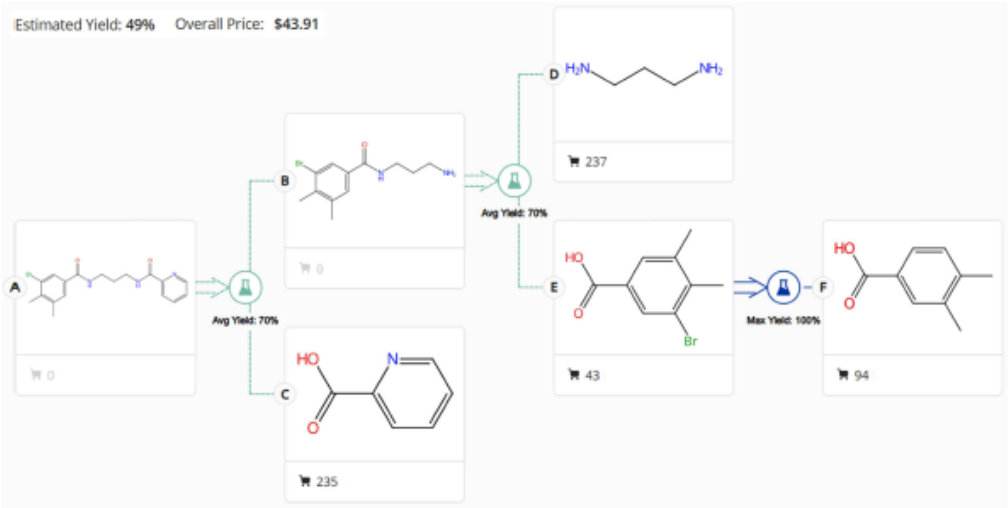



**FIGURE 5** | Retrosynthetic analysis of compound **7** intermediate relevant to PROTAC synthesis. The black cart indicates that the product is commercially available and the number of vendors. The blue reaction symbol indicates that the reaction has been already published in literature, while the green one indicates that the reaction is possible, although do not already accomplished using the specific reagents. The scheme was generated through the "retrosynthesis tool" available on CAS-SciFinder web platform.

## 3 | Conclusions

The integrated computational and biophysical approaches employed in this study demonstrated a promising strategy for the discovery of small-molecule binders targeting SARS-CoV-2 Mpro, with potential application in the rational design of innovative targeted protein degradation therapeutics, such as PROTACs. In particular, using an SVM-based virtual screening, we prioritized 180 candidate binders covering a huge chemical space, comprising approximately two million compounds retrieved form MolPort, Asinex, and ChEMBL database. The subsequent validation, using MST, identified three promising hits with binding affinities in the low micromolar range, with compound **7** exhibiting a $K_d$ value of $2.8 \pm 0.9\,\mu M$. The lack of enzymatic inhibition at the tested concentration of $20\,\mu M$ strongly suggested a non-competitive binding mode, prompting us to investigate the possibility of alternative interaction sites on the protein surface. Further computational analyses revealed consistent and energetically favorable interactions within an accessory pocket, spatially distinct from the catalytic site. Finally, we proposed the structure of next-generation PROTACs directed against SARS-CoV-2 Mpro based on the scaffolds of the compounds here disclosed. This strategy not only expands the chemical space for antiviral drug development but also holds potential for improved safety profiles by reducing off-target effects and associated toxicities typically observed with conventional inhibitors.

## 4 | Materials and Methods

### 4.1 | SVM Model

The model was trained on the PostEra COVID-19 Moonshot dataset, an open-access collection of small molecules first released in March 2020, in which each compound is annotated with an experimentally measured $IC_{50}$ [33]. All compounds were imported as SMILES and converted into 3D molecular structures using RDKit (www.rdkit.org/docs/api-docs.html). This preprocessing included the addition of explicit hydrogens and the generation of low-energy conformers using the ETKDG algorithm [46]. We then computed 1444 one- and two-dimensional descriptors with PaDEL-Descriptor, capturing topological indices, electronic features, atomic-property distributions, and physicochemical parameters [47]. Given the high dimensionality of the descriptor space relative to the number of compounds, and to prevent overfitting during model training, we implemented a feature selection strategy. This approach combined random forest (RF) importance ranking with recursive feature elimination and fivefold cross-validation (RF-RFE-CV), using the scikit-learn library [48–50]. Through this process, we identified seven descriptors with the highest predictive value: *AATS6i* and *ATSC7m* (2D autocorrelation descriptors encoding the spatial distribution of atomic properties such as mass and polarizability), *VE1_DzZ* (a Barysz matrix descriptor capturing electronic and topological information), three Burden-modified eigenvalues (*SpMax2_Bhm*, *SpMax1_Bhv*, and *SpMax2_Bhv*), and *CrippenLogP*, which estimates lipophilicity. Overall, the selected descriptors reflect molecular features essential for ligand–target interactions, such as shape complementarity, electrostatic properties, and hydrophobicity. The reduced descriptor set was finally used to train the final SVM model. Hyperparameters were optimized via grid search using scikit-learn's GridSearchCV, resulting in the selection of $C = 1.0$ and $\gamma = 1.0$. The model demonstrated strong generalization performance, achieving 88% classification accuracy and 75% precision on an independent test set. In this context, $C$ controls the trade-off between maximizing the margin and minimizing classification errors, while $\gamma$ defines the influence of individual data points on the decision boundary. Careful optimization of these parameters was essential to avoid overfitting given the limited dataset size, and to ensure robust generalization of the model.

### 4.2 | MST Experiments

The binding of small molecules to the SARS-CoV-2 Mpro enzyme was assessed using the Monolith NT.115 instrument (NanoTemper Technologies GmbH, Munich, Germany). Recombinant His-tagged SARS-CoV-2 Mpro (GeneTex, Irvine, CA, USA) was fluorescently labeled using the His-Tag Labeling Kit RED-tris-NTA 2nd Generation (MO-L018, NanoTemper Technologies GmbH, Munich, Germany) for 30 min at room temperature. A fixed concentration of red-labeled SARS-CoV-2 Mpro (35 nM) was mixed with each of 180 small molecules at a fixed concentration ($50\,\mu M$) to perform a binding check, using the "Expert Mode" of the MO. Control v1.6 software. The β-amido boronic acid compound **3ea** was included as a positive control [35]. The binding was determined by monitoring compound concentration-dependent changes in the normalized fluorescence (Fnorm), focusing on the 2.5-second time point in all binding assays. A response amplitude (RA), defined as the Fnorm difference between the ligand-bound and unbound states of SARS-CoV-2 Mpro, was considered indicative of binding when RA ≥ 1.5. The compounds that tested positive in the binding check assay underwent further assessment through the binding affinity experiment on the SARS-CoV-2 Mpro enzyme to generate a complete $K_d$ curve. Initially, the protein was labeled using the same protocol, and then a fixed concentration of labeled SARS-CoV-2 Mpro (35 nM) was mixed with 16 1:1 serial dilutions of small molecules, ranging in concentration from $250\,\mu M$ to 7.6 nM. Throughout both binding check and binding affinity MST experiments, the enzyme and ligands were incubated for 15–50 min at room temperature, employing a medium MST power (40%) and an excitation power of 80% at 25°C, and using standard capillaries (NanoTemper Technologies GmbH, München, Germany). Both interacting species were dissolved in PBS-T buffer (phosphate-buffered saline + 0.05% Tween 20) from NanoTemper Technologies (GmbH, München, Germany) and 2.5% dimethyl sulfoxide (DMSO) for molecular biology (Product No. D8418; Sigma-Aldrich, Saint Louis, USA). The Monolith software MO. Affinity Analysis v2.3 was used to generate the full analysis report, while the figures were generated using GraphPad Prism software v8.0.2 (GraphPad, Boston, USA).

### 4.3 | Protein Expression and Purification of SARS-CoV-2 Mpro

The protein was expressed and purified as described previously [51]. Briefly, the pE-SUMO vector (Kindly provided by Prof. Dr. Harald Schwalbe, University of Frankfurt), encoding the SARS-CoV-2 Mpro protein with a hexahistidine ($His_6$) tag attached to

a small ubiquitin-like modifier (SUMO) tag at the N-terminus, was transformed into competent *Escherichia coli* (E. coli) BL21 Gold (DE3) cells (Agilent Technologies, Santa Clara, CA, USA). The cells were cultured at 37°C with shaking at 160 rpm in Luria Broth (LB) medium supplemented with 100 µg/mL ampicillin. Overexpression was induced by adding 0.2 mM isopropyl-β-D-thiogalactopyranosid (IPTG), and the culture was incubated overnight at 18°C. The cells were then harvested by centrifugation at 4700 rpm for 1 h at 4°C. The resulting pellet was resuspended in 100 mL of lysis buffer (50 mM NaPi pH 7.5, 300 mM NaCl, 5 mM imidazole, 5% (v/v) glycerol, 10 mM β-ME), which was supplemented with RNase, DNase, and lysozyme. The suspension was sonicated for 10 cycles of 30 s at 60% power on ice, with 10-s pauses between each pulse. To remove insoluble material, the lysate was centrifuged at 20,000 rpm for 1 h at 4°C.

The protein purification was carried out using an ÄKTA start protein purification system (GE Healthcare, Chicago, IL, USA). The supernatant was loaded onto a HisTrap HP 5 mL column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany), which was pre-equilibrated with the lysis buffer. The column was washed with five column volumes (CV) of the same buffer. Elution was achieved by applying a linear gradient of elution buffer (50 mM NaPi (pH 7.5), 300 mM NaCl, 500 mM imidazole, 5% (v/v) glycerol, 10 mM β-ME). To exchange the elution buffer, the solution was passed through an Amicon Ultra-15 Centrifugal filter (Millipore, Billerica, MA, USA) with a 10 kDa cutoff, replacing it with Ulp-1 cleavage buffer (50 mM NaPi pH 7.0, 300 mM NaCl, 10 mM β-ME, 5% (v/v) glycerol). For tag removal, 250 units of $His_6$-tagged Ulp-1 protease (Sigma-Aldrich, St. Louis, MO, USA) were added, and the reaction proceeded overnight at 4°C. The mixture was loaded onto a HisTrap HP 5 mL column, collecting the flow-through to remove the Ulp-1 protease and the tags. Lastly, the protein underwent size exclusion chromatography (SEC) on a HiLoad 16/600 Superdex 75 pg column (Cytiva Europe GmbH, Freiburg im Breisgau, Germany), pre-equilibrated with SEC buffer (25 mM NaPi pH 7.5, 150 mM NaCl, 2 mM DTT). Afterward, the protease was exchanged into 10% (v/v) glycerol, aliquoted, flash frozen in liquid nitrogen, and stored at –80°C until further use.

### 4.4 | Enzymatic Assays

Inhibitory activities of the compounds against SARS-CoV-2 Mpro were determined as described previously by cleavage of a fluorescence resonance energy transfer (FRET) peptide substrate (Dabcyl-KTSAVLQSGFRKME-Edans), purchased from Genescript (New Jersey, USA) [52].

Fluorescence measurements were carried out in white flat-bottom 96-well plates (Greiner Bio-One, Kremsmünster, Upper Austria) using a Tecan Spark 10M plate reader (Tecan Group Ltd., Männedorf, Switzerland). Each well contained a final volume of 200 µL, consisting of 180 µL assay buffer (20 mM TRIS-HCl pH 7.5, 20 mM NaCl, 0.1 mM EDTA, 1 mM DTT), 5 µL enzyme solution (final concentration 75 nM), 10 µL inhibitor in DMSO (final concentration 20 µM) or pure DMSO as a negative control, and 5 µL substrate solution in DMSO (final concentration 25 µM). The reaction was monitored for 10 min at 25°C, with fluorescence readings taken at 30-s intervals. The substrate was excited at 335 nm, and fluorescence emission was recorded at 493 nm. The remaining enzymatic activity was determined by comparing the substrate hydrolysis rate in reaction mixtures containing the inhibitors to that of the DMSO control. Only the first 5 min of the fluorescence curves (Supporting Information S1: Figure S3) were used for the calculation to obtain the linear part.

### 4.5 | Computational Studies

The SARS-CoV-2 Mpro crystal structure used in this study was retrieved from the Protein Data Bank (PDB accession code: 6W63), in which the protein is bound to a potent broad-spectrum non-covalent inhibitor (namely, X77) [53]. Protein structure refinement was carried out using the "*Protein Preparation Wizard*" tool of Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA). The protocol included: (i) addition of hydrogen atoms and reconstruction of any missing side chains; (ii) optimization of hydrogen-bonding networks; and (iii) restrained energy minimization of all atoms to a maximum RMSD value of 0.3 Å, applying the OPLS4 force field. Co-crystallized ligands were removed before preparation. Binding site prediction was performed using the "*SiteMap*" tool of Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA), considering all protein atoms and employing default parameters. Compounds **7**, **8**, and **9** were manually designed and prepared using the "*LigPrep*" tool of Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA), applying the OPLS4 force field and Epik Classic for ionization state prediction at a pH of 7 ± 2. The docking grid was generated based on the residues defining the Site1 and Site2 pockets, namely Gln107, Pro108, Gln110, Asn151, Asp153, Tyr154, Val202, Asn203, Glu240, His246, Ile249, Thr292, Phe294, Asp295, and Arg298. The dimensions of the bounding box were set to encompass both binding sites. Docking calculations were performed using the "*Glide*" tool available in Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA) with the XP mode enabled. The number of poses to generate per ligand was set to 500, and the energy threshold for rejecting minimized poses was set to 0.5 kcal/mol. The reliability of the docking calculations was assessed by performing a self-docking calculation of AT7519 compound, in which the top-ranked binding pose was compared with its crystallographic conformation (PDB ID: 7AGA [37]), yielding an excellent agreement with an RMSD of 2.16 Å (Supporting Information S1: Figure S11). The three best docking poses of each compound were carefully examined to assess potential differences in binding modes, and the RMSD of the ligands were calculated after aligning the protein backbone atoms (Supporting Information S1: Figure S12). Since each compound essentially exhibited a single predominant binding pose, only the top-ranked pose of each compound/Mpro complex was immersed in an orthorhombic box of TIP3P water molecules using the "*System Builder*" tool in Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA), with box boundaries set to 10 Å from the protein surface. System neutrality was achieved by adding the appropriate number of counterions. Following solvation and energy minimization, each system underwent to three independent 500 ns MD simulations using the "*Desmond*" algorithm implemented in the

Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA) under constant conditions of 300 K and 1 atm. The protein/ligand stability (Cα RMSD plots) throughout the simulation was evaluated using the "*Simulation Interactions Diagram*" tool available in Maestro software suite (Release 2025-1, Schrödinger LLC, New York, NY, USA). Conformational clustering of ligand structures obtained from MD simulations was carried out using the GROMOS algorithm by Daura et al. [54], as implemented in the GROMACS software package (version 2024.3) [55]. Several clustering trials were conducted to determine an optimal RMSD cutoff, enabling clear discrimination among ligand conformations while reducing the prevalence of isolated (singleton) clusters (Supporting Information S1: Table S3).


## Acknowledgments

The MST experiments were accomplished thanks to the generous financial support of Dr. Ugo Perricone, Fondazione RI.MED (Via Bandiera 11, 90133 Palermo, Italy). G.G. and E.M.A.F. would like to thank INDACO for providing high-performance computing resources and support.

## Funding

The authors received no specific funding for this work.


## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

Data will be made available on request.

## References


1. M. Mazinani and B. J. Rude, "The Novel Zoonotic Coronavirus Disease 2019 (COVID-19) Pandemic: Health Perspective on the Outbreak," *Journal of Healthcare Quality Research* 36 (2021): 47–51, https://doi.org/10.1016/j.jhqr.2020.09.004.

2. K. Anand, "Structure of Coronavirus Main Proteinase Reveals Combination of a Chymotrypsin Fold With an Extra Alpha-Helical Domain," *EMBO Journal* 21 (2002): 3213–3224, https://doi.org/10.1093/emboj/cdf327.

3. Q. Hu, Y. Xiong, G. H. Zhu, et al., "The SARS-CoV-2 Main Protease (Mpro): Structure, Function, and Emerging Therapies for COVID-19," *MedComm* 3 (2022): e151, https://doi.org/10.1002/mco2.151.

4. S. Ning, B. Yu, Y. Wang, and F. Wang, "SARS-CoV-2: Origin, Evolution, and Targeting Inhibition," *Frontiers in Cellular and Infection Microbiology* 11 (2021): 676451, https://doi.org/10.3389/fcimb.2021.676451.

5. G. Navitha Reddy, A. Jogvanshi, S. Natkwadi, and R. Sonti, "Nirmatrelvir and Ritonavir Combination: An Antiviral Therapy for COVID-19," *Expert Review of Anti-Infective Therapy* 21 (2023): 943–955, https://doi.org/10.1080/14787210.2023.2241638.

6. B. Amani and B. Amani, "Efficacy and Safety of Nirmatrelvir/Ritonavir (Paxlovid) for COVID-19: A Rapid Review and Meta-Analysis," *Journal of Medical Virology* 95 (2023): e28441, https://doi.org/10.1002/jmv.28441.

7. C. A. Ramos-Guzmán, J. J. Ruiz-Pernía, and I. Tuñón, "Unraveling the SARS-CoV-2 Main Protease Mechanism Using Multiscale Methods," *ACS Catalysis* 10 (2020): 12544–12554, https://doi.org/10.1021/acscatal.0c03420.

8. S. A. Amin, S. Banerjee, K. Ghosh, S. Gayen, and T. Jha, "Protease Targeted COVID-19 Drug Discovery and Its Challenges: Insight into Viral Main Protease (Mpro) and Papain-Like Protease (PLpro) Inhibitors," *Bioorganic & Medicinal Chemistry* 29 (2021): 115860, https://doi.org/10.1016/j.bmc.2020.115860.

9. A. Clyde, S. Galanie, D. W. Kneller, et al., "High-Throughput Virtual Screening and Validation of a SARS-CoV-2 Main Protease Noncovalent Inhibitor," *Journal of Chemical Information and Modeling* 62 (2022): 116–128, https://doi.org/10.1021/acs.jcim.1c00851.

10. S. Gao, K. Sylvester, L. Song, et al., "Discovery and Crystallographic Studies of Trisubstituted Piperazine Derivatives as Non-Covalent SARS-CoV-2 Main Protease Inhibitors With High Target Specificity and Low Toxicity," *Journal of Medicinal Chemistry* 65 (2022): 13343–13364, https://doi.org/10.1021/acs.jmedchem.2c01146.

11. S. H. Han, C. M. Goins, T. Arya, et al., "Structure-Based Optimization of ML300-Derived, Noncovalent Inhibitors Targeting the Severe Acute Respiratory Syndrome Coronavirus 3CL Protease (SARS-CoV-2 3CLpro)," *Journal of Medicinal Chemistry* 65 (2022): 2880–2904, https://doi.org/10.1021/acs.jmedchem.1c00598.

12. L. Song, S. Gao, B. Ye, et al., "Medicinal Chemistry Strategies Towards the Development of Non-Covalent SARS-CoV-2 Mpro Inhibitors," *Acta Pharmaceutica Sinica B* 14, no. 1 (2024): 87–109, https://doi.org/10.1016/j.apsb.2023.08.004.

13. H. Kim, D. Hauner, J. A. Laureanti, K. Agustin, S. Raugei, and N. Kumar, "Mechanistic Investigation of SARS-CoV-2 Main Protease to Accelerate Design of Covalent Inhibitors," *Scientific Reports* 12 (2022): 21037, https://doi.org/10.1038/s41598-022-23570-6.

14. Y. N. Lamb, "Nirmatrelvir Plus Ritonavir: First Approval," *Drugs* 82 (2022): 585–591, https://doi.org/10.1007/s40265-022-01692-5.

15. H. Liu, A. Zask, F. Forouhar, et al., "Development of Small Molecule Non-Covalent Coronavirus 3CL Protease Inhibitors From DNA-Encoded Chemical Library Screening," *Nature Communications* 16 (2025): 152, https://doi.org/10.1038/s41467-024-55421-5.

16. Y.-Q. Xiao, J. Long, S.-S. Zhang, Y.-Y. Zhu, and S.-X. Gu, "Non-Peptidic Inhibitors Targeting SARS-CoV-2 Main Protease: A Review," *Bioorganic Chemistry* 147 (2024): 107380, https://doi.org/10.1016/j.bioorg.2024.107380.

17. N. E. Zhou, S. Tang, X. Bian, et al., "An Oral Non-Covalent Non-Peptidic Inhibitor of SARS-CoV-2 Mpro Ameliorates Viral Replication and Pathogenesis In Vivo," *Cell Reports* 43 (2024): 114929, https://doi.org/10.1016/j.celrep.2024.114929.

18. C. M. N. Allerton, J. T. Arcari, L. M. Aschenbrenner, et al., "A Second-Generation Oral SARS-CoV-2 Main Protease Inhibitor Clinical Candidate for the Treatment of COVID-19," *Journal of Medicinal Chemistry* 67 (2024): 13550–13571, https://doi.org/10.1021/acs.jmedchem.3c02469.

19. N. Mekni, C. Coronnello, T. Langer, M. D. Rosa, and U. Perricone, "Support Vector Machine as a Supervised Learning for the Prioritization of Novel Potential SARS-COV-2 Main Protease Inhibitors," *International Journal of Molecular Sciences* 22 (2021): 7714, https://doi.org/10.3390/ijms22147714.

20. C. Y. Liew, X. H. Ma, X. Liu, and C. W. Yap, "SVM Model for Virtual Screening of Lck Inhibitors," *Journal of Chemical Information and Modeling* 49 (2009): 877–885, https://doi.org/10.1021/ci800387z.

21. A. Fatima, A. M. Geethakumari, W. S. Ahmed, and K. H. Biswas, "A Potential Allosteric Inhibitor of SARS-CoV-2 Main Protease (Mpro) Identified Through Metastable State Analysis," *Frontiers in Molecular Biosciences* 11 (2024): 1451280, https://doi.org/10.3389/fmolb.2024.1451280.

22. J. S. Schneekloth, F. N. Fonseca, M. Koldobskiy, et al., "Chemical Genetic Control of Protein Levels: Selective In Vivo Targeted Degradation," *Journal of the American Chemical Society* 126 (2004): 3748–3754, https://doi.org/10.1021/ja039025z.

23. X. Sun, H. Gao, Y. Yang, et al., "Protacs: Great Opportunities for Academia and Industry," *Signal Transduction and Targeted Therapy* 4 (2019): 64, https://doi.org/10.1038/s41392-019-0101-6.

24. J. Adams, "The Proteasome: Structure, Function, and Role in the Cell," *Cancer Treatment Reviews* 29 (2003): 3–9, https://doi.org/10.1016/S0305-7372(03)00081-1.

25. M. Konstantinidou, J. Li, B. Zhang, et al., "PROTACs– a Game-Changing Technology," *Expert Opinion on Drug Discovery* 14 (2019): 1255–1268, https://doi.org/10.1080/17460441.2019.1659242.

26. Y. R. Alugubelli, Z. Z. Geng, K. S. Yang, et al., "A Systematic Exploration of Boceprevir-Based Main Protease Inhibitors as SARS-CoV-2 Antivirals," *European Journal of Medicinal Chemistry* 240 (2022): 114596, https://doi.org/10.1016/j.ejmech.2022.114596.

27. Y. R. Alugubelli, J. Xiao, K. Khatua, et al., "Discovery of First-In-Class Protac Degraders of SARS-CoV-2 Main Protease," *Journal of Medicinal Chemistry* 67 (2024): 6495–6507, https://doi.org/10.1021/acs.jmedchem.3c02416.

28. K. S. Yang, X. R. Ma, Y. Ma, et al., "A Quick Route to Multiple Highly Potent SARS-CoV-2 Main Protease Inhibitors**," *ChemMedChem* 16 (2021): 942–948, https://doi.org/10.1002/cmdc.202000924.

29. J. Desantis, A. Bazzacco, M. Eleuteri, et al., "Design, Synthesis, and Biological Evaluation of First-in-Class Indomethacin-Based PROTACs Degrading SARS-CoV-2 Main Protease and With Broad-Spectrum Antiviral Activity," *European Journal of Medicinal Chemistry* 268 (2024): 116202, https://doi.org/10.1016/j.ejmech.2024.116202.

30. M. Benchekroun, "The Advent of Directed Protein Degraders in Drug Discovery," *Future Drug Discovery* 1 (2019): 2, https://doi.org/10.4155/fdd-2019-0019.

31. J. M. Kelm, D. S. Pandey, E. Malin, et al., "PROTAC'ting Oncoproteins: Targeted Protein Degradation for Cancer Therapy," *Molecular Cancer* 22 (2023): 62, https://doi.org/10.1186/s12943-022-01707-5.

32. W. M. Czarnecki, S. Podlewska, and A. J. Bojarski, "Robust Optimization of SVM Hyperparameters in the Classification of Bioactive Compounds," *Journal of Cheminformatics* 7 (2015): 38, https://doi.org/10.1186/s13321-015-0088-0.

33. M. L. Boby, D. Fearon, M. Ferla, et al., "Open Science Discovery of Potent Noncovalent SARS-CoV-2 Main Protease Inhibitors," *Science* 382 (2023): eabo7201, https://doi.org/10.1126/science.abo7201.

34. E. M. A. Fassi, M. Garofalo, J. Sgrignani, et al., "Focused Design of Novel Cyclic Peptides Endowed With GABARAP-Inhibiting Activity," *International Journal of Molecular Sciences* 23 (2022): 5070, https://doi.org/10.3390/ijms23095070.

35. E. M. A. Fassi, M. Manenti, A. Citarella, et al., "Computational Design, Synthesis, and Biophysical Evaluation of β-Amido Boronic Acids as SARS-CoV-2 Mpro Inhibitors," *Molecules* 28 (2023): 2356, https://doi.org/10.3390/molecules28052356.

36. Y. Mao, L. Yu, R. Yang, L. Qu, and P. B. Harrington, "A Novel Method for the Study of Molecular Interaction by Using Microscale Thermophoresis," *Talanta* 132 (2015): 894–901, https://doi.org/10.1016/j.talanta.2014.09.038.

37. S. Günther, P. Y. A. Reinke, Y. Fernández-García, et al., "X-Ray Screening Identifies Active Site and Allosteric Inhibitors of SARS-CoV-2 Main Protease," *Science* 372 (2021): 642–646, https://doi.org/10.1126/science.abf7945.

38. P. G. Wyatt, A. J. Woodhead, V. Berdini, et al., "Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H- Pyrazole-3-Carboxamide (AT7519), a Novel Cyclin Dependent Kinase Inhibitor Using Fragment-Based X-Ray Crystallography and Structure Based Drug Design," *Journal of Medicinal Chemistry* 51 (2008): 4986–4999, https://doi.org/10.1021/jm800382h.

39. L. Alzyoud, M. A. Ghattas, and N. Atatreh, "Allosteric Binding Sites of the SARS-CoV-2 Main Protease: Potential Targets for Broad-Spectrum Anti-Coronavirus Agents," *Drug Design, Development and Therapy* 16 (2022): 2463–2478, https://doi.org/10.2147/DDDT.S370574.

40. C. Ma, M. D. Sacco, B. Hurst, et al., "Boceprevir, GC-376, and Calpain Inhibitors II, XII Inhibit SARS-CoV-2 Viral Replication by Targeting the Viral Main Protease," *Cell Research* 30 (2020): 678–692, https://doi.org/10.1038/s41422-020-0356-z.

41. S. K. Samrat, J. Xu, X. Xie, et al., "Allosteric Inhibitors of the Main Protease of SARS-COV-2," *Antiviral Research* 205 (2022): 105381, https://doi.org/10.1016/j.antiviral.2022.105381.

42. M. P. Repasky, M. Shelley, and R. A. Friesner, "Flexible Ligand Docking With Glide," *Current Protocols in Bioinformatics* 18 (2007): Unit 8.12, https://doi.org/10.1002/0471250953.bi0812s18.

43. J. Shi, J. Sivaraman, and J. Song, "Mechanism for Controlling the Dimer-Monomer Switch and Coupling Dimerization to Catalysis of the Severe Acute Respiratory Syndrome Coronavirus 3C-Like Protease," *Journal of Virology* 82 (2008): 4620–4629, https://doi.org/10.1128/jvi.02680-07.

44. T. M. Nguyen, V. Sreekanth, A. Deb, et al., "Proteolysis-Targeting Chimeras With Reduced Off-Targets," *Nature Chemistry* 16 (2024): 218–228, https://doi.org/10.1038/s41557-023-01379-8.

45. A. Pasteka, E. Diamanti, E. Ulfassi, and M. Laura Bolognesi, "Click Chemistry and Targeted Degradation: A Winning Combination for Medicinal Chemists?," *ChemMedChem* 18 (2023): e202300422, https://doi.org/10.1002/cmdc.202300422.

46. S. Riniker and G. A. Landrum, "Better Informed Distance Geometry: Using What We Know to Improve Conformation Generation," *Journal of Chemical Information and Modeling* 55 (2015): 2562–2574, https://doi.org/10.1021/acs.jcim.5b00654.

47. C. W. Yap, "Padel-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints," *Journal of Computational Chemistry* 32 (2011): 1466–1474, https://doi.org/10.1002/jcc.21707.

48. B. Gregorutti, B. Michel, and P. Saint-Pierre, "Correlation and Variable Importance in Random Forests," *Statistics and Computing* 27 (2017): 659–678, https://doi.org/10.1007/s11222-016-9646-1.

49. H. Ravishankar, R. Madhavan, R. Mullick, T. Shetty, L. Martinelli, and S. E. Joel, "Recursive Feature Elimination for Biomarker Discovery in Resting-State Functional Connectivity," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBS, 2016), https://doi.org/10.1109/EMBC.2016.7591621.

50. Q. Zhou, H. Zhou, Q. Zhou, F. Yang, and L. Luo, "Structure Damage Detection Based on Random Forest Recursive Feature Elimination," *Mechanical Systems and Signal Processing* 46 (2014): 82–90, https://doi.org/10.1016/j.ymssp.2013.12.013.

51. N. Altincekic, S. M. Korn, N. S. Qureshi, et al., "Large-Scale Recombinant Production of the SARS-CoV-2 Proteome for High-Throughput and Structural Biology Applications," *Frontiers in Molecular Biosciences* 8 (2021): 653148, https://doi.org/10.3389/fmolb.2021.653148.

52. G. Amendola, R. Ettari, S. Previti, et al., "Lead Discovery of SARS-CoV-2 Main Protease Inhibitors Through Covalent Docking-Based Virtual Screening," *Journal of Chemical Information and Modeling* 61 (2021): 2062–2073, https://doi.org/10.1021/acs.jcim.1c00184.

53. A. Luttens, H. Gullberg, E. Abdurakhmanov, et al., "Ultralarge Virtual Screening Identifies SARS-CoV-2 Main Protease Inhibitors With Broad-Spectrum Activity Against Coronaviruses," *Journal of the American Chemical Society* 144 (2022): 2905–2920, https://doi.org/10.1021/jacs.1c08402.

54. X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. Van Gunsteren, and A. E. Mark, "Peptide Folding: When Simulation Meets Experiment," *Angewandte Chemie International Edition* 38 (1999): 236–240, https://doi.org/10.1002/(sici)1521-3773(19990115)38:1/2<236::aid-anie236>3.0.co;2-m.

55. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, "GROMACS: Fast, Flexible, and Free," *Journal of Computational Chemistry* 26 (2005): 1701–1718, https://doi.org/10.1002/jcc.20291.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section.
SuppInfo_final. ArchPharm_SupplMat_InChI.

# 7. Bibliography


1) Clarke, L. A., & Amaral, M. D. (2023). What Can RNA-Based Therapy Do for Monogenic Diseases?. Pharmaceutics, 15(1), 260. https://doi.org/10.3390/pharmaceutics15010260

2) Zhu, Y., Zhu, L., Wang, X., & Jin, H. (2022). RNA-based therapeutics: an overview and prospectus. Cell death & disease, 13(7), 644. https://doi.org/10.1038/s41419-022-05075-2

3) Kim, J., Hu, C., Moufawad El Achkar, C., Black, L. E., Douville, J., Larson, A., Pendergast, M. K., Goldkind, S. F., Lee, E. A., Kuniholm, A., Soucy, A., Vaze, J., Belur, N. R., Fredriksen, K., Stojkovska, I., Tsytsykova, A., Armant, M., DiDonato, R. L., Choi, J., Cornelissen, L., … Yu, T. W. (2019). Patient-Customized Oligonucleotide Therapy for a Rare Genetic Disease. The New England journal of medicine, 381(17), 1644–1652. https://doi.org/10.1056/NEJMoa1813279

4) Naeem, S., Zhang, J., Zhang, Y., & Wang, Y. (2024). Nucleic acid therapeutics: Past, present, and future. Molecular therapy. Nucleic acids, 36(1), 102440. https://doi.org/10.1016/j.omtn.2024.102440

5) Dana, H., Chalbatani, G. M., Mahmoodzadeh, H., Karimloo, R., Rezaiean, O., Moradzadeh, A., Mehmandoost, N., Moazzen, F., Mazraeh, A., Marmari, V., Ebrahimi, M., Rashno, M. M., Abadi, S. J., & Gharagouzlo, E. (2017). Molecular Mechanisms and Biological Functions of siRNA. International journal of biomedical science : IJBS, 13(2), 48–57.

6) Geary, R. S., Henry, S. P., & Grillone, L. R. (2002). Fomivirsen: clinical pharmacology and potential drug interactions. Clinical pharmacokinetics, 41(4), 255–260. https://doi.org/10.2165/00003088-200241040-00002

7) Santos, R. D., Raal, F. J., Donovan, J. M., & Cromwell, W. C. (2015). Mipomersen preferentially reduces small low-density lipoprotein particle number in patients with hypercholesterolemia. Journal of clinical lipidology, 9(2), 201–209. https://doi.org/10.1016/j.jacl.2014.12.008

8) Mercuri, E., Darras, B. T., Chiriboga, C. A., Day, J. W., Campbell, C., Connolly, A. M., Iannaccone, S. T., Kirschner, J., Kuntz, N. L., Saito, K., Shieh, P. B., Tulinius, M., Mazzone, E. S., Montes, J., Bishop, K. M., Yang, Q., Foster, R., Gheuens, S., Bennett, C. F., Farwell, W., … CHERISH Study Group (2018). Nusinersen versus Sham Control in Later-Onset Spinal Muscular Atrophy. The New England journal of medicine, 378(7), 625–635. https://doi.org/10.1056/NEJMoa1710504.

9) Mendell, J. R., Rodino-Klapac, L. R., Sahenk, Z., Roush, K., Bird, L., Lowes, L. P., Alfano, L., Gomez, A. M., Lewis, S., Kota, J., Malik, V., Shontz, K., Walker, C. M., Flanigan, K. M., Corridore, M., Kean, J. R., Allen, H. D., Shilling, C., Melia, K. R., Sazani, P., … Eteplirsen Study Group (2013). Eteplirsen for the treatment of Duchenne muscular dystrophy. Annals of neurology, 74(5), 637–647. https://doi.org/10.1002/ana.23982.

10) Benson, M. D., Waddington-Cruz, M., Berk, J. L., Polydefkis, M., Dyck, P. J., Wang, A. K., Planté-Bordeneuve, V., Barroso, F. A., Merlini, G., Obici, L., Scheinberg, M., Brannagan, T. H., 3rd, Litchy,

W. J., Whelan, C., Drachman, B. M., Adams, D., Heitner, S. B., Conceição, I., Schmidt, H. H., Vita, G., … Coelho, T. (2018). Inotersen Treatment for Patients with Hereditary Transthyretin Amyloidosis. The New England journal of medicine, 379(1), 22–31. https://doi.org/10.1056/NEJMoa1716793.

11) Scaglioni, D., Catapano, F., Ellis, M., Torelli, S., Chambers, D., Feng, L., Beck, M., Sewry, C., Monforte, M., Harriman, S., Koenig, E., Malhotra, J., Popplewell, L., Guglieri, M., Straub, V., Mercuri, E., Servais, L., Phadke, R., Morgan, J., & Muntoni, F. (2021). The administration of antisense oligonucleotide golodirsen reduces pathological regeneration in patients with Duchenne muscular dystrophy. Acta neuropathologica communications, 9(1), 7. https://doi.org/10.1186/s40478-020-01106-1

12) Clemens, P. R., Rao, V. K., Connolly, A. M., Harper, A. D., Mah, J. K., McDonald, C. M., Smith, E. C., Zaidman, C. M., Nakagawa, T., CINRG DNHS Investigators, & Hoffman, E. P. (2022). Long-Term Functional Efficacy and Safety of Viltolarsen in Patients with Duchenne Muscular Dystrophy. Journal of neuromuscular diseases, 9(4), 493–501. https://doi.org/10.3233/JND-220811.

13) Shirley M. (2021). Casimersen: First Approval. Drugs, 81(7), 875–879. https://doi.org/10.1007/s40265-021-01512-2.

14) Coelho, T., Marques, W., Jr, Dasgupta, N. R., Chao, C. C., Parman, Y., França, M. C., Jr, Guo, Y. C., Wixner, J., Ro, L. S., Calandra, C. R., Kowacs, P. A., Berk, J. L., Obici, L., Barroso, F. A., Weiler, M., Conceição, I., Jung, S. W., Buchele, G., Brambatti, M., Chen, J., … NEURO-TTRansform Investigators (2023). Eplontersen for Hereditary Transthyretin Amyloidosis With Polyneuropathy. JAMA, 330(15), 1448–1458. https://doi.org/10.1001/jama.2023.18688

15) Kristen, A. V., Ajroud-Driss, S., Conceição, I., Gorevic, P., Kyriakides, T., & Obici, L. (2019). Patisiran, an RNAi therapeutic for the treatment of hereditary transthyretin-mediated amyloidosis. Neurodegenerative disease management, 9(1), 5–23. https://doi.org/10.2217/nmt-2018-0033

16) Balwani, M., Sardh, E., Ventura, P., Peiró, P. A., Rees, D. C., Stölzel, U., Bissell, D. M., Bonkovsky, H. L., Windyga, J., Anderson, K. E., Parker, C., Silver, S. M., Keel, S. B., Wang, J. D., Stein, P. E., Harper, P., Vassiliou, D., Wang, B., Phillips, J., Ivanova, A., … ENVISION Investigators (2020). Phase 3 Trial of RNAi Therapeutic Givosiran for Acute Intermittent Porphyria. The New England journal of medicine, 382(24), 2289–2301. https://doi.org/10.1056/NEJMoa1913147.

17) Liebow, A., Li, X., Racie, T., Hettinger, J., Bettencourt, B. R., Najafian, N., Haslett, P., Fitzgerald, K., Holmes, R. P., Erbe, D., Querbes, W., & Knight, J. (2017). An Investigational RNAi Therapeutic Targeting Glycolate Oxidase Reduces Oxalate Production in Models of Primary Hyperoxaluria. Journal of the American Society of Nephrology : JASN, 28(2), 494–503. https://doi.org/10.1681/ASN.2016030338.

18) Lamb Y. N. (2021). Inclisiran: First Approval. Drugs, 81(3), 389–395. https://doi.org/10.1007/s40265-021-01473-6.

19) Goldfarb, D. S., Lieske, J. C., Groothoff, J., Schalk, G., Russell, K., Yu, S., & Vrhnjak, B. (2023). Nedosiran in primary hyperoxaluria subtype 3: results from a phase I, single-dose study (PHYOX4). Urolithiasis, 51(1), 80. https://doi.org/10.1007/s00240-023-01453-3.

20) Gragoudas, E. S., Adamis, A. P., Cunningham, E. T., Jr, Feinsod, M., Guyer, D. R., & VEGF Inhibition Study in Ocular Neovascularization Clinical Trial Group (2004). Pegaptanib for neovascular age-related macular degeneration. The New England journal of medicine, 351(27), 2805–2816. https://doi.org/10.1056/NEJMoa042760.

21) Jaffe, G. J., Westby, K., Csaky, K. G., Monés, J., Pearlman, J. A., Patel, S. S., Joondeph, B. C., Randolph, J., Masonson, H., & Rezaei, K. A. (2021). C5 Inhibitor Avacincaptad Pegol for Geographic Atrophy Due to Age-Related Macular Degeneration: A Randomized Pivotal Phase 2/3 Trial. Ophthalmology, 128(4), 576–586. https://doi.org/10.1016/j.ophtha.2020.08.027.

22) Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Pérez Marc, G., Moreira, E. D., Zerbini, C., Bailey, R., Swanson, K. A., Roychoudhury, S., Koury, K., Li, P., Kalina, W. V., Cooper, D., Frenck, R. W., Jr, Hammitt, L. L., Türeci, Ö., … C4591001 Clinical Trial Group (2020). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. The New England journal of medicine, 383(27), 2603–2615. https://doi.org/10.1056/NEJMoa2034577.

23) Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., Diemert, D., Spector, S. A., Rouphael, N., Creech, C. B., McGettigan, J., Khetan, S., Segall, N., Solis, J., Brosz, A., Fierro, C., Schwartz, H., Neuzil, K., Corey, L., Gilbert, P., … COVE Study Group (2021). Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. The New England journal of medicine, 384(5), 403–416. https://doi.org/10.1056/NEJMoa2035389

24) Frangoul, H., Locatelli, F., Sharma, A., Bhatia, M., Mapara, M., Molinari, L., Wall, D., Liem, R. I., Telfer, P., Shah, A. J., Cavazzana, M., Corbacioglu, S., Rondelli, D., Meisel, R., Dedeken, L., Lobitz, S., de Montalembert, M., Steinberg, M. H., Walters, M. C., Eckrich, M. J., … CLIMB SCD-121 Study Group (2024). Exagamglogene Autotemcel for Severe Sickle Cell Disease. The New England journal of medicine, 390(18), 1649–1662. https://doi.org/10.1056/NEJMoa2309676.

