UNIVERSITÀ DI PARMA

# UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA
in "Scienze Chimiche"

CICLO XXXVI

## *Peptide Nucleic Acid (PNA)-mediated self-assembly for nano-biotechnology*

Coordinatore:
Chiar.mo Prof.ssa Alessia Bacchi

Tutore:
Chiar.mo Prof. Roberto Corradini

Co-tutore:
Dr. Alex Manicardi

Dottorando: Sabrina Capodaglio

Anni Accademici: 2020/2021-2022/2023

*To Irene*

*"Quanto manca alla vetta?*
*Tu sali e non pensarci"*
*Friedrich Nietzsche*

## *Table of Contents*

# General abstract

This PhD work focused on the use of Peptide Nucleic Acids (PNAs) as scaffold for nanofabrication. Using different approaches, PNA strands were employed as building blocks to create supramolecular self-assembled architectures with the final aim of producing biotechnological applications (*Figure 1*). The main part of the work was devoted to the assembly of new nanostructures with controlled positioning of functional moieties as described in *Chapter 2* and *Chapter 3*. The work described in *Chapter 2* and *Chapter 5* was carried out in collaboration with Prof. Hanadi Sleiman's group at McGill University in Montreal (CA), where I spent several months as training secondment in the frame of the MSCA-RISE Nano-OligoMed Project.

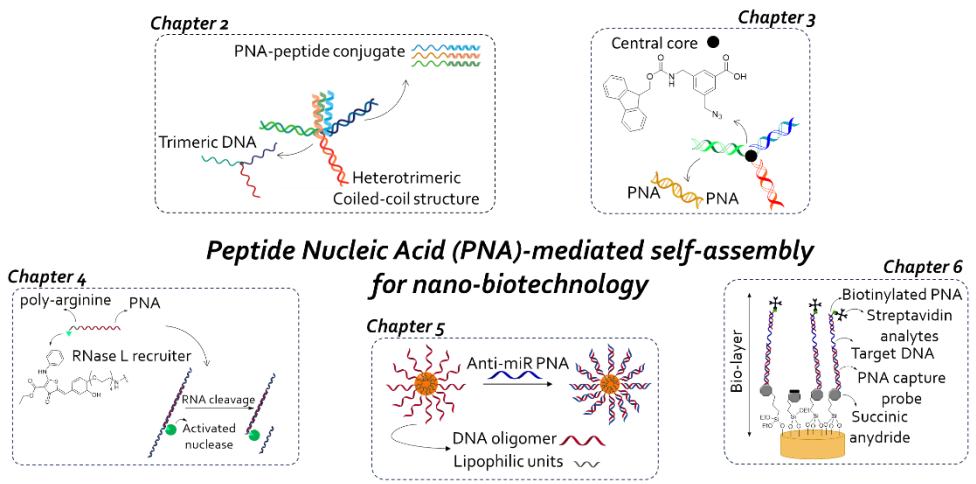



*Figure 1. Schematic representation of the five projects reported in this PhD thesis and relative chapters.*

*"Assembly of Heterotrimeric Coiled-Coil structures through DNA:PNA-Peptide interaction"* (*Chapter 2*) the creation of a multi-component platform able to drive the formation of heterotrimeric coiled-coil structures of short peptide domains through DNA:PNA interactions is described. The peptide sequences were designed to mimic the coordination site of plastocyanin (a small copper-binding protein with electron transfer properties), thus obtaining a mini-metalloprotein. Whereas *"Heterotrimeric*

*Coiled-Coil structures through PNA:PNA-Peptide interaction"* (*Chapter 3*), a similar assembly able to induce the formation of heterotrimeric coiled-coil system was designed exploiting only PNA-based probes. For this purpose, in *Chapter 3* is reported the synthesis of a tri-functionalized rigid molecule, designed to be used as a branching point enabling the synthesis of the trimeric PNA-template, and the study of the thermal stability of the PNA:PNA segments in order to minimize the probes length and obtain a stable supramolecular construct.

The ability of PNA to bind complementary nucleic acids forcing the interaction with other biological components is further developed in *Chapter 4*, by exploring the use of PNAs as novel approach in RNA-targeting therapies. In fact, in *"PNAs and Nuclease-recruiting PNAs as potential anti-COVID-19 approach"* (*Chapter 4*) a series of PNAs were synthetized to target different region of SARS-CoV-2 virus and further functionalized with a small molecule able to recruit RNase L in order to achieve a selective RNA cleavage.

In *"Design, assembly, and characterization of hybrid DNA/PNA nanospheres" (Chapter 5)* a different type of PNA-containing self-assembly is reported. In this case the assembly of nanospheres was obtained by conjugation of a polar DNA stretches with lipophilic oligomeric tails. The nanospheres thus obtained were decorated with PNA or PNA-peptide conjugates using a self-assembly process. The resulting PNA-containing nanospheres exhibited some modifications in properties, such as stability and resistance to degradation, compared to the DNA-only nanospheres. The major effort in this chapter was the study of the nanosphere stability following the introduction of charges on the PNA chain.

In the last chapter, *"DNA detection with Hollow Core optical fiber"* (*Chapter 6*), an example of PNA-mediated self-assembly for analytical purposes is described. Here, the PNA components were used as capture and reporting probes in label-free advanced detection schemes. In particular, in this chapter the very first data on the derivatization of a special class of micro-structured optical fibers with PNA using a newly developed lab-in-fiber optical biosensors response are described.

# *Chapter 1*

## *1.0 Introduction*

### 1.1 Nucleic Acids: DNA and RNA

Nucleic acids are biological macromolecules containing the genetic information of living organisms. The two principal nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), both of considerable importance in the biological field.

The DNA was isolated for the first time by Friedrich Miescher, a Swiss biochemist, in 1869 from with blood cells. However, the deoxyribonucleic acid is a biopolymer and, as all the polymers, it is composed by monomeric units (*Figure 1.1*). In specific, this monomeric unit are formed by five-membered-ring sugar (deoxyribose), a nitrogenous base and phosphate group and, it was called nucleotide. Four different nucleobases were determined in the DNA, explicitly adenine (A), cytosine (C), guanine (G) and thymine (T). Furthermore, these bases are divided into two groups, purines (A, G) and pyrimidines (C, T) and the purines are formed by double-ring bases instead the pyrimidines by single-ring bases. Regarding the monomeric unit, the nucleobase of every monomer is linked to 1′ carbon of the deoxyribose forming a glycosidic bond. A nucleotide is formed through a phosphodiester bond between a phosphate residue and nucleoside 5′ or 3′ (less frequently 2′) -OH group. The polymer DNA is formed by linkage of mononucleotides through 3′-phosphate-5′ phosphodiester bonds. The grow of DNA chain is therefore directional, and its sequence is normally written from 5′-terminous to 3′.[1]

The three-dimensional structure of deoxyribonucleic acid was unveiled by James D. Watson and Francis H. C. Crick in 1953[2], also with the help given by the data of X-ray diffraction obtained by Rosalind Franklin. These two researchers, an American biologist, and an English physicist respectively, have reported in their work that the DNA present a double helical structure where the two polynucleotide strands grown in opposite direction (anti-parallel orientation) around the same axis. Each chain is formed by mononucleotides linked by the 3′-5′ phosphodiester bonds, right-handed

double helical structure of DNA is caused by the capacity of the two complementary chains to interact together through hydrogen bonds formation. These non-covalent interactions, named Watson-Crick base pairing, are specific, since adenine can be paired with thymine (two hydrogen bonds are formed) and guanine can pair with cytosine (three hydrogen bonds are generated). The hydrogen bonds are individually weak forces but, taking into consideration the not usually short length of DNA double helix, their contribution to the thermodynamic stability is high. However, other important weak attractive forces that play a role for the stabilization of the DNA double helical structures are the van der Waals forces given by the stacking interaction (ππ) between adjacent bases. These stacking interactions are weaker between two pyrimidines (single-ring bases) and stronger between two purines (double-ring bases). Furthermore, in the most common DNA double helix conformation, called B-DNA, the nucleobases are positioned inside the structure and perpendicularly to the central axis and the sugar-phosphate backbone, being highly hydrophilic, is in the external part of the structure in contact with the aqueous environment forming two grooves, named minor and major groove.

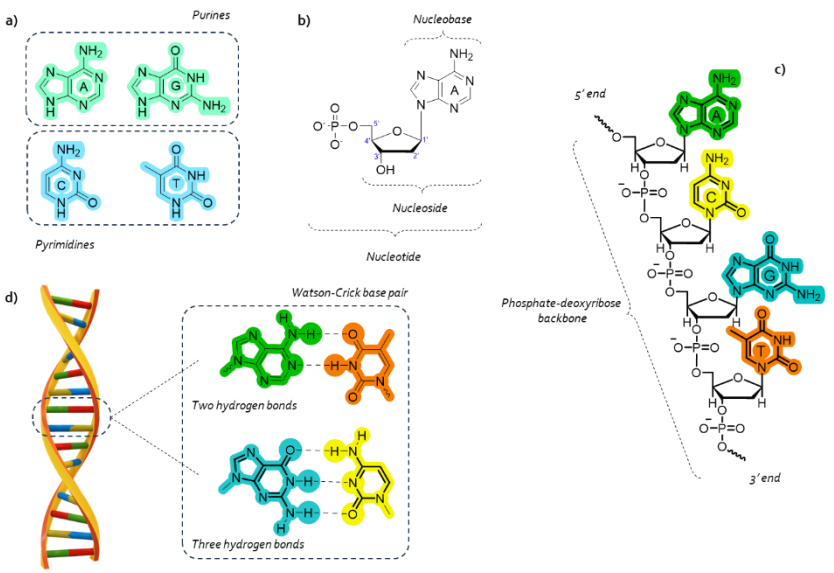



***Figure 1.1.*** *DNA structure a) Purines (adenine, A and guanine, G) and pyrimidines (cytosine, C and thymine, T) nucleobases. b) Nucleoside and nucleotide structure. c) DNA strand structure. d) Watson-Crick base pair interaction.*

The other previously mentioned nucleic acid is the ribonucleic acid (RNA) able to regulate the function of cells and genes in all the living organism. It presents some difference from DNA. In fact, as show in *Figure 1.2*, the uracil nucleobases took the place of thymine, while the sugar-phosphate backbone is formed by the sugar ribose and not deoxyribose. The difference in the chemical structure between the deoxyribose and the ribose is the presence of the hydroxyl group on the carbon in position 2′ which contributes to make the RNA chemically more unstable than the DNA, being cleavage of the phosphodiester bond in presence of alkali is more rapid.[3] Unlike DNA, that usually present in the double-stranded helix, the RNA is preferentially single polynucleotide chain. Furthermore, a network of several interactions leads to have different three-dimensional structures of RNA molecules. The principal building blocks of the secondary structure of RNA is the Watson–Crick base pairs, the stacking interaction, the mis-matched pairs, and the unpaired bases that together connect for the formation of different RNA secondary structures, like helical duplex, internal loops, hairpin loops, bulges or bulge loops, and junctions.[4] Consequently, the combination of these types of secondary structures in different way carries out principally kissing loops, pseudoknot region, multi-helical junctions triple-stranded, quadruplex and other as RNA tertiary structures.[5]

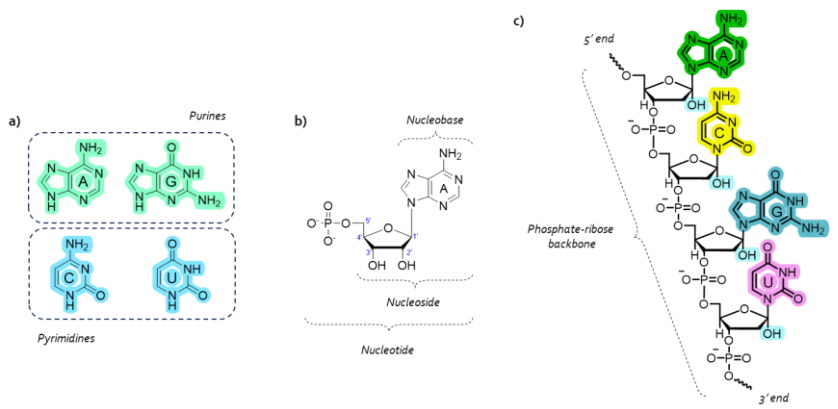



***Figure 1.2.*** *RNA structure a) Purines (adenine, A and guanine, G) and pyrimidines (cytosine, C and uracil, U) nucleobases. b) Nucleoside and nucleotide structure. c) RNA strand structure.*

Several types of RNA exist in the cell, the basic are messenger (mRNA), ribosomal (rRNA), transfer (tRNA), small nuclear RNA, and microRNA (miRNA).

### 1.1.1 DNA and RNA Nanotechnology

The discovery of the DNA and the subsequent study of its chemical and physical properties has led to the development of nanotechnology based on DNA, generating several structural motifs like tiles, multi-way junctions, DNA origami, cage structures and hybrid DNA-polymer. Nanometre-scale objects made of synthetic DNA were assembled to build DNA nanostructures, that have revolutionized the nanoscience and nanotechnology fields for technology and biological applications. This possibility to connect small building blocks engages non-covalent interactions including Watson-Crick base pairing (hydrogen bonding), hydrophobic interactions and van der Waals interactions, which can be considered as predictable and programmable interactions able to form symmetric and asymmetric nanoscale architectures with a specific shape and dimension.

The pioneer of the assembly of DNA strands for the construction of nanostructures was Nadrian C. Seeman in 1980. In fact, two years later, he published the realization of a scaffold based on branched junctions, named Holliday Junctions, of complementary DNA strands connected by sticky ends.[6] After that, a three-dimensional DNA cube was realized by specific branched sequences of oligomeric nucleic acids able to form in a directional Holliday junction and sticky ends connection.[7] To obtain a rigid scaffold and a simpler prediction of the DNA strands conformations, a DNA double-crossover (DX) molecule was designed. In this case, the single strands exchange present in the Holliday junction was replaced by two DNA double helices generated by sharing two DNA oligomers.[8] Over the years, the development of several DNA tiles (like double-crossover, triple-crossover, cross motif, four-helix, eight-helix, and twelve-helix planar tiles, parallelogram DNA junctions, three- or six-point star motifs, T-junctions) became the building blocks to the self-assembly of 2D and 3D networks based on DNA.[9,10] An example of 3D nanostructured was published by Nadrian Seeman, in specific he designed and self-assembled the first 3D crystal structure of DNA tensegrity triangle (~4 Å resolution). This DNA motif present a higher rigidity because three different DNA helices can

form the tensegrity triangle connected to each other in pairs by branched junctions.[11]

Two opposite methods to self-assembly symmetrical and asymmetrical nanostructure based on DNA were developed: *top-down* and *bottom-up*. Regarding the first method, the desired dimension was reached starting from huge structures, modified by synthetic DNA segments. Instead, in the *bottom-up* method the self-assembly of DNA nanostructures is obtained by designing the appropriate sequences of oligonucleotide sequences to drive the self-assembly process. This second method gave the possibility to obtain nanoscale objects even smaller size with respect the *top-down* method.[12]

In 2006, Paul W.K. Rothemund published a new method to the construction of nanostructures based on DNA, the *DNA origami* technique where short synthetic oligonucleotides used as *staple strands* can drive the folding of long DNA strand in a construction of a 2D scaffold with a specific crossover region and final shape. To design and synthetize the DNA origami a set of cylinders, used as representation of DNA double helix, are organized to create the desired shape structure. These helices are kept together with a crossover section. Then, a long single stranded scaffold was folded with the *staples strands* to generate the specific Watson-Crick interactions and to obtain, consequently, the desired 2D shapes. In the Rothemund's work, a genome of M13mp18 virus was used as a long single strand of DNA.[13] Subsequently, a similar approach to form DNA nanostructures was developed and it was named *single-stranded tile* (SST). In this method, a high number of short oligonucleotide strands can be self-assembled, without the necessity of a long DNA scaffold, to generate complex shape.[14]

These 2D planar structures can be assembly in 3D DNA origami with single-layered design in two different ways. In one the 2D planar structure and the staples are unpaired with the consequently formation of well-cut pieces (like triangles and squares) able to match together for 3D DNA origami assembly. In the other strategy, instead, the 3D DNA origami were generated with the introduction of a crossover and a linkage between the single-layers. Another type of single-layered design is the

wireframe single-layered design, based on the presence of DNA double-crossover molecules, where the organization of DNA double helices are in a compact and parallel motif. In addition, to enhance the rigidity of the 3D DNA origami, multilayered designs were developed, based on the connection between several single-layered DNA nanostructures.

Complex DNA origami structures with great size and molecular weight were assembled, using different approaches like extended scaffold, complex staples, hierarchical assembly, surface assisted and templated tile.

Moreover, small synthetic organic and inorganic molecules can be introduced inside the DNA strands to generate a building block for even more complex structural DNA nanostructures. Some of these small molecules can produce additional driving forces in the assembly of nanostructures in a directional and unique manner. For example, a 2D cyclic hexamer was synthetized using a rigid small organic molecule at the vertex position to confer directionality at this well-defined structure designed to include six DNA double stranded. In addition, this architecture was used as a scaffold for the arrangement of gold nanoparticles.[15]

Additionally, a benzene scaffold to construct a branched molecules able to link together three different oligonucleotide strands with threefold rotational axis and a horizontal plane of symmetry ($C_{3h}$ symmetry) was used by von Kiedrowski and colleagues. This organic linker, orthogonally protected and with a specific patterns of symmetry elements inside the structure, was employed for the construction of tetrahedral nanostructures.[16,17]

Branched nucleic acids can be synthesized using an asymmetric branching unit in solid-phase synthesis but could also be produced by a replica process (printing) starting from a preformed DNA template. Sleiman and co-workers used a templating 3-way junction to assemble three oligonucleotide strands in a defined order. The terminal part of these was then 'stapled' using a triazide with a benzene ring as central core to connect the three DNA strands in a covalent way using copper-catalyzed alkyne-azide "click" chemistry. The tripodal DNA obtained was used to

"print" these three DNA strands in a similar process by assembling the three complementary oligonucleotides and then 'stapling' them into a tripodal structure. In addition, both normal and reverse path of DNA strands synthesis were employed, and it is also possible to synthetize a scaffold with different lengths of DNA strands. Also, this branched template can be used as building blocks to create more complex 2D and 3D DNA nanostructures, exploiting methods enabling to transfer DNA strand patterns from a DNA scaffold to other materials.[18–20]

#### *1.1.1.1 Application of DNA nanostructures*

The evolution over the years of the design, manipulation, and assembly of DNA strands in nanostructures has led to their use in different fields (from chemistry, biochemistry, biomedicine, materials to engineering) and in several applications, like for example building blocks of molecular machines, drug delivery, smart materials able to respond at specific stimuli, detection of nucleic acids, and assembling of inorganic nanostructures or proteins.[12,21,22]

Over the years, the development of self-assembly and nanotechnology based on RNA strands has also gained interest. In fact, RNA unit was used as building blocks to synthetize 2D and 3D artificial nanoarchitectures.[23] RNA molecule can be considered like DNA molecule under the design and the manipulation to form complex nanostructure, it resembles proteins if the characteristics as flexibility and diversity in function are considered. RNA molecules can be self-assembled in different structural motifs taking advantage by their ability to form canonical and non-canonical base pairing.[24] To self-assemble RNA strands in nanometer-scale RNA architectures different methods were studied. RNA architectonics self-assembly, single-strand RNA assembly, RNA/DNA hybrid self-assembly, and co-transcriptional assembly are the main strategies used. Regarding the first strategies, a study of the intermolecular interactions in natural RNA molecules is carried out to generate complex 2D and subsequently 3D nanoarchitectures based on synthetic RNA. In the single-strand RNA assembly, a set of unstructured RNA strands were combined to

obtain complex architectures based on Watson-Crick base interactions and tail-tail interactions.[25] In the RNA/DNA hybrid self-assembly, instead, a long RNA scaffold was connect using several DNA staples. The last strategy was focused on the connection between the synthetic RNA strands with the natural RNA produced inside the cells.[26]

## 1.2 Peptide nucleic acid (PNA)

Peptide Nucleic Acids (PNAs) are compounds first described by Peter Nielsen and coworkers in 1991 at Copenhagen University, which are synthetic polyamide analogues of DNA.[27] The PNA structure, in fact, consists in the repetition of *N*-(2-aminoethyl)glycine units connected by amide bonds, which replace the deoxyribose or ribose phosphate backbone respectively of DNA or RNA (negatively charged). This change led to an uncharged backbone, thus avoiding phenomena of electrostatic repulsion upon hybridization. At the backbone, through methylenecarbonyl linker, the same purines (A, G) and pyrimidines (C, T) DNA nucleobases are linked (*Figure 1.3*). [28]

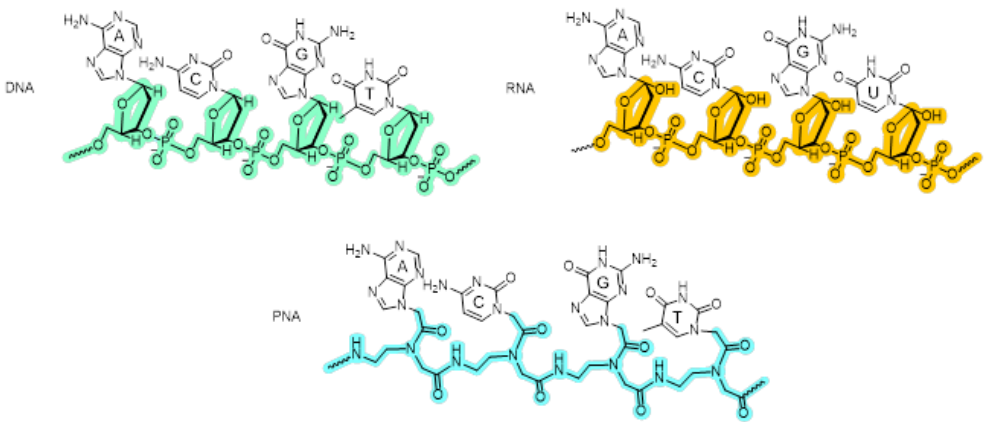



***Figure 1.3.*** *Structure of DNA, RNA and PNA, highlighting the difference in backbone.*

On the contrary of natural nucleic acid that are prone to degradation by nuclease due to their phosphate backbone, or peptides that are degraded by proteases and peptidases, both nuclease and protease are not able to degrade the PNA oligomers due to their unnatural structure. Furthermore, they have an excellent chemical and biochemical stability. In addition, since PNAs present an achiral pseudopeptide backbone, the enantiomeric issues are overcome, at least for unmodified PNAs.[29]

## 1.2.1 Properties of PNAs

### *1.2.1.1 PNA hybridizations with nucleic acid*

Originally PNA was conceived in order to recognize a double stranded DNA and to generate a DNA:PNA:DNA triplex through Hoogsteen hydrogen interactions. Surprisingly, they were shown instead to form very stable PNA:DNA:PNA triplexes. Later on, it was shown that the similarly in the length of repeating unit between the PNA and DNA or RNA backbone allowed to obtain very stable duplexes and triplexes with DNA or RNA oligomers.[30] This hybridization is based on Watson-Crick base pair interactions and these duplexes are more stable than the relative nucleic acid structures. The higher stability is attributed to the neutral PNA backbone which, unlike DNA and RNA, does not give rise to repulsive electrostatic interactions upon binding to complementary nucleic acids.[31] The polyamide backbone of PNAs, furthermore, makes the hybridization less dependent by the salt concentration.[32]

The PNA strands are capable to bind DNA or RNA oligomers both in parallel and antiparallel orientation.[33] Hybridization in *parallel* is by convention than in which the N-terminal moiety of PNA oligomer is paired with the 5'-end of the complementary DNA or RNA oligonucleotide and consequently the C-terminal moiety of PNA sequence is direct to the 3'-end of the corresponding DNA or RNA strand. On the contrary, in antiparallel hybridization the N-terminal moiety of PNA strand faces the 3'-end of complementary DNA or RNA (*Figure 1.4*). In general, the melting temperature of PNA/DNA antiparallel duplex is 1°C per base pair higher respect the DNA/DNA duplex,[34,35] and is reported to be more stable than the parallel binding.[36]

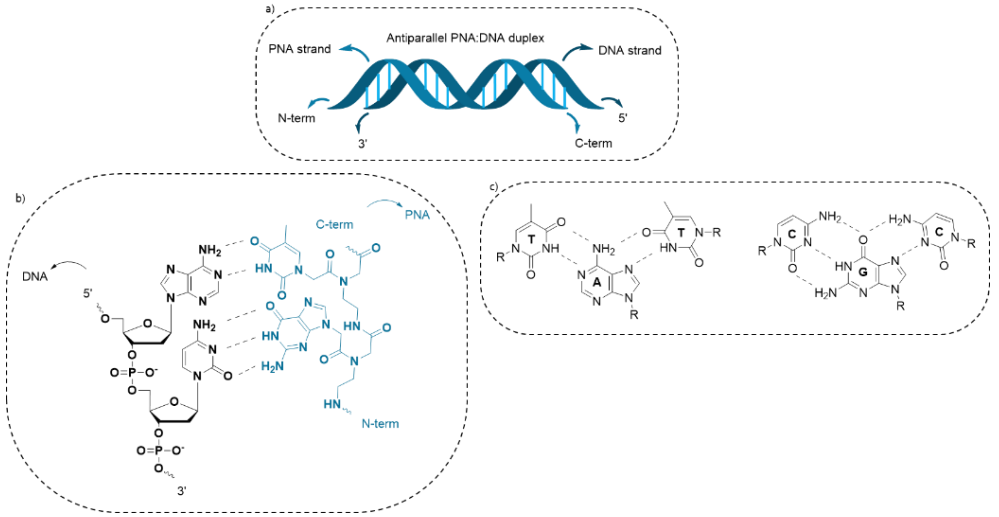



***Figure 1.4.*** *a) Antiparallel PNA-DNA hybridization; b) Watson-Crick interactions between DNA and PNA; ) Hoogsteen interactions.*

The thermal stability of PNA:RNA duplexes was reported to be higher than PNA:DNA duplexes. The conformation found for the helical structure formed between PNA and complementary RNA was similar to A-form, which is the preferred one for RNA at neutral pH and physiological salt concentrations and it is a right-handed duplex with a shorter distance between the base pairs.[31]

Furthermore, the PNA oligomers can hybridize with complementary nucleic acid sequences to generate (PNA)2-DNA triplex, thanks to the formation of Watson-Crick and Hoogsteen hydrogen bond interactions. This triplex with single stranded DNA can form when the PNAs present a high pyrimidine:purine ratio while the DNA or RNA is composed principally by purine bases.

Single-stranded PNA can interact with double stranded DNA (dsDNA) to generate different complexes based on oligomers compositions (ratio between pyrimidines and purines). In addition to PNA-dsDNA triplex, a strand displacement process can also be induced by PNA, giving rise to different processes, like duplex invasion, triplex invasion, double-duplex invasion, and tail-clamp. The triplex invasion complex was formed in presence of a homopyrimidine PNA strand that bind the double strand complementary nucleic acid through Watson-Crick hydrogen bonds and a second PNA oligomer replace the homopyrimidine DNA strand, forming a P-

loop, through Hoogsteen interactions. The ionic strength is a key factor in the formation of complexes by strand displacement process.[37,38]

Four-stranded of DNA oligomers composed by guanine bases units can form a G quadruplex structures under Watson-Crick base pairing and Hoogsteen interactions. The stability of these complexes can be also influenced by the presence of cations, like sodium and potassium. A G quadruplex based on the hybridization between DNA and PNA oligomers rich in guanine base units was reported. In fact, a $PNA_2$-$DNA_2$ quadruplex was formed by the hybridization of homologous G-rich PNA and DNA strands. This hybrid structure seems to be less dependent by the ionic strength than the corresponding G quadruplex composed by only DNA oligomers because the distance between the anionic phosphate backbone of different DNA strands is higher and consequently the electrostatic repulsions are lower.[39] After this study, the synthesis of G quadruplex structures based on PNA oligomers was expanded and different types of these structures were developed, like hybrid G quadruplex PNA-RNA[40] and G quadruplex based only PNAs.[41]

*1.2.1.2 Complexes based only on PNAs*

The fact that PNAs, artificial biopolymer, are able to bind complementary PNA sequences, forming stable duplexes that present double helical structures like DNA or RNA duplexes, demonstrated that the sugar-phosphate backbone of nucleic acids is not strictly necessary to form double helical architectures.[42]

Complexes based on nucleic acid are principally less stable than the corresponding duplexes formed by a PNA strand and its complementary PNA sequence because not only the electrostatic repulsion are missing, but also other stabilizing interactions are optimized.[43] PNA-PNA duplexes present a P-type helical structure with 18 base pairs per turn and these helices are around 28 Å wide. Both right- and left-handed helices can be found when duplexes are formed in the absence of any chiral bias. The two PNA strands are close, resulting in the formation of a tight minor groove and a large major groove.[44]

In addition, PNA:PNA duplex in antiparallel orientation was reported to be more stable than the corresponding parallel one.[45]

Also, triplex base only on PNA strands were studied and PNA-PNA-PNA structures generated with T-A-T triplets formed by Watson-Crick-Hoogsteen interactions was observed.[46] PNA was shown also to be able to form i-*motif* like natural nucleic acids.[47] The *i-motif*, that can be a building block to the construction of nanoscale architectures[48], is a motif constituted by cytosine-rich oligonucleotides. In specific, the *i-motifs* are generated by cytosine and protonated cytosine (C-$C^+$) base pairs in two parallel duplexes that then can be interleaved in an antiparallel orientation.[49]

Over the years, studies concerning the stability of the complexes based on PNAs were performed using melting temperature (Tm) measurements. In general, the Tm is the temperature where the 50% of the duplex are in the double helical state and the other 50% are in dissociated state. One important factor that strongly influence the melting temperatures of nucleic acid duplexes or PNA duplexes is the composition of the chains, given that the presence of a higher G-C base pairs respect A-T base pairs requires more energy to separate the double-helices, being higher the number of hydrogen bonds that they can be formed. Furthermore, these measurements can be carried out using both UV-Vis and circular dichroism (CD, for PNA complexes containing chiral centers), although the two results obtained by the two techniques might be different because the conformational rearrangement before the dissociation of the double helical may differ depending on the instrument used.[50]

#### *1.2.1.3 Solubility and stability of PNAs*

The charge-neutral backbone of PNA sequence led to this artificial nucleic acid to be poorly soluble in water if compared to the natural nucleic acids. However, the PNAs solubility depends on the composition of the sequences in terms of purine/pyrimidine ratio, on the length of the oligomers and on their tendency to form aggregates in solution.[28] While DNA duplexes are stable only in aqueous

solutions, the PNA-PNA duplexes are hydrophilic and stable in aqueous solutions, and being neutral, they also demonstrate good stability in aprotic polar solvents, such as dimethylformamide (DMF).

Furthermore, the PNAs solubility in water can be increased with the addition of positive charges on the backbone, or conjugation with positive charged lysine residues either C- or N-terminus. [51,52]

Compared to their natural counterparts (DNA and RNA), the PNAs exhibit remarkable chemical stability. In fact, their demonstrate higher stability properties over a wide range of pH.[28] PNAs are more stable in acidic environments, although they also exhibit moderate stability in basic conditions.[53]

### *1.2.2 Synthesis of PNAs by solid-phase synthesis (SPS)*

Peptide nucleic acid oligomers are synthetized using standard solid-phase manual or automatic synthesis (SPS), a synthetic methodology that was developed by Bruce Merrifield in the 1963 for peptide synthesis.[54] In fact, the solid-phase peptide synthesis (SPPS) enables growth of peptide chain on solid support, and as a result the separation of the synthetic product from by-products, that can be generated during the multistep synthesis, turn out to be simpler and more efficient. This easy separation is allowed because all the reactions occur in heterogeneous phase, using a solid support that is a co-polymer (generally formed by styrene and 1% of m-divinylbenzene as a cross-linking agent) completely insoluble in all the solutions for the synthesis and it can be isolated trough filtration.

Two principal strategies were developed for the synthesis of PNA oligomers, one based on tert-butyloxycarbonyl (Boc)/ benzyloxycarbonyl (Cbz) monomers and the other on 9-fluorenylmethoxycarbonyl (Fmoc)/ benzhydryloxycarbonyl (Bhoc) monomers (*Figure 1.5*). *Boc* and *Fmoc* are temporary protecting groups of the primary amino function of PNA monomers, instead *Cbz* and *Bhoc* are semi-permanent protecting groups of exocyclic amino function of the nucleobases used

to avoid side reaction and they are respectively orthogonal respect the *Fmoc* and *Boc* protecting groups.

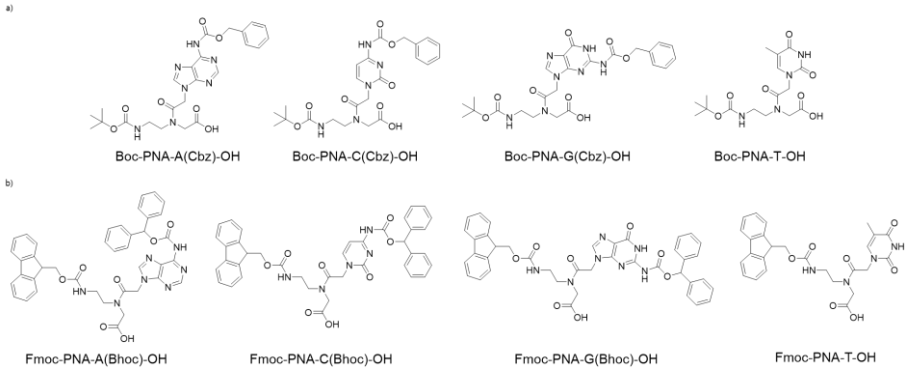



***Figure 1.5**. a) PNA monomers protected Boc/Cbz. b) Fmoc/Bhoc PNA monomers.*

The two strategies are different for the synthesis conditions given that the Boc protecting group is acid sensitive while the Fmoc is base sensitive, and thus the linking group to the solid support depends by the strategy employed. As an example, the methylbenzhydryl amine (MBHA), cleavable only under harsh acidic conditions, can be used as the linker for the solid support in *Boc* strategy, whereas the Rink Amide linker, which is cleavable under milder acidic conditions, can be employed in *Fmoc* strategy. The first step for the solid phase oligomers synthesis is the solid support pre-loaded with the first monomer or with an amino acid adequately protected. This allows also to control the density of the functional groups, and eventually to prevent excessive interactions between the growing peptide chains. As an example, 0.2 mmol/gr is the general loading employed for the PNA synthesis regardless of the chosen strategy.

Swelling of the solid support with apolar solvents, to increase the surface area of the solid support, is mandatory for a successful synthesis, since it ensures a better diffusion of the reagents and consequently a higher chemical conversion.

The chain growth in the SPS method consists of three principal reaction steps: deprotection, coupling and capping, that are carried out in a cyclic manner by

introducing the various PNA monomers, until obtaining the oligomer with the desired sequence.

The protocol is repeated cyclically until reaching the desired product, and after that the oligomer is cleaved from the solid support using appropriate cleavage cocktails that depends on the type of resin but also on the presence of sensitive residues. During the cleavage, also the semi-permanent protecting group of the exocyclic amino moiety of nucleobases are removed.

Precipitation of the PNA and purification by RP-IP HPLC provide pure PNA products. Assessment of the identity of the product can be obtained by mass spectrometry analysis. UV-Visible measurements are normally used for quantifying the final product.[55,56]

A schematic representation of PNA synthesis on solid phase support is reported in *Figure 1.6*.

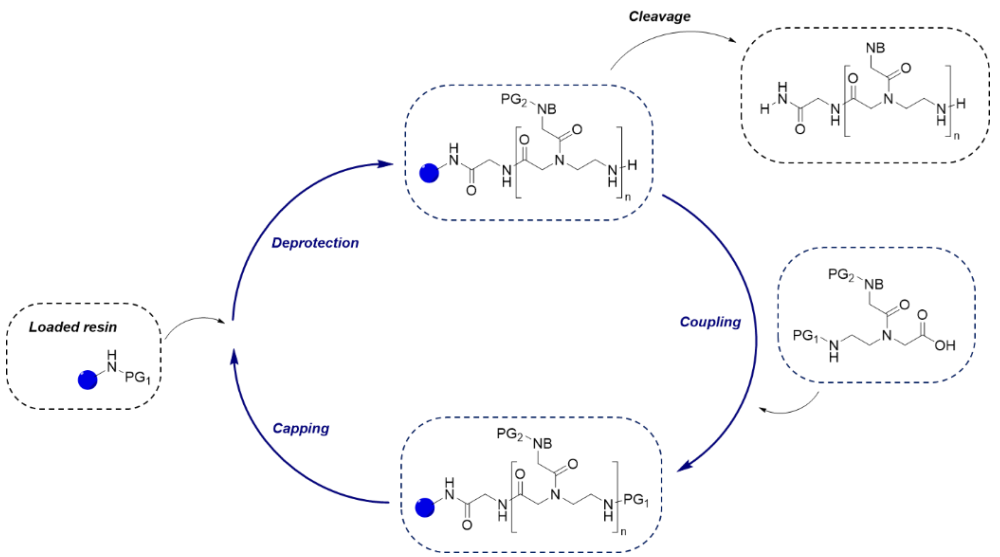



***Figure 1.6.*** *Schematic representation of PNA synthesis on solid phase synthesis.*

The chemistry of the various steps will be discussed in the following text.

Removal of Fmoc protective group is normally carried out under basic conditions a basic sensitive temporary protective group, a solution of 20% piperidine in DMF is normally used (*Figure 1.7*). Using Fmoc group, the UV-absorbance of deprotection

solution containing dibenzofulvene can be employed to evaluate the loading of the Fmoc-protected residue inserted, and thus evaluate the efficiency of incorporation of the monomer. Boc deprotection can be carried out under acidic conditions, typically by a solution of trifluoroacetic acid (TFA)/m-cresol (90:10). The m-cresol is an aromatic compound derived from phenol used as scavenger in the acid deprotection solution to avoid tert-butyl cation alkylate nucleophiles or aromatic groups.

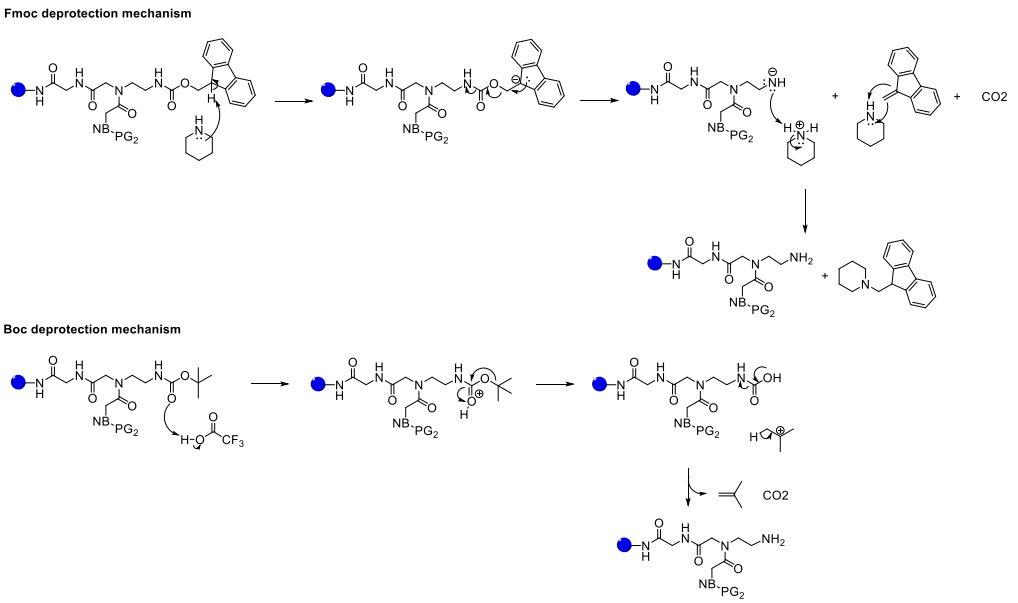



***Figure 1.7.*** *a) Fmoc deprotection mechanism, b) Boc deprotection mechanism.*

One crucial step to allow the growth of the chain is the coupling reaction which takes place after that the carboxylic acid of the entering monomer undergoes an activation step by a coupling agent to yield an active ester. Since the active esters formed in this step are sensitive to water, rigorous anhydrification of the solvent, usually *N,N*-dimethylformamide (DMF) or N-methylpyrrolidone (NMP) and elimination of nitrogen nucleophiles (e.g. dimethylamine traces) is very important for good overall yields of the synthesis.

In this step, it is important the choice of the coupling reagent and the principal classes of coupling reagents used are carbodiimides, uronium salt, and phosphonium salt.

Take into consideration the carbodiimides groups, able to active the carboxylic group *in situ* before the amide bond, dicyclohexylcarbodiimide (DCC), diisopropylcarbodiimide (DIC) and (N-(3-dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride (EDC·HCl) are the common coupling reagents used. Using DCC, N-N′-dicyclohexylurea is formed as precipitate during the progress of coupling reaction as by-product. This undesirable compound is insoluble in most organic solvents and, consequently, the use of DCC on solid phase synthesis appear to be inappropriate because the precipitation of this by-product would block the filtration of the solution. Therefore, DCC is principally used as coupling reagent in solution. Instead, the coupling reaction with the DIC leads to the formation of N.N′-diisopropylurea that is more soluble in DMF or NMP. Also, the EDC·HCl is used for chemistry in aqueous solution given that this coupling reagent and its urea by-product are water soluble, and it is possible to remove the excess of reagent and its undesirable product during the work-up step. The activation of carboxylic moiety in presence of carbodiimides involves the formation of O-acylisourea, a compound highly active that can rearrange in N-acylurea (inactive molecule), lowering drastically the reaction yield (*Figure 1.8*).

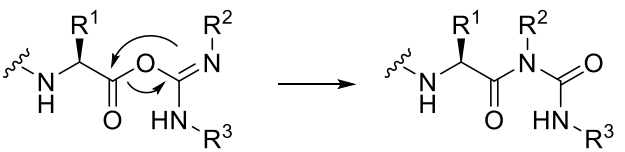



***Figure 1.8.*** *Formation of N-acylurea from O-acylisourea via O→N acyl migration.*

To increase the yield and to reduce this isomerization, the carbodiimides coupling reagent can be used in combination with additives, like *N*-hydroxysuccinimide (NHS), 1-hydroxy-1H-benzotriazole (HOBt) and 3-hydroxy-1,2,3-benzotriazin-4-one

(DhBtOH). The reaction mechanism of the coupling in presence of DIC and DhBtOH is reported in *Figure 1.9*.

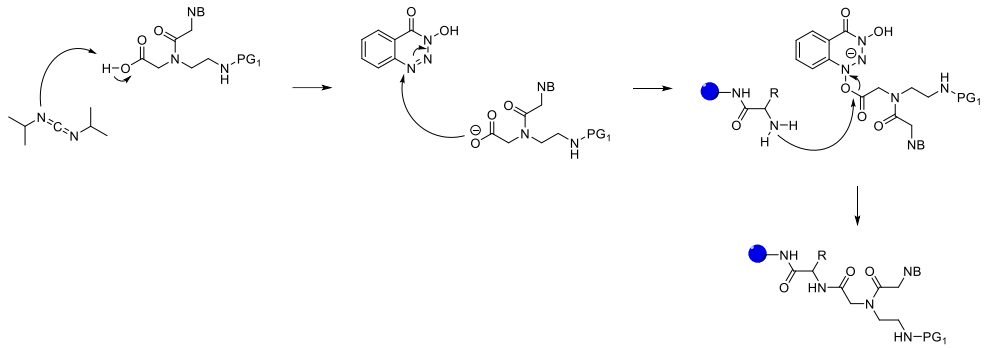


***Figure 1.9.*** *Coupling reaction mechanism with DIC and DhBtOH.*

N-hydroxysuccinimide esters are moderately stable and can be isolated. The DhBtOH esters turned out to be more reactive than HOBt-derived ones.[57]

Being the HOBt hazardous and explosive, acidic oximes were studied as alternatives, and the 2-cyano-2-(hydroxyimino)acetate (OxymaPure) was selected to use as coupling additive in carbodiimide-mediated amide bond formation on solid support.

Uronium and phosphonium salts are also very efficient coupling reagents. In these compounds the hydroxybenzotriazole or similar moieties are incorporated in a highly reactive electrophilic structure. The two principal coupling reagents based on uranium salts used are 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU) and 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), although the latter is the most employed coupling reagent both for solid-phase and solution reactions. Uronium coupling reagent if used in excess or added in the presence of a free amino moiety of the growing chain can lead to the formation of guanidinium derivatives that stable under the conditions used on solid-phase synthesis.[58] The formation of guanidinium with uronium type of coupling reagents is shown in *Figure 1.10*.

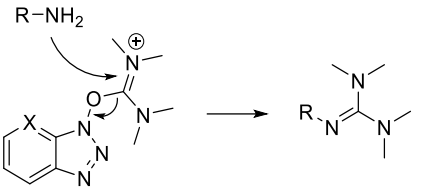


***Figure 1.10.*** *Side reaction: formation of guanidinium with uronium/guanidinium type of coupling reagents.*

The first phosphonium salt proposed as a coupling agent was benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate (BOP) but since it led to the formation of hexamethylphosphorotriamide (HMPA), a highly toxic by-product, it was then replaced by (benzotriazol-1-yloxy)tris(pyrrolidino)phosphonium hexafluorophosphate (PyBOP). PyBOP can be considered an excellent coupling reagent given that is not subjected to the guanylation reaction described above.

In presence of both uronium and phosphonium salts and a tertiary base, like N,N-diisopropylethylamine (DIPEA), the active ester can be generated by two successive substitution reactions, ad depicted in *Figure 1.11* for HBTU. The active ester is then used for amide bond formation in the coupling with the growing PNA chain.[59]

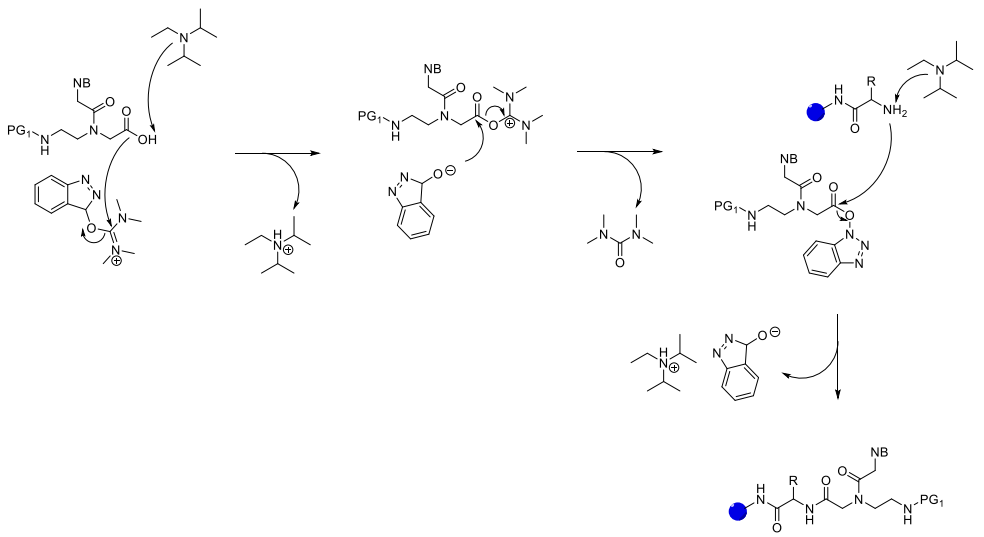


***Figure 1.11.*** *Coupling mechanism of PNA monomer on growing oligomer in presence of HBTU as coupling reagent and DIPEA.*

Stand-alone coupling reagents which include the previously mentioned OxymaPure and based both uronium or phosphonium salts are also available, like (1-cyano-2-ethoxy-2-oxoethylideneaminooxy) dimethylamino-morpholino-carbenium hexafluorophospate (COMU)[60], and (1-cyano-2-ethoxy-2-oxoethylideneaminooxy)-tris-pyrrolidino-phosphonium hexafluorophosphate (PyOxim).[61] Both present a higher stability and coupling efficiency, also overcoming the racemization problems (in peptides). In addition, it is possible to use COMU as stand-alone coupling reagent for microwave accelerated solid phase synthesis.[62,63] Other coupling reagents based on OxymaPure scaffold, like K-Oxyma, were also developed.[64,65]

The capping is used to prevent the growth of undesirable sequences caused by incomplete reactions during the coupling step. Using acetylation of the unreacted primary amine, truncated sequences are blocked and consequently the elimination of these in the purification can be less difficult. Capping is normally performed by an excess of acetic anhydride solution.