25) Damase, T. R., Sukhovershin, R., Boada, C., Taraballi, F., Pettigrew, R. I., & Cooke, J. P. (2021). The Limitless Future of RNA Therapeutics. Frontiers in bioengineering and biotechnology, 9, 628137. https://doi.org/10.3389/fbioe.2021.628137

26) Jahanafrooz, Z., Baradaran, B., Mosafer, J., Hashemzaei, M., Rezaei, T., Mokhtarzadeh, A., & Hamblin, M. R. (2020). Comparison of DNA and mRNA vaccines against cancer. Drug discovery today, 25(3), 552–560. https://doi.org/10.1016/j.drudis.2019.12.003

27) Liu, A., & Wang, X. (2022). The Pivotal Role of Chemical Modifications in mRNA Therapeutics. Frontiers in cell and developmental biology, 10, 901510. https://doi.org/10.3389/fcell.2022.901510.

28) Duan, Q., Hu, T., Zhu, Q., Jin, X., Chi, F., & Chen, X. (2022). How far are the new wave of mRNA drugs from us? mRNA product current perspective and future development. Frontiers in immunology, 13, 974433. https://doi.org/10.3389/fimmu.2022.974433

29) Qin, S., Tang, X., Chen, Y., Chen, K., Fan, N., Xiao, W., Zheng, Q., Li, G., Teng, Y., Wu, M., & Song, X. (2022). mRNA-based therapeutics: powerful and versatile tools to combat diseases. Signal transduction and targeted therapy, 7(1), 166. https://doi.org/10.1038/s41392-022-01007-w.

30) Karim, M. E., Haque, S. T., Al-Busaidi, H., Bakhtiar, A., Tha, K. K., Holl, M. M. B., & Chowdhury, E. H. (2022). Scope and challenges of nanoparticle-based mRNA delivery in cancer treatment. Archives of pharmacal research, 45(12), 865–893. https://doi.org/10.1007/s12272-022-01418-x

31) Sahin, U., Karikó, K., & Türeci, Ö. (2014). mRNA-based therapeutics--developing a new class of drugs. Nature reviews. Drug discovery, 13(10), 759–780. https://doi.org/10.1038/nrd4278

32) Youssef, M., Hitti, C., Puppin Chaves Fulber, J., & Kamen, A. A. (2023). Enabling mRNA Therapeutics: Current Landscape and Challenges in Manufacturing. Biomolecules, 13(10), 1497. https://doi.org/10.3390/biom13101497

33) Kondratyeva, L. G., Dyachkova, M. S., & Galchenko, A. V. (2022). The Origin of Genetic Code and Translation in the Framework of Current Concepts on the Origin of Life. Biochemistry. Biokhimiia, 87(2), 150–169. https://doi.org/10.1134/S0006297922020079

34) Chaudhary, N., Weissman, D., & Whitehead, K. A. (2021). mRNA vaccines for infectious diseases: principles, delivery and clinical translation. Nature reviews. Drug discovery, 20(11), 817–838. https://doi.org/10.1038/s41573-021-00283-5

35) Szabó, G. T., Mahiny, A. J., & Vlatkovic, I. (2022). COVID-19 mRNA vaccines: Platforms and current developments. Molecular therapy : the journal of the American Society of Gene Therapy, 30(5), 1850–1868. https://doi.org/10.1016/j.ymthe.2022.02.016

36) Pardi, N., Hogan, M. J., Porter, F. W., & Weissman, D. (2018). mRNA vaccines - a new era in vaccinology. Nature reviews. Drug discovery, 17(4), 261–279. https://doi.org/10.1038/nrd.2017.243.

37) Sayour, E. J., Boczkowski, D., Mitchell, D. A., & Nair, S. K. (2024). Cancer mRNA vaccines: clinical advances and future opportunities. Nature reviews. Clinical oncology, 21(7), 489–500. https://doi.org/10.1038/s41571-024-00902-1

38) Lang, F., Schrörs, B., Löwer, M., Türeci, Ö., & Sahin, U. (2022). Identification of neoantigens for individualized therapeutic cancer vaccines. Nature reviews. Drug discovery, 21(4), 261–282. https://doi.org/10.1038/s41573-021-00387-y

39) Kranz, L. M., Diken, M., Haas, H., Kreiter, S., Loquai, C., Reuter, K. C., Meng, M., Fritz, D., Vascotto, F., Hefesha, H., Grunwitz, C., Vormehr, M., Hüsemann, Y., Selmi, A., Kuhn, A. N., Buck, J., Derhovanessian, E., Rae, R., Attig, S., Diekmann, J., … Sahin, U. (2016). Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. Nature, 534(7607), 396–401. https://doi.org/10.1038/nature18300.

40) Sahin, U., Derhovanessian, E., Miller, M., Kloke, B. P., Simon, P., Löwer, M., Bukur, V., Tadmor, A. D., Luxemburger, U., Schrörs, B., Omokoko, T., Vormehr, M., Albrecht, C., Paruzynski, A., Kuhn, A. N., Buck, J., Heesch, S., Schreeb, K. H., Müller, F., Ortseifer, I., … Türeci, Ö. (2017). Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. Nature, 547(7662), 222–226. https://doi.org/10.1038/nature23003

41) Palmer, C. D., Rappaport, A. R., Davis, M. J., Hart, M. G., Scallan, C. D., Hong, S. J., Gitlin, L., Kraemer, L. D., Kounlavouth, S., Yang, A., Smith, L., Schenk, D., Skoberne, M., Taquechel, K., Marrali, M., Jaroslavsky, J. R., Nganje, C. N., Maloney, E., Zhou, R., Navarro-Gomez, D., … Jooss, K. (2022). Individualized, heterologous chimpanzee adenovirus and self-amplifying mRNA neoantigen vaccine for advanced metastatic solid tumors: phase 1 trial interim results. Nature medicine, 28(8), 1619–1629. https://doi.org/10.1038/s41591-022-01937-6.

42) Rojas, L. A., Sethna, Z., Soares, K. C., Olcese, C., Pang, N., Patterson, E., Lihm, J., Ceglia, N., Guasp, P., Chu, A., Yu, R., Chandra, A. K., Waters, T., Ruan, J., Amisaki, M., Zebboudj, A., Odgerel, Z., Payne, G., Derhovanessian, E., Müller, F., … Balachandran, V. P. (2023). Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer. Nature, 618(7963), 144–150. https://doi.org/10.1038/s41586-023-06063-y.

43) Dolgin E. (2023). Personalized cancer vaccines pass first major clinical test. Nature reviews. Drug discovery, 22(8), 607–609. https://doi.org/10.1038/d41573-023-00118-5

44) Krienke, C., Kolb, L., Diken, E., Streuber, M., Kirchhoff, S., Bukur, T., Akilli-Öztürk, Ö., Kranz, L. M., Berger, H., Petschenka, J., Diken, M., Kreiter, S., Yogev, N., Waisman, A., Karikó, K., Türeci, Ö., & Sahin, U. (2021). A noninflammatory mRNA vaccine for treatment of experimental autoimmune encephalomyelitis. Science (New York, N.Y.), 371(6525), 145–153. https://doi.org/10.1126/science.aay3638

45) Karikó, K., Muramatsu, H., Welsh, F.A., Ludwig, J., Kato, H., Akira, S., and Weissman, D. (2008). Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. Mol. Ther. 16, 1833–1840.

46) Leppek, K., Byeon, G. W., Kladwang, W., Wayment-Steele, H. K., Kerr, C. H., Xu, A. F., Kim, D. S., Topkar, V. V., Choe, C., Rothschild, D., Tiu, G. C., Wellington-Oguri, R., Fujii, K., Sharma, E., Watkins, A. M., Nicol, J. J., Romano, J., Tunguz, B., Diaz, F., Cai, H., … Das, R. (2022). Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. Nature communications, 13(1), 1536. https://doi.org/10.1038/s41467-022-28776-w

47) Dubin, D. T., & Stollar, V. (1975). Methylation of Sindbis virus "26S" messenger RNA. Biochemical and biophysical research communications, 66(4), 1373–1379. https://doi.org/10.1016/0006-291x(75)90511-2.

48) Zhang, Y., Zhang, L. S., Dai, Q., Chen, P., Lu, M., Kairis, E. L., Murugaiah, V., Xu, J., Shukla, R. K., Liang, X., Zou, Z., Cormet-Boyaka, E., Qiu, J., Peeples, M. E., Sharma, A., He, C., & Li, J. (2022). 5-methylcytosine ($m^5C$) RNA modification controls the innate immune response to virus infection by regulating type I interferons. Proceedings of the National Academy of Sciences of the United States of America, 119(42), e2123338119. https://doi.org/10.1073/pnas.2123338119.

49) Nance, K. D., & Meier, J. L. (2021). Modifications in an Emergency: The Role of N1-Methylpseudouridine in COVID-19 Vaccines. ACS central science, 7(5), 748–756. https://doi.org/10.1021/acscentsci.1c00197

50) Boo, S. H., & Kim, Y. K. (2020). The emerging role of RNA modifications in the regulation of mRNA stability. Experimental & molecular medicine, 52(3), 400–408. https://doi.org/10.1038/s12276-020-0407-z.

51) Guo, G., Pan, K., Fang, S., Ye, L., Tong, X., Wang, Z., Xue, X., & Zhang, H. (2021). Advances in mRNA 5-methylcytosine modifications: Detection, effectors, biological functions, and clinical relevance. Molecular therapy. Nucleic acids, 26, 575–593. https://doi.org/10.1016/j.omtn.2021.08.020.

52) Weng, Y., Li, C., Yang, T., Hu, B., Zhang, M., Guo, S., Xiao, H., Liang, X. J., & Huang, Y. (2020). The challenge and prospect of mRNA therapeutics landscape. Biotechnology advances, 40, 107534. https://doi.org/10.1016/j.biotechadv.2020.107534

53) Ripoll, M., Bernard, M. C., Vaure, C., Bazin, E., Commandeur, S., Perkov, V., Lemdani, K., Nicolaï, M. C., Bonifassi, P., Kichler, A., Frisch, B., & Haensler, J. (2022). An imidazole modified lipid confers enhanced mRNA-LNP stability and strong immunization properties in mice and non-human primates. Biomaterials, 286, 121570. https://doi.org/10.1016/j.biomaterials.2022.121570

54) Travieso, T., Li, J., Mahesh, S., Mello, J. D. F. R. E., & Blasi, M. (2022). The use of viral vectors in vaccine development. NPJ vaccines, 7(1), 75. https://doi.org/10.1038/s41541-022-00503-y.

55) Liu, C. S., Park, C., Ngo, T., Saikumar, J., Palmer, C. R., Shahnaee, A., Romanow, W. J., & Chun, J. (2024). RNA Isoform Diversity in Human Neurodegenerative Diseases. eNeuro, 11(12), ENEURO.0296-24.2024. https://doi.org/10.1523/ENEURO.0296-24.2024

56) Magaña Rodriguez, J. R., Guerra-Rebollo, M., Borrós, S., & Fornaguera, C. (2024). Nucleic acid-loaded poly(beta-aminoester) nanoparticles for cancer nano-immuno therapeutics: the good, the bad, and the future. Drug delivery and translational research, 14(12), 3477–3493. https://doi.org/10.1007/s13346-024-01585-y.

57) Ferreira, D., & Rodrigues, L. R. (2025). RNA-based therapeutics: Cutting-edge advances in clinical research. Advanced Therapeutics, 8(8), e00082. https://doi.org/10.1002/adtp.202500082.

58) Acar, R., Paydaş, S., Yıldırım, M., Kılıçarslan, E., Sahın, U., Dogan, A., Guven, D. C., Ekıncı, O., Tıglıoglu, M., Erdogan, I., Elıbol, T., Kızıloz, H., Aykan, M. B., Sayın, S., Kaptan, K., Soydan, E., Gokmen, A., Esen, R., Barısta, I., Albayrak, M., … Karadurmus, N. (2023). Treatment options in primary mediastinal B cell lymphoma patients, retrospective multicentric analysis; a Turkısh oncology group study. Journal of cancer research and therapeutics, 19(Supplement), S138–S144. https://doi.org/10.4103/jcrt.jcrt_355_22

59) Muñoz, F. M., Sher, L. D., Sabharwal, C., Gurtman, A., Xu, X., Kitchin, N., Lockhart, S., Riesenberg, R., Sexter, J. M., Czajka, H., Paulsen, G. C., Maldonado, Y., Walter, E. B., Talaat, K. R., Englund, J. A., Sarwar, U. N., Hansen, C., Iwamoto, M., Webber, C., Cunliffe, L., … C4591007 Clinical Trial Group (2023). Evaluation of BNT162b2 Covid-19 Vaccine in Children Younger than 5 Years of Age. The New England journal of medicine, 388(7), 621–634. https://doi.org/10.1056/NEJMoa2211031

60) Simões, E. A. F., Klein, N. P., Sabharwal, C., Gurtman, A., Kitchin, N., Ukkonen, B., Korbal, P., Zou, J., Xie, X., Sarwar, U. N., Xu, X., Lockhart, S., Cunliffe, L., Lu, C., Ma, H., Swanson, K. A., Koury, K., Shi, P. Y., Cooper, D., Türeci, Ö., … Gruber, W. C. (2023). Immunogenicity and Safety of a Third COVID-19 BNT162b2 mRNA Vaccine Dose in 5- to 11-Year Olds. Journal of the Pediatric Infectious Diseases Society, 12(4), 234–238. https://doi.org/10.1093/jpids/piad015

61) Verbeke, R., Lentacker, I., De Smedt, S. C., & Dewitte, H. (2021). The dawn of mRNA vaccines: The COVID-19 case. Journal of controlled release : official journal of the Controlled Release Society, 333, 511–520. https://doi.org/10.1016/j.jconrel.2021.03.043

62) Grippin, A. J., Marconi, C., Copling, S., Li, N., Braun, C., Woody, C., Young, E., Gupta, P., Wang, M., Wu, A., Jeong, S. D., Soni, D., Weidert, F., Xie, C., Goldenberg, E., Kim, A., Zhao, C., DeVries, A., Castillo, P., Lohray, R., … Lin, S. H. (2025). SARS-CoV-2 mRNA vaccines sensitize tumours to immune checkpoint blockade. Nature, 647(8089), 488–497. https://doi.org/10.1038/s41586-025-09655-y

63) NCBI. (2017). RNA interference (RNAi). National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/probe/docs/techrnai/

64) Agrawal, N., Dasaradhi, P. V., Mohmmed, A., Malhotra, P., Bhatnagar, R. K., & Mukherjee, S. K. (2003). RNA interference: biology, mechanism, and applications. Microbiology and molecular biology reviews : MMBR, 67(4), 657–685. https://doi.org/10.1128/MMBR.67.4.657-685.2003

65) Collotta, D., Bertocchi, I., Chiapello, E., & Collino, M. (2023). Antisense oligonucleotides: a novel Frontier in pharmacological strategy. Frontiers in pharmacology, 14, 1304342. https://doi.org/10.3389/fphar.2023.1304342.

66) Hu, B., Zhong, L., Weng, Y., Peng, L., Huang, Y., Zhao, Y., & Liang, X. J. (2020). Therapeutic siRNA: state of the art. Signal transduction and targeted therapy, 5(1), 101. https://doi.org/10.1038/s41392-020-0207-x

67) Roberts, T. C., Langer, R., & Wood, M. J. A. (2020). Advances in oligonucleotide drug delivery. Nature reviews. Drug discovery, 19(10), 673–694. https://doi.org/10.1038/s41573-020-0075-7

68) Wittrup, A., & Lieberman, J. (2015). Knocking down disease: a progress report on siRNA therapeutics. Nature reviews. Genetics, 16(9), 543–552. https://doi.org/10.1038/nrg3978

69) Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., & Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature, 411(6836), 494–498. https://doi.org/10.1038/35078107

70) Whelan, D. R., Hiscox, T. J., Rood, J. I., Bambery, K. R., McNaughton, D., & Wood, B. R. (2014). Detection of an en masse and reversible B- to A-DNA conformational transition in prokaryotes in response to desiccation. Journal of the Royal Society, Interface, 11(97), 20140454. https://doi.org/10.1098/rsif.2014.0454.

71) DiMaio, F., Yu, X., Rensen, E., Krupovic, M., Prangishvili, D., & Egelman, E. H. (2015). Virology. A virus that infects a hyperthermophile encapsidates A-form DNA. Science (New York, N.Y.), 348(6237), 914–917. https://doi.org/10.1126/science.aaa4181

72) Choi, J., & Majima, T. (2011). Conformational changes of non-B DNA. Chemical Society reviews, 40(12), 5893–5909. https://doi.org/10.1039/c1cs15153c.

73) Fakhr, E., Zare, F., & Teimoori-Toolabi, L. (2016). Precise and efficient siRNA design: a key point in competent gene silencing. Cancer gene therapy, 23(4), 73–82. https://doi.org/10.1038/cgt.2016.4.

74) Kim, D. H., Behlke, M. A., Rose, S. D., Chang, M. S., Choi, S., & Rossi, J. J. (2005). Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. Nature biotechnology, 23(2), 222–226. https://doi.org/10.1038/nbt1051

75) Bramsen, J. B., Laursen, M. B., Damgaard, C. K., Lena, S. W., Babu, B. R., Wengel, J., & Kjems, J. (2007). Improved silencing properties using small internally segmented interfering RNAs. Nucleic acids research, 35(17), 5886–5897. https://doi.org/10.1093/nar/gkm548

76) Alterman, J. F., Godinho, B. M. D. C., Hassler, M. R., Ferguson, C. M., Echeverria, D., Sapp, E., Haraszti, R. A., Coles, A. H., Conroy, F., Miller, R., Roux, L., Yan, P., Knox, E. G., Turanov, A. A., King, R. M., Gernoux, G., Mueller, C., Gray-Edwards, H. L., Moser, R. P., Bishop, N. C., … Khvorova, A. (2019). A divalent siRNA chemical scaffold for potent and sustained modulation of gene expression throughout the central nervous system. Nature biotechnology, 37(8), 884–894. https://doi.org/10.1038/s41587-019-0205-0

77) Lima, W. F., Prakash, T. P., Murray, H. M., Kinberger, G. A., Li, W., Chappell, A. E., Li, C. S., Murray, S. F., Gaus, H., Seth, P. P., Swayze, E. E., & Crooke, S. T. (2012). Single-stranded siRNAs activate RNAi in animals. Cell, 150(5), 883–894. https://doi.org/10.1016/j.cell.2012.08.014

78) Byrne, M., Tzekov, R., Wang, Y., Rodgers, A., Cardia, J., Ford, G., Holton, K., Pandarinathan, L., Lapierre, J., Stanney, W., Bulock, K., Shaw, S., Libertine, L., Fettes, K., Khvorova, A., Kaushal, S., & Pavco, P. (2013). Novel hydrophobically modified asymmetric RNAi compounds (sd-rxRNA) demonstrate robust efficacy in the eye. Journal of ocular pharmacology and therapeutics : the official journal of the Association for Ocular Pharmacology and Therapeutics, 29(10), 855–864. https://doi.org/10.1089/jop.2013.0148

79) Paunovska, K., Loughrey, D., & Dahlman, J. E. (2022). Drug delivery systems for RNA therapeutics. Nature reviews. Genetics, 23(5), 265–280. https://doi.org/10.1038/s41576-021-00439-4.

80) Song, E., Zhu, P., Lee, S. K., Chowdhury, D., Kussman, S., Dykxhoorn, D. M., Feng, Y., Palliser, D., Weiner, D. B., Shankar, P., Marasco, W. A., & Lieberman, J. (2005). Antibody mediated in vivo delivery of small interfering RNAs via cell-surface receptors. Nature biotechnology, 23(6), 709–717. https://doi.org/10.1038/nbt1101.

81) Ahn, I., Kang, C. S., & Han, J. (2023). Where should siRNAs go: applicable organs for siRNA drugs. Experimental & molecular medicine, 55(7), 1283–1292. https://doi.org/10.1038/s12276-023-00998-y.

82) Springer, A. D., & Dowdy, S. F. (2018). GalNAc-siRNA Conjugates: Leading the Way for Delivery of RNAi Therapeutics. Nucleic acid therapeutics, 28(3), 109–118. https://doi.org/10.1089/nat.2018.0736.

83) Zhang, X., Goel, V., & Robbie, G. J. (2020). Pharmacokinetics of Patisiran, the First Approved RNA Interference Therapy in Patients With Hereditary Transthyretin-Mediated Amyloidosis. Journal of clinical pharmacology, 60(5), 573–585. https://doi.org/10.1002/jcph.1553

84) Adams, D., Tournev, I. L., Taylor, M. S., Coelho, T., Planté-Bordeneuve, V., Berk, J. L., González-Duarte, A., Gillmore, J. D., Low, S.-C., Sekijima, Y., Obici, L., Chen, C., Badri, P., Arum, S. M., Vest, J., & Polydefkis, M. (2023). Efficacy and safety of vutrisiran for patients with hereditary transthyretin-mediated amyloidosis with polyneuropathy: A randomized clinical trial. Amyloid, 30(1), 18–26. https://doi.org/10.1080/13506129.2022.2091985

85) Garrelfs, S. F., Frishberg, Y., Hulton, S. A., Koren, M. J., O'Riordan, W. D., Cochat, P., Deschênes, G., Shasha-Lavsky, H., Saland, J. M., van't Hoff, W. G., Fuster, D. G., Magen, D., Moochhala, S. H., Schalk, G., Simkova, E., Groothoff, J. W., Sas, D. J., Meliambro, K. A., Lu, J., Sweetser, M. T., Garg, P. P., Vaishnaw, A. K., Gansner, J. M., McGregor, T. L., & Lieske, J. C. (2021). Lumasiran, an RNAi therapeutic for primary hyperoxaluria type 1. The New England Journal of Medicine, 384(13), 1216–1226. https://doi.org/10.1056/NEJMoa2021712.

86) Syed Y. Y. (2023). Nedosiran: First Approval. Drugs, 83(18), 1729–1733. https://doi.org/10.1007/s40265-023-01976-4.

87) Young, G., Srivastava, A., Kavakli, K., Ross, C., Sathar, J., You, C. W., Tran, H., Sun, J., Wu, R., Poloskey, S., Qiu, Z., Kichou, S., Andersson, S., Mei, B., & Rangarajan, S. (2023). Efficacy and safety of fitusiran prophylaxis in people with haemophilia A or haemophilia B with inhibitors (ATLAS-INH): a multicentre, open-label, randomised phase 3 trial. Lancet (London, England), 401(10386), 1427–1437. https://doi.org/10.1016/S0140-6736(23)00284-2

88) Scott L. J. (2020). Givosiran: First Approval. Drugs, 80(3), 335–339. https://doi.org/10.1007/s40265-020-01269-0.

89) Huang, Y., Zheng, S., Guo, Z., de Mollerat du Jeu, X., Liang, X. J., Yang, Z., Zhang, H. Y., Gao, S., & Liang, Z. (2022). Ionizable liposomal siRNA therapeutics enables potent and persistent treatment of Hepatitis B. Signal transduction and targeted therapy, 7(1), 38. https://doi.org/10.1038/s41392-021-00859-y

90) https://www.sanofi.com/en/media-room/press-releases/2024/2024-06-21-05-00-00-2902104

91) https://ir.arrowheadpharma.com/news-releases/news-release-details/arrowhead-pharmaceuticals-submits-new-drug-application-us-fda

92) Ferreira, D., Santos-Pereira, C., Costa, M., Afonso, J., Yang, S., Hensel, J., McAndrews, K. M., Longatto-Filho, A., Fernandes, R., Melo, J. B., Baltazar, F., Moreira, J. N., Kalluri, R., & Rodrigues, L. R. (2023). Exosomes modified with anti-MEK1 siRNA lead to an effective silencing of triple negative breast cancer cells. Biomaterials advances, 154, 213643. https://doi.org/10.1016/j.bioadv.2023.213643

93) Silva, R., Ferreira, D., & Rodrigues, L. R. (2022). Exosome-based delivery of RNAi leads to breast cancer inhibition. Journal of Drug Delivery Science and Technology, 78, 103931. https://doi.org/10.1016/j.jddst.2022.103931

94) O'Brien, J., Hayder, H., Zayed, Y., & Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. Frontiers in endocrinology, 9, 402. https://doi.org/10.3389/fendo.2018.00402.

95) Lee, J. H., Kang, S., Ahn, M., Jang, H., & Min, D. H. (2018). Development of Dual-Pore Coexisting Branched Silica Nanoparticles for Efficient Gene-Chemo Cancer Therapy. Small (Weinheim an der Bergstrasse, Germany), 14(7), 10.1002/smll.201702564. https://doi.org/10.1002/smll.201702564

96) Lam, J. K., Chow, M. Y., Zhang, Y., & Leung, S. W. (2015). siRNA Versus miRNA as Therapeutics for Gene Silencing. Molecular therapy. Nucleic acids, 4(9), e252. https://doi.org/10.1038/mtna.2015.23

97) McAndrews, K. M., Xiao, F., Chronopoulos, A., LeBleu, V. S., Kugeratski, F. G., & Kalluri, R. (2021). Exosome-mediated delivery of CRISPR/Cas9 for targeting of oncogenic $Kras^{G12D}$ in pancreatic cancer. Life science alliance, 4(9), e202000875. https://doi.org/10.26508/lsa.202000875

98) Alvarez-Erviti, L., Seow, Y., Yin, H., Betts, C., Lakhal, S., & Wood, M. J. (2011). Delivery of siRNA to the mouse brain by systemic injection of targeted exosomes. Nature biotechnology, 29(4), 341–345. https://doi.org/10.1038/nbt.1807

99) Hermeking H. (2010). The miR-34 family in cancer and apoptosis. Cell death and differentiation, 17(2), 193–199. https://doi.org/10.1038/cdd.2009.56.

100) Zhang, J. G., Wang, J. J., Zhao, F., Liu, Q., Jiang, K., & Yang, G. H. (2010). MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC). Clinica chimica acta; international journal of clinical chemistry, 411(11-12), 846–852. https://doi.org/10.1016/j.cca.2010.02.074

101) Morishita, A., Oura, K., Tadokoro, T., Fujita, K., Tani, J., & Masaki, T. (2021). MicroRNAs in the Pathogenesis of Hepatocellular Carcinoma: A Review. Cancers, 13(3), 514. https://doi.org/10.3390/cancers13030514

102) Sparmann, A., & Vogel, J. (2023). RNA-based medicine: from molecular mechanisms to therapy. The EMBO journal, 42(21), e114760. https://doi.org/10.15252/embj.2023114760.

103) Meka, A. K., Niu, Y., Karmakar, S., Hartono, S. B., Zhang, J., Lin, C. X. C., Zhang, H., Whittaker, A., Jack, K., Yu, M., & Yu, C. (2016). Facile synthesis of large-pore bicontinuous cubic mesoporous silica nanoparticles for intracellular gene delivery. ChemNanoMat, 2(3), 220–225. https://doi.org/10.1002/cnma.201600021

104) Zhang, B. C., Luo, B. Y., Zou, J. J., Wu, P. Y., Jiang, J. L., Le, J. Q., Zhao, R. R., Chen, L., & Shao, J. W. (2020). Co-delivery of Sorafenib and CRISPR/Cas9 Based on Targeted Core-Shell Hollow Mesoporous Organosilica Nanoparticles for Synergistic HCC Therapy. ACS applied materials & interfaces, 12(51), 57362–57372. https://doi.org/10.1021/acsami.0c17660

105) Bancos, S., Stevens, D. L., & Tyner, K. M. (2014). Effect of silica and gold nanoparticles on macrophage proliferation, activation markers, cytokine production, and phagocytosis in vitro. International journal of nanomedicine, 10, 183–206. https://doi.org/10.2147/IJN.S72580.

106) Hong, D. S., Kang, Y. K., Borad, M., Sachdev, J., Ejadi, S., Lim, H. Y., Brenner, A. J., Park, K., Lee, J. L., Kim, T. Y., Shin, S., Becerra, C. R., Falchook, G., Stoudemire, J., Martin, D., Kelnar, K., Peltier, H., Bonato, V., Bader, A. G., Smith, S., … Beg, M. S. (2020). Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. British journal of cancer, 122(11), 1630–1637. https://doi.org/10.1038/s41416-020-0802-1.

107) Huang Y. (2017). Preclinical and Clinical Advances of GalNAc-Decorated Nucleic Acid Therapeutics. Molecular therapy. Nucleic acids, 6, 116–132. https://doi.org/10.1016/j.omtn.2016.12.003

108) Kaikkonen, M. U., Lam, M. T., & Glass, C. K. (2011). Non-coding RNAs as regulators of gene expression and epigenetics. Cardiovascular research, 90(3), 430–440. https://doi.org/10.1093/cvr/cvr097

109) Statello, L., Guo, C. J., Chen, L. L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. Nature reviews. Molecular cell biology, 22(2), 96–118. https://doi.org/10.1038/s41580-020-00315-9

110) Patil, V. S., Zhou, R., & Rana, T. M. (2014). Gene regulation by non-coding RNAs. Critical reviews in biochemistry and molecular biology, 49(1), 16–32. https://doi.org/10.3109/10409238.2013.844092

111) Gao, N., Li, Y., Li, J., Gao, Z., Yang, Z., Li, Y., Liu, H., & Fan, T. (2020). Long Non-Coding RNAs: The Regulatory Mechanisms, Research Strategies, and Future Directions in Cancers. Frontiers in oncology, 10, 598817. https://doi.org/10.3389/fonc.2020.598817

112) Bu, T., Li, L., & Tian, J. (2023). Unlocking the role of non-coding RNAs in prostate cancer progression: exploring the interplay with the Wnt signaling pathway. Frontiers in pharmacology, 14, 1269233. https://doi.org/10.3389/fphar.2023.1269233

113) Chen, B., Dragomir, M. P., Yang, C., Li, Q., Horst, D., & Calin, G. A. (2022). Targeting non-coding RNAs to overcome cancer therapy resistance. Signal transduction and targeted therapy, 7(1), 121. https://doi.org/10.1038/s41392-022-00975-3

114) Razin, S. V., & Gavrilov, A. A. (2021). Non-coding RNAs in chromatin folding and nuclear organization. Cellular and molecular life sciences : CMLS, 78(14), 5489–5504. https://doi.org/10.1007/s00018-021-03876-w

115) Frías-Lasserre, D., & Villagra, C. A. (2017). The Importance of ncRNAs as Epigenetic Mechanisms in Phenotypic Variation and Organic Evolution. Frontiers in microbiology, 8, 2483. https://doi.org/10.3389/fmicb.2017.02483

116) Zong, Y., Wang, X., Cui, B., Xiong, X., Wu, A., Lin, C., & Zhang, Y. (2023). Decoding the regulatory roles of non-coding RNAs in cellular metabolism and disease. Molecular therapy : the journal of the American Society of Gene Therapy, 31(6), 1562–1576. https://doi.org/10.1016/j.ymthe.2023.04.012

117) Ramón Y Cajal, S., Segura, M. F., & Hümmer, S. (2019). Interplay Between ncRNAs and Cellular Communication: A Proposal for Understanding Cell-Specific Signaling Pathways. Frontiers in genetics, 10, 281. https://doi.org/10.3389/fgene.2019.00281

118) Di, X., Gao, X., Peng, L., Ai, J., Jin, X., Qi, S., Li, H., Wang, K., & Luo, D. (2023). Cellular mechanotransduction in health and diseases: from molecular mechanism to therapeutic targets. Signal transduction and targeted therapy, 8(1), 282. https://doi.org/10.1038/s41392-023-01501-9

119) Ghafouri-Fard, S., Aghabalazade, A., Shoorei, H., Majidpoor, J., Taheri, M., & Mokhtari, M. (2021). The Impact of lncRNAs and miRNAs on Apoptosis in Lung Cancer. Frontiers in oncology, 11, 714795. https://doi.org/10.3389/fonc.2021.714795

120) Viereck, J., Bührke, A., Foinquinos, A., Chatterjee, S., Kleeberger, J. A., Xiao, K., Janssen-Peters, H., Batkai, S., Ramanujam, D., Kraft, T., Cebotari, S., Gueler, F., Beyer, A. M., Schmitz, J., Bräsen, J. H., Schmitto, J. D., Gyöngyösi, M., Löser, A., Hirt, M. N., Eschenhagen, T., … Thum, T. (2020). Targeting muscle-enriched long non-coding RNA H19 reverses pathological cardiac hypertrophy. European heart journal, 41(36), 3462–3474. https://doi.org/10.1093/eurheartj/ehaa519

121) Kawaguchi, S., Moukette, B., Hayasaka, T., Haskell, A. K., Mah, J., Sepúlveda, M. N., Tang, Y., & Kim, I. M. (2023). Noncoding RNAs as Key Regulators for Cardiac Development and Cardiovascular Diseases. Journal of cardiovascular development and disease, 10(4), 166. https://doi.org/10.3390/jcdd10040166

122) Arun, G., Diermeier, S. D., & Spector, D. L. (2018). Therapeutic Targeting of Long Non-Coding RNAs in Cancer. Trends in molecular medicine, 24(3), 257–277. https://doi.org/10.1016/j.molmed.2018.01.001

123) Dias, N., Dheur, S., Nielsen, P. E., Gryaznov, S., Van Aerschot, A., Herdewijn, P., Hélène, C., & Saison-Behmoaras, T. E. (1999). Antisense PNA tridecamers targeted to the coding region of Ha-ras mRNA arrest polypeptide chain elongation. Journal of molecular biology, 294(2), 403–416. https://doi.org/10.1006/jmbi.1999.3277

124) Hagedorn, P. H., Persson, R., Funder, E. D., Albæk, N., Diemer, S. L., Hansen, D. J., Møller, M. R., Papargyri, N., Christiansen, H., Hansen, B. R., Hansen, H. F., Jensen, M. A., & Koch, T. (2018). Locked nucleic acid: modality, diversity, and drug discovery. Drug discovery today, 23(1), 101–114. https://doi.org/10.1016/j.drudis.2017.09.018

125) Crooke, S. T., Wang, S., Vickers, T. A., Shen, W., & Liang, X. H. (2017). Cellular uptake and trafficking of antisense oligonucleotides. Nature biotechnology, 35(3), 230–237. https://doi.org/10.1038/nbt.3779

126) Cerritelli, S. M., & Crouch, R. J. (2009). Ribonuclease H: the enzymes in eukaryotes. The FEBS journal, 276(6), 1494–1505. https://doi.org/10.1111/j.1742-4658.2009.06908.x

127) Liang, X. H., Sun, H., Nichols, J. G., & Crooke, S. T. (2017). RNase H1-Dependent Antisense Oligonucleotides Are Robustly Active in Directing RNA Cleavage in Both the Cytoplasm and the Nucleus. Molecular therapy : the journal of the American Society of Gene Therapy, 25(9), 2075–2092. https://doi.org/10.1016/j.ymthe.2017.06.002.

128) Lennox, K. A., & Behlke, M. A. (2016). Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. Nucleic acids research, 44(2), 863–877. https://doi.org/10.1093/nar/gkv1206

129) Rinaldi, C., & Wood, M. J. A. (2018). Antisense oligonucleotides: the next frontier for treatment of neurological disorders. Nature reviews. Neurology, 14(1), 9–21. https://doi.org/10.1038/nrneurol.2017.148.

130) Liang, X. H., Sun, H., Shen, W., Wang, S., Yao, J., Migawa, M. T., Bui, H. H., Damle, S. S., Riney, S., Graham, M. J., Crooke, R. M., & Crooke, S. T. (2017). Antisense oligonucleotides targeting translation inhibitory elements in 5' UTRs can selectively increase protein levels. Nucleic acids research, 45(16), 9528–9546. https://doi.org/10.1093/nar/gkx632

131) Ward, A. J., Norrbom, M., Chun, S., Bennett, C. F., & Rigo, F. (2014). Nonsense-mediated decay as a terminating mechanism for antisense oligonucleotides. Nucleic acids research, 42(9), 5871–5879. https://doi.org/10.1093/nar/gku184.

132) Baker, B. F., Lot, S. S., Kringel, J., Cheng-Flournoy, S., Villiet, P., Sasmor, H. M., Siwkowski, A. M., Chappell, L. L., & Morrow, J. R. (1999). Oligonucleotide-europium complex conjugate designed to cleave the 5' cap structure of the ICAM-1 transcript potentiates antisense activity in cells. Nucleic acids research, 27(6), 1547–1551. https://doi.org/10.1093/nar/27.6.1547.

133) Vickers, T. A., Wyatt, J. R., Burckin, T., Bennett, C. F., & Freier, S. M. (2001). Fully modified 2' MOE oligonucleotides redirect polyadenylation. Nucleic acids research, 29(6), 1293–1299. https://doi.org/10.1093/nar/29.6.1293.

134) Melton, D. A. (1985). Injected anti-sense RNAs specifically block messenger RNA translation in vivo. Proceedings of the National Academy of Sciences of the United States of America, 82(1), 144–148. https://doi.org/10.1073/pnas.82.1.144

135) Eckstein F. (2014). Phosphorothioates, essential components of therapeutic oligonucleotides. Nucleic acid therapeutics, 24(6), 374–387. https://doi.org/10.1089/nat.2014.0506.

136) Wang, G., Gunic, E., Girardet, J. L., & Stoisavljevic, V. (1999). Conformationally locked nucleosides. Synthesis and hybridization properties of oligodeoxynucleotides containing 2',4'-C-bridged 2'-deoxynucleosides. Bioorganic & medicinal chemistry letters, 9(8), 1147–1150. https://doi.org/10.1016/s0960-894x(99)00146-8.

137) Eckstein, F. (1966). Nucleoside phosphorothioates. Journal of the American Chemical Society, 88(18), 4292–4294. https://doi.org/10.1021/ja00970a054.

138) Martin, P. (1995). Ein neuer Zugang zu 2′-O-Alkylribonucleosiden und Eigenschaften deren Oligonucleotide. Helvetica Chimica Acta, 78, 486–504. https://doi.org/10.1002/hlca.19950780219

139) Nielsen, P. E., Egholm, M., Berg, R. H., & Buchardt, O. (1991). Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. Science (New York, N.Y.), 254(5037), 1497–1500. https://doi.org/10.1126/science.1962210.

140) Hudziak, R. M., Barofsky, E., Barofsky, D. F., Weller, D. L., Huang, S. B., & Weller, D. D. (1996). Resistance of morpholino phosphorodiamidate oligomers to enzymatic degradation. Antisense & nucleic acid drug development, 6(4), 267–272. https://doi.org/10.1089/oli.1.1996.6.267.

141) Obika, S., Nanbu, D., Hari, Y., Andoh, J.-I., Morio, K.-I., Doi, T., & Imanishi, T. (1998). Stability and structural features of the duplexes containing nucleoside analogues with a fixed N-type conformation, 2′-O,4′-C-methyleneribonucleosides. Tetrahedron Letters, 39(30), 5401–5404. https://doi.org/10.1016/S0040-4039(98)01084-3.

142) Tanaka, K., Okuda, T., Kasahara, Y., & Obika, S. (2020). Base-modified aptamers obtained by cell-internalization SELEX facilitate cellular uptake of an antisense oligonucleotide. Molecular therapy. Nucleic acids, 23, 440–449. https://doi.org/10.1016/j.omtn.2020.11.016

143) Yin, H., Moulton, H. M., Seow, Y., Boyd, C., Boutilier, J., Iverson, P., & Wood, M. J. (2008). Cell-penetrating peptide-conjugated antisense oligonucleotides restore systemic muscle and cardiac dystrophin expression and function. Human molecular genetics, 17(24), 3909–3918. https://doi.org/10.1093/hmg/ddn293.

144) Perry, C. M., & Balfour, J. A. B. (1999). Fomivirsen. Drugs, 57(3), 375–380. https://doi.org/10.2165/00003495-199957030-00010.

145) Raal, F. J., Santos, R. D., Blom, D. J., Marais, A. D., Charng, M. J., Cromwell, W. C., Lachmann, R. H., Gaudet, D., Tan, J. L., Chasan-Taber, S., Tribble, D. L., Flaim, J. D., & Crooke, S. T. (2010). Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: a randomised, double-blind, placebo-controlled trial. Lancet (London, England), 375(9719), 998–1006. https://doi.org/10.1016/S0140-6736(10)60284-X.

146) Paik, J., & Duggan, S. (2019). Volanesorsen: First Global Approval. Drugs, 79(12), 1349–1354. https://doi.org/10.1007/s40265-019-01168-z.

147) Mathew, V., & Wang, A. K. (2019). Inotersen: new promise for the treatment of hereditary transthyretin amyloidosis. Drug Design, Development and Therapy, 13, 1515–1525. https://doi.org/10.2147/DDDT.S162913

148) Miller, T. M., Cudkowicz, M. E., Genge, A., Shaw, P. J., Sobue, G., Bucelli, R. C., Chiò, A., Van Damme, P., Ludolph, A. C., Glass, J. D., Andrews, J. A., Babu, S., Benatar, M., McDermott, C. J., Cochrane, T., Chary, S., Chew, S., Zhu, H., Wu, F., Nestorov, I., Graham, D., Sun, P., McNeill, M., Fanning, L., Ferguson, T. A., & Fradette, S. (2022). Trial of antisense oligonucleotide tofersen for SOD1 amyotrophic lateral sclerosis. The New England Journal of Medicine, 387(12), 1099–1110. https://doi.org/10.1056/NEJMoa2204705.

149) Stroes, E. S. G., Alexander, V. J., Karwatowska-Prokopczuk, E., Hegele, R. A., Arca, M., Ballantyne, C. M., Soran, H., Prohaska, T. A., Xia, S., Ginsberg, H. N., Witztum, J. L., & Tsimikas, S. (2024). Olezarsen, acute pancreatitis, and familial chylomicronemia syndrome. The New England Journal of Medicine, 390(19), 1781–1792. https://doi.org/10.1056/NEJMoa2400201.

150) Lim, K. R., Maruyama, R., & Yokota, T. (2017). Eteplirsen in the treatment of Duchenne muscular dystrophy. Drug design, development and therapy, 11, 533–545. https://doi.org/10.2147/DDDT.S97635

151) Heo Y. A. (2020). Golodirsen: First Approval. Drugs, 80(3), 329–333. https://doi.org/10.1007/s40265-020-01267-2.

152) Wagner, K. R., Kuntz, N. L., Koenig, E., East, L., Upadhyay, S., Han, B., & Shieh, P. B. (2021). Safety, tolerability, and pharmacokinetics of casimersen in patients with Duchenne muscular dystrophy amenable to exon 45 skipping: A randomized, double-blind, placebo-controlled, dose-titration trial. Muscle & nerve, 64(3), 285–292. https://doi.org/10.1002/mus.27347

153) Neil, E. E., & Bisaccia, E. K. (2019). Nusinersen: A Novel Antisense Oligonucleotide for the Treatment of Spinal Muscular Atrophy. The journal of pediatric pharmacology and therapeutics: JPPT: the official journal of PPAG, 24(3), 194–203. https://doi.org/10.5863/1551-6776-24.3.194.

154) https://ir.ionis.com/news-releases/news-release-details/tryngolzatm-olezarsen-approved-us-first-ever-treatment-adults

155) https://ir.ionis.com/news-releases/news-release-details/ionis-announces-fda-acceptance-new-drug-application-donidalorsen

156) https://neurology.ionis.com/sites/default/files/2024-10/Fused%20in%20Sarcoma%20Amyotrophic%20Lateral%20Sclerosis%20Clinical%20Trial%20Handout.pdf

157) Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., Marzluff, W. F., & Sharpless, N. E. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA (New York, N.Y.), 19(2), 141–157. https://doi.org/10.1261/rna.035667.112.

158) Liang, D., & Wilusz, J. E. (2014). Short intronic repeat sequences facilitate circular RNA production. Genes & development, 28(20), 2233–2247. https://doi.org/10.1101/gad.251926.114.

159) Tang, T. H., Rozhdestvensky, T. S., d'Orval, B. C., Bortolin, M. L., Huber, H., Charpentier, B., Branlant, C., Bachellerie, J. P., Brosius, J., & Hüttenhofer, A. (2002). RNomics in Archaea reveals a further link between splicing of archaeal introns and rRNA processing. Nucleic acids research, 30(4), 921–930. https://doi.org/10.1093/nar/30.4.921.

160) Chen, L. L., & Yang, L. (2015). Regulation of circRNA biogenesis. RNA biology, 12(4), 381–388. https://doi.org/10.1080/15476286.2015.1020271.

161) Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., & Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. Nature, 495(7441), 384–388. https://doi.org/10.1038/nature11993.

162) Zang, J., Lu, D., & Xu, A. (2020). The interaction of circRNAs and RNA-binding proteins: An important part of circRNA maintenance and function. Journal of Neuroscience Research, 98(1), 87–97. https://doi.org/10.1002/jnr.24356.

163) Yao, D., Lin, S., Chen, S., & Wang, Z. (2022). circHIPK3 regulates cell proliferation and migration by sponging microRNA-124 and regulating serine/threonine kinase 3 expression in esophageal squamous cell carcinoma. Bioengineered, 13(4), 9767–9780. https://doi.org/10.1080/21655979.2022.2060776.

164) Wesselhoeft, R. A., Kowalski, P. S., Parker-Hale, F. C., Huang, Y., Bisaria, N., & Anderson, D. G. (2019). RNA Circularization Diminishes Immunogenicity and Can Extend Translation Duration In Vivo. Molecular cell, 74(3), 508–520.e4. https://doi.org/10.1016/j.molcel.2019.02.015.

165) Breuer, J., Barth, P., Noe, Y., Shalamova, L., Goesmann, A., Weber, F., & Rossbach, O. (2022). What goes around comes around: Artificial circular RNAs bypass cellular antiviral responses. Molecular therapy. Nucleic acids, 28, 623–635. https://doi.org/10.1016/j.omtn.2022.04.017.

166) He, A. T., Liu, J., Li, F., & Yang, B. B. (2021). Targeting circular RNAs as a therapeutic approach: current strategies and challenges. Signal transduction and targeted therapy, 6(1), 185. https://doi.org/10.1038/s41392-021-00569-5

167) Wang, Z., Ma, K., Cheng, Y., Abraham, J. M., Liu, X., Ke, X., Wang, Z., & Meltzer, S. J. (2019). Synthetic circular multi-miR sponge simultaneously inhibits miR-21 and miR-93 in esophageal carcinoma. Laboratory investigation; a journal of technical methods and pathology, 99(10), 1442–1453. https://doi.org/10.1038/s41374-019-0273-2.

168) Liu, X., Abraham, J. M., Cheng, Y., Wang, Z., Wang, Z., Zhang, G., Ashktorab, H., Smoot, D. T., Cole, R. N., Boronina, T. N., DeVine, L. R., Talbot, C. C., Jr, Liu, Z., & Meltzer, S. J. (2018). Synthetic Circular RNA Functions as a miR-21 Sponge to Suppress Gastric Carcinoma Cell Proliferation. Molecular therapy. Nucleic acids, 13, 312–321. https://doi.org/10.1016/j.omtn.2018.09.010.

169) Chen, R., Wang, S. K., Belk, J. A., Amaya, L., Li, Z., Cardenas, A., Abe, B. T., Chen, C. K., Wender, P. A., & Chang, H. Y. (2023). Engineering circular RNA for enhanced protein production. Nature biotechnology, 41(2), 262–272. https://doi.org/10.1038/s41587-022-01393-0.

170) Wu, D. P., Zhao, Y. D., Yan, Q. Q., Liu, L. L., Wei, Y. S., & Huang, J. L. (2023). Circular RNAs: emerging players in brain aging and neurodegenerative diseases. The Journal of Pathology, 259(1), 1–9. https://doi.org/10.1002/path.6021.

171) Wang, W., Sun, L., Huang, M. T., Quan, Y., Jiang, T., Miao, Z., & Zhang, Q. (2023). Regulatory circular RNAs in viral diseases: applications in diagnosis and therapy. RNA biology, 20(1), 847–858. https://doi.org/10.1080/15476286.2023.2272118

172) Schult, P., Roth, H., Adams, R. L., Mas, C., Imbert, L., Orlik, C., Ruggieri, A., Pyle, A. M., & Lohmann, V. (2018). microRNA-122 amplifies hepatitis C virus translation by shaping the structure of the internal ribosomal entry site. Nature communications, 9(1), 2613. https://doi.org/10.1038/s41467-018-05053-3.

173) Zhou, Z., Zhang, R., Li, X., Zhang, W., Zhan, Y., Lang, Z., Tao, Q., Yu, J., Yu, S., Yu, Z., & Zheng, J. (2024). Circular RNA cVIM promotes hepatic stellate cell activation in liver fibrosis via miR-122-5p/miR-9-5p-mediated TGF-β signaling cascade. Communications biology, 7(1), 113. https://doi.org/10.1038/s42003-024-05797-3

174) Amaya, L., Grigoryan, L., Li, Z., Lee, A., Wender, P. A., Pulendran, B., & Chang, H. Y. (2023). Circular RNA vaccine induces potent T cell responses. Proceedings of the National Academy of Sciences of the United States of America, 120(20), e2302191120. https://doi.org/10.1073/pnas.2302191120.

175) Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. Nature, 346(6287), 818–822. https://doi.org/10.1038/346818a0.

176) Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science (New York, N.Y.), 249(4968), 505–510. https://doi.org/10.1126/science.2200121

177) Dunn, M. R., Jimenez, R. M., & Chaput, J. C. (2017). Analysis of aptamer discovery and technology. Nature Reviews Chemistry, 1, 76. https://doi.org/10.1038/s41570-017-0076

178) Irvine, D., Tuerk, C., & Gold, L. (1991). SELEXION. Systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. Journal of molecular biology, 222(3), 739–761. https://doi.org/10.1016/0022-2836(91)90509-5.

179) Wilson, D. S., & Szostak, J. W. (1999). In vitro selection of functional nucleic acids. Annual review of biochemistry, 68, 611–647. https://doi.org/10.1146/annurev.biochem.68.1.611.

180) Yan, J., Xiong, H., Cai, S., Wen, N., He, Q., Liu, Y., Peng, D., & Liu, Z. (2019). Advances in aptamer screening technologies. Talanta, 200, 124–144. https://doi.org/10.1016/j.talanta.2019.03.015.

181) Narayan, C., , Kwon, J., , Kim, C., , Kim, S. J., , & Jang, S. K., (2020). Virus-based SELEX (viro-SELEX) allows development of aptamers targeting knotty proteins. The Analyst, 145(4), 1473–1482. https://doi.org/10.1039/c9an01943j.

182) Avci-Adali, M., Metzger, M., Perle, N., Ziemer, G., & Wendel, H. P. (2010). Pitfalls of cell-systematic evolution of ligands by exponential enrichment (SELEX): existing dead cells during in vitro selection anticipate the enrichment of specific aptamers. Oligonucleotides, 20(6), 317–323. https://doi.org/10.1089/oli.2010.0253.

183) Zhou, J., & Rossi, J. (2017). Aptamers as targeted therapeutics: current potential and challenges. Nature Reviews. Drug discovery, 16(6), 440. https://doi.org/10.1038/nrd.2017.86

184) Zhou, J., & Rossi, J. J. (2014). Cell-type-specific, aptamer-functionalized agents for targeted disease therapy. Molecular Therapy — Nucleic Acids, 3, e169. https://doi.org/10.1038/mtna.2014.21

185) Mi, J., Liu, Y., Rabbani, Z. N., Yang, Z., Urban, J. H., Sullenger, B. A., & Clary, B. M. (2010). In vivo selection of tumor-targeting RNA motifs. Nature Chemical Biology, 6(1), 22–24. https://doi.org/10.1038/nchembio.277

186) Yunn, N. O., Lee, J., Lee, H. S., Oh, E. J., Park, M., Park, S., Jin, S. Y., Shin, E., Lee, J. W. Y., Kim, Y., Bae, S. S., & Ryu, S. H. (2022). An aptamer agonist of the insulin receptor acts as a positive or negative allosteric modulator, depending on its concentration. Experimental & molecular medicine, 54(4), 531–541. https://doi.org/10.1038/s12276-022-00760-w.

187) Spöring, M., Boneberg, R., & Hartig, J. S. (2020). Aptamer-Mediated Control of Polyadenylation for Gene Expression Regulation in Mammalian Cells. ACS synthetic biology, 9(11), 3008–3018. https://doi.org/10.1021/acssynbio.0c00222

188) Spöring, M., Finke, M., & Hartig, J. S. (2020). Aptamers in RNA-based switches of gene expression. Current opinion in biotechnology, 63, 34–40. https://doi.org/10.1016/j.copbio.2019.11.008.

189) Vogel, M., Weigand, J. E., Kluge, B., Grez, M., & Suess, B. (2018). A small, portable RNA device for the control of exon skipping in mammalian cells. Nucleic acids research, 46(8), e48. https://doi.org/10.1093/nar/gky062

190) Suess, B., Hanson, S., Berens, C., Fink, B., Schroeder, R., & Hillen, W. (2003). Conditional gene expression by controlling translation with tetracycline-binding aptamers. Nucleic acids research, 31(7), 1853–1858. https://doi.org/10.1093/nar/gkg285.

191) Healy, J. M., Lewis, S. D., Kurz, M., Boomer, R. M., Thompson, K. M., Wilson, C., & McCauley, T. G. (2004). Pharmacokinetics and biodistribution of novel aptamer compositions. Pharmaceutical research, 21(12), 2234–2246. https://doi.org/10.1007/s11095-004-7676-4

192) Musumeci, D., & Montesarchio, D. (2012). Polyvalent nucleic acid aptamers and modulation of their activity: a focus on the thrombin-binding aptamer. Pharmacology & therapeutics, 136(2), 202–215. https://doi.org/10.1016/j.pharmthera.2012.07.011.

193) Guo P. (2010). The emerging field of RNA nanotechnology. Nature nanotechnology, 5(12), 833–842. https://doi.org/10.1038/nnano.2010.231

194) Dass, C. R., Saravolac, E. G., Li, Y., & Sun, L. Q. (2002). Cellular uptake, distribution, and stability of 10-23 deoxyribozymes. Antisense & nucleic acid drug development, 12(5), 289–299. https://doi.org/10.1089/108729002761381276.

195) Lee, C. H., Lee, S. H., Kim, J. H., Noh, Y. H., Noh, G. J., & Lee, S. W. (2015). Pharmacokinetics of a Cholesterol-conjugated Aptamer Against the Hepatitis C Virus (HCV) NS5B Protein. Molecular therapy. Nucleic acids, 4(10), e254. https://doi.org/10.1038/mtna.2015.30.

196) Burmeister, P. E., Lewis, S. D., Silva, R. F., Preiss, J. R., Horwitz, L. R., Pendergrast, P. S., McCauley, T. G., Kurz, J. C., Epstein, D. M., Wilson, C., & Keefe, A. D. (2005). Direct in vitro selection of a 2'-O-methyl aptamer to VEGF. Chemistry & biology, 12(1), 25–33. https://doi.org/10.1016/j.chembiol.2004.10.017.

197) Boomer, R. M., Lewis, S. D., Healy, J. M., Kurz, M., Wilson, C., & McCauley, T. G. (2005). Conjugation to polyethylene glycol polymer promotes aptamer biodistribution to healthy and inflamed tissues. Oligonucleotides, 15(3), 183–195. https://doi.org/10.1089/oli.2005.15.183.

198) Willis, M. C., Collins, B. D., Zhang, T., Green, L. S., Sebesta, D. P., Bell, C., Kellogg, E., Gill, S. C., Magallanez, A., Knauer, S., Bendele, R. A., Gill, P. S., & Janjić, N. (1998). Liposome-anchored vascular endothelial growth factor aptamers. Bioconjugate chemistry, 9(5), 573–582. https://doi.org/10.1021/bc980002x.

199) Zhou, J., Soontornworajit, B., Martin, J., Sullenger, B. A., Gilboa, E., & Wang, Y. (2009). A hybrid DNA aptamer-dendrimer nanomaterial for targeted cell labeling. Macromolecular bioscience, 9(9), 831–835. https://doi.org/10.1002/mabi.200900046.

200) Jo, H., & Ban, C. (2016). Aptamer-nanoparticle complexes as powerful diagnostic and therapeutic tools. Experimental & molecular medicine, 48(5), e230. https://doi.org/10.1038/emm.2016.44.

201) Bauer, M., Strom, M., Hammond, D. S., & Shigdar, S. (2019). Anything You Can Do, I Can Do Better: Can Aptamers Replace Antibodies in Clinical Diagnostic Applications? Molecules (Basel, Switzerland), 24(23), 4377. https://doi.org/10.3390/molecules24234377.

202) https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm

203) Zhu, G., & Chen, X. (2018). Aptamer-based targeted therapy. Advanced drug delivery reviews, 134, 65–78. https://doi.org/10.1016/j.addr.2018.08.005.

204) Di Mauro, V., Lauta, F. C., Modica, J., Appleton, S. L., De Franciscis, V., & Catalucci, D. (2023). Diagnostic and Therapeutic Aptamers: A Promising Pathway to Improved Cardiovascular Disease Management. JACC. Basic to translational science, 9(2), 260–277. https://doi.org/10.1016/j.jacbts.2023.06.013

205) Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell, 35(3 Pt 2), 849–857. https://doi.org/10.1016/0092-8674(83)90117-4.

206) Sullivan S. M. (1994). Development of ribozymes for gene therapy. The Journal of investigative dermatology, 103(5 Suppl), 85S–89S. https://doi.org/10.1038/jid.1994.15.

207) Forster, A. C., & Symons, R. H. (1987). Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. Cell, 49(2), 211–220. https://doi.org/10.1016/0092-8674(87)90562-9.

208) Hampel, A., Tritz, R., Hicks, M., & Cruz, P. (1990). 'Hairpin' catalytic RNA model: evidence for helices and sequence requirement for substrate RNA. Nucleic acids research, 18(2), 299–304. https://doi.org/10.1093/nar/18.2.299.

209) Kashani-Sabet M. (2002). Ribozyme therapeutics. The journal of investigative dermatology. Symposium proceedings, 7(1), 76–78. https://doi.org/10.1046/j.1523-1747.2002.19642.x.

210) Sarver, N., Cantin, E. M., Chang, P. S., Zaia, J. A., Ladne, P. A., Stephens, D. A., & Rossi, J. J. (1990). Ribozymes as potential anti-HIV-1 therapeutic agents. Science (New York, N.Y.), 247(4947), 1222–1225. https://doi.org/10.1126/science.2107573.

211) Weng, D. E., Masci, P. A., Radka, S. F., Jackson, T. E., Weiss, P. A., Ganapathi, R., Elson, P. J., Capra, W. B., Parker, V. P., Lockridge, J. A., Cowens, J. W., Usman, N., & Borden, E. C. (2005). A phase I clinical trial of a ribozyme-based angiogenesis inhibitor targeting vascular endothelial growth factor receptor-1 for patients with refractory solid tumors. Molecular cancer therapeutics, 4(6), 948–955. https://doi.org/10.1158/1535-7163.MCT-04-0210.

212) Morrow, P. K., Murthy, R. K., Ensor, J. D., Gordon, G. S., Margolin, K. A., Elias, A. D., Urba, W. J., Weng, D. E., Rugo, H. S., & Hortobagyi, G. N. (2012). An open-label, phase 2 trial of RPI.4610 (Angiozyme) in the treatment of metastatic breast cancer. Cancer, 118(17), 4098–4104. https://doi.org/10.1002/cncr.26730

213) Heidenreich, O., Benseler, F., Fahrenholz, A., & Eckstein, F. (1994). High activity and stability of hammerhead ribozymes containing 2'-modified pyrimidine nucleosides and phosphorothioates. The Journal of biological chemistry, 269(3), 2131–2138.

214) Czapik, T., Piasecka, J., Kierzek, R., & Kierzek, E. (2022). Structural variants and modifications of hammerhead ribozymes targeting influenza A virus conserved structural motifs. Molecular therapy. Nucleic acids, 29, 64–74. https://doi.org/10.1016/j.omtn.2022.05.035

215) Bulcha, J. T., Wang, Y., Ma, H., Tai, P. W. L., & Gao, G. (2021). Viral vector platforms within the gene therapy landscape. Signal transduction and targeted therapy, 6(1), 53. https://doi.org/10.1038/s41392-021-00487-6

216) Asmamaw Mengstie M. (2022). Viral Vectors for the in Vivo Delivery of CRISPR Components: Advances and Challenges. Frontiers in bioengineering and biotechnology, 10, 895713. https://doi.org/10.3389/fbioe.2022.895713

217) Xiao, W., Chirmule, N., Berta, S. C., McCullough, B., Gao, G., & Wilson, J. M. (1999). Gene therapy vectors based on adeno-associated virus type 1. Journal of Virology, 73(5), 3994–4003. https://doi.org/10.1128/JVI.73.5.3994-4003.1999

218) Wang, D., Tai, P. W. L., & Gao, G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. Nature Reviews. Drug discovery, 18(5), 358–378. https://doi.org/10.1038/s41573-019-0012-9

219) Suoranta, T., Laham-Karam, N., & Ylä-Herttuala, S. (2022). Strategies to improve safety profile of AAV vectors. Frontiers in molecular medicine, 2, 1054069. https://doi.org/10.3389/fmmed.2022.1054069.

220) Au, H. K. E., Isalan, M., & Mielcarek, M. (2022). Gene Therapy Advances: A Meta-Analysis of AAV Usage in Clinical Settings. Frontiers in medicine, 8, 809118. https://doi.org/10.3389/fmed.2021.809118.

221) Wang, J. H., Gessler, D. J., Zhan, W., Gallagher, T. L., & Gao, G. (2024). Adeno-associated virus as a delivery vector for gene therapy of human diseases. Signal transduction and targeted therapy, 9(1), 78. https://doi.org/10.1038/s41392-024-01780-w

222) Hauck, B., & Xiao, W. (2003). Characterization of tissue tropism determinants of adeno-associated virus type 1. Journal of virology, 77(4), 2768–2774. https://doi.org/10.1128/jvi.77.4.2768-2774.2003.

223) Ronzitti, G., Gross, D. A., & Mingozzi, F. (2020). Human Immune Responses to Adeno-Associated Virus (AAV) Vectors. Frontiers in immunology, 11, 670. https://doi.org/10.3389/fimmu.2020.00670.

224) Sant'Anna, T. B., & Araujo, N. M. (2022). Adeno-associated virus infection and its impact in human health: an overview. Virology journal, 19(1), 173. https://doi.org/10.1186/s12985-022-01900-4.

225) Kreppel, F., & Hagedorn, C. (2022). Episomes and Transposases-Utilities to Maintain Transgene Expression from Nonviral Vectors. Genes, 13(10), 1872. https://doi.org/10.3390/genes13101872

226) Greig, J. A., Martins, K. M., Breton, C., Lamontagne, R. J., Zhu, Y., He, Z., White, J., Zhu, J. X., Chichester, J. A., Zheng, Q., Zhang, Z., Bell, P., Wang, L., & Wilson, J. M. (2024). Integrated vector genomes may contribute to long-term expression in primate liver after AAV administration. Nature biotechnology, 42(8), 1232–1242. https://doi.org/10.1038/s41587-023-01974-7.

227) Chandler, R. J., Sands, M. S., & Venditti, C. P. (2017). Recombinant Adeno-Associated Viral Integration and Genotoxicity: Insights from Animal Models. Human gene therapy, 28(4), 314–322. https://doi.org/10.1089/hum.2017.009

228) Penaud-Budloo, M., Le Guiner, C., Nowrouzi, A., Toromanoff, A., Chérel, Y., Chenuaud, P., Schmidt, M., von Kalle, C., Rolling, F., Moullier, P., & Snyder, R. O. (2008). Adeno-associated virus vector genomes persist as episomal chromatin in primate muscle. Journal of virology, 82(16), 7875–7885. https://doi.org/10.1128/JVI.00649-08.

229) Connolly J. B. (2002). Lentiviruses in gene therapy clinical research. Gene therapy, 9(24), 1730–1734. https://doi.org/10.1038/sj.gt.3301893

230) White, M., Whittaker, R., Gándara, C., & Stoll, E. A. (2017). A Guide to Approaching Regulatory Considerations for Lentiviral-Mediated Gene Therapies. Human gene therapy methods, 28(4), 163–176. https://doi.org/10.1089/hgtb.2017.096.

231) Ciuffi A. (2008). Mechanisms governing lentivirus integration site selection. Current gene therapy, 8(6), 419–429. https://doi.org/10.2174/156652308786848021.

232) Dong, W., & Kantor, B. (2021). Lentiviral Vectors for Delivery of Gene-Editing Systems Based on CRISPR/Cas: Current State and Perspectives. Viruses, 13(7), 1288. https://doi.org/10.3390/v13071288.

233) Luis A. (2020). The Old and the New: Prospects for Non-Integrating Lentiviral Vector Technology. Viruses, 12(10), 1103. https://doi.org/10.3390/v12101103.

234) Jung, H. N., Lee, S. Y., Lee, S., Youn, H., & Im, H. J. (2022). Lipid nanoparticles for delivery of RNA therapeutics: Current status and the role of in vivo imaging. Theranostics, 12(17), 7509–7531. https://doi.org/10.7150/thno.77259

235) Eygeris, Y., Gupta, M., Kim, J., & Sahay, G. (2022). Chemistry of Lipid Nanoparticles for RNA Delivery. Accounts of chemical research, 55(1), 2–12. https://doi.org/10.1021/acs.accounts.1c00544.

236) Mukai, H., Ogawa, K., Kato, N., & Kawakami, S. (2022). Recent advances in lipid nanoparticles for delivery of nucleic acid, mRNA, and gene editing-based therapeutics. Drug metabolism and pharmacokinetics, 44, 100450. https://doi.org/10.1016/j.dmpk.2022.100450.

237) Swetha, K., Kotla, N. G., Tunki, L., Jayaraj, A., Bhargava, S. K., Hu, H., Bonam, S. R., & Kurapati, R. (2023). Recent Advances in the Lipid Nanoparticle-Mediated Delivery of mRNA Vaccines. Vaccines, 11(3), 658. https://doi.org/10.3390/vaccines11030658.

238) Schober, G. B., Story, S., & Arya, D. P. (2024). A careful look at lipid nanoparticle characterization: analysis of benchmark formulations for encapsulation of RNA cargo size gradient. Scientific reports, 14(1), 2403. https://doi.org/10.1038/s41598-024-52685-1.

239) Mendes, B. B., Conniot, J., Avital, A., Yao, D., Jiang, X., Zhou, X., Sharf-Pauker, N., Xiao, Y., Adir, O., Liang, H., Shi, J., Schroeder, A., & Conde, J. (2022). Nanodelivery of nucleic acids. Nature reviews. Methods primers, 2, 24. https://doi.org/10.1038/s43586-022-00104-y

240) Sristi, Almalki, W. H., Karwasra, R., Gupta, G., Singh, S., Sharma, A., Sahebkar, A., & Kesharwani, P. (2024). Advances in the polymeric nanoparticulate delivery systems for RNA therapeutics. Progress in molecular biology and translational science, 204, 219–248. https://doi.org/10.1016/bs.pmbts.2024.01.001.

241) Jiang, X., Abedi, K., & Shi, J. (2021). Polymeric nanoparticles for RNA delivery. In Reference Module in Materials Science and Materials Engineering (Vol. 3, pp. 555–573). Elsevier. https://doi.org/10.1016/B978-0-12-822425-0.00017-8

242) Khaliq, N. U., Lee, J., Kim, J., Kim, Y., Yu, S., Kim, J., Kim, S., Sung, D., & Kim, H. (2023). Mesoporous Silica Nanoparticles as a Gene Delivery Platform for Cancer Therapy. Pharmaceutics, 15(5), 1432. https://doi.org/10.3390/pharmaceutics15051432.

243) Kazemzadeh, P., Sayadi, K., Toolabi, A., Sayadi, J., Zeraati, M., Chauhan, N. P. S., & Sargazi, G. (2022). Structure-Property Relationship for Different Mesoporous Silica Nanoparticles and its Drug Delivery Applications: A Review. Frontiers in chemistry, 10, 823785. https://doi.org/10.3389/fchem.2022.823785.

244) Elizarova, T. N., Antopolsky, M. L., Novichikhin, D. O., Skirda, A. M., Orlov, A. V., Bragina, V. A., & Nikitin, P. I. (2023). A Straightforward Method for the Development of Positively Charged Gold Nanoparticle-Based Vectors for Effective siRNA Delivery. Molecules (Basel, Switzerland), 28(8), 3318. https://doi.org/10.3390/molecules28083318.

245) Cai, Y., Liu, Z., Wang, H., Meng, H., & Cao, Y. (2024). Mesoporous Silica Nanoparticles Mediate SiRNA Delivery for Long-Term Multi-Gene Silencing in Intact Plants. Advanced science (Weinheim, Baden-Wurttemberg, Germany), 11(9), e2301358. https://doi.org/10.1002/advs.202301358.

246) Luther, D. C., Huang, R., Jeon, T., Zhang, X., Lee, Y. W., Nagaraj, H., & Rotello, V. M. (2020). Delivery of drugs, proteins, and nucleic acids using inorganic nanoparticles. Advanced drug delivery reviews, 156, 188–213. https://doi.org/10.1016/j.addr.2020.06.020.

247) Xu, L., Shao, Z., Fang, X., Xin, Z., Zhao, S., Zhang, H., Zhang, Y., Zheng, W., Yu, X., Zhang, Z., & Sun, L. (2024). Exploring precision treatments in immune-mediated inflammatory diseases: Harnessing the infinite potential of nucleic acid delivery. Exploration (Beijing, China), 5(1), 20230165. https://doi.org/10.1002/EXP.20230165

248) Kon, E., Elia, U., & Peer, D. (2022). Principles for designing an optimal mRNA lipid nanoparticle vaccine. Current opinion in biotechnology, 73, 329–336. https://doi.org/10.1016/j.copbio.2021.09.016

249) Kwon, H., Kim, M., Seo, Y., Moon, Y. S., Lee, H. J., Lee, K., & Lee, H. (2018). Emergence of synthetic mRNA: In vitro synthesis of mRNA and its applications in regenerative medicine. Biomaterials, 156, 172–193. https://doi.org/10.1016/j.biomaterials.2017.11.034

250) Eckstein F. (2000). Phosphorothioate oligodeoxynucleotides: what is their origin and what is unique about them?. Antisense & nucleic acid drug development, 10(2), 117–121. https://doi.org/10.1089/oli.1.2000.10.117

251) Geary, R. S., Yu, R. Z., & Levin, A. A. (2001). Pharmacokinetics of phosphorothioate antisense oligodeoxynucleotides. Current opinion in investigational drugs (London, England : 2000), 2(4), 562–573..

252) Crooke S. T. (1999). Molecular mechanisms of action of antisense drugs. Biochimica et biophysica acta, 1489(1), 31–44. https://doi.org/10.1016/s0167-4781(99)00148-7.

253) Stein, C. A., Subasinghe, C., Shinozuka, K., & Cohen, J. S. (1988). Physicochemical properties of phosphorothioate oligodeoxynucleotides. Nucleic acids research, 16(8), 3209–3221. https://doi.org/10.1093/nar/16.8.3209.

254) Verma, S., & Eckstein, F. (1998). Modified oligonucleotides: synthesis and strategy for users. Annual review of biochemistry, 67, 99–134. https://doi.org/10.1146/annurev.biochem.67.1.99

255) Rifai, A., Brysch, W., Fadden, K., Clark, J., & Schlingensiepen, K. H. (1996). Clearance kinetics, biodistribution, and organ saturability of phosphorothioate oligodeoxynucleotides in mice. The American journal of pathology, 149(2), 717–725.

256) Watanabe, T. A., Geary, R. S., & Levin, A. A. (2006). Plasma protein binding of an antisense oligonucleotide targeting human ICAM-1 (ISIS 2302). Oligonucleotides, 16(2), 169–180. https://doi.org/10.1089/oli.2006.16.169.

257) Guvakova, M. A., Yakubov, L. A., Vlodavsky, I., Tonkinson, J. L., & Stein, C. A. (1995). Phosphorothioate oligodeoxynucleotides bind to basic fibroblast growth factor, inhibit its binding to cell surface receptors, and remove it from low affinity binding sites on extracellular matrix. The Journal of biological chemistry, 270(6), 2620–2627. https://doi.org/10.1074/jbc.270.6.2620.

258) Rockwell, P., O'Connor, W. J., King, K., Goldstein, N. I., Zhang, L. M., & Stein, C. A. (1997). Cell-surface perturbations of the epidermal growth factor and vascular endothelial growth factor receptors by phosphorothioate oligodeoxynucleotides. Proceedings of the National Academy of Sciences of the United States of America, 94(12), 6523–6528. https://doi.org/10.1073/pnas.94.12.6523.

259) Teplova, M., Minasov, G., Tereshko, V., Inamati, G. B., Cook, P. D., Manoharan, M., & Egli, M. (1999). Crystal structure and improved antisense properties of 2'-O-(2-methoxyethyl)-RNA. Nature structural biology, 6(6), 535–539. https://doi.org/10.1038/9304

260) Baker, B. F., Lot, S. S., Condon, T. P., Cheng-Flournoy, S., Lesnik, E. A., Sasmor, H. M., & Bennett, C. F. (1997). 2'-O-(2-Methoxy)ethyl-modified anti-intercellular adhesion molecule 1 (ICAM-1) oligonucleotides selectively increase the ICAM-1 mRNA level and inhibit formation of the ICAM-1 translation initiation complex in human umbilical vein endothelial cells. The Journal of biological chemistry, 272(18), 11994–12000. https://doi.org/10.1074/jbc.272.18.11994.

261) Koshkin A.A., Singh S.K., Nielsen P., Rajwanshi V.K., Kumar R., Meldgaard M., Olsen C.E., Wengel J. LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and

unprecedented nucleic acid recognition. Tetrahedron. 1998;54:3607–3630. 10.1016/S0040-4020(98)00094-5

262) Obad, S., dos Santos, C. O., Petri, A., Heidenblad, M., Broom, O., Ruse, C., Fu, C., Lindow, M., Stenvang, J., Straarup, E. M., Hansen, H. F., Koch, T., Pappin, D., Hannon, G. J., & Kauppinen, S. (2011). Silencing of microRNA families by seed-targeting tiny LNAs. Nature genetics, 43(4), 371–378. https://doi.org/10.1038/ng.786.

263) Summerton J. E. (2007). Morpholino, siRNA, and S-DNA compared: impact of structure and mechanism of action on off-target effects and sequence specificity. Current topics in medicinal chemistry, 7(7), 651–660. https://doi.org/10.2174/156802607780487740

264) Iversen P. L. (2001). Phosphorodiamidate morpholino oligomers: favorable properties for sequence-specific gene inactivation. Current opinion in molecular therapeutics, 3(3), 235–238..

265) Khvorova, A., & Watts, J. K. (2017). The chemical evolution of oligonucleotide therapies of clinical utility. Nature biotechnology, 35(3), 238–248. https://doi.org/10.1038/nbt.3765

266) Crooke, S. T., Baker, B. F., Crooke, R. M., & Liang, X. H. (2021). Antisense technology: an overview and prospectus. Nature reviews. Drug discovery, 20(6), 427–453. https://doi.org/10.1038/s41573-021-00162-z.

267) Haraszti, R. A., Roux, L., Coles, A. H., Turanov, A. A., Alterman, J. F., Echeverria, D., Godinho, B. M. D. C., Aronin, N., & Khvorova, A. (2017). 5′-Vinylphosphonate improves tissue accumulation and efficacy of conjugated siRNAs in vivo. Nucleic acids research, 45(13), 7581–7592. https://doi.org/10.1093/nar/gkx507.

268) Uhlmann, E., & Peyman, A. (1990). Antisense oligonucleotides: A new therapeutic principle. Chemical Reviews, 90(4), 543–584. https://doi.org/10.1021/cr00102a001.

269) Micklefield J. (2001). Backbone modification of nucleic acids: synthesis, structure and therapeutic applications. Current medicinal chemistry, 8(10), 1157–1179. https://doi.org/10.2174/0929867013372391.

270) Krützfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., & Stoffel, M. (2005). Silencing of microRNAs in vivo with 'antagomirs'. Nature, 438(7068), 685–689. https://doi.org/10.1038/nature04303.

271) Amarzguioui, M., Holen, T., Babaie, E., & Prydz, H. (2003). Tolerance for mutations and chemical modifications in a siRNA. Nucleic acids research, 31(2), 589–595. https://doi.org/10.1093/nar/gkg147.

272) Soutschek, J., Akinc, A., Bramlage, B., Charisse, K., Constien, R., Donoghue, M., Elbashir, S., Geick, A., Hadwiger, P., Harborth, J., John, M., Kesavan, V., Lavine, G., Pandey, R. K., Racie, T., Rajeev, K. G., Röhl, I., Toudjarska, I., Wang, G., Wuschko, S., … Vornlocher, H. P. (2004).

Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. Nature, 432(7014), 173–178. https://doi.org/10.1038/nature03121.

273) Hall, A. H., Wan, J., Shaughnessy, E. E., Ramsay Shaw, B., & Alexander, K. A. (2004). RNA interference using boranophosphate siRNAs: structure-activity relationships. Nucleic acids research, 32(20), 5991–6000. https://doi.org/10.1093/nar/gkh936

274) Pradeep, S. P., Malik, S., Slack, F. J., & Bahal, R. (2023). Unlocking the potential of chemically modified peptide nucleic acids for RNA-based therapeutics. RNA (New York, N.Y.), 29(4), 434–445. https://doi.org/10.1261/rna.079498.122.

275) Setten, R. L., Rossi, J. J., & Han, S. P. (2019). The current state and future directions of RNAi-based therapeutics. Nature reviews. Drug discovery, 18(6), 421–446. https://doi.org/10.1038/s41573-019-0017-4.

276) Adachi, H., Hengesbach, M., Yu, Y. T., & Morais, P. (2021). From Antisense RNA to RNA Modification: Therapeutic Potential of RNA-Based Technologies. Biomedicines, 9(5), 550. https://doi.org/10.3390/biomedicines9050550.

277) Humphreys, S. C., Thayer, M. B., Campbell, J., Chen, W. L. K., Adams, D., Lade, J. M., & Rock, B. M. (2020). Emerging siRNA Design Principles and Consequences for Biotransformation and Disposition in Drug Development. Journal of medicinal chemistry, 63(12), 6407–6422. https://doi.org/10.1021/acs.jmedchem.9b01839.

278) Yi, C., & Pan, T. (2011). Cellular dynamics of RNA modification. Accounts of chemical research, 44(12), 1380–1388. https://doi.org/10.1021/ar200057m.

279) Peacock, H., Kannan, A., Beal, P. A., & Burrows, C. J. (2011). Chemical modification of siRNA bases to probe and enhance RNA interference. The Journal of organic chemistry, 76(18), 7295–7300. https://doi.org/10.1021/jo2012225.

280) Rohloff, J. C., Gelinas, A. D., Jarvis, T. C., Ochsner, U. A., Schneider, D. J., Gold, L., & Janjic, N. (2014). Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. Molecular therapy. Nucleic acids, 3(10), e201. https://doi.org/10.1038/mtna.2014.49

281) L. A. Moran, H. R. Horton, K. G. Scrimgeour, M. D. Perry, Principles of Biochemistry, 5th ed., Pearson, Boston, 2012

282) Trujillo, C. A., Nery, A. A., Alves, J. M., Martins, A. H., & Ulrich, H. (2007). Development of the anti-VEGF aptamer to a therapeutic agent for clinical ophthalmology. Clinical ophthalmology (Auckland, N.Z.), 1(4), 393–402..

283) Schmidt, K. S., Borkowski, S., Kurreck, J., Stephens, A. W., Bald, R., Hecht, M., Friebe, M., Dinkelborg, L., & Erdmann, V. A. (2004). Application of locked nucleic acids to improve aptamer in vivo stability and targeting function. Nucleic acids research, 32(19), 5757–5765. https://doi.org/10.1093/nar/gkh862

284) Cheung, Y. W., Röthlisberger, P., Mechaly, A. E., Weber, P., Levi-Acobas, F., Lo, Y., Wong, A. W. C., Kinghorn, A. B., Haouz, A., Savage, G. P., Hollenstein, M., & Tanner, J. A. (2020). Evolution of abiotic cubane chemistries in a nucleic acid aptamer allows selective recognition of a malaria biomarker. Proceedings of the National Academy of Sciences of the United States of America, 117(29), 16790–16798. https://doi.org/10.1073/pnas.2003267117

285) Ondruš, M., Sýkorová, V., Bednárová, L., Pohl, R., & Hocek, M. (2020). Enzymatic synthesis of hypermodified DNA polymers for sequence-specific display of four different hydrophobic groups. Nucleic acids research, 48(21), 11982–11993. https://doi.org/10.1093/nar/gkaa999.

286) Espinasse, A., Lembke, H. K., Cao, A. A., & Carlson, E. E. (2020). Modified nucleoside triphosphates in bacterial research for in vitro and live-cell applications. RSC chemical biology, 1(5), 333–351. https://doi.org/10.1039/d0cb00078g.

287) Kimoto, M., Yamashige, R., Matsunaga, K., Yokoyama, S., & Hirao, I. (2013). Generation of high-affinity DNA aptamers using an expanded genetic alphabet. Nature biotechnology, 31(5), 453–457. https://doi.org/10.1038/nbt.2556

288) Zhang, L., Yang, Z., Sefah, K., Bradley, K. M., Hoshika, S., Kim, M. J., Kim, H. J., Zhu, G., Jiménez, E., Cansiz, S., Teng, I. T., Champanhac, C., McLendon, C., Liu, C., Zhang, W., Gerloff, D. L., Huang, Z., Tan, W., & Benner, S. A. (2015). Evolution of functional six-nucleotide DNA. Journal of the American Chemical Society, 137(21), 6734–6737. https://doi.org/10.1021/jacs.5b02251

289) Kimoto, M., , & Hirao, I., (2020). Genetic alphabet expansion technology by creating unnatural base pairs. Chemical Society reviews, 49(21), 7602–7626. https://doi.org/10.1039/d0cs00457j

290) McKenzie, L. K., El-Khoury, R., Thorpe, J. D., Damha, M. J., & Hollenstein, M. (2021). Recent progress in non-native nucleic acid modifications. Chemical Society reviews, 50(8), 5126–5164. https://doi.org/10.1039/d0cs01430c

291) Freund, N., Fürst, M. J. L. J., & Holliger, P. (2022). New chemistries and enzymes for synthetic genetics. Current opinion in biotechnology, 74, 129–136. https://doi.org/10.1016/j.copbio.2021.11.004,

292) Hoffmann, S., Hoos, J., Klussmann, S., & Vonhoff, S. (2011). RNA aptamers and spiegelmers: synthesis, purification, and post-synthetic PEG conjugation. Current protocols in nucleic acid chemistry, Chapter 4, . https://doi.org/10.1002/0471142700.nc0446s46

293) Shi, H., He, X., Cui, W., Wang, K., Deng, K., Li, D., & Xu, F. (2014). Locked nucleic acid/DNA chimeric aptamer probe for tumor diagnosis with improved serum stability and extended imaging window in vivo. Analytica Chimica Acta, 812, 138–144. https://doi.org/10.1016/j.aca.2013.12.023

294) Shan, G. (2010). RNA interference as a gene knockdown technique. International Journal of Biochemistry & Cell Biology, 42(8), 1243–1251. https://doi.org/10.1016/j.biocel.2009.04.023.

295) Vaught, J. D., Bock, C., Carter, J., Fitzwater, T., Otis, M., Schneider, D., Rolando, J., Waugh, S., Wilcox, S. K., & Eaton, B. E. (2010). Expanding the chemistry of DNA for in vitro selection. Journal of the American Chemical Society, 132(12), 4141–4151. https://doi.org/10.1021/ja908035g

296) Gold, L., Ayers, D., Bertino, J., Bock, C., Bock, A., Brody, E. N., Carter, J., Dalby, A. B., Eaton, B. E., Fitzwater, T., Flather, D., Forbes, A., Foreman, T., Fowler, C., Gawande, B., Goss, M., Gunn, M., Gupta, S., Halladay, D., Heil, J., … Zichi, D. (2010). Aptamer-based multiplexed proteomic technology for biomarker discovery. PloS one, 5(12), e15004. https://doi.org/10.1371/journal.pone.0015004

297) Mutisya, D., Selvam, C., Kennedy, S. D., & Rozners, E. (2011). Synthesis and properties of triazole-linked RNA. Bioorganic & medicinal chemistry letters, 21(11), 3420–3422. https://doi.org/10.1016/j.bmcl.2011.03.111.