Once obtained the desired PNA, it is cleaved by the resin and only at this moment also the semi-permanent protecting group of the exocyclic amino moiety of nucleobases are removed. In general, the cleavage cocktail used for the Boc strategy is formed by TFA:TFMSA:m-cresol:thioanisole (6:2:1:19 instead for the Fmoc synthesis is composed by TFA:m-cresol (9:1)

Nowadays, the world of chemistry also looks at the production of high amount of waste, in particular toxic waste, generated by the chemical reactions and at the consequently possibility to replace the hazardous substances and solvent to chemical products that are green, safe and, environmentally friendly. Unfortunately, the solid phase synthesis cannot be ruled out by the enormous quantity of toxic solvents and reagents employed and the resulting toxic waste produced. However, the research is evolving to improve the solid phase synthesis protocols towards the green chemistry direction. In fact, several studies show the possibility to decrease the use of DMF, NMP and DCM (dichloromethane), that are the three principal toxic organic solvents used for the solid phase synthesis, or to replace them with

appropriate green solvents, like ethyl acetate (EtOAc), DMC (dimethyl carbonate), PC (propylene carbonate), GVL (γ-valerolactone), NBP (N-Butyl-2-pyrrolidone) and 2-MeTHF (2-methyl-tetrahydrofuran). Although, the yield and the product purity seem to be comparable, some problems regarding the solubility of monomers , the removal of reagents (like piperidine in Fmoc synthesis protocol) and the higher price of these solvents are encountered. Changing the solvents becomes consequently important to also study the resin swelling, being mandatory the optimal swelling of the solid support to allow a better diffusion of reagent and to obtain a higher chemical conversion on the solid phase synthesis.[66] Furthermore, the solid phase synthesis of dipeptides and tripeptide solvent free through ball-milling methodology was reported by F. Lamaty et All. in 2009.[67] To optimize the formation of amide bond toward the green approach, also greener and safer coupling reagents were deepened, like coupling reagents based on OxymaPure scaffold or coupling reagents able to perform the reaction in water, as EDC·HCl. Remaining in aqueous environments, peptide bond synthesis can be also formed using enzymes, that with a specific enzymatic synthesis can be able to create the peptide bond and not to hydrolyze it.[68,69]

Beside the development of growing optimal microwave peptide synthesizers, the synthesis of PNA was also performed through an automatic flow technology, including a system to storage the solution, three HPLC pumps and corresponding valves, loop and reactor with temperature control, UV-Visible detector, waste reservoir and a computer as central control.[70]

### 1.2.3 Orthogonality issues for PNA-conjugates

In 1977, Barany and Merrifield defined an orthogonal system as "*a set of completely independent classes of protecting groups. In a system of this kind, each class of groups can be removed in any order and in the presence of all other classes".*[71] Orthogonality is a valuable characteristic in chemical synthesis as it enables the efficient and precise creation of complex molecules by accommodating various chemical reactions without undesired interactions.[72]

Although for PNA synthesis the problem of orthogonality of the protecting groups has been solved with the appropriate combinations present in the commercially available monomers, the conjugation of PNA with peptides and other functional molecules, the synthesis of PNA-DNA chimeras and of nanostructures based on PNAs require further optimization.[73]

Beside the *Cbz* and *Bhoc*, the trityl (*Trt*) group and its derivatives has been employed in orthogonal solid-phase PNA synthesis. In fact, the monomethoxytrityl (*Mmt*) has been used as semi-permanent protecting groups of exocyclic amino function of the nucleobases in combination with Fmoc group used as temporary protecting groups of the primary amino function of PNA monomers. These two protecting groups are orthogonal because the *Mmt* group is acid sensitive.[74] The combination of the *Fmoc* group with the *Mmt* group provides monomers that can be cleaved under mildly acidic conditions and can be used for the synthesis of PNA-DNA chimeras, since deprotection by TFA (used for *Bhoc*) would result in DNA depurination. On the contrary, the *Boc(Z)*-monomers are not suitable for the synthesis of PNA-DNA chimeras because TFA deprotection and strong acidic conditions (using TFA/HF) required for cleavage from the resin and base deprotection.[45] Another derivative of *Trt* group is the 4-methyl trityl (*Mtt)* group that is renowned for its orthogonality with respect to the Fmoc group because it can be selectively removed under mildly acidic conditions. In fact, this orthogonality of Mtt and Fmoc protecting groups was used to develop PNA-conjugates for PNA-encoded synthesis (PES).[75]. In addition, using Fmoc-Lys(Mtt)-OH [76] some PNA-peptides conjugates could be synthetized with lysine as a branching point.[77,78]

Another protecting group suitable for solid-phase synthesis, orthogonal and compatible with the *Fmoc* strategies is the 1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl (*Dde*) protecting group. The Dde group is defined as quasi-orthogonal respect the Fmoc amino protected group because its orthogonality depends by the deprotection solution employed, as illustrated in *Figure 1.12* with Fmoc-Lys(Dde)-OH.[79]

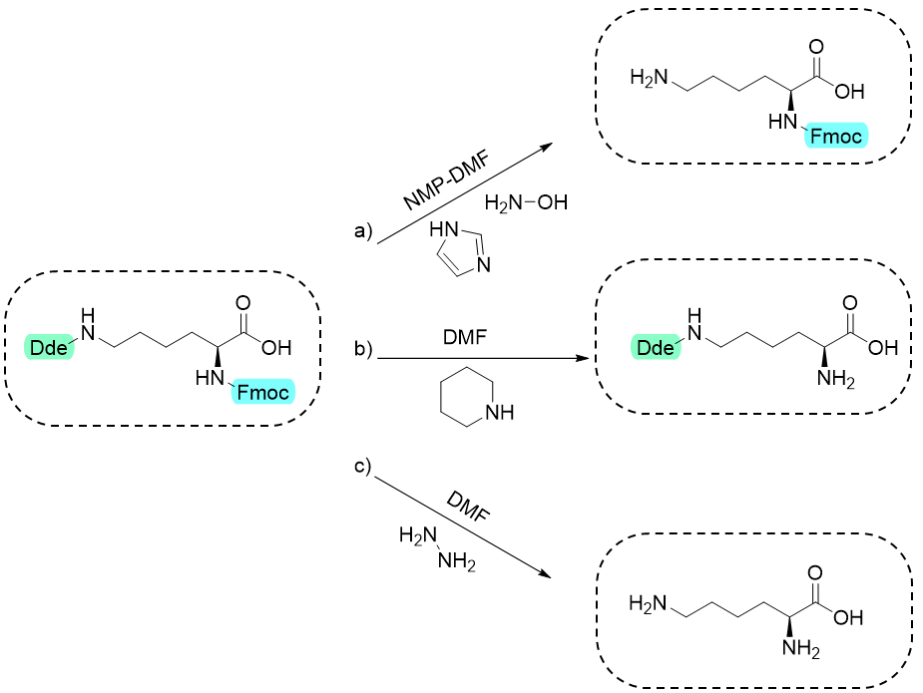



***Figure 1.12.*** *Different methodologies to deprotection of Fmoc-Lys-Dde-OH. a) Dde deprotection with a solution of hydroxylamine hydrochloride, imidazole, dry NMP and dry DMF; b) Fmoc deprotection with a solution of piperidine in DMF; c) Fmoc and Dde deprotection with a solution containing a strong base as hydrazine in DMF.*

Using a solution containing hydroxylamine, a mild and orthogonal deprotection of *Dde* in the presence of Fmoc group can be carried out. This choice of reagent proved effective in a versatile synthesis process for PNA-peptide conjugates, that can be formed both with modified *Dde/Mmt* PNA monomers or with the commercially available *Fmoc/Dde* amino acids.[80,81]

Another strategy that has been developed involves the use of the azide group to mask amino functional group, without affecting other functional groups, and thus can be considered equivalent to an orthogonal protecting group.[82] In fact, the azide function allows to avoid the presence of amino groups, especially in sensitive substrates, preventing unwanted reactions and subsequently it can be efficiently reduced (for example by Staudinger reduction or during Staudinger ligation[83]), allowing amine functionality to become accessible to bioconjugation reactions; the azide can in addition be used to perform azide-alkyne "click" reactions,[84] providing a method for biorthogonal functionalization.[85] The azide group has also been used

as masked amino-function in modified PNA monomer used in solid phase synthesis and in combination with other specific orthogonal protecting groups in the construction of multifunctional PNAs.[86] The combination of different orthogonal groups is important for obtaining scaffolds in which several PNAs segments are connected through branched units, like in nucleic acid field.[18] For PNA-peptides conjugates, both linear and orthogonal strategies were employed using manual or automatic[87] solid phase synthesis; some PNA-peptide conjugates can be used to improve intracellular delivery,[88] but might also be used as a scaffold to create self-assembled complex architectures that, enabling the control over the spatial organization of peptides, can be used as protein mimics.[89,90]

### 1.2.4 Bioconjugation via click and ligation reactions

PNA-peptide and PNA-bioconjugate structures can be obtained, in addition to SPPS, by convergent synthesis using ligation or "click" chemistry.[91,92]

The notion of orthogonal reactivity can be extended to chemical ligation [93] and click reactions. Both processes, not inherently orthogonal, can be made orthogonal by carefully selecting specific functional groups that can react with each other without interference with other components. An example of orthogonal reactivity is a peptide with cysteine residues and azide/alkyne groups. Consequently, controlling the reaction conditions, it is possible to choose if perform the native chemical ligation (NCL) or the click chemistry based on the copper-catalyzed azide-alkyne cycloaddition (CuAAC) in the same peptide sequence. In NCL, a thiol group, typically derived from a N-terminal cysteine residue, selectively reacts with a specific activated carboxylic group (thioester) of another peptide, resulting in the formation of a thioester intermediate, and then of amide bond. Instead, in CuAAC an azide group reacts with an alkyne group in the presence of a copper catalyst. Both types of reactions can be carried out without disturbing the functional groups of biomolecules and therefore are examples of "biorthogonal" reactions. Another example of biorthogonal ligation chemistry that can be performed in biological contexts is the Staudinger-Bertozzi ligation where an azide functional group and a

phosphine compound are involved to produce a stable amine and phosphine oxide. Azides and phosphines are relatively rare in biological molecules, making this reaction compatible with complex biological systems. It can occur selectively in the presence of other functional groups, such as amines, thiols, and carboxylates, which are more commonly found in biomolecules.[94] These studies on orthogonal chemistry are further broadened within supramolecular chemistry, where non-covalent interactions can be intentionally designed to create complex orthogonal systems.

### 1.2.5 PNAs in gene therapeutics

Peptide nucleic acids showed potential as candidates for designing gene therapeutic drugs. Their successful utilization depends on clearly defined targets and a thoroughly understood mechanism for delivering them into cells. The strong binding affinity of PNAs for its DNA or RNA target, while simultaneously evading enzymatic degradation, has proven it to be a powerful instrument for controlling gene expression at transcription, processing, and translation levels.[95,96] The principal strategies that can be employed in the development of gene therapeutic drugs are antigene and antisense strategies.

In the antigene strategy, PNAs can disrupt the transcription processes by forming stable duplex and triplex structures able to give rise to strand-invasion and strand displacement in dsDNA. These complexes can introduce structural obstacles that impede the proper functioning of RNA polymerase, thus blocking transcription to pre-mRNA.[97] For example, Nielsen et al. have shown that an 8-mer PNA (T8) can inhibit the activity of phage $T_3$ polymerase.[98]

However, the strand invasion and the formation of strand displacement complexes in presence of PNAs are relatively slow at physiological salt concentrations, and this feature represents a significant drawback for utilizing PNAs as an antigene agents.[99] To overcome the problems, several enhancements to PNA have been explored, such as chemically connecting the ends of Watson-Crick and Hoogsteen PNA strands, introducing pH-independent pseudoisocytosines into the Hoogsteen strand[100],

adding positively charged lysine residues to PNA strands[101], or adding interacalators[102]. These modifications can significantly accelerate the association rates with double-stranded DNA.[103]

As an example, antigene PNAs directed towards the NMYC oncogene and conjugated to a nuclear localization signal (NLS) peptide has proven to be effective *in vivo* in animal models.[104]

In the antisense strategy, the formation of hybrids with both pre-mRNA and RNA can prevent translation, and PNA directed towards pre-mRNA can block aberrant splicing. PNA can also form hybrids with complementary sequences in microRNA (miRNA), consequently impeding its incorporation into the RNA-induced silencing complex (RISC), though this strategy is most often referred to as anti-miR or anti-microRNA rather than antisense.

MicroRNAs are brief (18-25 nucleotides in length), non-coding RNAs found in various tissue and cell types, and they play a role in repressing the expression of specific target genes. In fact, miRNAs are transcribed and processed from precursors called pri-miRNA, pre-miRNA, and mature miRNA. Mature miRNAs are incorporated into the RISC complex and guide it to target mRNAs to regulate their expression.[97]

The primary mechanism of oligonucleotide antisense activity predominantly involves ribonuclease-H (RNase-H)-mediated degradation of the mRNA/oligonucleotide hybrid, thus reducing mRNA levels. In contrast, PNAs do not serve as substrates for RNase-H or other RNases. Consequently, PNAs primarily exert their antisense effects through steric interference. Thus, higher concentrations of PNA are required to achieve reasonable antisense activity compared to traditional oligonucleotides, but their chemical and biological characteristics make them a powerful tool for antisense applications.[105,106]

In 1992, the use of PNAs as antisense agents was initially showed. By employing nuclear microinjection, a 15-mer PNA designed to target the translation initiation site of SV40 large T antigen mRNA effectively hindered transcription in cellular

extracts. This inhibition exhibited both sequence-specificity and was reliant on the dosage applied.[107]

Antisense PNAs have been developed with the aim, for example, of inhibiting RNA splicing in the expression of the human cancer gene mdm2 in choriocarcinoma (JAR) cells. These antisense PNAs targeted the 5' or 3' splice sites in intron 2 or the 3' splice site in intron 3 of mdm2 pre-mRNA. Cellular data revealed a reduction in mdm2 protein levels and a simultaneous increase in tumor suppressor p53 levels for a specific PNA sequence. However, when this sequence was combined with the DNA-damaging agent camptothecin (CPT), cell growth was inhibited more effectively than with CPT alone.[108]

In addition, an antisense peptide nucleic acid was engineered to specifically target the pseudoknot structure within the SARS-CoV frameshifting region. This PNA was then evaluated for its effectiveness in inhibiting -1 ribosomal frameshifting (-1 PRF) and SARS-CoV replication and the results showed a reduction of it in a sequence-specific way. An $IC_{50}$ of 4.4 μM in cell culture was necessary to down-regulates SARS-CoV replication.[109]

Another PNA with antisense activity directed, instead, to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was reported. This PNA sequence was design to target the 5′ UTR region of SARS-Cov-2 and the results demonstrated that the viral infection was inhibited over 95% in live SARS-Cov-2 assays with an $IC_{50}$ of 0.8 μM.[70]

PNA can however have greater advantages over oligonucleotides as potential modulators of miRNA, since their binding is strong regardless of the secondary structure of the RNA strand, and Rnase-H activation is not required.[110–112] The exceptionally strong binding affinity of PNAs to RNA, coupled with their remarkable chemical and enzymatic stability, particularly in the case of backbone-modified sequences, positions them as excellent candidates for miRNA inhibitors. The possibility to develop PNAs designed to target miRNAs exhibit several notable effects. First, they can reduce the accumulation of miRNAs, essentially by limiting their bioavailability within the cell. Second, PNAs interfere with the normal functions

of miRNAs, which consequently leads to an increase in the expression of miRNA target mRNAs, and thus protein production. Lastly, and perhaps most importantly, they induce substantial disruption in the biological effects that are normally regulated by miRNAs. Thus, they are presently considered as powerful tools for modulating miRNA activity.[113]

For example, two anti-miR PNAs conjugated with an octa-arginine moiety and targeting miR-221 and miR-222 are synthetized and employed in U251, U373 and T98G glioma cell lines. This study revealed that when both anti-miR-221 and anti-miR-222 PNAs were administered together, a significantly enhanced pro-apoptotic effect was observed.[114]

In such an effort, several research have highlighted that anti-miR PNAs can be a tool on the control of the growth of tumor cells in various cancer types.[115–118]

In view of their possible use as therapeutic agents, over the years, several PNAs backbone and nucleobases modifications and some delivery systems were developed to increase affinity and induce greater specificity towards target DNA and RNA sequences, and to overcome problems such as penetration of PNA strands inside the cells.[103,119–122]

### 1.2.6 Delivery of PNAs

The successful delivery of peptide nucleic acids is essential for their application in diverse fields such as gene therapy, antisense therapy, and theragnostic. To ensure that PNAs effectively reach their designated cells or tissues, is crucial to select an appropriate delivery method is crucial. In fact, the choice of delivery strategy hinges on the specific application, the intended target cell or tissue, and the desired therapeutic outcome. Ongoing research endeavors are focused on innovating and refining delivery approaches to enhance the efficiency, selectivity, and safety of PNA delivery across various biomedical applications.

Just as over the years there has been significant progress in the field of nucleic acid delivery, including the use of conjugates and of nanomaterials. [123] In fact, several approaches for the delivery of PNAs and modified backbone PNAs (*Figure 1.13*) are

listed, going from direct introduction into cell cultures to conjugation with organic molecules, peptides, and polymers or to incorporation in inorganic structures.[124,125] Mechanical and electrical techniques (like microinjection or electroporation) are employed to incorporate into cells unmodified PNAs, showing that potentially the internalization could occur, but this was a labor-intensive procedure highly invadable, feasible for small-scale setups.[107,126,127]

Regarding instead modified PNAs, research has demonstrated that chemical alterations to the PNA backbone, introducing positive charge characteristics, can enhance cellular uptake. One instance of this is the incorporation of positively charged guanidinium groups at either the C-2 or C-5 position of the PNA backbone.[128–130]

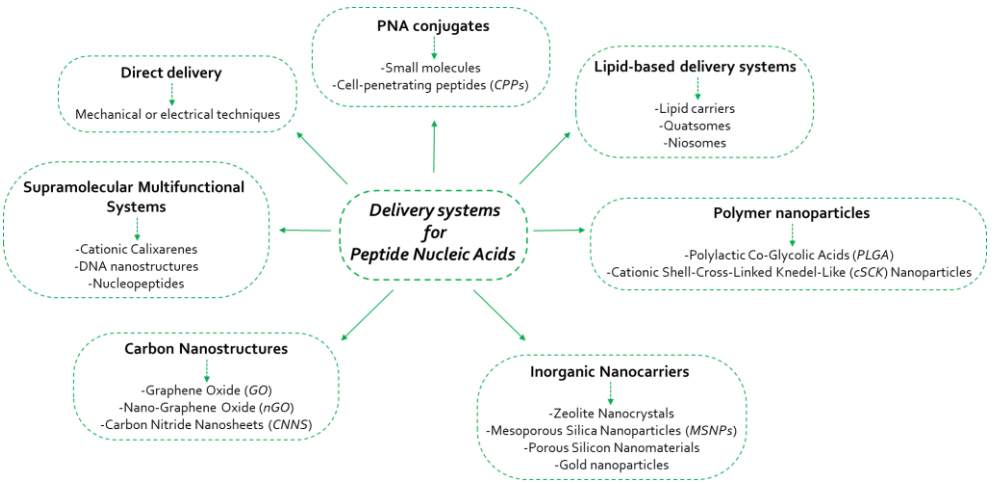



***Figure 1.13.*** *Several delivery systems of Peptide Nucleic Acids.*

Over the years, PNAs has been conjugated with small molecules and macromolecules like lipophilic adamantly groups[131], triphenylphosponium[132], N-acetyl galactosamine[133,134], guanidinium moiety[135], phospholipid units[136] and others, to target different cells and tissues. However, a highly promising strategy extensively investigated for the cellular delivery of PNA entails linking them with cell-penetrating peptides (CPPs), also referred to as protein translocation domains (PTD) or Trojan peptides (TP). Cationic CPPs and amphipathic peptides are the two

principal groups of this delivery through peptide conjugations. The first are general constituted by amino acids with positive charges and include poly-lysine and polyarginine domains. Instead, the amphipathic peptides are short amino acid sequences possessing both hydrophobic and hydrophilic regions in their structure. These structural characteristics enable these modified peptides to engage with cell membranes, facilitating their ability to traverse the lipid bilayer and enter cells.[137,138]

In the literature, several examples are reported that highlight how PNA-CPP conjugates can be a potential tool for overcoming delivery-related issues associated with PNAs[139–145], with the goal of further optimization to avoid high levels of toxicity in in vivo applications (one of the main limitations of such systems).[146]

Lipid carriers are also conjugated with PNAs to optimize the delivery for specific applications and cell types. However, this conjugation was performed with the help, for example, of DNA strands[147], anionic peptides[148], and lysine modified monomers[149] to overcome the problems related the compatibility between these two systems. Also, others lipid-based delivery systems have been created, including quatsomes[150] and niosomes[151].

For PNAs delivery, some polymer nanoparticles like Polylactic Co-Glycolic Acids (PLGA)[152], and cationic Shell-Cross-Linked Knedel-Like (cSCK) nanoparticles[153] have been employed. The polylactic Co-Glycolic Acid (PLGA) is non-toxic and biodegradable polymer, can undergo hydrolysis under physiological conditions, and it finds extensive application in drug delivery.[154,155] Instead, cationic Shell-Cross-Linked Knedel-Like (cSCK) nanoparticles are a type of nanoscale particle constructed using co-polymers and characterized by a cationic outer shell.[156,157]

Inside the inorganic nanocarriers, some different groups are investigated to improve the PNA delivery, like zeolite nanocrystals[158], mesoporous silica nanoparticles (MSNPs)[159], porous silicon nanomaterials[160], and gold nanoparticles[161].

On the other hand, carbon-based nanocarriers have shown significant potential for the delivery systems for various types of therapeutic molecules, including PNAs. Carbon-based nanocarriers, such as graphene oxide (GO)[162,163], nanographene oxide

(nGO)[164], and carbon nitride nanosheets (CNNS)[165] were found to be attractive for PNA delivery and sensing application.

Instead of resorting to chemical modifications, a particularly advantageous approach for PNA delivery involves employing carriers capable of non-covalent and reversible interactions with the cargo, like cation calixarenes. In fact, for example, an argininocalix[4]arene was reported to the delivery of PNAs inside the cells.[166]

DNA-based nanostructures (such as DNA tetrahedron)[12] and nucleopeptides (artificial macromolecules)[167] have been investigated as delivery systems for PNAs.

### 1.2.7 PNAs as diagnostic tools

The exploration of novel diagnostic applications for PNAs in the field of biosensing is a flourishing field of research, thanks to the strong affinity of peptide nucleic acids for complementary natural nucleic acid, DNA or RNA and, most importantly, to their high discriminating ability towards fully complementary sequences. PNAs have been utilized for detecting genetic mutations and analyzing mismatches in Watson-Crick hydrogen bonding. The remarkable characteristics of PNAs have led to a significant upsurge in the development of optical and electrochemical biosensors.[168]

In particular, the attachment of PNA as diagnostic probes onto the surfaces of electrochemical and optical transducers has enabled the development of very efficient biosensors for the detection of various analytes, such as viral RNA, bacterial DNA, circulating DNA for the so-called *liquid biopsy* strategy, and for the detection of point mutations or single-nucleotide polymorphisms (SNPs).[169]

#### *1.2.7.1 Optical biosensors based on PNAs*

Optical biosensors provide significant rewards compared to traditional analytical methods since they allow for the direct, real-time, and label-free detection of a wide array of biological and chemical compounds. Optical biosensors represent advanced analytical instruments that utilize light-guiding technologies as their transduction mechanism. They harness the characteristics of light to detect biological or chemical

analytes. Alterations in the optical properties of light when interacting with these analytes result in modifications of the output spectrum or the electrical signal detected by the photodetector. [170]

Fluorescence-based PNA sensing dates to 1999. This approach involved employing a PNA oligomer labeled with a cyanine dye known as thiazole orange. Upon binding to complementary DNA, this system resulted in a remarkable ~50-fold increase in fluorescence compared to a free probe.[171] Notably, this technique found primary application in real-time PCR quantification and the differentiation of single-point mutations.[172,173]

PNA probes have been used in parallel sensing using microarray technology that require labelling of DNA, but label-free techniques are nowadays mainly developed. These include surface plasmon resonance (SPR), optical resonators, waveguide, and optical fiber sensors.

Optical fiber-based biosensors offer a vast range of applications in both chemical and biochemical sensing, thanks to their inherent advantages, including compactness, potential for miniaturization, high compatibility with optoelectronic devices, and the capability to perform multiplexing and remote measurements.

A novel approach was introduced for the implementation of a label-free DNA sensor by Candiani et Al., using tilted fiber Bragg grafting (TFBGs) functionalized with a PNA as probe for direct label-free DNA detection. The system demonstrated specificity exclusively when the fiber was functionalized, displaying excellent sensitivity even for solutions with extremely low concentrations. Furthermore, experiments conducted with mismatched DNA strands illustrated the sensor's ability to discern a single nucleotide polymorphism (SNP) within the DNA strand.[174]

In a second moment, the same group functionalized micro-structured optical fibers (MOFs) or photonic crystal fibers (PCFs) with a PNA probe, a DNA strand of the Cystic Fibrosis (CF) gene containing a point mutation, and oligonucleotide-functionalized gold nanoparticles (ON-AuNPs). In this way, a sandwich-like system for signal amplification was created, showing a wavelength shift in spectral measurements.[175] Then, Bertucci et al employed the same type of optical fiber but

modified with a PNA probe to detect genomic DNA linked to a genetically modified Roundup Ready soy gene segment, eliminating the need of amplification techniques. Their research showed the feasibility of directly measuring genomic DNA through infrared light reflection with PNA probes and the use of streptavidin-coated gold nanoparticles to increase the signal.[176]

Given that photonic crystal fibers have proven to be a promising tool for the development of DNA and protein detection sensors with high sensitivity.[177] Hollow-Core Tube Lattice Fibers (HC-TLFs) which exploit properties of photonic crystals and confine light in hollow cores of the fiber's inner surface were also used.[178,179] Giovanardi et al reported the functionalization of HC-TLFs with biotin probe for the detection of streptavidin analytes, forming a bio-layer able to generate a redshift on the fiber transmission spectrum comparable with the dimension of the streptavidin protein.[180]

*1.2.7.2 Electrochemical biosensors based on PNAs*

Capable of transducing a bio-molecular recognition event into a measurable electric signal, which may be voltammetric, potentiometric, or impedimetric an electrochemical biosensor is a device that functions as a converter. In fact, an electrochemical biosensor combines a biological recognition element with an electrochemical transducer, offering either quantitative or semi-quantitative data of various analytes.[181,182]

In 1997, Wang et al. provided one of the earliest instances of PNA utilization as a bioreceptor. They employed a 15-mer PNA in an electrochemical biosensor to detect a specific mutation in the p53 gene. However, the electrochemical quantification of DNA and the discrimination of mismatches are made possible by monitoring the hybridization event between the PNA probe and the target DNA.[183]

Over the years, PNAs as probe towards several analytes are employed in voltametric biosensors. For example, square wave voltammetry was used to record the electrochemical alterations during the hybridization process between a PNA molecule covalently bonded to a quinone-based electroactive polymer and its

complementary DNA target.[184] Furthermore, for the detection of complementary oligonucleotides, Aoki and Umezawa showed a sensitive biosensor. This biosensor involved modifying a gold electrode with self-assembled monolayers of a 13-mer PNA probe and 8-amino-1-octanethiol.[185]

PNAs' properties as capture probes have also been harnessed for microRNA detection.[186] In fact, a PNA-based biosensor with gold electrode as a platform to fixate a thiolated PNA was designed to detect miRNA-145.[187] In addition, a PNA-modified gold electrode combined with the target-catalyzed hairpin assembly (CHA) in a square wave voltammetric sensor was reported to identify miRNA21 and miRNA155 within cancer cells.[188]

A recent study introduced an innovative electrochemical magneto-genosensing assay. This method involves immobilizing PNA capture probes on magnetic microbeads (m-MBs) for the detection of both mutated and wild-type KRAS oncogene DNA.[189]

## 1.3 Self-assembly of peptides

### 1.3.1 Peptide synthesis

Peptides can be defined as a short polymer of amino acids with molecular weight maximum around 5000 Da (< 50 amino acid residues), an important feature that distinguishes peptides from proteins, although a sharp boundary between these two classes is not present.[190]

The first peptide was synthesized by the German chemist Emil Fisher around 1900[191], a small peptide composed by glycine residues, and consequently he marked a pivotal moment in the study of proteins and their constituents. However, the evolution of peptide chemistry happened thanks to several factors: the introduction of amino protecting groups, starting from Bergmann and Zervas[192] in 1932 with the synthesis of N-carbobenzoxy (*Cbz*) as reversible amino protecting group; the development of more efficient coupling reagents, as N,N'-dicyclohexylcarbodiimide used by Sheehan and Hess (1955)[193] to improve the methodology of forming peptide or other amide bonds; the invention of solid-phase peptide synthesis (SPPS) strategies by R. B. Merrifield in 1963, where the growth of peptide chain occurs from C-terminal end to N-terminal end on solid support. This strategy is based on the use of protecting groups for $N_\alpha$-amino group and for side-chain functional groups of amino acids.[54] The peptide synthesis, therefore, has been gradually advanced over the years through the development of further protecting groups for $N_\alpha$-amino and of new activating agents for the carboxyl functions. Orthogonality, as previously discussed allows to remove efficiently and selectively several protecting groups on the same peptide, whenever a high degree of compatibility between the different protecting groups is present.[194,195]

Regarding the $N_\alpha$-amino protecting groups different types are developed based on several deprotection conditions, like acid-labile, base-labile, Pd-labile, or thio-labile protecting groups. In the category of acid-labile protecting groups, that can be deprotected under milder acid conditions than the Boc group, we can cite as examples 2-(4-biphenyl)isopropoxycarbonyl (Bpoc), triphenylmethyl (trityl, Trt),

and α,α-dimethyl-3,5-dimethoxybenzyloxycarbonyl (Ddz). As example of base-labile groups 2-[4-(methylsulphonyl)phenylsulphonyl] ethyloxycarbonyl (Mpc), 2,2-bis-(4'-nitrophenyl) ethyloxycarbonyl (Bnpeoc), 2-(2,4-dinitrophenyl) ethyloxycarbonyl (Dnpeoc), and 2-(4-nitrophenylsulfonyl) ethoxycarbonyl (Nsc) can be mentioned. Other two important base-labile protecting groups are Dde and 1-(4,4-dimethyl-2,6-dioxocyclohexylidine)-3-methylbutyl (ivDde) that lead to the formation of imidazole derivative as a side-product during the deprotection reaction. The allyloxycarbonyl (Alloc) group is commonly used as a protecting group in peptide synthesis and can be cleaved by a palladium-catalyzed reaction. The dithiasuccinoyl (Dts) is instead a thio-labile protecting group since it can be deprotected efficiently using strong acids and light, but it can also be selectively cleaved under gentle conditions through thiolysis.[196]

### 1.3.2 Peptide conformations

The sequence of amino acid provides the primary structure of polypeptide chain and contains the information about the folding of the protein, which is closely related to its biological properties and functions. The primary structure is the most basic level of protein or peptide structure and serves as the foundation for higher-order structures like secondary, tertiary, and quaternary structures. The secondary structure is the arrangement of amino acids in localized regions of a polypeptide, and the hydrogen bond between the peptide backbone atoms play an important role in the stabilization of these types of polypeptide structures. These structures correspond to a state of minimum internal energy achieved by maximizing hydrogen bond formation and minimizing steric repulsion between the side chain groups.[197]
The principal secondary structures are *α-helices* and *β-sheets*. An *α-helix* secondary structure involves hydrogen bonds regularly spaced between carbonyl oxygen (hydrogen bond acceptor) and the amino group (hydrogen bond donor) of each third residue in the same polypeptide chain. This type of secondary structure present 3.6 residues per turn and the side chain of amino acids points outwards from the helix,

and the peptide chain coils in a right-handed spiral manner. A *β-sheet* secondary structure is formed by hydrogen bonds between two or more *β-strands*, in specific between carbonyl groups of one *β-strand* and amino groups of another chain, creating a sheet-like structure. This is the reason because in the *β-pleated sheet*, a portion of the chain cannot make hydrogen bonds with itself, such as it happens in the *α-helix* secondary structure, but it can only bind with another adjacent β-strand. *β-sheets* can be organized in two different orientations: parallel and antiparallel, according to the nature of hydrogen bonds.[198]

Following *α-helices* and *β-strands*, the *β-turn* emerges as the third most important secondary structure in proteins and the most common type in the category of *turns* or *loops* (including also γ-turns, α-turns and Ω-loops). A *β-turn* is defined by a nearly 180° turn comprising four amino acid residues and this structural feature enables the peptide chain to alter its direction. In fact, it is frequently stabilized by a hydrogen bond between the carbonyl oxygen of the first residue and the amide hydrogen of the fourth residue (Type I *β-turn*). However, different types of *β-turn* are identified and classified based on the specific amino acid residues involved and the geometry of the turn, in specific the dihedral angles of the central residues, although the classification of *β-turns* in proteins continues to be a field undergoing continuous development.[199–201]

Another type of secondary structure is $3_{10}$*-helix*, that includes approximately 15-20% of helices in proteins. The typical $3_{10}$*-helix* present three residues per helical turn and the amino group of an amino acid forms a hydrogen bond with the carbonylic group of the amino acid three residues earlier, obtaining a repetition of $i$, $i + 3$ hydrogen bonding and forming a 10-atom ring. In addition, $3_{10}$*-helices* are shorter and less stable as compared to *α-helices* as they are generally composed of 3.5 amino acid residues.[202,203] Being shorter, $3_{10}$*-helices* tend to have difficulty forming super-secondary or tertiary structures, and generally, this type of secondary structure is observed in short stretches of natural peptide and protein chains. However, by leveraging de novo protein design, it is possible to assemble more complex structures based on $3_{10}$*-helix* peptides.[204]

A polypeptide chain that does not exhibit either of the previously described organized secondary structures can be referred to as a *random coil*. Random coil structures are not considered properly as a secondary structure in proteins, but this term is used to describe the disordered and unstructured polypeptide chain, lacking a secondary structure and a repeating or regular pattern in their backbone conformation.

Elements of secondary structure can be combined into specific arrangements known as *super-secondary structures or motif*, which are neither regular nor repetitive. Super-secondary structures can be simple if they consist of a maximum of three elements, but also complex super-secondary structures have been studied. Among these, the most common ones are *beta-alpha-beta motif* (*βαβ motif*), *β-hairpin*, *β-barrel*, and alpha-alpha motif (*αα-motif)*. The *βαβ motif* consists of a β-sheet (usually composed of β-strands) followed by an α-helix and then another *β-sheet*. In *β-hairpin motif*, instead, a short section of β-sheet structure forms a hairpin-like bend and, in fact, in the *β-hairpin* two consecutive *β-strand* elements exhibit an antiparallel orientation and are separated by a short loop region (hairpin). *B-barrel* is characterized by a cylindrical arrangement of beta strands, which are connected by hydrogen bonds, forming a closed, barrel-like structure. On the other hand, *αα-motif* involves the presence of two alpha helices in close proximity, often interacting with each other or with other components.[205]

The regular folding of long peptide chain gives rise to a three-dimensional tertiary structure, predetermined and dependent on the primary structure. While peptides are shorter than proteins, they can still exhibit various types of tertiary structures. The quaternary structure determines how two or more polypeptide chains associate through non-covalent interactions.[206,207]

## 1.4 De novo protein design

A protein is a complex and large molecule made up of amino acids residues, which determine their function. Being able to act as molecular machines, proteins play a crucial role in the structure, function, and regulation of cells, tissues, and organs within living organisms.

The functions that proteins naturally mediate are countless, and these properties arise from a suitable combination of amino acid sequences that determine the three-dimensional structure of the protein itself. Consequently, three-dimensional structure of protein can be calculated and predicted if its primary sequence is recognized. However, eventually changes in any level of protein structure can affect its activity and biological role.[208] This is the reason aware that the folding of the proteins, and their consequently biological activity, is under continuous study both in biochemistry, biotechnology, medicine, and material science fields, in combination with the protein design.

The principal factors engaged in the protein folding are hydrogen bonds, entropic effect, electrostatic interactions, hydrophobic interaction, covalent bond but also backbone and side chains of amino acid and interactions with metal ions. Protein folding is a delicate balance of these interactions, and the protein seeks to adopt a conformation that minimizes its free energy, thus achieving a stable, lower-energy state.[209]

Protein design is focuses on the creation and modification of proteins to obtain specific functions or properties. The approach has made significant advances in recent years, thanks to improvements in understanding structures and functions of proteins, as well as to the development of computational tools (base on calculation artificial intelligence (AI) predictions) and laboratory techniques. In fact, different concepts of chemistry, physics, biology, and bioinformatics are necessary in protein design. Two distinct approaches in the field of protein engineering will be discussed here: *protein redesign* and *de novo protein design*, each with its own objectives and methodologies.

In specific, *protein redesign* is focused on the modification and optimization of natural protein to obtain specific functions and applications. In fact, it is concentrates on modifying existing protein structure or sequence to improve or alter the connection inside the secondary structure (changing the amino acid sequence, side chains, or backbone) and consequently their properties, like stability, binding affinity, or enzymatic activity. In *protein redesign*, being that involves modifying of existing proteins, the amino acids forming the polypeptide chains are coded and usually these proteins are macromolecules with relatively low symmetry. [210]

*Protein redesign* is used in biotechnology, drug development, and enzyme engineering to improve or customize the properties of natural proteins. For example, a redesign of a GTPase-activating protein (GAP) in a way that enhances its ability to activate oncogenic Ras proteins, to understand and potentially target cancer-related signaling pathways, was carried out.[211] In another work, instead, a computational effort to redesign the Beta-27 Fab antibody to enhance its ability to bind to the Omicron variant of the SARS-CoV-2 virus more effectively than the human ACE2 receptor was performed.[212] Furthermore, also artificial nanostructures were synthetized redesigning nanoarchitectures of natural protein. This process involves an initial examination of the crucial interfaces between subunits within the targeted protein building block. Subsequently, the specific sites that ensure the protein nanostructures assembly are identified and, at the end, the modification of various types of interactions at these critical sites are introduced.[213]

*De novo protein design*, instead, focuses the attention on the creation of new proteins presenting novel predefined structures and functions that do not exist in nature, but that may have practical applications in various filed, like chemistry, biochemistry, medicine, and nanoscience. In fact, in this folding approach any pre-existing template is known and the creation of a new protein starting from the organization of amino acids (building blocks) following principles of protein folding and stability. Computational methods and algorithms (based on factors like energetics, geometry, and interactions) to predict the three-dimensional structure of a designed

protein based on the desired function have been employed. The objective of computation methods is to identify, when provided with a structure and function specified by the protein designers, one or a small number of amino acid sequences with minimal energy that can reliably adopt the intended structure while executing the desired function. Nowadays, using *de novo protein design* approach, it is possible in fact to design new proteins with desired pre-determined three-dimensional structures and functions. This new development was possible thanks to previous studies that have led to the understanding of proteins folding. The initial skeleton of novels proteins design can be constructed from short peptide fragments that assemble like the secondary structures of proteins, employing parametric equations to predict their geometry. However, the number of alternative conformations for a set of peptide sequences, though large, is limited. Within this set, a small portion is selected, characterized by sequences capable of aggregating through suitable central "core" packing and intramolecular hydrogen bonds among polar residues. Attention is focused on possible protein skeletons that exhibit the desired structure in a state of lower energy. This lowest energy state can be achieved by imposing appropriate constraints on the structure, such as the exposure of polar residues to the solvent in certain types of structures and the adjustment of the flexibility of the polypeptide chain by replacing different amino acid residues.[209]

The development of de novo protein design can be divided over the years into three distinct phases. The initial period involved the use of physical models to perform manual protein design. The second phase relied on computational design controlled by fundamental physicochemical principles. In the third phase, the development of backbone fragments from the Protein Data Bank (PDB) and the proliferation of crystallographic structure data have contributed to the improvement of de novo protein design. Nowadays, the tools of these three historical phases were connected to enhance the development of de novo protein design.[214] In the first phase, an example that involves manual protein design through physical method was carried out by Gutte et al. in 1979, designing and synthetizing an artificial polypeptide consisting of 34 residues able to bind nucleic acid.[215]

Regarding, instead, the second phase, the design based on α-helices used for the creation of helical bundles family was performed. Subsequently, after development of novels helical bundles structures, became feasible to introduce specific functions into these, like some modification to bind various metal ions and small substrates. In a helical bundle the α-helices do not form a compact packing, leading to the generation of cavities that can bind small molecules. Some of the *de novo* protein design focused the attention also on metalloproteins[216], helical bundles serving as catalysts and regulators of protein-protein interactions[217], and helical bundles able to interact with cofactors.[218] For example, starting from a simple and unnatural peptide sequence formed by only three different amino acids (glutamate, lysine and leucine), and then strategically placing histidines to bind hemes through bis-histidine coordination, a four α-helical bundles was formed. After that, a glutamate residue was introduced to induce distal histidine strain, thereby enhancing $O_2$ binding. To avoid the heme oxidation and keep stable the oxygen binding, loops were introduced at the end of peptide strands, reducing the mobility of the helices and the exposure to water. In this way, an oxygen transport protein was designed and synthetized. [219]

After helical bundles structures, more challenging tertiary architectures were designed, such as αβ protein structures, β-sheets, β-barrels[220], β-hairpin[221], coiled-coil[222], repeat units of proteins and others.

This type of design became possible thanks to the use of molecular machines, and then with the develop of software packages for protein structure prediction and design.[214,223]

Monte Carlo methods are commonly used in computational protein design to explore conformational space and optimize protein structures. One of the well-known software tools that utilize Monte Carlo simulations for protein design is Rosetta, a powerful tool for computational protein design. Rosetta uses a combination of Monte Carlo methods and energy minimization techniques to perform functions as the de novo protein design, protein-protein interface design, and protein-ligand docking, searching for low-energy conformations.[224–226] Other

software tools for de novo protein design are Foldit (software tool for protein design and structure prediction)[227], Modeller (program for comparative protein structure modeling and optimization), Talaris (tool for detection of low-energy sequences starting from a protein structure), and Avogadro (editor and visualization tool for visualization of protein structures and design of new molecules), but also other *de novo* protein design programs are reported in literature.[228–232]

A central component of these software is the scoring function. In de novo protein design, a scoring function is a mathematical or computational model used to evaluate the designed protein structure. It considers various factors and energy terms to estimate the stability, energetics, and overall suitability of the designed protein. The scoring function assigns a numerical score or energy value to a proposed protein structure, with lower scores indicating more favorable or better-designed structures.[223]

Constructing de novo proteins with repeated units lead to have a high symmetry on the final design, obtaining a reduction of the spatial dimensions of each sequence of the proteins. This internal symmetry is a relevant feature in closed structures in which the final unit overlaps with the initial one, like in α-helical toroids[233] and triose phosphate isomerase (TIM) barrel[234] structures. In alpha-helical toroids, multiple alpha-helices are arranged in a circular or toroidal shape, creating a closed, ring-like structure, while in TIM barrel structures the repetition of α-β-α-β alternating units was observed.

In addition, de novo protein design turns attention on the membrane protein design, concerning the engineering or modification of proteins to be functional in the context of biological membranes. Membrane proteins are a diverse class of proteins that play essential roles in various cellular processes, including signal transduction, transport of molecules across membranes, and cell adhesion. The development of de novo membrane proteins design allowed a better comprehension of folding and functions of these membrane proteins, for example of trans-membrane (TM) group formed by β-barrels and helical bundles building blocks. Furthermore, the design of the membrane proteins led to understand that the hydrophobic effect, driven by the

tendency of nonpolar molecules to cluster together in a hydrophilic environment, is enough by itself to stabilize the folding of membrane proteins.

De novo proteins can be generated through different strategies. One is the blueprint strategy where the creation of a detailed blueprint from scratch with specific lengths and orientations can led to create a protein. Another strategy, instead, combines fragments or protein domains found in nature, which can be found in Protein Data Bank. The introduction of short fragments from PBD used as a building-blocks for the backbone of de novo proteins has allowed to further expand this field (third phase of de novo protein design development). Using building blocks from PBD led to the design of TOP7, featured by αβ-motif that showed an impressive thermal stability.[235] AbDesign is a different method to assembly de novo protein, which separates different protein families into fragments, which are subsequently recombined to obtain the tertiary structure and the desired protein function.[209,214,223] Thus, the combination of small fragments gave the opportunity to design de novo protein, that consequently can be assembled in a supramolecular model.[236]

### 1.4.1 Coiled-coil structures

Coiled-coil motif are biological macromolecules discovered independently by Linus Pauling and Robert B. Corey[222] and F. H. C. Crick[237,238] around 1952. Both observed that a group of fibrous proteins, like keratin, myosin, and fibrinogen, defined by William Astbury exhibited a common structure composed of α-helices oriented in the same direction and twisted around each other with consequent change in curvature of the axis of the α-helix, causing a tilt angle of the packing rather than a perfect parallel interaction. For L. Pauling and R. B. Corey these structures were helical compounds and they do not provide any information regarding the arrangement of the side chains. They underlined that the distortion of the axis of the α-helix can be attributed to different distances on the hydrogen bond formations, caused by the interactions between side chains or by the steric hindrance. However, this distortion was attributed to the presence of a specific pattern of amino acids, necessary for maintaining the structural integrity of alpha-helices within the protein.

Instead, F. H. C. Crick called these structures as coiled-coils, focusing the attention on the interactions of the side chains as driving force of his model. In this way, he was able to recognize that α-helices, when interacting with each other at an angle of 20°, could systematically weave their side chains, repeating the same interaction every seven residues (two turns of the helix). The interactions between the chains were defined as *knobs-into-holes*, such as to provide the energy required for the distortion of the α-helix. Furthermore, both carried out that also three α-helices can interact together around an axis.

The first coiled-coil structure with two α-helices was published by Hodges et al. in 1981. It was two strands, formed by 43 amino acid residues, able to dimerize through a disulfide bridge under aerobic oxidation thank to the presence of a cysteine residue at the N-terminal. Each strand was formed by $AB_4C$ base structure were A, B and C are three different building blocks formed by seven amino acids. In specific, A was constituted by *KCAELEG*, B from *KLEALEG* and C was set up by *KLEALEGK*.[239]

*Coiled-coils* are formed by the rolling of two or more α-helices to create left-handed quaternary structure, called a *supercoil*. The most frequently encountered form of coiled coil is the left-handed one and it was defined as canonical coiled-coil, but also right-handed coiled-coil structures have been studied, and it was considered as a part of non-canonical coiled-coils.[240] However, the two or more α-helices can roll up around a central axis both parallel and antiparallel orientations. In canonical coiled-coils, the helices exhibit a periodicity of seven residues, represented as $(abcdefg)_n$, where n indicates the number of times the heptad repeat occurs. In addition, the *a* and *d* amino acid residues present a hydrophobic character, instead the *b*, *c*, *e*, *f* domains are hydrophilic. These characteristics allowed to stabilize the coiled-coil with an internal hydrophobic core located at the internal junction of the helices, where hydrophobic or non-polar residues are concentrated, and with electrostatic interactions between the helices generated by hydrophilic or polar residues that have a strong affinity with aqueous environment. In *Figure 1.14* an example of these packing in presence of two and three strands of peptide chains is reported.