298) Diafa, S., & Hollenstein, M. (2015). Generation of Aptamers with an Expanded Chemical Repertoire. Molecules (Basel, Switzerland), 20(9), 16643–16671. https://doi.org/10.3390/molecules200916643.

299) Sau, S. P., & Hrdlicka, P. J. (2012). C2'-pyrene-functionalized triazole-linked DNA: universal DNA/RNA hybridization probes. The Journal of organic chemistry, 77(1), 5–16. https://doi.org/10.1021/jo201845z.

300) Sacca, B., Lacroix, L., & Mergny, J. L. (2005). The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. Nucleic Acids Research, 33(4), 1182–1192. https://doi.org/10.1093/nar/gki257.

301) Eremeeva, E., Abramov, M., Margamuljana, L., & Herdewijn, P. (2017). Base-Modified Nucleic Acids as a Powerful Tool for Synthetic Biology and Biotechnology. Chemistry (Weinheim an der Bergstrasse, Germany), 23(40), 9560–9576. https://doi.org/10.1002/chem.201700679S. Qian, D. Chang, S. He, Y. Li, Anal. Chim. Acta 2022, 1196, 339511

302) Sgallová, R., & Curtis, E. A. (2021). Secondary Structure Libraries for Artificial Evolution Experiments. Molecules (Basel, Switzerland), 26(6), 1671. https://doi.org/10.3390/molecules26061671.

303) Pobanz, K., & Lupták, A. (2016). Improving the odds: Influence of starting pools on in vitro selection outcomes. Methods (San Diego, Calif.), 106, 14–20. https://doi.org/10.1016/j.ymeth.2016.04.021.

304) Gevertz, J., Gan, H. H., & Schlick, T. (2005). In vitro RNA random pools are not structurally diverse: a computational analysis. RNA (New York, N.Y.), 11(6), 853–863. https://doi.org/10.1261/rna.7271405

305) Carothers, J. M., Oestreich, S. C., Davis, J. H., & Szostak, J. W. (2004). Informational complexity and functional activity of RNA structures. Journal of the American Chemical Society, 126(16), 5130–5137. https://doi.org/10.1021/ja031504a.

306) Ruff, K. M., Snyder, T. M., & Liu, D. R. (2010). Enhanced functional potential of nucleic acid aptamer libraries patterned to increase secondary structure. Journal of the American Chemical Society, 132(27), 9453–9464. https://doi.org/10.1021/ja103023m.

307) Chizzolini, F., Passalacqua, L. F. M., Oumais, M., Dingilian, A. I., Szostak, J. W., & Lupták, A. (2020). Large Phenotypic Enhancement of Structured Random RNA Pools. Journal of the American Chemical Society, 142(4), 1941–1951. https://doi.org/10.1021/jacs.9b11396.

308) Luo, X., McKeague, M., Pitre, S., Dumontier, M., Green, J., Golshani, A., Derosa, M. C., & Dehne, F. (2010). Computational approaches toward the design of pools for the in vitro selection of complex aptamers. RNA (New York, N.Y.), 16(11), 2252–2262. https://doi.org/10.1261/rna.2102210.

309) Röthlisberger, P., & Hollenstein, M. (2018). Aptamer chemistry. Advanced drug delivery reviews, 134, 3–21. https://doi.org/10.1016/j.addr.2018.04.007

310) Janik, B., Kotick, M. P., Kreiser, T. H., Reverman, L. F., Sommer, R. G., & Wilson, D. P. (1972). Synthesis and properties of poly 2'-fluoro-2'-deoxyuridylic acid. Biochemical and biophysical research communications, 46(3), 1153–1160. https://doi.org/10.1016/s0006-291x(72)80095-0

311) Hobbs, J., Sternbach, H., Sprinzl, M., & Eckstein, F. (1973). Polynucleotides containing 2'-amino-2'-deoxyribose and 2'-azido-2'-deoxyribose. Biochemistry, 12(25), 5138–5145. https://doi.org/10.1021/bi00749a018

312) Geary, R. S., Watanabe, T. A., Truong, L., Freier, S., Lesnik, E. A., Sioufi, N. B., Sasmor, H., Manoharan, M., & Levin, A. A. (2001). Pharmacokinetic properties of 2'-O-(2-methoxyethyl)-modified oligonucleotide analogs in rats. The Journal of pharmacology and experimental therapeutics, 296(3), 890–897..

313) Ng, E. W., Shima, D. T., Calias, P., Cunningham, E. T., Jr, Guyer, D. R., & Adamis, A. P. (2006). Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. Nature reviews. Drug discovery, 5(2), 123–132. https://doi.org/10.1038/nrd1955.

314) Vinores S. A. (2006). Pegaptanib in the treatment of wet, age-related macular degeneration. International journal of nanomedicine, 1(3), 263–268.

315) Gawande, B. N., Rohloff, J. C., Carter, J. D., von Carlowitz, I., Zhang, C., Schneider, D. J., & Janjic, N. (2017). Selection of DNA aptamers with two modified bases. Proceedings of the National Academy of Sciences of the United States of America, 114(11), 2898–2903. https://doi.org/10.1073/pnas.1615475114

316) Renders, M., Miller, E., Lam, C. H., & Perrin, D. M. (2017). Whole cell-SELEX of aptamers with a tyrosine-like side chain against live bacteria. Organic & biomolecular chemistry, 15(9), 1980–1989. https://doi.org/10.1039/c6ob02451c

317) Battersby, T. R., Ang, D. N., Burgstaller, P., Jurczyk, S. C., Bowser, M. T., Buchanan, D. D., Kennedy, R. T., & Benner, S. A. (1999). Quantitative analysis of receptors for adenosine nucleotides obtained via in vitro selection from a library incorporating a cationic nucleotide analog. Journal of the American Chemical Society, 121(42), 9781–9789. https://doi.org/10.1021/ja9816436

318) Eremeeva, E., & Herdewijn, P. (2019). Reprint of: Non Canonical Genetic Material. Current opinion in biotechnology, 60, 259–267. https://doi.org/10.1016/j.copbio.2019.11.009.

319) Ouaray, Z., Benner, S. A., Georgiadis, M. M., & Richards, N. G. J. (2020). Building better polymerases: Engineering the replication of expanded genetic alphabets. The Journal of biological chemistry, 295(50), 17046–17059. https://doi.org/10.1074/jbc.REV120.013745

320) Alamudi, S. H., Kimoto, M., & Hirao, I. (2021). Uptake mechanisms of cell-internalizing nucleic acid aptamers for applications as pharmacological agents. RSC medicinal chemistry, 12(10), 1640–1649. https://doi.org/10.1039/d1md00199j.

321) Biondi, E., & Benner, S. A. (2018). Artificially Expanded Genetic Information Systems for New Aptamer Technologies. Biomedicines, 6(2), 53. https://doi.org/10.3390/biomedicines6020053.

322) Biondi, E., Lane, J. D., Das, D., Dasgupta, S., Piccirilli, J. A., Hoshika, S., Bradley, K. M., Krantz, B. A., & Benner, S. A. (2016). Laboratory evolution of artificially expanded DNA gives redesignable aptamers that target the toxic form of anthrax protective antigen. Nucleic acids research, 44(20), 9565–9577. https://doi.org/10.1093/nar/gkw890.

323) Sefah, K., Yang, Z., Bradley, K. M., Hoshika, S., Jiménez, E., Zhang, L., Zhu, G., Shanker, S., Yu, F., Turek, D., Tan, W., & Benner, S. A. (2014). In vitro selection with artificial expanded genetic information systems. Proceedings of the National Academy of Sciences of the United States of America, 111(4), 1449–1454. https://doi.org/10.1073/pnas.1311778111.

324) Zumrut, H., Yang, Z., Williams, N., Arizala, J., Batool, S., Benner, S. A., & Mallikaratchy, P. (2020). Ligand-Guided Selection with Artificially Expanded Genetic Information Systems against TCR-CD3ε. Biochemistry, 59(4), 552–562. https://doi.org/10.1021/acs.biochem.9b00919.

325) Flamme, M., Figazzolo, C., Gasser, G., & Hollenstein, M. (2021). Enzymatic construction of metal-mediated nucleic acid base pairs. Metallomics : integrated biometal science, 13(4), mfab016. https://doi.org/10.1093/mtomcs/mfab016.

326) Flamme, M., Röthlisberger, P., Levi-Acobas, F., Chawla, M., Oliva, R., Cavallo, L., Gasser, G., Marlière, P., Herdewijn, P., & Hollenstein, M. (2020). Enzymatic Formation of an Artificial Base Pair Using a Modified Purine Nucleoside Triphosphate. ACS chemical biology, 15(11), 2872–2884. https://doi.org/10.1021/acschembio.0c00396.

327) Nakama, T., , Takezawa, Y., , & Shionoya, M., (2021). Site-specific polymerase incorporation of consecutive ligand-containing nucleotides for multiple metal-mediated base pairing. Chemical communications (Cambridge, England), 57(11), 1392–1395. https://doi.org/10.1039/d0cc07771b

328) Funai, T., , Tagawa, C., , Nakagawa, O., , Wada, S. I., , Ono, A., , & Urata, H., (2020). Enzymatic formation of consecutive thymine-$Hg^{II}$-thymine base pairs by DNA polymerases. Chemical communications (Cambridge, England), 56(80), 12025–12028. https://doi.org/10.1039/d0cc04423g

329) Kaul, C., Müller, M., Wagner, M., Schneider, S., & Carell, T. (2011). Reversible bond formation enables the replication and amplification of a crosslinking salen complex as an orthogonal base pair. Nature Chemistry, 3(10), 794–800. https://doi.org/10.1038/nchem.1117.

330) Nakama, T., Takezawa, Y., Sasaki, D., & Shionoya, M. (2020). Allosteric regulation of DNAzyme activities through intrastrand transformation induced by Cu(II)-mediated artificial base pairing. Journal of the American Chemical Society, 142(22), 10153–10162. https://doi.org/10.1021/jacs.0c03129

331) Ni, S., Yao, H., Wang, L., Lu, J., Jiang, F., Lu, A., & Zhang, G. (2017). Chemical modifications of nucleic acid aptamers for therapeutic purposes. International Journal of Molecular Sciences, 18(8), 1683. https://doi.org/10.3390/ijms18081683.

332) Ni, S., Zhuo, Z., Pan, Y., Yu, Y., Li, F., Liu, J., Wang, L., Wu, X., Li, D., Wan, Y., Zhang, L., Yang, Z., Zhang, B.-T., Lu, A., & Zhang, G. (2020). Recent progress in aptamer discoveries and modifications for therapeutic applications. ACS Applied Materials & Interfaces, 13(8), 9500–9519. https://doi.org/10.1021/acsami.0c05750.

333) Kovacevic, K. D., Gilbert, J. C., & Jilma, B. (2018). Pharmacokinetics, pharmacodynamics and safety of aptamers. Advanced Drug Delivery Reviews, 134, 36–50. https://doi.org/10.1016/j.addr.2018.10.008

334) Lee, C. H., Lee, S.-H., Kim, J. H., Noh, Y.-H., Noh, G.-J., & Lee, S.-W. (2015). Pharmacokinetics of a cholesterol-conjugated aptamer against the Hepatitis C Virus (HCV) NS5B protein. Molecular Therapy — Nucleic Acids, 4(10), e254. https://doi.org/10.1038/mtna.2015.30

335) Schmidt, P. G., Campbell, K. M., Hinds, K. D., & Cook, G. P. (2007). PEGylated bioactive molecules in biodegradable polymer microparticles. Expert opinion on biological therapy, 7(9), 1427–1436. https://doi.org/10.1517/14712598.7.9.1427

336) Wohlgemuth, R. (2007). Interfacing biocatalysis and organic synthesis. Journal of Chemical Technology and Biotechnology, 82(12), 1055–1062. https://doi.org/10.1002/jctb.1761

337) Hatano, A., Wakana, H., Terado, N., Kojima, A., Nishioka, C., Iizuka, Y., Imaizumi, T., & Uehara, S. (2019). Bio-catalytic synthesis of unnatural nucleosides possessing a large functional group such as a fluorescent molecule by purine nucleoside phosphorylase. Catalysis Science & Technology, 9(18), 5122–5129. https://doi.org/10.1039/C9CY01063G.

338) Kolb, H. C., Finn, M. G., & Sharpless, K. B. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angewandte Chemie (International ed. in English), 40(11), 2004–2021. https://doi.org/10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5

339) Zhang, J., Chen, B., Gan, C., Sun, H., Zhang, J., & Feng, L. (2023). A Comprehensive Review of Small Interfering RNAs (siRNAs): Mechanism, Therapeutic Targets, and Delivery Strategies for Cancer Therapy. International journal of nanomedicine, 18, 7605–7635. https://doi.org/10.2147/IJN.S436038.

340) Lauffer, M. C., van Roon-Mom, W., Aartsma-Rus, A., & N = 1 Collaborative (2024). Possibilities and limitations of antisense oligonucleotide therapies for the treatment of monogenic disorders. Communications medicine, 4(1), 6. https://doi.org/10.1038/s43856-023-00419-1

341) Kang, H., Ga, Y. J., Kim, S. H., Cho, Y. H., Kim, J. W., Kim, C., & Yeh, J. Y. (2023). Small interfering RNA (siRNA)-based therapeutic applications against viruses: principles, potential, and challenges. Journal of biomedical science, 30(1), 88. https://doi.org/10.1186/s12929-023-00981-9

342) Lefferts, J. W., Bierlaagh, M. C., Kroes, S., Nieuwenhuijze, N. D. A., Sonneveld van Kooten, H. N., Niemöller, P. J., Verburg, T. F., Janssens, H. M., Muilwijk, D., van Beuningen, S. F. B., van der Ent, C. K., & Beekman, J. M. (2023). CFTR Function Restoration upon Elexacaftor/Tezacaftor/Ivacaftor Treatment in Patient-Derived Intestinal Organoids with Rare CFTR Genotypes. International journal of molecular sciences, 24(19), 14539. https://doi.org/10.3390/ijms241914539

343) Townshend, R. J. L., Eismann, S., Watkins, A. M., Rangan, R., Karelina, M., Das, R., & Dror, R. O. (2021). Geometric deep learning of RNA structure. Science (New York, N.Y.), 373(6558), 1047–1051. https://doi.org/10.1126/science.abe5650

344) Hwang, H., Jeon, H., Yeo, N., & Baek, D. (2024). Big data and deep learning for RNA biology. Experimental & molecular medicine, 56(6), 1293–1321. https://doi.org/10.1038/s12276-024-01243-w.

345) Allen, D., Rosenberg, M., & Hendel, A. (2021). Using Synthetically Engineered Guide RNAs to Enhance CRISPR Genome Editing Systems in Mammalian Cells. Frontiers in genome editing, 2, 617910. https://doi.org/10.3389/fgeed.2020.617910

346) Sinha S., Vohora D., editors. Pharmaceutical Medicine and Translational Clinical Research. Elsevier Inc.; Amsterdam, The Netherlands: 2018. Drug discovery and development: ISBN: 978-0-12-802103-3

347) Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. British journal of pharmacology, 162(6), 1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x

348) Zhong, F., Xing, J., Li, X., Liu, X., Fu, Z., Xiong, Z., Lu, D., Wu, X., Zhao, J., Tan, X., Li, F., Luo, X., Li, Z., Chen, K., Zheng, M., & Jiang, H. (2018). Artificial intelligence in drug design. Science China. Life sciences, 61(10), 1191–1204. https://doi.org/10.1007/s11427-018-9342-2

349) Hou, T., & Xu, X. (2004). Recent development and application of virtual screening in drug discovery: an overview. Current pharmaceutical design, 10(9), 1011–1033. https://doi.org/10.2174/1381612043452721.

350) Yu, W., & MacKerell, A. D., Jr (2017). Computer-Aided Drug Design Methods. Methods in molecular biology (Clifton, N.J.), 1520, 85–106. https://doi.org/10.1007/978-1-4939-6634-9_5

351) Gurung, A. B., Ali, M. A., Lee, J., Farah, M. A., & Al-Anazi, K. M. (2021). An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. BioMed research international, 2021, 8853056. https://doi.org/10.1155/2021/8853056.

352) Macalino, S. J., Gosu, V., Hong, S., & Choi, S. (2015). Role of computer-aided drug design in modern drug discovery. Archives of pharmacal research, 38(9), 1686–1701. https://doi.org/10.1007/s12272-015-0640-5

353) Duch, W., Swaminathan, K., & Meller, J. (2007). Artificial intelligence approaches for rational drug design and discovery. Current pharmaceutical design, 13(14), 1497–1508. https://doi.org/10.2174/138161207780765954

354) Batool, M., Ahmad, B., & Choi, S. (2019). A Structure-Based Drug Discovery Paradigm. International journal of molecular sciences, 20(11), 2783. https://doi.org/10.3390/ijms20112783

355) Huang, H.-J., Yu, H. W., Chen, C.-Y., et al. (2010). Current developments of computer-aided drug design. Journal of the Taiwan Institute of Chemical Engineers, 41(6), 623–635. https://doi.org/10.1016/j.jtice.2010.03.017

356) Zhang, Y., Luo, M., Wu, P., Wu, S., Lee, T. Y., & Bai, C. (2022). Application of Computational Biology and Artificial Intelligence in Drug Design. International journal of molecular sciences, 23(21), 13568. https://doi.org/10.3390/ijms232113568

357) Lionta, E., Spyrou, G., Vassilatis, D. K., & Cournia, Z. (2014). Structure-based virtual screening for drug discovery: principles, applications and recent advances. Current topics in medicinal chemistry, 14(16), 1923–1938. https://doi.org/10.2174/1568026614666140929124445.

358) Kalyaanamoorthy, S., & Chen, Y. P. (2011). Structure-based drug design to augment hit discovery. Drug discovery today, 16(17-18), 831–839. https://doi.org/10.1016/j.drudis.2011.07.006

359) Wang, X., Song, K., Li, L., & Chen, L. (2018). Structure-Based Drug Design Strategies and Challenges. Current topics in medicinal chemistry, 18(12), 998–1006. https://doi.org/10.2174/1568026618666180813152921.

360) Lee, J., Wu, S., Zhang, Y. (2009). Ab Initio Protein Structure Prediction. In: Rigden, D.J. (eds) From Protein Structure to Function with Bioinformatics. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-9058-5_1.

361) Vyas, V. K., Ukawala, R. D., Ghate, M., & Chintha, C. (2012). Homology modeling a fast tool for drug discovery: current perspectives. Indian journal of pharmaceutical sciences, 74(1), 1–17. https://doi.org/10.4103/0250-474X.102537

362) Lemer, C. M., Rooman, M. J., & Wodak, S. J. (1995). Protein structure prediction by threading methods: evaluation of current techniques. Proteins, 23(3), 337–355. https://doi.org/10.1002/prot.340230308.

363) Muhammed, M. T., & Aki-Yalcin, E. (2019). Homology modeling in drug discovery: Overview, current applications, and future perspectives. Chemical biology & drug design, 93(1), 12–20. https://doi.org/10.1111/cbdd.13388.

364) Eswar, N., Eramian, D., Webb, B., Shen, M. Y., & Sali, A. (2008). Protein structure modeling with MODELLER. Methods in molecular biology (Clifton, N.J.), 426, 145–159. https://doi.org/10.1007/978-1-60327-058-8_8

365) Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis, 18(15), 2714–2723. https://doi.org/10.1002/elps.1150181505

366) Sánchez, R., & Sali, A. (1999). ModBase: a database of comparative protein structure models. Bioinformatics (Oxford, England), 15(12), 1060–1061. https://doi.org/10.1093/bioinformatics/15.12.1060

367) Harrell, C. R., & Price, R. N. (1998). Simulation modeling and optimization using ProModel. In D. J. Medeiros, E. F. Watson, J. S. Carson, & M. S. Manivannan (Eds.), Proceedings of the 1998 Winter Simulation Conference (pp. 191–197). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1145/268437.268586.

368) Xu, J., Jiao, F., & Yu, L. (2008). Protein structure prediction using threading. In M. J. Zaki & C. Bystroff (Eds.), Protein structure prediction (pp. 91–121). Humana Press. https://doi.org/10.1007/978-1-59745-574-9_4

369) Zhang Y. (2008). I-TASSER server for protein 3D structure prediction. BMC bioinformatics, 9, 40. https://doi.org/10.1186/1471-2105-9-40

370) Shi, J., Blundell, T. L., & Mizuguchi, K. (2001). FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. Journal of molecular biology, 310(1), 243–257. https://doi.org/10.1006/jmbi.2001.4762

371) Lobley, A., Sadowski, M. I., & Jones, D. T. (2009). pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. Bioinformatics (Oxford, England), 25(14), 1761–1767. https://doi.org/10.1093/bioinformatics/btp302

372) Kelley, L. A., & Sternberg, M. J. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nature protocols, 4(3), 363–371. https://doi.org/10.1038/nprot.2009.2

373) Yousef, M., Abdelkader, T., & El-Bahnasy, K. (2019). Performance comparison of ab initio protein structure prediction methods. Ain Shams Engineering Journal, 10(4), 713–719. https://doi.org/10.1016/j.asej.2019.03.004.

374) Rohl, C. A., Strauss, C. E., Misura, K. M., & Baker, D. (2004). Protein structure prediction using Rosetta. Methods in enzymology, 383, 66–93. https://doi.org/10.1016/S0076-6879(04)83004-0

375) Wlodawer, A., & Vondrasek, J. (1998). Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. Annual review of biophysics and biomolecular structure, 27, 249–284. https://doi.org/10.1146/annurev.biophys.27.1.249

376) Clark D. E. (2006). What has computer-aided molecular design ever done for drug discovery?. Expert opinion on drug discovery, 1(2), 103–110. https://doi.org/10.1517/17460441.1.2.103

377) Anderson, A. C. (2003). The process of structure-based drug design. Chemistry & Biology, 10(9), 787–797. https://doi.org/10.1016/j.chembiol.2003.09.002.

378) Grover, S., Apushkin, M. A., & Fishman, G. A. (2006). Topical dorzolamide for the treatment of cystoid macular edema in patients with retinitis pigmentosa. American Journal of Ophthalmology, 141(5), 850–858. https://doi.org/10.1016/j.ajo.2005.12.030.

379) Marrakchi, H., Lanéelle, G., & Quémard, A. (2000). InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II. Microbiology, 146(2), 289–296. https://doi.org/10.1099/00221287-146-2-289.

380) Dadashpour, S., Tuylu Kucukkilinc, T., Unsal Tan, O., Ozadali, K., Irannejad, H., & Emami, S. (2015). Design, synthesis and in vitro study of 5,6-diaryl-1,2,4-triazine-3-ylthioacetate derivatives as COX-2 and β-amyloid aggregation inhibitors. Archiv der Pharmazie, 348(3), 179–187. https://doi.org/10.1002/ardp.201400400.

381) Miller, Z., Kim, K.-S., Lee, D.-M., et al. (2015). Proteasome inhibitors with pyrazole scaffolds from structure-based virtual screening. Journal of Medicinal Chemistry, 58(4), 2036–2041. https://doi.org/10.1021/jm501344n.

382) Pan, L., Gardner, C. L., Pagliai, F. A., Gonzalez, C. F., & Lorca, G. L. (2017). Identification of the tolfenamic acid binding pocket in PrbP from Liberibacter asiaticus. Frontiers in Microbiology, 8, 1591. https://doi.org/10.3389/fmicb.2017.01591.

383) Binkowski, T. A., Naghibzadeh, S., & Liang, J. (2003). CASTp: Computed atlas of surface topography of proteins. Nucleic Acids Research, 31(13), 3352–3355. https://doi.org/10.1093/nar/gkg512.

384) Volkamer, A., Kuhn, D., Rippmann, F., & Rarey, M. (2012). DoGSiteScorer: A web server for automatic binding site prediction, analysis and druggability assessment. Bioinformatics, 28(15), 2074–2075. https://doi.org/10.1093/bioinformatics/bts310.

385) Sun, J., & Chen, K. (2017). NSiteMatch: Prediction of binding sites of nucleotides by identifying the structure similarity of local surface patches. Computational and Mathematical Methods in Medicine, 2017, 5471607. https://doi.org/10.1155/2017/5471607

386) Tan, K. P., Varadarajan, R., & Madhusudhan, M. S. (2011). DEPTH: A web server to compute depth and predict small-molecule binding cavities in proteins. Nucleic Acids Research, 39(suppl_2), W242–W248. https://doi.org/10.1093/nar/gkr356

387) Zhu, H., & Pisabarro, M. T. (2011). MSPocket: An orientation-independent algorithm for the

detection of ligand binding pockets. Bioinformatics, 27(3), 351–358. https://doi.org/10.1093/bioinformatics/btq672

388) Huang, B. (2009). MetaPocket: A meta approach to improve protein ligand binding site prediction. OMICS: A Journal of Integrative Biology, 13(4), 325–330. https://doi.org/10.1089/omi.2009.0045

389) Laurie, A. T. R., & Jackson, R. M. (2005). Q-SiteFinder: An energy-based method for the prediction of protein-ligand binding sites. Bioinformatics, 21(9), 1908–1916. https://doi.org/10.1093/bioinformatics/bti315

390) Lavecchia, A., & Di Giovanni, C. (2013). Virtual screening strategies in drug discovery: a critical review. Current medicinal chemistry, 20(23), 2839–2860. https://doi.org/10.2174/09298673113209990001.

391) Cheng, T., Li, Q., Zhou, Z., Wang, Y., & Bryant, S. H. (2012). Structure-based virtual screening for drug discovery: a problem-centric review. The AAPS journal, 14(1), 133–141. https://doi.org/10.1208/s12248-012-9322-0

392) Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B. A., Wang, J., Yu, B., Zhang, J., & Bryant, S. H. (2016). PubChem Substance and Compound databases. Nucleic acids research, 44(D1), D1202–D1213. https://doi.org/10.1093/nar/gkv951

393) Irwin, J. J., & Shoichet, B. K. (2005). ZINC--a free database of commercially available compounds for virtual screening. Journal of chemical information and modeling, 45(1), 177–182. https://doi.org/10.1021/ci049714+

394) https://mcule.com/

395) Pence, H. E., & Williams, A. (2010). ChemSpider: An online chemical information resource. Journal of Chemical Education, 87(11), 1123–1124. https://doi.org/10.1021/ed100697w

396) Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic acids research, 40(Database issue), D1100–D1107. https://doi.org/10.1093/nar/gkr777

397) Wishart, D. S., Knox, C., Guo, A. C., Cheng, D., Shrivastava, S., Tzur, D., Gautam, B., & Hassanali, M. (2008). DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic acids research, 36(Database issue), D901–D906. https://doi.org/10.1093/nar/gkm958

398) Lipinski C. A. (2004). Lead- and drug-like compounds: the rule-of-five revolution. Drug discovery today. Technologies, 1(4), 337–341. https://doi.org/10.1016/j.ddtec.2004.11.007

399) Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports, 7, 42717. https://doi.org/10.1038/srep42717

400) Athanasiadis, E., Cournia, Z., & Spyrou, G. (2012). ChemBioServer: a web-based pipeline for filtering, clustering and visualization of chemical compounds used in drug discovery. Bioinformatics (Oxford, England), 28(22), 3002–3003. https://doi.org/10.1093/bioinformatics/bts551

401) Lagorce, D., Sperandio, O., Galons, H., Miteva, M. A., & Villoutreix, B. O. (2008). FAF-Drugs2: free ADME/tox filtering tool to assist drug discovery and chemical biology projects. BMC bioinformatics, 9, 396. https://doi.org/10.1186/1471-2105-9-396

402) Kalliokoski, T., Salo, H. S., Lahtela-Kakkonen, M., & Poso, A. (2009). The effect of ligand-based tautomer and protomer prediction on structure-based virtual screening. Journal of chemical information and modeling, 49(12), 2742–2748. https://doi.org/10.1021/ci900364w.

403) Sadowski, J., Rudolph, C., & Gasteiger, J. (1992). The generation of 3D models of host-guest complexes. Analytica Chimica Acta, 265(2), 233-241. https://doi.org/10.1016/0003-2670(92)85029-6

404) Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of computational chemistry, 30(16), 2785–2791. https://doi.org/10.1002/jcc.21256.

405) Schrödinger, LLC. (2013). LigPrep [Software]. New York, NY.

406) Chemical Computing Group Inc. (2013). Molecular Operating Environment (MOE) [Software]. Montreal, QC, Canada.

407) Scienomics SARL. (2014). MAPS, version 3.4 [Software]. Paris, France.

408) DISI.. Preparing the ligand. Retrieved from http://wiki.uoft.bkslab.org/index.php/Preparing_the_ligand

409) Accelrys. Pipeline Pilot: Combinatorial library design chemistry component. Retrieved from www.accelrys.com

410) HyperCube. Hyperchem [Software]. Gainesville, FL.

411) Paggi, J. M., Pandit, A., & Dror, R. O. (2024). The Art and Science of Molecular Docking. Annual review of biochemistry, 93(1), 389–410. https://doi.org/10.1146/annurev-biochem-030222-120000

412) Agrawal, P., Singh, H., Srivastava, H. K., Singh, S., Kishore, G., & Raghava, G. P. S. (2019). Benchmarking of different molecular docking methods for protein-peptide docking. BMC bioinformatics, 19(Suppl 13), 426. https://doi.org/10.1186/s12859-018-2449-y

413) Martin, S. J., Chen, I. J., Chan, A. W. E., & Foloppe, N. (2020). Modelling the binding mode of macrocycles: Docking and conformational sampling. Bioorganic & medicinal chemistry, 28(1), 115143. https://doi.org/10.1016/j.bmc.2019.115143

414) Morris, G. M., Huey, R., Lindstrom, W., et al. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry, 30(16), 2785–2791. https://doi.org/10.1002/jcc.21256

415) Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31(2), 455–461. https://doi.org/10.1002/jcc.21334

416) Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. Journal of Molecular Biology, 267(3), 727–748. https://doi.org/10.1006/jmbi.1996.0897

417) Wu, G., Robertson, D. H., Brooks, C. L., III, & Vieth, M. (2003). Detailed analysis of grid-based molecular docking: A case study of CDOCKER—A CHARMm-based MD docking algorithm. Journal of Computational Chemistry, 24(13), 1549–1562. https://doi.org/10.1002/jcc.10306

418) Kramer, B., Rarey, M., & Lengauer, T. (1999). Evaluation of the FLEXX incremental construction algorithm for protein-ligand docking. Proteins: Structure, Function, and Genetics, 37(2), 228–241. https://doi.org/10.1002/(SICI)1097-0134(19991101)37:2<228::AID-PROT8>3.0.CO;2-8

419) Jain, A. N. (2003). Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. Journal of Medicinal Chemistry, 46(4), 499–511. https://doi.org/10.1021/jm020406h

420) Friesner, R. A., Banks, J. L., Murphy, R. B., et al. (2004). Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of Medicinal Chemistry, 47(7), 1739–1749. https://doi.org/10.1021/jm0306430

421) Allen, W. J., Balius, T. E., Mukherjee, S., et al. (2015). DOCK 6: Impact of new features and current docking performance. Journal of Computational Chemistry, 36(15), 1132–1156. https://doi.org/10.1002/jcc.23905

422) Grosdidier, A., Zoete, V., & Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Research, 39(suppl), W270–W277. https://doi.org/10.1093/nar/gkr366

423) Yan, Y., Tao, H., He, J., & Huang, S. Y. (2020). The HDOCK server for integrated protein-protein docking. Nature protocols, 15(5), 1829–1852. https://doi.org/10.1038/s41596-020-0312-x

424) Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V., & Mee, R. P. (1997). Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. Journal of computer-aided molecular design, 11(5), 425–445. https://doi.org/10.1023/a:1007996124545

425) Adeshina, Y. O., Deeds, E. J., & Karanicolas, J. (2020). Machine learning classification can reduce false positives in structure-based virtual screening. Proceedings of the National Academy of Sciences of the United States of America, 117(31), 18477–18488. https://doi.org/10.1073/pnas.2000585117

426) Li, J., Fu, A., & Zhang, L. (2019). An Overview of Scoring Functions Used for Protein-Ligand Interactions in Molecular Docking. Interdisciplinary sciences, computational life sciences, 11(2), 320–328. https://doi.org/10.1007/s12539-019-00327-w

427) Koes, D. R., Baumgartner, M. P., & Camacho, C. J. (2013). Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. Journal of chemical information and modeling, 53(8), 1893–1904. https://doi.org/10.1021/ci300604z

428) Gilks, W. R., Richardson, S., & Spiegelhalter, D. (Eds.). (1996). Markov Chain Monte Carlo in Practice. Chapman & Hall/CRC

429) Ravindranath, P. A., Forli, S., Goodsell, D. S., Olson, A. J., & Sanner, M. F. (2015). AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. PLoS computational biology, 11(12), e1004586. https://doi.org/10.1371/journal.pcbi.1004586

430) Coleman, R. G., Carchia, M., Sterling, T., Irwin, J. J., & Shoichet, B. K. (2013). Ligand pose and orientational sampling in molecular docking. PloS one, 8(10), e75992. https://doi.org/10.1371/journal.pone.0075992

431) Enyedy, I. J., & Egan, W. J. (2008). Can we use docking and scoring for hit-to-lead optimization?. Journal of computer-aided molecular design, 22(3-4), 161–168. https://doi.org/10.1007/s10822-007-9165-4

432) Ramírez, D., & Caballero, J. (2016). Is It Reliable to Use Common Molecular Docking Methods for Comparing the Binding Affinities of Enantiomer Pairs for Their Protein Target?. International journal of molecular sciences, 17(4), 525. https://doi.org/10.3390/ijms17040525

433) Che, T., Majumdar, S., Zaidi, S. A., Ondachi, P., McCorvy, J. D., Wang, S., Mosier, P. D., Uprety, R., Vardy, E., Krumm, B. E., Han, G. W., Lee, M. Y., Pardon, E., Steyaert, J., Huang, X. P., Strachan, R. T., Tribo, A. R., Pasternak, G. W., Carroll, F. I., Stevens, R. C., … Roth, B. L. (2018).

Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor. Cell, 172(1-2), 55–67.e15. https://doi.org/10.1016/j.cell.2017.12.011

434) Song, G., Yang, D., Wang, Y., de Graaf, C., Zhou, Q., et al. (2017). Human GLP-1 receptor transmembrane domain structure in complex with allosteric modulators. Nature, 546, 312–315. https://doi.org/10.1038/nature22378

435) Zhang, K., Zhang, J., Gao, Z.-G., Zhang, D., Zhu, L., et al. (2014). Structure of the human P2Y12 receptor in complex with an antithrombotic drug. Nature, 509, 115–118. https://doi.org/10.1038/nature13083

436) Taniguchi, R., Inoue, A., Sayama, M., Uwamizu, A., Yamashita, K., Hirata, K., Yoshida, M., Tanaka, Y., Kato, H. E., Nakada-Nakura, Y., Otani, Y., Nishizawa, T., Doi, T., Ohwada, T., Ishitani, R., Aoki, J., & Nureki, O. (2017). Structural insights into ligand recognition by the lysophosphatidic acid receptor $LPA_6$. Nature, 548(7667), 356–360. https://doi.org/10.1038/nature23448

437) El Daibani, A., Paggi, J. M., Kim, K., Laloudakis, Y. D., Popov, P., Bernhard, S. M., Krumm, B. E., Olsen, R. H. J., Diberto, J., Carroll, F. I., Katritch, V., Wünsch, B., Dror, R. O., & Che, T. (2023). Molecular mechanism of biased signaling at the kappa opioid receptor. Nature communications, 14(1), 1338. https://doi.org/10.1038/s41467-023-37041-7

438) Powers, A. S., Pham, V., Burger, W. A. C., Thompson, G., Laloudakis, Y., Barnes, N. W., Sexton, P. M., Paul, S. M., Christopoulos, A., Thal, D. M., Felder, C. C., Valant, C., & Dror, R. O. (2023). Structural basis of efficacy-driven ligand selectivity at GPCRs. Nature chemical biology, 19(7), 805–814. https://doi.org/10.1038/s41589-022-01247-5

439) McCorvy, J. D., Butler, K. V., Kelly, B., Rechsteiner, K., Karpiak, J., Betz, R. M., Kormos, B. L., Shoichet, B. K., Dror, R. O., Jin, J., & Roth, B. L. (2018). Structure-inspired design of β-arrestin-biased ligands for aminergic GPCRs. Nature chemical biology, 14(2), 126–134. https://doi.org/10.1038/nchembio.2527

440) Kelly, B., Hollingsworth, S. A., Blakemore, D. C., Owen, R. M., Storer, R. I., Swain, N. A., Aydin, D., Torella, R., Warmus, J. S., & Dror, R. O. (2021). Delineating the Ligand-Receptor Interactions That Lead to Biased Signaling at the μ-Opioid Receptor. Journal of chemical information and modeling, 61(7), 3696–3707. https://doi.org/10.1021/acs.jcim.1c00585

441) Lei, T., Hu, Z., Ding, R., Chen, J., Li, S., Zhang, F., Pu, X., & Zhao, N. (2020). Exploring the Activation Mechanism of a Metabotropic Glutamate Receptor Homodimer via Molecular Dynamics Simulation. ACS chemical neuroscience, 11(2), 133–145. https://doi.org/10.1021/acschemneuro.9b00425

442) Hollingsworth, S. A., & Dror, R. O. (2018). Molecular Dynamics Simulation for All. Neuron, 99(6), 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011

443) Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., Grubmüller, H., & MacKerell, A. D., Jr (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nature methods, 14(1), 71–73. https://doi.org/10.1038/nmeth.4067

444) D'Amore, L., Hahn, D. F., Dotson, D. L., Horton, J. T., Anwar, J., Craig, I., Fox, T., Gobbi, A., Lakkaraju, S. K., Lucas, X., Meier, K., Mobley, D. L., Narayanan, A., Schindler, C. E. M., Swope, W. C., In 't Veld, P. J., Wagner, J., Xue, B., & Tresadern, G. (2022). Collaborative Assessment of Molecular Geometries and Energies from the Open Force Field. Journal of chemical information and modeling, 62(23), 6094–6104. https://doi.org/10.1021/acs.jcim.2c01185

445) Cournia, Z., Allen, B., & Sherman, W. (2017). Relative Binding Free Energy Calculations in Drug Discovery: Recent Advances and Practical Considerations. Journal of chemical information and modeling, 57(12), 2911–2937. https://doi.org/10.1021/acs.jcim.7b00564

446) Li, Z., Huang, Y., Wu, Y., Chen, J., Wu, D., Zhan, C. G., & Luo, H. B. (2019). Absolute Binding Free Energy Calculation and Design of a Subnanomolar Inhibitor of Phosphodiesterase-10. Journal of medicinal chemistry, 62(4), 2099–2111. https://doi.org/10.1021/acs.jmedchem.8b01763

447) Cournia, Z., Allen, B. K., Beuming, T., Pearlman, D. A., Radak, B. K., & Sherman, W. (2020). Rigorous Free Energy Simulations in Virtual Screening. Journal of chemical information and modeling, 60(9), 4153–4169. https://doi.org/10.1021/acs.jcim.0c00116

448) Heinzelmann, G., & Gilson, M. K. (2021). Automation of absolute protein-ligand binding free energy calculations for docking refinement and compound evaluation. Scientific reports, 11(1), 1116. https://doi.org/10.1038/s41598-020-80769-1

449) Miller, E. B., Murphy, R. B., Sindhikara, D., Borrelli, K. W., Grisewood, M. J., Ranalli, F., Dixon, S. L., Jerome, S., Boyles, N. A., Day, T., Ghanakota, P., Mondal, S., Rafi, S. B., Troast, D. M., Abel, R., & Friesner, R. A. (2021). Reliable and Accurate Solution to the Induced Fit Docking Problem for Protein-Ligand Binding. Journal of chemical theory and computation, 17(4), 2630–2639. https://doi.org/10.1021/acs.jctc.1c00136

450) Hsu, D. J., Davidson, R. B., Sedova, A., & Glaser, J. (2023). tinyIFD: A High-Throughput Binding Pose Refinement Workflow Through Induced-Fit Ligand Docking. Journal of chemical information and modeling, 63(11), 3438–3447. https://doi.org/10.1021/acs.jcim.2c01530

451) Durrant, J. D., & McCammon, J. A. (2010). NNScore: a neural-network-based scoring function for the characterization of protein-ligand complexes. Journal of chemical information and modeling, 50(10), 1865–1871. https://doi.org/10.1021/ci100244v

452) Feinberg, E. N., Sur, D., Wu, Z., Husic, B. E., Mai, H., Li, Y., Sun, S., Yang, J., Ramsundar, B., & Pande, V. S. (2018). PotentialNet for Molecular Property Prediction. ACS central science, 4(11), 1520–1530. https://doi.org/10.1021/acscentsci.8b00507

453) Francoeur, P. G., Masuda, T., Sunseri, J., Jia, A., Iovanisci, R. B., Snyder, I., & Koes, D. R. (2020). Three-Dimensional Convolutional Neural Networks and a Cross-Docked Data Set for Structure-Based Drug Design. Journal of chemical information and modeling, 60(9), 4200–4215. https://doi.org/10.1021/acs.jcim.0c00411

454) Jiménez J, Škalič M, Martínez-Rosell G, De Fabritiis G. 2018. KDEEP: protein–ligand absolute binding affinity prediction via 3D-convolutional neural networks. J. Chem. Inform. Model. 58:287–96

455) Wallach, I., Dzamba, M., & Heifets, A. (2015). AtomNet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery (arXiv:1510.02855) [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1510.02855

456) Ballester, P. J., & Mitchell, J. B. (2010). A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. Bioinformatics (Oxford, England), 26(9), 1169–1175. https://doi.org/10.1093/bioinformatics/btq112

457) Mark, A. E., & van Gunsteren, W. F. (1994). Decomposition of the free energy of a system in terms of specific interactions. Implications for theoretical and experimental studies. Journal of molecular biology, 240(2), 167–176. https://doi.org/10.1006/jmbi.1994.1430

458) Suriana, P., Paggi, J. M., & Dror, R. O. (2023). FlexVDW: A machine learning approach to account for protein flexibility in ligand docking (arXiv:2303.11494) [Preprint]. arXiv. https://arxiv.org/abs/2303.11494

459) Masters, M. R., Mahmoud, A. H., Wei, Y., & Lill, M. A. (2023). Deep Learning Model for Efficient Protein-Ligand Docking with Implicit Side-Chain Flexibility. Journal of chemical information and modeling, 63(6), 1695–1707. https://doi.org/10.1021/acs.jcim.2c01436

460) Liu, Z., Su, M., Han, L., Liu, J., Yang, Q., Li, Y., & Wang, R. (2017). Forging the Basis for Developing Protein-Ligand Interaction Scoring Functions. Accounts of chemical research, 50(2), 302–309. https://doi.org/10.1021/acs.accounts.6b00491

461) Yang, Y., Yao, K., Repasky, M. P., Leswing, K., Abel, R., Shoichet, B. K., & Jerome, S. V. (2021). Efficient Exploration of Chemical Space with Docking and Deep Learning. Journal of chemical theory and computation, 17(11), 7106–7119. https://doi.org/10.1021/acs.jctc.1c00810

462) Gentile, F., Yaacoub, J. C., Gleave, J., Fernandez, M., Ton, A. T., Ban, F., Stern, A., & Cherkasov, A. (2022). Artificial intelligence-enabled virtual screening of ultra-large chemical

libraries with deep docking. Nature protocols, 17(3), 672–697. https://doi.org/10.1038/s41596-021-00659-2

463) Graff, D. E., Shakhnovich, E. I., & Coley, C. W. (2021). Accelerating high-throughput virtual screening through molecular pool-based active learning. Chemical science, 12(22), 7866–7881. https://doi.org/10.1039/d0sc06805e

464) Thomas, M., Bender, A., & de Graaf, C. (2023). Integrating structure-based approaches in generative molecular design. Current opinion in structural biology, 79, 102559. https://doi.org/10.1016/j.sbi.2023.102559

465) Spiegel, J. O., & Durrant, J. D. (2020). AutoGrow4: an open-source genetic algorithm for de novo drug design and lead optimization. Journal of cheminformatics, 12(1), 25. https://doi.org/10.1186/s13321-020-00429-4

466) Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert opinion on drug discovery, 10(5), 449–461. https://doi.org/10.1517/17460441.2015.1032936

467) Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J. Z. H., & Hou, T. (2019). End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. Chemical reviews, 119(16), 9478–9508. https://doi.org/10.1021/acs.chemrev.9b00055

468) Christ, C. D., Mark, A. E., & van Gunsteren, W. F. (2010). Basic ingredients of free energy calculations: a review. Journal of computational chemistry, 31(8), 1569–1582. https://doi.org/10.1002/jcc.21450.