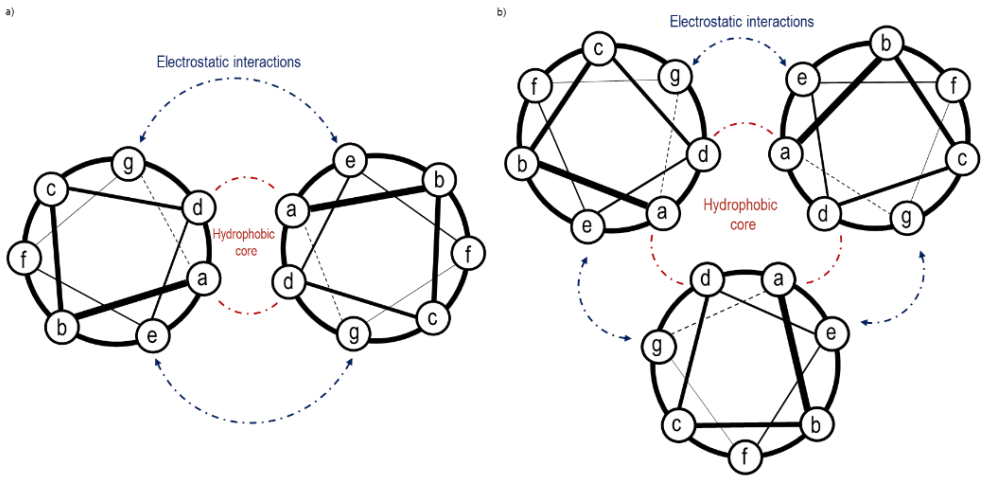



***Figure 1.14.*** *Representation of hydrophobic core and electrostatic interactions generated by a) ) $(abcdefg)_2$ and b) $(abcdefg)_3$ in a coiled-coil motifs.*

Being the seven repeated residues accommodated on two turns of the helix of the left-handed coiled-coil, the periodicity is 3.5 residues, which is less than that of a traditional α-helix, which is 3.6. Instead, the periodicity of the right-handed coiled-coil is 3.67, given that the eleven repeated residues are arranged in three helical turns. Moreover, several terms have been selected to describe the structure of coiled-coils. In fact, *pitch* is used to define the distance of a complete turn, while *pitch angle* is related to the angle generated between each single helical and the central axis. *Helix-crossing angle*, instead, is employed to explain the angle generated between two spatially close helices. These factors are affected by the number of oligomers involved in the formation of coiled-coil systems. Indeed, since the hydrophobic residues *a* and *d* tend to form a hydrophobic core, the *pitch angle* and the *helix-crossing angle* differ based on the number of helices involved, and consequently, on the structures that can be generated.[241] Nowadays, the development of coiled-coil architectures was focused not only on dimer, trimers, or tetramers structures (containing two, three or four helices, respectively) but also to the design of these structures in a more complex way with a greater number of helices, such as from pentamers to dodecamers.[242] Furthermore, coiled-coils can be either homomeric or heteromeric, if respectively formed by α-helices of the same peptide sequence or of different sequences. The orientation of these structures is

parallel if the involved helices have the same C-moiety direct toward the N-moiety, and they are instead antiparallel if the direction are not the same for all the helices. The maximization of non-covalent interactions represents the key factor to obtain coiled-coils in parallel or antiparallel orientations.[243]

General guidelines for explaining the interactions involved in left-handed coiled-coils are highlighted in the *Peptide Velcro* (PV) theory.[244] It outlines the three structural components required for the creation of specific coiled-coils architectures and it was one of the initial approaches in the rational design of such coiled-coils.

The starting point of this theory was that the positions *a* and *d* in the heptad repeated should be characterized by hydrophobic residues like leucine, isoleucine, or valine able to form hydrophobic and van der Waals interactions. Next, electrostatic interactions between the helices should be formed thanks to the presence of charged residues in *e* and *g* positions, like those between lysine and glutamic acid. These amino acid residues in *e* and *g* positions contribute significantly to determine both the preference for forming homo or hetero coiled-coils and the parallel and antiparallel orientation of the helices. Finally, hydrophilic residues should potentially be preferred in *b*, *c*, and *f* positions because they form the solvent-exposed helical surfaces.

Generally, in canonical coiled-coil structures, isoleucine at position *a* and leucine at position *d* lead to the formation of parallel dimers. Conversely, in presence of isoleucine both *a* and *d* position the assembly of trimeric coiled-coils is favored, while four-helices coiled-coil are preferred when leucine residue is at position *a* and *d* position has the isoleucine residue.

Being these characteristics of the peptide Velcro theory not totally mandatory, over the years, several studies were carried out changing some amino acid residues with different features in the three discussed structural components of coiled-coils to obtain more complex structures and various functions.[245]

For example, in fact, from the canonical structure *hpphppp* (where *h* means hydrophobic, instead *p* is related to polar) called Type N, other three typologies of coiled-coil structures were proposed: Type 1, Type 2 and, Type 3. The first type can be

found under two different repeated heptads, like *hpphpph* or *hpphhpp.* The second, instead, is given by *hpphhph* base structure, while the Type 3 can be *hphhphp* or *hhphphp*.[246]

Stable self-assembled of α-helices in canonical coiled-coil structures are characterized by three or four heptad repeats (21-28 residues of amino acids), even at low concentration. Instead, shorter sequences hardly manage to self-assemble into coiled-coils because the number of interactions turn out to be too small to compensate the entropic factor for the distortion of the α-helices, and in fact coiled-coil structures formed by α-helices with 14 amino acids residues tend to form oligomers.[247,248] Although homomeric coiled-coil structures can be formed by controlling the pH, even in the presence of only two heptads, such structures are found to be unstable at room temperature over an extended period.[249]

To compensate for the absence of non-covalent interactions in the assembly of short α-helices in coiled-coils architectures, covalent bonds can be added between the helices. These minimal coiled-coil structures have resulted from various studies that connected relatively short alpha-helices through non-covalent interactions.[250] For example, non-natural mini-proteins CovCore were obtained by addition of a multivalent (bicyclic or tricyclic) crosslinker based on xylyl or mesityl groups to connect the helices.[251]

Other examples where the peptide strands are linked together under a covalent way through a framework structure are represented by the template-assembled synthetic protein (TASP) molecules.[252,253]

Furthermore, parallel, and antiparallel dimers but also trimers of coiled-coil architectures were generated with short peptide oligomers, not able to self-assemble in solution in a coiled-coil motif, through the introduction of a covalent bond which replaced non-covalent interactions between the helices. The covalent bond was formed by bis-triazole bridges generated with an azide-alkyne cycloaddition reaction in presence of copper as a catalyst.[254,255]

Coiled-coil structures have also garnered significant interest in the field of nanotechnology, such as in the assembly of nanotubes through the control of α-

helices self-assembly.[256] However, heterodimeric coiled-coil structure was compatible to interact also with the DNA strands, acting as a driving force for the generation of supramolecular polymer. In fact, DNA-peptide conjugates connected by azido-alkyne cycloaddition reaction, was assembled to construct a 1D nanofibers thanks to the ability of peptide domains to assembly in a coiled-coil structures.[257] In addition, peptide-DNA conjugates are also capable to self-assembling in a complex homotrimer coiled-coil structure as a three-way junction to be used in the formation of nanostructures as DNA cage or to link DNA tiles.[258] This coiled-coil origami field was also expanded to the design and to self-assembly of de novo polyhedral protein, like a tetrahedron, a four-sided pyramid, a triangular prism, and a trigonal bipyramid cage.[259] In addition, it has been studied that the heterodimeric bipyramid architectures can also integrate a proteolysis-triggered structural switch. [260]

Over the years, the repertoire of the coiled-coil structures was expanded in a multiformity architectures, like fibers, tubes, sheets, spirals, hydrogels, and rings. This was possible thank to the acquired knowledge of the delicate equilibrium between attractive and repulsive interactions which are at the foundation of the folding and assembly of coiled-coil structures. Furthermore, they can include not only α-helices but also π-turns and β-strands, such as local alterations in the final coiled-coil structures.[261] Changing the periodicity of the coiled-coil structure with the introduction of 3 or 4 residues, α-helices can be altered with the formation of short β-strand. This led to the formation of a more complex super-secondary structure that include both coiled-coil and β-layer.[262]

Coiled-coil motifs, in specific coiled-coil homodimers, were also used to functionalize the surface of gold nanoparticle (AuNPs), forming a supramolecular complex available to be used as energy-transfer based enzymatic sensor.[263]

Although aromatic amino acids have not been found in high quantities in natural coiled-coil structures, several studies have highlighted that they can be included in these structures as they can contribute to their stabilization through stacking interactions and hydrogen bonds. In fact, phenylalanine (Phe), tyrosine (Tyr),

tryptophane (Trp) are incorporated in a different way to generate coiled-coil structures.[264–267]

Furthermore, the assembling of peptides with precise spatial positioning can be carried out using a DNA scaffold as a template. For example, the Wengel's group presented the self-assembly of short (three heptad) homomeric coiled-coil peptides through oligonucleotide triplex formation. This trimeric complex has a high degree of α-helicity, and the length of the involved peptides influences the stability of the formed coiled-coil structures. To conjugate the peptide to the three oligonucleotides, authors carried out alkyne–azide cycloaddition reactions, exploiting the efficiency of orthogonal chemistry.[268] In a second moment, the attention of the Wengel's group of self-assembly of peptide-oligonucleotide conjugates (POCs), where the peptides form trimeric coiled-coil bundles and the oligonucleotides generate a triple helix, was focused on how the chirality and polarity of the peptide and oligonucleotide sequences regulate the stability of the resulting peptide-oligonucleotide conjugates.[269]

However, the assembly of the de novo peptides can be carried out through DNA origami architectures with covalent linkages. Indeed, the creation of peptide-oligonucleotide conjugates can facilitate the controlled assembly of significant DNA-peptide structures. The peptides' interaction to form coiled coils drives the dimerization of these DNA nanostructures.[270]

Due to their unique properties and versatility, coiled-coil structures have a broad range of applications under multiple fields. Coiled-coils are extensively used in structural and molecular biology to gain insights into protein three-dimensional structures and their interactions. They find applications in biotechnology and nanotechnology, where complex coiled-coil architectures with functional properties can be designed. Material sciences benefit from coiled-coils for material functionalization, and coiled-coils are also significant in drug discovery efforts.[246]

Furthermore, to detect coiled-coil structures, numerous computational programs have been devised, like Socket[271], Coils[272], Paircoil[273], and Multicoil[274].

The formation of coiled-coil structures can be studied by different techniques, employing circular dichroism (CD), fluorimeter, UV-Visible spectrophotometer, mass spectrometry, and NMR spectroscopy. Regarding the typical signature at the CD of α-helix, it exhibits two negative bands at 222 and 208 nm, accompanied by a prominent positive band at approximately 192 nm. Also, the typical CD signature of coiled-coils presents the same negative bands at 222 and 208 nm.[275–277]

Understanding the thermodynamics of self-assembly coiled-coil structures, especially the dissociation constant ($K_D$), is crucial for the practical use of engineered coiled coils. Using the circular dichroism, it is possible to determinate the $K_D$ of the coiled-coil structures through comprehensive analysis of thermal melting curves or by utilizing the midpoint of the thermal unfolding transition at various concentrations. However, these procedures can be time-intensive, demanding multiple meticulous melting experiments with strict control over concentration and peptide ratios. For the determination of $K_D$ regarding coiled-coil structures, the method originally conceived by Marky and Breslauer, and after adapted for coiled coils by Woolfson at Al. can be employed.[278]

Fluorescence-based determination of $K_D$ represents the established norm in studies involving protein-protein and protein-small molecule interactions. This approach may provide improved user-friendliness and potentially enhanced precision compared to the existing techniques employed in coiled-coil systems. In fact, the fluorescence quenching experiments can be employed to determinate the $K_D$ of the coiled-coil formation. However, it is important to use a fluorescent marker that is unable to perturb the assembly of coiled-coil structures.[279]

### 1.4.2 Metalloproteins design

Metalloproteins are macromolecules belong to a protein class that incorporates metal ions as essential components, given that these can contribute to the structural integrity or provide specific functions. Within living organisms, metalloproteins participate in diverse biological processes, contributing to a broad spectrum of functions, like electron transfer[280], gas transport[281], the transport and storage of

transition metal ions[282], as well as catalyzing intricate chemical transformations[283]. Remarkably, these diverse functions are achieved using a restricted set of readily available terrestrial metals and ligands that can be synthesized by biological processes.

Metalloproteins design was focused on the comprehension of the interplay between the secondary and tertiary structures of proteins and the intended primary structure surrounding the metal center. However, in the metal-protein interaction there are two possibilities of complex stabilization, given that the folding of the protein can dictate a specific coordination geometry for the metal ion, or conversely, the coordination of metal ion can influence the geometric arrangement of the protein's tertiary or quaternary structure.

Since the design of metalloproteins is a subcategory of protein design, it too has been divided into redesign and de novo protein design categories. Nevertheless, for every functionally designed metalloprotein system, there is a need for both a robust protein scaffold, and an integrated metal-binding site.[284]

Regarding the metalloproteins redesign category, the correlation between structure and function of a native proteins is deepened to develop novel metal-binding site, to increase the stability, and to generate new functionalities in pre-existing proteins.

However, to introduce a novel metal-binding site in a natural protein, it is necessary to avoid steric hindrance that could disrupt the stability of the folded protein, and to design a coordination geometry able to accommodate the desired metal ions.

Over the years, several protein redesigns were described, like that Zinc Finger (ZFs) proteins to, for example, optimize the interaction between these small protein domains with the DNA, to include a novel metal-binding site or to change the side chains of their structural backbone, and so on. Furthermore, another example of redesign regarded the modification of proteins with heme centers.

However, the introduction of new functionalities, of redox characteristics, and of diverse metal-binding properties are very interesting since the protein redesign can effectively alter both the inner and outer coordination spheres of metal ions.

On the other site, the de novo metalloprotein design is an approach to create novel metalloprotein or mini-metalloprotein with specific tertiary or quaternary structures and functionality, starting from the bottom up and availing of computational methods. In this way the interaction between the metal-ion and the protein was further investigated, and the introduction of specific coordination site in de novo protein scaffold, previously discussed, and the consequent control was performed, such as the creation of imitative functional metalloenzymes. The de novo metalloprotein design focused the attention to develop structures based on both α-helices (like coiled-coils) and β-sheet, but also on αβ-motifs. However, these de novo metalloproteins can be able to bind diverse metal ions, including toxic heavy metals (Hg(II), Cd(II), Pb(II), and, As(II).[285]

Several studies were also published regarding the design of de novo metalloproteins with electron transfer sites and catalytic activity, since at the core of many natural energy transfer processes lie electron transfer reactions.[286,287]

Pecoraro et Al. developed a series of artificial metalloproteins based on TRI family peptides, that can form three-stranded coiled-coil structures. Furthermore, some derivatives of this family were designed to bind metal ion, like Hg(II) and Cd(II), thanks to the introduction of thiol-rich banding site.[288] Subsequently, a TRI family peptide derivative with three histidine residues binding site both for Cu(I) and Cu(II) was developed. This artificial coiled-coil system has proven to be capable of accommodate both copper ions and the reduction of Cu(II) to Cu(I) can be observed in presence of ascorbate, resulting also a system with nitrite reductase activity in aqueous solution.[289] However, modifying these TRI family peptides with a binding site to Zn(II), artificial metalloenzymes with carbonic anhydrase (CA) activity were studied.[290]

In addition, a copper metalloproteins with electron transfer (ET) activity were developed. The de novo design of one of these complex structures derives from TRI-family peptides and it was modified to have a specific binding site that mimic the plastocyanin, given the fact that two histidine residues, one cysteine and one methionine were incorporated.[291,292]

## 1.5 References


1. S. Minchin, J. Lodge. Understanding biochemistry: Structure and function of nucleic acids. *Essays in Biochemistry*. Portland Press Ltd 2019, pp 433–456.
2. J.D. WATSON, F.H.C. CRICK. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, 171 (4356), 737–738.
3. R.O.Y. MARKHAM, J.D. SMITH. Structure of Ribonucleic Acid. *Nature* **1951**, 168 (4271), 406–408.
4. J. Nowakowski, I. Tinoco. RNA Structure and Stability. *Semin Virol* **1997**, 8 (3), 153–165.
5. S.E. Butcher, A.M. Pyle. The Molecular Interactions That Stabilize RNA Tertiary Structure: RNA Motifs, Patterns, and Networks. *Acc Chem Res* **2011**, 44 (12), 1302–1311.
6. N.C. Seeman. Nucleic acid junctions and lattices. *J Theor Biol* **1982**, 99 (2), 237–247.
7. J. Chen, N.C. Seeman. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **1991**, 350 (6319), 631–633.
8. T.J. Fu, N.C. Seeman. DNA double-crossover molecules. *Biochemistry* **1993**, 32 (13), 3211–3220.
9. C. Lin, Y. Liu, S. Rinker, H. Yan. DNA Tile Based Self-Assembly: Building Complex Nanoarchitectures. *ChemPhysChem* **2006**, 7 (8), 1641–1647.
10. N.C. Seeman. An Overview of Structural DNA Nanotechnology. *Mol Biotechnol* **2007**, 37 (3), 246–257.
11. J. Zheng, J.J. Birktoft, Y. Chen, et al. From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **2009**, 461 (7260), 74–77.
12. N.C. Seeman, H.F. Sleiman. DNA nanotechnology. *Nat Rev Mater* **2017**, 3 (1), 17068.
13. P.W.K. Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature* **2006**, 440 (7082), 297–302.
14. B. Wei, M. Dai, P. Yin. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **2012**, 485 (7400), 623–626.
15. F.A. Aldaye, H.F. Sleiman. Sequential Self-Assembly of a DNA Hexagon as a Template for the Organization of Gold Nanoparticles. *Angewandte Chemie International Edition* **2006**, 45 (14), 2204–2209.
16. C. Panagiotidis, S. Kath-Schorr, G. von Kiedrowski. Flexibility of C3h-Symmetrical Linkers in Tris-oligonucleotide-Based Tetrahedral Scaffolds. *ChemBioChem* **2016**, 17 (3), 254–259.
17. C.K. McLaughlin, G.D. Hamblin, H.F. Sleiman. Supramolecular DNA assembly. *Chem Soc Rev* **2011**, 40 (12), 5647–5656.
18. T. Trinh, D. Saliba, C. Liao, et al. "Printing" DNA Strand Patterns on Small Molecules with Control of Valency, Directionality, and Sequence. *Angewandte Chemie International Edition* **2019**, 58 (10), 3042–3047.
19. F.J. Rizzuto, T. Trinh, H.F. Sleiman. Molecular Printing with DNA Nanotechnology. *Chem* **2020**, 6 (7), 1560–1574.
20. D. Saliba, T. Trinh, C. Lachance-Brais, et al. Asymmetric patterning drives the folding of a tripodal DNA nanotweezer. *Chem Sci* **2021**, 13, 74–80.

21. F. Hong, F. Zhang, Y. Liu, H. Yan. DNA Origami: Scaffolds for Creating Higher Order Structures. *Chem Rev* **2017**, 117 (20), 12584–12640.
22. W. Ma, Y. Zhan, Y. Zhang, et al. The biological applications of DNA nanomaterials: current challenges and future directions. *Signal Transduct Target Ther* **2021**, 6 (1), 351.
23. P. Guo. The emerging field of RNA nanotechnology. *Nat Nanotechnol* **2010**, 5 (12), 833–842.
24. D. Jasinski, F. Haque, D.W. Binzel, P. Guo. Advancement of the Emerging Field of RNA Nanotechnology. *ACS Nano* **2017**, 11 (2), 1142–1164.
25. J. Yu, Z. Liu, W. Jiang, G. Wang, C. Mao. De novo design of an RNA tile that self-assembles into a homo-octameric nanoprism. *Nat Commun* **2015**, 6 (1), 5724.
26. W.W. Grabow, L. Jaeger. RNA Self-Assembly and RNA Nanotechnology. *Acc Chem Res* **2014**, 47 (6), 1871–1880.
27. P.E. Nielsen, M. Egholm, R.H. Berg, O. Buchardt. Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide. *Science (1979)* **1991**, 254 (5037), 1497–1500.
28. B. Hyrup, P.E. Nielsen. Peptide Nucleic Acids (PNA): Synthesis, properties and potential applications. *Bioorg Med Chem* **1996**, 4 (1), 5–23.
29. V. V Demidov, N.E. Broude, I. V Lavrentieva-Smolina, H. Kuhn, M.D. Frank-Kamenetskii. An Artificial Primosome: Design, Function, and Applications. *ChemBioChem* **2001**, 2 (2), 133–139.
30. S. V Smulevitch, C.G. Simmons, J.C. Norton, T.W. Wise, D.R. Corey. Enhancement of strand invasion by oligonucleotides through manipulation of backbone charge. *Nat Biotechnol* **1996**, 14 (13), 1700–1704.
31. M. Egholm, O. Buchardt, L. Christensen, et al. PNA hybridizes to complementary oligonucleotides obeying the Watson–Crick hydrogen-bonding rules. *Nature* **1993**, 365 (6446), 566–568.
32. F. Pellestor, P. Paulasova. The peptide nucleic acids (PNAs), powerful tools for molecular genetics and cytogenetics. *European Journal of Human Genetics* **2004**, 12 (9), 694–700.
33. M. Eriksson, P.E. Nielsen. Solution structure of a peptide nucleic acid–DNA duplex. *Nat Struct Biol* **1996**, 3 (5), 410–413.
34. O. Buchardt, M. Egholm, R.H. Berg, P.E. Nielsen. Peptide nucleic acids and their potential applications in biotechnology. *Trends Biotechnol* **1993**, 11 (9), 384–386.
35. V. Menchise, G. De Simone, T. Tedeschi, et al. Insights into peptide nucleic acid (PNA) structural features: The crystal structure of a d-lysine-based chiral PNA–DNA duplex. *Proceedings of the National Academy of Sciences* **2003**, 100 (21), 12021–12026.
36. S. Sen, L. Nilsson. Molecular Dynamics of Duplex Systems Involving PNA: Structural and Dynamical Consequences of the Nucleic Acid Backbone. *J Am Chem Soc* **1998**, 120 (4), 619–631.
37. P.E. Nielsen. PNA technology. *Mol Biotechnol* **2004**, 26 (3), 233–248.
38. P.E. Nielsen. Peptide Nucleic Acids (PNA) in Chemical Biology and Drug Discovery. *Chem Biodivers* **2010**, 7 (4), 786–804.
39. B. Datta, C. Schmitt, B.A. Armitage. Formation of a PNA2–DNA2 Hybrid Quadruplex. *J Am Chem Soc* **2003**, 125 (14), 4111–4118.

40. V.L. Marin, B.A. Armitage. Hybridization of Complementary and Homologous Peptide Nucleic Acid Oligomers to a Guanine Quadruplex-Forming RNA. *Biochemistry* **2006**, 45 (6), 1745–1754.
41. Y. Krishnan-Ghosh, E. Stephens, S. Balasubramanian. A PNA4 Quadruplex. *J Am Chem Soc* **2004**, 126 (19), 5944–5945.
42. P. Wittung, P.E. Nielsen, O. Buchardt, M. Egholm, B. Norde´n. DNA-like double helix formed by peptide nucleic acid. *Nature* **1994**, 368 (6471), 561–563.
43. M. Jasiński, J. Miszkiewicz, M. Feig, J. Trylska. Thermal Stability of Peptide Nucleic Acid Complexes. *J Phys Chem B* **2019**, 123 (39), 8168–8177.
44. H. Rasmussen, J.S. Kastrup, J.N. Nielsen, J.M. Nielsen, P.E. Nielsen. Crystal structure of a peptide nucleic acid (PNA) duplex at 1.7 Å resolution. *Nat Struct Biol* **1997**, 4 (2), 98–101.
45. E. Uhlmann, A. Peyman, G. Breipohl, D.W. Will. PNA: Synthetic Polyamide Nucleic Acids with Unusual Binding Properties. *Angewandte Chemie International Edition* **1998**, 37 (20), 2796–2823.
46. P. Wittung, P. Nielsen, B. Nordén. Observation of a PNA–PNA–PNA Triplex. *J Am Chem Soc* **1997**, 119 (13), 3189–3190.
47. Y. Krishnan-Ghosh, E. Stephens, S. Balasubramanian. PNA forms an i-motif. *Chemical Communications* **2005**, No. 42, 5278–5280.
48. D. Liu, S. Balasubramanian. A Proton-Fuelled DNA Nanomachine. *Angewandte Chemie International Edition* **2003**, 42 (46), 5734–5736.
49. M. Guéron, J.-L. Leroy. The i-motif in nucleic acids. *Curr Opin Struct Biol* **2000**, 10 (3), 326–331.
50. R. Corradini, T. Tedeschi, S. Sforza, R. Marchelli. Electronic Circular Dichroism of Peptide Nucleic Acids and their Analogues. In *Comprehensive Chiroptical Spectroscopy*; **2012**; pp 587–614.
51. P.E. Nielsen, M. Egholm, O. Buchardt. Peptide nucleic acid (PNA). A DNA mimic with a peptide backbone. *Bioconjug Chem* **1994**, 5 (1), 3–7.
52. S. Shakeel, S. Karim, A. Ali. Peptide nucleic acid (PNA) — a review. *Journal of Chemical Technology & Biotechnology* **2006**, 81, 892–899.
53. K.R.B. Singh, P. Sridevi, R.P. Singh. Potential applications of peptide nucleic acid in biomedical domain. *Engineering Reports* **2020**, 2 (9), e12238.
54. R.B. Merrifield. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J Am Chem Soc* **1963**, 85 (14), 2149–2154.
55. K.J. Jensen. Solid-Phase Peptide Synthesis: An Introduction. In *Peptide Synthesis and Applications*; Jensen, K. J., Tofteng Shelton, P., Pedersen, S. L., Eds.; Humana Press, Totowa, NJ, **2013**; pp 1–21.
56. A.Y. Shaikh, F. Björkling, P.E. Nielsen, H. Franzyk. Optimized Synthesis of Fmoc/Boc-Protected PNA Monomers and their Assembly into PNA Oligomers. *European J Org Chem* **2021**, 2021 (19), 2792–2801.
57. D.H. RICH, J. SINGH. Chapter 5 - The Carbodiimide Method. In *Major Methods of Peptide Bond Formation*; GROSS, E., MEIENHOFER, J., Eds.; Academic Press, **1979**; Vol. 1, pp 241–261.

58. Y. Yang. Chapter 5 - Side Reactions Upon Amino Acid/Peptide Carboxyl Activation. In *Side Reactions in Peptide Synthesis*; Yang, Y., Ed.; Academic Press, Oxford, **2016**; pp 95–118.
59. E. Valeur, M. Bradley. Amide bond formation: beyond the myth of coupling reagents. *Chem Soc Rev* **2009**, 38 (2), 606–631.
60. A. El-Faham, R.S. Funosas, R. Prohens, F. Albericio. COMU: A Safer and More Effective Replacement for Benzotriazole-Based Uronium Coupling Reagents. *Chemistry – A European Journal* **2009**, 15 (37), 9404–9416.
61. R. Subirós-Funosas, A. El-Faham, F. Albericio. PyOxP and PyOxB: the Oxyma-based novel family of phosphonium salts. *Org Biomol Chem* **2010**, 8 (16), 3665–3673.
62. T.I. Al-Warhi, H.M.A. Al-Hazimi, A. El-Faham. Recent development in peptide coupling reagents. *Journal of Saudi Chemical Society* **2012**, 16 (2), 97–116.
63. F. Albericio, A. El-Faham. Choosing the Right Coupling Reagent for Peptides: A Twenty-Five-Year Journey. *Org Process Res Dev* **2018**, 22 (7), 760–772.
64. P. Cherkupally, G.A. Acosta, L. Nieto-Rodriguez, et al. K-Oxyma: a Strong Acylation-Promoting, 2-CTC Resin-Friendly Coupling Additive. *European J Org Chem* **2013**, 2013 (28), 6372–6378.
65. A. El-Faham, F. Albericio. Peptide Coupling Reagents, More than a Letter Soup. *Chem Rev* **2011**, 111 (11), 6557–6602.
66. L. Ferrazzano, D. Corbisiero, G. Martelli, et al. Green Solvent Mixtures for Solid-Phase Peptide Synthesis: A Dimethylformamide-Free Highly Efficient Synthesis of Pharmaceutical-Grade Peptides. *ACS Sustain Chem Eng* **2019**, 7 (15), 12867–12877.
67. V. Declerck, P. Nun, J. Martinez, F. Lamaty. Solvent-Free Synthesis of Peptides. *Angewandte Chemie International Edition* **2009**, 48 (49), 9318–9321.
68. K.G. Varnava, V. Sarojini. Making Solid-Phase Peptide Synthesis Greener: A Review of the Literature. *Chem Asian J* **2019**, 14 (8), 1088–1097.
69. Y.E. Jad, A. Kumar, A. El-Faham, B.G. de la Torre, F. Albericio. Green Transformation of Solid-Phase Peptide Synthesis. *ACS Sustain Chem Eng* **2019**, 7 (4), 3671–3683.
70. C. Li, A.J. Callahan, K.S. Phadke, et al. Automated Flow Synthesis of Peptide–PNA Conjugates. *ACS Cent Sci* **2022**, 8 (2), 205–213.
71. G. Barany, R.B. Merrifield. A new amino protecting group removable by reduction. Chemistry of the dithiasuccinoyl (Dts) function. *J Am Chem Soc* **1977**, 99 (22), 7363–7365.
72. C.-H. Wong, S.C. Zimmerman. Orthogonality in organic, polymer, and supramolecular chemistry: from Merrifield to click chemistry. *Chemical Communications* **2013**, 49 (17), 1679–1695.
73. K.P. Nandhini, D. Al Shaer, F. Albericio, B.G. de la Torre. The challenge of peptide nucleic acid synthesis. *Chem Soc Rev* **2023**, 52 (8), 2764–2789.
74. G. Breipohl, J. Knolle, D. Langner, G.J. O'malley, E.Dr. Uhlmann. Synthesis of Polyamide Nucleic Acids (PNAs) Using a Novel Fmoc/Mmt Protecting-Group Combination. *ChemInform* **1996**, 27.
75. D. Chouikhi, M. Ciobanu, C. Zambaldo, et al. Expanding the Scope of PNA-Encoded Synthesis (PES): Mtt-Protected PNA Fully Orthogonal to Fmoc Chemistry and a Broad Array of Robust Diversity-Generating Reactions. *Chemistry – A European Journal* **2012**, 18 (40), 12698–12704.

76. A. ALETRAS, K. BARLOS, D. GATOS, S. KOUTSOGIANNI, P. MAMOS. Preparation of the very acid-sensitive Fmoc-Lys(Mtt)-OH Application in the synthesis of side-chain to side-chain cyclic peptides and oligolysine cores suitable for the solid-phase assembly of MAPs and TASPs. *Int J Pept Protein Res* **1995,** 45 (5), 488–496.
77. S. Pensato, M. Renda, F. Leccia, et al. PNA zipper as a dimerization tool: Development of a bZip mimic. *Biopolymers* **2010,** 93 (5), 434–441.
78. Y. Altrichter, O. Seitz. Simultaneous Targeting of Two Master Regulators of Apoptosis with Dual-Action PNA– and DNA–Peptide Conjugates. *Bioconjug Chem* **2020**, 31 (8), 1928–1937.
79. K. AUGUSTYNS, W. KRAAS, G. JUNG. Investigation on the stability of the Dde protecting group used in peptide synthesis: migration to an unprotected lysine1. *The Journal of Peptide Research* **1998**, 51 (2), 127–133.
80. J.J. Díaz-Mochón, L. Bialy, M. Bradley. Full Orthogonality between Dde and Fmoc: The Direct Synthesis of PNA–Peptide Conjugates. *Org Lett* **2004,** 6 (7), 1127–1129.
81. L. Bialy, J.J. Díaz-Mochón, E. Specker, L. Keinicke, M. Bradley. Dde-protected PNA monomers, orthogonal to Fmoc, for the synthesis of PNA–peptide conjugates. *Tetrahedron* **2005,** 61 (34), 8295–8305.
82. F. Debaene, N. Winssinger. Azidopeptide Nucleic Acid. An Alternative Strategy for Solid-Phase Peptide Nucleic Acid (PNA) Synthesis. *Org Lett* **2003**, 5 (23), 4445–4447.
83. C. Bednarek, I. Wehl, N. Jung, U. Schepers, S. Bräse. The Staudinger Ligation. *Chem Rev* **2020**, 120 (10), 4301–4354.
84. W.H. Binder, C. Kluger. Azide—Alkyne "Click" Reactions: Applications in Material Science and Organic Synthesis. *ChemInform* **2007**, 38 (4).
85. F. Müggenburg, S. Müller. Azide-Modified Nucleosides as Versatile Tools for Bioorthogonal Labeling and Functionalization. *The Chemical Record* **2022**, 22 (5), e202100322.
86. A. Manicardi, A. Bertucci, A. Rozzi, R. Corradini. A Bifunctional Monomer for On-Resin Synthesis of Polyfunctional PNAs and Tailored Induced-Fit Switching Probes. *Org Lett* **2016**, 18 (21), 5452–5455.
87. N. Svensen, J.J. Díaz-Mochón, M. Bradley. Microwave-assisted orthogonal synthesis of PNA–peptide conjugates. *Tetrahedron Lett* **2008**, 49 (46), 6498–6500.
88. J.J. Díaz-Mochón, L. Bialy, J. Watson, R.M. Sánchez-Martín, M. Bradley. Synthesis and cellular uptake of cell delivering PNA–peptide conjugates. *Chemical Communications* **2005**, No. 26, 3316–3318.
89. B.E. Hubrich, P.M. Menzel, B. Kugler, U. Diederichsen. Synthesis of PNA-Peptide Conjugates as Functional SNARE Protein Mimetics. In *Peptide Nucleic Acids: Methods and Protocols*; Nielsen, P. E., Ed.; Springer US, New York, NY, **2020**; pp 61–74.
90. A. Erben, O. Seitz. Synthesis of Nucleic Acid–Peptide Conjugates Targeted to Proteins. *Isr J Chem* **2011**, 51 (8–9), 876–884.
91. F. Burlina, D.D. Dixson, R.P. Doyle, et al. Orthogonal ligation: a three piece assembly of a PNA–peptide–PNA conjugate. *Chemical Communications* **2008**, No. 24, 2785–2787.
92. M.C. de Koning, G.A. van der Marel, M. Overhand. Synthetic developments towards PNA–peptide conjugates. *Curr Opin Chem Biol* **2003,** 7 (6), 734–740.

93. J.P. Tam, Q. Yu, Z. Miao. Orthogonal ligation strategies for peptide and protein. *Peptide Science* **1999,** 51 (5), 311–332.
94. E. Saxon, C.R. Bertozzi. Cell Surface Engineering by a Modified Staudinger Reaction. *Science (1979)* **2000,** 287 (5460), 2007–2010.
95. P.E. Nielsen, M. Egholm, R.H. Berg, O. Buchardt. Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide. *Science (1979)* **1991,** 254 (5037), 1497–1500.
96. V.L. Marin, S. Roy, B.A. Armitage. Recent advances in the development of peptide nucleic acid as a gene-targeted drug. *Expert Opin Biol Ther* **2004,** 4 (3), 337–348.
97. S. Shakeel, S. Karim, A. Ali. Peptide nucleic acid (PNA) — a review. *Journal of Chemical Technology & Biotechnology* **2006,** 81 (6), 892–899.
98. P.E. Nielsen, M. Egholm, O. Buchardt. Sequence-specific transcription arrest by peptide nucleic acid bound to the DNA template strand. *Gene* **1994,** 149 (1), 139–145.
99. M. Pooga, T. Land, T. Bartfai, Ü. Langel. PNA oligomers as tools for specific modulation of gene expression. *Biomol Eng* **2001,** 17 (6), 183–192.
100. M. Egholm, L. Christensen, K.L. Deuholm, et al. Efficient pH-independent sequence-specific DNA binding by pseudoisocytosine-containing bis-PNA. *Nucleic Acids Res* **1995,** 23 (2), 217–222.
101. P.E. Nielsen. Peptide nucleic acids as therapeutic agents. *Curr Opin Struct Biol* **1999,** 9 (3), 353–357.
102. B. Armitage, T. Koch, H. Frydenlund, H. Orum, G.B. Schuster. Peptide nucleic acid (PNA)/DNA hybrid duplexes: intercalation by an internally linked anthraquinone. *Nucleic Acids Res* **1998,** 26 (3), 715–720.
103. J.D.R. Perera, K.E.W. Carufe, P.M. Glazer. Peptide nucleic acids and their role in gene regulation and editing. *Biopolymers* **2021,** 112 (12), e23460.
104. R. Tonelli, A. McIntyre, C. Camerin, et al. Antitumor activity of sustained N-myc reduction in rhabdomyosarcomas and transcriptional block by antigene therapy. *Clinical Cancer Research* **2012,** 18 (3), 796–807.
105. D.A. Dean. Peptide nucleic acids: versatile tools for gene therapy strategies. *Adv Drug Deliv Rev* **2000,** 44 (2), 81–95.
106. F. Pellestor, P. Paulasova. The peptide nucleic acids (PNAs), powerful tools for molecular genetics and cytogenetics. *European Journal of Human Genetics* **2004,** 12 (9), 694–700.
107. J.C. Hanvey, N.J. Peffer, J.E. Bisi, et al. Antisense and Antigene Properties of Peptide Nucleic Acids. *Science (1979)* **1992,** 258 (5087), 1481–1485.
108. T. Shiraishi, J. Eysturskarð, P.E. Nielsen. Modulation of mdm2 pre-mRNA splicing by 9-aminoacridine-PNA (peptide nucleic acid) conjugates targeting intron-exon junctions; **2010;** Vol. 10.
109. D.G. Ahn, W. Lee, J.K. Choi, et al. Interference of ribosomal frameshifting by antisense peptide nucleic acids suppresses SARS coronavirus replication. *Antiviral Res* **2011,** 91 (1), 1–10.
110. M.M. Fabani, C. Abreu-Goodger, D. Williams, et al. Efficient inhibition of miR-155 function in vivo by peptide nucleic acids. *Nucleic Acids Res* **2010,** 38 (13), 4466–4475.

111. G.D. Ivanova, M.M. Fabani, A.A. Arzumanov, et al. PNA-peptide conjugates as intracellular gene control agents. *Nucleic Acids Symp Ser (Oxf)* **2008**, No. 52, 31–32.
112. A.G. Torres, M.M. Fabani, E. Vigorito, M.J. Gait. MicroRNA fate upon targeting with anti-miRNA oligonucleotides as revealed by an improved Northern-blot-based method for miRNA detection. *RNA* **2011**, 17 (5), 933–943.
113. E. Fabbri, E. Brognara, M. Borgatti, et al. miRNA therapeutics: Delivery and biological activity of peptide nucleic acids targeting miRNAs. *Epigenomics*. December 2011, pp 733–745.
114. E. Brognara, E. Fabbri, G. Montagner, et al. High levels of apoptosis are induced in human glioma cell lines by co-administration of peptide nucleic acids targeting miR-221 and miR-222. *Int J Oncol* **2016**, 48 (3), 1029–1038.
115. S. Malik, J. Lim, F.J. Slack, D.T. Braddock, R. Bahal. Next generation miRNA inhibition using short anti-seed PNAs encapsulated in PLGA nanoparticles. *Journal of Controlled Release* **2020**, 327, 406–419.
116. T. Soudah, S. Khawaled, R.I. Aqeilan, E. Yavin. AntimiR-155 Cyclic Peptide–PNA Conjugate: Synthesis, Cellular Uptake, and Biological Activity. *ACS Omega* **2019**, 4 (9), 13954–13961.
117. A. Gupta, E. Quijano, Y. Liu, et al. Anti-tumor Activity of miniPEG-γ-Modified PNAs to Inhibit MicroRNA-210 for Cancer Therapy. *Mol Ther Nucleic Acids* **2017**, 9, 111–119.
118. J. Gasparello, C. Papi, R. Gambari, et al. High Levels of Apoptosis Are Induced in the Human Colon Cancer HT-29 Cell Line by Co-Administration of Sulforaphane and a Peptide Nucleic Acid Targeting miR-15b-5p. *Nucleic Acid Ther* **2020**, 30 (3), 164–174.
119. M.E. Hansen, T. Bentin, P.E. Nielsen. High-affinity triplex targeting of double stranded DNA using chemically modified peptide nucleic acid oligomers. *Nucleic Acids Res* **2009**, 37 (13), 4498–4507.
120. C. Suparpprom, T. Vilaivan. Perspectives on conformationally constrained peptide nucleic acid (PNA): insights into the structural design, properties and applications. *RSC Chemical Biology*. Royal Society of Chemistry March 18, 2022, pp 648–697.
121. A. Manicardi, A. Calabretta, M. Bencivenni, et al. Affinity and selectivity of C2- and C5-substituted "chiral-box" PNA in solution and on microarrays. In *Chirality*; **2010**; Vol. 22.
122. S. Pensato, M. Saviano, N. Bianchi, et al. γ-Hydroxymethyl PNAs: Synthesis, interaction with DNA and inhibition of protein/DNA interactions. *Bioorg Chem* **2010**, 38 (5), 196–201.
123. B.B. Mendes, J. Conniot, A. Avital, et al. Nanodelivery of nucleic acids. *Nature Reviews Methods Primers*. Springer Nature December 1, 2022.
124. U. Koppelhus, P.E. Nielsen. C ellular delivery of peptide nucleic acid (PNA); **2003**; Vol. 55.
125. S. Volpi, U. Cancelli, M. Neri, R. Corradini. Multifunctional delivery systems for peptide nucleic acids. *Pharmaceuticals*. MDPI AG December 1, 2021, pp 1–31.
126. G. Wang, X. Xu, B. Pace, et al. Peptide nucleic acid (PNA) binding-mediated induction of human γ-globin gene expression. *Nucleic Acids Res* **1999**, 27 (13), 2806–2813.

127. S.Z. Hirschman, C.W. Chen. Peptide nucleic acids stimulate gamma interferon and inhibit the replication of the human immunodeficiency virus. *J Investig Med* **1996**, 44 (6), 347–351.
128. A. Dragulescu-Andrasi, P. Zhou, G. He, D.H. Ly. Cell-permeable GPNA with appropriate backbone stereochemistry and spacing binds sequence-specifically to RNA. *Chemical Communications* **2005**, No. 2, 244–246.
129. A. Dragulescu-Andrasi, S. Rapireddy, B.M. Frezza, et al. A Simple γ-Backbone Modification Preorganizes Peptide Nucleic Acid into a Helical Structure. *J Am Chem Soc* **2006**, 128 (31), 10258–10267.
130. A. Manicardi, R. Gambari, L. de Cola, R. Corradini. Preparation of Anti-miR PNAs for Drug Development and Nanomedicine. In *DNA Nanotechnology: Methods and Protocols*; Zuccheri, G., Ed.; Springer New York, New York, NY, **2018**; pp 49–63.
131. T. Ljungstrøm, H. Knudsen, P.E. Nielsen. Cellular Uptake of Adamantyl Conjugated Peptide Nucleic Acids. *Bioconjug Chem* **1999**, 10 (6), 965–972.
132. A. Muratovska, R.N. Lightowlers, R.W. Taylor, et al. Targeting peptide nucleic acid (PNA) oligomers to mitochondria within cells by conjugation to lipophilic cations:implications for mitochondrial DNA replication, expression and disease. *Nucleic Acids Res* **2001**, 29 (9), 1852–1863.
133. S.M.W. van Rossenberg, K.M. Sliedregt-Bol, P. Prince, et al. A Targeted Peptide Nucleic Acid To Down-Regulate Mouse Microsomal Triglyceride Transfer Protein Expression in Hepatocytes. *Bioconjug Chem* **2003**, 14 (6), 1077–1082.
134. E.A.L. Biessen, K. Sliedregt-Bol, P.A.Chr. 'T Hoen, et al. Design of a Targeted Peptide Nucleic Acid Prodrug To Inhibit Hepatic Human Microsomal Triglyceride Transfer Protein Expression in Hepatocytes. *Bioconjug Chem* **2002**, 13 (2), 295–302.
135. I.M. Gabas, P.E. Nielsen. Effective Cellular Delivery of Antisense Peptide Nucleic Acid by Conjugation to Guanidinylated Diaminobutanoic Acid-Based Peptide Dendrons. *Biomacromolecules* **2020**, 21 (2), 472–483.
136. G. Shen, H. Fang, Y. Song, et al. Phospholipid Conjugate for Intracellular Delivery of Peptide Nucleic Acids. *Bioconjug Chem* **2009**, 20 (9), 1729–1736.
137. M. Zorko, Ü. Langel. Cell-penetrating peptides: mechanism and kinetics of cargo delivery. *Adv Drug Deliv Rev* **2005**, 57 (4), 529–545.
138. D.M. Copolovici, K. Langel, E. Eriste, Ü. Langel. Cell-Penetrating Peptides: Design, Synthesis, and Applications. *ACS Nano* **2014**, 8 (3), 1972–1994.
139. T. Soudah, M. Mogilevsky, R. Karni, E. Yavin. CLIP6-PNA-Peptide Conjugates: Non-Endosomal Delivery of Splice Switching Oligonucleotides. *Bioconjug Chem* **2017**, 28 (12), 3036–3042.
140. N. Brodyagin, Y. Kataoka, I. Kumpina, D.W. McGee, E. Rozners. Cellular uptake of 2-aminopyridine-modified peptide nucleic acids conjugated with cell-penetrating peptides. *Biopolymers* **2022**, 113 (4), e23484.
141. R. Abes, A. Arzumanov, H. Moulton, et al. Arginine-rich cell penetrating peptides: Design, structure–activity, and applications to alter pre-mRNA splicing by steric-block oligonucleotides. *Journal of Peptide Science* **2008**, 14 (4), 455–460.
142. J.J. Turner, G.D. Ivanova, B. Verbeure, et al. Cell-penetrating peptide conjugates of peptide nucleic acids (PNA) as inhibitors of HIV-1 Tat-dependent trans -activation in cells. *Nucleic Acids Res* **2005**, 33 (21), 6837–6849.

143. B. Chaubey, S. Tripathi, V.N. Pandey. Single Acute-Dose and Repeat-Doses Toxicity of anti-HIV-1 PNATAR–Penetratin Conjugate after Intraperitoneal Administration to Mice. *Oligonucleotides* **2008**, 18 (1), 9–20.
144. R. Abes, A.A. Arzumanov, H.M. Moulton, et al. Cell-penetrating-peptide-based delivery of oligonucleotides: an overview. *Biochem Soc Trans* **2007**, 35 (4), 775–779.
145. T. Shiraishi, M. Ghavami, P.E. Nielsen. In Vitro Cellular Delivery of Peptide Nucleic Acid (PNA). In *Peptide Nucleic Acids: Methods and Protocols*; Nielsen, P. E., Ed.; Springer US, New York, NY, **2020**; pp 173–185.
146. H. Lv, S. Zhang, B. Wang, S. Cui, J. Yan. Toxicity of cationic lipids and cationic polymers in gene delivery. *Journal of Controlled Release* **2006**, 114 (1), 100–109.
147. S.E. Hamilton, C.G. Simmons, I.S. Kathiriya, D.R. Corey. Cellular delivery of peptide nucleic acids and inhibition of human telomerase. *Chem Biol* **1999**, 6, 343–351.
148. Y.M. Bae, M.H. Kim, G.S. Yu, et al. Enhanced splicing correction effect by an oligo-aspartic acid–PNA conjugate and cationic carrier complexes. *Journal of Controlled Release* **2014**, 175, 54–62.
149. Z. Zhang, Y. Liu, C. Jarreau, M.J. Welch, J.-S.A. Taylor. Nucleic acid-directed self-assembly of multifunctional gold nanoparticle imaging agents. *Biomater Sci* **2013**, 1 (10), 1055–1064.
150. N. Grimaldi, F. Andrade, N. Segovia, et al. Lipid-based nanovesicles for nanomedicine. *Chem Soc Rev* **2016**, 45 (23), 6520–6545.
151. S. Grijalvo, G. Puras, J. Zárate, et al. Cationic niosomes as non-viral vehicles for nucleic acids: Challenges and opportunities in gene delivery. *Pharmaceutics*. MDPI AG February 1, 2019.
152. A. Gupta, E. Quijano, Y. Liu, et al. Anti-tumor Activity of miniPEG-γ-Modified PNAs to Inhibit MicroRNA-210 for Cancer Therapy. *Mol Ther Nucleic Acids* **2017**, 9, 111–119.
153. H. Fang, K. Zhang, G. Shen, K.L. Wooley, J.-S.A. Taylor. Cationic Shell-Cross-Linked Knedel-like (cSCK) Nanoparticles for Highly Efficient PNA Delivery. *Mol Pharm* **2009**, 6 (2), 615–626.
154. H.K. Makadia, S.J. Siegel. Poly Lactic-co-Glycolic Acid (PLGA) as biodegradable controlled drug delivery carrier. *Polymers (Basel)* **2011**, 3 (3), 1377–1397.
155. F. Danhier, E. Ansorena, J.M. Silva, et al. PLGA-based nanoparticles: An overview of biomedical applications. *Journal of Controlled Release* **2012**, 161 (2), 505–522.
156. C.F. van Nostrum. Covalently cross-linked amphiphilic block copolymer micelles. *Soft Matter* **2011**, 7 (7), 3246–3259.
157. M. Elsabahy, K.L. Wooley. Design of polymeric nanoparticles for biomedical delivery applications. *Chem Soc Rev* **2012**, 41 (7), 2545–2561.
158. A. Bertucci, H. Lülf, D. Septiadi, et al. Intracellular Delivery of Peptide Nucleic Acid and Organic Molecules Using Zeolite-L Nanocrystals. *Adv Healthc Mater* **2014**, 3 (11), 1812–1817.
159. A. Bertucci, E.A. Prasetyanto, D. Septiadi, et al. Combined Delivery of Temozolomide and Anti-miR221 PNA Using Mesoporous Silica Nanoparticles Induces Apoptosis in Resistant Glioma Cells. *Small* **2015**, 11 (42), 5687–5695.