469) Jorgensen W. L. (2004). The many roles of computation in drug discovery. Science (New York, N.Y.), 303(5665), 1813–1818. https://doi.org/10.1126/science.1096361.

470) Shirts M. R. (2012). Best practices in free energy calculations for drug design. Methods in molecular biology (Clifton, N.J.), 819, 425–467. https://doi.org/10.1007/978-1-61779-465-0_26.

471) Christ, C. D., & Fox, T. (2014). Accuracy assessment and automation of free energy calculations for drug design. Journal of chemical information and modeling, 54(1), 108–120. https://doi.org/10.1021/ci4004199.

472) Williams-Noonan, B. J., Yuriev, E., & Chalmers, D. K. (2018). Free Energy Methods in Drug Design: Prospects of "Alchemical Perturbation" in Medicinal Chemistry. Journal of medicinal chemistry, 61(3), 638–649. https://doi.org/10.1021/acs.jmedchem.7b00681.

473) Zhou R. (2003). Free energy landscape of protein folding in water: explicit vs. implicit solvent. Proteins, 53(2), 148–161. https://doi.org/10.1002/prot.10483.

474) Lei, H., Wu, C., Liu, H., & Duan, Y. (2007). Folding free-energy landscape of villin headpiece subdomain from molecular dynamics simulations. Proceedings of the National Academy of Sciences of the United States of America, 104(12), 4925–4930. https://doi.org/10.1073/pnas.0608432104.

475) Otzen, D. E., Kristensen, O., Proctor, M., & Oliveberg, M. (1999). Structural changes in the transition state of protein folding: alternative interpretations of curved chevron plots. Biochemistry, 38(20), 6499–6511. https://doi.org/10.1021/bi982819j

476) Pitera, J. W., & van Gunsteren, W. F. (2001). One-step perturbation methods for solvation free energies of polar solutes. Journal of Physical Chemistry B, 105(45), 11264–11274. https://doi.org/10.1021/jp012003j

477) Grossfield, A., Ren, P., & Ponder, J. W. (2003). Ion solvation thermodynamics from simulation with a polarizable force field. Journal of the American Chemical Society, 125(50), 15671–15682. https://doi.org/10.1021/ja037005r

478) Noskov, S. Y., & Lim, C. (2001). Free energy decomposition of protein-protein interactions. Biophysical journal, 81(2), 737–750. https://doi.org/10.1016/S0006-3495(01)75738-4

479) Rietman, E. A., Platig, J., Tuszynski, J. A., & Lakka Klement, G. (2016). Thermodynamic measures of cancer: Gibbs free energy and entropy of protein-protein interactions. Journal of biological physics, 42(3), 339–350. https://doi.org/10.1007/s10867-016-9410-y.

480) Oostenbrink, B. C., Pitera, J. W., van Lipzig MM, Meerman, J. H., & van Gunsteren WF (2000). Simulations of the estrogen receptor ligand-binding domain: affinity of natural ligands and xenoestrogens. Journal of medicinal chemistry, 43(24), 4594–4605. https://doi.org/10.1021/jm001045d

481) Bishop, T. C., & Williams, K. Y. (2002). Molecular dynamics simulations of the estrogen receptor ligandbinding domain. Biophysical Journal, 82(2 Suppl.), 486a

482) Williams, K. Y., & Bishop, T. C. (2003). Free energy simulations of the estrogen receptor ligandbinding domain using MM/PBSA. Biophysical Journal, 84(2 Suppl.), 462a.

483) Fujitani, H., Tanida, Y., Ito, M., Jayachandran, G., Snow, C. D., Shirts, M. R., Sorin, E. J., & Pande, V. S. (2005). Direct calculation of the binding free energies of FKBP ligands. The Journal of chemical physics, 123(8), 084108. https://doi.org/10.1063/1.1999637

484) Ytreberg, F. M., Swendsen, R. H., & Zuckerman, D. M. (2006). Comparison of free energy methods for molecular systems. The Journal of chemical physics, 125(18), 184114. https://doi.org/10.1063/1.2378907.

485) Brandsdal, B. O., Osterberg, F., Almlöf, M., Feierberg, I., Luzhkov, V. B., & Aqvist, J. (2003). Free energy calculations and ligand binding. Advances in protein chemistry, 66, 123–158. https://doi.org/10.1016/s0065-3233(03)66004-3.

486) Meng, Y., Dashti, D. S., & Roitberg, A. E. (2011). Computing Alchemical Free Energy Differences with Hamiltonian Replica Exchange Molecular Dynamics (H-REMD) Simulations. Journal of chemical theory and computation, 7(9), 2721–2727. https://doi.org/10.1021/ct200153u

487) Chodera, J. D., Mobley, D. L., Shirts, M. R., Dixon, R. W., Branson, K., & Pande, V. S. (2011). Alchemical free energy methods for drug discovery: progress and challenges. Current opinion in structural biology, 21(2), 150–160. https://doi.org/10.1016/j.sbi.2011.01.011.

488) Abel, R., Wang, L., Mobley, D. L., & Friesner, R. A. (2017). A Critical Review of Validation, Blind Testing, and Real- World Use of Alchemical Protein-Ligand Binding Free Energy Calculations. Current topics in medicinal chemistry, 17(23), 2577–2585. https://doi.org/10.2174/1568026617666170414142131

489) Srinivasan, J., Miller, J., Kollman, P. A., & Case, D. A. (1998). Continuum solvent studies of the stability of RNA hairpin loops and helices. Journal of biomolecular structure & dynamics, 16(3), 671–682. https://doi.org/10.1080/07391102.1998.10508279

490) Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., & Cheatham, T. E., 3rd (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Accounts of chemical research, 33(12), 889–897. https://doi.org/10.1021/ar000033j. 1203

491) Srinivasan, J., Cheatham, T. E., III, Cieplak, P., Kollman, P. A., & Case, D. A. (1998). Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate–DNA helices. Journal of the American Chemical Society, 120(37), 9401–9409. https://doi.org/10.1021/ja981844+

492) Clark, A. J., Gindin, T., Zhang, B., Wang, L., Abel, R., Murret, C. S., Xu, F., Bao, A., Lu, N. J., Zhou, T., Kwong, P. D., Shapiro, L., Honig, B., & Friesner, R. A. (2017). Free Energy Perturbation Calculation of Relative Binding Free Energy between Broadly Neutralizing Antibodies and the gp120 Glycoprotein of HIV-1. Journal of molecular biology, 429(7), 930–947. https://doi.org/10.1016/j.jmb.2016.11.021.

493) Hirono, S., & Kollman, P. A. (1990). Calculation of the relative binding free energy of 2'GMP and 2'AMP to ribonuclease T1 using molecular dynamics/free energy perturbation approaches. Journal of molecular biology, 212(1), 197–209. https://doi.org/10.1016/0022-2836(90)90315-D

494) Reddy, M. R., & Erion, M. D. (2001). Calculation of relative binding free energy differences for fructose 1,6-bisphosphatase inhibitors using the thermodynamic cycle perturbation approach. Journal of the American Chemical Society, 123(26), 6246–6252. https://doi.org/10.1021/ja0103288

495) Mutyala, R., Reddy, R. N., Sumakanth, M., Reddanna, P., & Reddy, M. R. (2007). Calculation of relative binding affinities of fructose 1,6-bisphosphatase mutants with adenosine monophosphate using free energy perturbation method. Journal of computational chemistry, 28(5), 932–937. https://doi.org/10.1002/jcc.20617.

496) Kamath, S., Coutinho, E., & Desai, P. (1999). Calculation of relative binding free energy difference of DHFR inhibitors by a finite difference thermodynamic integration (FDTI) approach. Journal of biomolecular structure & dynamics, 16(6), 1239–1244. https://doi.org/10.1080/07391102.1999.10508331.

497) Wu, K. W., Chen, P. C., Wang, J., & Sun, Y. C. (2012). Computation of relative binding free energy for an inhibitor and its analogs binding with Erk kinase using thermodynamic integration MD simulation. Journal of computer-aided molecular design, 26(10), 1159–1169. https://doi.org/10.1007/s10822-012-9606-6.

498) Lawrenz, M., Baron, R., Wang, Y., & McCammon, J. A. (2012). Independent-Trajectory Thermodynamic Integration: a practical guide to protein-drug binding free energy calculations using distributed computing. Methods in molecular biology (Clifton, N.J.), 819, 469–486. https://doi.org/10.1007/978-1-61779-465-0_27

499) Wang, C. X., Shi, Y. Y., Zhou, F., & Wang, L. (1993). Thermodynamic integration calculations of binding free energy difference for Gly-169 mutation in subtilisin BPN'. Proteins, 15(1), 5–9. https://doi.org/10.1002/prot.340150103

500) Swanson, J. M., Henchman, R. H., & McCammon, J. A. (2004). Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. Biophysical journal, 86(1 Pt 1), 67–74. https://doi.org/10.1016/S0006-3495(04)74084-9

501) Genheden, S., & Ryde, U. (2012). Comparison of end-point continuum-solvation methods for the calculation of protein-ligand binding free energies. Proteins, 80(5), 1326–1342. https://doi.org/10.1002/prot.24029.

502) Sham, Y. Y., Chu, Z. T., Tao, H., & Warshel, A. (2000). Examining methods for calculations of binding free energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA calculations of ligands binding to an HIV protease. Proteins, 39(4), 393–407..

503) Aqvist, J., Medina, C., & Samuelsson, J. E. (1994). A new method for predicting binding affinity in computer-aided drug design. Protein engineering, 7(3), 385–391. https://doi.org/10.1093/protein/7.3.385

504) Adekoya, O. A., Willassen, N. P., & Sylte, I. (2006). Molecular insight into pseudolysin inhibition using the MM-PBSA and LIE methods. Journal of structural biology, 153(2), 129–144. https://doi.org/10.1016/j.jsb.2005.11.003

505) Hansson, T., Marelius, J., & Aqvist, J. (1998). Ligand binding affinity prediction by linear interaction energy methods. Journal of computer-aided molecular design, 12(1), 27–35. https://doi.org/10.1023/a:1007930623000.

506) Pan, P., Yu, H., Liu, Q., Kong, X., Chen, H., Chen, J., Liu, Q., Li, D., Kang, Y., Sun, H., Zhou, W., Tian, S., Cui, S., Zhu, F., Li, Y., Huang, Y., & Hou, T. (2017). Combating Drug-Resistant Mutants of Anaplastic Lymphoma Kinase with Potent and Selective Type-$I^{1/2}$ Inhibitors by Stabilizing Unique DFG-Shifted Loop Conformation. ACS central science, 3(11), 1208–1220. https://doi.org/10.1021/acscentsci.7b00419

507) Xu, W., Lau, Y. H., Fischer, G., Tan, Y. S., Chattopadhyay, A., de la Roche, M., Hyvönen, M., Verma, C., Spring, D. R., & Itzhaki, L. S. (2017). Macrocyclized Extended Peptides: Inhibiting the Substrate-Recognition Domain of Tankyrase. Journal of the American Chemical Society, 139(6), 2245–2256. https://doi.org/10.1021/jacs.6b10234.

508) Roca, C., Martinez-González, L., Daniel-Mozo, M., Sastre, J., Infantes, L., Mansilla, A., Chaves-Sanjuan, A., González-Rubio, J. M., Gil, C., Cañada, F. J., Martinez, A., Sanchez-Barrena, M. J., & Campillo, N. E. (2018). Deciphering the Inhibition of the Neuronal Calcium Sensor 1 and the Guanine Exchange Factor Ric8a with a Small Phenothiazine Molecule for the Rational Generation of Therapeutic Synapse Function Regulators. Journal of medicinal chemistry, 61(14), 5910–5921. https://doi.org/10.1021/acs.jmedchem.8b00088

509) Jiang, Z. Y., Lu, M. C., Xu, L. L., Yang, T. T., Xi, M. Y., Xu, X. L., Guo, X. K., Zhang, X. J., You, Q. D., & Sun, H. P. (2014). Discovery of potent Keap1-Nrf2 protein-protein interaction inhibitor based on molecular binding determinants analysis. Journal of medicinal chemistry, 57(6), 2736–2745. https://doi.org/10.1021/jm5000529.

510) Guan, Y., Sun, H., Li, Y., Pan, P., Li, D., & Hou, T. (2014). The competitive binding between inhibitors and substrates of HCV NS3/4A protease: a general mechanism of drug resistance. Antiviral research, 103, 60–70. https://doi.org/10.1016/j.antiviral.2014.01.010

511) Guan, Y., Sun, H., Pan, P., Li, Y., Li, D., & Hou, T. (2015). Exploring resistance mechanisms of HCV NS3/4A protease mutations to MK5172: insight from molecular dynamics simulations and free energy calculations. Molecular bioSystems, 11(9), 2568–2578. https://doi.org/10.1039/c5mb00394f.

512) Pan, P., Li, Y., Yu, H., Sun, H., & Hou, T. (2013). Molecular principle of topotecan resistance by topoisomerase I mutations through molecular modeling approaches. Journal of chemical information and modeling, 53(4), 997–1006. https://doi.org/10.1021/ci400066x.

513) Chen, F., Liu, H., Sun, H., Pan, P., Li, Y., Li, D., & Hou, T. (2016). Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. Physical chemistry chemical physics : PCCP, 18(32), 22129–22139. https://doi.org/10.1039/c6cp03670h.

514) Yan, Y., Yang, M., Ji, C. G., & Zhang, J. Z. H. (2017). Interaction Entropy for Computational Alanine Scanning. Journal of chemical information and modeling, 57(5), 1112–1122. https://doi.org/10.1021/acs.jcim.6b00734.

515) Feng, T., Chen, F., Kang, Y., Sun, H., Liu, H., Li, D., Zhu, F., & Hou, T. (2017). HawkRank: a new scoring function for protein-protein docking based on weighted energy terms. Journal of cheminformatics, 9(1), 66. https://doi.org/10.1186/s13321-017-0254-7

516) Shimba, N., Kamiya, N., & Nakamura, H. (2016). Model Building of Antibody-Antigen Complex Structures Using GBSA Scores. Journal of chemical information and modeling, 56(10), 2005–2012. https://doi.org/10.1021/acs.jcim.6b00066.

517) Maffucci, I., & Contini, A. (2016). Improved Computation of Protein-Protein Relative Binding Energies with the Nwat-MMGBSA Method. Journal of chemical information and modeling, 56(9), 1692–1704. https://doi.org/10.1021/acs.jcim.6b00196.

518) Simões, I. C., Costa, I. P., Coimbra, J. T., Ramos, M. J., & Fernandes, P. A. (2017). New Parameters for Higher Accuracy in the Computation of Binding Free Energy Differences upon Alanine Scanning Mutagenesis on Protein-Protein Interfaces. Journal of chemical information and modeling, 57(1), 60–72. https://doi.org/10.1021/acs.jcim.6b00378

519) Li, M., Petukh, M., Alexov, E., & Panchenko, A. R. (2014). Predicting the Impact of Missense Mutations on Protein-Protein Binding Affinity. Journal of chemical theory and computation, 10(4), 1770–1780. https://doi.org/10.1021/ct401022c.

520) Xu, L., Li, Y., Li, D., Xu, P., Tian, S., Sun, H., Liu, H., & Hou, T. (2015). Exploring the binding mechanisms of MIF to CXCR2 using theoretical approaches. Physical chemistry chemical physics : PCCP, 17(5), 3370–3382. https://doi.org/10.1039/c4cp05095a.

521) Xu, L., Li, Y., Sun, H., Li, D., & Hou, T. (2013). Structural basis of the interactions between CXCR4 and CXCL12/SDF-1 revealed by theoretical approaches. Molecular bioSystems, 9(8), 2107–2117. https://doi.org/10.1039/c3mb70120d.

522) Hou, T., Li, N., Li, Y., & Wang, W. (2012). Characterization of domain-peptide interaction interface: prediction of SH3 domain-mediated protein-protein interaction network in yeast by generic structure-based models. Journal of proteome research, 11(5), 2982–2995. https://doi.org/10.1021/pr3000688

523) Salmaso, V., Sturlese, M., Cuzzolin, A., & Moro, S. (2017). Exploring Protein-Peptide Recognition Pathways Using a Supervised Molecular Dynamics Approach. Structure (London, England : 1993), 25(4), 655–662.e2. https://doi.org/10.1016/j.str.2017.02.009.

524) Ivanov, S. M., Huber, R. G., Warwicker, J., & Bond, P. J. (2016). Energetics and Dynamics Across the Bcl-2-Regulated Apoptotic Pathway Reveal Distinct Evolutionary Determinants of Specificity and Affinity. Structure (London, England : 1993), 24(11), 2024–2033. https://doi.org/10.1016/j.str.2016.09.006.

525) Hou, T., Li, Y., & Wang, W. (2011). Prediction of peptides binding to the PKA RIIalpha subunit using a hierarchical strategy. Bioinformatics (Oxford, England), 27(13), 1814–1821. https://doi.org/10.1093/bioinformatics/btr294.

526) Hard, R., Li, N., He, W., Ross, B., Mo, G. C. H., Peng, Q., Stein, R. S. L., Komives, E., Wang, Y., Zhang, J., & Wang, W. (2018). Deciphering and engineering chromodomain-methyllysine peptide recognition. Science advances, 4(11), eaau1447. https://doi.org/10.1126/sciadv.aau1447.

527) Hou, T., Chen, K., McLaughlin, W. A., Lu, B., & Wang, W. (2006). Computational analysis and prediction of the binding motif and protein interacting partners of the Abl SH3 domain. PLoS Computational Biology, 2(1), e1. https://doi.org/10.1371/journal.pcbi.0020001

528) Peng, Y., Sun, L., Jia, Z., Li, L., & Alexov, E. (2018). Predicting protein-DNA binding free energy change upon missense mutations using modified MM/PBSA approach: SAMPDI webserver. Bioinformatics (Oxford, England), 34(5), 779–786. https://doi.org/10.1093/bioinformatics/btx698.

529) Ottaviani, A., Iacovelli, F., Idili, A., Falconi, M., Ricci, F., & Desideri, A. (2018). Engineering a responsive DNA triple helix into an octahedral DNA nanostructure for a reversible opening/closing switching mechanism: a computational and experimental integrated study. Nucleic acids research, 46(19), 9951–9959. https://doi.org/10.1093/nar/gky857

530) Soni, A., Khurana, P., Singh, T., & Jayaram, B. (2017). A DNA intercalation methodology for an efficient prediction of ligand binding pose and energetics. Bioinformatics (Oxford, England), 33(10), 1488–1496. https://doi.org/10.1093/bioinformatics/btx006.

531) Fogolari, F., Corazza, A., & Esposito, G. (2015). Accuracy assessment of the linear Poisson-Boltzmann equation and reparametrization of the OBC generalized Born model for nucleic acids and

nucleic acid-protein complexes. Journal of computational chemistry, 36(9), 585–596. https://doi.org/10.1002/jcc.23832.

532) Zhou, Q., Sun, X., Xia, X., Fan, Z., Luo, Z., Zhao, S., Shakhnovich, E., & Liang, H. (2017). Exploring the Mutational Robustness of Nucleic Acids by Searching Genotype Neighborhoods in Sequence Space. The journal of physical chemistry letters, 8(2), 407–414. https://doi.org/10.1021/acs.jpclett.6b02769.

533) Morrison, E. A., Bowerman, S., Sylvers, K. L., Wereszczynski, J., & Musselman, C. A. (2018). The conformation of the histone H3 tail inhibits association of the BPTF PHD finger with the nucleosome. eLife, 7, e31481. https://doi.org/10.7554/eLife.31481

534) Rajagopalan, M., Balasubramanian, S., & Ramaswamy, A. (2017). Insights into the RNA binding mechanism of human L1-ORF1p: a molecular dynamics study. Molecular bioSystems, 13(9), 1728–1743. https://doi.org/10.1039/c7mb00358g.

535) Orr, A. A., Gonzalez-Rivera, J. C., Wilson, M., Bhikha, P. R., Wang, D., Contreras, L. M., & Tamamis, P. (2018). A high-throughput and rapid computational method for screening of RNA post-transcriptional modifications that can be recognized by target proteins. Methods (San Diego, Calif.), 143, 34–47. https://doi.org/10.1016/j.ymeth.2018.01.015.

536) Zhang, N., Chen, Y., Zhao, F., Yang, Q., Simonetti, F. L., & Li, M. (2018). PremPDI estimates and interprets the effects of missense mutations on protein-DNA interactions. PLoS computational biology, 14(12), e1006615. https://doi.org/10.1371/journal.pcbi.1006615

537) Bowerman, S., & Wereszczynski, J. (2016). Effects of MacroH2A and H2A.Z on Nucleosome Dynamics as Elucidated by Molecular Dynamics Simulations. Biophysical journal, 110(2), 327–337. https://doi.org/10.1016/j.bpj.2015.12.015

538) Beuerle, M. G., Dufton, N. P., Randi, A. M., & Gould, I. R. (2016). Molecular dynamics studies on the DNA-binding process of ERG. Molecular bioSystems, 12(12), 3600–3610. https://doi.org/10.1039/c6mb00506c

539) Xu, L., Kong, R., Zhu, J., Sun, H., & Chang, S. (2016). Unraveling the conformational determinants of LARP7 and 7SK small nuclear RNA by theoretical approaches. Molecular bioSystems, 12(8), 2613–2621. https://doi.org/10.1039/c6mb00252h

540) Sun, H., Li, Y., Shen, M., Tian, S., Xu, L., Pan, P., Guan, Y., & Hou, T. (2014). Assessing the performance of MM/PBSA and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. Physical chemistry chemical physics : PCCP, 16(40), 22035–22045. https://doi.org/10.1039/c4cp03179b

541) Rastelli, G., Del Rio, A., Degliesposti, G., & Sgobba, M. (2010). Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. Journal of computational chemistry, 31(4), 797–810. https://doi.org/10.1002/jcc.21372.

542) Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. Journal of computational chemistry, 32(5), 866–877. https://doi.org/10.1002/jcc.21666.

543) Lindström, A., Edvinsson, L., Johansson, A., Andersson, C. D., Andersson, I. E., Raubacher, F., & Linusson, A. (2011). Postprocessing of docked protein-ligand complexes using implicit solvation models. Journal of chemical information and modeling, 51(2), 267–282. https://doi.org/10.1021/ci100354x

544) Steinbrecher, T., Case, D. A., & Labahn, A. (2006). A multistep approach to structure-based drug design: studying ligand binding at the human neutrophil elastase. Journal of medicinal chemistry, 49(6), 1837–1844. https://doi.org/10.1021/jm0505720

545) Thompson, D. C., Humblet, C., & Joseph-McCarthy, D. (2008). Investigation of MM-PBSA rescoring of docking poses. Journal of chemical information and modeling, 48(5), 1081–1091. https://doi.org/10.1021/ci700470c.

546) Zhang, X., Wong, S. E., & Lightstone, F. C. (2014). Toward fully automated high performance computing drug discovery: a massively parallel virtual screening pipeline for docking and molecular mechanics/generalized Born surface area rescoring to improve enrichment. Journal of chemical information and modeling, 54(1), 324–337. https://doi.org/10.1021/ci4005145.

547) Greenidge, P. A., Kramer, C., Mozziconacci, J. C., & Sherman, W. (2014). Improving docking results via reranking of ensembles of ligand poses in multiple X-ray protein conformations with MM-GBSA. Journal of chemical information and modeling, 54(10), 2697–2717. https://doi.org/10.1021/ci5003735

548) Zhu, T., Lee, H., Lei, H., Jones, C., Patel, K., Johnson, M. E., & Hevener, K. E. (2013). Fragment-based drug discovery using a multidomain, parallel MD-MM/PBSA screening protocol. Journal of chemical information and modeling, 53(3), 560–572. https://doi.org/10.1021/ci300502h

549) Karami, M., Jalali, C., & Mirzaie, S. (2017). Combined virtual screening, MMPBSA, molecular docking and dynamics studies against deadly anthrax: An in silico effort to inhibit Bacillus anthracis nucleoside hydrolase. Journal of theoretical biology, 420, 180–189. https://doi.org/10.1016/j.jtbi.2017.03.010

550) Zoete, V., Irving, M. B., & Michielin, O. (2010). MM-GBSA binding free energy decomposition and T cell receptor engineering. Journal of molecular recognition : JMR, 23(2), 142–152. https://doi.org/10.1002/jmr.1005

551) Zoete, V., & Michielin, O. (2007). Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein-protein association using MM-GBSA: application to the TCR-p-MHC complex. Proteins, 67(4), 1026–1047. https://doi.org/10.1002/prot.21395

552) Gilson, M. K., & Honig, B. (1988). Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. Proteins, 4(1), 7–18. https://doi.org/10.1002/prot.340040104

553) Weis, A., Katebzadeh, K., Söderhjelm, P., Nilsson, I., & Ryde, U. (2006). Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. Journal of medicinal chemistry, 49(22), 6596–6606. https://doi.org/10.1021/jm0608210

554) Su, P. C., Tsai, C. C., Mehboob, S., Hevener, K. E., & Johnson, M. E. (2015). Comparison of radii sets, entropy, QM methods, and sampling on MM-PBSA, MM-GBSA, and QM/MM-GBSA ligand binding energies of F. tularensis enoyl-ACP reductase (FabI). Journal of computational chemistry, 36(25), 1859–1873. https://doi.org/10.1002/jcc.24011.

555) Stoica, I., Sadiq, S. K., & Coveney, P. V. (2008). Rapid and accurate prediction of binding free energies for saquinavir-bound HIV-1 proteases. Journal of the American Chemical Society, 130(8), 2639–2648. https://doi.org/10.1021/ja0779250

556) Genheden, S., & Ryde, U. (2010). How to obtain statistically converged MM/GBSA results. Journal of computational chemistry, 31(4), 837–846. https://doi.org/10.1002/jcc.21366

557) Wang, B., Li, L., Hurley, T. D., & Meroueh, S. O. (2013). Molecular recognition in a diverse set of protein-ligand interactions studied with molecular dynamics simulations and end-point free energy calculations. Journal of chemical information and modeling, 53(10), 2659–2670. https://doi.org/10.1021/ci400312v.

558) Kuhn, B., Gerber, P., Schulz-Gasch, T., & Stahl, M. (2005). Validation and use of the MM-PBSA approach for drug discovery. Journal of medicinal chemistry, 48(12), 4040–4048. https://doi.org/10.1021/jm049081q.

559) Xu, L., Sun, H., Li, Y., Wang, J., & Hou, T. (2013). Assessing the performance of MM/PBSA and MM/GBSA methods. 3. The impact of force fields and ligand charge models. The journal of physical chemistry. B, 117(28), 8408–8421. https://doi.org/10.1021/jp404160y. .

560) Genheden, S., Luchko, T., Gusarov, S., Kovalenko, A., & Ryde, U. (2010). An MM/3D-RISM approach for ligand binding affinities. The journal of physical chemistry. B, 114(25), 8505–8516. https://doi.org/10.1021/jp101461s

561) Corti D. S. (2001). Isothermal-isobaric ensemble for small systems. Physical review. E, Statistical, nonlinear, and soft matter physics, 64(1 Pt 2), 016128. https://doi.org/10.1103/PhysRevE.64.016128

562) Tama, F., & Sanejouand, Y. H. (2001). Conformational change of proteins arising from normal mode calculations. Protein engineering, 14(1), 1–6. https://doi.org/10.1093/protein/14.1.1

563) Brooks, B.R., Janežič, D., & Karplus, M. (1995). Harmonic analysis of large systems. I. Methodology. Journal of Computational Chemistry, 16.

564) Ma J. (2005). Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. Structure (London, England: 1993), 13(3), 373–380. https://doi.org/10.1016/j.str.2005.02.002

565) Karplus, M., & Kushick, J.N. (1981). Method for estimating the configurational entropy of macromolecules. Macromolecules, 14, 325-332.

566) Brooks, B., & Karplus, M. (1983). Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. Proceedings of the National Academy of Sciences of the United States of America, 80(21), 6571–6575. https://doi.org/10.1073/pnas.80.21.6571.

567) Schlitter, J. (1993). Estimation of absolute and relative entropies of macromolecules using the covariance matrix. Chemical Physics Letters, 215(6), 617–621 /doi.org/10.1016/0009-2614(93)89366-P

568) Ben-Shalom, I. Y., Pfeiffer-Marek, S., Baringhaus, K. H., & Gohlke, H. (2017). Efficient Approximation of Ligand Rotational and Translational Entropy Changes upon Binding for Use in MM-PBSA Calculations. Journal of chemical information and modeling, 57(2), 170–189. https://doi.org/10.1021/acs.jcim.6b00373

569) Genheden, S., Akke, M., & Ryde, U. (2014). Conformational Entropies and Order Parameters: Convergence, Reproducibility, and Transferability. Journal of chemical theory and computation, 10(1), 432–438. https://doi.org/10.1021/ct400747s

570) Hikiri, S., Yoshidome, T., & Ikeguchi, M. (2016). Computational Methods for Configurational Entropy Using Internal and Cartesian Coordinates. Journal of chemical theory and computation, 12(12), 5990–6000. https://doi.org/10.1021/acs.jctc.6b00563

571) Sharp K. (2013). Calculation of Molecular Entropies Using Temperature Integration. Journal of chemical theory and computation, 9(2), 1164–1172. https://doi.org/10.1021/ct300901x

572) Choi, H., Kang, H., & Park, H. (2015). Computational prediction of molecular hydration entropy with hybrid scaled particle theory and free-energy perturbation method. Journal of chemical theory and computation, 11(10), 4933–4942. https://doi.org/10.1021/acs.jctc.5b00325

573) Gyimesi, G., Závodszky, P., & Szilágyi, A. (2017). Calculation of Configurational Entropy Differences from Conformational Ensembles Using Gaussian Mixtures. Journal of chemical theory and computation, 13(1), 29–41. https://doi.org/10.1021/acs.jctc.6b00837

574) Homeyer, N., & Gohlke, H. (2012). Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. Molecular informatics, 31(2), 114–122. https://doi.org/10.1002/minf.201100135

575) Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. Nature reviews. Drug discovery, 18(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5.

576) Bychkov, D., Linder, N., Turkki, R., Nordling, S., Kovanen, P. E., Verrill, C., Walliander, M., Lundin, M., Haglund, C., & Lundin, J. (2018). Deep learning based tissue analysis predicts outcome in colorectal cancer. Scientific reports, 8(1), 3395. https://doi.org/10.1038/s41598-018-21758-3

577) Wang L., Ding J., Pan L., Cao D., Jiang H., and Ding X., Artificial intelligence facilitates drug design in the big data era, Chemometrics and Intelligent Laboratory Systems. (2019) 194, article 103850, https://doi.org/10.1016/j.chemolab.2019.103850, 2-s2.0-85072599263.

578) Hessler, G., & Baringhaus, K. H. (2018). Artificial Intelligence in Drug Design. Molecules (Basel, Switzerland), 23(10), 2520. https://doi.org/10.3390/molecules23102520

579) Schneider, P., Walters, W. P., Plowright, A. T., Sieroka, N., Listgarten, J., Goodnow, R. A., Jr, Fisher, J., Jansen, J. M., Duca, J. S., Rush, T. S., Zentgraf, M., Hill, J. E., Krutoholow, E., Kohler, M., Blaney, J., Funatsu, K., Luebkemann, C., & Schneider, G. (2020). Rethinking drug design in the artificial intelligence era. Nature reviews. Drug discovery, 19(5), 353–364. https://doi.org/10.1038/s41573-019-0050-3

580) Paul, D., Sanap, G., Shenoy, S., Kalyane, D., Kalia, K., & Tekade, R. K. (2021). Artificial intelligence in drug discovery and development. Drug Discovery Today, 26(1), 80–93. https://doi.org/10.1016/j.drudis.2020.10.010

581) Patel, L., Shukla, T., Huang, X., Ussery, D. W., & Wang, S. (2020). Machine learning methods in drug discovery. Molecules, 25(22), 5277. https://doi.org/10.3390/molecules25225277

582) Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

583) Sammut, C., & Webb, G. I. (Eds.). (2011). Encyclopedia of machine learning. Springer Science & Business Media.

584) Cortés, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018.

585) https://appinventiv.com/blog/ai-in-drug-discovery/

586) Gütlein, M., & Kramer, S. (2016). Filtered circular fingerprints improve either prediction or runtime performance while retaining interpretability. Journal of cheminformatics, 8, 60. https://doi.org/10.1186/s13321-016-0173-z

587) Jeon, J., Nim, S., Teyra, J., Datti, A., Wrana, J. L., Sidhu, S. S., Moffat, J., & Kim, P. M. (2014). A systematic approach to identify novel cancer drug targets using machine learning, inhibitor design and high-throughput screening. Genome medicine, 6(7), 57. https://doi.org/10.1186/s13073-014-0057-7

588) Ferrero, E., Dunham, I., & Sanseau, P. (2017). In silico prediction of novel therapeutic targets using gene-disease association data. Journal of translational medicine, 15(1), 182. https://doi.org/10.1186/s12967-017-1285-6

589) Riniker, S., Wang, Y., Jenkins, J. L., & Landrum, G. A. (2014). Using information from historical high-throughput screens to predict active compounds. Journal of chemical information and modeling, 54(7), 1880–1891. https://doi.org/10.1021/ci500190p

590) Godinez, W. J., Hossain, I., Lazic, S. E., Davies, J. W., & Zhang, X. (2017). A multi-scale convolutional neural network for phenotyping high-content cellular images. Bioinformatics (Oxford, England), 33(13), 2010–2019. https://doi.org/10.1093/bioinformatics/btx069

591) Olsen, T. G., Jackson, B. H., Feeser, T. A., Kent, M. N., Moad, J. C., Krishnamurthy, S., Lunsford, D. D., & Soans, R. E. (2018). Diagnostic Performance of Deep Learning Algorithms Applied to Three Common Diagnoses in Dermatopathology. Journal of pathology informatics, 9, 32. https://doi.org/10.4103/jpi.jpi_31_18

592) Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 15(56), 1929–1958. https://doi.org/10.5555/2627435.2670313

593) Costa, P. R., Acencio, M. L., & Lemke, N. (2010). A machine learning approach for genome-wide prediction of morbid and druggable human genes based on systems-level data. BMC genomics, 11 Suppl 5(Suppl 5), S9. https://doi.org/10.1186/1471-2164-11-S5-S9

594) Mamoshina, P., Volosnikova, M., Ozerov, I. V., Putin, E., Skibina, E., Cortese, F., & Zhavoronkov, A. (2018). Machine Learning on Human Muscle Transcriptomic Data for Biomarker

Discovery and Tissue-Specific Drug Target Identification. Frontiers in genetics, 9, 242. https://doi.org/10.3389/fgene.2018.00242.

595) Leung, M. K., Xiong, H. Y., Lee, L. J., & Frey, B. J. (2014). Deep learning of the tissue-regulated splicing code. Bioinformatics (Oxford, England), 30(12), i121–i129. https://doi.org/10.1093/bioinformatics/btu277

596) Iorio, F., Knijnenburg, T. A., Vis, D. J., Bignell, G. R., Menden, M. P., Schubert, M., ... & Garnett, M. J. (2016). A landscape of pharmacogenomic interactions in cancer. Cell, 166(3), 740-754

597) Lengauer, T. (2007). Bioinformatics—from genomes to therapies. Bioinformatics-From Genomes to Therapies, 1-24.

598) LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

599) Rifaioglu, A. S., Atas, H., Martin, M. J., Cetin-Atalay, R., Atalay, V., & Doğan, T. (2019). Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. Briefings in bioinformatics, 20(5), 1878–1912. https://doi.org/10.1093/bib/bby061

600) Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., & Blaschke, T. (2018). The rise of deep learning in drug discovery. Drug discovery today, 23(6), 1241–1250. https://doi.org/10.1016/j.drudis.2018.01.039

601) Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. Science (New York, N.Y.), 313(5786), 504–507. https://doi.org/10.1126/science.1127647

602) Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E., & Svetnik, V. (2015). Deep neural nets as a method for quantitative structure-activity relationships. Journal of chemical information and modeling, 55(2), 263–274. https://doi.org/10.1021/ci500747n.

603) Olivecrona, M., Blaschke, T., Engkvist, O., & Chen, H. (2017). Molecular de-novo design through deep reinforcement learning. Journal of cheminformatics, 9(1), 48. https://doi.org/10.1186/s13321-017-0235-x

604) Ramsundar, B., Kearnes, S., Riley, P., Webster, D., Konerding, D., & Pande, V. (2015). Massively Multitask Networks for Drug Discovery. arXiv:1502.02072.

605) Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM (2018) Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. PLOS ONE 13(8): e0202344. https://doi.org/10.1371/journal.pone.0202344.

606) Koscielny, S. (2010). Why most gene expression signatures of tumors have not been useful in the clinic. Science Translational Medicine, 2(14), 14ps2. https://doi.org/10.1126/scitranslmed.3000313

607) Odell, S. G., Lazo, G. R., Woodhouse, M. R., Hane, D. L., & Sen, T. Z. (2017). The art of curation at a biological database: principles and application. Current Plant Biology, 11, 2-11..

608) Behrends, C., Sowa, M. E., Gygi, S. P., & Harper, J. W. (2010). Network organization of the human autophagy system. Nature, 466(7302), 68–76. https://doi.org/10.1038/nature09204

609) Xie, Z., & Klionsky, D. J. (2007). Autophagosome formation: core machinery and adaptations. Nature cell biology, 9(10), 1102–1109. https://doi.org/10.1038/ncb1007-1102.

610) Rabinowitz, J. D., & White, E. (2010). Autophagy and metabolism. Science (New York, N.Y.), 330(6009), 1344–1348. https://doi.org/10.1126/science.1193497.

611) Marx J. (2006). Autophagy: is it cancer's friend or foe?. Science (New York, N.Y.), 312(5777), 1160–1161. https://doi.org/10.1126/science.312.5777.1160

612) Sugawara, K., Suzuki, N. N., Fujioka, Y., Mizushima, N., Ohsumi, Y., & Inagaki, F. (2004). The crystal structure of microtubule-associated protein light chain 3, a mammalian homologue of Saccharomyces cerevisiae Atg8. Genes to cells : devoted to molecular & cellular mechanisms, 9(7), 611–618. https://doi.org/10.1111/j.1356-9597.2004.00750.x

613) Zhou, M., & Wang, R. (2013). Small-molecule regulators of autophagy and their potential therapeutic applications. ChemMedChem, 8(5), 694–707. https://doi.org/10.1002/cmdc.201200560

614) Akin, D., Wang, S. K., Habibzadegah-Tari, P., Law, B., Ostrov, D., Li, M., Yin, X. M., Kim, J. S., Horenstein, N., & Dunn, W. A., Jr (2014). A novel ATG4B antagonist inhibits autophagy and has a negative impact on osteosarcoma tumors. Autophagy, 10(11), 2021–2035. https://doi.org/10.4161/auto.32229

615) Fu, Y., Hong, L., Xu, J., Zhong, G., Gu, Q., Gu, Q., Guan, Y., Zheng, X., Dai, Q., Luo, X., Liu, C., Huang, Z., Yin, X. M., Liu, P., & Li, M. (2019). Discovery of a small molecule targeting autophagy via ATG4B inhibition and cell death of colorectal cancer cells in vitro and in vivo. Autophagy, 15(2), 295–311. https://doi.org/10.1080/15548627.2018.1517073.

616) Bosc, D., Vezenkov, L., Bortnik, S., An, J., Xu, J., Choutka, C., Hannigan, A. M., Kovacic, S., Loo, S., Clark, P. G. K., Chen, G., Guay-Ross, R. N., Yang, K., Dragowska, W. H., Zhang, F., Go, N. E., Leung, A., Honson, N. S., Pfeifer, T. A., Gleave, M., … Young, R. N. (2018). A new quinoline-based chemical probe inhibits the autophagy-related cysteine protease ATG4B. Scientific reports, 8(1), 11653. https://doi.org/10.1038/s41598-018-29900-x

617) Wild, P., McEwan, D. G., & Dikic, I. (2014). The LC3 interactome at a glance. Journal of cell science, 127(Pt 1), 3–9. https://doi.org/10.1242/jcs.140426

618) Amaravadi, R. K., Lippincott-Schwartz, J., Yin, X. M., Weiss, W. A., Takebe, N., Timmer, W., DiPaola, R. S., Lotze, M. T., & White, E. (2011). Principles and current strategies for targeting autophagy for cancer treatment. Clinical cancer research : an official journal of the American Association for Cancer Research, 17(4), 654–666. https://doi.org/10.1158/1078-0432.CCR-10-2634

619) Schaaf, M. B., Keulers, T. G., Vooijs, M. A., & Rouschop, K. M. (2016). LC3/GABARAP family proteins: autophagy-(un)related functions. FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 30(12), 3961–3978. https://doi.org/10.1096/fj.201600698R

620) M. Lubas, L.M. Harder, C. Kumsta, I. Tiessen, M. Hansen, J.S. Andersen, A. H. Lund, L.B. Frankel, eIF5A is required for autophagy by mediating ATG3 translation, EMBO Rep. 19 (2018) e46072, https://doi.org/10.15252/ embr.201846072.

621) Hu, G., McQuiston, T., Bernard, A., Park, Y. D., Qiu, J., Vural, A., Zhang, N., Waterman, S. R., Blewett, N. H., Myers, T. G., Maraia, R. J., Kehrl, J. H., Uzel, G., Klionsky, D. J., & Williamson, P. R. (2015). A conserved mechanism of TOR-dependent RCK-mediated mRNA degradation regulates autophagy. Nature cell biology, 17(7), 930–942. https://doi.org/10.1038/ncb3189

622) Gatica, D., Hu, G., Liu, X., Zhang, N., Williamson, P. R., & Klionsky, D. J. (2019). The Pat1-Lsm Complex Stabilizes ATG mRNA during Nitrogen Starvation-Induced Autophagy. Molecular cell, 73(2), 314–324.e4. https://doi.org/10.1016/j.molcel.2018.11.002

623) Tang, H. W., Hu, Y., Chen, C. L., Xia, B., Zirin, J., Yuan, M., Asara, J. M., Rabinow, L., & Perrimon, N. (2018). The TORC1-Regulated CPA Complex Rewires an RNA Processing Network to Drive Autophagy and Metabolic Reprogramming. Cell metabolism, 27(5), 1040–1054.e8. https://doi.org/10.1016/j.cmet.2018.02.023

624) Wong, Y. C., & Holzbaur, E. L. (2014). Optineurin is an autophagy receptor for damaged mitochondria in parkin-mediated mitophagy that is disrupted by an ALS-linked mutation. Proceedings of the National Academy of Sciences of the United States of America, 111(42), E4439–E4448. https://doi.org/10.1073/pnas.1405752111.

625) McLendon, P. M., Ferguson, B. S., Osinska, H., Bhuiyan, M. S., James, J., McKinsey, T. A., & Robbins, J. (2014). Tubulin hyperacetylation is adaptive in cardiac proteotoxicity by promoting autophagy. Proceedings of the National Academy of Sciences of the United States of America, 111(48), E5178–E5186. https://doi.org/10.1073/pnas.1415589111.

626) Chen, M., Hong, M. J., Sun, H., Wang, L., Shi, X., Gilbert, B. E., Corry, D. B., Kheradmand, F., & Wang, J. (2014). Essential role for autophagy in the maintenance of immunological memory against influenza infection. Nature medicine, 20(5), 503–510. https://doi.org/10.1038/nm.3521

627) Quan, W., Lim, Y. M., & Lee, M. S. (2012). Role of autophagy in diabetes and endoplasmic reticulum stress of pancreatic β-cells. Experimental & molecular medicine, 44(2), 81–88. https://doi.org/10.3858/emm.2012.44.2.030

628) Sarparanta, J., García-Macia, M., & Singh, R. (2017). Autophagy and Mitochondria in Obesity and Type 2 Diabetes. Current diabetes reviews, 13(4), 352–369. https://doi.org/10.2174/1573399812666160217122530.

629) Mao, Y., Yu, F., Wang, J., Guo, C., & Fan, X. (2016). Autophagy: a new target for nonalcoholic fatty liver disease therapy. Hepatic medicine : evidence and research, 8, 27–37. https://doi.org/10.2147/HMER.S98120.

630) Giannopoulos, S., Bozkus, C. C., Zografos, E., Athanasiou, A., Bongiovanni, A. M., Doulaveris, G., Bakoyiannis, C. N., Theodoropoulos, G. E., Zografos, G. C., Witkin, S. S., & Orfanelli, T. (2022). Targeting Both Autophagy and Immunotherapy in Breast Cancer Treatment. Metabolites, 12(10), 966. https://doi.org/10.3390/metabo12100966.

631) Pérez-Hernández, M., Arias, A., Martínez-García, D., Pérez-Tomás, R., Quesada, R., & Soto-Cerrato, V. (2019). Targeting Autophagy for Cancer Treatment and Tumor Chemosensitization. Cancers, 11(10), 1599. https://doi.org/10.3390/cancers11101599.

632) Yu, G., & Klionsky, D. J. (2022). Life and Death Decisions-The Many Faces of Autophagy in Cell Survival and Cell Death. Biomolecules, 12(7), 866. https://doi.org/10.3390/biom12070866

633) MJacquet, M., Guittaut, M., Fraichard, A., & Despouy, G. (2021). The functions of Atg8-family proteins in autophagy and cancer: linked or unrelated?. Autophagy, 17(3), 599–611. https://doi.org/10.1080/15548627.2020.1749367

634) Huang, R., Xu, Y., Wan, W., Shou, X., Qian, J., You, Z., Liu, B., Chang, C., Zhou, T., Lippincott-Schwartz, J., & Liu, W. (2015). Deacetylation of nuclear LC3 drives autophagy initiation under starvation. Molecular cell, 57(3), 456–466. https://doi.org/10.1016/j.molcel.2014.12.013.

635) Nijholt, D. A., de Graaf, T. R., van Haastert, E. S., Oliveira, A. O., Berkers, C. R., Zwart, R., Ovaa, H., Baas, F., Hoozemans, J. J., & Scheper, W. (2011). Endoplasmic reticulum stress activates autophagy but not the proteasome in neuronal cells: implications for Alzheimer's disease. Cell death and differentiation, 18(6), 1071–1081. https://doi.org/10.1038/cdd.2010.176

636) Xie, M., Liu, J., Wang, Z., Sun, B., & Wang, J. (2020). Inhibitory effects of 5-heptadecylresorcinol on the proliferation of human MCF-7 breast cancer cells through modulating PI3K/Akt/mTOR pathway. Journal of Functional Foods, 69, 103946. https://doi.org/10.1016/j.jff.2020.103946.

637) Fu, R., Deng, Q., Zhang, H., Hu, X., Li, Y., Liu, Y., Hu, J., Luo, Q., Zhang, Y., Jiang, X., Li, L., Yang, C., & Gao, N. (2018). A novel autophagy inhibitor berbamine blocks SNARE-mediated autophagosome-lysosome fusion through upregulation of BNIP3. Cell death & disease, 9(2), 243. https://doi.org/10.1038/s41419-018-0276-8

638) Tang, J., Zhu, J., Ye, Y., Liu, Y., He, Y., Zhang, L., Tang, D., Qiao, C., Feng, X., Li, J., Kan, Y., Li, X., Jin, X., & Kong, D. (2019). Inhibition LC3B can increase chemosensitivity of ovarian cancer cells. Cancer cell international, 19, 199. https://doi.org/10.1186/s12935-019-0921-z

639) Wu, S., Sun, C., Tian, D., Li, Y., Gao, X., He, S., & Li, T. (2015). Expression and clinical significances of Beclin1, LC3 and mTOR in colorectal cancer. International journal of clinical and experimental pathology, 8(4), 3882–3891.

640) El-Mashed, S., O'Donovan, T. R., Kay, E. W., Abdallah, A. R., Cathcart, M. C., O'Sullivan, J., O'Grady, A., Reynolds, J., O'Reilly, S., O'Sullivan, G. C., & McKenna, S. L. (2015). LC3B globular structures correlate with survival in esophageal adenocarcinoma. BMC cancer, 15, 582. https://doi.org/10.1186/s12885-015-1574-5

641) Mortezavi, A., Salemi, S., Rupp, N. J., Rüschoff, J. H., Hermanns, T., Poyet, C., Randazzo, M., Simon, H. U., Moch, H., Sulser, T., Wild, P., & Eberli, D. (2017). Negative LC3b immunoreactivity in cancer cells is an independent prognostic predictor of prostate cancer specific death. Oncotarget, 8(19), 31765–31774. https://doi.org/10.18632/oncotarget.15986

642) Lai, K., Matthews, S., Wilmott, J. S., Killingsworth, M. C., Yong, J. L., Caixeiro, N. J., Wykes, J., Samakeh, A., Forstner, D., Lee, M., McGuinness, J., Niles, N., Hong, A., Ebrahimi, A., & Lee, C. S. (2018). Differences in LC3B expression and prognostic implications in oropharyngeal and oral cavity squamous cell carcinoma patients. BMC cancer, 18(1), 624. https://doi.org/10.1186/s12885-018-4536-x.

643) Holah, N. S., El-Dien, M. M. S., & Mahmoud, S. F. (2022). Expression of Autophagy Markers Beclin1 and LC3B in Prostatic Carcinoma: An Immunohistochemical Case-Control Study. Iranian journal of pathology, 17(1), 75–84. https://doi.org/10.30699/IJP.2021.530887.2649.

644) Mohammed, S. M., Elesawy, Y. F., Abd El Aziz, A. M., & Khairy, R. A. (2022). The Pathological Evaluation of Autophagy-Related Protein (LC3B) and Its Association with the Infiltration of Immune Cells in Glioma. Asian Pacific journal of cancer prevention : APJCP, 23(5), 1777–1784. https://doi.org/10.31557/APJCP.2022.23.5.1777.

645) Danish, F., Qureshi, M. A., Mirza, T., Amin, W., Sufiyan, S., Naeem, S., Arshad, F., & Mughal, N. (2024). Investigating the Association between the Autophagy Markers LC3B, SQSTM1/p62, and DRAM and Autophagy-Related Genes in Glioma. International journal of molecular sciences, 25(1), 572. https://doi.org/10.3390/ijms25010572.

646) Jiang, T., & Wu, Z. (2018). Immunohistochemical assessment of autophagic protein LC3B and p62 levels in glioma patients. International journal of clinical and experimental pathology, 11(2), 862–868.

647) Cj, P., Hv, E., Vijayakurup, V., R Menon, G., Nair, S., & Gopala, S. (2019). High LC3/Beclin Expression Correlates with Poor Survival in Glioma: a Definitive Role for Autophagy as Evidenced

by In Vitro Autophagic Flux. Pathology oncology research : POR, 25(1), 137–148. https://doi.org/10.1007/s12253-017-0310-7.

648) Liu, J. L., Chen, F. F., Lung, J., Lo, C. H., Lee, F. H., Lu, Y. C., & Hung, C. H. (2014). Prognostic significance of p62/SQSTM1 subcellular localization and LC3B in oral squamous cell carcinoma. British journal of cancer, 111(5), 944–954. https://doi.org/10.1038/bjc.2014.355.

649) Ladoire, S., Enot, D., Senovilla, L., Ghiringhelli, F., Poirier-Colame, V., Chaba, K., Semeraro, M., Chaix, M., Penault-Llorca, F., Arnould, L., Poillot, M. L., Arveux, P., Delaloge, S., Andre, F., Zitvogel, L., & Kroemer, G. (2016). The presence of LC3B puncta and HMGB1 expression in malignant cells correlate with the immune infiltrate in breast cancer. Autophagy, 12(5), 864–875. https://doi.org/10.1080/15548627.2016.1154244

650) Kouno, T., Mizuguchi, M., Tanida, I., Ueno, T., Kanematsu, T., Mori, Y., Shinoda, H., Hirata, M., Kominami, E., & Kawano, K. (2005). Solution structure of microtubule-associated protein light chain 3 and identification of its functional subdomains. The Journal of Biological Chemistry, 280(26), 24610–24617. https://doi.org/10.1074/jbc.M413565200

651) Koukourakis, M. I., Kalamida, D., Giatromanolaki, A., Zois, C. E., Sivridis, E., Pouliliou, S., Mitrakas, A., Gatter, K. C., & Harris, A. L. (2015). Autophagosome Proteins LC3A, LC3B and LC3C Have Distinct Subcellular Distribution Kinetics and Expression in Cancer Cell Lines. PloS one, 10(9), e0137675. https://doi.org/10.1371/journal.pone.0137675.

652) Lv, M., Wang, C., Li, F., Peng, J., Wen, B., Gong, Q., Shi, Y., & Tang, Y. (2017). Structural insights into the recognition of phosphorylated FUNDC1 by LC3B in mitophagy. Protein & cell, 8(1), 25–38. https://doi.org/10.1007/s13238-016-0328-8.

653) Maruyama, Y., Sou, Y. S., Kageyama, S., Takahashi, T., Ueno, T., Tanaka, K., Komatsu, M., & Ichimura, Y. (2014). LC3B is indispensable for selective autophagy of p62 but not basal autophagy. Biochemical and biophysical research communications, 446(1), 309–315. https://doi.org/10.1016/j.bbrc.2014.02.093

654) Ma, D., Panda, S., & Lin, J. D. (2011). Temporal orchestration of circadian autophagy rhythm by C/EBPβ. The EMBO journal, 30(22), 4642–4651. https://doi.org/10.1038/emboj.2011.322

655) Polager, S., Ofir, M., & Ginsberg, D. (2008). E2F1 regulates autophagy and the transcription of autophagy genes. Oncogene, 27(35), 4860–4864. https://doi.org/10.1038/onc.2008.117

656) Li, D. D., Wang, L. L., Deng, R., Tang, J., Shen, Y., Guo, J. F., Wang, Y., Xia, L. P., Feng, G. K., Liu, Q. Q., Huang, W. L., Zeng, Y. X., & Zhu, X. F. (2009). The pivotal role of c-Jun NH2-terminal kinase-mediated Beclin 1 expression during anticancer agents-induced autophagy in cancer cells. Oncogene, 28(6), 886–898. https://doi.org/10.1038/onc.2008.441

657) Sun, T., Li, D., Wang, L., Xia, L., Ma, J., Guan, Z., Feng, G., & Zhu, X. (2011). c-Jun NH2-terminal kinase activation is essential for up-regulation of LC3 during ceramide-induced autophagy

in human nasopharyngeal carcinoma cells. Journal of translational medicine, 9, 161. https://doi.org/10.1186/1479-5876-9-161

658) Seo, Y. K., Jeon, T. I., Chong, H. K., Biesinger, J., Xie, X., & Osborne, T. F. (2011). Genome-wide localization of SREBP-2 in hepatic chromatin predicts a role in autophagy. Cell metabolism, 13(4), 367–375. https://doi.org/10.1016/j.cmet.2011.03.005

659) JNieto-Torres, J. L., Encalada, S. E., & Hansen, M. (2021). LC3B phosphorylation: autophagosome's ticket for a ride toward the cell nucleus. Autophagy, 17(10), 3266–3268. https://doi.org/10.1080/15548627.2021.1961073

660) Shrestha, B. K., Skytte Rasmussen, M., Abudu, Y. P., Bruun, J. A., Larsen, K. B., Alemu, E. A., Sjøttem, E., Lamark, T., & Johansen, T. (2020). NIMA-related kinase 9-mediated phosphorylation of the microtubule-associated LC3B protein at Thr-50 suppresses selective autophagy of p62/sequestosome 1. The Journal of biological chemistry, 295(5), 1240–1260. https://doi.org/10.1074/jbc.RA119.010068

661) Pankiv, S., Alemu, E. A., Brech, A., Bruun, J. A., Lamark, T., Overvatn, A., Bjørkøy, G., & Johansen, T. (2010). FYCO1 is a Rab7 effector that binds to LC3 and PI3P to mediate microtubule plus end-directed vesicle transport. The Journal of cell biology, 188(2), 253–269. https://doi.org/10.1083/jcb.200907015.

662) Olsvik, H. L., Lamark, T., Takagi, K., Larsen, K. B., Evjen, G., Øvervatn, A., Mizushima, T., & Johansen, T. (2015). FYCO1 Contains a C-terminally Extended, LC3A/B-preferring LC3-interacting Region (LIR) Motif Required for Efficient Maturation of Autophagosomes during Basal Autophagy. The Journal of biological chemistry, 290(49), 29361–29374. https://doi.org/10.1074/jbc.M115.686915.

663) Wilkinson, D. S., Jariwala, J. S., Anderson, E., Mitra, K., Meisenhelder, J., Chang, J. T., Ideker, T., Hunter, T., Nizet, V., Dillin, A., & Hansen, M. (2015). Phosphorylation of LC3 by the Hippo kinases STK3/STK4 is essential for autophagy. Molecular cell, 57(1), 55–68. https://doi.org/10.1016/j.molcel.2014.11.019.

664) Sakurai, S., Tomita, T., Shimizu, T., & Ohto, U. (2017). The crystal structure of mouse LC3B in complex with the FYCO1 LIR reveals the importance of the flanking region of the LIR motif. Acta crystallographica. Section F, Structural biology communications, 73(Pt 3), 130–137. https://doi.org/10.1107/S2053230X17001911.

665) Jiang, H., Cheng, D., Liu, W., Peng, J., & Feng, J. (2010). Protein kinase C inhibits autophagy and phosphorylates LC3. Biochemical and biophysical research communications, 395(4), 471–476. https://doi.org/10.1016/j.bbrc.2010.04.030.

666) Cherra, S. J., 3rd, Kulich, S. M., Uechi, G., Balasubramani, M., Mountzouris, J., Day, B. W., & Chu, C. T. (2010). Regulation of the autophagy protein LC3 by phosphorylation. The Journal of cell biology, 190(4), 533–539. https://doi.org/10.1083/jcb.201002108

667) Shabb J. B. (2001). Physiological substrates of cAMP-dependent protein kinase. Chemical reviews, 101(8), 2381–2411. https://doi.org/10.1021/cr0002361

668) Lee, I. H., & Finkel, T. (2009). Regulation of autophagy by the p300 acetyltransferase. The Journal of biological chemistry, 284(10), 6322–6328. https://doi.org/10.1074/jbc.M807135200.

669) Zhou, B., & Rabinovitch, M. (1998). Microtubule involvement in translational regulation of fibronectin expression by light chain 3 of microtubule-associated protein 1 in vascular smooth muscle cells. Circulation Research, *83*(5), 481–489. https://doi.org/10.1161/01.res.83.5.481

670) Ying, L., Lau, A., Alvira, C. M., West, R., Cann, G. M., Zhou, B., Kinnear, C., Jan, E., Sarnow, P., Van de Rijn, M., & Rabinovitch, M. (2009). LC3-mediated fibronectin mRNA translation induces fibrosarcoma growth by increasing connective tissue growth factor. Journal of cell science, 122(Pt 9), 1441–1451. https://doi.org/10.1242/jcs.025957

671) Corley, M., Burns, M. C., & Yeo, G. W. (2020). How RNA-binding proteins interact with RNA: molecules and mechanisms. Molecular Cell, *78*(1), 9–29. https://doi.org/10.1016/j.molcel.2020.03.011

672) Bayer, T. S., Booth, L. N., Knudsen, S. M., & Ellington, A. D. (2005). Arginine-rich motifs present multiple interfaces for specific binding by RNA. RNA, *11*(12), 1848–1857. https://doi.org/10.1261/rna.2167605

673) Thandapani, P., O'Connor, T. R., Bailey, T. L., & Richard, S. (2013). Defining the RGG/RG motif. Molecular Cell, *50*(5), 613–623. https://doi.org/10.1016/j.molcel.2013.05.021

674) Fang, S., Zhang, L., Liu, Y., Xu, W., Wu, W., Huang, Z., Wang, X., Liu, H., Sun, Y., Zhang, R., Peng, B., Liu, X., Sun, X., Yu, J., Chan, F. K. L., Ng, S. C., Wong, S. H., Wang, M. H. T., Gin, T., Joynt, G. M., … Xia, J. (2021). Lysosome activation in peripheral blood mononuclear cells and prognostic significance of circulating LC3B in COVID-19. Briefings in bioinformatics, 22(2), 1466–1475. https://doi.org/10.1093/bib/bbab043

675) Hwang, H. J., Ha, H., Lee, B. S., Kim, B. H., Song, H. K., & Kim, Y. K. (2022). LC3B is an RNA-binding protein to trigger rapid mRNA degradation during autophagy. Nature communications, 13(1), 1436. https://doi.org/10.1038/s41467-022-29139-1

676) Hwang, H. J., & Kim, Y. K. (2022). The role of LC3B in autophagy as an RNA-binding protein. Autophagy, 19(3), 1028–1030. https://doi.org/10.1080/15548627.2022.2111083

677) Zhou, B., Boudreau, N., Coulber, C., Hammarback, J., & Rabinovitch, M. (1997). Microtubule-associated protein 1 light chain 3 is a fibronectin mRNA-binding protein linked to mRNA translation in lamb vascular smooth muscle cells. Journal of Clinical Investigation, *100*(12), 3070–3082. https://doi.org/10.1172/JCI119861

678) Mark, P., & Nilsson, L. (2001). Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. The Journal of Physical Chemistry A, 105(43), 9954–9960. https://doi.org/10.1021/jp003020w

679) Loncharich, R. J., Brooks, B. R., & Pastor, R. W. (1992). Langevin dynamics of peptides: the frictional dependence of isomerization rates of N-acetylalanyl-N'-methylamide. Biopolymers, 32(5), 523–535. https://doi.org/10.1002/bip.360320508

680) Åqvist, J., Wennerström, P., Nervall, M., Bjelic, S., & Brandsdal, B. O. (2004). Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. Chemical Physics Letters, *384*(4–6), 288–294. https://doi.org/10.1016/j.cplett.2003.12.039

681) Tian, C., Kasavajhala, K., Belfon, K. A. A., Raguette, L., Huang, H., Migues, A. N., Bickel, J., Wang, Y., Pincay, J., Wu, Q., & Simmerling, C. (2020). ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. Journal of chemical theory and computation, 16(1), 528–552. https://doi.org/10.1021/acs.jctc.9b00591

682) Zgarbová, M., Otyepka, M., Sponer, J., Mládek, A., Banáš, P., Cheatham, T. E., 3rd, & Jurečka, P. (2011). Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. Journal of chemical theory and computation, 7(9), 2886–2902. https://doi.org/10.1021/ct200162x

683) Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. Journal of Chemical Theory and Computation, *9*(7), 3084–3095. https://doi.org/10.1021/ct400341p

684) Pedretti, A., Mazzolari, A., Gervasoni, S., Fumagalli, L., & Vistoli, G. (2021). The VEGA suite of programs: an versatile platform for cheminformatics and drug design projects. Bioinformatics (Oxford, England), 37(8), 1174–1175. https://doi.org/10.1093/bioinformatics/btaa774

685) Huang, L., & Zhang, C. (2021). Microscale Thermophoresis (MST) to Detect the Interaction Between Purified Protein and Small Molecule. Methods in molecular biology (Clifton, N.J.), 2213, 187–193. https://doi.org/10.1007/978-1-0716-0954-5_17

686) Goodwin, G. H., Sanders, C., & Johns, E. W. (1973). A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. European journal of biochemistry, 38(1), 14–19. https://doi.org/10.1111/j.1432-1033.1973.tb03026.x.

687) Wang, S., & Zhang, Y. (2020). HMGB1 in inflammation and cancer. Journal of hematology & oncology, 13(1), 116. https://doi.org/10.1186/s13045-020-00950-x

688) Thomas, J. O., & Travers, A. A. (2001). HMG1 and 2, and related 'architectural' DNA-binding proteins. Trends in biochemical sciences, 26(3), 167–174. https://doi.org/10.1016/s0968-0004(01)01801-1.

689) Ivics, Z., Kaufman, C. D., Zayed, H., Miskey, C., Walisko, O., & Izsvák, Z. (2004). The Sleeping Beauty transposable element: evolution, regulation and genetic applications. Current issues in molecular biology, 6(1), 43–55.

690) Malarkey, C. S., & Churchill, M. E. (2012). The high mobility group box: the ultimate utility

player of a cell. Trends in biochemical sciences, 37(12), 553–562. https://doi.org/10.1016/j.tibs.2012.09.003

691) Stros M. (2010). HMGB proteins: interactions with DNA and chromatin. Biochimica et biophysica acta, 1799(1-2), 101–113. https://doi.org/10.1016/j.bbagrm.2009.09.008

692) Kang, R., Chen, R., Zhang, Q., Hou, W., Wu, S., Cao, L., Huang, J., Yu, Y., Fan, X. G., Yan, Z., Sun, X., Wang, H., Wang, Q., Tsung, A., Billiar, T. R., Zeh, H. J., 3rd, Lotze, M. T., & Tang, D. (2014). HMGB1 in health and disease. Molecular aspects of medicine, 40, 1–116. https://doi.org/10.1016/j.mam.2014.05.001

693) Ren, W., Zhao, L., Sun, Y., Wang, X., & Shi, X. (2023). HMGB1 and Toll-like receptors: potential therapeutic targets in autoimmune diseases. Molecular medicine (Cambridge, Mass.), 29(1), 117. https://doi.org/10.1186/s10020-023-00717-3

694) Bianchi, M. E., Falciola, L., Ferrari, S., & Lilley, D. M. (1992). The DNA binding site of HMG1 protein is composed of two similar segments (HMG boxes), both of which have counterparts in other eukaryotic regulatory proteins. The EMBO journal, 11(3), 1055–1063. https://doi.org/10.1002/j.1460-2075.1992.tb05144.x.

695) Hardman, C. H., Broadhurst, R. W., Raine, A. R., Grasser, K. D., Thomas, J. O., & Laue, E. D. (1995). Structure of the A-domain of HMG1 and its interaction with DNA as studied by heteronuclear three- and four-dimensional NMR spectroscopy. Biochemistry, 34(51), 16596–16607. https://doi.org/10.1021/bi00051a007

696) Ohndorf, U. M., Rould, M. A., He, Q., Pabo, C. O., & Lippard, S. J. (1999). Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. Nature, 399(6737), 708–712. https://doi.org/10.1038/21460

697) Stott, K., Tang, G. S., Lee, K. B., & Thomas, J. O. (2006). Structure of a complex of tandem HMG boxes and DNA. Journal of molecular biology, 360(1), 90–104. https://doi.org/10.1016/j.jmb.2006.04.059

698) Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R., Laue, E. D., & Thomas, J. O. (1993). Structure of the HMG box motif in the B-domain of HMG1. The EMBO journal, 12(4), 1311–1319. https://doi.org/10.1002/j.1460-2075.1993.tb05776.x

699) Abdul-Razzak, K. K., Denton, M. L., Cox, D. J., & Reeck, G. R. (1989). Isolation and characterization of folded fragments released by Staphylococcal aureus proteinase from the non-histone chromosomal protein HMG-1. Biochimica et biophysica acta, 996(1-2), 125–131. https://doi.org/10.1016/0167-4838(89)90104-0

700) Webb, M., & Thomas, J. O. (1999). Structure-specific binding of the two tandem HMG boxes of HMG1 to four-way junction DNA is mediated by the A domain. Journal of molecular biology, 294(2), 373–387. https://doi.org/10.1006/jmbi.1999.3150.

701) Naglova, H., & Bucova, M. (2012). HMGB1 and its physiological and pathological roles. Bratislavske lekarske listy, 113(3), 163–171. https://doi.org/10.4149/bll_2012_039

702) He, S. J., Cheng, J., Feng, X., Yu, Y., Tian, L., & Huang, Q. (2017). The dual role and therapeutic potential of high-mobility group box 1 in cancer. Oncotarget, 8(38), 64534–64550. https://doi.org/10.18632/oncotarget.17885

703) Tang, D., Kang, R., Livesey, K. M., Cheh, C. W., Farkas, A., Loughran, P., Hoppe, G., Bianchi, M. E., Tracey, K. J., Zeh, H. J., 3rd, & Lotze, M. T. (2010). Endogenous HMGB1 regulates autophagy. The Journal of cell biology, 190(5), 881–892. https://doi.org/10.1083/jcb.200911078

704) Huebener, P., Gwak, G. Y., Pradere, J. P., Quinzii, C. M., Friedman, R., Lin, C. S., Trent, C. M., Mederacke, I., Zhao, E., Dapito, D. H., Lin, Y., Goldberg, I. J., Czaja, M. J., & Schwabe, R. F. (2014). High-mobility group box 1 is dispensable for autophagy, mitochondrial quality control, and organ function in vivo. Cell metabolism, 19(3), 539–547. https://doi.org/10.1016/j.cmet.2014.01.014

705) Xu, T., Jiang, L., & Wang, Z. (2018). The progression of HMGB1-induced autophagy in cancer biology. OncoTargets and therapy, 12, 365–377. https://doi.org/10.2147/OTT.S185876

706) Merenmies, J., Pihlaskari, R., Laitinen, J., Wartiovaara, J., & Rauvala, H. (1991). 30-kDa heparin-binding protein of brain (amphoterin) involved in neurite outgrowth. Amino acid sequence and localization in the filopodia of the advancing plasma membrane. The Journal of biological chemistry, 266(25), 16722–16729.

707) Fuentes, E., Rojas, A., & Palomo, I. (2014). Role of multiligand/RAGE axis in platelet activation. Thrombosis research, 133(3), 308–314. https://doi.org/10.1016/j.thromres.2013.11.007

708) Maugeri, N., Franchini, S., Campana, L., Baldini, M., Ramirez, G. A., Sabbadini, M. G., Rovere-Querini, P., & Manfredi, A. A. (2012). Circulating platelets as a source of the damage-associated molecular pattern HMGB1 in patients with systemic sclerosis. Autoimmunity, 45(8), 584–587. https://doi.org/10.3109/08916934.2012.719946

709) Passalacqua, M., Zicca, A., Sparatore, B., Patrone, M., Melloni, E., & Pontremoli, S. (1997). Secretion and binding of HMG1 protein to the external surface of the membrane are required for murine erythroleukemia cell differentiation. FEBS letters, 400(3), 275–279. https://doi.org/10.1016/s0014-5793(96)01402-0

710) Hanspal, M., & Hanspal, J. S. (1994). The association of erythroblasts with macrophages promotes erythroid proliferation and maturation: a 30-kD heparin-binding protein is involved in this contact. Blood, 84(10), 3494–3504.

711) Parkkinen, J., & Rauvala, H. (1991). Interactions of plasminogen and tissue plasminogen activator (t-PA) with amphoterin. Enhancement of t-PA-catalyzed plasminogen activation by amphoterin. The Journal of biological chemistry, 266(25), 16730–16735.

712) Ciucci, A., Gabriele, I., Percario, Z. A., Affabris, E., Colizzi, V., & Mancino, G. (2011). HMGB1 and cord blood: its role as immuno-adjuvant factor in innate immunity. PloS one, 6(8), e23766. https://doi.org/10.1371/journal.pone.0023766

713) Klune, J. R., Dhupar, R., Cardinal, J., Billiar, T. R., & Tsung, A. (2008). HMGB1: endogenous danger signaling. Molecular medicine (Cambridge, Mass.), 14(7-8), 476–484. https://doi.org/10.2119/2008-00034.Klune

714) Yang, H., Antoine, D. J., Andersson, U., & Tracey, K. J. (2013). The many faces of HMGB1: molecular structure-functional activity in inflammation, apoptosis, and chemotaxis. Journal of leukocyte biology, 93(6), 865–873. https://doi.org/10.1189/jlb.1212662

715) Pisetsky D. S. (2014). The expression of HMGB1 on microparticles released during cell activation and cell death in vitro and in vivo. Molecular medicine (Cambridge, Mass.), 20(1), 158–

163. https://doi.org/10.2119/molmed.2014.00014

716) Scaffidi, P., Misteli, T., & Bianchi, M. E. (2002). Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. Nature, 418(6894), 191–195. https://doi.org/10.1038/nature00858.

717) Magna, M., & Pisetsky, D. S. (2014). The role of HMGB1 in the pathogenesis of inflammatory and autoimmune diseases. Molecular medicine (Cambridge, Mass.), 20(1), 138–146. https://doi.org/10.2119/molmed.2013.00164

718) Andersson, U., Yang, H., & Harris, H. (2018). Extracellular HMGB1 as a therapeutic target in inflammatory diseases. Expert opinion on therapeutic targets, 22(3), 263–277. https://doi.org/10.1080/14728222.2018.1439924

719) Andersson, U., Wang, H., Palmblad, K., Aveberger, A. C., Bloom, O., Erlandsson-Harris, H., Janson, A., Kokkola, R., Zhang, M., Yang, H., & Tracey, K. J. (2000). High mobility group 1 protein (HMG-1) stimulates proinflammatory cytokine synthesis in human monocytes. The Journal of experimental medicine, 192(4), 565–570. https://doi.org/10.1084/jem.192.4.565

720) Fiuza, C., Bustin, M., Talwar, S., Tropea, M., Gerstenberger, E., Shelhamer, J. H., & Suffredini, A. F. (2003). Inflammation-promoting activity of HMGB1 on human microvascular endothelial cells. Blood, 101(7), 2652–2660. https://doi.org/10.1182/blood-2002-05-1300

721) Liu, G., Wang, J., Park, Y. J., Tsuruta, Y., Lorne, E. F., Zhao, X., & Abraham, E. (2008). High mobility group protein-1 inhibits phagocytosis of apoptotic neutrophils through binding to phosphatidylserine. Journal of immunology (Baltimore, Md. : 1950), 181(6), 4240–4246. https://doi.org/10.4049/jimmunol.181.6.4240.

722) Janko, C., Filipović, M., Munoz, L. E., Schorn, C., Schett, G., Ivanović-Burmazović, I., & Herrmann, M. (2014). Redox modulation of HMGB1-related signaling. Antioxidants & redox signaling, 20(7), 1075–1085. https://doi.org/10.1089/ars.2013.5179

723) Venereau, E., Schiraldi, M., Uguccioni, M., & Bianchi, M. E. (2013). HMGB1 and leukocyte migration during trauma and sterile inflammation. Molecular immunology, 55(1), 76–82. https://doi.org/10.1016/j.molimm.2012.10.037

724) Bianchi, M. E., Crippa, M. P., Manfredi, A. A., Mezzapelle, R., Rovere Querini, P., & Venereau, E. (2017). High-mobility group box 1 protein orchestrates responses to tissue damage via inflammation, innate and adaptive immunity, and tissue repair. Immunological reviews, 280(1), 74–82. https://doi.org/10.1111/imr.12601

725) Andersson, U., & Tracey, K. J. (2011). HMGB1 is a therapeutic target for sterile inflammation and infection. Annual review of immunology, 29, 139–162. https://doi.org/10.1146/annurev-immunol-030409-101323

726) Yang, H., Hreggvidsdottir, H. S., 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, K. J. (2010). A critical cysteine is required for HMGB1 binding to Toll-like receptor 4 and activation of macrophage cytokine release. Proceedings of the National Academy of Sciences of the United States of America, 107(26), 11942–11947. https://doi.org/10.1073/pnas.1003893107

727) Chen, R., Zou, J., Zhong, X., Li, J., Kang, R., & Tang, D. (2024). HMGB1 in the interplay

between autophagy and apoptosis in cancer. Cancer letters, 581, 216494. https://doi.org/10.1016/j.canlet.2023.216494

728) Shi, X., Yu, L., Zhang, Y., Liu, Z., Zhang, H., Zhang, Y., Liu, P., & Du, P. (2020). Glycyrrhetinic acid alleviates hepatic inflammation injury in viral hepatitis disease via a HMGB1-TLR4 signaling pathway. International immunopharmacology, 84, 106578. https://doi.org/10.1016/j.intimp.2020.106578

729) Andersson, U., Ottestad, W., & Tracey, K. J. (2020). Extracellular HMGB1: a therapeutic target in severe pulmonary inflammation including COVID-19?. Molecular medicine (Cambridge, Mass.), 26(1), 42. https://doi.org/10.1186/s10020-020-00172-4

730) Li, Z. C., Cheng, G. Q., Hu, K. Z., Li, M. Q., Zang, W. P., Dong, Y. Q., Wang, W. L., & Liu, Z. D. (2011). Correlation of synovial fluid HMGB-1 levels with radiographic severity of knee osteoarthritis. Clinical and investigative medicine. Medecine clinique et experimentale, 34(5), E298. https://doi.org/10.25011/cim.v34i5.15673

731) Oktayoglu, P., Em, S., Tahtasiz, M., Bozkurt, M., Ucar, D., Yazmalar, L., Nas, K., Yardımeden, I., Cevik, F., Celik, Y., & Mete, N. (2013). 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. Rheumatology international, 33(5), 1327–1331. https://doi.org/10.1007/s00296-012-2578-y

732) Palmblad, K., Sundberg, E., Diez, M., Söderling, R., Aveberger, A. C., Andersson, U., & Harris, H. E. (2007). Morphological characterization of intra-articular HMGB1 expression during the course of collagen-induced arthritis. Arthritis research & therapy, 9(2), R35. https://doi.org/10.1186/ar2155

733) af Klint, E., Grundtman, C., Engström, M., Catrina, A. I., Makrygiannakis, D., Klareskog, L., Andersson, U., & Ulfgren, A. K. (2005). Intraarticular glucocorticoid treatment reduces inflammation in synovial cell infiltrations more efficiently than in synovial blood vessels. Arthritis and rheumatism, 52(12), 3880–3889. https://doi.org/10.1002/art.21488

734) Hofmann, M. A., Drury, S., Hudson, B. I., Gleason, M. R., Qu, W., Lu, Y., Lalla, E., Chitnis, S., Monteiro, J., Stickland, M. H., Bucciarelli, L. G., Moser, B., Moxley, G., Itescu, S., Grant, P. J., Gregersen, P. K., Stern, D. M., & Schmidt, A. M. (2002). RAGE and arthritis: the G82S polymorphism amplifies the inflammatory response. Genes and immunity, 3(3), 123–135. https://doi.org/10.1038/sj.gene.6363861

735) Van de Wouwer, M., Plaisance, S., De Vriese, A., Waelkens, E., Collen, D., Persson, J., Daha, M. R., & Conway, E. M. (2006). The lectin-like domain of thrombomodulin interferes with complement activation and protects against arthritis. Journal of thrombosis and haemostasis : JTH, 4(8), 1813–1824. https://doi.org/10.1111/j.1538-7836.2006.02033.x

736) Zetterström, C. K., Jiang, W., Wähämaa, H., Ostberg, T., Aveberger, A. C., Schierbeck, H., Lotze, M. T., Andersson, U., Pisetsky, D. S., & Erlandsson Harris, H. (2008). Pivotal advance: inhibition of HMGB1 nuclear translocation as a mechanism for the anti-rheumatic effects of gold sodium thiomalate. Journal of leukocyte biology, 83(1), 31–38. https://doi.org/10.1189/jlb.0507323

737) Hamada, T., Torikai, M., Kuwazuru, A., Tanaka, M., Horai, N., Fukuda, T., Yamada, S., Nagayama, S., Hashiguchi, K., Sunahara, N., Fukuzaki, K., Nagata, R., Komiya, S., Maruyama, I., Fukuda, T., & Abeyama, K. (2008). Extracellular high mobility group box chromosomal protein 1 is a coupling factor for hypoxia and inflammation in arthritis. Arthritis and rheumatism, 58(9), 2675–2685. https://doi.org/10.1002/art.23729

738) Han, J., Zhong, J., Wei, W., Wang, Y., Huang, Y., Yang, P., Purohit, S., Dong, Z., Wang, M. H., She, J. X., Gong, F., Stern, D. M., & Wang, C. Y. (2008). Extracellular high-mobility group box 1 acts as an innate immune mediator to enhance autoimmune progression and diabetes onset in NOD mice. Diabetes, 57(8), 2118–2127. https://doi.org/10.2337/db07-1499

739) Lee, H. J., Li, C. W., Hammerstad, S. S., Stefan, M., & Tomer, Y. (2015). Immunogenetics of autoimmune thyroid diseases: A comprehensive review. Journal of autoimmunity, 64, 82–90. https://doi.org/10.1016/j.jaut.2015.07.009

740) Peng, S., Li, C., Wang, X., Liu, X., Han, C., Jin, T., Liu, S., Zhang, X., Zhang, H., He, X., Xie, X., Yu, X., Wang, C., Shan, L., Fan, C., Shan, Z., & Teng, W. (2016). Increased Toll-Like Receptors Activity and TLR Ligands in Patients with Autoimmune Thyroid Diseases. Frontiers in immunology, 7, 578. https://doi.org/10.3389/fimmu.2016.00578

741) Li, C., Peng, S., Liu, X., Han, C., Wang, X., Jin, T., Liu, S., Wang, W., Xie, X., He, X., Zhang, H., Shan, L., Fan, C., Shan, Z., & Teng, W. (2017). Glycyrrhizin, a Direct HMGB1 Antagonist, Ameliorates Inflammatory Infiltration in a Model of Autoimmune Thyroiditis via Inhibition of TLR2-HMGB1 Signaling. Thyroid : official journal of the American Thyroid Association, 27(5), 722–731. https://doi.org/10.1089/thy.2016.0432

742) Catez, F., Yang, H., Tracey, K. J., Reeves, R., Misteli, T., & Bustin, M. (2004). Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. Molecular and cellular biology, 24(10), 4321–4328. https://doi.org/10.1128/MCB.24.10.4321-4328.2004

743) Müller, S., Ronfani, L., & Bianchi, M. E. (2004). Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function. Journal of internal medicine, 255(3), 332–343. https://doi.org/10.1111/j.1365-2796.2003.01296.x

744) Ni, P., Zhang, Y., Liu, Y., Lin, X., Su, X., Lu, H., Shen, H., Xu, W., Xu, H., & Su, Z. (2015). HMGB1 silence could promote MCF-7 cell apoptosis and inhibit invasion and metastasis. International journal of clinical and experimental pathology, 8(12), 15940–15946.