160. K.R. Beavers, J.W. Mares, C.M. Swartz, et al. In Situ Synthesis of Peptide Nucleic Acids in Porous Silicon for Drug Delivery and Biosensing. *Bioconjug Chem* **2014,** 25 (7), 1192–1197.
161. E. Ghaffari, S.E. Rezatofighi, M.R. Ardakani, S. Rastegarzadeh. Delivery of antisense peptide nucleic acid by gold nanoparticles for the inhibition of virus replication. *Nanomedicine* **2019,** 14 (14), 1827–1840.
162. J. Kim, S.-J. Park, D.-H. Min. Emerging Approaches for Graphene Oxide Biosensor. *Anal Chem* **2017,** 89 (1), 232–248.
163. J. Lee, J. Kim, S. Kim, D.-H. Min. Biosensors based on graphene oxide and its biomedical application. *Adv Drug Deliv Rev* **2016,** 105, 275–287.
164. A. Baek, Y.M. Baek, H.-M. Kim, B.-H. Jun, D.-E. Kim. Polyethylene Glycol-Engrafted Graphene Oxide as Biocompatible Materials for Peptide Nucleic Acid Delivery into Cells. *Bioconjug Chem* **2018,** 29 (2), 528–537.
165. X. Liao, Q. Wang, H. Ju. A peptide nucleic acid-functionalized carbon nitride nanosheet as a probe for in situ monitoring of intracellular microRNA. *Analyst* **2015,** 140 (12), 4245–4252.
166. J. Gasparello, A. Manicardi, A. Casnati, et al. Efficient cell penetration and delivery of peptide nucleic acids by an argininocalix[4]arene. *Sci Rep* **2019,** 9 (1).
167. S. Tomassi, C. Ieranò, M.E. Mercurio, et al. Cationic nucleopeptides as novel non-covalent carriers for the delivery of peptide nucleic acid (PNA) and RNA oligomers. *Bioorg Med Chem* **2018,** 26 (9), 2539–2550.
168. M. Moccia, A. Antonacci, M. Saviano, et al. Emerging technologies in the design of peptide nucleic acids (PNAs) based biosensors. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. November 1, 2020.
169. A. Saadati, S. Hassanpour, M. de la Guardia, et al. Recent advances on application of peptide nucleic acids as a bioreceptor in biosensors development. *TrAC - Trends in Analytical Chemistry*. Elsevier B.V. May 1, 2019, pp 56–68.
170. P. Damborský, J. Švitel, J. Katrlík. Optical biosensors. *Essays Biochem* **2016,** 60 (1), 91–100.
171. N. Svanvik, G. Westman, D. Wang, M. Kubista. Light-Up Probes: Thiazole Orange-Conjugated Peptide Nucleic Acid for Detection of Target Nucleic Acid in Homogeneous Solution. *Anal Biochem* **2000,** 281 (1), 26–35.
172. P. Wolffs, R. Knutsson, R. Sjöback, P. Rådström. PNA-Based Light-Up Probes for Real-Time Detection of Sequence-Specific PCR Products. *Biotechniques* **2001,** 31 (4), 766–771.
173. J. Isacsson, H. Cao, L. Ohlsson, et al. Rapid and specific detection of PCR products using light-up probes. *Mol Cell Probes* **2000,** 14 5, 321–328.
174. A. Candiani, M. Sozzi, A. Cucinotta, et al. Optical fiber ring cavity sensor for label-free DNA detection. *IEEE Journal on Selected Topics in Quantum Electronics* **2012,** 18 (3), 1176–1183.
175. A. Candiani, A. Bertucci, S. Giannetti, et al. Label-free DNA biosensor based on a peptide nucleic acid-functionalized microstructured optical fiber-Bragg grating. *J Biomed Opt* **2013,** 18 (5), 057004.

176. A. Bertucci, A. Manicardi, A. Candiani, et al. Detection of unamplified genomic DNA by a PNA-based microstructured optical fiber (MOF) Bragg-grating optofluidic system. *Biosens Bioelectron* **2015**, 63, 248–254.
177. M. Barozzi, A. Manicardi, A. Vannucci, et al. Optical Fiber Sensors for Label-Free DNA Detection. *Journal of Lightwave Technology* **2017**, 35 (16), 3461–3472.
178. F. Giovanardi, A. Cucinotta, L. Vincetti. Inhibited coupling guiding hollow fibers for label-free DNA detection. *Opt Express* **2017**, 25 (21), 26215–26220.
179. F. Giovanardi, A. Cucinotta, A. Rozzi, et al. Protein Detection Using Hollow-Core Tube Lattice Fibers. In *Advanced Photonics 2018 (BGPP, IPR, NP, NOMA, Sensors, Networks, SPPCom, SOF)*; OSA Technical Digest (online); Optica Publishing Group, Zurich, **2018**; p SeM3E.4.
180. F. Giovanardi, A. Cucinotta, A. Rozzi, et al. Hollow Core Inhibited Coupling Fibers for Biological Optical Sensing. *Journal of Lightwave Technology* **2019**, 37 (11), 2598–2604.
181. D.R. Thévenot, K. Toth, R.A. Durst, G.S. Wilson, D.R. Thévenot. Electrochemical biosensors: recommended definitions and classification; **2001**; Vol. 16.
182. M. Labib, E.H. Sargent, S.O. Kelley. Electrochemical Methods for the Analysis of Clinically Relevant Biomolecules. *Chemical Reviews*. American Chemical Society August 24, 2016, pp 9001–9090.
183. J. Wang, P.E. Nielsen, M. Jiang, et al. Mismatch-Sensitive Hybridization Detection by Peptide Nucleic Acids Immobilized on a Quartz Crystal Microbalance. *Anal Chem* **1997**, 69 (24), 5200–5202.
184. S. Reisberg, L.A. Dang, Q.A. Nguyen, et al. Label-free DNA electrochemical sensor based on a PNA-functionalized conductive polymer. *Talanta* **2008**, 76 (1), 206–210.
185. H. Aoki, Y. Umezawa. Trace analysis of an oligonucleotide with a specific sequence using PNA-based ion-channel sensors. *Analyst* **2003**, 128 (6), 681–685.
186. E. Cadoni, A. Manicardi, A. Madder. PNA-based microRNA detection methodologies. *Molecules*. MDPI AG March 2, 2020.
187. P. Jolly, M.R. Batistuti, A. Miodek, et al. Highly sensitive dual mode electrochemical platform for microRNA detection. *Sci Rep* **2016**, 6.
188. P. Fu, S. Xing, M. Xu, Y. Zhao, C. Zhao. Peptide nucleic acid-based electrochemical biosensor for simultaneous detection of multiple microRNAs from cancer cells with catalytic hairpin assembly amplification. *Sens Actuators B Chem* **2020**, 305, 127545.
189. S. Fortunati, C. Giliberti, M. Giannetto, et al. A highly sensitive electrochemical magneto-genosensing assay for the specific detection of a single nucleotide variation in the KRAS oncogene in human plasma. *Biosens Bioelectron X* **2023**, 15.
190. Mklos Bodanszky. Peptide Chemistry; Springer-Verlag, Ed.; **1988**.
191. E. Fischer, E. Fourneau. Ueber einige Derivate des Glykocolls. *Berichte der deutschen chemischen Gesellschaft* **1901**, 34 (2), 2868–2877.
192. M. Bergmann, L. Zervas. Über ein allgemeines Verfahren der Peptid-Synthese. *Berichte der deutschen chemischen Gesellschaft (A and B Series)* **1932**, 65 (7), 1192–1201.
193. J.C. Sheehan, G.P. Hess. A New Method of Forming Peptide Bonds. *J Am Chem Soc* **1955**, 77 (4), 1067–1068.

194. F. Albericio. Orthogonal protecting groups for Nα-amino and C-terminal carboxyl functions in solid-phase peptide synthesis. *Peptide Science* **2000**, 55 (2), 123–139.
195. G.B. Fields. Introduction to peptide synthesis. *Curr Protoc Protein Sci* **2002**, Chapter 18, Unit-18.1.
196. F. Albericio. Orthogonal protecting groups for Nα-amino and C-terminal carboxyl functions in solid-phase peptide synthesis. *Peptide Science* **2000**, 55 (2), 123–139.
197. N. Trier, P. Hansen, G. Houen. Peptides, Antibodies, Peptide Antibodies and More. *Int J Mol Sci* **2019**, 20 (24), 6289.
198. D. Eisenberg. The discovery of the α-helix and β-sheet, the principal structural features of proteins. *Proceedings of the National Academy of Sciences* **2003**, 100 (20), 11207–11210.
199. A.M.C. Marcelino, L.M. Gierasch. Roles of β-turns in protein folding: From peptide models to protein engineering. *Biopolymers* **2008**, 89 (5), 380–391.
200. A.G. De Brevern. Extension of the classical classification of β-turns. *Sci Rep* **2016**, 6.
201. R. Zhang, M.C. Stahr, M.A. Kennedy. Introduction of a new scheme for classifying β-turns in protein structures. *Proteins: Structure, Function and Bioinformatics* **2022**, 90 (1), 110–122.
202. M.E. Karpen, P.L. De Haseth, K.E. Neet. Differences in the amino acid distributions of 310-helices and α-helices. *Protein Science* **1992**, 1 (10), 1333–1342.
203. R. Armen, D.O.V. Alonso, V. Daggett. The role of α-, 3 10 -, and π-helix in helix→coil transitions . *Protein Science* **2003**, 12 (6), 1145–1157.
204. P. Kumar, N.G. Paterson, J. Clayden, D.N. Woolfson. De novo design of discrete, stable 310-helix peptide assemblies. *Nature* **2022**, 607 (7918), 387–392.
205. D. Voet, J.G. Voet, C.W. Pratt. Principles of Biochemistry. *John Wiley & Sons.* **2008**.
206. D.J. Dietzen. 13 - Amino Acids, Peptides, and Proteins. In *Principles and Applications of Molecular Diagnostics*; Rifai, N., Horvath, A. R., Wittwer, C. T., Eds.; Elsevier, **2018**; pp 345–380.
207. J. Venkatraman, S.C. Shankaramma, P. Balaram. Design of Folded Peptides. *Chem Rev* **2001**, 101 (10), 3131–3152.
208. G. Caetano-Anollés, M. Wang, D. Caetano-Anollés, J.E. Mittenthal. The origin, evolution and structure of the protein world. *Biochemical Journal* **2009**, 417 (3), 621–637.
209. P.-S. Huang, S.E. Boyken, D. Baker. The coming of age of de novo protein design. *Nature* **2016**, 537 (7620), 320–327.
210. L. Regan. Protein redesign. *Curr Opin Struct Biol* **1999**, 9 (4), 494–499.
211. D. Berta, S. Gehrke, K. Nyíri, B.G. Vértessy, E. Rosta. Mechanism-Based Redesign of GAP to Activate Oncogenic Ras. *J Am Chem Soc* **2023**, 145 (37), 20302–20310.
212. W. Treewattanawong, T. Sitthiyotha, S. Chunsrivirot. Computational redesign of Beta-27 Fab with substantially better predicted binding affinity to the SARS-CoV-2 Omicron variant than human ACE2 receptor. *Sci Rep* **2023**, 13 (1), 15476.
213. C. Lv, X. Zhang, Y. Liu, et al. Redesign of protein nanocages: the way from 0D, 1D, 2D to 3D assembly. *Chem Soc Rev* **2021**, 50 (6), 3957–3989.
214. I. V Korendovych, W.F. DeGrado. De novo protein design, a retrospective. *Q Rev Biophys* **2020**, 53, e3.

215. B. Gutte, M. Däumigen, E. Wittschieber. Design, synthesis and characterisation of a 34-residue polypeptide that interacts with nucleic acids. *Nature* **1979**, 281 (5733), 650–655.
216. A. Lombardi, C.M. Summa, S. Geremia, et al. Retrostructural analysis of metalloproteins: Application to the design of a minimal model for diiron proteins. *Proceedings of the National Academy of Sciences* **2000**, 97 (12), 6298–6305.
217. S.H. Hong, T. Nguyen, P. Arora. Design and Synthesis of Crosslinked Helix Dimers as Protein Tertiary Structure Mimics. *Curr Protoc* **2022**, 2 (1), e315.
218. H.C. Fry, A. Lehmann, L.E. Sinks, et al. Computational de Novo Design and Characterization of a Protein That Selectively Binds a Highly Hyperpolarizable Abiological Chromophore. *J Am Chem Soc* **2013**, 135 (37), 13914–13926.
219. R.L. Koder, J.L.R. Anderson, L.A. Solomon, et al. Design and engineering of an O2 transport protein. *Nature* **2009**, 458 (7236), 305–309.
220. J. Dou, A.A. Vorobieva, W. Sheffler, et al. De novo design of a fluorescence-activating β-barrel. *Nature* **2018**, 561 (7724), 485–491.
221. A.G. Cochran, N.J. Skelton, M.A. Starovasnik. Tryptophan zippers: Stable, monomeric β-hairpins. *Proceedings of the National Academy of Sciences* **2001**, 98 (10), 5578–5583.
222. L. PAULING, R.B. COREY. Compound Helical Configurations of Polypeptide Chains: Structure of Proteins of the α-Keratin Type. *Nature* **1953**, 171 (4341), 59–61.
223. X. Pan, T. Kortemme. Recent advances in de novo protein design: Principles, methods, and applications. *Journal of Biological Chemistry* **2021**, 296, 100558.
224. Y. Song, F. DiMaio, R.Y.-R. Wang, et al. High-Resolution Comparative Modeling with RosettaCM. *Structure* **2013**, 21 (10), 1735–1742.
225. C.A. Rohl, C.E.M. Strauss, K.M.S. Misura, D. Baker. Protein Structure Prediction Using Rosetta. In *Methods in Enzymology*; Academic Press, **2004**; Vol. 383, pp 66–93.
226. J.K. Leman, B.D. Weitzner, S.M. Lewis, et al. Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat Methods* **2020**, 17 (7), 665–680.
227. D. Baker. Protein folding, structure prediction and design. *Biochem Soc Trans* **2014**, 42 (2), 225–229.
228. J.L. Watson, D. Juergens, N.R. Bennett, et al. De novo design of protein structure and function with RFdiffusion. *Nature* **2023**, 620 (7976), 1089–1100.
229. Z. Du, H. Su, W. Wang, et al. The trRosetta server for fast and accurate protein structure prediction. *Nat Protoc* **2021**, 16 (12), 5634–5651.
230. C.W. Wood, M. Bruning, A.Á. Ibarra, et al. CCBuilder: an interactive web-based tool for building, designing and assessing coiled-coil protein assemblies. *Bioinformatics* **2014**, 30 (21), 3029–3035.
231. C. Negron, A.E. Keating. Chapter Eight - Multistate Protein Design Using CLEVER and CLASSY. In *Methods in Enzymology*; Keating, A. E., Ed.; Academic Press, **2013**; Vol. 523, pp 171–190.
232. J. AU - Smadbeck, M.B. AU - Peterson, G.A. AU - Khoury, M.S. AU - Taylor, C.A. AU - Floudas. Protein WISDOM: A Workbench for In silico De novo Design of BioMolecules. *JoVE* **2013**, No. 77, e50476.
233. L. Doyle, J. Hallinan, J. Bolduc, et al. Rational design of α-helical tandem repeat proteins with closed architectures. *Nature* **2015**, 528 (7583), 585–588.

234. P.-S. Huang, K. Feldmeier, F. Parmeggiani, et al. De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy. *Nat Chem Biol* **2016**, 12 (1), 29–34.
235. B. Kuhlman, G. Dantas, G.C. Ireton, et al. Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science (1979)* **2003**, 302 (5649), 1364–1368.
236. J.L. Beesley, D.N. Woolfson. The de novo design of α-helical peptides for supramolecular self-assembly. *Curr Opin Biotechnol* **2019**, 58, 175–182.
237. F.H.C. CRICK. Is α-Keratin a Coiled Coil? *Nature* **1952**, 170 (4334), 882–883.
238. F.H.C. Crick. The packing of α-helices: simple coiled-coils. *Acta Crystallogr* **1953**, 6 (8–9), 689–697.
239. R.S. Hodges, A.K. Saund, P.C. Chong, S. St.-Pierre, R.E. Reid. Synthetic model for two-stranded alpha-helical coiled-coils. Design, synthesis, and characterization of an 86-residue analog of tropomyosin. *J Biol Chem* **1981**, 256 3, 1214–1224.
240. J. Stetefeld, M. Jenny, T. Schulthess, et al. Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat Struct Biol* **2000**, 7 (9), 772–776.
241. D.N. Woolfson. Understanding a protein fold: The physics, chemistry, and biology of α-helical coiled coils. *Journal of Biological Chemistry* **2023**, 299 (4).
242. D.N. Woolfson. Coiled-Coil Design: Updated and Upgraded. In *Fibrous Proteins: Structures and Mechanisms*; Parry, D. A. D., Squire, J. M., Eds.; Springer International Publishing, Cham, **2017**; pp 35–61.
243. A. Lupas. Coiled coils: new structures and new functions. *Trends Biochem Sci* **1996**, 21 (10), 375–382.
244. E.K. O'Shea, K.J. Lumb, P.S. Kim. Peptide 'Velcro': Design of a heterodimeric coiled coil. *Current Biology* **1993**, 3 (10), 658–667.
245. J.M. Mason, K.M. Arndt. Coiled Coil Domains: Stability, Specificity, and Biological Implications. *ChemBioChem* **2004**, 5 (2), 170–176.
246. J. Utterström, S. Naeimipour, R. Selegård, D. Aili. Coiled coil-based therapeutics and drug delivery systems. *Adv Drug Deliv Rev* **2021**, 170, 26–43.
247. J.R. Litowski, R.S. Hodges. Designing heterodimeric two-stranded α-helical coiled-coils: the effect of chain length on protein folding, stability and specificity. *The Journal of Peptide Research* **2001**, 58 (6), 477–492.
248. D.N. Woolfson. The Design of Coiled-Coil Structures and Assemblies. In *Advances in Protein Chemistry*; Academic Press, **2005**; Vol. 70, pp 79–112.
249. H. Dong, J.D. Hartgerink. Short Homodimeric and Heterodimeric Coiled Coils. *Biomacromolecules* **2006**, 7 (3), 691–695.
250. L.K. Henchey, A.L. Jochim, P.S. Arora. Contemporary strategies for the stabilization of peptides in the α-helical conformation. *Curr Opin Chem Biol* **2008**, 12 (6), 692–697.
251. B. Dang, H. Wu, V.K. Mulligan, et al. De novo design of covalently constrained mesosize protein scaffolds with unique tertiary structures. *Proc Natl Acad Sci U S A* **2017**, 114 (41), 10852–10857.
252. P.E. Dawson, S.B.H. Kent. Convenient total synthesis of a 4-helix template-assembled synthetic protein (TASP) molecule by chemoselective ligation. *J Am Chem Soc* **1993**, 115 (16), 7263–7266.
253. M. Mutter, P. Dumy, P. Garrouste, et al. Template assembled synthetic proteins (TASP) as functional mimetics of proteins. *Angewandte Chemie* **1996**, 35, 1482–1485.

254. M.G. Wuo, A.B. Mahon, P.S. Arora. An Effective Strategy for Stabilizing Minimal Coiled Coil Mimetics. *J Am Chem Soc* **2015**, 137 (36), 11618–11621.
255. M.G. Wuo, S.H. Hong, A. Singh, P.S. Arora. Synthetic Control of Tertiary Helical Structures in Short Peptides. *J Am Chem Soc* **2018**, 140 (47), 16284–16290.
256. F. Thomas, N.C. Burgess, A.R. Thomson, D.N. Woolfson. Controlling the Assembly of Coiled–Coil Peptide Nanotubes. *Angewandte Chemie International Edition* **2016**, 55 (3), 987–991.
257. A. Buchberger, C.R. Simmons, N.E. Fahmi, R. Freeman, N. Stephanopoulos. Hierarchical Assembly of Nucleic Acid/Coiled-Coil Peptide Nanostructures. *J Am Chem Soc* **2020**, 142 (3), 1406–1416.
258. A. Buchberger, M. Al-Amin, C.R. Simmons, N. Stephanopoulos. Self-Assembly of Hybrid Peptide–DNA Nanostructures using Homotrimeric Coiled-Coil/Nucleic Acid Building Blocks. *ChemBioChem* **2023**, 24 (17), e202300223.
259. A. Ljubetič, F. Lapenta, H. Gradišar, et al. Design of coiled-coil protein-origami cages that self-assemble in vitro and in vivo. *Nat Biotechnol* **2017**, 35 (11), 1094–1101.
260. F. Lapenta, J. Aupič, M. Vezzoli, et al. Self-assembly and regulation of protein cages from pre-organised coiled-coil modules. *Nat Commun* **2021**, 12 (1), 939.
261. A.N. Lupas, J. Bassler. Coiled Coils – A Model System for the 21st Century. *Trends Biochem Sci* **2017**, 42 (2), 130–140.
262. M.D. Hartmann, C.T. Mendler, J. Bassler, et al. α/β coiled coils. *Elife* **2016**, 5, e11861.
263. M.N. Creyer, M. Retout, Z. Jin, W. Yim, J. V Jokerst. Ligation of Gold Nanoparticles with Self-Assembling, Coiled-Coil Peptides. *J Phys Chem B* **2023**, 127 (37), 8009–8018.
264. J. Liu, W. Yong, Y. Deng, N.R. Kallenbach, M. Lu. Atomic structure of a tryptophan-zipper pentamer. *Proceedings of the National Academy of Sciences* **2004**, 101 (46), 16156–16161.
265. Y. Sakurai, T. Mizuno, H. Hiroaki, J.-I. Oku, T. Tanaka. Optimization of aromatic side chain size complementarity in the hydrophobic core of a designed coiled-coil. *The Journal of Peptide Research* **2005**, 66 (6), 387–394.
266. J. Liu, Q. Zheng, Y. Deng, N.R. Kallenbach, M. Lu. Conformational Transition between Four and Five-stranded Phenylalanine Zippers Determined by a Local Packing Interaction. *J Mol Biol* **2006**, 361 (1), 168–179.
267. G.G. Rhys, W.M. Dawson, J.L. Beesley, et al. How Coiled-Coil Assemblies Accommodate Multiple Aromatic Residues. *Biomacromolecules* **2021**, 22 (5), 2010–2019.
268. C. Lou, N.J. Christensen, M.C. Martos-Maldonado, et al. Folding Topology of a Short Coiled-Coil Peptide Structure Templated by an Oligonucleotide Triplex. *Chemistry – A European Journal* **2017**, 23 (39), 9297–9305.
269. C. Lou, J.T. Boesen, N.J. Christensen, et al. Self-Assembly of DNA–Peptide Supermolecules: Coiled-Coil Peptide Structures Templated by d-DNA and l-DNA Triplexes Exhibit Chirality-Independent but Orientation-Dependent Stabilizing Cooperativity. *Chemistry – A European Journal* **2020**, 26 (25), 5676–5684.
270. J. Jin, E.G. Baker, C.W. Wood, et al. Peptide Assembly Directed and Quantified Using Megadalton DNA Nanostructures. *ACS Nano* **2019**, 13 (9), 9927–9935.

271. J. Walshaw, D.N. Woolfson. SOCKET: a program for identifying and analysing coiled-coil motifs within protein structures11Edited by J. Thornton. *J Mol Biol* **2001**, 307 (5), 1427–1450.
272. A. Lupas, M. Van Dyke, J. Stock. Predicting Coiled Coils from Protein Sequences. *Science (1979)* **1991**, 252 (5009), 1162–1164.
273. B. Berger, D.B. Wilson, E. Wolf, et al. Predicting coiled coils by use of pairwise residue correlations. *Proceedings of the National Academy of Sciences* **1995**, 92 (18), 8259–8263.
274. E. Wolf, P.S. Kim, B. Berger. MultiCoil: A program for predicting two-and three-stranded coiled coils. *Protein Science* **1997**, 6 (6), 1179–1189.
275. M.C. Manning, R.W. Woody. Theoretical CD studies of polypeptide helices: Examination of important electronic and geometric factors. *Biopolymers* **1991**, 31 (5), 569–586.
276. C. Toniolo, A. Polese, F. Formaggio, M. Crisma, J. Kamphuis. Circular Dichroism Spectrum of a Peptide 310-Helix. *J Am Chem Soc* **1996**, 118 (11), 2744–2745.
277. A. Micsonai, F. Wien, L. Kernya, et al. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences* **2015**, 112 (24), E3095–E3103.
278. F. Thomas, A.L. Boyle, A.J. Burton, D.N. Woolfson. A Set of de Novo Designed Parallel Heterodimeric Coiled Coils with Quantified Dissociation Constants in the Micromolar to Sub-nanomolar Regime. *J Am Chem Soc* **2013**, 135 (13), 5161–5166.
279. M.D. Watson, I. Peran, D.P. Raleigh. A Non-perturbing Probe of Coiled Coil Formation Based on Electron Transfer Mediated Fluorescence Quenching. *Biochemistry* **2016**, 55 (26), 3685–3691.
280. J.R. Winkler, H.B. Gray. Electron Flow through Metalloproteins. *Chem Rev* **2014**, 114 (7), 3369–3380.
281. T. Shimizu, D. Huang, F. Yan, et al. Gaseous O2, NO, and CO in Signal Transduction: Structure and Function Relationships of Heme-Based Gas Sensors and Heme-Redox Sensors. *Chem Rev* **2015**, 115 (13), 6491–6533.
282. E. Theil, K. Raymond. Transition-Metal Storage, Transport, and Biomineralization. **2023**.
283. J. Liu, S. Chakraborty, P. Hosseinzadeh, et al. Metalloproteins Containing Cytochrome, Iron–Sulfur, or Copper Redox Centers. *Chem Rev* **2014**, 114 (8), 4366–4469.
284. M.J. Chalkley, S.I. Mann, W.F. DeGrado. De novo metalloprotein design. *Nat Rev Chem* **2022**, 6 (1), 31–50.
285. F. Yu, V.M. Cangelosi, M.L. Zastrow, et al. Protein Design: Toward Functional Metalloenzymes. *Chem Rev* **2014**, 114 (7), 3495–3578.
286. A.J. Doerr, G.L. McLendon. Design, Folding, and Activities of Metal-Assembled Coiled Coil Proteins. *Inorg Chem* **2004**, 43 (25), 7916–7925.
287. K.J. Koebke, T.B.J. Pinter, W.C. Pitts, V.L. Pecoraro. Catalysis and Electron Transfer in De Novo Designed Metalloproteins. *Chem Rev* **2022**, 122 (14), 12046–12109.
288. O. Iranzo, D. Ghosh, V.L. Pecoraro. Assessing the integrity of designed homomeric parallel three-stranded coiled coils in the presence of metal ions. *Inorg Chem* **2006**, 45 (25), 9959–9973.

289. M. Tegoni, F. Yu, M. Bersellini, J.E. Penner-Hahn, V.L. Pecoraro. Designing a functional type 2 copper center that has nitrite reductase activity within α-helical coiled coils. *Proceedings of the National Academy of Sciences* **2012**, 109 (52), 21234–21239.
290. A.G. Tebo, V.L. Pecoraro. Artificial metalloenzymes derived from three-helix bundles. *Curr Opin Chem Biol* **2015**, 25, 65–70.
291. J.S. Plegaria, M. Duca, C. Tard, et al. De Novo Design and Characterization of Copper Metallopeptides Inspired by Native Cupredoxins. *Inorg Chem* **2015**, 54 (19), 9470–9482.
292. J.S. Plegaria, C. Herrero, A. Quaranta, V.L. Pecoraro. Electron transfer activity of a de novo designed copper center in a three-helix bundle fold. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2016**, 1857 (5), 522–530.

# *Chapter 2*

## *2.0 Assembly of Heterotrimeric Coiled-Coil structures through DNA:PNA-Peptide interaction*

### 2.1 Abstract

In this chapter is reported the creation of a platform for the assembly of heterotrimeric coiled-coil structures, formed through DNA:PNA-peptides interactions, that can mimic the coordination site of plastocyanin (a small copper-binding protein involved in electron transfer processes, where the copper ion is coordinated through the side chains of the two histidines, a cysteine, and a methionine), attaining a supramolecular mini-metalloprotein.[1] This can be achieved through the interaction between a tripodal DNA[2] and three different PNA-peptide conjugates, where the trimeric DNA is used as a template containing three different arms complementary to each specific PNA sequence. The peptide moiety connected orthogonally to the PNA sequence are designed to obtain a heterotrimeric coiled-coil with short peptide strands. The DNA strands, thanks to their very stable DNA:PNA complex formation, ensure the spatial proximity of three different peptides domains, allowing the formation of the heterotrimeric coiled-coil system (*Figure 2.1*).

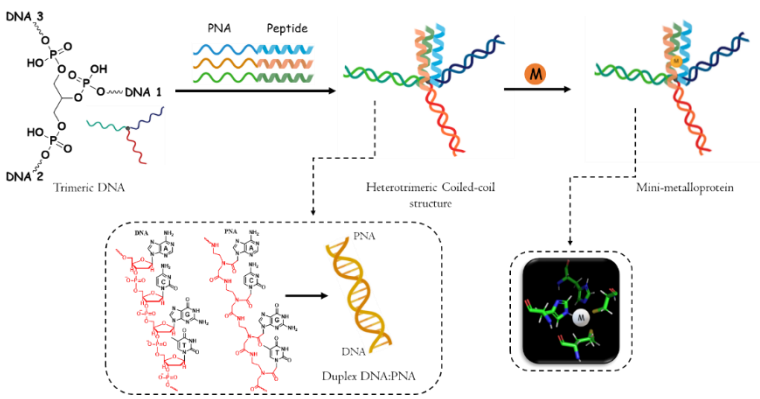



***Figure 2.1.*** *Schematic representation of the template directed assembly of heterotrimeric coiled-coil structures.*

This project was carried out in collaboration with Prof. Matteo Tegoni from University of Parma, and Prof. Hanadi Sleiman and Fangzhou Zhao from McGill University (MGU) with a secondment period of four months at MGU.

## 2.2 Introduction

Coiled-coil systems are biological motifs that play a crucial role in the structural and functional features of many proteins, in fact, they were described for the first time by Pauling and Corey[3] and by Crick[4] as principal components of fibrous proteins, such as keratin, myosin, and fibrinogen. Coiled-coils systems consist of two or more α-helical peptides, that are arranged into superhelices. A heptameric repeating units of amino acids dictates the organization of the peptides into several types of coiled-coil structures.[5] The heptad repeats, usually denoted as $(abcdefg)_n$, where hydrophobic residues are typically located in the 'a' and 'd' positions to create a hydrophobic core, while charged residues occupy the 'e' and 'g' positions, forming electrostatic interactions.[6,7] The peptide sequences may be identical or different (homomeric or heteromeric, respectively), arranged in a parallel or antiparallel orientation, thus giving rise to different coiled-coil combinations.[8] Advances of coiled-coil architectures described in recent years are not limited solely to dimeric, trimeric, or tetrameric structures (comprising two, three, or four helices, respectively), but are extended to the intricate design of structures with a higher number of helices, ranging from pentamers to dodecamers and beyond.[9]

To spontaneously form coiled-coils, three or more heptad repeats, equivalent to more than 21 amino acids are required.[5] In fact, long peptide sequences are usually designed with a specific portion of residues available to self-assemble to create pre-designed complex scaffold. Furthermore, coiled-coil motifs can be form from peptide fragments designed from scratch, in a process called *de novo protein design*,[10] and they can achieve additional stability through the chelation of metal ions or by covalently connecting them with organic molecules.[11–13]

Conversely, shorter peptide sequences struggle to spontaneously form coiled-coils since they have too few interactions to counterbalance the enthalpic/entropic effects caused by the association and distortion of the α-helices, and in fact coiled-coil structures formed by α-helices with 14 amino acids residues tend to form oligomers.[14,15] To offset the lack of non-covalent interactions in the assembly of short

α-helices within coiled-coil architectures, covalent bonds can be introduced between the helices.[16–18]

The peptide strands can be also linked together through synergistic supramolecular templating effects. This is exemplified using Template-Assembled Synthetic Proteins (TASPs) to support the creation of intricate protein structures.[19–22]

A template for assembling peptides with precise spatial arrangement can be obtained using DNA based scaffolds. The chemical and physical properties of DNA were found to be of great interest in nanotechnology, for generating several structural motifs like tiles,[23–25] multi-way junctions,[26–28] DNA origami,[29,30] cage structures and hybrid DNA-polymer.[31–33] These motifs can be connected generating DNA nanostructures through Watson-Crick base-pairing interactions in a predictable and programmable fashion to create symmetric and asymmetric nanoscale architectures with specific shapes and dimensions.[34–36]

By exploiting these properties, DNA-peptide conjugates were assembled in supramolecular polymer (1D nanofibers) thanks to the ability of peptide domains to assembly in a heterodimeric coiled-coil structure.[37] Furthermore, DNA-peptides conjugates possess the capability to self-assemble into a complex homotrimer coiled-coil structure, resembling a three-way junction. This structural feature is valuable for constructing nanostructures such as DNA cages or connecting DNA tiles.[38]

The assembling of peptides with precise spatial positioning, using a DNA scaffold as a template, was first illustrated by Wengel's group that showed the self-organization of homomeric coiled-coil peptides (three heptads) via the creation of oligonucleotide triplexes. This trimeric assembly exhibits a pronounced α-helical conformation, with the stability of the resulting structures being dependent on the length of the peptides involved.[39,40]

The assembly of *de novo* designed peptides can be achieved using DNA origami structures that feature covalent connections. In fact, the development of peptide-oligonucleotide conjugates can streamline the controlled construction of peptide self-assembly.

Coiled-coil motifs are of interest also in the *de novo* design and creation of metalloprotein, a method for generating entirely new metalloproteins or mini-metalloproteins with precise tertiary or quaternary structures and desired functions, starting from scratch and harnessing computational techniques.[11]

In fact, for example, Pecoraro et Al. developed a series of artificial metalloproteins based on TRI family peptides, that can form three-stranded coiled-coil structures.[41] Some derivatives of this family were found to bind metal ions, like Hg(II), Cd(II), Cu(II)/Cu(I), and Zn(I) with specific catalytic, mimicking the properties of nitrite reductase, carbonic anhydrase, or electron transfer activities.[42–44]

Aiming at obtaining electron transfer (ET) activity, copper metalloproteins were engineered by initially utilizing TRI-family peptides and subsequently altering their sequences to introduce a specific binding site that mimics plastocyanin. This involved the incorporation of two histidine residues, one cysteine, and one methionine.[45,46]

By combining the DNA-templated assembly with coiled coil design and model metalloprotein de novo design, together with the PNA properties and their easy conjugation with peptides, the design and synthesis of heterotrimeric coiled coil with minimal peptide components could be realized. The work to design and synthesize these systems is described in the present chapter.

## 2.3 Result and discussion

During a secondment period at Mc Gill University (MGU), the synthesis of trimeric systems for generating coiled-coil structures by joining peptide and DNA segments was carried out. The results obtained using a tripodal DNA template and DNA segments linked to different peptides is described in a paper submitted for publication and presently under revision. This work, which will not be described in detail in this thesis being part of a parallel thesis at Mc Gill University (Fangzhou Zhao, PhD Thesis), showed the successful assembly of three different peptides which could be forced to stay in close proximity and then chemically 'stapled', demonstrating that the template effect of the tripodal DNA could be used to induce peptide-peptide interactions with rational design, and to create chemical libraries composed of different combinations of peptides.

In the present work, we instead used PNA-peptide conjugates which could be assembled on a similar tripodal DNA template to generate minimal coiled-coil structures with metal binding abilities (*Figure 2.2).*

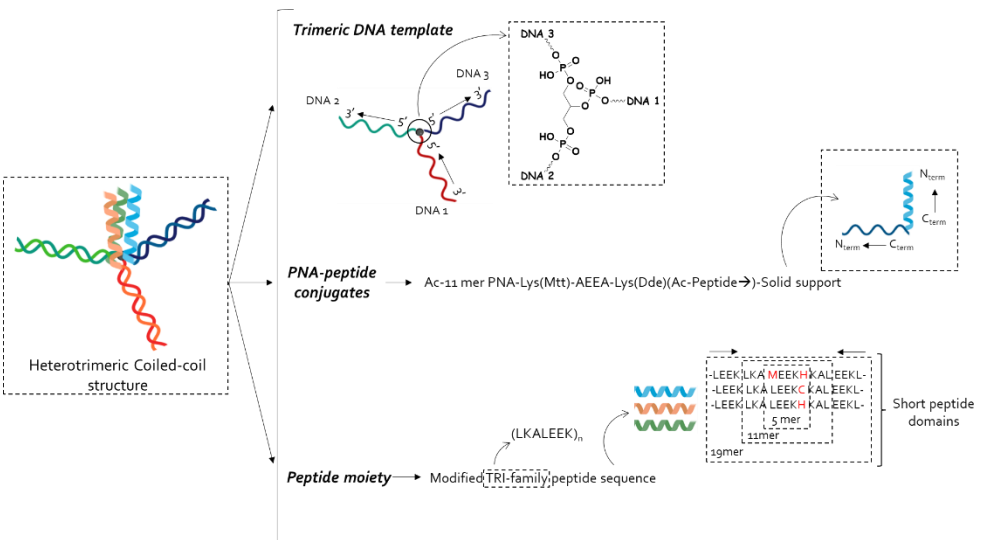



***Figure 2.2.*** *Design of heterotrimeric Coiled-Coil structures through DNA:PNA-Peptide interaction.*

The heterotrimeric coiled-coil systems were generated by producing the DNA tripodal template and a series of three different PNA-peptides conjugates designed

to obtain three different DNA:PNA duplexes able to drive the formation of the desired scaffold.

The peptide domains were designed to generate coiled-coil structures, starting from TRI peptide sequences, and substituting some of the amino acids with those required for the coordination of copper ions (more specifically, two histidine residue, one cysteine and one methionine), and to create a coordination model mimicking plastocyanin. These specific amino acid residues can modulate the ox-redox properties of the total construct. Taking into consideration that, unlike the long peptide sequences, that are usually designed with a specific portion of residue available to self-assemble in solution to create a coiled-coil systems, the short peptide sequences cannot form these scaffolds in solution. This can be attributed to the fact that the driving forces of coiled-coil formation are weak interactions, and consequently, the number of these interactions is a decisive factor.

### 2.3.1 Assembly and conformations of peptides components

First, to study if, in absence of a rigid template, the peptide domains are suitable to self-assembly spontaneously in solution to generate coiled-coil structures, several peptide domains extrapolated from TRI-family peptides were designed, synthetized, and characterized by circular dichroism.

Four different lengths (22mer, 19mer, 11mer and 5mer), obtained cutting the original tetrameric TRI-family peptide ($g(lkaleek)_4g$), were synthetized to study the formation of coiled-coil structures without a template. Each peptide domain was synthetized in original orientation respect the TRI peptides (A), but also by reversing the order of the amino acid sequence (B). The peptide sequences are reported in *Table 2.1* and a schematic representation of the coiled-coil structure is reported in *Figure 2.3*.

***Table 2.1.*** *TRI-family peptide sequences synthetized.*

| Entry | Name | Sequence | MW (g/mol) | Yield (%) |
|---|---|---|---|---|
| 1 | 22merA | ABA-lkaleeklkaleeklkaleekl-$NH_2$ | 2727,23 | 15 |
| 2 | 22merB | ABA-lkeelaklkeelaklkeelakl-$NH_2$ | 2727,23 | 20 |
| 3 | 19merA | ABA-leeklkaleeklkaleekl-$NH_2$ | 2414,82 | 18 |
| 4 | 19merB | ABA-lkeelaklkeelaklkeel-$NH_2$ | 2414,82 | 6 |
| 5 | 11merA | ABA-lkaleeklkal-$NH_2$ | 1415,72 | 25 |
| 6 | 11merB | ABA-laklkeelakl-$NH_2$ | 1415,72 | 41 |
| 7 | 5merA | ABA-leekl-$NH_2$ | 790,90 | 25 |
| 8 | 5merB | ABA-lkeel-$NH_2$ | 790,90 | 10 |

These peptides sequences were synthesized by solid phase peptide synthesis using Fmoc chemistry and were functionalized with 4-acetamidobenzoic acid (ABA) at the N-terminal moiety for the quantification of these sequences by UV-Visible spectrophotometer.

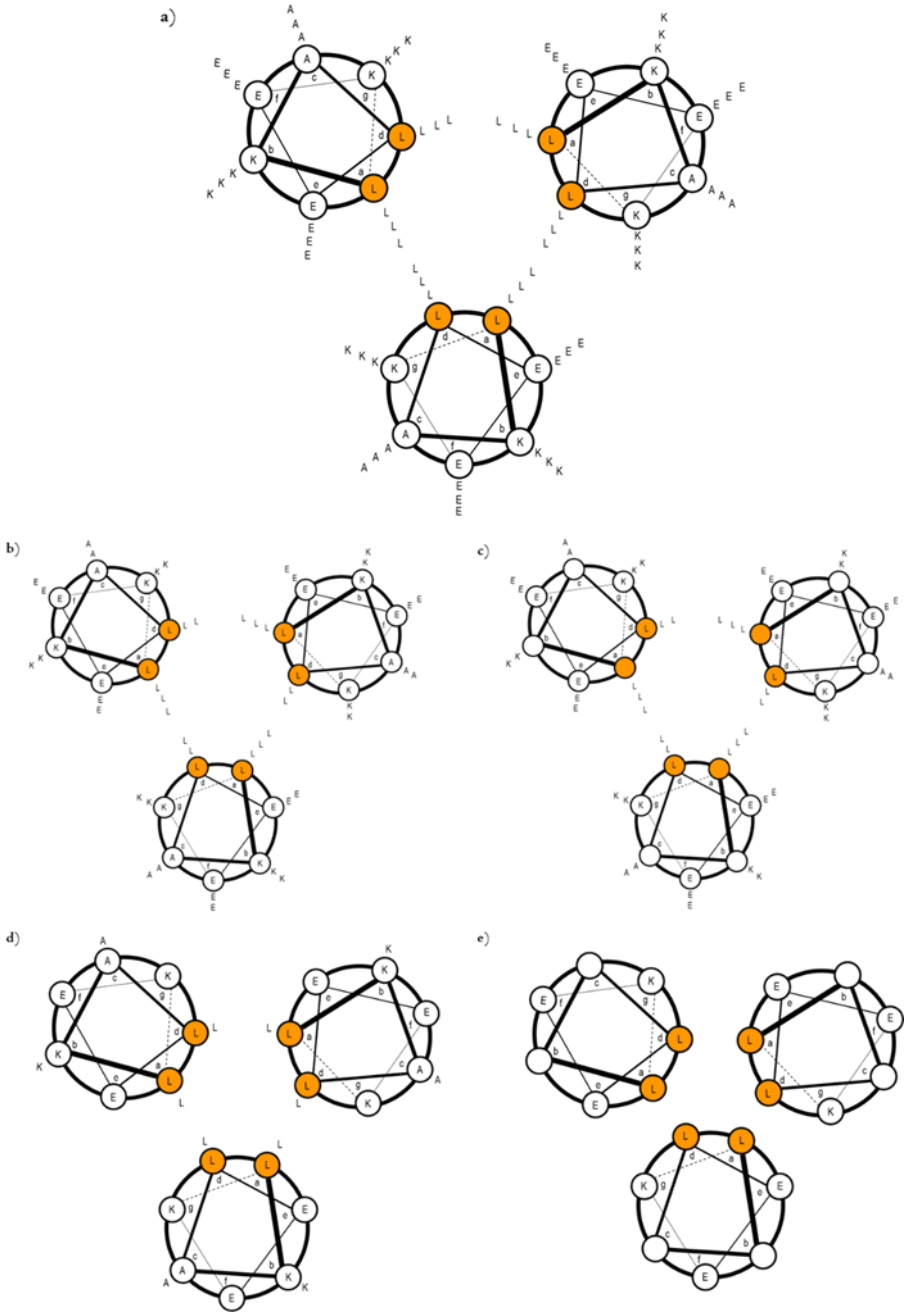


***Figure 2.3.*** *Schematic representation of mentioned peptide sequences to coiled-coil structure: a) original TRI-family peptide, b) 22mer, c) 19mer, d) 11mer, and e) 5mer.*

The experiments to assess the secondary structure of these peptides and to verify their ability to spontaneously self-assemble in solution were carried out at three

different concentrations (45 µM, 15 µM and 6 µM) in TRIS Base 10 mM (pH:7.2) using circular dichroism (CD).

A typical CD signature of α-helix is characterized by two negative bands, one at 208 nm and another at 222 nm, and when the ratio of band intensities at 222 nm and 208 nm is less than 1.0, it is regarded as a distinctive feature of these structures. Respect the signature of α-helix, CD signature of coiled-coil structure consists of the increase in negative direction of the band around 220 nm, and in particular the ratio between the band intensity at 222 and 208 nm higher than 1.0 is considered to be a characteristic of these structures. However, CD signature of the formation of the coiled coil systems was established in the literature.[47–49]

Considering the 22mer peptides (*22merA and 22merB*), the CD signature showed a coiled-coil structure at all the three concentrations examined in both orientations, as can be seen from *Figure 2.4a and 2.4b* relative to the first and second orientation, respectively.

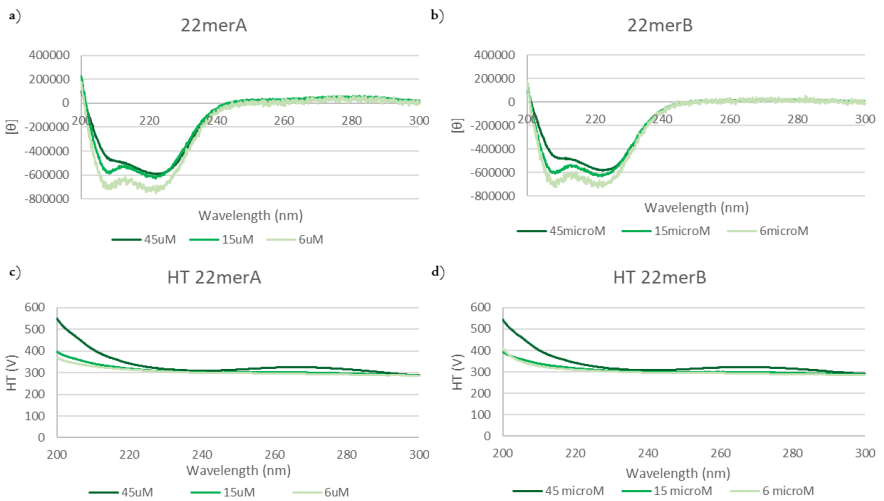



***Figure 2.4.*** *a) CD signature of 22merA to the three different concentrations, b) CD signature of 22merB to the three different concentrations, c) HT signature of 22merA to the three different concentrations, d) HT signature of 22merB to the three different concentrations.*

Absorption (measured by HT channel of the CD instrument) was excessively high for the 45 µM sample and suitable for the 15 µM and 6 µM samples for both the orientations, it is possible to calculate the ratio 222/208 (characteristic of coiled-coil

systems formation) only for the last two concentration mentioned. In fact, the ratio 222/208 is 1.10 and 1.05, respectively for 15 µM and 6 µM of *22merA*, and 1.04 and 1.02 for *22merB*, confirming the coiled-coil formation. The HT signals of all the three concentrations relative to peptide sequences *22merA and 22merB* are reported in *Figure 2.4c and 2.4d*, respectively.

However, this excessive absorption for the 45 µM sample was observed for all the sequence reported in *Table 2.1* except for 5mer peptides.

A different result was obtained for the two 19mer peptide sequences, because only in one orientation (*19merA*), specifically that corresponding to the original TRI-family direction, the CD signature corresponds to coiled-coil formation moderately at all the three concentrations examined, as showed in *Figure 2.5a*. In addition, the ratio 222/208 is 0.98 and 0.98 for 15 µM and 6 µM of *19merA* respectively, which are lower value than those reported for 22mer peptide sequences, but are close to 1, suggesting a partial formation of the coiled-coil.

Regarding the other 19mer peptide sequence (*19merB, Figure 2.5b*), the CD spectrum was found to be compatible with that of an α-helix, where the band at 208 nm is more negative with respect to that at 222 nm. The difference between these two peptide sequences lead to consider that, beyond the length, also the orientation of the sequence and the different specific interactions which can be established are important.

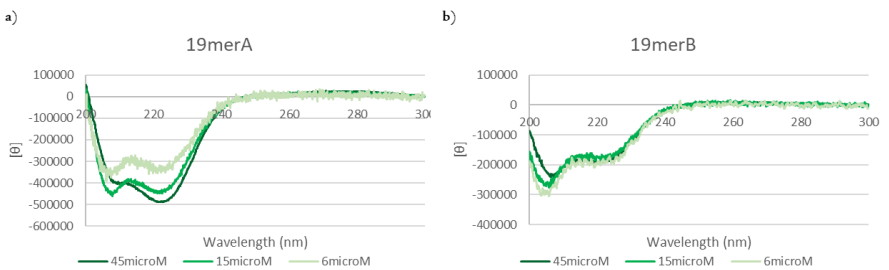



***Figure 2.5.*** *a) CD signature of 19merA to the three different concentrations, b) CD signature of 19merB to the three different concentrations.*

Similar results as for *19merB* were observed for *11mer* and *5mer* peptide sequences at all the three concentrations taken into consideration and in each the orientations.