745) Wang, L. L., Meng, Q. H., Jiao, Y., Xu, J. Y., Ge, C. M., Zhou, J. Y., Rosen, E. M., Wang, H. C., & Fan, S. J. (2012). High-mobility group boxes mediate cell proliferation and radiosensitivity via retinoblastoma-interaction-dependent and -independent mechanisms. Cancer biotherapy & radiopharmaceuticals, 27(5), 329–335. https://doi.org/10.1089/cbr.2012.1199

746) Liu, Y., Yan, W., Tohme, S., Chen, M., Fu, Y., Tian, D., Lotze, M., Tang, D., & Tsung, A. (2015). Hypoxia induced HMGB1 and mitochondrial DNA interactions mediate tumor growth in hepatocellular carcinoma through Toll-like receptor 9. Journal of hepatology, 63(1), 114–121. https://doi.org/10.1016/j.jhep.2015.02.009

747) Jiang, J., Wang, G. Z., Wang, Y., Huang, H. Z., Li, W. T., & Qu, X. D. (2018). Hypoxia-induced HMGB1 expression of HCC promotes tumor invasiveness and metastasis via regulating

macrophage-derived IL-6. Experimental cell research, 367(1), 81–88. https://doi.org/10.1016/j.yexcr.2018.03.025

748) Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. Cell, 144(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

749) Yang, Y., Yang, L., Jiang, S., Yang, T., Lan, J., Lei, Y., Tan, H., & Pan, K. (2020). HMGB1 mediates lipopolysaccharide-induced inflammation via interacting with GPX4 in colon cancer cells. Cancer cell international, 20, 205. https://doi.org/10.1186/s12935-020-01289-6

750) Yuan, S., Liu, Z., Xu, Z., Liu, J., & Zhang, J. (2020). High mobility group box 1 (HMGB1): a pivotal regulator of hematopoietic malignancies. Journal of hematology & oncology, 13(1), 91. https://doi.org/10.1186/s13045-020-00920-3

751) Mukherjee, A., & Vasquez, K. M. (2020). Targeting Chromosomal Architectural HMGB Proteins Could Be the Next Frontier in Cancer Therapy. Cancer research, 80(11), 2075–2082. https://doi.org/10.1158/0008-5472.CAN-19-3066

752) Liu, Y., Dong, Y., Kong, L., Shi, F., Zhu, H., & Yu, J. (2018). Abscopal effect of radiotherapy combined with immune checkpoint inhibitors. Journal of hematology & oncology, 11(1), 104. https://doi.org/10.1186/s13045-018-0647-8

753) Shi, Q., Wang, Y., Dong, W., Song, E., & Song, Y. (2019). Polychlorinated biphenyl quinone-induced signaling transition from autophagy to apoptosis is regulated by HMGB1 and p53 in human hepatoma HepG2 cells. Toxicology letters, 306, 25–34. https://doi.org/10.1016/j.toxlet.2019.02.002

754) Liang, W. J., Yang, H. W., Liu, H. N., Qian, W., & Chen, X. L. (2020). HMGB1 upregulates NF-kB by inhibiting IKB-α and associates with diabetic retinopathy. Life sciences, 241, 117146. https://doi.org/10.1016/j.lfs.2019.117146

755) Dong, H., Zhang, L., & Liu, S. (2022). Targeting HMGB1: An available Therapeutic Strategy for Breast Cancer Therapy. International journal of biological sciences, 18(8), 3421–3434. https://doi.org/10.7150/ijbs.73504

756) Jiao, Y., Wang, H. C., & Fan, S. J. (2007). Growth suppression and radiosensitivity increase by HMGB1 in breast cancer. Acta pharmacologica Sinica, 28(12), 1957–1967. https://doi.org/10.1111/j.1745-7254.2007.00669.x

757) De Azevedo, J., Mourtada, J., Bour, C., Devignot, V., Schultz, P., Borel, C., Pencreach, E., Mellitzer, G., Gaiddon, C., & Jung, A. C. (2022). The EXTREME Regimen Associating Cetuximab and Cisplatin Favors Head and Neck Cancer Cell Death and Immunogenicity with the Induction of an Anti-Cancer Immune Response. Cells, 11(18), 2866. https://doi.org/10.3390/cells11182866

758) Li, C., Zhang, Y., Cheng, X., Yuan, H., Zhu, S., Liu, J., Wen, Q., Xie, Y., Liu, J., Kroemer, G., Klionsky, D. J., Lotze, M. T., Zeh, H. J., Kang, R., & Tang, D. (2018). PINK1 and PARK2 Suppress Pancreatic Tumorigenesis through Control of Mitochondrial Iron-Mediated Immunometabolism. Developmental cell, 46(4), 441–455.e8. https://doi.org/10.1016/j.devcel.2018.07.012

759) Hubert, P., Roncarati, P., Demoulin, S., Pilard, C., Ancion, M., Reynders, C., Lerho, T., Bruyere, D., Lebeau, A., Radermecker, C., Meunier, M., Nokin, M. J., Hendrick, E., Peulen, O.,

Delvenne, P., & Herfs, M. (2021). Extracellular HMGB1 blockade inhibits tumor growth through profoundly remodeling immune microenvironment and enhances checkpoint inhibitor-based immunotherapy. Journal for immunotherapy of cancer, 9(3), e001966. https://doi.org/10.1136/jitc-2020-001966

760) Troitskaya, O., Varlamov, M., Nushtaeva, A., Richter, V., & Koval, O. (2020). Recombinant Lactaptin Induces Immunogenic Cell Death and Creates an Antitumor Vaccination Effect in Vivo with Enhancement by an IDO Inhibitor. Molecules (Basel, Switzerland), 25(12), 2804. https://doi.org/10.3390/molecules25122804

761) Pasquereau-Kotula, E., Habault, J., Kroemer, G., & Poyet, J. L. (2018). The anticancer peptide RT53 induces immunogenic cell death. PloS one, 13(8), e0201220. https://doi.org/10.1371/journal.pone.0201220

762) Wang, Y., Xie, W., Humeau, J., Chen, G., Liu, P., Pol, J., Zhang, Z., Kepp, O., & Kroemer, G. (2020). Autophagy induction by thiostrepton improves the efficacy of immunogenic chemotherapy. Journal for immunotherapy of cancer, 8(1), e000462. https://doi.org/10.1136/jitc-2019-000462

763) Kajioka, H., Kagawa, S., Ito, A., Yoshimoto, M., Sakamoto, S., Kikuchi, S., Kuroda, S., Yoshida, R., Umeda, Y., Noma, K., Tazawa, H., & Fujiwara, T. (2021). Targeting neutrophil extracellular traps with thrombomodulin prevents pancreatic cancer metastasis. Cancer letters, 497, 1–13. https://doi.org/10.1016/j.canlet.2020.10.015

764) Tang, D., Kang, R., Livesey, K. M., Zeh, H. J., 3rd, & Lotze, M. T. (2011). High mobility group box 1 (HMGB1) activates an autophagic response to oxidative stress. Antioxidants & redox signaling, 15(8), 2185–2195. https://doi.org/10.1089/ars.2010.3666

765) Guo, X., He, D., Zhang, E., Chen, J., Chen, Q., Li, Y., Yang, L., Yang, Y., Zhao, Y., Wang, G., He, J., & Cai, Z. (2018). HMGB1 knockdown increases MM cell vulnerability by regulating autophagy and DNA damage repair. Journal of experimental & clinical cancer research : CR, 37(1), 205. https://doi.org/10.1186/s13046-018-0883-3

766) Nakamura, T., Okui, T., Hasegawa, K., Ryumon, S., Ibaragi, S., Ono, K., Kunisada, Y., Obata, K., Masui, M., Shimo, T., & Sasaki, A. (2020). High mobility group box 1 induces bone pain associated with bone invasion in a mouse model of advanced head and neck cancer. Oncology reports, 44(6), 2547–2558. https://doi.org/10.3892/or.2020.7788

767) Yi, C., Zhang, X., Li, H., Chen, G., Zeng, B., Li, Y., Wang, C., He, Y., Chen, X., Huang, Z., & Yu, D. (2021). EPHB4 Regulates the Proliferation and Metastasis of Oral Squamous Cell Carcinoma through the HMGB1/NF-κB Signalling Pathway. Journal of Cancer, 12(20), 5999–6011. https://doi.org/10.7150/jca.59331

768) Jiao, D., Zhang, J., Chen, P., Guo, X., Qiao, J., Zhu, J., Wang, L., Lu, Z., & Liu, Z. (2021). HN1L promotes migration and invasion of breast cancer by up-regulating the expression of HMGB1. Journal of cellular and molecular medicine, 25(1), 397–410. https://doi.org/10.1111/jcmm.16090

769) Liu, N., Wu, Y., Wen, X., Li, P., Lu, F., & Shang, H. (2021). Chronic stress promotes acute myeloid leukemia progression through HMGB1/NLRP3/IL-1β signaling pathway. Journal of

molecular medicine (Berlin, Germany), 99(3), 403–414. https://doi.org/10.1007/s00109-020-02011-9

770) Wang, W., Zhu, H., Zhang, H., Zhang, L., Ding, Q., & Jiang, H. (2014). Targeting HMGB1 inhibits bladder cancer cells bioactivity by lentivirus-mediated RNA interference. Neoplasma, 61(6), 638–646. https://doi.org/10.4149/neo_2014_079

771) Zhang, J., Kou, Y. B., Zhu, J. S., Chen, W. X., & Li, S. (2014). Knockdown of HMGB1 inhibits growth and invasion of gastric cancer cells through the NF-κB pathway in vitro and in vivo. International journal of oncology, 44(4), 1268–1276. https://doi.org/10.3892/ijo.2014.2285

772) Elangovan, I., Thirugnanam, S., Chen, A., Zheng, G., Bosland, M. C., Kajdacsy-Balla, A., & Gnanasekar, M. (2012). Targeting receptor for advanced glycation end products (RAGE) expression induces apoptosis and inhibits prostate tumor growth. Biochemical and biophysical research communications, 417(4), 1133–1138. https://doi.org/10.1016/j.bbrc.2011.12.060

773) Sun, E., Liu, K., Zhao, K., & Wang, L. (2019). Serine/threonine kinase 32C is overexpressed in bladder cancer and contributes to tumor progression. Cancer biology & therapy, 20(3), 307–320. https://doi.org/10.1080/15384047.2018.1529098

774) Zhang, D. Y., Zou, X. J., Cao, C. H., Zhang, T., Lei, L., Qi, X. L., Liu, L., & Wu, D. H. (2018). Identification and Functional Characterization of Long Non-coding RNA MIR22HG as a Tumor Suppressor for Hepatocellular Carcinoma. Theranostics, 8(14), 3751–3765. https://doi.org/10.7150/thno.22493

775) Zheng, M., Xu, L., Wei, C., & Guan, W. (2023). CircRTN1 stimulates HMGB1 to regulate the malignant progression of papillary thyroid cancer by sponging miR-101-3p. Hormones (Athens, Greece), 22(2), 281–293. https://doi.org/10.1007/s42000-023-00440-y

776) Chen, L., Chi, K., Xiang, H., & Yang, Y. (2020). Circ_0032821 Facilitates Gastric Cancer Cell Proliferation, Migration, Invasion and Glycolysis by Regulating MiR-1236-3p/HMGB1 Axis. Cancer management and research, 12, 9965–9976. https://doi.org/10.2147/CMAR.S270164

777) Liu, Z., Dou, C., Yao, B., Xu, M., Ding, L., Wang, Y., Jia, Y., Li, Q., Zhang, H., Tu, K., Song, T., & Liu, Q. (2016). Methylation-mediated repression of microRNA-129-2 suppresses cell aggressiveness by inhibiting high mobility group box 1 in human hepatocellular carcinoma. Oncotarget, 7(24), 36909–36923. https://doi.org/10.18632/oncotarget.9377

778) Sun, S., Ma, J., Xie, P., Wu, Z., & Tian, X. (2020). Hypoxia-responsive miR-141-3p is involved in the progression of breast cancer via mediating the HMGB1/HIF-1α signaling pathway. The journal of gene medicine, 22(10), e3230. https://doi.org/10.1002/jgm.3230

779) Xu, Q., Liao, B., Hu, S., Zhou, Y., & Xia, W. (2021). Circular RNA 0081146 facilitates the progression of gastric cancer by sponging miR-144 and up-regulating HMGB1. Biotechnology letters, 43(4), 767–779. https://doi.org/10.1007/s10529-020-03058-x

780) Liu, P. L., Liu, W. L., Chang, J. M., Chen, Y. H., Liu, Y. P., Kuo, H. F., Hsieh, C. C., Ding, Y. S., Chen, W. W., & Chong, I. W. (2017). MicroRNA-200c inhibits epithelial-mesenchymal transition, invasion, and migration of lung cancer by targeting HMGB1. PloS one, 12(7), e0180844. https://doi.org/10.1371/journal.pone.0180844

781) Shan, G., Tang, T., Xia, Y., & Qian, H. J. (2020). Long non-coding RNA NEAT1 promotes bladder progression through regulating miR-410 mediated HMGB1. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 121, 109248. https://doi.org/10.1016/j.biopha.2019.109248

782) Zhang, J. L., Zheng, H. F., Li, K., & Zhu, Y. P. (2022). miR-495-3p depresses cell proliferation and migration by downregulating HMGB1 in colorectal cancer. World journal of surgical oncology, 20(1), 101. https://doi.org/10.1186/s12957-022-02500-w

783) Yang, G., Xu, Q., Wan, Y., Zhang, L., Wang, L., & Meng, F. (2021). Circ-CSPP1 knockdown suppresses hepatocellular carcinoma progression through miR-493-5p releasing-mediated HMGB1 downregulation. Cellular signalling, 86, 110065. https://doi.org/10.1016/j.cellsig.2021.110065

784) Liu, Y., Guo, R., Qiao, Y., Han, L., & Liu, M. (2020). LncRNA NNT-AS1 contributes to the cisplatin resistance of cervical cancer through NNT-AS1/miR-186/HMGB1 axis. Cancer cell international, 20, 190. https://doi.org/10.1186/s12935-020-01278-9

785) Wu, D., Zhang, T., Wang, J., Zhou, J., Pan, H., & Qu, P. (2019). Long noncoding RNA NNT-AS1 enhances the malignant phenotype of bladder cancer by acting as a competing endogenous RNA on microRNA-496 thereby increasing HMGB1 expression. Aging, 11(24), 12624–12640. https://doi.org/10.18632/aging.102591

786) Tian, L., Wang, Z. Y., Hao, J., & Zhang, X. Y. (2019). miR-505 acts as a tumor suppressor in gastric cancer progression through targeting HMGB1. Journal of cellular biochemistry, 120(5), 8044–8052. https://doi.org/10.1002/jcb.28082

787) Feng X. E. (2021). miR-548b Suppresses Melanoma Cell Growth, Migration, and Invasion by Negatively Regulating Its Target Gene HMGB1. Cancer biotherapy & radiopharmaceuticals, 36(2), 189–201. https://doi.org/10.1089/cbr.2019.3507

788) Qiu, Z., Pan, X. X., & You, D. Y. (2020). LncRNA DSCAM-AS1 promotes non-small cell lung cancer progression via regulating miR-577/HMGB1 axis. Neoplasma, 67(4), 871–879. https://doi.org/10.4149/neo_2020_190826N821

789) Zhang, S., Long, J., & Hu, Y. (2020). Long noncoding RNA LINC00205 enhances the malignant characteristics of retinoblastoma by acting as a molecular sponge of microRNA-665 and consequently increasing HMGB1 expression. Biochemical and biophysical research communications, 526(2), 396–403. https://doi.org/10.1016/j.bbrc.2020.03.083

790) Ding, S., Zhang, G., Gao, Y., Chen, S., & Cao, C. (2020). Circular RNA hsa_circ_0005909 modulates osteosarcoma progression via the miR-936/HMGB1 axis. Cancer cell international, 20, 305. https://doi.org/10.1186/s12935-020-01399-1

791) Ding, C., Yu, H., Shi, C., Shi, T., Qin, H., & Cui, Y. (2019). MiR-let-7e inhibits invasion and magration and regulates HMGB1 expression in papillary thyroid carcinoma. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 110, 528–536. https://doi.org/10.1016/j.biopha.2018.11.057

792) Li, P., Ren, K., Liang, Y. Y., Liu, J. K., Liang, Z. W., & Zhang, Y. F. (2020). Aloin promotes cell apoptosis by targeting HMGB1-TLR4-ERK axis in human melanoma cells. EXCLI journal, 19, 641–651.

793) Buoncervello, M., Borghi, P., Romagnoli, G., Spadaro, F., Belardelli, F., Toschi, E., & Gabriele, L. (2012). Apicidin and docetaxel combination treatment drives CTCFL expression and HMGB1 release acting as potential antitumor immune response inducers in metastatic breast cancer cells. Neoplasia (New York, N.Y.), 14(9), 855–867. https://doi.org/10.1593/neo.121020

794) Shang, Z., Li, Y., Zhang, M., Tian, J., Han, R., Shyr, C. R., Messing, E., Yeh, S., Niu, Y., & Chang, C. (2015). Antiandrogen Therapy with Hydroxyflutamide or Androgen Receptor Degradation Enhancer ASC-J9 Enhances BCG Efficacy to Better Suppress Bladder Cancer Progression. Molecular cancer therapeutics, 14(11), 2586–2594. https://doi.org/10.1158/1535-7163.MCT-14-1055-T

795) Genoud, V., Espinoza, F. I., Marinari, E., Rochemont, V., Dietrich, P. Y., McSheehy, P., Bachmann, F., Lane, H. A., & Walker, P. R. (2021). Treating ICB-resistant glioma with anti-CD40 and mitotic spindle checkpoint controller BAL101553 (lisavanbulin). JCI insight, 6(18), e142980. https://doi.org/10.1172/jci.insight.142980

796) Kines, R. C., Thompson, C. D., Spring, S., Li, Z., de Los Pinos, E., Monks, S., & Schiller, J. T. (2021). Virus-Like Particle-Drug Conjugates Induce Protective, Long-lasting Adaptive Antitumor Immunity in the Absence of Specifically Targeted Tumor Antigens. Cancer immunology research, 9(6), 693–706. https://doi.org/10.1158/2326-6066.CIR-19-0974

797) Wang, M., Zhao, L., Tong, D., Yang, L., Zhu, H., Li, Q., & Zhang, F. (2019). BET bromodomain inhibitor JQ1 promotes immunogenic cell death in tongue squamous cell carcinoma. International immunopharmacology, 76, 105921. https://doi.org/10.1016/j.intimp.2019.105921

798) Erkes, D. A., Cai, W., Sanchez, I. M., Purwin, T. J., Rogers, C., Field, C. O., Berger, A. C., Hartsough, E. J., Rodeck, U., Alnemri, E. S., & Aplin, A. E. (2020). Mutant BRAF and MEK Inhibitors Regulate the Tumor Immune Microenvironment via Pyroptosis. Cancer discovery, 10(2), 254–269. https://doi.org/10.1158/2159-8290.CD-19-0672

799) Florêncio, K. G. D., Edson, E. A., Fernandes, K. S. D. S., Luiz, J. P. M., Pinto, F. D. C. L., Pessoa, O. D. L., Cunha, F. Q., Machado-Neto, J. A., & Wilke, D. V. (2022). Chromomycin $A_5$ induces bona fide immunogenic cell death in melanoma. Frontiers in immunology, 13, 941757. https://doi.org/10.3389/fimmu.2022.941757

800) Hossain, D. M. S., Javaid, S., Cai, M., Zhang, C., Sawant, A., Hinton, M., Sathe, M., Grein, J., Blumenschein, W., Pinheiro, E. M., & Chackerian, A. (2018). Dinaciclib induces immunogenic cell death and enhances anti-PD1-mediated tumor suppression. The Journal of clinical investigation, 128(2), 644–654. https://doi.org/10.1172/JCI94586

801) Hufnagel, S., Xu, H., Colemam, M. F., Valdes, S. A., Liu, K. A., Hursting, S. D., & Cui, Z. (2022). 4-(N)-Docosahexaenoyl 2', 2'-difluorodeoxycytidine induces immunogenic cell death in colon and pancreatic carcinoma models as a single agent. Cancer chemotherapy and pharmacology, 89(1), 59–69. https://doi.org/10.1007/s00280-021-04367-2

802) Saleh, M. A., Antar, S. A., Abdo, W., Ashour, A., & Zaki, A. A. (2023). Genistin modulates high-mobility group box protein 1 (HMGB1) and nuclear factor kappa-B (NF-κB) in Ehrlich-ascites-carcinoma-bearing mice. Environmental science and pollution research international, 30(1), 966–978. https://doi.org/10.1007/s11356-022-22268-6

803) Barnoud, T., Leung, J. C., Leu, J. I., Basu, S., Poli, A. N. R., Parris, J. L. D., Indeglia, A., Martynyuk, T., Good, M., Gnanapradeepan, K., Sanseviero, E., Moeller, R., Tang, H. Y., Cassel, J., Kossenkov, A. V., Liu, Q., Speicher, D. W., Gabrilovich, D. I., Salvino, J. M., George, D. L., … Murphy, M. E. (2020). A Novel Inhibitor of HSP70 Induces Mitochondrial Toxicity and Immune Cell Recruitment in Tumors. Cancer research, 80(23), 5270–5281. https://doi.org/10.1158/0008-5472.CAN-20-0397

804) Huang, S. W., Wang, S. T., Chang, S. H., Chuang, K. C., Wang, H. Y., Kao, J. K., Liang, S. M., Wu, C. Y., Kao, S. H., Chen, Y. J., & Shieh, J. J. (2020). Imiquimod Exerts Antitumor Effects by Inducing Immunogenic Cell Death and Is Enhanced by the Glycolytic Inhibitor 2-Deoxyglucose. The Journal of investigative dermatology, 140(9), 1771–1783.e6. https://doi.org/10.1016/j.jid.2019.12.039

805) Song, X., Zhu, S., Xie, Y., Liu, J., Sun, L., Zeng, D., Wang, P., Ma, X., Kroemer, G., Bartlett, D. L., Billiar, T. R., Lotze, M. T., Zeh, H. J., Kang, R., & Tang, D. (2018). JTC801 Induces pH-dependent Death Specifically in Cancer Cells and Slows Growth of Tumors in Mice. Gastroenterology, 154(5), 1480–1493. https://doi.org/10.1053/j.gastro.2017.12.004

806) Yuan, Y., Wang, Z., Yang, R., Qian, T., & Zhou, Q. (2019). Naphthyl quinoxaline thymidine conjugate is a potent anticancer agent post UVA activation and elicits marked inhibition of tumor growth through vaccination. European journal of medicinal chemistry, 171, 255–264. https://doi.org/10.1016/j.ejmech.2019.03.051

807) Pellegrini, L., Xue, J., Larson, D., Pastorino, S., Jube, S., Forest, K. H., Saad-Jube, Z. S., Napolitano, A., Pagano, I., Negi, V. S., Bianchi, M. E., Morris, P., Pass, H. I., Gaudino, G., Carbone, M., & Yang, H. (2017). HMGB1 targeting by ethyl pyruvate suppresses malignant phenotype of human mesothelioma. Oncotarget, 8(14), 22649–22661. https://doi.org/10.18632/oncotarget.15152

808) Kang, R., Tang, D., Schapiro, N. E., Loux, T., Livesey, K. M., Billiar, T. R., Wang, H., Van Houten, B., Lotze, M. T., & Zeh, H. J. (2014). The HMGB1/RAGE inflammatory pathway promotes pancreatic tumor growth by regulating mitochondrial bioenergetics. Oncogene, 33(5), 567–577. https://doi.org/10.1038/onc.2012.631

809) Liang, X., Chavez, A. R., Schapiro, N. E., Loughran, P., Thorne, S. H., Amoscato, A. A., Zeh, H. J., Beer-Stolz, D., Lotze, M. T., & de Vera, M. E. (2009). Ethyl pyruvate administration inhibits hepatic tumor growth. Journal of leukocyte biology, 86(3), 599–607. https://doi.org/10.1189/jlb.0908578

810) Kam, A. Y. F., Piryani, S. O., McCall, C. M., Park, H. S., Rizzieri, D. A., & Doan, P. L. (2019). Targeting High Mobility Group Box-1 (HMGB1) Promotes Cell Death in Myelodysplastic Syndrome. Clinical cancer research : an official journal of the American Association for Cancer Research, 25(13), 4155–4167. https://doi.org/10.1158/1078-0432.CCR-18-3517

811) Li, Y., Zhang, H., Li, Q., Zou, P., Huang, X., Wu, C., & Tan, L. (2020). CDK12/13 inhibition induces immunogenic cell death and enhances anti-PD-1 anticancer activity in breast cancer. Cancer letters, 495, 12–21. https://doi.org/10.1016/j.canlet.2020.09.011

812) Song, S., Lee, J. Y., Ermolenko, L., Mazumder, A., Ji, S., Ryu, H., Kim, H., Kim, D. W., Lee, J. W., Dicato, M., Christov, C., Schnekenburger, M., Cerella, C., Gérard, D., Orlikova-Boyer,

B., Al-Mourabit, A., & Diederich, M. (2020). Tetrahydrobenzimidazole TMQ0153 triggers apoptosis, autophagy and necroptosis crosstalk in chronic myeloid leukemia. Cell death & disease, 11(2), 109. https://doi.org/10.1038/s41419-020-2304-8

813) Combès, E., Andrade, A. F., Tosi, D., Michaud, H. A., Coquel, F., Garambois, V., Desigaud, D., Jarlier, M., Coquelle, A., Pasero, P., Bonnefoy, N., Moreaux, J., Martineau, P., Del Rio, M., Beijersbergen, R. L., Vezzio-Vie, N., & Gongora, C. (2019). Inhibition of Ataxia-Telangiectasia Mutated and RAD3-Related (ATR) Overcomes Oxaliplatin Resistance and Promotes Antitumor Immunity in Colorectal Cancer. Cancer research, 79(11), 2933–2946. https://doi.org/10.1158/0008-5472.CAN-18-2807

814) Zhang, H., Wang, J., Li, J., Zhou, X., Yin, L., Wang, Y., Gu, Y., Niu, X., Yang, Y., Ji, H., & Zhang, Q. (2021). HMGB1 is a key factor for tamoxifen resistance and has the potential to predict the efficacy of CDK4/6 inhibitors in breast cancer. Cancer science, 112(4), 1603–1613. https://doi.org/10.1111/cas.14813

815) Kofla, G., Radecke, C., Frentsch, M., Walther, W., Stintzing, S., Riess, H., Bullinger, L., & Na, I. K. (2022). Conventional amphotericin B elicits markers of immunogenic cell death on leukemic blasts, mediates immunostimulatory effects on phagocytic cells, and synergizes with PD-L1 blockade. Oncoimmunology, 11(1), 2068109. https://doi.org/10.1080/2162402X.2022.2068109

816) Yang, H., Pellegrini, L., Napolitano, A., Giorgi, C., Jube, S., Preti, A., Jennings, C. J., De Marchis, F., Flores, E. G., Larson, D., Pagano, I., Tanji, M., Powers, A., Kanodia, S., Gaudino, G., Pastorino, S., Pass, H. I., Pinton, P., Bianchi, M. E., & Carbone, M. (2015). Aspirin delays mesothelioma growth by inhibiting HMGB1-mediated tumor progression. Cell death & disease, 6(6), e1786. https://doi.org/10.1038/cddis.2015.153

817) Patnaik, A., Swanson, K. D., Csizmadia, E., Solanki, A., Landon-Brace, N., Gehring, M. P., Helenius, K., Olson, B. M., Pyzer, A. R., Wang, L. C., Elemento, O., Novak, J., Thornley, T. B., Asara, J. M., Montaser, L., Timmons, J. J., Morgan, T. M., Wang, Y., Levantini, E., Clohessy, J. G., … Cantley, L. C. (2017). Cabozantinib Eradicates Advanced Murine Prostate Cancer by Activating Antitumor Innate Immunity. Cancer discovery, 7(7), 750–765. https://doi.org/10.1158/2159-8290.CD-16-0778

818) Karsch-Bluman, A., Amoyav, B., Friedman, N., Shoval, H., Schwob, O., Ella, E., Wald, O., & Benny, O. (2017). High mobility group box 1 antagonist limits metastatic seeding in the lungs via reduction of cell-cell adhesion. Oncotarget, 8(20), 32706–32721. https://doi.org/10.18632/oncotarget.16188

819) Xue, J., Patergnani, S., Giorgi, C., Suarez, J., Goto, K., Bononi, A., Tanji, M., Novelli, F., Pastorino, S., Xu, R., Caroccia, N., Dogan, A. U., Pass, H. I., Tognon, M., Pinton, P., Gaudino, G., Mak, T. W., Carbone, M., & Yang, H. (2020). Asbestos induces mesothelial cell transformation via HMGB1-driven autophagy. Proceedings of the National Academy of Sciences of the United States of America, 117(41), 25543–25552. https://doi.org/10.1073/pnas.2007622117

820) Wang, C., Peng, G., Huang, H., Liu, F., Kong, D. P., Dong, K. Q., Dai, L. H., Zhou, Z., Wang, K. J., Yang, J., Cheng, Y. Q., Gao, X., Qu, M., Wang, H. R., Zhu, F., Tian, Q. Q., Liu, D., Cao, L., Cui, X. G., Xu, C. L., … Sun, Y. H. (2018). Blocking the Feedback Loop between Neuroendocrine

Differentiation and Macrophages Improves the Therapeutic Effects of Enzalutamide (MDV3100) on Prostate Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research, 24(3), 708–723. https://doi.org/10.1158/1078-0432.CCR-17-2446

821) Keyvani-Ghamsari, S., Rabbani-Chadegani, A., Sargolzaei, J., & Shahhoseini, M. (2017). Effect of irinotecan on HMGB1, MMP9 expression, cell cycle, and cell growth in breast cancer (MCF-7) cells. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine, 39(4), 1010428317698354. https://doi.org/10.1177/1010428317698354

822) Razmi, M., Rabbani-Chadegani, A., Hashemi-Niasari, F., & Ghadam, P. (2018). Lithium chloride attenuates mitomycin C induced necrotic cell death in MDA-MB-231 breast cancer cells via HMGB1 and Bax signaling. Journal of trace elements in medicine and biology: organ of the Society for Minerals and Trace Elements (GMS), 48, 87–96. https://doi.org/10.1016/j.jtemb.2018.03.011

823) Sequeira, G. R., Sahores, A., Dalotto-Moreno, T., Perrotta, R. M., Pataccini, G., Vanzulli, S. I., Polo, M. L., Radisky, D. C., Sartorius, C. A., Novaro, V., Lamb, C. A., Rabinovich, G. A., Salatino, M., & Lanari, C. (2021). Enhanced Antitumor Immunity via Endocrine Therapy Prevents Mammary Tumor Relapse and Increases Immune Checkpoint Blockade Sensitivity. Cancer research, 81(5), 1375–1387. https://doi.org/10.1158/0008-5472.CAN-20-1441

824) Li, C., Sun, H., Wei, W., Liu, Q., Wang, Y., Zhang, Y., Lian, F., Liu, F., Li, C., Ying, K., Huo, H., Qi, Z., & Li, B. (2020). Mitoxantrone triggers immunogenic prostate cancer cell death via p53-dependent PERK expression. Cellular oncology (Dordrecht, Netherlands), 43(6), 1099–1116. https://doi.org/10.1007/s13402-020-00544-2

825) Huang, Z., & Hu, H. (2021). Arginine Deiminase Induces Immunogenic Cell Death and Is Enhanced by N-acetylcysteine in Murine MC38 Colorectal Cancer Cells and MDA-MB-231 Human Breast Cancer Cells In Vitro. Molecules (Basel, Switzerland), 26(2), 511. https://doi.org/10.3390/molecules26020511

826) Zhu, H., Shan, Y., Ge, K., Lu, J., Kong, W., & Jia, C. (2020). Oxaliplatin induces immunogenic cell death in hepatocellular carcinoma cells and synergizes with immune checkpoint blockade therapy. Cellular oncology (Dordrecht, Netherlands), 43(6), 1203–1214. https://doi.org/10.1007/s13402-020-00552-2

827) Schaer, D. A., Geeganage, S., Amaladas, N., Lu, Z. H., Rasmussen, E. R., Sonyi, A., Chin, D., Capen, A., Li, Y., Meyer, C. M., Jones, B. D., Huang, X., Luo, S., Carpenito, C., Roth, K. D., Nikolayev, A., Tan, B., Brahmachary, M., Chodavarapu, K., Dorsey, F. C., … Novosiadly, R. D. (2019). The Folate Pathway Inhibitor Pemetrexed Pleiotropically Enhances Effects of Cancer Immunotherapy. Clinical cancer research: an official journal of the American Association for Cancer Research, 25(23), 7175–7188. https://doi.org/10.1158/1078-0432.CCR-19-0433

828) Booth, L., Roberts, J. L., Poklepovic, A., & Dent, P. (2017). [pemetrexed + sildenafil], via autophagy-dependent HDAC downregulation, enhances the immunotherapy response of NSCLC cells. Cancer biology & therapy, 18(9), 705–714. https://doi.org/10.1080/15384047.2017.1362511

829) Shan, C. K., Du, Y. B., Zhai, X. T., Wang, Y. X., Li, Y., Gong, J. H., Ge, Z. J., Liu, X. J., & Zhen, Y. S. (2021). Pingyangmycin enhances the antitumor efficacy of anti-PD-1 therapy associated

with tumor-infiltrating $CD8^+$ T cell augmentation. Cancer chemotherapy and pharmacology, 87(3), 425–436. https://doi.org/10.1007/s00280-020-04209-7

830) Di Grazia, A., Laudisi, F., Di Fusco, D., Franzè, E., Ortenzi, A., Monteleone, I., Monteleone, G., & Stolfi, C. (2020). Rafoxanide Induces Immunogenic Death of Colorectal Cancer Cells. Cancers, 12(5), 1314. https://doi.org/10.3390/cancers12051314

831) Zhang, J., Shen, L., Li, X., Song, W., Liu, Y., & Huang, L. (2019). Nanoformulated Codelivery of Quercetin and Alantolactone Promotes an Antitumor Response through Synergistic Immunogenic Cell Death for Microsatellite-Stable Colorectal Cancer. ACS nano, 13(11), 12511–12524. https://doi.org/10.1021/acsnano.9b02875

832) Li, C., Zhang, Y., Yan, S., Zhang, G., Wei, W., Qi, Z., & Li, B. (2021). Alternol triggers immunogenic cell death via reactive oxygen species generation. Oncoimmunology, 10(1), 1952539. https://doi.org/10.1080/2162402X.2021.1952539

833) Gu, X., Zhu, L. Y., Xu, Z. Y., & Shen, K. P. (2021). Astragaloside IV and Saponins of Rhizoma Polygonati Cure Cyclophosphamide-Induced Myelosuppression in Lung Adenocarcinoma via Down-Regulating miR-142-3p. Frontiers in oncology, 11, 630921. https://doi.org/10.3389/fonc.2021.630921

834) Hashemi-Niasari, F., Rabbani-Chadegani, A., Razmi, M., & Fallah, S. (2018). Synergy of theophylline reduces necrotic effect of berberine, induces cell cycle arrest and PARP, HMGB1, Bcl-2 family mediated apoptosis in MDA-MB-231 breast cancer cells. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 106, 858–867. https://doi.org/10.1016/j.biopha.2018.07.019

835) D'Eliseo, D., Manzi, L., & Velotti, F. (2013). Capsaicin as an inducer of damage-associated molecular patterns (DAMPs) of immunogenic cell death (ICD) in human bladder cancer cells. Cell stress & chaperones, 18(6), 801–808. https://doi.org/10.1007/s12192-013-0422-2

836) Zhu, C., Fang, Z., Peng, L., Gao, F., Peng, W., & Song, F. (2022). Curcumin Suppresses the Progression of Colorectal Cancer by Improving Immunogenic Cell Death Caused by Irinotecan. Chemotherapy, 67(4), 211–222. https://doi.org/10.1159/000518121

837) Kam, N. W., Wu, K. C., Dai, W., Wang, Y., Yan, L. Y. C., Shakya, R., Khanna, R., Qin, Y., Law, S., Lo, A. W. I., Lee, V. H. F., Guan, X. Y., & Kwong, D. L. (2022). Peritumoral B cells drive proangiogenic responses in HMGB1-enriched esophageal squamous cell carcinoma. Angiogenesis, 25(2), 181–203. https://doi.org/10.1007/s10456-021-09819-0

838) Shinde-Jadhav, S., Mansure, J. J., Rayes, R. F., Marcq, G., Ayoub, M., Skowronski, R., Kool, R., Bourdeau, F., Brimo, F., Spicer, J., & Kassouf, W. (2021). Role of neutrophil extracellular traps in radiation resistance of invasive bladder cancer. Nature communications, 12(1), 2776. https://doi.org/10.1038/s41467-021-23086-z

839) Kang, R., Xie, Y., Zhang, Q., Hou, W., Jiang, Q., Zhu, S., Liu, J., Zeng, D., Wang, H., Bartlett, D. L., Billiar, T. R., Zeh, H. J., 3rd, Lotze, M. T., & Tang, D. (2017). Intracellular HMGB1 as a novel tumor suppressor of pancreatic cancer. Cell research, 27(7), 916–932. https://doi.org/10.1038/cr.2017.51

840) Waki, K., & Yamada, A. (2016). Blockade of high mobility group box 1 augments antitumor T-cell response induced by peptide vaccination as a co-adjuvant. Cancer science, 107(12), 1721–1729. https://doi.org/10.1111/cas.13084

841) Sun, D., Zou, Y., Song, L., Han, S., Yang, H., Chu, D., Dai, Y., Ma, J., O'Driscoll, C. M., Yu, Z., & Guo, J. (2022). A cyclodextrin-based nanoformulation achieves co-delivery of ginsenoside Rg3 and quercetin for chemo-immunotherapy in colorectal cancer. Acta pharmaceutica Sinica. B, 12(1), 378–393. https://doi.org/10.1016/j.apsb.2021.06.005

842) Li, X., Zheng, J., Chen, S., Meng, F. D., Ning, J., & Sun, S. L. (2021). Oleandrin, a cardiac glycoside, induces immunogenic cell death via the PERK/elF2α/ATF4/CHOP pathway in breast cancer. Cell death & disease, 12(4), 314. https://doi.org/10.1038/s41419-021-03605-y

843) Prieto, K., Cao, Y., Mohamed, E., Trillo-Tinoco, J., Sierra, R. A., Urueña, C., Sandoval, T. A., Fiorentino, S., Rodriguez, P. C., & Barreto, A. (2019). Polyphenol-rich extract induces apoptosis with immunogenic markers in melanoma cells through the ER stress-associated kinase PERK. Cell death discovery, 5, 134. https://doi.org/10.1038/s41420-019-0214-2

844) Wang, S., Wang, G., Wu, W., Xu, Z., Yang, J., Cao, M., Wang, Q., Wang, J., Yang, C., & Zhang, W. (2022). Autophagy activation by dietary piceatannol enhances the efficacy of immunogenic chemotherapy. Frontiers in immunology, 13, 968686. https://doi.org/10.3389/fimmu.2022.968686

845) Zhang, Y., Yang, S., Yang, Y., & Liu, T. (2019). Resveratrol induces immunogenic cell death of human and murine ovarian carcinoma cells. Infectious agents and cancer, 14, 27. https://doi.org/10.1186/s13027-019-0247-4

846) Lin, T. J., Lin, H. T., Chang, W. T., Mitapalli S, P., Hsiao, P. W., Yin, S. Y., & Yang, N. S. (2015). Shikonin-enhanced cell immunogenicity of tumor vaccine is mediated by the differential effects of DAMP components. Molecular cancer, 14, 174. https://doi.org/10.1186/s12943-015-0435-9

847) Jiang, W., Chen, M., Xiao, C., Yang, W., Qin, Q., Tan, Q., Liang, Z., Liao, X., Mao, A., & Wei, C. (2019). Triptolide Suppresses Growth of Breast Cancer by Targeting HMGB1 in Vitro and in Vivo. Biological & pharmaceutical bulletin, 42(6), 892–899. https://doi.org/10.1248/bpb.b18-00818

848) Zhou, P., Zheng, Z. H., Wan, T., Wu, J., Liao, C. W., & Sun, X. J. (2021). Vitexin Inhibits Gastric Cancer Growth and Metastasis through HMGB1-mediated Inactivation of the PI3K/AKT/mTOR/HIF-1α Signaling Pathway. Journal of gastric cancer, 21(4), 439–456. https://doi.org/10.5230/jgc.2021.21.e40

849) Sánchez-Giraldo, R., Acosta-Reyes, F. J., Malarkey, C. S., Saperas, N., Churchill, M. E., & Campos, J. L. (2015). Two high-mobility group box domains act together to underwind and kink DNA. Acta crystallographica. Section D, Biological crystallography, 71(Pt 7), 1423–1432. https://doi.org/10.1107/S1399004715007452

850) Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J., Honig, B., Shaw, D. E., & Friesner, R. A. (2004). A hierarchical approach to all-atom protein loop prediction. Proteins, 55(2), 351–367. https://doi.org/10.1002/prot.10613

851) Contini, A., Erba, E., Bondavalli, V., Barbiroli, A., Gelmi, M. L., & Romanelli, A. (2021). Morpholino-based peptide oligomers: Synthesis and DNA binding properties. Biochemical and biophysical research communications, 549, 8–13. https://doi.org/10.1016/j.bbrc.2021.02.087

852) Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., & Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science (New York, N.Y.), 220(4599), 868–871. https://doi.org/10.1126/science.6189183

853) Gallo, R. C., Sarin, P. S., Gelmann, E. P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V. S., Mann, D., Sidhu, G. D., Stahl, R. E., Zolla-Pazner, S., Leibowitch, J., & Popovic, M. (1983). Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). Science (New York, N.Y.), 220(4599), 865–867. https://doi.org/10.1126/science.6601823

854) Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Oleske, J., & Safai, B. (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science (New York, N.Y.), 224(4648), 500–503. https://doi.org/10.1126/science.6200936

855) Montagnier, L., Chermann, J. C., Barré-Sinoussi, F., Klatzmann, D., Wain-Hobson, S., Alizon, M., Clavel, F., Brun-Vezinet, F., Vilmer, E., & Rouzioux, C. (1984). Lymphadenopathy associated virus and its etiological role in AIDS. Princess Takamatsu symposia, 15, 319–331.

856) Chiu, I. M., Yaniv, A., Dahlberg, J. E., Gazit, A., Skuntz, S. F., Tronick, S. R., & Aaronson, S. A. (1985). Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses. Nature, 317(6035), 366–368. https://doi.org/10.1038/317366a0

857) Luciw, P. A. (1996). Human immunodeficiency virus and its replication. In B. N. Fields, D. M. Knipe, & P. M. Howley (Eds.), Fields virology (3rd ed., pp. 1881–1952). Lippincott-Raven

858) Fanales-Belasio, E., Raimondo, M., Suligoi, B., & Buttò, S. (2010). HIV virology and pathogenetic mechanisms of infection: a brief overview. Annali dell'Istituto superiore di sanita, 46(1), 5–14. https://doi.org/10.4415/ANN_10_01_02

859) Klimas, N., Koneru, A. O., & Fletcher, M. A. (2008). Overview of HIV. Psychosomatic medicine, 70(5), 523–530. https://doi.org/10.1097/PSY.0b013e31817ae69f

860) World Health Organization. HIV. https://www.who.int/data/gho/data/themes/hiv-aids

861) Lekkerkerker, A. N., van Kooyk, Y., & Geijtenbeek, T. B. (2006). Viral piracy: HIV-1 targets dendritic cells for transmission. Current HIV research, 4(2), 169–176. https://doi.org/10.2174/157016206776055020.

862) Chitwood, D. D., McCoy, C. B., Inciardi, J. A., McBride, D. C., Comerford, M., Trapido, E., McCoy, H. V., Page, J. B., Griffin, J., & Fletcher, M. A. (1990). HIV seropositivity of needles from shooting galleries in south Florida. American journal of public health, 80(2), 150–152.

https://doi.org/10.2105/ajph.80.2.150

863) Shaw, G. M., & Hunter, E. (2012). HIV transmission. Cold Spring Harbor perspectives in medicine, 2(11), a006965. https://doi.org/10.1101/cshperspect.a006965

864) Menéndez-Arias L. (2002). Targeting HIV: antiretroviral therapy and development of drug resistance. Trends in pharmacological sciences, 23(8), 381–388. https://doi.org/10.1016/s0165-6147(02)02054-0

865) Sarafianos, S. G., Marchand, B., Das, K., Himmel, D. M., Parniak, M. A., Hughes, S. H., & Arnold, E. (2009). Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. Journal of molecular biology, 385(3), 693–713. https://doi.org/10.1016/j.jmb.2008.10.071

866) Preston, B. D., Poiesz, B. J., & Loeb, L. A. (1988). Fidelity of HIV-1 reverse transcriptase. Science (New York, N.Y.), 242(4882), 1168–1171. https://doi.org/10.1126/science.2460924

867) Ho D. D. (1997). Perspectives series: host/pathogen interactions. Dynamics of HIV-1 replication in vivo. The Journal of clinical investigation, 99(11), 2565–2567. https://doi.org/10.1172/JCI119443

868) Smith, D. M., Richman, D. D., & Little, S. J. (2005). HIV superinfection. The Journal of infectious diseases, 192(3), 438–444. https://doi.org/10.1086/431682

869) IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. (2012). Human immunodeficiency virus-1. In Biological agents (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, No. 100B). International Agency for Research on Cancer. https://www.ncbi.nlm.nih.gov/books/NBK304351

870) Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., & Baumeister, K. (1985). Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature, 313(6000), 277–284. https://doi.org/10.1038/313277a0

871) Wain-Hobson S. (1989). HIV genome variability in vivo. AIDS (London, England), 3 Suppl 1, S13–S18. https://doi.org/10.1097/00002030-198901001-00003

872) Cassan, E., Arigon-Chifolleau, A. M., Mesnard, J.-M., Gross, A., & Gascuel, O. (2016). Concomitant emergence of the antisense protein gene of HIV-1 and of the pandemic. Proceedings of the National Academy of Sciences of the United States of America, 113(41), 11537–11542. https://doi.org/10.1073/pnas.1605739113.

873) Sertznig, H., Hillebrand, F., Erkelenz, S., Schaal, H., & Widera, M. (2018). Behind the scenes of HIV-1 replication: Alternative splicing as the dependency factor on the quiet. Virology, 516, 176–188. https://doi.org/10.1016/j.virol.2018.01.011.

874) German Advisory Committee Blood (Arbeitskreis Blut), Subgroup 'Assessment of Pathogens Transmissible by Blood' (2016). Human Immunodeficiency Virus (HIV). Transfusion medicine and hemotherapy: offizielles Organ der Deutschen Gesellschaft fur Transfusionsmedizin und Immunhamatologie, 43(3), 203–222. https://doi.org/10.1159/000445852.

875) Turner, B. G., & Summers, M. F. (1999). Structural biology of HIV-1. Journal of Molecular Biology, 285(1), 1–32. https://doi.org/10.1006/jmbi.1998.2354

876) Burnie, J., & Guzzo, C. (2019). The incorporation of host proteins into the external HIV-1 envelope. Viruses, 11(1), 85. https://doi.org/10.3390/v11010085

877) Zhu, P., Liu, J., Bess, J., Jr, Chertova, E., Lifson, J. D., Grisé, H., Ofek, G. A., Taylor, K. A., & Roux, K. H. (2006). Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature, 441(7095), 847–852. https://doi.org/10.1038/nature04817

878) Bukrinskaya A. G. (2004). HIV-1 assembly and maturation. Archives of virology, 149(6), 1067–1082. https://doi.org/10.1007/s00705-003-0281-8

879) Crispin, M., Ward, A. B., & Wilson, I. A. (2018). Structure and Immune Recognition of the HIV Glycan Shield. Annual review of biophysics, 47, 499–523. https://doi.org/10.1146/annurev-biophys-060414-034156

880) Briggs, J. A. G., Simon, M. N., Gross, I., Kräusslich, H. G., Fuller, S. D., Vogt, V. M., & Johnson, M. C. (2004). The stoichiometry of Gag protein in HIV-1. Nature Structural & Molecular Biology, 11(7), 672–675. https://doi.org/10.1038/nsmb785.

881) Freed, E. O., & Martin, M. A. (2007). HIVs and their replication. In B. N. Fields, D. M. Knipe, & P. M. Howley (Eds.), Fields virology (5th ed., pp. 2107–2186). Wolters Kluwer Health / Lippincott Williams & Wilkins.

882) https://scienceofhiv.org/life-cycle

883) Alkhatib, G., & Berger, E. A. (2007). HIV coreceptors: from discovery and designation to new paradigms and promise. European journal of medical research, 12(9), 375–384.

884) Engelman, A., & Cherepanov, P. (2012). The structural biology of HIV-1: Mechanistic and therapeutic insights. Nature Reviews Microbiology, 10(4), 279–290. https://doi.org/10.1038/nrmicro2747

885) Kong, R., Xu, K., Zhou, T., Acharya, P., Lemmin, T., Liu, K., Ozorowski, G., Soto, C., Taft, J. D., Bailer, R. T., et al. (2016). Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. Science, 352(6287), 828–833. https://doi.org/10.1126/science.aae0474.

886) Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., & Hendrickson, W. A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature, 393(6686), 648–659. https://doi.org/10.1038/31405.

887) Gallo, S. A., Finnegan, C. M., Viard, M., Raviv, Y., Dimitrov, A., Rawat, S. S., Puri, A., Durell, S., & Blumenthal, R. (2003). The HIV Env-mediated fusion reaction. Biochimica et Biophysica Acta (Biomembranes), 1614(1), 36–50. https://doi.org/10.1016/S0005-2736(03)00161-5

888) Zila, V., Margiotta, E., Turoňová, B., Müller, T. G., Zimmerli, C. E., Mattei, S., Allegretti, M., Börner, K., Rada, J., Müller, B., et al. (2021). Cone-shaped HIV-1 capsids are transported through intact nuclear pores. Cell, 184(4), 1032–1046.e18. https://doi.org/10.1016/j.cell.2021.01.025.

889) Dharan, A., Bachmann, N., Talley, S., Zwikelmaier, V., & Campbell, E. M. (2020). Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. Nature Microbiology, 5(8), 1088–1095. https://doi.org/10.1038/s41564-020-0735-8

890) Davis, A. J., Carr, J. M., Bagley, C. J., Powell, J., Warrilow, D., Harrich, D., Burrell, C. J., & Li, P. (2008). Human immunodeficiency virus type-1 reverse transcriptase exists as post-translationally modified forms in virions and cells. Retrovirology, 5, 115. https://doi.org/10.1186/1742-4690-5-115.

891) Onafuwa-Nuga, A., & Telesnitsky, A. (2009). The remarkable frequency of human immunodeficiency virus type 1 genetic recombination. Microbiology and Molecular Biology Reviews, 73(3), 451–480. https://doi.org/10.1128/MMBR.00012-09

892) Pommier, Y., Pilon, A. A., Bajaj, K., Mazumder, A., & Neamati, N. (1997). HIV-1 integrase as a target for antiviral drugs. Antiviral Chemistry & Chemotherapy, 8(6), 463–483. https://doi.org/10.1177/095632029700800601

893) Ruelas, D. S., & Greene, W. C. (2013). An integrated overview of HIV-1 latency. Cell, 155(3), 519–529. https://doi.org/10.1016/j.cell.2013.09.044.

894) Verdikt, R., Hernalsteens, O., & Van Lint, C. (2021). Epigenetic mechanisms of HIV-1 persistence. Vaccines, 9(5), 514. https://doi.org/10.3390/vaccines9050514.

895) Karn, J., & Stoltzfus, C. M. (2012). Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harbor Perspectives in Medicine, 2(2), a006916. https://doi.org/10.1101/cshperspect.a006916.

896) Dutilleul, A., Rodari, A., & Van Lint, C. (2020). Depicting HIV-1 transcriptional mechanisms: A summary of what we know. Viruses, 12(12), 1385. https://doi.org/10.3390/v12121385

897) Murphy, R. E., & Saad, J. S. (2020). The interplay between HIV-1 Gag binding to the plasma membrane and Env incorporation. Viruses, 12(5), 548. https://doi.org/10.3390/v12050548

898) Murakami, T., Ablan, S., Freed, E. O., & Tanaka, Y. (2004). Regulation of human immunodeficiency virus type 1 Env-mediated membrane fusion by viral protease activity. Journal of Virology, 78(2), 1026–1031. https://doi.org/10.1128/JVI.78.2.1026-1031.2004

899) Arts, E. J., & Hazuda, D. J. (2012). HIV-1 antiretroviral drug therapy. Cold Spring Harbor Perspectives in Medicine, 2(4), a007161. https://doi.org/10.1101/cshperspect.a007161

900) Kemnic, T. R., & Gulick, P. G. (n.d.). HIV antiretroviral therapy. Retrieved October 21, 2021, from http://www.ncbi.nlm.nih.gov/pubmed/30020680

901) Ghosh, R. K., Ghosh, S. M., & Chawla, S. (2011). Recent advances in antiretroviral drugs. Expert Opinion on Pharmacotherapy, 12(1), 31–46. https://doi.org/10.1517/14656566.2010.509345

902) Adamson, C. S. (2012). Protease-mediated maturation of HIV: Inhibitors of protease and the maturation process. Molecular Biology International, 2012, 604261. https://doi.org/10.1155/2012/604261

903) Jamjian, M. C., & McNicholl, I. R. (2004). Enfuvirtide: First fusion inhibitor for treatment of HIV infection. American Journal of Health-System Pharmacy, 61(12), 1242–1247. https://doi.org/10.1093/ajhp/61.12.1242

904) Qi, B., Fang, Q., Liu, S., Hou, W., Li, J., Huang, Y., & Shi, J. (2020). Advances of CCR5 antagonists: From small molecules to macromolecules. European Journal of Medicinal Chemistry, 208, 112819. https://doi.org/10.1016/j.ejmech.2020.112819

905) Meanwell, N. A., Krystal, M. R., Nowicka-Sans, B., Langley, D. R., Conlon, D. A., Eastgate, M. D., Grasela, D. M., Timmins, P., Wang, T., & Kadow, J. F. (2018). Inhibitors of HIV-1 attachment: The discovery and development of temsavir and its prodrug fostemsavir. Journal of Medicinal Chemistry, 61(1), 62–80. https://doi.org/10.1021/acs.jmedchem.7b01337

906) Beccari, M. V., Mogle, B. T., Sidman, E. F., Mastro, K. A., Asiago-Reddy, E., & Kufel, W. D. (2019). Ibalizumab, a novel monoclonal antibody for the management of multidrug-resistant HIV-1 infection. Antimicrobial Agents and Chemotherapy, 63(3), e00110–e00119. https://doi.org/10.1128/AAC.00110-19

907) https://hivinfo.nih.gov/

908) Maenza, J., & Flexner, C. (1998). Combination antiretroviral therapy for HIV infection. American Family Physician, 57(12), 2789–2798.

909) Warnke, D., Barreto, J., & Temesgen, Z. (2007). Therapeutic review: Antiretroviral drugs. Journal of Clinical Pharmacology, 47(12), 1570–1579. https://doi.org/10.1177/0091270007308034

910) Panel on Antiretroviral Guidelines for Adults and Adolescents. (2021). Guidelines for the use of antiretroviral agents in adults and adolescents with HIV. U.S. Department of Health and Human Services. https://clinicalinfo.hiv.gov/sites/default/files/guidelines/archive/AdultandAdolescentGL_2021_08_16.pdf

911) Palella, F. J., Delaney, K. M., Moorman, A. C., Loveless, M. O., Fuhrer, J., Satten, G. A., Aschman, D. J., & Holmberg, S. D. (1998). Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. New England Journal of Medicine, 338(13), 853–860. https://doi.org/10.1056/NEJM199803263381301

912) Moore, R. D., & Chaisson, R. E. (1999). Natural history of HIV infection in the era of combination antiretroviral therapy. AIDS, 13(14), 1933–1942. https://doi.org/10.1097/00002030-199910010-00017

913) Cohen, M. S., Chen, Y. Q., McCauley, M., Gamble, T., Hosseinipour, M. C., Kumarasamy, N., Hakim, J. G., Kumwenda, J., Grinsztejn, B., Pilotto, J. H., Godbole, S. V., Mehendale, S., Chariyalertsak, S., Santos, B. R., Mayer, K. H., Hoffman, I. F., Eshleman, S. H., Piwowar-Manning, E., Wang, L., … HPTN 052 Study Team. (2011). Prevention of HIV-1 infection with early antiretroviral therapy. New England Journal of Medicine, 365(6), 493–505. https://doi.org/10.1056/NEJMoa1105243

914) Cohen, M. S., Chen, Y. Q., McCauley, M., Gamble, T., Hosseinipour, M. C., Kumarasamy, N., Hakim, J. G., Kumwenda, J., Grinsztejn, B., Pilotto, J. H., Godbole, S. V., Chariyalertsak, S., Santos, B. R., Mayer, K. H., Hoffman, I. F., Eshleman, S. H., Piwowar-Manning, E., Cottle, L., Zhang, X. C., Makhema, J., … HPTN 052 Study Team. (2016). Antiretroviral therapy for the prevention of HIV-1 transmission. New England Journal of Medicine, 375(9), 830–839. https://doi.org/10.1056/NEJMoa1600693

915) Wu, J., Yan, P., & Archibald, C. (2007). Modelling the evolution of drug resistance in the presence of antiviral drugs. BMC Public Health, 7, 300. https://doi.org/10.1186/1471-2458-7-300

916) Mohri, H., Singh, M. K., Ching, W. T. W., & Ho, D. D. (1993). Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the blood of treated and untreated patients. Proceedings of the National Academy of Sciences, 90(1), 25–29. https://doi.org/10.1073/pnas.90.1.25

917) Kirschner, D. E., & Webb, G. F. (1997). Understanding drug resistance for monotherapy treatment of HIV infection. Bulletin of Mathematical Biology, 59(4), 763–785. https://doi.org/10.1007/BF02458429

918) Pennings, P. S. (2013). HIV drug resistance: Problems and perspectives. Infectious Disease Reports, 5(S1), e5. https://doi.org/10.4081/idr.2013.s1.e5

919) Richman, D. D. (1993). HIV drug resistance. Annual Review of Pharmacology and Toxicology, 33, 149–164. https://doi.org/10.1146/annurev.pa.33.040193.001053

920) Hunt, P. W., Martin, J. N., Sinclair, E., Bredt, B., Hagos, E., Lampiris, H., & Deeks, S. G. (2003). T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency virus-infected patients with sustained viral suppression during antiretroviral therapy. The Journal of Infectious Diseases, 187(10), 1534–1543. https://doi.org/10.1086/374786

921) Jain, V., Hartogensis, W., Bacchetti, P., Hunt, P. W., Hatano, H., Sinclair, E., Epling, L., Lee, T. H., Busch, M. P., McCune, J. M., Pilcher, C. D., Hecht, F. M., & Deeks, S. G. (2013). Antiretroviral therapy initiated within 6 months of HIV infection is associated with lower T-cell activation and smaller HIV reservoir size. The Journal of Infectious Diseases, 208(8), 1202–1211. https://doi.org/10.1093/infdis/jit311

922) Strategies for Management of Antiretroviral Therapy (SMART) Study Group, El-Sadr, W. M., Lundgren, J., Neaton, J. D., Gordin, F., Abrams, D., Arduino, R. C., Babiker, A., Burman, W., Clumeck, N., Cohen, C. J., Cohn, D., Cooper, D., Darbyshire, J., Emery, S., Fätkenheuer, G., Gazzard, B., Grund, B., Hoy, J., Klingman, K., … Rappoport, C. (2006). CD4+ count-guided interruption of antiretroviral treatment. The New England journal of medicine, 355(22), 2283–2296. https://doi.org/10.1056/NEJMoa062360

923) Thomas, R., Galanakis, C., Vézina, S., Longpré, D., Boissonnault, M., Huchet, E., Charest, L., Murphy, D., Trottier, B., & Machouf, N. (2015). Adherence to post-exposure prophylaxis (PEP) and incidence of HIV seroconversion in a major North American cohort. PLoS ONE, 10, e0142534. https://doi.org/10.1371/journal.pone.0142534

924) Rey, D. (2011). Post-exposure prophylaxis for HIV infection. Expert Review of Anti-Infective Therapy, 9(4), 431–442. https://doi.org/10.1586/eri.11.20

925) Eakle, R., Venter, F., & Rees, H. (2018). Pre-exposure prophylaxis (PrEP) in an era of stalled HIV prevention: Can it change the game? Retrovirology, 15, 29. https://doi.org/10.1186/s12977-018-0408-3

926) Dau, B., & Holodniy, M. (2009). Novel targets for antiretroviral therapy: Clinical progress to date. Drugs, 69, 31–50. https://doi.org/10.2165/00003495-200969010-00003

927) Cunha, R. F., Simões, S., Carvalheiro, M., Pereira, J. M. A., Costa, Q., & Ascenso, A. (2021). Novel antiretroviral therapeutic strategies for HIV. Molecules, 26, 5305. https://doi.org/10.3390/molecules26175305

928) Cutrell, A. G., Schapiro, J. M., Perno, C. F., Kuritzkes, D. R., Quercia, R., Patel, P., Polli, J. W., Dorey, D., Wang, Y., Wu, S., … et al. (2021). Exploring predictors of HIV-1 virologic failure to long-acting cabotegravir and rilpivirine: A multivariable analysis. AIDS, 35, 1333–1342. https://doi.org/10.1097/QAD.0000000000002883

929) Rizzardini, G., Overton, E. T., Orkin, C., Swindells, S., Arasteh, K., Górgolas Hernández-Mora, M., Pokrovsky, V., Girard, P.-M., Oka, S., Andrade-Villanueva, J. F., … et al. (2020). Long-acting injectable cabotegravir + rilpivirine for HIV maintenance therapy: Week 48 pooled analysis of Phase 3 ATLAS and FLAIR trials. JAIDS Journal of Acquired Immune Deficiency Syndromes, 85, 498–506. https://doi.org/10.1097/QAI.0000000000002466

930) Li, X.-D., Liu, L., & Cheng, L. (2018). Identification of thienopyridine carboxamides as selective binders of HIV-1 trans activation response (TAR) and Rev response element (RRE) RNAs. Organic & Biomolecular Chemistry, 16, 9191–9196. https://doi.org/10.1039/C8OB02753F

931) Chavali, S. S., Mali, S. M., Jenkins, J. L., Fasan, R., & Wedekind, J. E. (2020). Co-crystal structures of HIV TAR RNA bound to lab-evolved proteins show key roles for arginine relevant to the design of cyclic peptide TAR inhibitors. Journal of Biological Chemistry, 295, 16470–16486. https://doi.org/10.1074/jbc.RA120.015444

932) Melidis, L., Styles, I. B., & Hannon, M. J. (2021). Targeting structural features of viral genomes with a nano-sized supramolecular drug. Chemical Science, 12, 7174–7184. https://doi.org/10.1039/D1SC00933H

933) Sosic, A., Olivato, G., Carraro, C., Göttlich, R., Fabris, D., & Gatto, B. (2021). Bis-3-chloropiperidines targeting TAR RNA as a novel strategy to impair the HIV-1 nucleocapsid protein. Molecules, 26, 1874. https://doi.org/10.3390/molecules26071874

934) Li, Y., Liu, M., Chen, L. F., & Chen, R. (2018). P-TEFb: Finding its ways to release promoter-proximally paused RNA polymerase II. Transcription, 9, 88–94. https://doi.org/10.1080/21541264.2017.1281864

935) Dai, Y., Wynn, J. E., Peralta, A. N., Sherpa, C., Jayaraman, B., Li, H., Verma, A., Frankel, A. D., Le Grice, S. F., & Santos, W. L. (2018). Discovery of a branched peptide that recognizes the Rev response element (RRE) RNA and blocks HIV-1 replication. Journal of Medicinal Chemistry, 61, 9611–9620. https://doi.org/10.1021/acs.jmedchem.8b01076

936) Medina-Trillo, C., Sedgwick, D. M., Herrera, L., Beltrán, M., Moreno, Á., Barrio, P., Bedoya, L. M., Alcamí, J., Fustero, S., & Gallego, J. (2020). Nucleic acid recognition and antiviral activity of 1,4-substituted terphenyl compounds mimicking all faces of the HIV-1 Rev protein positively-charged α-helix. Scientific Reports, 10, 7190. https://doi.org/10.1038/s41598-020-64120-2

937) Dietz, J., Koch, J., Kaur, A., Raja, C., Stein, S., Grez, M., Pustowka, A., Mensch, S., Ferner, J., Möller, L., Bannert, N., Tampé, R., Divita, G., Mély, Y., Schwalbe, H., & Dietrich, U. (2008). Inhibition of HIV-1 by a peptide ligand of the genomic RNA packaging signal Psi. ChemMedChem, 3(5), 749–755. https://doi.org/10.1002/cmdc.200700194

938) Warui, D. M., & Baranger, A. M. (2009). Identification of specific small molecule ligands for stem loop 3 ribonucleic acid of the packaging signal Ψ of human immunodeficiency virus-1. Journal of Medicinal Chemistry, 52, 5462–5473. https://doi.org/10.1021/jm900599v

939) Ingemarsdotter, C. K., Zeng, J., Long, Z., Lever, A. M. L., & Kenyon, J. C. (2018). An RNA-binding compound that stabilizes the HIV-1 gRNA packaging signal structure and specifically blocks HIV-1 RNA encapsidation. Retrovirology, 15, 25. https://doi.org/10.1186/s12977-018-0407-4

940) Robinson, H. L. (2018). HIV/AIDS vaccines: 2018. Clinical Pharmacology & Therapeutics, 104, 1062–1073. https://doi.org/10.1002/cpt.1208

941) Cohn, L. B., Chomont, N., & Deeks, S. G. (2020). The biology of the HIV-1 latent reservoir and implications for cure strategies. Cell Host & Microbe, 27, 519–530. https://doi.org/10.1016/j.chom.2020.03.014

942) Douek, D. C., Kwong, P. D., & Nabel, G. J. (2006). The rational design of an AIDS vaccine. Cell, 124, 677–681. https://doi.org/10.1016/j.cell.2006.02.005

943) Mona Sadat, L., Seyed Mehdi, S., & Amitis, R. (2018). HIV-1 immune evasion: The main obstacle toward a successful vaccine. Archives of Asthma, Allergy & Immunology, 2, 13–15. https://doi.org/10.29328/journal.aaai.1001013

944) Rolland, M. (2016). HIV-1 immune evasion—A threat to effective vaccines? Nature Medicine, 22, 580–581. https://doi.org/10.1038/nm.4119

945) Barouch, D. H. (2008). Challenges in the development of an HIV-1 vaccine. Nature, 455, 613–619. https://doi.org/10.1038/nature07352

946) Cohen, K. W., & Frahm, N. (2017). Current views on the potential for development of a HIV vaccine. Expert Opinion on Biological Therapy, 17, 295–303. https://doi.org/10.1080/14712598.2017.1282457

947) Zolla-Pazner, S. (2004). Identifying epitopes of HIV-1 that induce protective antibodies. Nature Reviews Immunology, 4, 199–210. https://doi.org/10.1038/nri1307

948) Pantophlet, R., & Burton, D. R. (2006). GP120: Target for neutralizing HIV-1 antibodies. Annual Review of Immunology, 24, 739–769. https://doi.org/10.1146/annurev.immunol.24.021605.090557

949) Mascola, J. R., & Montefiori, D. C. (2010). The role of antibodies in HIV vaccines. Annual Review of Immunology, 28, 413–444. https://doi.org/10.1146/annurev-immunol-030409-101256

950) Burton, D. R., & Hangartner, L. (2016). Broadly neutralizing antibodies to HIV and their role in vaccine design. Annual Review of Immunology, 34, 635–659. https://doi.org/10.1146/annurev-immunol-041015-055515.

951) Saphire, E. O., Parren, P. W. H. I., Pantophlet, R., Zwick, M. B., Morris, G. M., Rudd, P. M., Dwek, R. A., Stanfield, R. L., Burton, D. R., & Wilson, I. A. (2001). Crystal structure of a neutralizing human IgG against HIV-1: A template for vaccine design. Science, 293, 1155–1159. https://doi.org/10.1126/science.1061692

952) Binley, J. M., Wrin, T., Korber, B., Zwick, M. B., Wang, M., Chappey, C., Stiegler, G., Kunert, R., Zolla-Pazner, S., Katinger, H., … et al. (2004). Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. Journal of Virology, 78, 13232–13252. https://doi.org/10.1128/JVI.78.23.13232-13252.2004

953) Gray, E. S., Moore, P. L., Choge, I. A., Decker, J. M., Bibollet-Ruche, F., Li, H., Leseka, N., Treurnicht, F., Mlisana, K., Shaw, G. M., … et al. (2007). Neutralizing antibody responses in acute human immunodeficiency virus type 1 subtype C infection. Journal of Virology, 81, 6187–6196. https://doi.org/10.1128/JVI.00239-07

954) Ghosn, J., Taiwo, B., Seedat, S., Autran, B., & Katlama, C. (2018). HIV. Lancet, 392, 685–697. https://doi.org/10.1016/S0140-6736(18)31311-4

955) Tuerk, C., MacDougal, S., & Gold, L. (1992). RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase. Proceedings of the National Academy of

Sciences of the United States of America, 89(15), 6988–6992. https://doi.org/10.1073/pnas.89.15.6988

956) Chen, H., & Gold, L. (1994). Selection of high-affinity RNA ligands to reverse transcriptase: Inhibition of cDNA synthesis and RNase H activity. Biochemistry, 33(29), 8746–8756. https://doi.org/10.1021/bi00195a016

957) Schneider, D. J., Feigon, J., Hostomsky, Z., & Gold, L. (1995). High-affinity ssDNA inhibitors of the reverse transcriptase of type 1 human immunodeficiency virus. Biochemistry, 34(29), 9599–9610. https://doi.org/10.1021/bi00029a037

958) Burke, D. H., Scates, L., Andrews, K., & Gold, L. (1996). Bent pseudoknots and novel RNA inhibitors of type 1 human immunodeficiency virus (HIV-1) reverse transcriptase. Journal of Molecular Biology, 264(4), 650–666. https://doi.org/10.1006/jmbi.1996.0667

959) Kensch, O., Connolly, B. A., Steinhoff, H. J., McGregor, A., Goody, R. S., & Restle, T. (2000). HIV-1 reverse transcriptase-pseudoknot RNA aptamer interaction has a binding affinity in the low picomolar range coupled with high specificity. Journal of Biological Chemistry, 275(24), 18271–18278. https://doi.org/10.1074/jbc.M001309200

960) Andreola, M. L., Pileur, F., Calmels, C., Ventura, M., Tarrago-Litvak, L., Toulmé, J. J., & Litvak, S. (2001). DNA aptamers selected against the HIV-1 RNase H display in vitro antiviral activity. Biochemistry, 40(34), 10087–10094. https://doi.org/10.1021/bi0108599

961) Chaloin, L., Lehmann, M. J., Sczakiel, G., & Restle, T. (2002). Endogenous expression of a high-affinity pseudoknot RNA aptamer suppresses replication of HIV-1. Nucleic Acids Research, 30(18), 4001–4008. https://doi.org/10.1093/nar/gkf522

962) Joshi, P. J., North, T. W., & Prasad, V. R. (2005). Aptamers directed to HIV-1 reverse transcriptase display greater efficacy over small hairpin RNAs targeted to viral RNA in blocking HIV-1 replication. Molecular Therapy, 11(5), 677–686. https://doi.org/10.1016/j.ymthe.2005.01.013

963) Kissel, J. D., Held, D. M., Hardy, R. W., & Burke, D. H. (2007). Single-stranded DNA aptamer RT1t49 inhibits RT polymerase and RNase H functions of HIV type 1, HIV type 2, and SIVCPZ RTs. AIDS Research and Human Retroviruses, 23(5), 699–708. https://doi.org/10.1089/aid.2006.0262

964) DeStefano, J. J., & Nair, G. R. (2008). Novel aptamer inhibitors of human immunodeficiency virus reverse transcriptase. Oligonucleotides, 18(2), 133–144. https://doi.org/10.1089/oli.2008.0103

965) Michalowski, D., Chitima-Matsiga, R., Held, D. M., & Burke, D. H. (2008). Novel bimodular DNA aptamers with guanosine quadruplexes inhibit phylogenetically diverse HIV-1 reverse transcriptases. Nucleic Acids Research, 36(22), 7124–7135. https://doi.org/10.1093/nar/gkn891

966) Mei, H., Liao, J. Y., Jimenez, R. M., Wang, Y., Bala, S., McCloskey, C., Switzer, C., & Chaput, J. C. (2018). Synthesis and evolution of a threose nucleic acid aptamer bearing 7-deaza-7-substituted guanosine residues. Journal of the American Chemical Society, 140(17), 5706–5713. https://doi.org/10.1021/jacs.7b13031

967) Gruenke, P. R., Alam, K. K., Singh, K., & Burke, D. H. (2020). 2'-fluoro-modified pyrimidines enhance affinity of RNA oligonucleotides to HIV-1 reverse transcriptase. RNA, 26(11), 1667–1679. https://doi.org/10.1261/rna.077008.120

968) Dunn, M. R., McCloskey, C. M., Buckley, P., Rhea, K., & Chaput, J. C. (2020). Generating biologically stable TNA aptamers that function with high affinity and thermal stability. Journal of the American Chemical Society, 142(17), 7721–7724. https://doi.org/10.1021/jacs.0c00641

969) Allen, P., Worland, S., & Gold, L. (1995). Isolation of high-affinity RNA ligands to HIV-1 integrase from a random pool. Virology, 209(2), 327–336. https://doi.org/10.1006/viro.1995.1264

970) Ojwang, J. O., Buckheit, R. W., Pommier, Y., Mazumder, A., De Vreese, K., Esté, J. A., Reymen, D., Pallansch, L. A., Lackman-Smith, C., & Wallace, T. L. (1995). T30177, an oligonucleotide stabilized by an intramolecular guanosine octet, is a potent inhibitor of laboratory strains and clinical isolates of human immunodeficiency virus type 1. Antimicrobial Agents and Chemotherapy, 39(11), 2426–2435. https://doi.org/10.1128/AAC.39.11.2426

971) Mazumder, A., Neamati, N., Ojwang, J. O., Sunder, S., Rando, R. F., & Pommier, Y. (1996). Inhibition of the human immunodeficiency virus type 1 integrase by guanosine quartet structures. Biochemistry, 35(43), 13762–13771. https://doi.org/10.1021/bi960541u

972) de Soultrait, V. R., Lozach, P. Y., Altmeyer, R., Tarrago-Litvak, L., Litvak, S., & Andréola, M. L. (2002). DNA aptamers derived from HIV-1 RNase H inhibitors are strong anti-integrase agents. Journal of Molecular Biology, 324(2), 195–203. https://doi.org/10.1016/s0022-2836(02)01064-1

973) Pang, K. M., Castanotto, D., Li, H., Scherer, L., & Rossi, J. J. (2018). Incorporation of aptamers in the terminal loop of shRNAs yields an effective and novel combinatorial targeting strategy. Nucleic Acids Research, 46(1), e6. https://doi.org/10.1093/nar/gkx980

974) Rose, K. M., Alves Ferreira-Bravo, I., Li, M., Craigie, R., Ditzler, M. A., Holliger, P., & DeStefano, J. J. (2019). Selection of 2'-deoxy-2'-fluoroarabino nucleic acid (FANA) aptamers that bind HIV-1 integrase with picomolar affinity. ACS Chemical Biology, 14(10), 2166–2175. https://doi.org/10.1021/acschembio.9b00237

975) Wyatt, J. R., Vickers, T. A., Roberson, J. L., Buckheit, R. W., Jr, Klimkait, T., DeBaets, E., Davis, P. W., Rayner, B., Imbach, J. L., & Ecker, D. J. (1994). Combinatorially selected guanosine-quartet structure is a potent inhibitor of human immunodeficiency virus envelope-mediated cell fusion. Proceedings of the National Academy of Sciences of the United States of America, 91(4), 1356–1360. https://doi.org/10.1073/pnas.91.4.1356

976) N. Sayer, J. Ibrahim, K. Turner, A. Tahiri-Alaoui, W. James, Structural characterization of a 2′F-RNA aptamer that binds a HIV-1 SU glycoprotein, gp120, Biochemical and Biophysical Research Communications, Volume 293, Issue 3, 2002, Pages 924-931, ISSN 0006-291X, https://doi.org/10.1016/S0006-291X(02)00308-X.

977) Khati, M., Schüman, M., Ibrahim, J., Sattentau, Q., Gordon, S., & James, W. (2003). Neutralization of infectivity of diverse R5 clinical isolates of human immunodeficiency virus type 1 by gp120-binding 2'F-RNA aptamers. Journal of virology, 77(23), 12692–12698. https://doi.org/10.1128/jvi.77.23.12692-12698.2003

978) Smith, D., Collins, B. D., Heil, J., & Koch, T. H. (2003). Sensitivity and specificity of photoaptamer probes. Molecular & cellular proteomics: MCP, 2(1), 11–18. https://doi.org/10.1074/mcp.m200059-mcp200

979) Dey, A. K., Griffiths, C., Lea, S. M., & James, W. (2005). Structural characterization of an anti-gp120 RNA aptamer that neutralizes R5 strains of HIV-1. RNA (New York, N.Y.), 11(6), 873–884. https://doi.org/10.1261/rna.7205405

980) Cohen, C., Forzan, M., Sproat, B., Pantophlet, R., McGowan, I., Burton, D., & James, W. (2008). An aptamer that neutralizes R5 strains of HIV-1 binds to core residues of gp120 in the CCR5 binding site. Virology, 381(1), 46–54. https://doi.org/10.1016/j.virol.2008.08.025

981) Zhao, N., Pei, S. N., Parekh, P., Salazar, E., & Zu, Y. (2014). Blocking interaction of viral gp120 and CD4-expressing T cells by single-stranded DNA aptamers. The international journal of biochemistry & cell biology, 51, 10–18. https://doi.org/10.1016/j.biocel.2014.03.008

982) Mufhandu, H. T., Gray, E. S., Madiga, M. C., Tumba, N., Alexandre, K. B., Khoza, T., Wibmer, C. K., Moore, P. L., Morris, L., & Khati, M. (2012). UCLA1, a synthetic derivative of a gp120 RNA aptamer, inhibits entry of human immunodeficiency virus type 1 subtype C. Journal of virology, 86(9), 4989–4999. https://doi.org/10.1128/JVI.06893-11.

983) Zhou, J., Li, H., Zhang, J., Piotr, S., & Rossi, J. (2011). Development of cell-type specific anti-HIV gp120 aptamers for siRNA delivery. Journal of visualized experiments : JoVE, (52), 2954. https://doi.org/10.3791/2954.

984) Joubert, M. K., Kinsley, N., Capovilla, A., Sewell, B. T., Jaffer, M. A., & Khati, M. (2010). A modeled structure of an aptamer-gp120 complex provides insight into the mechanism of HIV-1 neutralization. Biochemistry, 49(28), 5880–5890. https://doi.org/10.1021/bi100301k

985) Zhou, J., Swiderski, P., Li, H., Zhang, J., Neff, C. P., Akkina, R., & Rossi, J. J. (2009). Selection, characterization and application of new RNA HIV gp 120 aptamers for facile delivery of Dicer substrate siRNAs into HIV infected cells. Nucleic acids research, 37(9), 3094–3109. https://doi.org/10.1093/nar/gkp185

986) Sepehri Zarandi, H., Behbahani, M., & Mohabatkar, H. (2020). In Silico Selection of Gp120 ssDNA Aptamer to HIV-1. SLAS discovery : advancing life sciences R & D, 25(9), 1087–1093. https://doi.org/10.1177/2472555220923331

987) Shehata, M., Casalino, L., Duquette, M., Chen, S., Flaherty, A., Villa, E., & Amaro, R. E. (2025). N-Glycans Modulate HIV-1 Env Conformational Plasticity. bioRxiv, 2025-03.

988) Hoinka J, Backofen R, Przytycka TM. AptaSUITE: A Full-Featured Bioinformatics Framework for the Comprehensive Analysis of Aptamers from HT-SELEX Experiments. Mol Ther Nucleic Acids. 2018 Jun 1;11:515-517. doi: 10.1016/j.omtn.2018.04.006.

989) Zuker M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic acids research, 31(13), 3406–3415. https://doi.org/10.1093/nar/gkg595

990) Waterman, M.S., & Smith, T.F. (1978). RNA secondary structure: a complete mathematical analysis. Bellman Prize in Mathematical Biosciences, 42, 257-266.

991) Zhang, Y., Xiong, Y., Yang, C., & Xiao, Y. (2024). 3dRNA/DNA: 3D Structure Prediction from RNA to DNA. Journal of molecular biology, 436(17), 168742. https://doi.org/10.1016/j.jmb.2024.168742

992) Xiong, Y., Zhang, Y., Wang, J., & Xiao, Y. (2023). Using 3dRNA/DNA for RNA and DNA 3D Structure Prediction and Evaluation. Current protocols, 3(5), e770. https://doi.org/10.1002/cpz1.770

993) Wang, J., Mao, K., Zhao, Y., Zeng, C., Xiang, J., Zhang, Y., & Xiao, Y. (2017). Optimization of RNA 3D structure prediction using evolutionary restraints of nucleotide-nucleotide interactions from direct coupling analysis. Nucleic acids research, 45(11), 6299–6309. https://doi.org/10.1093/nar/gkx386

994) Zhao, Y., Huang, Y., Gong, Z., Wang, Y., Man, J., & Xiao, Y. (2012). Automated and fast building of three-dimensional RNA structures. Scientific reports, 2, 734. https://doi.org/10.1038/srep00734

995) Wang, J., Wang, J., Huang, Y., & Xiao, Y. (2019). 3dRNA v2.0: An Updated Web Server for RNA 3D Structure Prediction. International journal of molecular sciences, 20(17), 4116. https://doi.org/10.3390/ijms20174116

996) Capriotti, E., Norambuena, T., Marti-Renom, M. A., & Melo, F. (2011). All-atom knowledge-based potential for RNA structure prediction and assessment. Bioinformatics (Oxford, England), 27(8), 1086–1093. https://doi.org/10.1093/bioinformatics/btr093

997) Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., & Mackerell, A. D., Jr (2010). CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. Journal of computational chemistry, 31(4), 671–690. https://doi.org/10.1002/jcc.21367

998) Phillips, J. C., Hardy, D. J., Maia, J. D. C., Stone, J. E., Ribeiro, J. V., Bernardi, R. C., Buch, R., Fiorin, G., Hénin, J., Jiang, W., McGreevy, R., Melo, M. C. R., Radak, B. K., Skeel, R. D., Singharoy, A., Wang, Y., Roux, B., Aksimentiev, A., Luthey-Schulten, Z., Kalé, L. V., … Tajkhorshid, E. (2020). Scalable molecular dynamics on CPU and GPU architectures with NAMD. The Journal of chemical physics, 153(4), 044130. https://doi.org/10.1063/5.0014475

999) Liu, H., & Hou, T. (2016). CaFE: a tool for binding affinity prediction using end-point free energy methods. Bioinformatics (Oxford, England), 32(14), 2216–2218. https://doi.org/10.1093/bioinformatics/btw215

1000) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proceedings of the National Academy of Sciences of the United States of America, 98(18), 10037–10041. https://doi.org/10.1073/pnas.181342398

1001) Scheurer, M., Rodenkirch, P., Siggel, M., Bernardi, R. C., Schulten, K., Tajkhorshid, E., & Rudack, T. (2018). PyContact: Rapid, Customizable, and Visual Analysis of Noncovalent Interactions in MD Simulations. Biophysical journal, 114(3), 577–583. https://doi.org/10.1016/j.bpj.2017.12.003