These peptide sequences, therefore, are not able to form alone a super-secondary structure in solution because they are too short. The results are reported in the Appendix in *Figures 2.A1.*

In order to confirm the formation of coiled-coil structures, melting experiments were carried out for the peptides *22merA*, *22merB* and *19merB*, at 6 μM as concentration in TRIS Base 10mM (pH:7.2). The first two peptides, as reported above, can form coiled-coil structure whereas the last one cannot. By recording a spectrum every 10°C, from 20°C to 90°C (melting) and from 90°C to 20°C (annealing), a decrease of both the bands (222 nm and 208 nm) for the *22merA* and *22merB* peptides was observed, but the band at 222 nm decreases more rapidly, confirming the formation of coiled-coil structures. The variation of CD measurements for peptide sequence *22merA* during melting is showed in *Figure 2.6a*, while the annealing result in *Figure 2.6b.* Corresponding spectra for *22merB* sample are reported in the Appendix in *Figure 2.A2.*

The melting variation observed means that, as the temperature increases, there is a breakdown of coiled-coil system formed spontaneously in solution, since the CD signature of coiled-coil is replaced by that of an α-helix. The ratio 222/208, in addition, confirm this progressive desegregation of the coiled-coil structure in both these peptides during the self-melting, as shown as an example in *Figure 2.6c* (relative to *22merA*).

By contrast, the self-melting and annealing experiments relative to *19merB* only show the CD signature characteristic of α-helix from 20°C to 90°C and reverse (*Figure 2.7a and 2.7b*). The 222/208 ratio for the self-melting experiment varies from 0.94 to 0.85 from 20°C to 60°C and then remain constant at 0.85 right up to 90°C (*Figure 2.7c*).

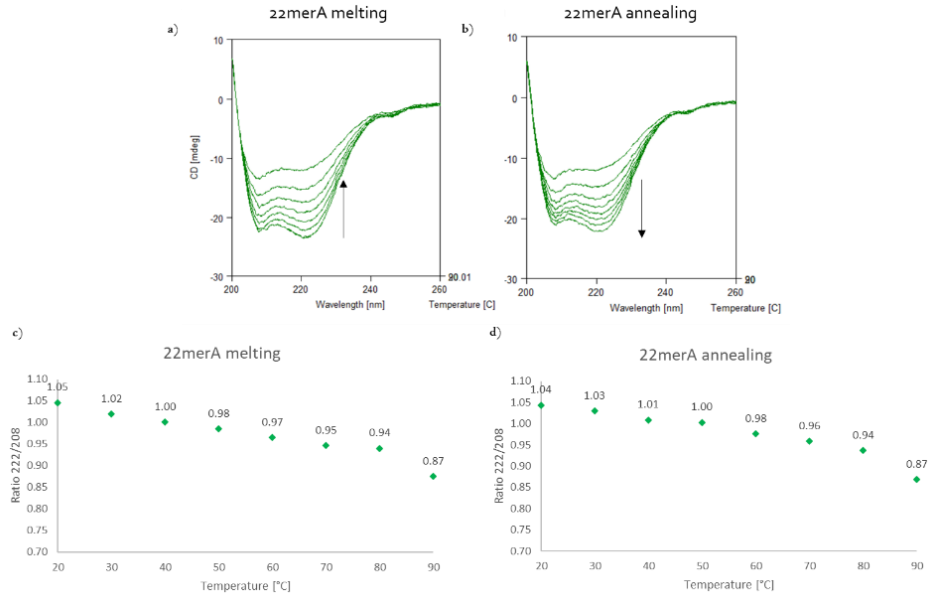



***Figure 2.6.*** *a) Self-melting of 22merA, recording a spectrum every 10°C, from 20°C to 90°C. b) annealing of 22merA, recording a spectrum every 10°C, from 90°C to 10°C, c) decrease of ratio 222/208 during the self-melting to 22merA, d) increase of ratio 222/208 during the annealing to 22merA.*

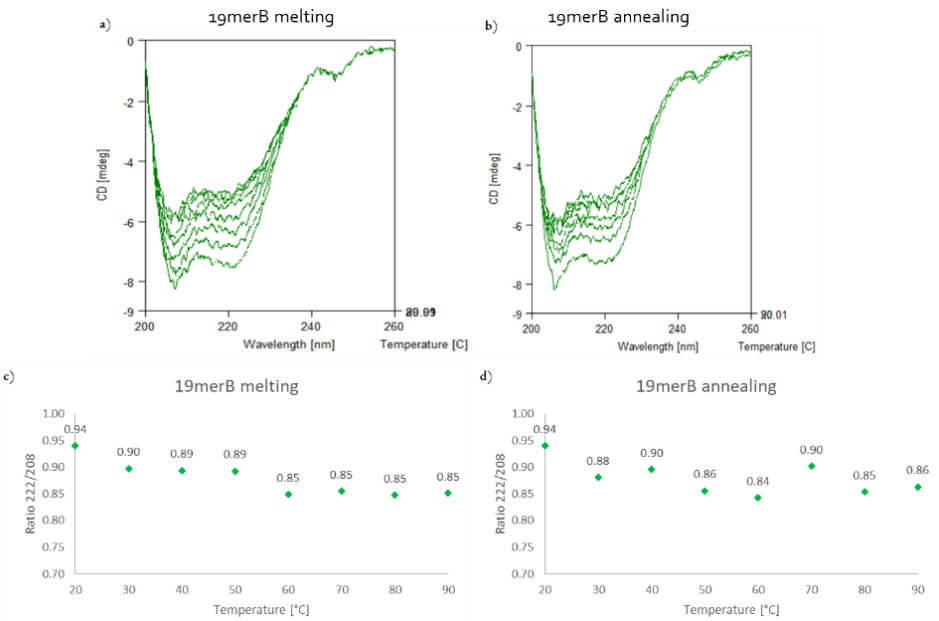



***Figure 2.7.*** *a) Self-melting of 19merB, recording a spectrum every 10°C, from 20°C to 90°C. b) annealing of 19merB, recording a spectrum every 10°C, from 90°C to 10°C, c) decrease of ratio 222/208 during the self-melting to 19merB, , d) increase of ratio 222/208 during the annealing to 19merB.*

*2.3.1.1 Assembly and conformations of peptide components in the presence of copper(II)*

The peptide sequences, reported in *Table 2.1*, were used also to investigate if, even in the absence of specific amino acids that can coordinate a copper(II), they could interact with the copper ion and undergo to a change in the secondary structure and/or assembly leading to observe changes in CD spectra. These experiments were conducted at 6 µM as concentration, in HEPES buffer 1mM (pH:7.4) from 1 eq to 50 eq of copper (II) sulfate.

Regarding the data of the two 19mer peptide sequences (*Figure 2.8a and 2.8b, respectively*): the *19merA*, that formerly showed a CD signature corresponds to coiled-coil formation with the ratio 222/208 for sample at 6 µM near to one, in present of 50 eq of copper(II) the formation of coiled-coil is detected, since the 222/208 ratio is 1.57.For the *19merB* that showed a CD signature of α-helix, instead, the presence of 50 eq of copper(II) seems to induce the formation of a coiled-coil structure.

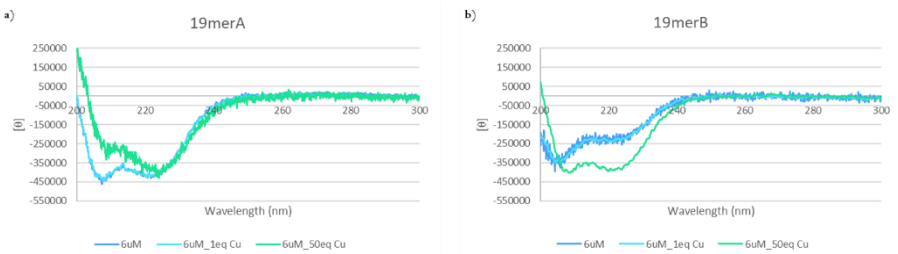



***Figure 2.8.*** *a) CD results of 19merA in present of 1 eq and 50 eq of Copper(II) sulfate, b) CD results of 19merB in present of 1 eq and 50 eq of Copper(II) sulfate.*

To investigate this phenomenon a Cu titration (1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate) of a 6 µM sample was performed (*Figure 2.9).* Until 25 eq of Cu(II) the CD signature correspond to α-helix while with 50 eq of Cu(II) a change of CD signature is observed (more like a coiled-coil signature). The 222/208 ratio also is in line with this variation, being 1.09. To examine this interaction should be explore the titration between 25 eq and 50 eq to understand when the secondary structure of the peptide sequence in solution moved from α-helix structure to coiled-coil

complex. However, these quantities of Cu(II) are very high respect to the desired copper concentration and out of the scope of this work.

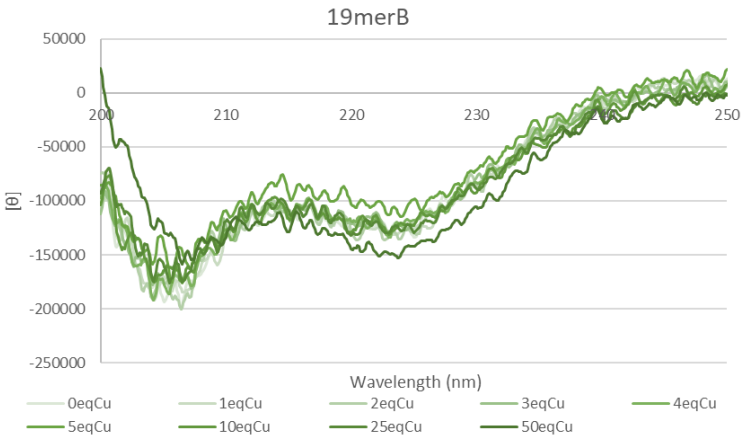



***Figure 2.9.*** *Titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate of Entry 30.*

However, by analyzing, instead, the CD spectra concerning the 22mer, 11mer and 5mer peptide sequences in both the orientations, no variation was observed if compared to experiments without the copper ion. More in detail, the presence of Cu(II) does not interfere with the formation of the coiled-coil in the long peptide (22mer) sequences, and it does not promote the formation of coiled-coil structures with short sequences (11mer and 5mer). The results were reported in Appendix (*Figure 2.A3*).

### 2.3.2 Assembly and conformations of modified peptide components mixtures

The previous 19mer, 11mer and 5mer peptide sequences from TRI-family peptide was also modified introducing a coordination side for the copper ion, in particular for those present in the plastocyanin model described by Pecoraro et al.[45,46] In fact, four different modifications of amino acid residues can be noted in the sequences reported in *Table 2.2*, where they are highlighted in red. Cysteine, histidine and methionine residues have been introduced into each series formed by three peptide sequences of the same length and orientation. In specific, two residues of leucine were replaced with one methionine and one histidine residues in one peptide

sequence. Only one leucine was changed with a cysteine in another sequences, and, in the last modified peptide sequence, a histidine residue has substituted a leucine. A schematic representation of these modified TRI-family peptide sequences in a coiled-coil structure was reported in *Figure 2.10*.

***Table 2.2.*** *Modified TRI-family peptide sequences.*

| Entry | Name | Sequence | MW (g/mol) | Yield (%) |
|---|---|---|---|---|
| 9 | m-5merA1 | ABA-meekh-$NH_2$ | 832.92 | 21 |
| 10 | m-5merA2 | ABA-leekc-$NH_2$ | 780.88 | 15 |
| 11 | m-5merA3 | ABA-leekh-$NH_2$ | 814.88 | 23 |
| 12 | m-11merA1 | ABA-lkameekhkal-$NH_2$ | 1457.74 | 13 |
| 13 | m-11merA2 | ABA-lkaleekckal-$NH_2$ | 1405.70 | 19 |
| 14 | m-11merA3 | ABA-lkaleekhkal-$NH_2$ | 1439.70 | 40 |
| 15 | m-19merA1 | ABA-leeklkameekhkaleekl-$NH_2$ | 2456.84 | 18 |
| 16 | m-19merA2 | ABA-leeklkaleekckaleekl-$NH_2$ | 2404.80 | 20 |
| 17 | m-19merA3 | ABA-leeklkaleekhkaleekl-$NH_2$ | 2438.80 | 42 |
| 18 | m-5merB1 | ABA-hkeem-$NH_2$ | 832.92 | 26 |
| 19 | m-5merB2 | ABA-ckeel-$NH_2$ | 780.88 | 64 |
| 20 | m-5merB3 | ABA-hkeel-$NH_2$ | 814.88 | 15 |
| 21 | m-11merB1 | ABA-lakhkeemakl-$NH_2$ | 1457.74 | 23 |
| 22 | m-11merB2 | ABA-lakckeelakl-$NH_2$ | 1405.70 | 33 |
| 23 | m-11merB3 | ABA-lakhkeelakl-$NH_2$ | 1439.70 | 6 |
| 24 | m-19merB1 | ABA-lkeelakhkeemaklkeel-$NH_2$ | 2456.84 | 10 |
| 25 | m-19merB2 | ABA-lkeelakckeelaklkeel-$NH_2$ | 2404.80 | 19 |
| 26 | m-19merB3 | ABA-lkeelakheelaklkeel-$NH_2$ | 2438.80 | 12 |

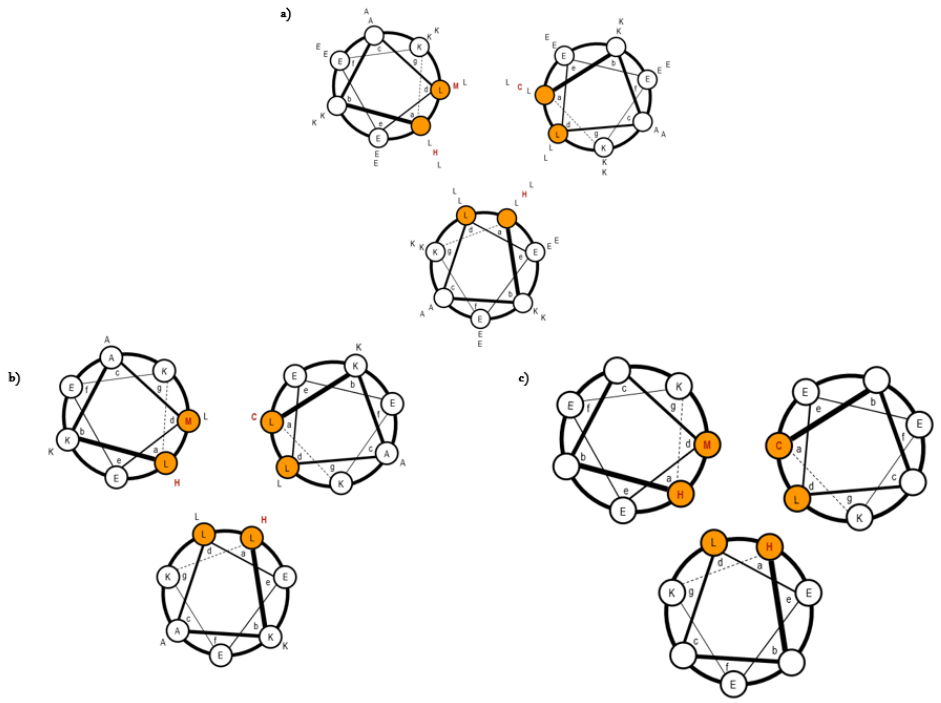


***Figure 2.10.*** *Schematic representation of modifies TRI-family peptide sequences in coiled-coil structure.*

Also in this case, CD measurements at three different final concentrations (45 µM, 15 µM and 6 µM) of pure modified peptide sequences or a mixture (considering a 15 µM, 5 µM and 2 µM respectively as a contribution of each single peptide sequences) in TRIS Base 10mM (pH:7.2) were carried out to study the formation of both homo- and hetero-aggregates.

The 19mer peptide sequences in the same orientation as TRI family peptide (*m-19merA1, m-19merA2 and m-19merA3*), the measurements at CD showed an α-helix signature both in samples containing individual peptide sequences and in combinations with each other at all the three concentrations tested, as shown in *Figure 2.11* (single samples) and *Figure 2.12a* (combination).

From data in *Figure 2.11* it can be noted that samples at 45 µM of all the three different peptides present a signature with a higher 222/208 ratio, but, due to excessive absorbance of the sample at 208, distortion of the CD signal might also occur. However, θ at 222 nm remains the same compared to 15 µM and 6 µM,

confirming the presence of only α-helix structure in solution at all the concentrations.

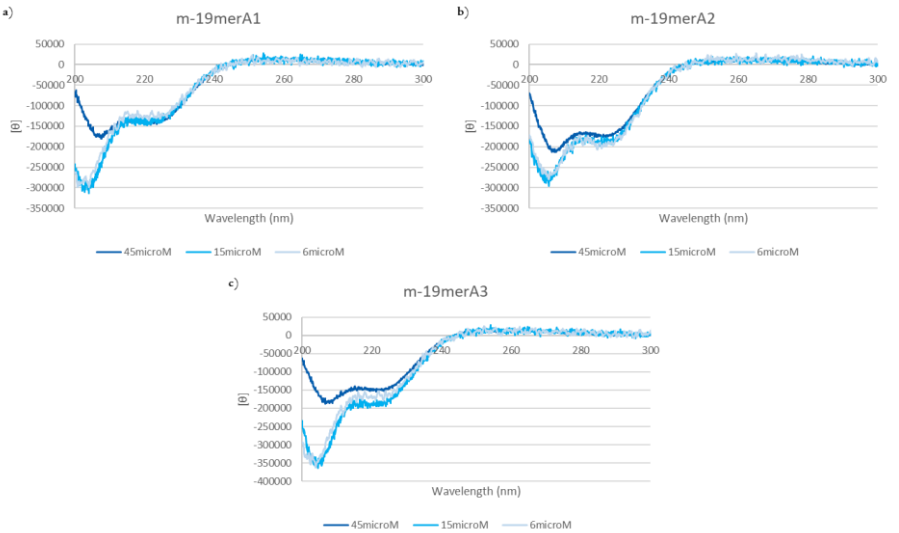



***Figure 2.11.*** *CD measurements at three different concentrations, in specific 45 μM, 15 μM and 6 μM, of a) m-19merA1 b) m-19merA2 and c) m-19merA3 as single separate sequences in solution.*

From *Figure 2.12b-d* it is possible to compare the CD signals of the mixture of all three peptides with the sum of those of the single peptides in solution. Since these are substantially overlapping, there is no evidence of aggregation from these data. In addition, ratios of 222/208 nm CD intensity higher than 1 were not found.

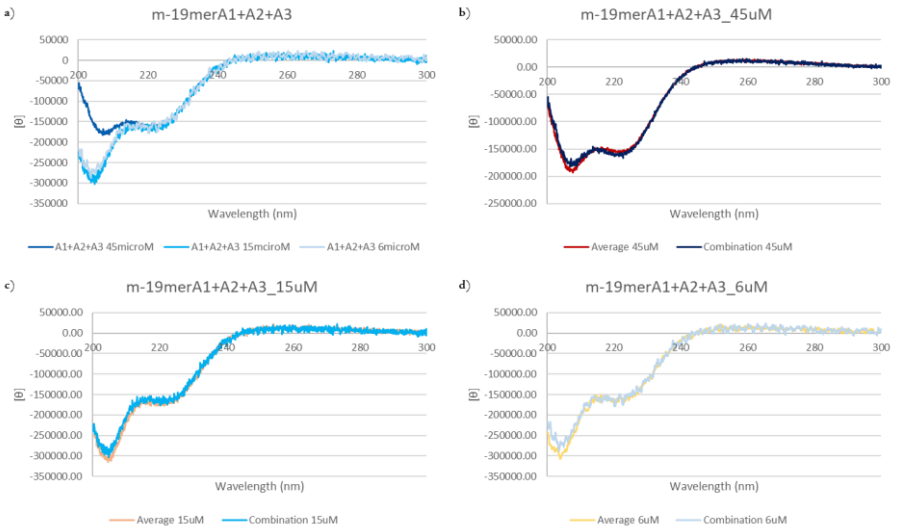



***Figure 2.12.*** *a) CD measurements of combination of m-19merA1, m-19mer A2 and m-19mer A3 at three different final concentrations (45 μM, 15 μM and 6 μM); Comparison between combination spectrum and sum of each individual sequence at b) 45 μM, c) 15 μM and d) 6 μM.*

Regarding the other 19mer modified peptide sequence (*m-19merB1, m-19merB2 and m-19merB3*), where the sequence is reversed compared to TRI family peptides, the CD measurements gave the same results of previous sequences, except for one peptide (*m-19merB2*). This diverges only for the sample at 45 μM concentration (*Figure 2.13a*) and given that θ value at 222 nm becomes more negative respect the other two concentrations, a large concentration of peptides in solution can lead to the formation of hetero-aggregates. The spectra concerning the other two 19mer peptides are reported in the Appendix in *Figure 2.A4*.

Similar results were obtained by matching the three different concentrations of combination measurements (*m-19merB1, m-19merB2 and m-19merB3 mix*), given that a slight difference relative to the 222 nm band in the sample at 45 μM respect the other two concentrations can be noticed (*Figure 2.13b*).

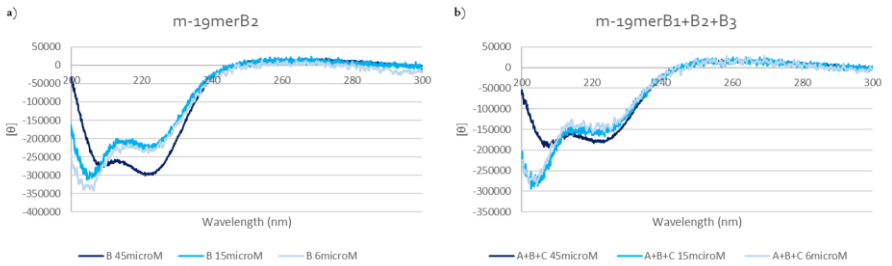



***Figure 2.13. a)*** *CD measurements at three different concentrations (45 μM, 15 μM and 6μM) of m-19merB2. b) CD measurements of combination of m-19merB1, m-19merB2 and m-19merB3 at three different final concentrations (45 μM, 15 μM and 6 μM).*

As concerns the 11mer and 5mer peptides, the CD results showed an α-helix structure both when they were analyzed individually and in combination. This confirms that the sequences are too short to form spontaneously either a homo or hetero coiled-coil structure. CD data are given in the Appendix from *Figure 2.A5 to Figure 2.A8.*

### *2.3.2.1 Assembly and conformations of modified peptide components in the presence of copper(II)*

To assess the effect of copper(II) on the peptide part of the final templated assembly, the modified peptides reported in sequences of *Table 2.2* were analyzed also in the present of copper(II) sulfate to evaluate the effect of eventual coordination of the copper ion. A titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of copper(II) sulfate in a sample including three different modified peptides with the same length and orientation at 6 μM as final concentration (hence, 2 μM of each single peptide) was carried out. CD experiments were performed in HEPES buffer 1 mM (pH:7.4).

Overall, the ratio in CD signals at 222/208 nm was found to be lower than 1 for all the lengths of modified peptide samples analyzed.

The titrations regarding the 19mer peptides, in both orientations, showed a progressive decrease, towards positive values of molar ellipticity of the 222 nm and

208 nm bands, although more pronounced in the former band. The spectra of both 19mer series are reported in *Figure 2.14*.

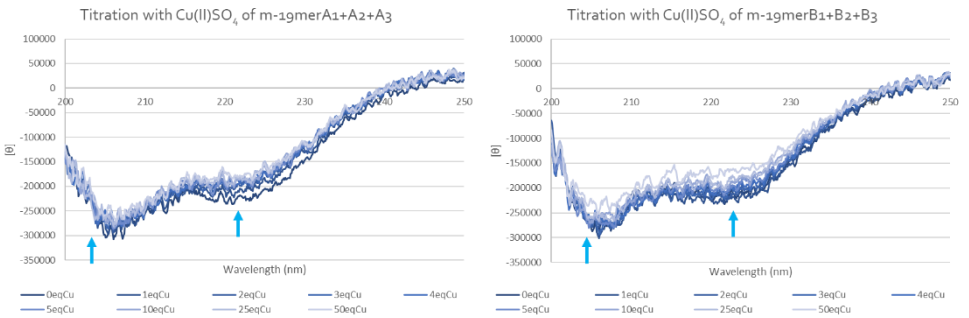



***Figure 2.14.*** *Titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate of 19mer modified peptides a) orientation A and b) orientation B.*

As it is possible to see from the *Figure 2.15*, after the addition of only 1 eq of Cu(II) is present a slight change in both bands, not consistent with the induction of coiled-coil structures as measured by the CD intensities at 222/208 nm. A similar titration with addition of only water confirmed that it is the presence of Cu(II) to induce the change of the CD signal, since no variation was observed in this case (*Figure 2.15*).

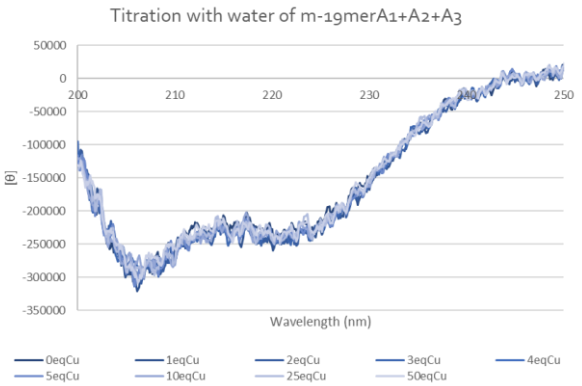



***Figure 2.15.*** *Titration with water of orientation A 19mer modified peptides.*

Therefore, the presence of Cu(II) ion for the 19mer modified peptides do not induce self-assembly to generate a coiled-coil structure but the metal addition leads to a slight change in the conformation/aggregation.

A different behavior was observed with 11mer modified peptides, given that after the addition of 1 eq of Cu(II), an increase toward negative values of molar ellipticity of the band at 222 nm was observed. The CD signal then remained constant for the subsequent additions (*Figure 2.16*). However, the CD intensity ratio at 222/208 remained less than one, thus suggesting that Cu(II) could not induce the formation of aggregates.

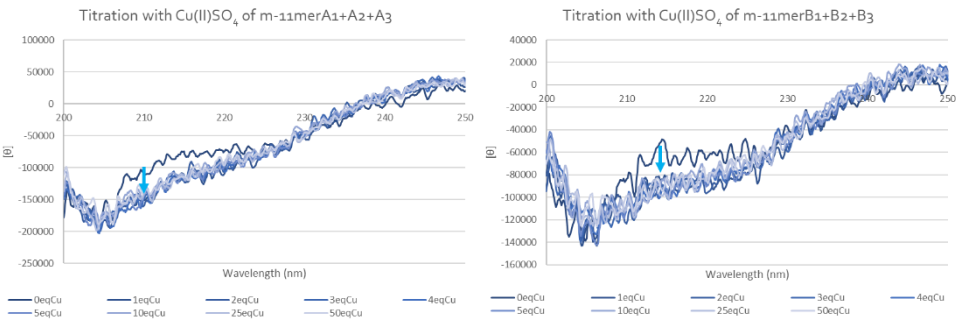



***Figure 2.16.*** *Titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate of 11mer modified peptides a) orientation A and b) orientation B.*

By analyzing the CD results regarding 5mer modified peptides in both the orientations no variation compared to the spectra measures without the copper ion were observed (*Figure 2.A9*). Despite the modifications with amino acids containing coordinating groups, these sequences remain excessively short to self-assemble and effectively bind copper(II).

### 2.3.3 Design, assembly, and conformations of PNA-peptide conjugates

A library of PNA-peptide conjugates was designed and synthesized with the aim of forming a PNA:DNA duplex in an antiparallel configuration. This means that the C-term moiety of PNA-peptide sequences is directed to the central core of the system as well as the 5'-term of the DNA sequences. In this way, the peptide domains of the antiparallel conjugates are placed towards the center of the system. The PNA-peptide conjugates are synthesized using an orthogonal chemistry on solid-phase

manual synthesis strategy. In fact, the base structure of all PNA-peptide sequences was the following:

PNA sequences-Lys(Mtt)-Spacer-Lys(Dde)(Peptide→)-Solid support,

where the several peptide sequences were linked to the side chain of one lysine at the C-term of the corresponding PNA domains.

To study the minimal length necessary to create a heterotrimeric coiled-coil system, three types of modified peptide domains were produced with different lengths: 5-mer, 11-mer and 19-mer. In addition, each peptide moiety was synthetized in both possible orientation in relation to the corresponding PNA domain.

In addition, the three different PNA sequences lacking the peptide domains, were synthetized to be used as a control for the assembly of DNA:PNA-peptides platform.

A summary of all the PNAs is reported in *Table 2.3*.

***Table 2.3.*** *Sequence of PNA-peptides synthetized to create the mini-protein assembly.*

| Entry | PNA-peptide | Sequence | MW (g/mol) | Yield (%) |
|---|---|---|---|---|
| 27 | 1A-1 | Ac-GGGCATGATCT-k(Ac)-AEEA-k(Ac-meekh→)-$NH_2$ | 4216.28 | 15 |
| 28 | 1A-2 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k(Ac-leekc→)-$NH_2$ | 4124.22 | 13 |
| 29 | 1A-3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(Ac-leekh→)-$NH_2$ | 4182.24 | 18 |
| 30 | 2A-1 | Ac-GGGCATGATCT-k(Ac)-AEEA-k (Ac-lkameekhkal→)-$NH_2$ | 4841.10 | 12 |
| 31 | 2A-2 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k (Ac-lkaleekckal→)-$NH_2$ | 4749.04 | 11 |
| 32 | 2A-3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k (Ac-lkaleekhkal→)-$NH_2$ | 4807.06 | 16 |
| 33 | 3A-1 | Ac-GGGCATGATCT-k(Ac)-AEEA-k (Ac-leeklkameekhkaleekl→)-$NH_2$ | 5840.20 | 8 |
| 34 | 3A-2 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k (Ac-leeklkaleekckaleekl→)-$NH_2$ | 5748.14 | 4 |
| 35 | 3A-3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k (Ac-leeklkaleekhkaleekl→)-$NH_2$ | 5806.16 | 7 |
| 36 | 1B-1 | Ac-GGGCATGATCT-k(Ac)-AEEA-k(Ac-hkeem→)-$NH_2$ | 4216.28 | 7 |
| 37 | 1B-2 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k(Ac-ckeel→)-$NH_2$ | 4124.22 | 10 |
| 38 | 1B-3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(Ac-hkeel→)-$NH_2$ | 4182.24 | 10 |

| | | | | |
|---|---|---|---|---|
| 39 | 2B-1 | Ac-GGGCATGATCT-k(Ac)-AEEA-k (Ac-lakhkeemakl→)-$NH_2$ | 4841.10 | 10 |
| 40 | 2B-2 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k (Ac-lakckeelakl→)-$NH_2$ | 4749.04 | 14 |
| 41 | 2B-3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k (Ac-lakhkeelakl→)-$NH_2$ | 4807.06 | 22 |
| 42 | 3B-1 | Ac-GGGCATGATCT-k(Ac)-AEEA-k (Ac-lkeelakhkeemaklkeel→)-$NH_2$ | 5840.20 | 5 |
| 43 | 3B-2 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k (Ac-lkeelakckeelaklkeel→)-$NH_2$ | 5748.14 | 3 |
| 44 | 3B-3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k (Ac-lkeelakhkeelaklkeel→)-$NH_2$ | 5806.16 | 7 |
| 45 | PNA 1 | Ac-GGGCATGATCT-k(Ac)-AEEA-k(Ac)-$NH_2$ | 3561.55 | 11 |
| 46 | PNA 2 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k(Ac)-$NH_2$ | 3521.53 | 16 |
| 47 | PNA 3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(Ac)-$NH_2$ | 3545.55 | 14 |
| Capital letters: PNA monomers, Lower case: amino acid | | | | |

Subsequently, CD measurements were carried out with PNA-peptide conjugates, listed in *Table 2.3*, to study if these sequences are capable to self-assemble into coiled-coil structures without the presence of the rigid trimeric template. In this case, it is possible to understand if PNA domains, providing additional weak interactions, can favor the formation of homo or hetero coiled-coil structures, even in the absence of Watson-Crick pairing.

CD measurements of each single PNA-peptide conjugate at 6 μM in HEPES buffer 1 mM (pH:7.4) were performed. In addition, the three PNA-peptide conjugates with the same length and orientation were analyzed in combination in the same solution with 6 μM total concentration (therefore 2 μM of each single PNA-peptide) to evaluate the possible formation of homo- and hetero-aggregates, with the same approach described in the previous section for the peptides.

Analyzing the CD data of PNA-peptide conjugates with 19mer peptide domains in the orientation A (corresponding to *3A-1, 3A-2 and 3A-3*), in the mixture of all three components a CD signal corresponding to the sum of individual contribution was observed (*Figure 2.17*). As for the peptides, no evidence for hetero-coiled coil structure could be detected. For all PNA-peptide conjugates a signature more reminiscent of coiled-coil formation rather than a typical α-helix secondary structure

seems to be present both in samples containing a single PNA-peptide conjugate that in the one with the combinations of them. However, one of the three PNA-peptide conjugates, the one with the cysteine modification (3A-2), showed a more intense CD signal (*Figure 2.17a*). This significant difference in CD signature can be attributed to a possible dimerization of the cysteine residue or to the ability of this specific sequences to self-assemble in solution as aggregate structures. In addition, also in this case, it should be emphasized that excessive absorbance was present in this case (from HT channel of the CD instrument) at 208 nm and therefore the CD intensity of the corresponding band is not reliable. This consideration can be extended also to the CD results, later discussed, carried out with the other PNA-peptide sequences on *Table 2.3*.

Interestingly, no CD signal was detected in the 240-280 nm region, typical of the nucleobase chromophores. This implies that: a) the presence of a chiral peptide moiety conjugated to the PNA do not affect the conformation of the single strand PNA part. b) the formation of PNA:PNA partial helices favoring association of the peptides is unlikely, since dsPNA segments could in principle have preferential helicity and hence a CD signal.

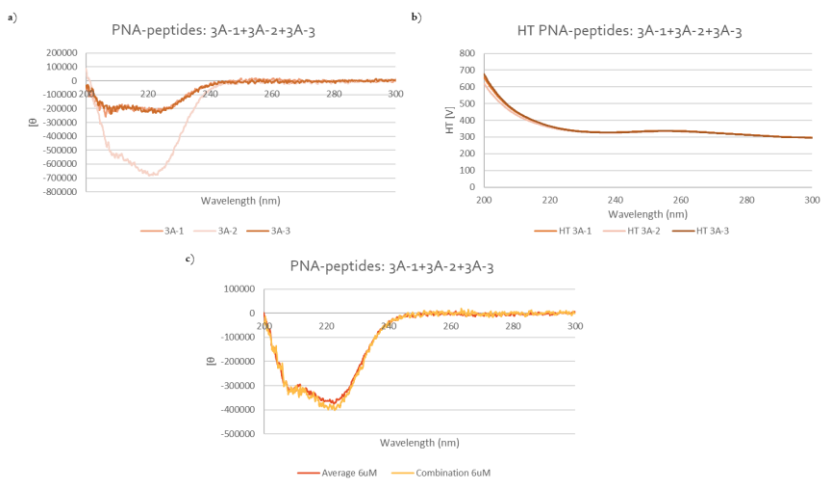



***Figure 2.17.*** *a) CD results of 3A-1, 3A-2, and 3A-3 as single separate PNA-peptide conjugates in solution at 6 µM as concentration; b) HT spectra of 3A-1, 3A-2, and 3A-3 as single separate PNA-peptide conjugates in solution. c) Comparison between CD spectrum obtained by the sample containing all the three PNA-peptide sequences with 19mer peptides domains in orientation A and the spectrum carried out by the mathematic elaboration of the results given by each individual sequence.*

Similar CD results (*Figure 2.18*) were obtained taking into consideration the same lengths of PNA-peptides but the opposite peptide domain orientation (corresponding to *3B-1, 3B-2, and 3B-3*). Nevertheless, in this case the three PNA-peptides analyzed do not shown significant difference with each other in terms of CD signature, which seems to be more like the formation of coiled-coil structures.

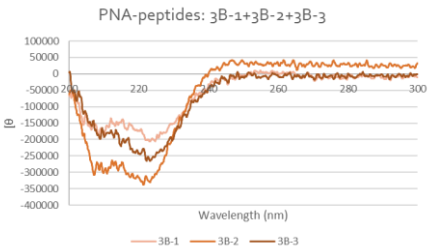



***Figure 2.18.*** *a) CD results of 3B-1, 3B-2, and 3B-3 as single separate PNA-peptide conjugates in solution at 6 μM as concentration.*

For the PNA-peptide conjugates with 11mer and 5mer peptides in both the orientations, the CD results showed the typical signature of α-helix structures. This observation can be turn out whether the conjugates are analyzed individually or in mixture, confirming that these sequences are too short to spontaneously form either homo- or hetero-coiled-coil structures in solution. CD data are reported in the Appendix *Figure 2.A10, 2.A11, 2.A12 and 2.A13*.

### *2.3.3.1 Assembly and conformations of PNA-peptide components in the presence of copper(II)*

A titration of the PNA-peptide conjugates with copper(II) sulfate was also carried out to study eventual changes due to the addition of PNA domains to the peptides described in *paragraph 2.3.2*. Also in this case, the titration was performed with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Cu(II) in a sample including three different PNA-peptide sequences with the same length and orientation at 6 μM as final concentration (hence, 2 μM as a contribution of each single PNA-peptide sequence). CD measurements were performed in HEPES buffer 1 mM (pH:7.4). For each titration

only the band a 222 nm was monitored, given that the absorbance at 208 nm turn out to be excessive (HT channel higher than 450V) to provide reliable data. All the HT spectra are reported showing this effect are reported in *Figure 2.A14* in Appendix. In the titrations of the PNA-conjugates with 19mer peptides in both the orientations (*Figure 2.19*), a progressive slight reduction to less negative values in the molar ellipticity were reported by averaging the values between 223 nm and 221 nm to suppress the noise. A similar behavior was observed also in the titration carried out with the 19mer peptides, in each orientation, discussed in *paragraph 2.3.2*. Also in this case, however, this behavior can be attributed to the formation of aggregates induced by the Cu(II).

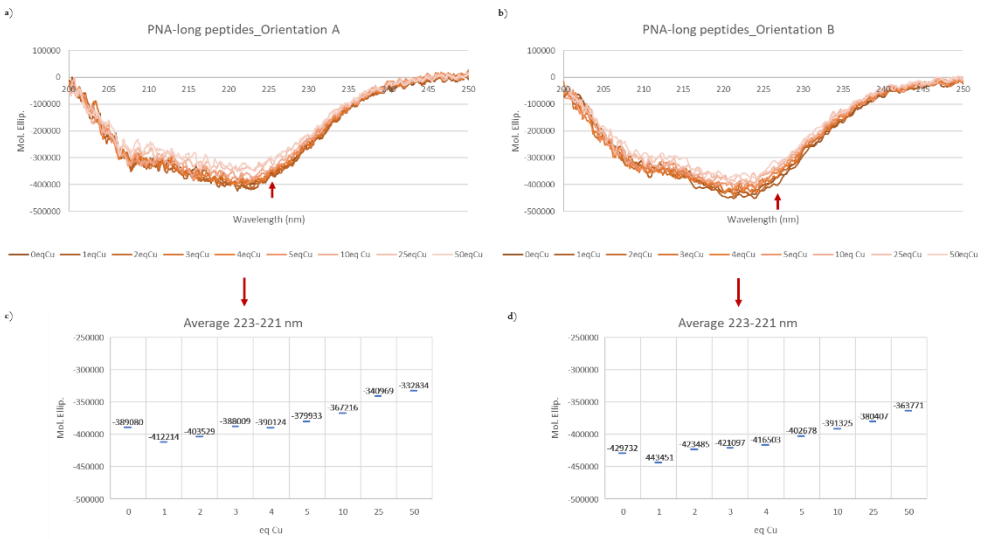



***Figure 2.19.*** *Titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate of PNA-peptide conjugates with 19mer peptide sequences a) orientation A and b) orientation B.*

Opposite behavior was noted, instead, for the titrations of the PNA-peptide conjugates with 11mer peptide sequences in both orientations since after the addition of 1 eq of Cu(II) an increase toward more negative values of molar ellipticity of the 222 nm band (*Figure 2.20*) was observed. Also, in this case a comparable response was observed during the titration conducted using the 11mer peptide sequences without the PNA domains in both the orientations. So, in this instance as

well, the introduction of copper metal contributes to the formation of aggregates in the solution, but without significant apparent role of the PNA part.

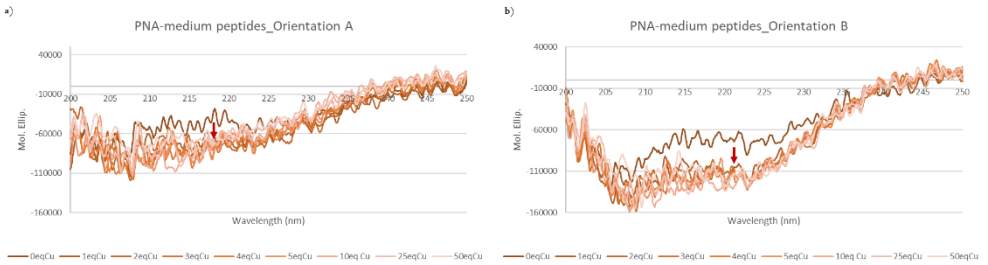



***Figure 2.20.*** *Titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate of PNA-peptide conjugates with 11mer peptide sequences a) orientation A and b) orientation B.*

In the CD data obtained for the titration of the PNA-peptide conjugates with 5mer peptide sequences, no significant changes were observed when Cu(II) ion was added (*Figure 2.A15*), confirming the fact that these sequences remain excessively short to self-assemble into a coiled-coil structures in solution and this behavior is not affected by the presence of a conjugated PNA segment or of copper(II).

### 2.3.4 Design and synthesis of trimeric DNA template and its variants

The DNA template synthesis was carried out at MGU using a synthetic protocol based on the standard phosphoramidite chemistry for DNA[50] and using as *central core* a small trifunctional molecule, namely 2-hydroxymethyl-1,3-propanediol (HMPD), bearing three orthogonal functional groups, which enabled the growth of three different DNA arms. In specific, the HMPD molecule was modified with a phosphoramidite building block (2-cyanoethyl, *N,N*-diisopropyl phsphoramidite, CED) and two orthogonal protecting groups: dimethoxytrityl (DMT) and levulinyl (Lev) group. A schematic representation of the trimeric DNA template synthesis and the structure of the modified HMPD derivative used a central core is shown in *Figure 2.21*. Standard phosphoramidites (5′-DMT and 3′-CED) were used for first 5′ to 3′ DNA synthesis (using 3′ OH to 5′ CED coupling) on controlled pore glass (CPG).

The HMPD monomer was then coupled to the terminal 5′ OH of the first segment. Reverse phosphonamidites (3′-DMT 5′-CED) were then used for the second branch synthesis, which proceeded in the 5′ to 3′ direction after DMT removal. After completion, the second DNA segment was capped and the levulinyl protecting group was removed by treatment with hydrazine hydrate in 1:1 pyridine/acetic acid. The third DNA segment was then synthesized in the 5′ to 3′ direction as for the second branch. The construct was then cleaved from the CPG using ammonium hydroxide and purified by a denaturing polyacrylamide gel electrophoresis (PAGE) analysis.

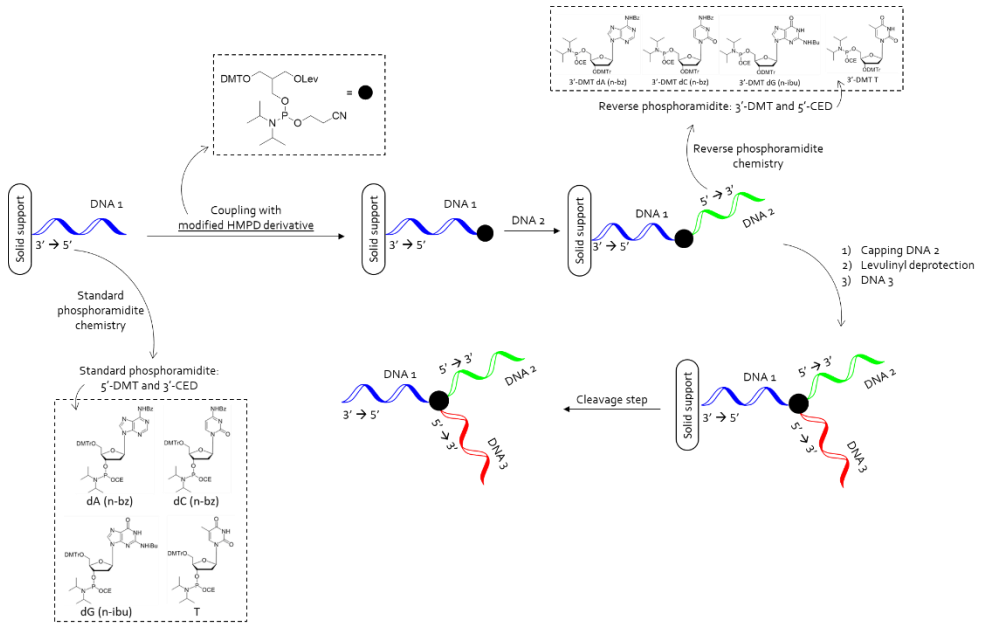



***Figure 2.21.*** *Trimeric DNA synthesis.*

The first tripodal template was synthetized to have all the three DNA sequences complementary to the PNA probes. The sequences of these trimeric DNA are reported in *Table 2.4* and the DNA template structure in *Figure 2.22*.

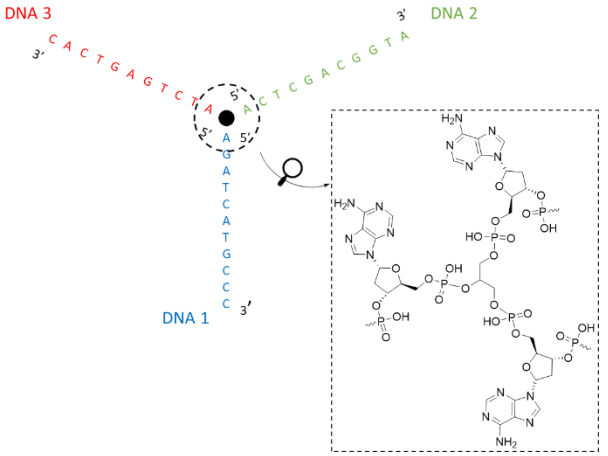



***Figure 2.22.*** *Trimeric DNA template structure.*

Subsequently, other two trimeric tripodal DNA were synthetized: a fluorescence version of the DNA bearing a cyanine 3 (Cy3) fluorophore (placed at the end of the sequence named *DNA 3)* and a DNA bearing a different *DNA3* sequence to be used as a mismatched arm (not complementary to any PNA sequence) as a negative control to demonstrate the specificity of PNA:DNA interactions (the sequence named *DNA 3* in *Table 2.4* was replaced with sequence called *mismatched arm*).

***Table 2.4.*** *DNA sequences.*

| Entry | Name | Sequence |
|---|---|---|
| 48 | DNA 1 | 5′–AGATCATGCCC–3′ |
| 49 | DNA 2 | 5′–ACTCGACGGTA–3′ |
| 50 | DNA 3 | 5′–ATCTGAGTCAC–3′ |
| 51 | DNA 3-fluo | 5′–ATCTGAGTCAC [$Cy_3$]–3′ |
| 52 | Mismatched arm (DNA 3′) | 5′–TCTTCTATACT–3′ |

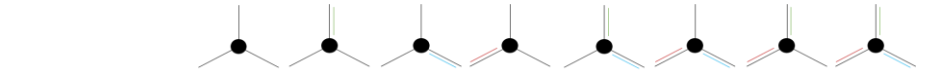
## 2.3.5 PAGE studies of the assembly process

The assembly of heterotrimeric coiled-coil systems with the fully matched trimeric DNA template (hereafter indicated as TP) and the several PNA-peptides conjugates was followed using PAGE analysis. Several screenings were carried out to explore the optimal PAGE conditions to use, by varying:

- TP:PNA-peptide ratio (1:1, 1:1.5, 1:2 and 1:5)
- running buffer (TBE (Tris-Borate-EDTA) both pH 8 and pH 6.5 and, TAMg (Tris-$MgCl_2$-glacial acetic acid) at pH 8)
- sample concentration: (1 μM:2 μM, 2.5 μM:5 μM, and 5 μM:10 μM of trimeric template and PNA-peptides respectively)
- staining conditions (GelRed, CyberGold and SilverStain)
- reaction buffer (water and different phosphate buffers saline (PBS). In specific, one formed by 100 mM of NaCl and 10 mM of phosphate, a second with 10 mM of NaCl and 1 mM of phosphate and a third with only 1 mM of phosphate).

The preliminary screenings led to the identification of the following parameters to perform optimal PAGE analyses:5% native PAGE with 1:2 as ratio of TP/PNA-peptides in TBE buffer (pH:6.5), and, in specific, with 1 μM and 2 μM concentrations for TP and PNA-peptide conjugates, respectively. All the samples are prepared in water and submitted to thermal annealing with a temperature ramp from 95 °C to 4 °C in 1h.

In a qualitative analysis, reported in *Figure 2.23*, the progressive changes in mobility were observed moving from the TP alone (*lane 1*) to the mixture of all components (*lane 8*), which can be attributed to the assembly of each component in solution, demonstrating the correct hybridization of the complex containing all the building blocks (TP and each triplet of PNA-peptides conjugates).



| TP | + | + | + | + | + | + | + | + |
|---|---|---|---|---|---|---|---|---|
| PNA-peptide 1 | - | + | - | - | + | + | - | + |
| PNA-peptide 2 | - | - | + | - | + | - | + | + |
| PNA-peptide 3 | - | - | - | + | - | + | + | + |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

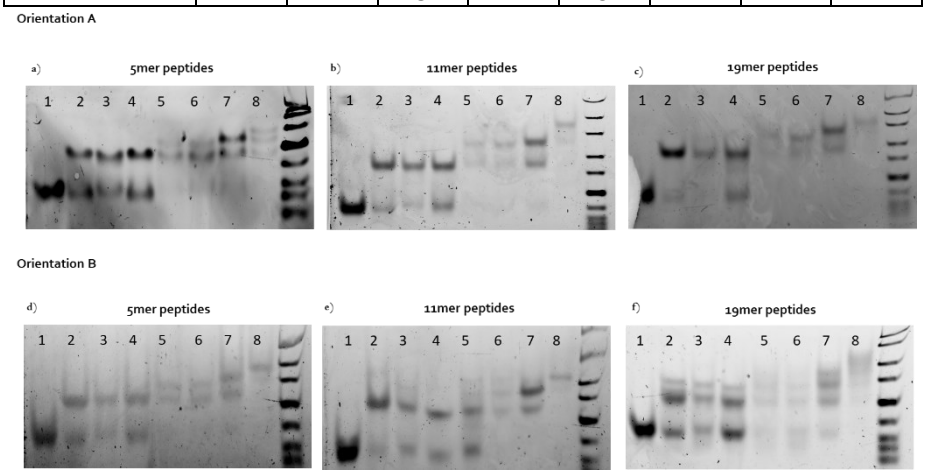
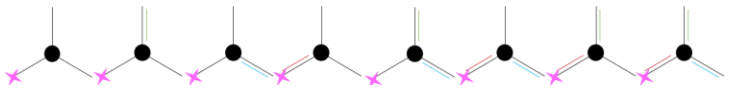



***Figure 2.23.*** *Qualitative PAGE experiments to study the progressive changes in mobility of the assembly. Each band of all the gels from the left to right corresponds to 1) Trimeric template (TP) alone; 2) TP+PNA-peptide 1; 3) TP+PNA-peptide 2; 4) TP+PNA-peptide 3; 5) TP+PNA-peptide 1+PNA-peptide 2; 6) TP+PNA-peptide 2+PNA-peptide 3; 7) TP+PNA-peptide 1+PNA-peptide 3; 8) TP+PNA-peptide 1+PNA-peptide 2+PNA-peptide 3. The first three gels on the top correspond to the orientation A of peptide sequences respectively (from left to right) to a) 5-mer, b) 11-mer and c) 19-mer as lengths of peptides moiety (from Entry 27 to Entry 35). The three gels on the bottom are respectively relative to d) 5-mer, e) 11-mer and f) 19-mer as lengths of peptides domains for the orientation B (from Entry 36 to Entry 47).*

Subsequently, the same experiment was performed with Cy3-labelled TP to achieve a selective visualization of the TP-containing components and the possibility to obtain an evaluation of the amount of complex formed during each hybridization step. These semi-quantitative experiments also confirmed the identity of the bands in the previously qualitative analysis (*Figure 2.24*).



| Cy3-labelled TP | + | + | + | + | + | + | + | + |
|---|---|---|---|---|---|---|---|---|
| PNA-peptide 1 | - | + | - | - | + | + | - | + |
| PNA-peptide 2 | - | - | + | - | + | - | + | + |
| PNA-peptide 3 | - | - | - | + | - | + | + | + |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

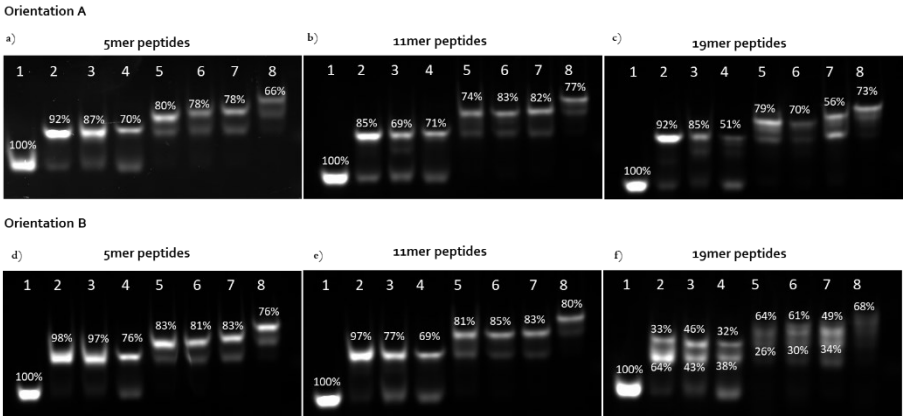
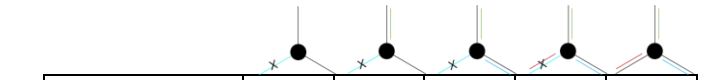


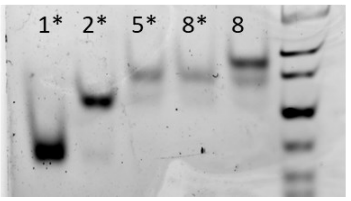

***Figure 2.24.*** *PAGE experiments for semi-quantitative evaluation of complex formed during each hybridization step and to confirm the identity of the bands in the qualitative PAGE analysis. Each band of all the gels from the left to right correspond to 1) Cy3-labelled trimeric template (Cy3-TP) alone; 2) Cy3-TP+PNA-peptide 1; 3) Cy3-TP+PNA-peptide 2; 4) Cy3-TP+PNA-peptide 3; 5) Cy3-TP+PNA-peptide 1+PNA-peptide 2; 6) Cy3-TP+PNA-peptide 2+PNA-peptide 3; 7) Cy3-TP+PNA-peptide 1+PNA-peptide 3; 8) Cy3-TP+PNA-peptide 1+PNA-peptide 2+PNA-peptide 3. The first three gels on the top correspond to the orientation A of peptide sequences respectively (from left to right) to a) 5-mer, b) 11-mer and c) 19-mer as lengths of peptides moiety (from Entry 27 to Entry 35). Instead, the three gels on the bottom are respectively relative to d) 5-mer, e) 11-mer and f) 19-mer as lengths of peptides domains for the orientation B (from Entry 36 to Entry 47). Fluorescence intensities are normalized with respect of that of TP alone (lane 1).*

The specificity of the assembly of the desired heterotrimeric coiled-coil system has been confirmed with the use of a mismatched TP (*Figure 2.25*). Indeed, only two of the three PNA-peptide conjugates were assembled on mismatched TP (*Fig. 2.25, lane 8**), while a complete architecture was obtained only by using the fully matched TP as a reference (*lane 8*). In addition, the two samples in lanes named 2* and 5* were added to observe the changes in mobility, which has remained constant from 5* to 8*.



| TP | - | - | - | - | + |
|---|---|---|---|---|---|
| Mismatched TP | + | + | + | + | - |
| PNA-peptide 1 | - | + | + | - | + |
| PNA-peptide 2 | - | - | + | + | + |
| PNA-peptide 3 | - | - | - | + | + |
| | 1* | 2* | 5* | 8* | 8 |

a) **Orientation A: 5mer peptides**




b) **Orientation B: 5mer peptides**

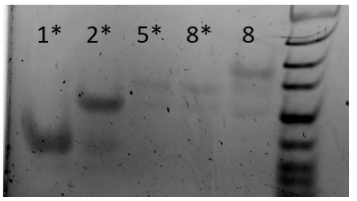
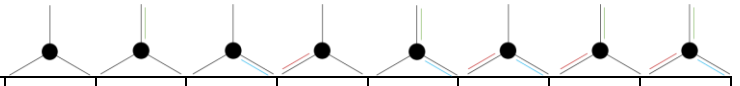



***Figure 2.25.*** *PAGE experiments carried out in presence of mismatched trimeric DNA template to test the specificity of the self-assembly of the desired complex. Each band of all the gels from the left to right correspond to 1*) Mismatched TP alone; 2*) Mismatched TP+PNA-peptide 1; 5*) Mismatched TP+PNA-peptide 1+PNA-peptide 2; 8*) Mismatched TP+PNA-peptide 1+PNA-peptide 2+PNA-peptide 3; 8) TP+PNA-peptide 1+PNA-peptide 2+PNA-peptide 3. a) PNA-short peptides series in the orientation A, b) PNA-short peptides conjugates in the orientation B (from Entry 27 to Entry 29 and from Entry 37 to 38, respectively).*

In addition, a PAGE experiments with the fully matched TP, Cy3-labelled trimeric template, or mismatched TP and the PNAs alone (without peptides domains) were performed to study the interaction between this trimeric DNA and the PNA sequences and to observe the mobility changes with respect to the previous samples that present the peptides moiety. From *Figure 2.26*, it is possible to observe that also in this case, in presence of fully-matched TP, the changes in mobility were observed for each hybridization step observed, moving from the TP alone (*lane 1*) to the mixture of all components without peptides domains (*lane 8*). In addition, the same result was achieved in the quantitative PAGE experiments (*Figure 2.27*). Also, the specificity of the self-assembly of the scaffold was confirmed by PAGE experiments with mismatched TP (*Figure 2.28*). However, small changes in mobility, in specific from the first to the second lane, were observed comparing these PAGE experiments performed without peptides domain compared to those of the PNA-peptide conjugates.



| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| TP | + | + | + | + | + | + | + | + |
| PNA 1 | - | + | - | - | + | + | - | + |
| PNA 2 | - | - | + | - | + | - | + | + |
| PNA 3 | - | - | - | + | - | + | + | + |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

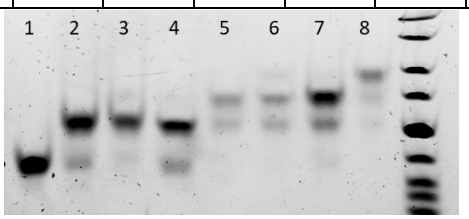
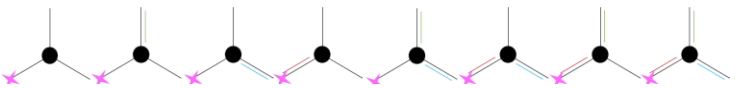



***Figure 2.26.*** *PAGE experiments with the fully matched TP and PNAs alone (without peptide domains).*



| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Cy3-labelled TP | + | + | + | + | + | + | + | + |
| PNA 1 | - | + | - | - | + | + | - | + |
| PNA 2 | - | - | + | - | + | - | + | + |
| PNA 3 | - | - | - | + | - | + | + | + |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

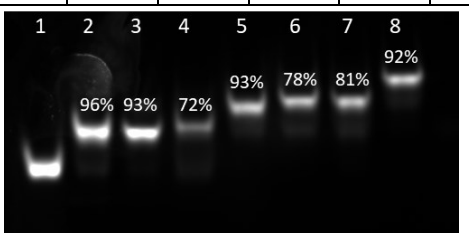
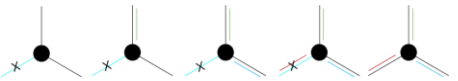



***Figure 2.27.*** *PAGE experiments with Cy-3labelled TP and PNAs alone (without peptide domains).*



| TP | - | - | - | - | + |
|---|---|---|---|---|---|
| Mismatched TP | + | + | + | + | - |
| PNA 1 | - | + | + | - | + |
| PNA 2 | - | - | + | + | + |
| PNA 3 | - | - | - | + | + |
| | 1* | 2* | 5* | 8* | 8 |

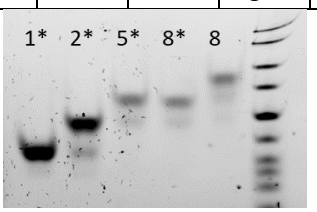



***Figure 2.28.*** *PAGE experiments with mismatched TP and PNAs alone (without peptide domains).*

## 2.3.6 CD studies of templated TP:PNA-peptide assembly

Since from previous data it was inferred that spontaneous aggregation to form coiled coil sequences is possible for longer peptides at higher concentrations, we focused our attention to the assembly formed by PNA-peptide with 11mer peptide sequences, to understand if the templated assembly could induce the formation of coiled-coil systems under conditions where neither the peptides nor the PNA-peptide conjugates can form them.

Therefore, the creation of heterotrimeric coiled-coil system within the assembly involving the PNA-11mer peptide conjugate in orientation B was investigated using circular dichroism (CD) analysis (*Figure 2.29)*. The experimental conditions of all the samples for CD measurements were consistent with those employed for PAGE analysis. This including maintaining a 1:2 as ratio between TP and PNA-Peptide (1 μM of TP and 2 μM of each PNA-Peptide), and the samples underwent thermal annealing treatment.

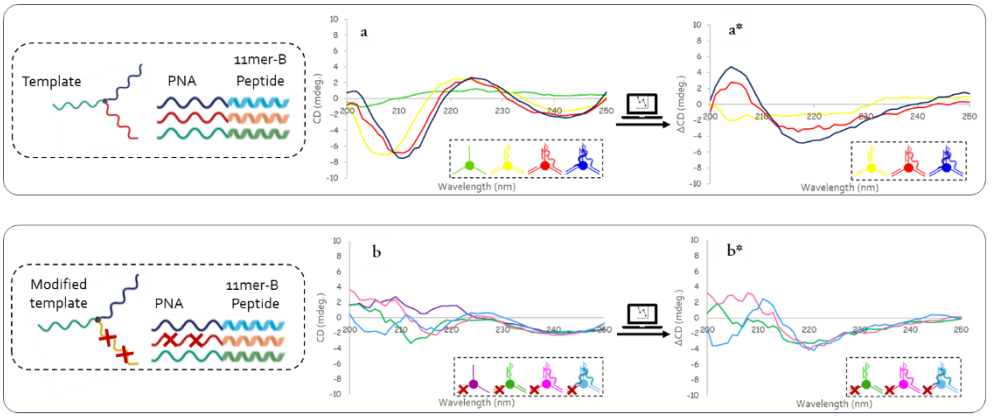



***Figure 2.29.*** *CD spectra showing the formation of heterotrimeric coiled-coil system with TP and PNA-11mer-peptide conjugates in Orientation B and CD spectra regarding a negative control; a* and b* are the difference spectra obtained by subtraction of the spectrum of the alone TP-PNA assembly (without the peptide part), thus indicating the contribute of the peptides in the assembly.*

CD spectra were collected for the assemblies formed by TP with one, two or three PNA-11mer peptide conjugates in orientation B (*Figure 2.29a*), as well as for one control sample. This control featured a modified (mismatched) template (*Figure 2.29b*). These measurements were conducted at 25 °C in water. In *Figure 2.29a*, a progressive increase in the intensity of the bands from naked tripodal DNA to heterotrimeric coiled-coil system was observed, particularly the negative peak around 208 nm, which is a specific signal associated with of α-helical structures. Since the CD spectrum also contains the contribution of the PNA:DNA duplexes, subtraction of the corresponding spectrum allows to analyze only the peptide part. By subtracting the CD signal of DNA:PNA complexes formed in presence of PNAs lacking the peptide domains from *Figure 2.29a and 2.29b*, it was possible to obtain the *Figure 2.29a* and 2.29b** in which emerge a negative band at 222 nm, a signature of coiled-coil formation. Notably, this negative band is more prominent in *Figure 2.29a**, thus confirming the formation of heterotrimeric coiled-coil system.

### *2.3.6.1 CD studied of templated TP:PNA-peptide assembly in the presence of copper(II)*

A titration experiment was conducted with 1, 2, 3, 4, 5, 10, 25 and 50 eq of copper(II) sulfate in a sample containing all the three distinct PNA-11mer peptide conjugates

in orientation B, each at a final concentration of 6 μM (resulting in a contribution of 2 μM for each individual PNA-peptide sequence), along with the trimeric DNA 1 μM concentration. CD measurements were performed in HEPES buffer 1 mM (pH:7.4). Moreover, also in this titration, absorbance (HT channel) was too high at 208 nm to obtain reliable CD results. Thus, only the band at 222 nm was monitored. The HT spectra are showed in *Figure 2.A16* in Appendix.

Despite the mathematical subtraction of contributions from the complex between trimeric DNA template and PNAs was not performed and consequently the typical CD signature of coiled-coil is not visible (*Figure 2.30)*, it possible to observe a slight increase toward negative values of molar ellipticity from spectrum without the Cu(II) addition to 3 eq of Cu(II). After that, additional equivalents of copper, up to 25 eq of Cu(II), seem to not bring further interactions able to stabilize the construct. Instead, at 50 eq of Cu(II) a dramatic change occurs in the CD spectrum: a more pronounced negative signal at 222 nm seems to indicate the formation of coiled-coil structures, but the typical signature of PNA:DNA duplex disappeared, suggesting drastic changes in the overall structures. The latter effect might be due to the effect of Cu(II) on the formation of higher-order aggregates or to the interaction of excess of copper(II) with DNA phosphates or the nucleobases.

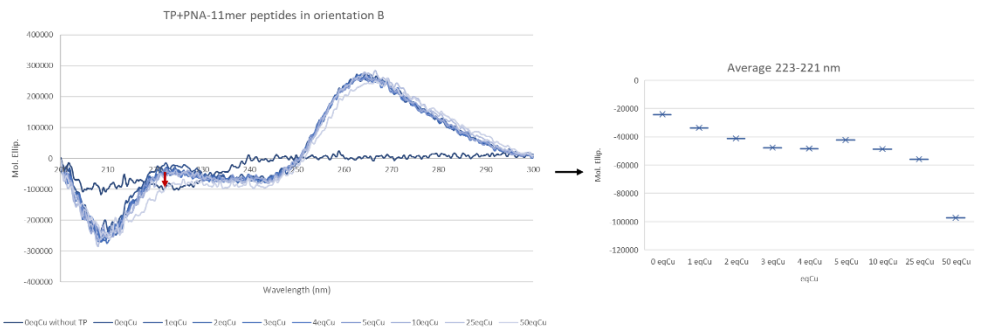



***Figure 2.30.*** *Titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate of the sample containing the trimeric template and three PNA-peptide conjugates with 11mer peptide domains in orientation B.*

With this CD experiments it was possible to show an effect of copper(II) which seemingly reinforce a coiled-coil system, but it was not possible to understand if the site of coordination for Cu(II) that was corresponding to the plastocyanin model. UV

and CD measurements of the Cu(II) chromophore could give more precise information about the donor atoms and hence evaluate if the coordination site corresponds to two histidines, a cysteine, and a methionine (available in peptide domains) as in the case of plastocyanin, but these types of measurements require higher concentrations, due to scarce sensitivity of the copper(II) chromophore. Higher concentrations however could lead to different type of aggregation, favoring higher order aggregates, and quantities of TP DNA that were not available. EPR or voltammetric studies also require concentrations in millimolar or at least sub-millimolar range and thus could not be pursued. For these reasons, exploiting the fluorescence quenching ability of copper(II), we chose fluorescence studies for the study of Cu(II) binding, since experiments could be performed in the same micromolar range used for PAGE and CD studies.

### 2.3.7 Fluorescence studies of templated TP:PNA-peptide assembly in the presence of copper (II)

To understand if Cu(II) can interact with its specific active site, formed by peptide domains, only when the total complex was assembled, fluorescence quenching experiments were carried out. In fact, titration experiments from 0.0 eq to 4.0 eq of $Cu(II)SO_4$ (with collection of the data each 0.2 eq of copper addition) were performed for several samples constituted by the target complex or, in other cases, by different building blocks as control experiments. In addition, the experimental conditions of all the samples for fluorescence quenching measurements were consistent with those employed for PAGE analysis and CD measurements previously discussed, in specific a 1:2 as ratio between TP and PNA-peptide conjugates (1 μM of TP and 2 μM of each PNA-peptide conjugate).

In a first series of experiments, TP+Cy3 was employed. The typical excitation wavelength for Cy3 is 550 nm, with a typical emission peak around 560-580 nm.
Three titrations were performed to study the fluorescence quenching behavior of the target architecture alone and with copper.
Initially, fluorescence emission spectra were collected to understand the contribution given by each single component. In fact, *Figure 2.31* highlights that each component have different effects on the fluorescence of TP-Cy3. *Figure 2.31a* shows that the addition of the first PNA-peptide with the methionine residue led to an increase of the fluorescence emission intensity, although it starts to decrease after the addition of the second PNA-peptide containing a cysteine residue and undergoes a sharp decrease after the addition of the last PNA-peptide conjugate (pairing to the labelled arm of TP-Cy3). This behavior was not observed in a parallel dilution experiment performed by adding water to TP-Cy3 (*Figure 2.31b*). However, in *Figure 2.31c*, in the PNA sequences it is possible to observe that after the additions of all the three PNAs the fluorescence emission intensity was decreased, and this effect was similarly observed for the TP:PNA-peptide assembly (*Figure 2.31a*). This PNA-induced quenching effect was observed only when PNA C (targeting the

labelled DNA branch) was added, and therefore is most likely due to the quenching effect of flanking nucleobases on Cy3 emission.

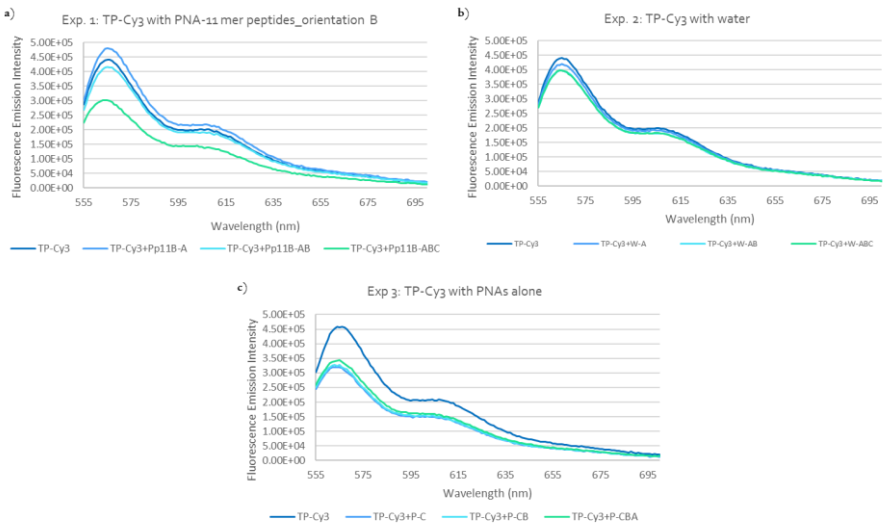



***Figure 2.31.*** *Fluorescence emission spectra collected after the addition of each building blocks of a) TP-Cy3 with all the three PNA-peptide conjugates (ABC) with 11mer peptide domains in orientation B; b) TP-Cy3 alone; c) TP-Cy3 with the PNAs without the peptide domain (ABC).*

After this, three titrations with copper(II) were performed (*Figure 2.32*). *Titration 1* was carried out with the TP+Cy3 in presence of the three PNA-peptide conjugates with 11mer as peptide domains in orientation B. Instead, *Titration 2* was formed using only TP+Cy3 and *Titration 3* was performed on the mixture of TP+Cy3 and the three different PNAs without the peptide domains. These preliminary tests were carried out in water.

Analyzing the data of titration from 0eq to 4.0 eq of $Cu(SO_4)$ obtained with the target complex in the sample, a similar fluorescence quenching was observed for the overall assembly (*Titration 1*) compared to the TP+Cy3:PNA assembly (*Titration 3*), but a higher fluorescence quenching was noted in the presence of only TP+Cy3 (*Titration 2*). However, since the Cy3 fluorophore is placed at the end of the DNA strand and thus far from the peptide core, it is possible that energy transfer by Förster mechanism from the fluorophore to Cu(II) is not effective in this case.

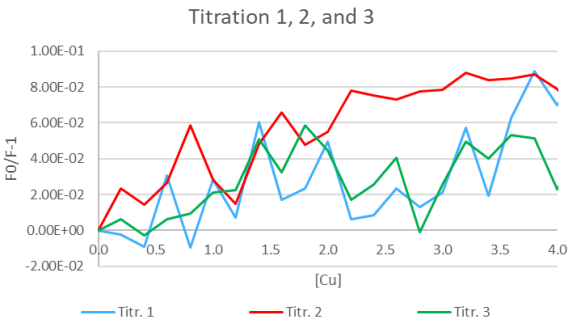



***Figure 2.32.*** *A comparison based on the F/Fo-1 values and equivalent of Cu(II) relative to titration 1, 2, and 3.*

To overcome this problem, two labelled versions of one sequence of PNA-peptide conjugate with 11mer in orientation B as peptide domains bearing a 5(6)-carboxytetramethylrhodamine (5(6)-TAMRA) were synthetized. After HPLC purification, the two sequences (one with 5-TAMRA and other with 6-TAMRA) were separated and for all the experiments only the sequences functionalized with 5-TAMRA were employed (the second peak in HPLC purification as reported in literature[51]). TAMRA is characterized by excitation at 552 nm and emission around 580-590 nm.

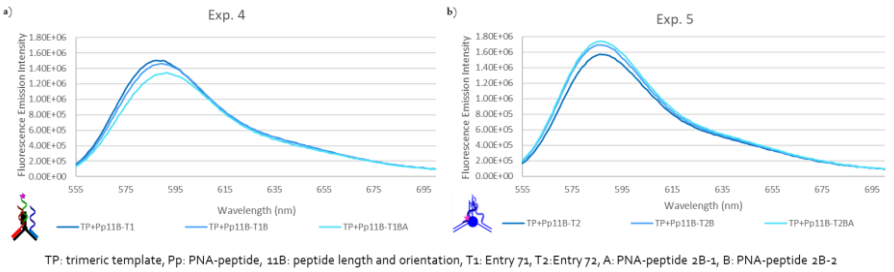
In one model, the coupling with the 5(6)-TAMRA was performed at the end of the peptide domain (*Entry 53*) while in another one the side chain of the lysine near the PNA sequence was selected as linking point (*Entry 54*). For the coupling with the fluorophore the PNA bearing the peptide sequence without methionine and cysteine residues was chosen as reported in *Table 2.5*. In addition, all the fluorescence quenching experiments discussed below were carried out in 10 mM of HEPES with NaCl 0.1 M at pH:7.4 as buffer.

***Table 2.5.*** *Two versions of PNA-peptide conjugates with 11mer in orientation B as peptide domains functionalized with 5-TAMRA and used for fluorescence quenching experiments.*

| Entry | Sigla | Sequence | MW (g/mol) | Yield (%) |
|---|---|---|---|---|
| 53 | T1 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(5-TAMRA-lakhkeelakl→)-$NH_2$ | 5177.46 | 14 |

| | | | | |
|---|---|---|---|---|
| 54 | T2 | Ac-GTGACTCAGAT-k(5-TAMRA)-AEEA-k(Ac-lakhkeelakl→)-$NH_2$ | 5177.46 | 9 |
| Capital letters: PNA monomers, Lower case: amino acid | | | | |

Fluorescence emission spectra in absence of copper(II) were collected to understand the contribution given by each single component added (*Figure 2.33*). Starting from the trimeric DNA and the *Entry 53* (*Exp. 4*) in one case or E*ntry 54* (*Exp. 5*) in the second experiment, similar intensities of fluorescence emission can be observed. However, after the addition of the others two PNA-peptide conjugates, it is possible to see a decrease in intensity for the *Exp. 4* with a red-shift, while only an increase of fluorescence emission in *Exp. 5*.




TP: trimeric template, Pp: PNA-peptide, 11B: peptide length and orientation, T1: Entry 71, T2:Entry 72, A: PNA-peptide 2B-1, B: PNA-peptide 2B-2

***Figure 2.33.*** *Fluorescence emission spectra collected after the addition of each building blocks of a) TP with Entry 53 and the other two PNA-peptide conjugates with 11mer peptide domains in orientation B; b) TP with Entry 54 and the other two PNA-peptide conjugates with 11mer peptide domains in orientation B.*

Quenching experiments were then carried out with addition of Cu(II) (*Figure 2.34a and 2.34b*). By comparing the data of the two titrations from 0 to 4.0 eq of $Cu(SO_4)$, a slightly more pronounced quenching was observed for the *Exp. 4* where the 5-TAMRA is attached at the $N_{term}$ of the peptide domain.

In addition, plotting the fluorescence emission intensities as a function of the equivalents of Cu(II), it can be noticed that there is a systematic dependence on copper(II) concentration and the largest quenching effect is observed for Exp. 4 (*Figure 2.34c*). In fact, from 0.0 eq to 4.0 eq of Cu(II), a 12% of fluorescence quenching was observed when the fluorophore was attached at the end of peptide domain, while only a 7% was noted when the 5-TAMRA is on the side chain of the lysine near

the PNA sequence. This difference can be attributed to the fact the fluorophore at the end of the peptide domain is closer to the specific coordination site of the copper. Furthermore, given this moderately more intense fluorescence quenching by design used in *Exp. 4*, the PNA-peptide conjugate with the 5-TAMRA at the $N_{term}$ of the peptide domain was selected for the next series of experiments.

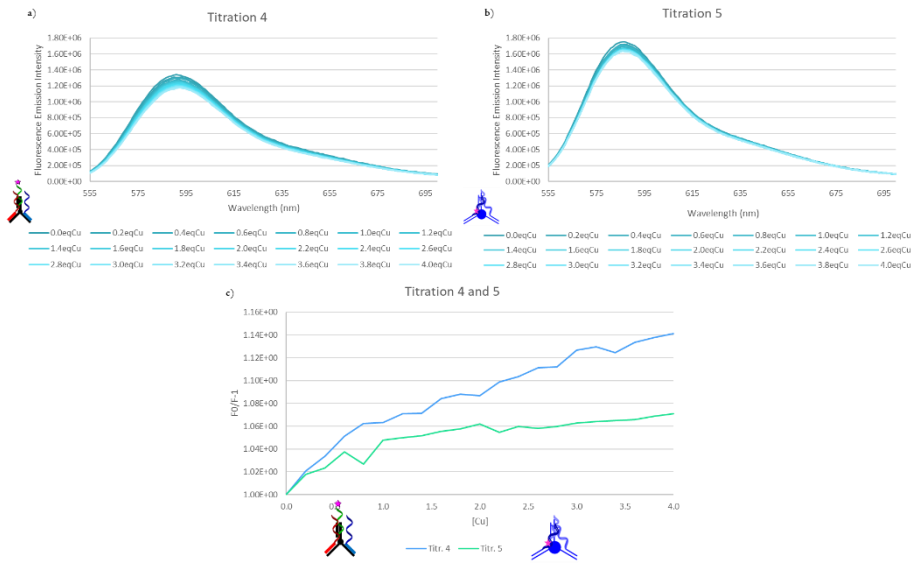



***Figure 2.34.*** *Fluorescence quenching titration of a) TP with Entry 53 and the other two PNA-peptide conjugates with 11mer peptide domains in orientation B; b) TP with Entry 54 and the other two PNA-peptide conjugates with 11mer peptide domains in orientation B; c) A comparison based on the F/Fo-1 values and equivalent of Cu(II) relative to titration 4 and 5.*

For this reason, in fact, the others PNA-peptide conjugates with different orientations and lengths in peptide domains were functionalized with 5(6)-TAMRA at the $N_{term}$ of the peptide domain, but as before only the sequences functionalized with 5-TAMRA were employed. All the sequences were listed in *Table 2.6*.

***Table 2.6.*** *PNA-peptide conjugates functionalized with TAMRA and used for fluorescence quenching experiments.*

| Entry | Sigla | Sequence | MW (g/mol) | Yield (%) |
|---|---|---|---|---|
| 55 | NM03-13 pk3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(5-TAMRA-leekh→)-$NH_2$ | 4552.65 | 8 |

| | | | | |
|---|---|---|---|---|
| 56 | NM03-09 pk2 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(5-TAMRA-lkaleekhkal→)-$NH_2$ | 5177.46 | 13 |
| 57 | CS02-193 pk2 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(5-TAMRA-leeklkaleekhkaleekl→)-$NH_2$ | 6176.56 | 4 |
| 58 | NM03-14 pk2 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(5-TAMRA-hkeel→)-$NH_2$ | 4552.65 | 12 |
| 59 | NM03-15 pk3 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(5-TAMRA-lkeelakhkeelaklkeel→)-$NH_2$ | 6176.56 | 5 |
| Capital letters: PNA monomers, Lower case: amino acid | | | | |

However, before moving toward the other heterotrimeric coiled-coil architectures with different peptide domains, the order addition of trimeric DNA and labelled PNA-peptide for the fluorimeter sample was tested. In fact, in the *experiment 4*, previously discussed, the first data was collected with the trimeric template and the labelled PNA-11mer peptide in orientation B conjugate. At this solution, then, was added first the PNA-peptide with the cysteine residue and subsequently the PNA-peptide sequence with the methionine residue. Instead, in the *experiment 6* (*Figure 2.35b*), the trimeric template was added at the solution containing already all three PNA-peptide conjugates of the same block, keeping the same order of addition used from *experiment 4*. In addition, a duplicate of *Exp. 4* was performed (*experiment 7*) and reported in *Figure 2.35c*.

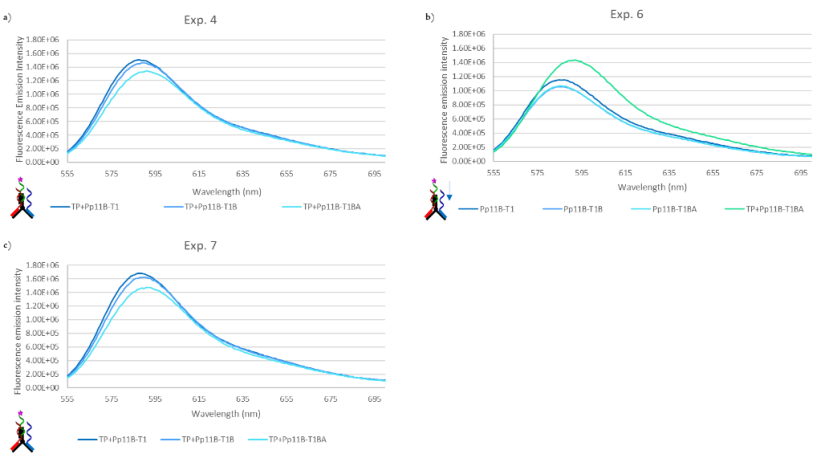



***Figure 2.35.*** *Fluorescence emission spectra collected after the addition of each building blocks of a) TP with Entry 53 and the other two PNA-peptide conjugates with 11mer peptide domains in orientation B; b) Entry 53, the other two PNA-peptide conjugates with 11mer peptide domains in orientation B and TP; c) duplicate of experiment 4.*

From *Figure 2.35*, some changes some additional information could be obtained by changing the order of addition of components. In fact, if the template was already present in the solution before the addition of PNA-peptide conjugates, only a small red shift can be observed with a slight fluorescence emission intensity decrease after the introduction of the last PNA-peptide conjugate (*Figure 2.35a and 2.35c*). Instead, the fluorescence intensity of the labelled PNA-peptide (*Entry 53*) was lower in the absence of TP (*Figure 2.35b*). No change was observed by addition of PNA-peptide B (2B-2, PNA-11mer peptide in B orientation with the cysteine residue), and a slight increase was observed when PNA-peptide A (*2B-1*, PNA-11mer peptide in B orientation with the methionine residue) was added. However, after the addition of the trimeric DNA as last component a larger fluorescence enhancement, accompanied by a bathochromic shift was observed, thus indicating that a different assembly was obtained.

After that, however, a titration with these two new discussed samples was carried out and the results are reported in *Figure 2.36*. Surprisingly, the two assemblies behaved in a different way in the Cu(II) titrations. In presence of the sample where the trimeric DNA was the last component added (*Titration 6*), a less pronounced fluorescence quenching was observed. On the other hand, the result obtained with the duplicate (*Titration 7*) were in line with those of *Titration 4*. This behavior seems to indicate that the final assembly might be dependent on the order of addition, i.e. that in the case of addition of PNA-peptides to the template, a metastable, more open structure can be obtained, whereas by allowing the PNA-peptide to interact without the template, the final assembly obtained by subsequent addition of TP reaches a more compact structure for which accessibility of the Cu(II) might be limited.

For this reason, subsequent experiments were carried out with an assembly obtained by the addition of trimeric template as a first component, and then the PNA-Peptide conjugates.

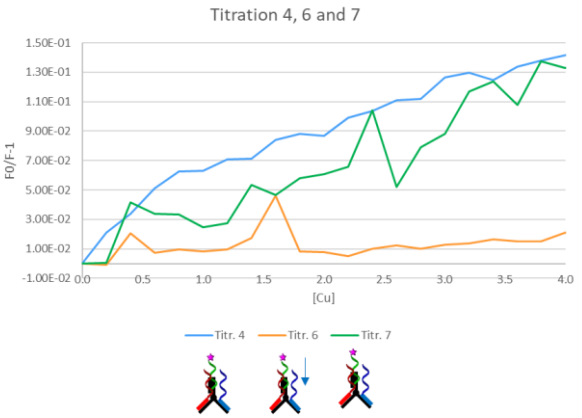



***Figure 2.36.*** *A comparison based on the F/Fo-1 values and equivalent of Cu(II) relative to titration 4, 6 and 7.*

Other heterotrimeric coiled-coil architectures with different peptide domains were also tested, using the same experimental conditions employed for the *Exp. 4*. Consequently, other five experiments were performed to study if the changes in terms of length and orientation of the peptide domains (11mer A, 19mer B, 19me A, 5mer B, and 5mer A) can be influence the fluorescence emission quenching. However, to make clear the composition of each sample for the fluorescence quenching titration experiments, the specific components are reported in *Table 2.7*.

***Table 2.7.*** *Sample composition of all the heterotrimeric coiled-coil structures for fluorescence emission titration experiments.*

| Titr. 4 | Titr. 8 | Titr. 9 | Titr. 10 | Titr. 11 | Titr. 12 |
|---|---|---|---|---|---|
| TP | TP | TP | TP | TP | TP |
| Entry 53 | Entry 56 | Entry 59 | Entry 57 | Entry 58 | Entry 55 |
| Entry 39 | Entry 30 | Entry 42 | Entry 33 | Entry 36 | Entry 27 |
| Entry 40 | Entry 31 | Entry 43 | Entry 34 | Entry 37 | Entry 28 |
|   |   |   |   |   |   |

*Figure 2.37*, all the results were plotted by relating the F/Fo-1 values with the equivalents of Cu(II). Comparing the data, it is possible to notice that the

fluorescence emission quenching seems to be present only when the PNA-peptide conjugates with 11mer in orientation B as peptide domains were employed. In fact, no significant quenching effect was observed for the other assemblies with different peptide lengths and orientation.

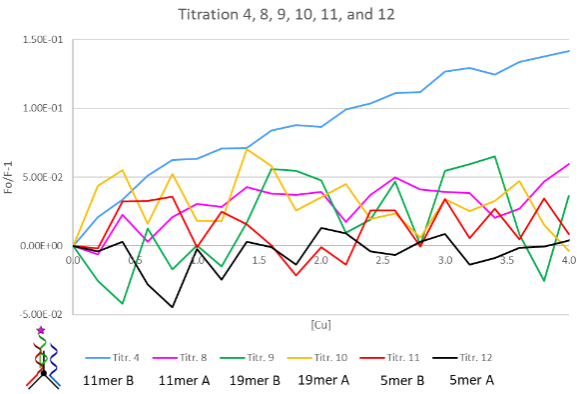



***Figure 2.37.*** *A comparison based on the F/Fo-1 values and equivalent of Cu(II) relative to titration 4, 8, 9, 10, 11, and 12.*

These results could be attributed to the higher distance between the coordination site for the Cu(II) and the fluorophore taking into consideration both the supramolecular structures within the 19mer peptide domains. Since the copper(II) binding sites are the same in these systems, these results are in line with the location of the copper(II) ion in the central part of the peptide. Instead, regarding the structures with shorter peptide domains (5mer), this trend could be explained through the inability of PNA-peptide in both orientations to form coiled-coil structures despite the presence of the rigid trimeric scaffold, being the peptide domains too short. Furthermore, an explanation of the titration carried out with the trimeric DNA and the PNA-peptide conjugates containing the 11mer in orientation A as peptide domains that showed a random behavior with respect to its counterpart in orientation B, could be found in the generation of minor number of electrostatic interactions.

The stability constants, reported in *Table 2.8*, were calculated using HypSpec program. The log β was calculated for the desired complex formed by trimeric DNA

and the PNA-peptide conjugates containing the 11mer in orientation B, where two replicates were performed. The data of the two replicates showed similar, but not the same log β, reason why a third replicate will be necessary, also because the standard deviation turned out to be slightly higher.

***Table 2.8.*** *Log beta of experiments 4 and 7.*

| | Log β | Standard deviation |
|---|---|---|
| Titration 4 | 6.60 | 0.2328 |
| Titration 7 | 6.00 | 0.1083 |

Taking into consideration only the complex formed by trimeric DNA and the PNA-peptide conjugates containing the 11mer in orientation B, given its fluorescence emission quenching (*Titration 4 and 7*), several controls were carried out to understand this interaction with Cu(II).

Maintaining fixed the labelled PNA-peptide conjugate (being the only component able to give a fluorescence signal), one of the other two unlabeled PNA-peptide sequence (and reverse) was replaced with the same PNA strand lacking the peptide domain (*Titration 13 and 14*). Furthermore, another titration (*Titr. 15*) was performed with the template DNA, the labelled PNA-peptide conjugate and two PNAs lacking the peptide domains. Then, also only the template DNA and the labelled PNA-peptide conjugate were tested (*Titration 16*).

In *Table 2.9* the list of each component used is reported, whereas in *Table 2.10* the components used in each experiment are reported.

***Table 2.9.*** *Components used for the first series of control experiments.*

| Name | | Entry | Sequences |
|---|---|---|---|
| TP | | Trimeric DNA | |
| Pp11B | T1 | Entry 53 | Ac-GTGACTCAGAT-k(Ac)-AEEA-k(5-TAMRA-lakhkeelakl→)$NH_2$ |
| | A | Entry 39 | Ac-GGGCATGATCT-k(Ac)-AEEA-k(Ac-lakhkeemakl→)-$NH_2$ |
| | B | Entry 40 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k(Ac-lakckeelakl→)-$NH_2$ |
| P | 1 | Entry 45 | Ac-GGGCATGATCT-k(Ac)-AEEA-k(Ac)-$NH_2$ |
| | 2 | Entry 46 | Ac-TACCGTCGAGT-k(Ac)-AEEA-k(Ac)-$NH_2$ |
| Capital letters: PNA monomers, Lower case: amino acid | | | |

***Table 12.10.*** *Composition of each titration from target complex (Titr. 4) to first series of control experiments (from Titr. 13 to Exp. 16).*

| | | Pp11B | | | P | |
|---|---|---|---|---|---|---|
| | TP | T1 | A | B | 1 | 2 |
| Titr. 4 | + | + | + | + | - | - |
| Titr. 13 | + | + | + | - | - | + |
| Titr. 14 | + | + | - | + | + | - |
| Titr. 15 | + | + | - | - | + | + |
| Titr. 16 | + | + | - | - | - | - |

The results of these fluorescence quenching titrations from 0.0 eq to 4.0 eq of Cu(II)$SO_4$ are shown in *Figure 2.38*, where the titration with the target complex is compared with the data obtained in presence of control samples. The fluorescence emission quenching observed for the complete assembly (*Titr. 4*), is more pronounced than that of the other experiments. In fact, a non-significant quenching (with random fluctuations) was observed for *Exp. 13* and *Exp. 14*. A systematic quenching was noted for the last two experiments, though amounting to less than 10% at 4.0 eq of Cu(II). The higher quenching observed in *Titrations 4* (and replicate T*itration 7*) compared to the other systems containing only some components could be attributed to the presence of all the four amino acid residues able to generate a specific coordination site for the copper ion near enough to fluorophore.

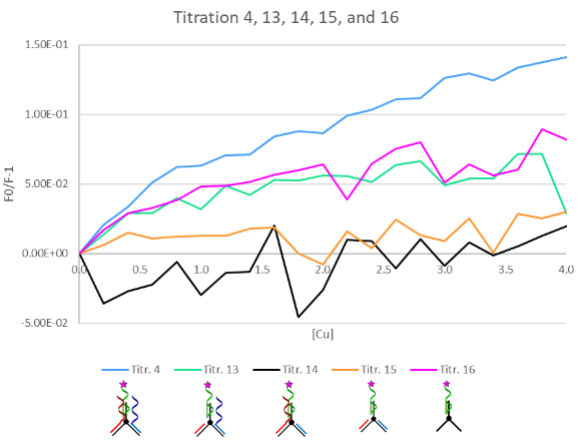



***Figure 2.38.*** *A comparison based on the F/Fo-1 values and equivalent of Cu(II) relative to titration 4, 13, 14, 15, and 16.*

A second block of control experiments was carried out, where the former was constituted by the labelled PNA-peptide conjugate (*Entry 53*) and the two PNA-peptide conjugate of the same block (*Entry 39 and 40*) without the DNA template, and the last by the labelled PNA-peptide conjugate (*Entry 53*) alone, in order to study the behavior of these components in absence of trimeric template and assess if the quenching observed could be due to the effect of copper (II) on single components or their combinations. These two controls are named *titration 17* (*Tirt. 17*) and 18 (*Tirt. 18*), respectively. However, the concentrations of the PNA-peptide conjugates used for both the experiments were maintained constant respect the previously analysis (2 μM for each strand). Comparing the data with the target complex (*Titration 4*), a similar trend of fluorescence emission intensity decrease can be observed by *Figure 2.39*.


***Figure 2.39.*** *A comparison based on the F/Fo-1 values and equivalent of Cu(II) relative to titration 4, 17, and 18.*

The quenching effect was similar in the two controls and in the complete assembly. This could be attributed to binding of the Cu(II) to the labelled peptide through histidine and other flanking amino acid side chains. Since these sequences are insufficiently long to spontaneously form either homo- or hetero-coiled-coil structures in solution, as previously reported by CD experiments, a similar effect was observed for the PNA-peptide alone and for the combination of all three sequences. Since both static and dynamic quenching might occur, the similar effect observed

for the single components could be the result of higher flexibility and higher accessibility of the fluorophore in the single component, which could counterbalance the stability of the complex.

To understand the stability of the assembly copper(II) complex compared to that of the peptide alone, a third block of control experiments was performed after the synthesis of the 11mer peptide strands in orientation B with 5(6)-TAMRA at the $N_{term}$. This labelled peptide sequence, used for fluorescence quenching analysis, is reported in *Table 2.11*.

***Table 2.11.*** *Labelled peptide sequence.*

| Entry | Sigla | Sequence | MW (g/mol) | Yield (%) |
|---|---|---|---|---|
| 60 | CS02-192 pk2 | 5-TAMRA-lakhkeelakl-$NH_2$ | 1690.98 | 24 |

Inside this third block of controls, three different experiments were conducted: a) labelled peptide sequence alone (*Titr. 19*); b) labelled peptide sequence with the trimeric template and the three PNAs without peptide domains listed in *Table 2.3* (*Trt. 20*); c) competition experiment with the labelled peptide (*Entry 60*) in the presence of the unlabeled assembly composed by the DNA template and the three PNA-11mer peptide (orientation B) conjugates reported in *Entry 39, 40* and *41* (*Trt. 21*). Furthermore, also for this block of controls the concentration regarding each component have not been changed. Consequently, the labelled peptide sequence was used at 2 μM as concentration, such as the PNA-peptide conjugates and the PNAs alone. Instead, the trimeric template was employed at 1 μM as concentration.

The fluorescence quenching titrations relative to these three control experiments are shown in *Figure 2.40*, where it is possible to observe that the labelled peptide alone undergoes a more efficient quenching effect when compared with the desired assembly (*Titration 4*, reported for comparison). The fluorescence quenching behavior obtained with the labelled peptide strand alone in presence of Cu(II), being higher also than the previously discussed experiment with the labelled PNA-peptide

conjugate alone (*Titration 18*), would seem to confirm that the copper ion is able to interact better with the fluorophore when super-secondary structures are not present in solution. This might be due to the higher flexibility of the free peptide that allows to bring the copper(II) in close proximity with the fluorophore. In fact, as showed before in CD experiments, this peptide sequence is too short to form coiled-coil structures in solution.

In addition, the fact that the quenching is slightly more evident in *Titration 21* respect *Titration 20* could be attributed to the formation of the heterotrimeric coiled-coil structure and the consequent competition of this for Cu(II) binding, thus confirming the Cu(II) binding ability of the assembly. However, due to the increased flexibility of the free peptide, enabling closer proximity between Cu(II) and the fluorophore, and the formation of the heterotrimeric coiled-coil structure, it is possible to account the slightly more pronounced quenching observed in *Titration 21* compared to *Titration 4*.

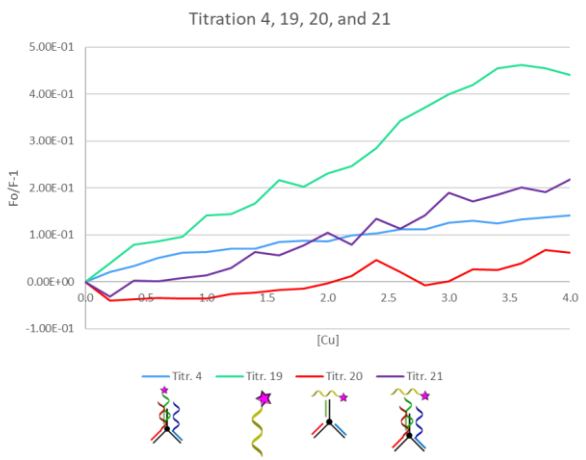



***Figure 2.40.*** *A comparison based on the F/Fo-1 values and equivalent of Cu(II) relative to titration 4, 19, 20 and 21.*

The last block of control experiments, formed by two distinct titrations, was carried out using, in one, the labelled PNA-peptide conjugate (*Entry 53*) and an unrelated 11mer linear DNA sequence, while in the other a complementary PNA strand was added. These controls, reported as T*itration 22* and *Titration 23* respectively, were

performed to understand the possible role of the copper binding by the DNA backbone. The DNA and PNA sequences were already present in our laboratory, and they are reported in *Table 2.12*.

***Table 2.12.*** *DNA and PNA sequences used for the last block of control experiments.*

| Entry | Sigla | Sequence |
|---|---|---|
| 61 | DNA Leoh 23.1 | 5'-CCACCGGAAAA-3' |
| 62 | PNA Leoh 23.1 | H-AEEA-AEEA-TTTCCGGTGG-$NH_2$ |

The concentrations of the labelled PNA-peptide conjugate used for both experiments were maintained constant with respect of the previously analysis (2 μM), instead the DNA oligomer, being that the trimeric DNA used for the desired complex is formed by three strands by 11mer in lengths and this linear DNA sequence is constituted by 11 nucleobases, was employed at 3 μM rather than 1 μM, to have a similar quantity of DNA in solution).

The results are reported in *Figure 2.41*, that shows as the fluorescence emission quenching in absence of the DNA:PNA duplexes in *Titration 22* is quite like the quenching obtained in the titration with the target construct (*Titr. 4*). Instead, the fluorescence quenching is higher in the *Titration 22*, where the DNA:PNA duplexes were formed. This behavior can be attributed to the interaction between the DNA backbone and the copper ion, given that when the DNA strand was alone in solution without the complementary PNA sequence the fluorescence emission quenching is lower, resulting in minor quantity of copper free able to interact with the fluorophore to give the quenching.

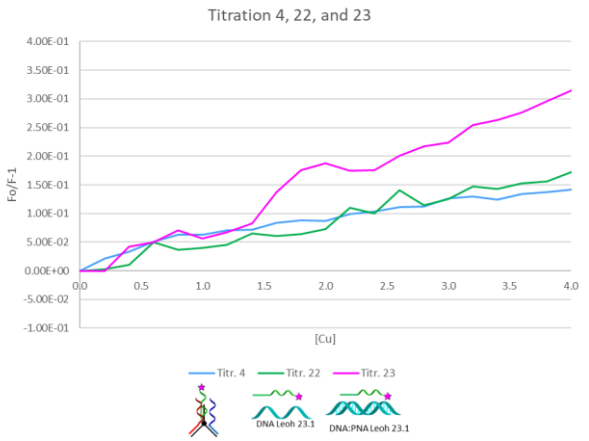



***Figure 2.41.*** *A comparison based on the F/Fo-1 values and equivalent of Cu(II) relative to titration 4, 22, and 23.*

## 2.4 Conclusions and future directions

The possibility to obtain heterotrimeric coiled-coil structure using a tripodal DNA template and PNA-peptide conjugates was demonstrated in this work. Binding of copper(II) ion was also demonstrated, though the nature of the binding site is not clarified. Careful controls of the behavior of single components and their combinations were in line with the obtainment of the designed assembly.

In PAGE experiments, qualitative and quantitative analysis of the assembly of heterotrimeric coiled-coil systems with the full matched trimeric DNA template or Cy3-labelled TP and the several PNA-peptides conjugates was performed, demonstrating the formation of the architecture containing all building blocks (TP and PNA-peptides conjugates). In addition, also the specificity of the PNA:DNA interactions that were used to arrange the heterotrimeric systems was demonstrated with the mismatched TP.

The formation of heterotrimeric coiled-coil systems was then studied by circular dichroism (CD) analysis for the assembly including PNA-11mer peptide conjugate in only one orientation. More in specific, CD spectra were collected for the assemblies formed by TP with one, two or three PNA-peptide conjugates and for negative controls including a mismatched. The results were consistent with the formation of a coiled-coil system. Detailed control experiments on the peptide components and on PNA-peptide conjugates allowed to focus on cases where spontaneous coiled-coil formation does not occur.

CD and Fluorescence titration experiments with copper(II) sulfate in a sample containing the trimeric template and all the three distinct PNA-11mer peptide conjugates suggested that copper(II) binding occurs. In particular, the fluorescence quenching of the assembly was found to be more pronounced when compared to its components. However, being that the fluorescence quenching obtained with the labelled peptide strand alone in presence of Cu(II) was higher than the overall assembly, suggests that when super-secondary structures are not present in solution the Cu(II) is more able to interact with the fluorophore. In addition, higher

fluorescence quenching is showed when the peptide domains are present on the PNAs respect when they are absent.

For the other five heterotrimeric coiled-coil architectures with different peptide domains the fluorescence emission quenching was not observed, although from PAGE experiments the formation of the assemblies was assessed. This behavior can be explained by higher distance between the coordination site for the Cu(II) and the fluorophore in presence of long peptide domains, or on the other side to the inability of PNA-peptides with shorter peptide domains to form coiled-coil structures despite the presence of the rigid trimeric scaffold, and the formation of minor number of electrostatic interactions with the 11mer in orientation A as peptide domains that cannot contribute to the stabilization of the coiled-coil formation.

In conclusion, in presence of the complex structure assembled with the trimeric template and the three PNA-11mer peptide conjugates in orientation B, some interactions with the Cu(II) were observed by fluorescence quenching experiments, but not sufficient to establish if this total complex presents a binding site as that necessary to behave as mini-metalloprotein. For this reason, other type of experiments and comparison should be performed to understand this important feature. In addition, increasing the concentration of the desired complex or optimizing the experimental setup, UV-visible measurements with the aim to see the band relative to the coordination of Cu(II) might become addressable, thus establishing if this heterotrimeric coiled-coil structure can be mimic the coordination site of plastocyanin (a small copper-binding protein with electron transfer properties).

## 2.5 Experimental section

### 2.5.1 Reagents and instrumentations

The reagents were bought from Merck, abcr, Carlo Erba, VWR, Fluka, TCI Europe, Link Technologies, Thermofisher, IDT, PolyOrg, LGC Biosearc, Chemgenes, BioShop and used without additional purification. Only for the solid phase synthesis, the DMF was purged with nitrogen flux to avoid the presence of dimethylamine and was dried over 4Å molecular sieves.

All the PNA, PNA-peptide and peptide sequences are synthesized by standard solid-phase manual synthesis or by automatic synthesizer Biotage Syro I. Purification was performed by HPLC (Agilent Technologies 1260 Infinity I) using a SepaChrom Vydamas® (C18, 5 µm, 300 Å, 10 x 250 mm) column. Gradient used were as follows: a) Gradient 1 (for 5 and 11mer peptides): 100% A for 1min, then from 0% to 50% B for 16min; b) Gradient 2 (for 19 and 22mer peptides): 100% A for 1min, then from 0% to 60% B for 19min; c) Gradient 3 (for PNA-5mer peptides): 100% A for 1min, then from 0% to 35% B for 22 min; d) Gradient 4 (for PNA-11mer peptides): 100% A for 1min, then from 0% to 45% B for 28min; e) Gradient 5 (for PNA-19mer peptides and labelled PNA-peptides): 100%A for 1 min, then from 0% to 50% B for 31 min; f) Gradient 6 (for PNAs): 100% A for 1 min, then from 0% to 30% B for 19 min. For all gradients: 4 mL/min as flow and A: water + 0.1% trifluoroacetic acid; B: acetonitrile + 0.1% trifluoroacetic acid, as eluents. Detector: UV set with wavelength: 260 nm.

After purification, the sequences were characterized by UPLC-MS (Waters Acquity Ultra Performance LC) using the following instrumental set-up: Waters Acquity ultra-performance LC E07SQD086W, with Waters SQ detector and ESI-interface equipped with Acquity UPLC BEH 300 (50 × 2.1 mm, 1.7 µm, C18). Chromatographic conditions: eluent A: water + 0.2% formic acid; eluent B: acetonitrile + 0.2% formic acid. Column temperature: 35 °C. Program: initial isocratic at 100% A (0.9 min), then linear gradient to 50% B (in 5.7 min). Final wash with 100% B for 1.2 min. Flow rate: 0.25 mL/min.

Some PNA-peptide conjugates and peptides sequences were characterized with Thermo LTQ Orbitrap XL detector and ESI-interface equipped with a Phenomenex Kinetex EVO (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic conditions: eluent A: water + 0.1% formic acid; eluent B: acetonitrile + 0.1% formic acid. Column temperature: 35 °C. Program: initial isocratic at 98% A (3 min), then linear gradient to 50% B (in 20 min) and after to 95% (in 1 min). Final wash with 98% B for 8 min. Flow rate: 0.20 mL/min. UV wavelength: 260 nm.

The concentration of the PNAs and PNA-peptide conjugates was calculated using Evolution 260 Bio UV-Visible spectrophotometer with Peltier thermos tatting accessories (Thermo Fisher Scientific SPE8W), following the UV-absorbance at 260 nm and assuming an additive contribution of all bases and any fluorophores. Instead, the concentration of peptide sequences was measured using the same instrument but following the UV-absorbance at 270 nm, typical of the 4-acetamidobenzoic acid (ABA).

The single scan CD measurements were performed with a JASCO J-1500 instrument (Serial No. D067261638) with PM-539 as detector and PTC-510 Peltier accessory for temperature control. Conditions: Measure range: 300-200 nm; Data pitch: 0.1 nm; Bandwidth: 1.00 nm; Start mode: immediately; Scanning mode: continuous; Scanning speed: 50 nm/min, Accumulations: 5. Sample and cell information: Volume sample: 1ml; Volume cell: 1.5ml; Path cell: 0.5cm. Processing: blank subtraction and smoothing correction (convolution width 5).

Experimental conditions to melting temperature CD measurements: Tm Setup: Start/end temp.: 20-90°C; Data interval: 0.1°C; Temp. Gradient: 0.5°C/min; Keep within +/- 0.1°C of the target temperature for 5 seconds; Control/monitor sensor: holder; End condition: return to start. Tm Parameters: Channels: CD and HT; CD scale: 20mdeg/0.05dOD; FL scale: 200mdeg/0.1 dOD; D.I.T.: 1 sec; Bandwidth: 1.00 nm; Wavelength: 222 nm. CD Parameters: Start/end: 260-200 nm; Data pitch: 0.1 nm; Scanning mode: continuous; Scanning speed: 50 nm/min; No. of accumulations: 9; concentration: 6 µmol/L.

The mean residue molar ellipticity was calculate for comparing the CD spectra of various constructs through the following equation:

$$\theta(deg \cdot cm^2 \cdot dmol^{-1}) = \frac{Ellipticity(mdeg) \times 10^5}{l \times c \times n}$$

where *Ellipticity* is the raw data by the instrument, *l*(cm) is the pathlength, *c*(μM) is the concentration of each sample and *n* is the number of amino acids in each structure.

The fluorescence quenching measurements were carried out on an Edinburgh FLS1000 fluorimeter. The instrument was equipped with Xenon lamp 750 nm blaze as source light path and PMT900 500 nm blaze as detector light path. The fluorescence quenching spectra were collected in a quartz cuvette using the following parameters: λexc: 552 nm, λemi: 578 nm, bandwidth: 1.0 nm, scan: 555-700 nm, step: 1 nm, dwell time: 0.3s and number of scans: 1.

BioAutomation MerMade MM6 DNA synthesizer was used to synthetize the trimeric DNA scaffolds through a using standard automated solid-phase oligonucleotide synthesis on a 1 μmol as scale. Polyacrylamide gel electrophoresis (PAGE) was employed for the purification of raw materials (20x20cm vertical Hoefer 600 electrophoresis equipment). Bio-Rad Laboratories ChemiDocTM MP System was used for the collection of gel images. The concentration of branching DNA trimers was calculated using NanoDrop Lite Spectrophotometer following the UV-absorbance at 260 nm and assuming an additive contribution of all the nucleobases and the fluorophore. The scaffolds were characterized by LC-ESI-MS using a Dionex Ultimate 3000 linked to a Bruker MaXis Impact™ QTOF.

Polyacrylamide gel electrophoresis (PAGE) was employed also to study the assembly of heterotrimeric coiled-coil systems. The thermal annealing was executed with an Eppendorf Mastercycler® 96 well thermocycler.

### 2.5.2 Synthesis of peptides and PNA-peptide conjugates

All the PNA-peptide conjugates, PNAs, and modified TRI-family peptide sequences were synthetized with standard solid-phase manual synthesis, using 9-

fluorenylmethoxycarbonyl (Fmoc) strategy while the TRI-family peptide sequences were obtained, using the same strategy, but with automatic synthesis through automatic synthesizer Syro I. The PNA monomers employed for the Fmoc strategy are Fmoc/Bhoc protected. The Fmoc amino acids employed for the peptide synthesis are Fomc-Ala-OH, Fmoc-Cys(tBu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Glu(OtBu)-OH, Fomc-His(trt)-OH, Fmoc-Leu-OH, Fmoc-Lys(Mtt)-OH, Fmoc-Lys(Dde)-OH, Fmoc-Lys(Boc)-OH, and Fmoc-Met(O)-OH.

For the PNA-peptides and PNA sequences, the H-Rink Amide Chemmatrix resin was loaded with Fmoc-Gly-OH to obtain 0.2 mmol/gr as theoretical loading. Instead, Fmoc-Rink Amide AM-Champion resin was used for all the peptide sequences with a loading of 0.64 mmol/g measured by UV-Visible measurements.

The principal steps to load the H-Rink Amide Chemmatrix resin were: a) swelling in DCM (1x1h), b) DMF dry wash, c) coupling (1x5h, activation for 10', using 1 eq of Fmoc-Gly-OH, 10 eq of DIC and 10 eq of DhBtOH in DMF dry as activation solution), d) DMF wash, e) capping (2x15', DMF:$Ac_2O$, 1:1), e) DMF wash and f) DCM wash. To measure the loading of the both the resins a UV/Visible technique was used, following the absorbance of divenzofulvene (DBF) formation after deprotection (30', Piperidine:DMF, 1:4) at 290 nm. The values at 290 nm and 400 nm (used as baseline absorbance) were used. The loading of the resins was calculated dividing the value obtained from the difference between the absorbance at 290 nm and at 400nm with a value obtained by multiplication of a conversion factor (1.65 g/mmol*mg) with the weighed resin mass.

The synthesis of the three PNA sequences was performed in a 30 µmol as scale to have the general formula: Fmoc-PNA-Lys(Mtt)-AEEA-Lys(Dde)-H-RinkAmide Chemmatrix resin. After loading the first tract (Lys(Mtt)-AEEA-Lys(Dde)-H RinkAmide Chemmatrix), three portions of the resin were separated to complete the three different PNA-peptide conjugates, each in a 4 µmol as a scale. The PNA part was grown after deprotection of Fmoc group in this tract. Then, after completion of the PNA sequence, capping the PNA part and removal of the Dde group on the first lysine, the peptide moiety was synthetized.

The modified TRI family peptides were synthesized in a 10 µmol as scale, while the TRI family peptides were synthesized in a 25 µmol as scale.

The principal steps and solutions for the standard solid-phase manual synthesis were: a) swelling in DCM (1x30'), b) deprotection (2x8', Piperidine:DMF, 1:4), c) DCM wash, d) DMF dry wash, e) Kaiser test (1', positive), f) coupling (1x30', activation for 2', using 5 eq of PNA monomer or amino acid residue, 4.9 eq of HBTU as coupling reagent and 10 eq of DIPEA in DMF dry as activation solution), g) DMF wash, h) Kaiser test (1', negative), i) capping (2x1', $Ac_2O$:DIPEA:DMF dry, 5:6:89), l) DMF wash, m) 5% DIPEA wash (2x2', DIPEA:DMF, 5:95), n) DMF wash and o) DCM wash.

The solutions used for the Kaiser test were the following: K1) 1g of ninidrin in 10 mL absolute ethanol, K2) 8g phenol in 2 mL of absolute ethanol, and K3) 0.2 mL aqueous solution 0.001M of KCN diluited at 10 mL with pyridine. To perform this test 2-3 drops of each solution were added in a tube where few grains of the resin were previously included with a capillary. After that, the tube was put in an oil bath at 100°C for 1 minute. The test is positive if the resin becomes blue.

For the solid-phase automatic synthesis of peptide sequences with Syro I the protocol does not present the activation step. In fact, the Fmoc protected amino acids, HBTU and DIPEA were added in this order in the reactor containing the resin and then mixed. The coupling step was performed once for 40', using 4 eq of monomer and coupling reagent, and 10 eq of DIPEA). In addition, the amino acids were dissolved in DMF dry to obtain a 0.5 mol/l as concentration. The concentration of HBTU in DMF dry was 0.43 mol/l, while for the DIPEA in NMP dry it was 2.0 mol/l. Unlike the manual protocol, 40% of piperidine as deprotection solution was used for the automatic peptide synthesis and, moreover, the 5% DIPEA wash, and the capping step were not performed.

Before starting the synthesis of the peptide domains, each PNA sequence underwent the following protocol: a) Fmoc deprotection (2x8', Piperidine:DMF, 1:4), b) DMF wash, c) capping (2x1', $Ac_2O$:DIPEA:DMF dry, 5:6:89), d) DMF wash, e) 5% DIPEA wash (2x2', DIPEA:DMF, 5:95), f) DMF wash, g) DCM wash, h) Dde deprotection, i) DMF wash, and l) DCM wash. The solution for the Dde deprotection

was constituted by 250 mg of hydroxylamine hydrochloride, 184 mg of imidazole, 1 mL of NMP dry, and 200 µl of DMF dry for a 5 µmol scale. The deprotection was repeated three times for 1 hour. After this synthetic steps, 1 µmol of each PNA-peptide conjugates was subjected at the subsequent protocol: a) Fmoc deprotection (2x8', Piperidine:DMF, 1:4), b) DMF wash, c) DCM wash, d) Mtt deprotection, e) DCM wash, f) DMF wash, g) capping (2x1', $Ac_2O$:DIPEA:DMF dry, 5:6:89), h) DMF wash, i) 5% DIPEA wash (2x2', DIPEA:DMF, 5:95), l) DMF wash, and m) DCM wash. The Mtt deprotection solution was 0.5% 1-hydroxybenzotriazole (BtOH) in hexafluoroisopropanol (HFIP):DCM (1:1) and the deprotection reaction was repeated three times for 3'.

For the PNA-peptide conjugates functionalized with 5(6)-carboxytetramethylrhodamine (5(6)-TAMRA) at the N-terminal moiety, the coupling was performed in 1 µmol of each target sequence, using 10 eq of 5(6)-TAMRA, 9.9 eq of HBTU and 20 eq of DIPEA (2' of activation and 1h of coupling). These sequences underwent the following synthetic steps: a) Fmoc deprotection (2x8', Piperidine:DMF, 1:4), b) DCM wash, c) DMF wash, d) coupling with 5(6)-TAMRA, f) DMF wash, g) DCM wash, h) Mtt deprotection, i) DCM wash, l) DMF wash, m) capping (2x1', $Ac_2O$:DIPEA:DMF dry, 5:6:89), n) DMF wash, o) 5% DIPEA wash (2x2', DIPEA:DMF, 5:95), p) DMF wash, and q) DCM wash.

Furthermore, the coupling between the TRI family peptide sequences or modified TRI family peptide oligomers with 4-acetamidobenzoic acid (ABA) at the N-terminal moiety was carried out in 5 µmol of each peptide sequence employing 5 eq of ABA, 4.9 eq of HBTU and 10 eq of DIPEA (2' of activation and 1h of coupling). The synthetic protocol used was constituted by a) Fmoc deprotection (2x8', Piperidine:DMF, 1:4), b) DCM wash, c) DMF wash, d) Coupling with ABA, e) DMF wash, and f) DCM wash.

For the cleavage step several cocktail solutions were employed to overcome synthetic problems and two cycles of 1h were performed to ensure a complete cleavage. After the first cycle the solution was filtered and collected in a falcon. Then, the resin was washed with only TFA to ensure a complete collection of the product, and the same thing was repeated at the end of the second cycle.

For the PNA-peptide conjugates with 5mer and 11mer peptide domains, modified TRI family peptides without methionine and cysteine residues in the sequences, and TRI family peptides a cleavage solution of TFA:m-cresol (9:1) was employed. Instead, for all the PNA-peptide conjugates and modified TRI family peptides with Met(O) residue, a solution of TFA:m-cresol (9:1) + 0.1M of tetrabutylammonium bromide ($Bu_4NBr$) in TFA was used. In this case the solution of 0.1M $Bu_4NBr$ in TFA was added in the falcon after the cleavage and before the precipitation, leaving the reaction to proceed for 30'. Furthermore, for the PNA-peptide conjugates and modified TRI family peptides with Cys(tBu) residue two cleavage steps were necessary. In the first step, a solution of TFA/m-cresol (9:1) was used to cleave the oligomers from the solid support. Subsequently, the oligomers were precipitated by adding diethyl ether and cooling the resulting solution for 3 hours at -18°C; the precipitate was then dried. A second treatment with a cleavage solution constituted by TFA/TFMSA/m-cresol/thioanisole (6:2:1:1) was then employed to recover further product from the resin. Instead, for the PNA-peptide conjugates and modified TRI family peptides containing a Cys(Trt) residue, to overcome at the cysteine oxidation, TFA:m-cresol:EDT (80:10:10) was used as cleavage solution. The last cleavage cocktail used was TFA:TIS:EDT + 50 eq of cysteamine-HCl for the PNA-peptide conjugates with 19mer of peptide domains without methionine and cysteine residues in the sequences to prevent the Bhoc alkylation.

After the reduction of TFA volume with nitrogen flow, all the oligomers were precipitated in diethyl ether in freezer for at least 2h. After removing diethyl ether, and drying the product, the latter was dissolved in double distilled water.

All the PNA-peptide conjugates and all the peptide sequences were purified by HPLC and characterized by UPLC-MS and UV-Visible techniques. For the PNA-peptide conjugates, the absorbance was measured at 260 nm and at 400 nm (baseline contribution). In addition, the molar absorptivity (ε) specific for each PNA was calculated considering 13700 $M^{-1}cm^{-1}$ for adenine monomer, 6600 $M^{-1}cm^{-1}$ for cytosine, 11700 $M^{-1}cm^{-1}$ for guanine and 8600 $M^{-1}cm^{-1}$ for thymine as molar absorptivity values of each single nucleobase, assuming an additive model. For the

ABA labelled peptides a value of 17989 $M^{-1}cm^{-1}$ as molar absorptivity at 270 nm was considered. The concentration of each product was calculated using a Lambert-Beer law, and from this the reaction yield was calculated using the overall volume of the solution.

### 2.5.3 Synthesis of branching DNA trimers

The three branching DNA trimers were synthetized through solid phase phosphoramidite oligo synthesis, using 1 µmol Universal 1000Å LCAA-CPG as a solid support. The first arm, in a manner that advanced from the 3' end to the 5' end, was carried out using a standard phosphoramidites (5'-DMT, 3'-CED). Subsequently, an asymmetrical branching CED phosphoramidite was attached to the DNA chain, and the synthesis of the second branch proceeded in the 5' to 3' direction and was carried out with using a reverse phosphoramidites (3'-DMT, 5'-CED). after removing the DMT group. The 3' end of the second arm was treated with a capping solution, and the levulinyl protecting group was eliminated using a 0.5M hydrazine monohydrate solution in a pyridine/acetic acid (1:1) mixture for 10 minutes. This allowed for the synthesis of the third arm using a reverse phosphoramidite. Subsequently, the strand was detached from the CPG support in ammonium hydroxide, leaving the solution to react overnight at 65°C in water bath. After that, the cleavage solution was removed with SpeedVac, and the branching DNA trimers were purified using 15% denaturing PAGE analysis. Finally, a second step of purification was performed to separate the urea with a size exclusion column (SEC), in specific a Sephadex G-100 DNA Grade.

The general steps of DNA automatic synthesis and the solution composition employed are the following: a) detritylation of the support-bound 3'-nucleoside (3% trichloroacetic acid in DCM), b) activation and coupling (0.05M phoshoramidite monomer and 0.25M 5-ethylthio tetrazole in ACN), c) capping A ($Ac_2O$:Lutidine:THF, 1:1:8), d) capping B (16% w/v N-methyl imidazole in ACN), e) oxidation (0.02M iodine in $H_2O$:pyridine:THF, 10:20.70), and f) detritylation at the end of the DNA synthesis (3% trichloroacetic acid in DCM). In addition, the coupling between the last

monomer of the first arm and the asymmetrical branching small molecule was performed in Glovebox, using 0.1M of the small molecule in ACN and 0.25M 5-ethylthio tetrazole in ACN as activator.

### 2.5.4 Polyacrylamide gel electrophoresis (PAGE) experiments

5% native PAGE analysis were carried out to study the assembly of heterotrimeric coiled-coil systems. Both TBE and TaMg buffer were used. In specific TAMg buffer was formed by 45 mM Tris and 12.5 mM $MgCl_2 \bullet 6H_2O$ with a pH of 8.0 adjusted with glacial acetic acid. Instead, the TBE buffer was constituted by 90 mM Tris, 90 mM boric acid, and 2 mM EDTA with a pH of 6.5. All the samples were prepared in water and submitted to thermal annealing with a temperature ramp from 95°C to 4°C in 1h.

### 2.5.5 Peptides, PNA-peptide conjugates and PNAs characterizations

Entry 1 (Figure 2.A19 in Appendix)
CS02-117 A1: ABA-lkaleeklkaleeklkaleekl-$NH_2$
yield: 15%; $t_r$: 5.76 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2727.23 [M], observed (from ESI-MS deconvolution):2726.0; ESI-MS: m/z observed (calcd): 910.0 (910.1) $[M+3H]^{3+}$, 682.8 (682.8) $[M+4H]^{4+}$, 546.5 (546.4) $[M+5H]^{5+}$.

Entry 2 (Figure 2.A20 in Appendix)
CS02-117 A2: ABA-lkeelaklkeelaklkeelakl-$NH_2$
yield: 20%; $t_r$: 5.60 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2727.23 [M], observed (from ESI-MS deconvolution):2726.0; ESI-MS: m/z observed (calcd): 910.0 (910.1) $[M+3H]^{3+}$, 682.8 (682.8) $[M+4H]^{4+}$, 546.4 (546.4) $[M+5H]^{5+}$.

Entry 3 (Figure 2.A21 in Appendix)
CS02-117 A3: ABA-leeklkaleeklkaleekl-$NH_2$
yield: 18%; $t_r$: 5.38 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2414.82 [M], observed (from ESI-MS deconvolution):2414.0; ESI-MS: m/z observed (calcd): 805.7 (805.9) $[M+3H]^{3+}$, 604.7 (604.7) $[M+4H]^{4+}$, 484.0 (483.9) $[M+5H]^{5+}$.

Entry 4 (Figure 2.A22 in Appendix)
CS02-117 A4: ABA-lkeelaklkeelaklkeel-$NH_2$

yield: 6%; $t_r$: 3.80 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2414.82 [M], observed (from ESI-MS deconvolution):2414.8; ESI-MS: m/z observed (calcd): 805.9 (805.9) $[M+3H]^{3+}$, 604.7 (604.7) $[M+4H]^{4+}$, 484.0 (483.9) $[M+5H]^{5+}$.

Entry 5 (Figure 2.A23 in Appendix)
CS02-117 C1: ABA-lkaleeklkal-$NH_2$
yield: 25%; $t_r$: 3.78 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 1415.72 [M]; ESI-MS: m/z observed (calcd): 708.3 (708.9) $[M+2H]^{2+}$, 472.8 (472.9) $[M+3H]^{3+}$.

Entry 6 (Figure 2.A24 in Appendix)
CS02-117 C2: ABA-laklkeelakl-$NH_2$
yield: 41%; $t_r$: 3.75 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 1415.72 [M]; ESI-MS: m/z observed (calcd): 708.6 (708.9) $[M+2H]^{2+}$, 472.8 (472.9) $[M+3H]^{3+}$.

Entry 7 (Figure 2.A25 in Appendix)
CS02-117 C3: ABA-leekl-$NH_2$
yield: 25%; $t_r$: 4.55 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 790.90 [M], observed: 791.5.

Entry 8 (Figure 2.A26 in Appendix)
CS02-117 C4: ABA-lkeel-$NH_2$
yield: 10%; $t_r$: 4.50 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 790.90 [M], observed: 791.5.

Entry 9 (Figure 2.A27 in Appendix)
MaCa01-04 A: ABA-meekh-$NH_2$
yield: 21%; $t_r$: 3.30 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 832.92 [M], observed:833.5; ESI-MS: m/z observed (calcd): 417.4 (417.5) $[M+2H]^{2+}$.

Entry 10 (Figure 2.A28 in Appendix)
MaCa01-04 B: ABA-leekc-$NH_2$
yield: 15%; $t_r$: 4.07 min; ε (270nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 780.88 [M], observed: 781.6.

Entry 11 (Figure 2.A29 in Appendix)
MaCa01-04 C: ABA-leekh-$NH_2$
yield: 23%; $t_r$: 3.61 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 814.88 [M], observed: 815.7; ESI-MS: m/z observed (calcd): 408.4 (408.4) $[M+2H]^{2+}$.

Entry 12 (Figure 2.A30 in Appendix)
MaCa01-05 A: ABA-lkameekhkal-NH2
yield: 13%; $t_r$: 3.87 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 1457.74 [M], observed (from ESI-MS deconvolution): 1457.2; ESI-MS: m/z observed (calcd): 729.8 (729.9) $[M+2H]^{2+}$, 486.9 (486.9) $[M+3H]^{3+}$, 365.6 (365.4) $[M+4H]^{4+}$.

Entry 13 (Figure 2.A31 in Appendix)
MaCa01-05 B: ABA-lkaleekckal-$NH_2$
yield: 19%; $t_r$: 4.92 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 1405.70 [M], observed (from ESI-MS deconvolution): 1406.0; ESI-MS: m/z observed (calcd): 703.7 (703.8) $[M+2H]^{2+}$, 469.4 (469.6) $[M+3H]^{3+}$.

Entry 14 (Figure 2.A32 in Appendix)
MaCa01-05 C: ABA-lkaleekhkal-$NH_2$
yield: 40%; $t_r$: 4.09 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 1439.70 [M], observed (from ESI-MS deconvolution): 1440.6; ESI-MS: m/z observed (calcd): 720.7 (720.8) $[M+2H]^{2+}$, 481.0 (480.9) $[M+3H]^{3+}$, 361.1 (360.9) $[M+4H]^{4+}$.

Entry 15 (Figure 2.A33 in Appendix)
MaCa01-06 A: ABA-leeklkameekhkaleekl-$NH_2$
yield: 18%; $t_r$: 4.17 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2456.84 [M], observed (from ESI-MS deconvolution): 2457.0; ESI-MS: m/z observed (calcd): 820.1 (819.9) $[M+3H]^{3+}$, 615.3 (615.2) $[M+4H]^{4+}$, 492.4 (492.4) $[M+5H]^{5+}$.

Entry 16 (Figure 2.A34 in Appendix)
MaCa01-06 B: ABA-leeklkaleekckaleekl-$NH_2$
yield: 20%; $t_r$: 5.23 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2404.80 [M], observed (from ESI-MS deconvolution): 2404.0; ESI-MS: m/z observed (calcd): 802.8 (802.6) $[M+3H]^{3+}$, 602.3 (602.2) $[M+4H]^{4+}$, 482.9 (481.9) $[M+5H]^{5+}$.

Entry 17 (Figure 2.A35 in Appendix)
MaCa01-06 C: ABA-leeklkaleekhkaleekl-$NH_2$
yield: 42%; $t_r$: 4.31 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2438.80 [M], observed (from ESI-MS deconvolution): 2438.2; ESI-MS: m/z observed (calcd): 814.2 (813.9) $[M+3H]^{3+}$, 610.8 (610.7) $[M+4H]^{4+}$, 488.8 (488.8) $[M+5H]^{5+}$.

Entry 18 (Figure 2.A36 in Appendix)
MaCa01-07 A: ABA-hkeem-$NH_2$
yield: 26%; $t_r$: 2.88 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 832.92 [M], observed:833.7; ESI-MS: m/z observed (calcd): 417.5 (417.5) $[M+2H]^{2+}$.

Entry 19 (Figure 2.A37 in Appendix)
MaCa01-07 B: ABA-ckeel-$NH_2$
yield: 64%; $t_r$: 3.80 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 780.88 [M], observed: 780.5.

Entry 20 (Figure 2.A38 in Appendix)
MaCa01-07 C: ABA-hkeel-$NH_2$
yield: 15%; $t_r$: 3.07 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 814.88 [M], observed: 815.5; ESI-MS: m/z observed (calcd): 408.3 (408.4) $[M+2H]^{2+}$.

Entry 21 (Figure 2.A39 in Appendix)
MaCa01-08 A: ABA-lakhkeemakl-$NH_2$
yield: 23%; $t_r$: 3.59 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 1457.74 [M], observed (from ESI-MS deconvolution): 1458.0; ESI-MS: m/z observed (calcd): 729.4 (729.9) $[M+2H]^{2+}$, 486.8 (486.9) $[M+3H]^{3+}$, 365.3 (365.4) $[M+4H]^{4+}$.

Entry 22 (Figure 2.A40 in Appendix)
MaCa01-08 B: ABA-lakckeelakl-$NH_2$
yield: 33%; $t_r$: 4.82 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 1405.70 [M], observed (from ESI-MS deconvolution): 1406.0; ESI-MS: m/z observed (calcd): 703.6 (703.8) $[M+2H]^{2+}$, 469.6 (469.6) $[M+3H]^{3+}$.

Entry 23 (Figure 2.A413 in Appendix)
MaCa01-08 C: ABA-lakhkeelakl-$NH_2$
yield: 6%; $t_r$: 3.71 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 1439.70 [M], observed (from ESI-MS deconvolution): 1439.4; ESI-MS: m/z observed (calcd): 720.6 (720.8) $[M+2H]^{2+}$, 480.9 (480.9) $[M+3H]^{3+}$, 361.0 (360.9) $[M+4H]^{4+}$.

Entry 24 (Figure 2.A42 in Appendix)
MaCa01-09 A: ABA-lkeelakhkeemaklkeel-$NH_2$
yield: 10%; $t_r$: 4.07 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2456.84 [M], observed (from ESI-MS deconvolution): 2457.0; ESI-MS: m/z observed (calcd): 1229.4 (1229.4) $[M+2H]^{2+}$, 819.8 (819.9) $[M+3H]^{3+}$, 615.1 (615.2) $[M+4H]^{4+}$, 492.4 (492.4) $[M+5H]^{5+}$.

Entry 25 (Figure 2.A43 in Appendix)
MaCa01-09 B: ABA-lkeelakckeelaklkeel-$NH_2$
yield: 19%; $t_r$: 4.27 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2404.80 [M], observed (from ESI-MS deconvolution): 2404.0; ESI-MS: m/z observed (calcd): 1203.0 (1203.4) $[M+2H]^{2+}$, 802.4 (802.6) $[M+3H]^{3+}$, 602.0 (602.2) $[M+4H]^{4+}$, 481.9 (481.9) $[M+5H]^{5+}$.

Entry 26 (Figure 2.A44 in Appendix)
MaCa01-09 C: ABA-leeklkaleekhkaleekl-$NH_2$
yield: 12%; $t_r$: 4.19 min; ε (270 nm): 17989 $M^{-1}cm^{-1}$; MW calculated: 2438.80 [M], observed (from ESI-MS deconvolution): 2437.8; ESI-MS: m/z observed (calcd): 1220.4 (1220.4) $[M+2H]^{2+}$, 813.8 (813.9) $[M+3H]^{3+}$, 610.6 (610.7) $[M+4H]^{4+}$, 488.7 (488.8) $[M+5H]^{5+}$.

Entry 27 (Figure 2.A45 in Appendix)
CS02-68A: Ac-GGGCATGATCT-Lys(Ac)-AEEA-Lys(Ac-meekh→)-$NH_2$
yield: 15%; $t_r$: 3.05 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 4216.28 [M], observed (from ESI-MS deconvolution): 4216.8; ESI-MS: m/z observed (calcd): 1055.1 (1055.1) $[M+4H]^{4+}$, 844.2 (844.3) $[M+5H]^{5+}$, 703.8 (703.7) $[M+6H]^{6+}$, 603.2 (603.3) $[M+7H]^{7+}$.

Entry 28 (Figure 2.A46 in Appendix)
CS02-68B: Ac-TACCGTCGAGT-Lys(Ac)-AEEA-Lys(Ac-leekc→)-$NH_2$
yield: 13%; $t_r$: 3.42 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 4124.22 [M], observed (from ESI-MS deconvolution): 4123.0; ESI-MS: m/z observed (calcd): 1375.8 (1375.7) $[M+3H]^{3+}$, 1032.1 (1032.1) $[M+4H]^{4+}$, 825.9 (825.8) $[M+5H]^{5+}$, 688.4 (688.4) $[M+6H]^{6+}$.

Entry 29 (Figure 2.A47 in Appendix)
CS02-68C: Ac-GTGACTCAGAT-Lys(Ac)-AEEA-Lys(Ac-leekh→)-$NH_2$
yield: 18%; $t_r$: 3.17 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; MW calculated: 4182.24 [M], observed (from ESI-MS deconvolution): 4182.8; ESI-MS: m/z observed (calcd): 1395.3 (1395.1) $[M+3H]^{3+}$, 1046.5 (1046.6) $[M+4H]^{4+}$, 837.5 (837.4) $[M+5H]^{5+}$, 698.1 (698.0) $[M+6H]^{6+}$, 598.5 (598.5) $[M+7H]^{7+}$, 523.9 (523.8) $[M+8H]^{8+}$.

Entry 30 (Figure 2.A48 in Appendix)
CS02-76A: Ac-GGGCATGATCT-Lys(Ac)-AEEA-Lys(Ac-lkameekhkal→)-$NH_2$
yield: 12%; $t_r$: 3.37 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 4841.10 [M], observed (from ESI-MS deconvolution): 4841.6; ESI-MS: m/z observed (calcd): 1211.1 (1211.3) $[M+4H]^{4+}$, 969.3 (969.2) $[M+5H]^{5+}$, 807.9 (807.8) $[M+6H]^{6+}$, 692.7 (692.6) $[M+7H]^{7+}$, 606.2 (606.1) $[M+8H]^{8+}$, 538.9 (538.9) $[M+9H]^{9+}$.

Entry 31 (Figure 2.A49 in Appendix)
CS02-76B: Ac-TACCGTCGAGT-Lys(Ac)-AEEA-Lys(Ac-lkaleekckal→)-$NH_2$
yield: 11%; $t_r$: 4.21 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 4749.0 [M], observed (from ESI-MS deconvolution): 4748.6; ESI-MS: m/z observed (calcd): 1188.3 (1188.3) $[M+4H]^{4+}$, 950.9 (950.8) $[M+5H]^{5+}$, 792.5 (792.5) $[M+6H]^{6+}$, 679.5 (679.4) $[M+7H]^{7+}$, 594.8 (594.6) $[M+8H]^{8+}$, 528.8 (528.7) $[M+9H]^{9+}$.

Entry 32 (Figure 2.A50 in Appendix)
CS02-76C: Ac-GTGACTCAGAT- Lys(Ac)-AEEA-Lys(Ac-lkaleekhkal→)-$NH_2$
yield: 16%; $t_r$: 3.49 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; MW calculated: 4807.06 [M], observed (from ESI-MS deconvolution): 4807.0; ESI-MS: m/z observed (calcd): 1202.6 (1202.8) $[M+4H]^{4+}$, 962.4 (962.4) $[M+5H]^{5+}$, 802.3 (802.2) $[M+6H]^{6+}$, 687.8 (687.7) $[M+7H]^{7+}$, 601.9 (601.9) $[M+8H]^{8+}$, 535.2 (535.1) $[M+9H]^{9+}$.

Entry 33 (Figure 2.A51 in Appendix)
CS02-90A: Ac-GGGCATGATCT-Lys(Ac)-AEEA-Lys(Ac-leeklkameekhkaleekl→)-$NH_2$
yield: 8%; $t_r$: 3.83 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 5840.2 [M], observed (from ESI-MS deconvolution): 5840.4; ESI-MS: m/z observed (calcd): 1169.2 (1169.0) $[M+5H]^{5+}$, 974.3 (974.4) $[M+6H]^{6+}$, 835.3 (635.3) $[M+7H]^{7+}$, 731.1 (731.0) $[M+8H]^{8+}$, 650.0 (649.9) $[M+9H]^{9+}$, 585.1 (585.0) $[M+10H]^{10+}$.

Entry 34 (Figure 2.A52 in Appendix)
CS02-90B: Ac-TACCGTCGAGT-Lys(Ac)-AEEA-Lys(Ac-leeklkaleekckaleekl→)-$NH_2$
yield: 4%; $t_r$: 4.79 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 5748.14 [M], observed (from ESI-MS deconvolution): 5747.8; ESI-MS: m/z observed (calcd): 1150.5 (1150.6) $[M+5H]^{5+}$, 959.0 (959.0) $[M+6H]^{6+}$, 822.0 (822.2) $[M+7H]^{7+}$, 719.5 (719.5) $[M+8H]^{8+}$, 639.7 (639.7) $[M+9H]^{9+}$, 575.8 (575.8) $[M+10H]^{10+}$, 523.7 (523.5) $[M+11H]^{11+}$.

Entry 35 (Figure 2.A53 in Appendix)
CS02-90C: Ac-GTGACTCAGAT- Lys(Ac)-AEEA-Lys(Ac-leeklkaleekhkaleekl→)-$NH_2$
yield: 7%; $t_r$: 3.95 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; MW calculated: 5806.16 [M], observed (from ESI-MS deconvolution): 5806.8; ESI-MS: m/z observed (calcd): 1162.3 (1162.2) $[M+5H]^{5+}$, 968.7 (968.7) $[M+6H]^{6+}$, 830.5 (830.5) $[M+7H]^{7+}$, 727.0 (726.8) $[M+8H]^{8+}$, 646.2 (646.1) $[M+9H]^{9+}$, 581.6 (581.6) $[M+10H]^{10+}$, 529.0 (528.8) $[M+11H]^{11+}$.

Entry 36 (Figure 2.A54 in Appendix)
CS02-80A: Ac-GGGCATGATCT-Lys(Ac)-AEEA-Lys(Ac-hkeem→)-$NH_2$
yield: 7%; $t_r$: 3.08 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 4216.28 [M], observed (from ESI-MS deconvolution): 4216.0; ESI-MS: m/z observed (calcd): 1055.2 (1055.1) $[M+4H]^{4+}$, 844.3 (844.3) $[M+5H]^{5+}$, 703.9 (703.7) $[M+6H]^{6+}$, 603.6 (603.3) $[M+7H]^{7+}$.

Entry 37 (Figure 2.A55 in Appendix)
CS02-80B: Ac-TACCGTCGAGT-Lys(Ac)-AEEA-Lys(Ac-ckeel→)-$NH_2$
yield: 10%; $t_r$: 3.39 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 4124.22 [M], observed (from ESI-MS deconvolution): 4123.8; ESI-MS: m/z observed (calcd): 1375.3 (1375.7) $[M+3H]^{3+}$, 1031.4 (1032.1) $[M+4H]^{4+}$, 825.7 (825.8) $[M+5H]^{5+}$, 688.3 (688.4) $[M+6H]^{6+}$.

Entry 38 (Figure 2.A56 in Appendix)
CS02-80C: Ac-GTGACTCAGAT- Lys(Ac)-AEEA-Lys(Ac-hkeel→)-$NH_2$
yield: 10%; $t_r$: 3.20 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; MW calculated: 4182.24 [M], observed (from ESI-MS deconvolution): 4182.2; ESI-MS: m/z observed (calcd): 1395.1 (1395.1) $[M+3H]^{3+}$, 1046.6 (1046.6) $[M+4H]^{4+}$, 837.4 (837.4) $[M+5H]^{5+}$, 698.1 (698.0) $[M+6H]^{6+}$, 598.6 (598.5) $[M+7H]^{7+}$, 523.7 (523.8) $[M+8H]^{8+}$.

Entry 39 (Figure 2.A57 in Appendix)
CS02-85A: Ac-GGGCATGATCT-Lys(Ac)-AEEA-Lys(Ac- lakhkeemakl→)-$NH_2$
yield: 10%; $t_r$: 3.32 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 4841.10 [M], observed (from ESI-MS deconvolution): 4841.8; ESI-MS: m/z observed (calcd): 1211.4 (1211.3) $[M+4H]^{4+}$, 969.2 (969.2) $[M+5H]^{5+}$, 807.8 (807.8) $[M+6H]^{6+}$, 692.5 (692.6) $[M+7H]^{7+}$, 606.4 (606.1) $[M+8H]^{8+}$, 538.8 (538.9) $[M+9H]^{9+}$.

Entry 40 (Figure 2.A58 in Appendix)
CS02-85B: Ac-TACCGTCGAGT-Lys(Ac)-AEEA-Lys(Ac- lakckeelakl→)-$NH_2$
yield: 14%; $t_r$: 4.11 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 4749.0 [M], observed (from ESI-MS deconvolution): 4749.2; ESI-MS: m/z observed (calcd): 1188.1 (1188.3) $[M+4H]^{4+}$, 950.9 (950.8) $[M+5H]^{5+}$, 792.6 (792.5) $[M+6H]^{6+}$, 679.5 (679.4) $[M+7H]^{7+}$, 594.7 (594.6) $[M+8H]^{8+}$, 528.9 (528.7) $[M+9H]^{9+}$.

Entry 41 (Figure 2.A59 in Appendix)
CS02-85C: Ac-GTGACTCAGAT- Lys(Ac)-AEEA-Lys(Ac- lakhkeelakl→)-$NH_2$
yield: 22%; $t_r$: 3.44 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; MW calculated: 4807.06 [M], observed (from ESI-MS deconvolution): 4807.4; ESI-MS: m/z observed (calcd): 1202.9 (1202.8) $[M+4H]^{4+}$, 962.5 (962.4) $[M+5H]^{5+}$, 802.2 (802.2) $[M+6H]^{6+}$, 687.7 (687.7) $[M+7H]^{7+}$, 601.9 (601.9) $[M+8H]^{8+}$, 535.1 (535.1) $[M+9H]^{9+}$.

Entry 42 (Figure 2.A60 in Appendix)
NM03-18: Ac-GGGCATGATCT-Lys(Ac)-AEEA-Lys(Ac-lkeelakhkeemaklkeel→)-$NH_2$
yield: 5%; $t_r$: 3.85 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 5840.2 [M], observed (from ESI-MS deconvolution): 5840.8; ESI-MS: m/z observed (calcd): 1169.2 (1169.0) $[M+5H]^{5+}$, 974.5 (974.4) $[M+6H]^{6+}$, 835.6 (635.3) $[M+7H]^{7+}$, 731.2 (731.0) $[M+8H]^{8+}$, 650.2 (649.9) $[M+9H]^{9+}$, 585.0 (585.0) $[M+10H]^{10+}$, 532.1 (531.9) $[M+11H]^{11+}$.

Entry 43 (Figure 2.A61 in Appendix)
CS02-103B: Ac-TACCGTCGAGT-Lys(Ac)-AEEA-Lys(Ac- lkeelakckeelaklkeel→)-$NH_2$
yield: 3%; $t_r$: 4.77 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 5748.14 [M], observed (from ESI-MS deconvolution): 5748.0; ESI-MS: m/z observed (calcd): 959.3 (959.0) $[M+6H]^{6+}$, 822.1 (822.2) $[M+7H]^{7+}$, 719.5 (719.5) $[M+8H]^{8+}$, 639.8 (639.7) $[M+9H]^{9+}$.

Entry 44 (Figure 2.A62 in Appendix)
CS02-103C: Ac-GTGACTCAGAT-Lys(Ac)-AEEA-Lys(Ac- lkeelakhkeelaklkeel→)-$NH_2$
yield: 7%; $t_r$: 3.99 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; MW calculated: 5806.16 [M], observed (from ESI-MS deconvolution): 5807.0; ESI-MS: m/z observed (calcd): 1162.3 (1162.2) $[M+5H]^{5+}$, 968.7 (968.7) $[M+6H]^{6+}$, 830.4 (830.5) $[M+7H]^{7+}$, 726.9 (726.8) $[M+8H]^{8+}$, 646.4 (646.1) $[M+9H]^{9+}$, 581.9 (581.6) $[M+10H]^{10+}$.

Entry 45 (Figure 2.A63 in Appendix)
CS02-112: Ac-GGGCATGATCT-Lys(Ac)-AEEA-Lys(Ac)-$NH_2$
yield: 11%; $t_r$: 3.54 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 3561.55 [M], observed (from ESI-MS deconvolution): 3562.0; ESI-MS: m/z observed (calcd): 1187.9 (1188.2) $[M+3H]^{3+}$, 891.5 (891.4) $[M+4H]^{4+}$, 713.3 (713.3) $[M+5H]^{5+}$, 594.6 (594.6) $[M+6H]^{6+}$.

Entry 46 (Figure 2.A64 in Appendix)
CS02-115 A1: Ac-TACCGTCGAGT-Lys(Ac)-AEEA-Lys(Ac)-$NH_2$
yield: 16%; $t_r$: 3.18 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 3521.53 [M], observed (from ESI-MS deconvolution): 3520.2; ESI-MS: m/z observed (calcd): 1174.8 (1174.8) $[M+3H]^{3+}$, 881.3 (881.4) $[M+4H]^{4+}$, 705.2 (705.3) $[M+5H]^{5+}$, 588.1 (587.9) $[M+6H]^{6+}$.

Entry 47 (Figure 2.A65 in Appendix)
CS02-115 A2: Ac-GTGACTCAGAT-Lys(Ac)-AEEA-Lys(Ac)-$NH_2$
yield: 14%; $t_r$: 3.55 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; MW calculated: 3545.55 [M], observed (from ESI-MS deconvolution): 3545.6; ESI-MS: m/z observed (calcd): 1182.7 (1182.8) $[M+3H]^{3+}$, 887.3 (887.4) $[M+4H]^{4+}$, 710.2 (710.1) $[M+5H]^{5+}$, 591.8 (591.9) $[M+6H]^{6+}$.

Entry 48 (Figure 2.A66 in Appendix)
CS02-139 pk2: Ac-GTGACTCAGAT- Lys(Ac)-AEEA-Lys(5-TAMRA-lakhkeelakl→)-$NH_2$
yield: 14%; $t_r$: 4.04 min; ε (260 nm): 147180 $M^{-1}cm^{-1}$; MW calculated: 5177.46 [M], observed (from ESI-MS deconvolution): 5177.4; ESI-MS: m/z observed (calcd): 864.1 (863.9) $[M+6H]^{6+}$, 740.9 (740.6) $[M+7H]^{7+}$, 648.2 (648.2) $[M+8H]^{8+}$, 576.4 (576.3) $[M+9H]^{9+}$, 518.7 (518.7) $[M+10H]^{10+}$.

Entry 49 (Figure 2.A67 in Appendix)
CS02-140 pk2: Ac-GTGACTCAGAT- Lys(5-TAMRA)-AEEA-Lys(Ac-lakhkeelakl→)-$NH_2$
yield: 9%; $t_r$: 3.87 min; ε (260 nm): 147180 $M^{-1}cm^{-1}$; MW calculated: 5177.46 [M], observed (from ESI-MS deconvolution): 5179.6; ESI-MS: m/z observed (calcd): 1036.9 (1036.5) $[M+5H]^{5+}$, 864.1 (863.9) $[M+6H]^{6+}$, 740.9 (740.6) $[M+7H]^{7+}$, 648.4 (648.2) $[M+8H]^{8+}$, 576.4 (576.3) $[M+9H]^{9+}$.

Entry 50 (Figure 2.A68 in Appendix)
NM03-13 pk3: Ac-GTGACTCAGAT-Lys(Ac)-AEEA-Lys(5-TAMRA-leekh→)-$NH_2$
yield: 8%; $t_r$: 3.95 min; ε (260 nm): 147180 $M^{-1}cm^{-1}$; MW calculated: 4552.65 [M], observed (from ESI-MS deconvolution): 4553.0; ESI-MS: m/z observed (calcd): 911.6 (911.5) $[M+5H]^{5+}$, 760.0 (759.8) $[M+6H]^{6+}$, 651.5 (651.4) $[M+7H]^{7+}$, 570.1 (570.1) $[M+8H]^{8+}$.

Entry 51 (Figure 2.A69 in Appendix)
NM03-09 pk2: Ac-GTGACTCAGAT-Lys(Ac)-AEEA-Lys(5-TAMRA-lkaleekhkal→)-$NH_2$
yield: 13%; $t_r$: 4.17 min; ε (260 nm): 147180 $M^{-1}cm^{-1}$; MW calculated: 5177.46 [M], observed (from ESI-MS deconvolution): 5178.6; ESI-MS: m/z observed (calcd): 864.5 (863.9) $[M+6H]^{6+}$, 740.6 (740.6) $[M+7H]^{7+}$, 648.4 (648.2) $[M+8H]^{8+}$, 576.5 (576.3) $[M+9H]^{9+}$.

Entry 52 (Figure 2.A70 in Appendix)
CS02-193 pk2: Ac-GTGACTCAGAT-Lys(Ac)-AEEA-Lys(5-TAMRA-leeklkaleekhkaleekl→)-$NH_2$
yield: 4%; $t_r$: 4.24 min; ε (260 nm): 147180 $M^{-1}cm^{-1}$; MW calculated: 6176.56 [M], observed (from ESI-MS deconvolution): 6177.4; ESI-MS: m/z observed (calcd): 1030.4 (1030.4) $[M+6H]^{6+}$, 883.2 (883.4) $[M+7H]^{7+}$, 773.3 (773.1) $[M+8H]^{8+}$, 687.5 (687.3) $[M+9H]^{9+}$, 618.8 (618.7) $[M+10H]^{10+}$, 563.3 (562.5) $[M+11H]^{11+}$.

Entry 53 (Figure 2.A71 in Appendix)
NM03-14 pk2: Ac-GTGACTCAGAT-Lys(Ac)-AEEA-Lys(5-TAMRA-hkeel→)-$NH_2$
yield: 12%; $t_r$: 3.68 min; ε (260 nm): 147180 $M^{-1}cm^{-1}$; MW calculated: 4552.65 [M], observed (from ESI-MS deconvolution): 4553.0; ESI-MS: m/z observed (calcd): 1139.6 (1139.2) $[M+4H]^{4+}$, 912.2 (911.5) $[M+5H]^{5+}$, 760.0 (759.8) $[M+6H]^{6+}$, 651.4 (651.4) $[M+7H]^{7+}$, 570.1 (570.1) $[M+8H]^{8+}$, 507.0 (506.8) $[M+9H]^{9+}$.

Entry 54 (Figure 2.A72 in Appendix)
NM03-15 pk3: Ac-GTGACTCAGAT-Lys(Ac)-AEEA-Lys(5-TAMRA-lkeelakhkeelaklkeel→)-$NH_2$
yield: 5%; $t_r$: 4.29 min; ε (260 nm): 147180 $M^{-1}cm^{-1}$; MW calculated: 6176.56 [M], observed (from ESI-MS deconvolution): 6177.6; ESI-MS: m/z observed (calcd): 1030.8 (1030.4) $[M+6H]^{6+}$, 883.6 (883.4) $[M+7H]^{7+}$, 773.3 (773.1) $[M+8H]^{8+}$, 687.5 (687.3) $[M+9H]^{9+}$, 618.9 (618.7) $[M+10H]^{10+}$, 563.0 (562.5) $[M+11H]^{11+}$.

Entry 55 (Figure 2.A73 in Appendix)
CS02-192 pk2: Ac-GTGACTCAGAT- Lys(Ac)-AEEA-Lys(5-TAMRA-lakhkeelakl→)-$NH_2$
yield: 24%; $t_r$: 4.32 min; ε (260 nm): 31980 $M^{-1}cm^{-1}$; MW calculated: 1690.98 [M], observed (from ESI-MS deconvolution): 1690.4; ESI-MS: m/z observed (calcd): 564.8 (564.7) $[M+3H]^{3+}$, 423.7 (423.7) $[M+4H]^{4+}$, 339.3 (339.2) $[M+5H]^{5+}$.

## 2.6 Appendix Chapter 2

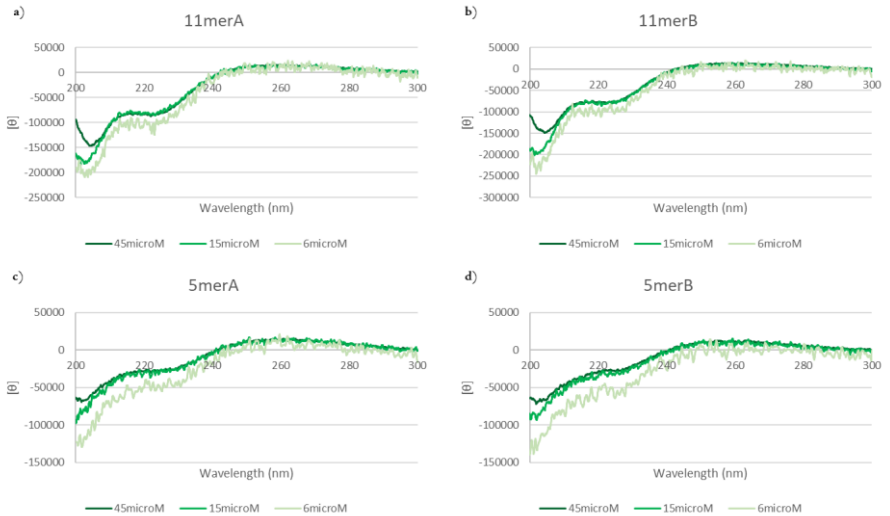



***Figure 2.A1.*** *a) CD signature of 11merA to the three different concentrations, b) CD signature of 11merB to the three different concentrations, c) CD signature of 5merA to the three different concentrations, d) CD signature of 5merB to the three different concentrations.*

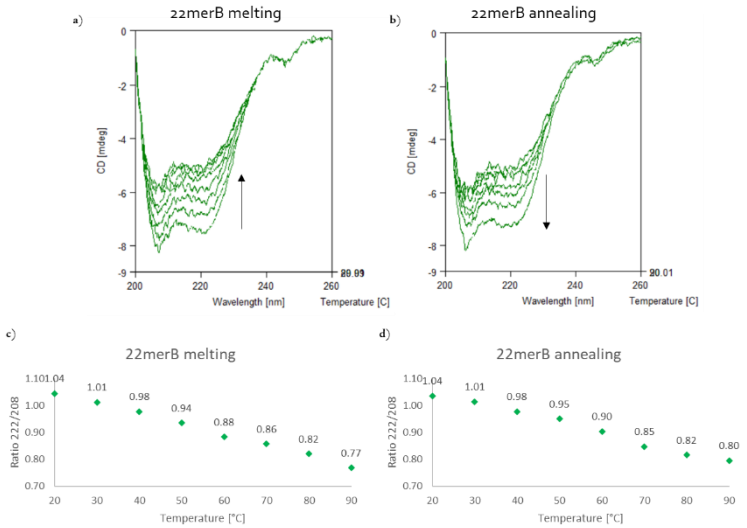



***Figure 2.A2.*** *a) Self-melting of 22merB, recording a spectrum every 10°C, from 20°C to 90°C. b) annealing of 22merB, recording a spectrum every 10°C, from 90°C to 10°C, c) decrease of ratio 222/208 during the self-melting to 22merB, d) increase of ratio 222/208 during the annealing to 22merB.*

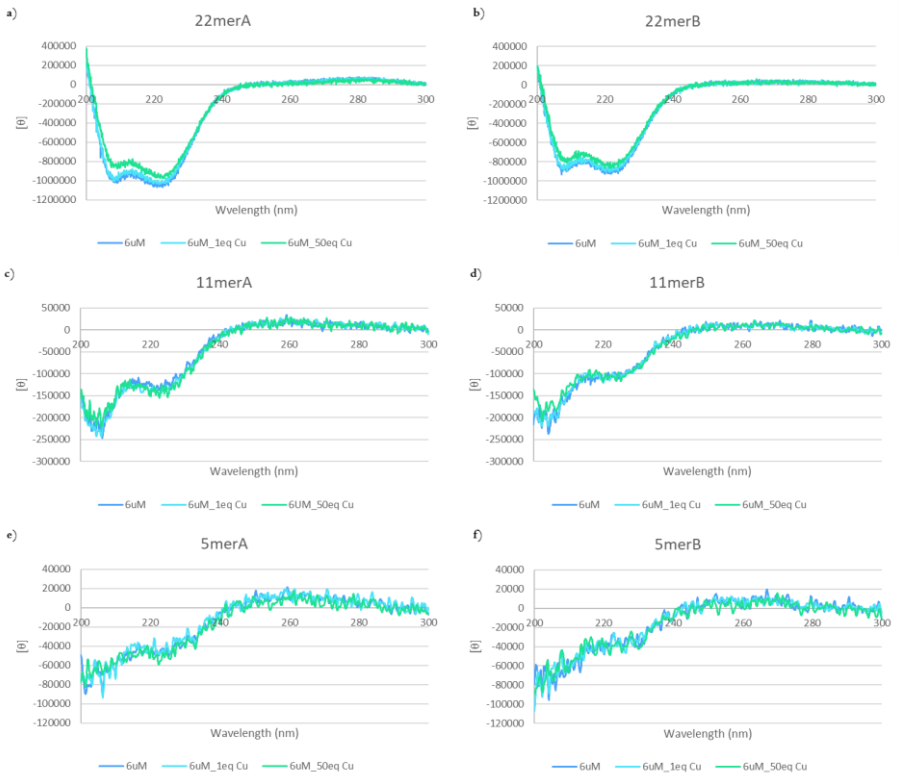



***Figure 2.A3.*** *a) CD results of 22merA in present of 1 eq and 50 eq of Copper(II) sulfate, b) CD results of 22merB in present of 1 eq and 5 oeq of Copper(II) sulfate, c) CD results of 11merA in present of 1 eq and 50 eq of Copper(II) sulfate, d) CD results of 11merB in present of 1 eq and 50 eq of Copper(II) sulfate, e) CD results of 5merA in present of 1 eq and 50 eq of Copper(II) sulfate, f) CD results of 5merB in present of 1 eq and 50 eq of Copper(II) sulfate.*

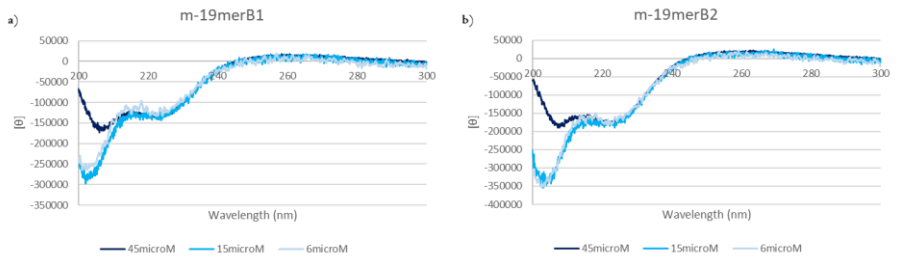



***Figure 2.A4.*** *CD measurements at three different concentrations, in specific 45 µM, 15 µM and 6 µM, of a) m-19merB1 and b) m-19merB3.*

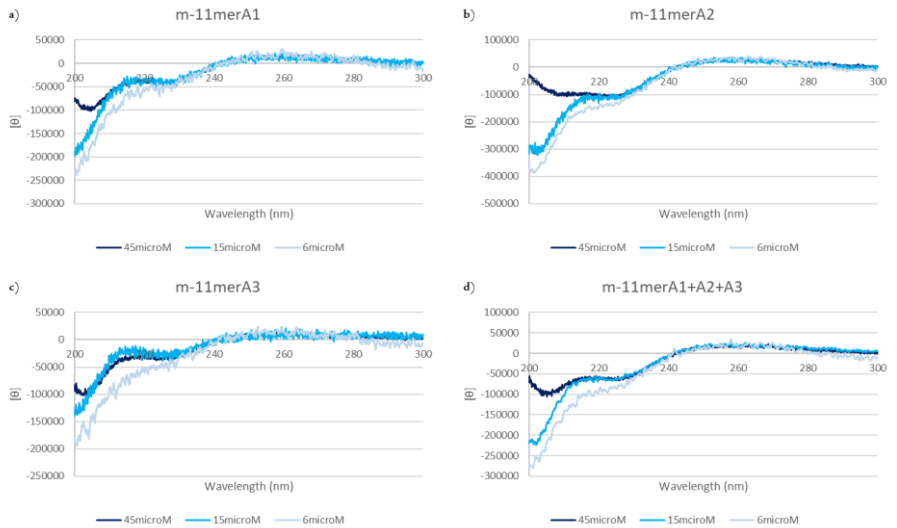



***Figure 2.A5.*** *CD measurements at three different concentrations (45μM, 15μM and 6μM) of a) m-11merA1, b) m-11merA2 and c) m-11merA3. d) CD measurements of combination of m-11merA1, m-11meA2 and m-11mer A3 at three different final concentrations (45μM, 15μM and 6μM).*

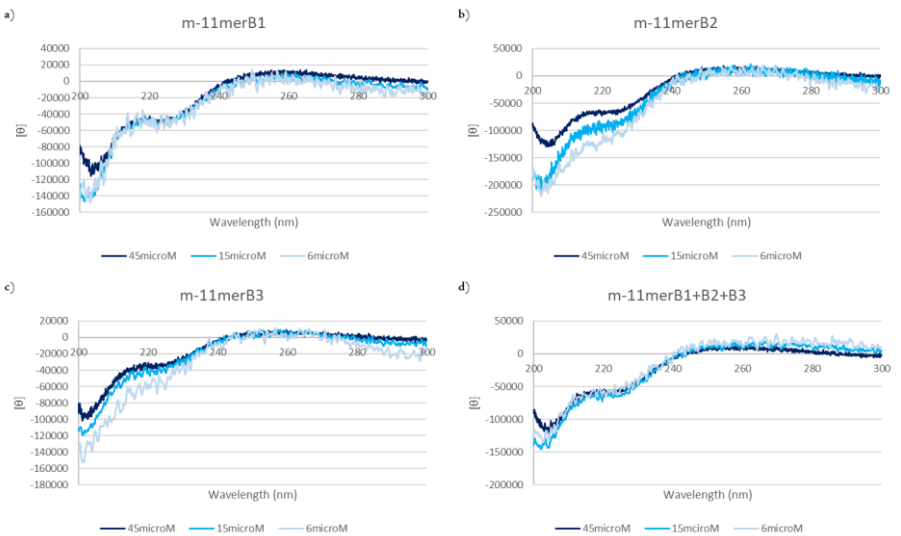



***Figure 2.A6.*** *CD measurements at three different concentrations (45 μM, 15 μM and 6 μM) of a) m-11merB1, b) m-11merB2 and c) m-11merB3. d) CD measurements of combination of m-11merB1, m-11mer B2 and m-11mer B3 at three different final concentrations (45 μM, 15 μM and 6 μM).*

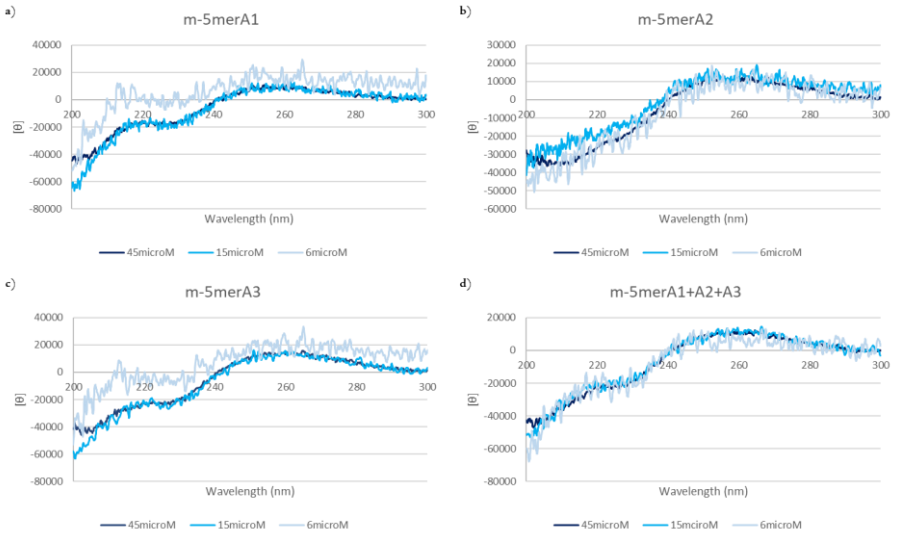



***Figure 2.A7.*** *CD measurements at three different concentrations (45 µM, 15 µM and 6 µM) of a) m-5merA1, b) m-5merA2 and c) m-5merA3. d) CD measurements of combination of m-5merA1, m-5merA2, and m-5mer A3 at three different final concentrations (45 µM, 15 µM and 6 µM).*

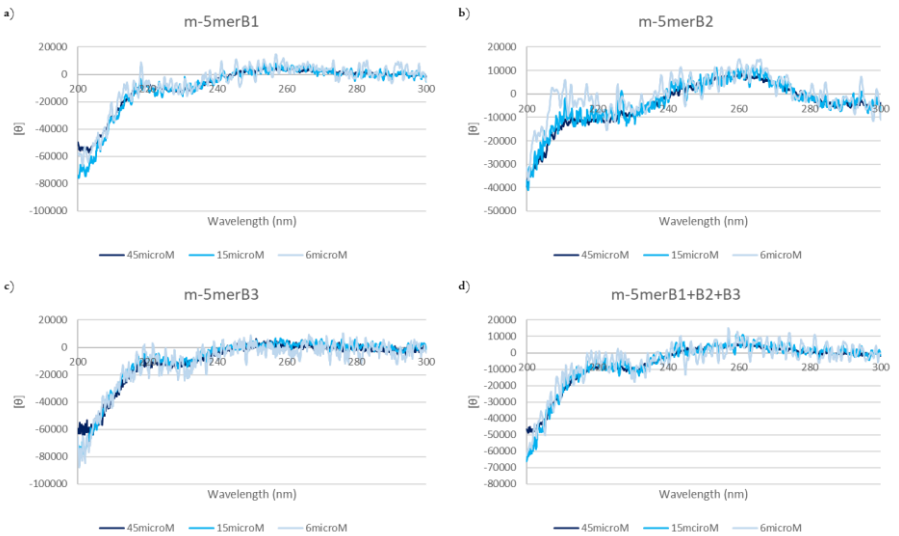



***Figure 2.A8.*** *CD measurements at three different concentrations (45 µM, 15 µM and 6 µM) of a) m-5merB1, b) m-5merB2 and c) m-5merB3. d) CD measurements of combination of m-5merB1, m-5merB2 and m-5merB3 at three different final concentrations (45 µM, 15 µM and 6 µM).*

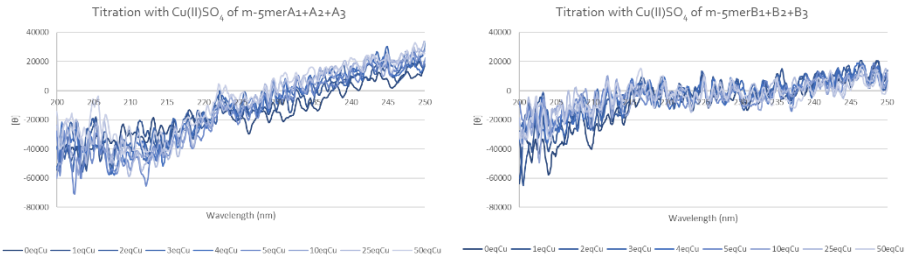



***Figure 2.A9.*** *Titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate of 5mer modified peptides a) orientation A and b) orientation B.*

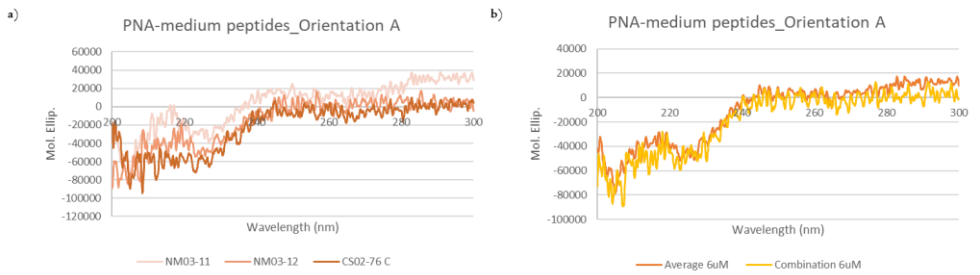



***Figure 2.A10.*** *a) CD results of 2A-1, 2A-2, and 2A-3 as single separate PNA-peptide conjugates in solution at 6 μM as concentration. b) Comparison between CD spectrum obtained by the sample containing all the three PNA-peptide sequences with 11mer peptides domains in orientation A and the spectrum carried out by the mathematic elaboration of the results given by each individual sequence.*

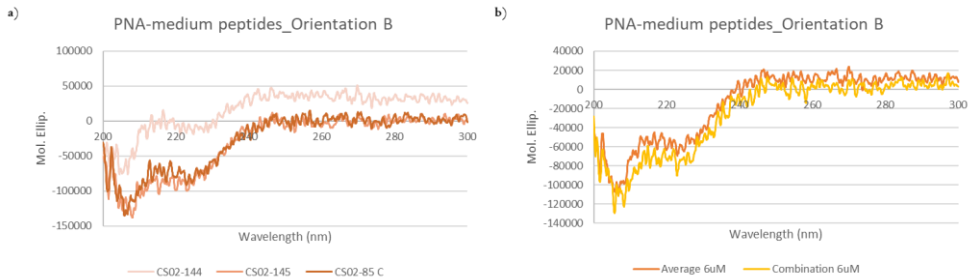



***Figure 2.A11.*** *a) CD results of 2B-1, 2B-2, and 2B-3 as single separate PNA-peptide conjugates in solution at 6 μM as concentration. b) Comparison between CD spectrum obtained by the sample containing all the three PNA-peptide sequences with 11mer peptides domains in orientation B and the spectrum carried out by the mathematic elaboration of the results given by each individual sequence.*

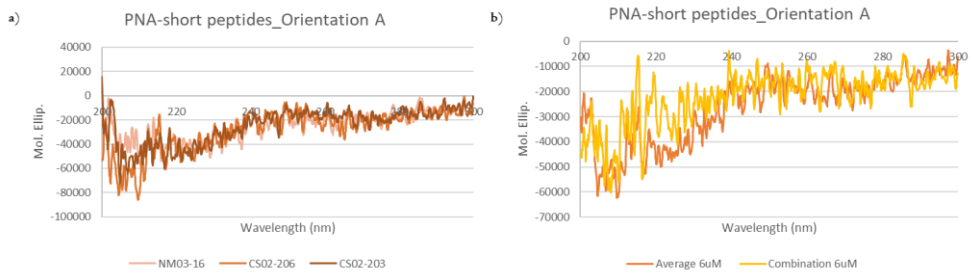



***Figure 2.A12.*** *a) CD results of 1A-1, 1A-2, and 1A-3 as single separate PNA-peptide conjugates in solution at 6 μM as concentration. b) Comparison between CD spectrum obtained by the sample containing all the*

*three PNA-peptide sequences with 5mer peptides domains in orientation A and the spectrum carried out by the mathematic elaboration of the results given by each individual sequence.*

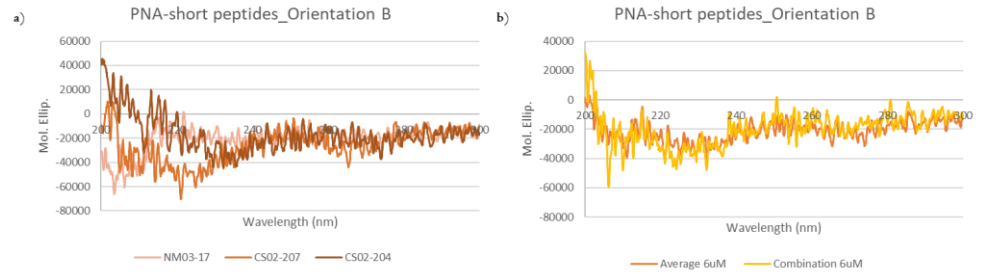



***Figure 2.A13.*** *a) CD results of 1B-1, 1B-2, and 1B-3 as single separate PNA-peptide conjugates in solution at 6 μM as concentration. b) Comparison between CD spectrum obtained by the sample containing all the three PNA-peptide sequences with 5mer peptides domains in orientation B and the spectrum carried out by the mathematic elaboration of the results given by each individual sequence.*

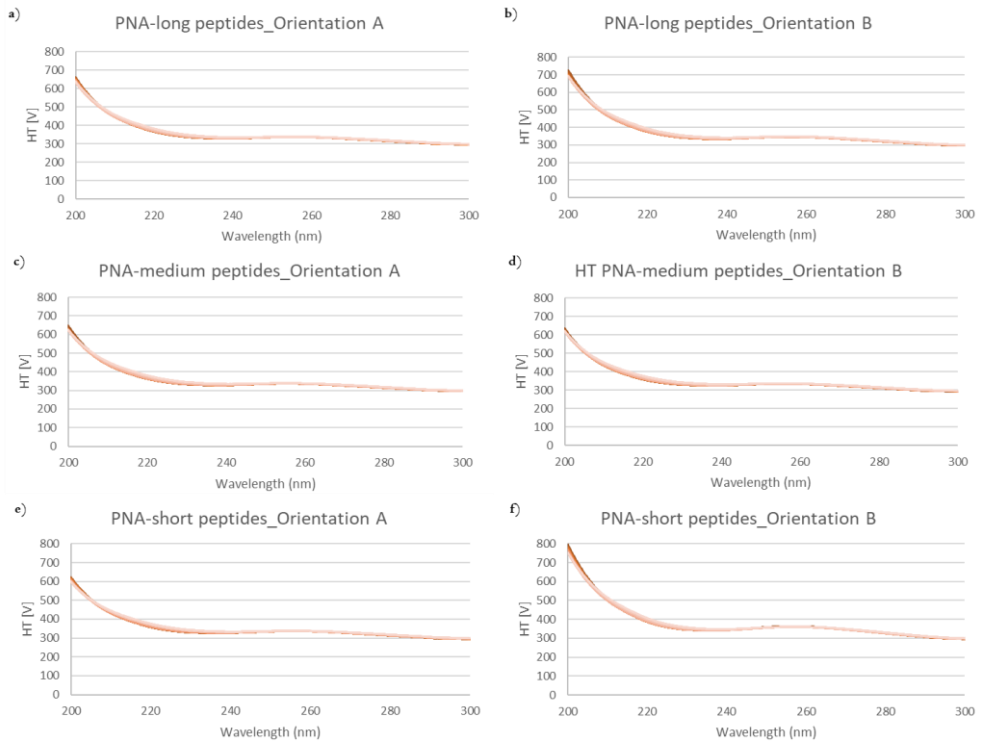



***Figure 2.A14.*** *HT signature spectra regarding a) titration with Cu(II) of PNA-peptide conjugates with 19mer peptide sequences in orientation A; b) titration with Cu(II) of PNA-peptide conjugates with 19mer peptide sequences in orientation B; c) titration with Cu(II) of PNA-peptide conjugates with 11mer peptide sequences in orientation A; d) titration with Cu(II) of PNA-peptide conjugates with 11mer peptide sequences in orientation B; e) titration with Cu(II) of PNA-peptide conjugates with 5mer peptide sequences in orientation A; f) titration with Cu(II) of PNA-peptide conjugates with 5mer peptide sequences in orientation B.*

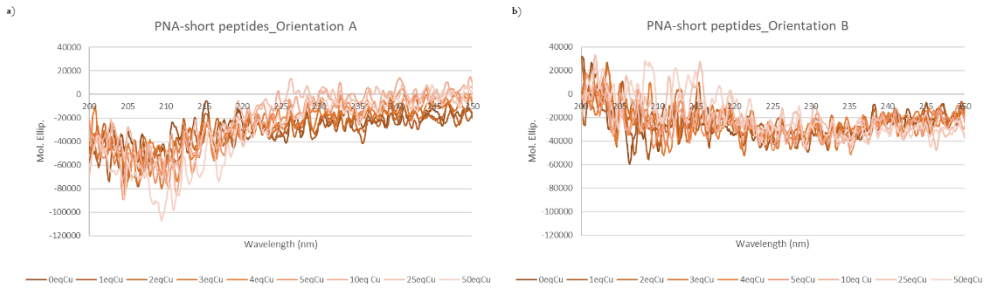



***Figure 2.A15.*** *Titration with 1, 2, 3, 4, 5, 10, 25 and 50 eq of Copper(II) sulfate of PNA-peptide conjugates with 5mer peptide sequences a) orientation A and b) orientation B.*

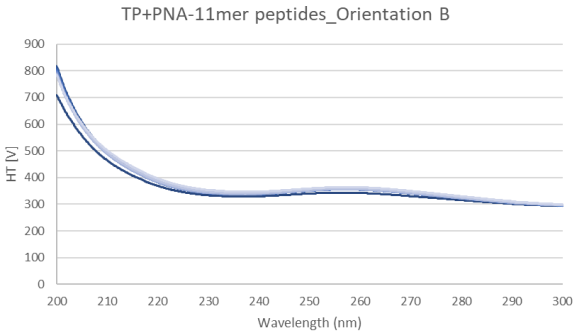



***Figure 2.A16.*** *HT signature spectrum regarding the titration with Cu(II) of total complex with TP and PNA-11mer peptide conjugates sequences in orientation B.*

## 2.6.1 TP and mismatched TP characterizations

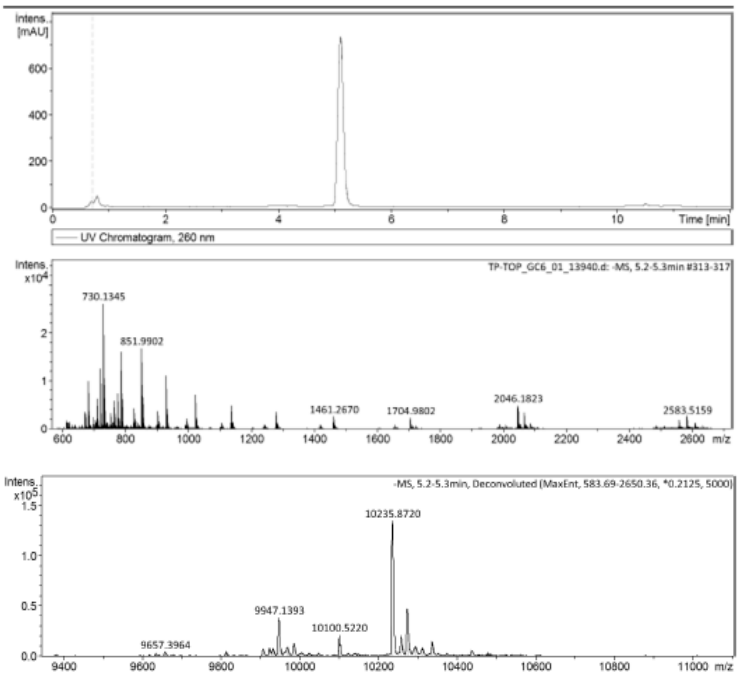



***Figure 2.A17.*** *UPLC-MS of TP: chromatogram, mass spectrum and deconvoluted spectrum.*

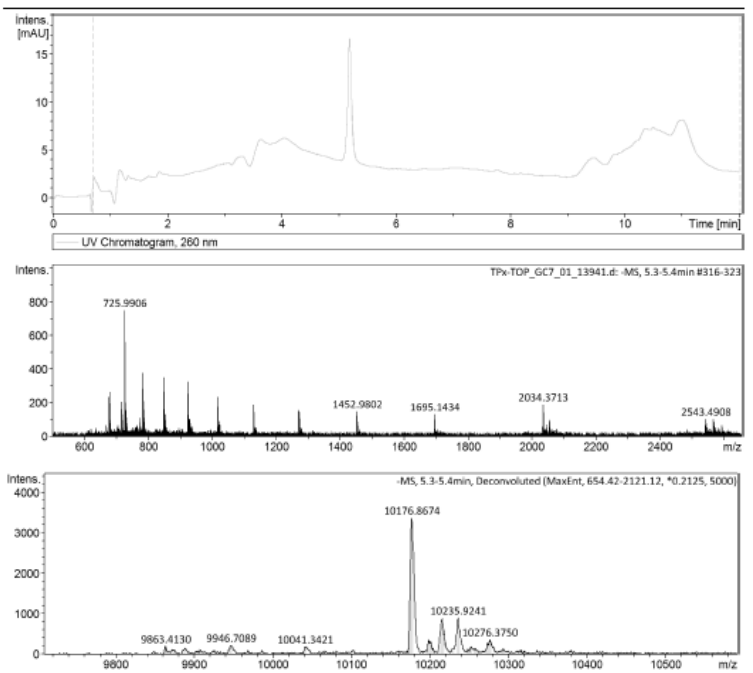



***Figure 2.A18.*** *UPLC-MS of mismatched TP: chromatogram, mass spectrum and deconvoluted spectrum.*

## 2.6.2 TRI-family peptides characterizations

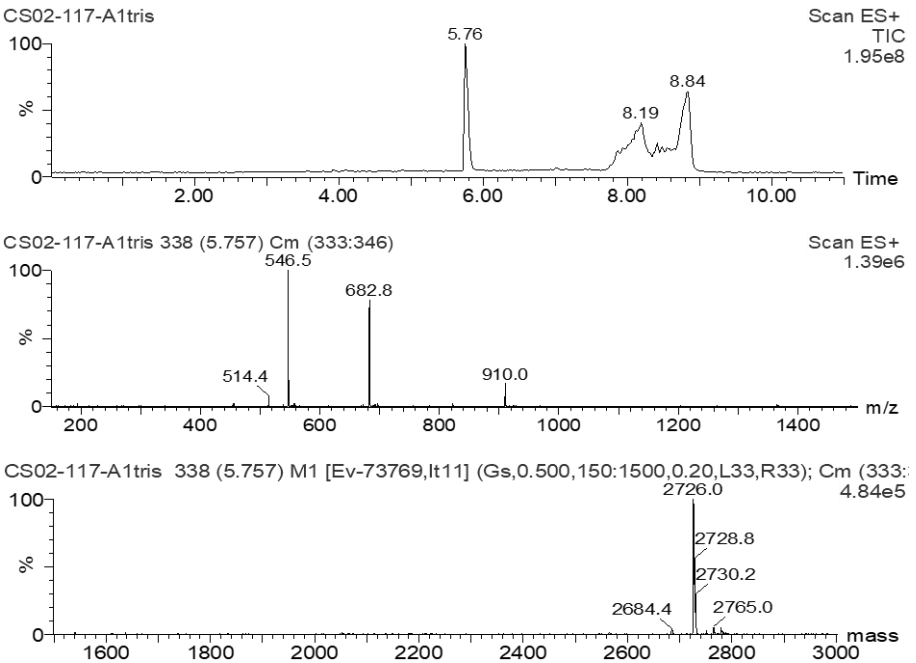



***Figure 2.A19.*** *UPLC-MS of CS02-117 A1: chromatogram, mass spectrum and deconvoluted spectrum.*

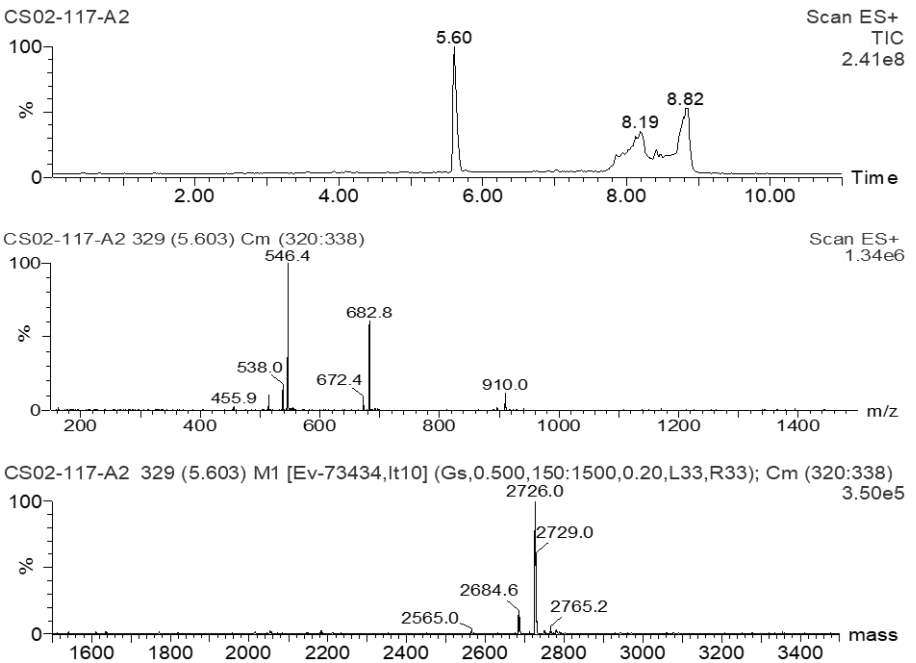



***Figure 2.A20.*** *UPLC-MS of CS02-117 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

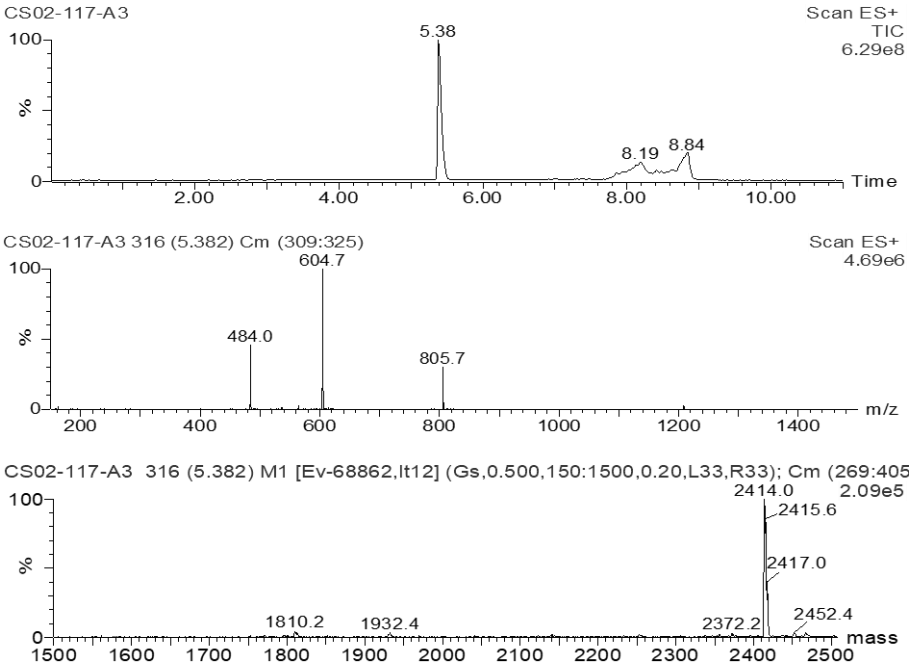



***Figure 2.A21.*** *UPLC-MS of CSo2-117 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

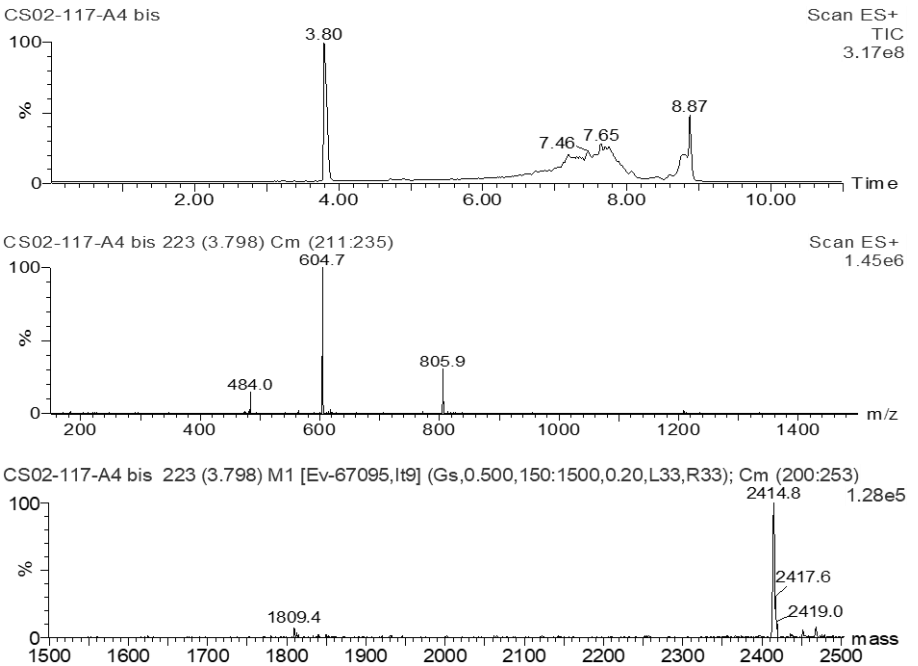



***Figure 2.A22.*** *UPLC-MS of CSo2-117 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

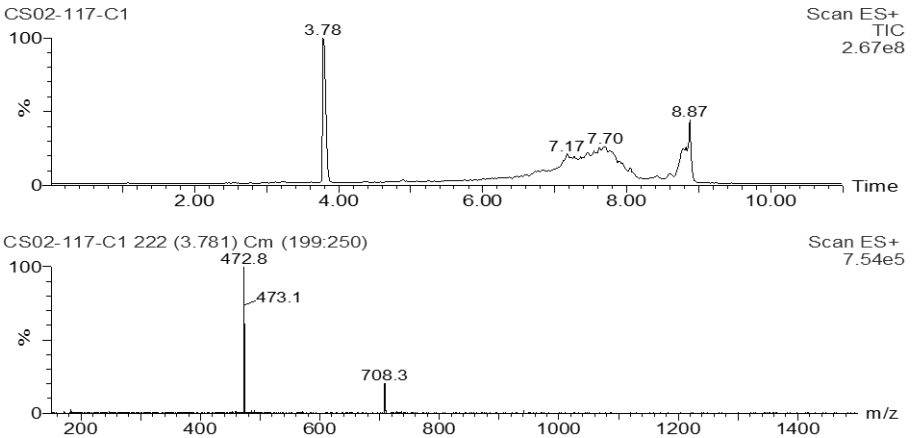



***Figure 2.A23.*** *UPLC-MS of CS02-117 C1: chromatogram and mass spectrum.*

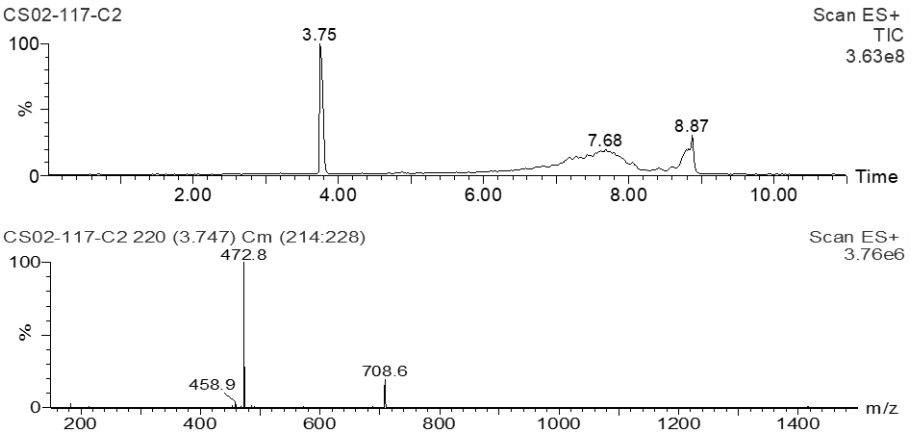



***Figure 2.A24.*** *UPLC-MS of CS02-117 C2: chromatogram and mass spectrum.*

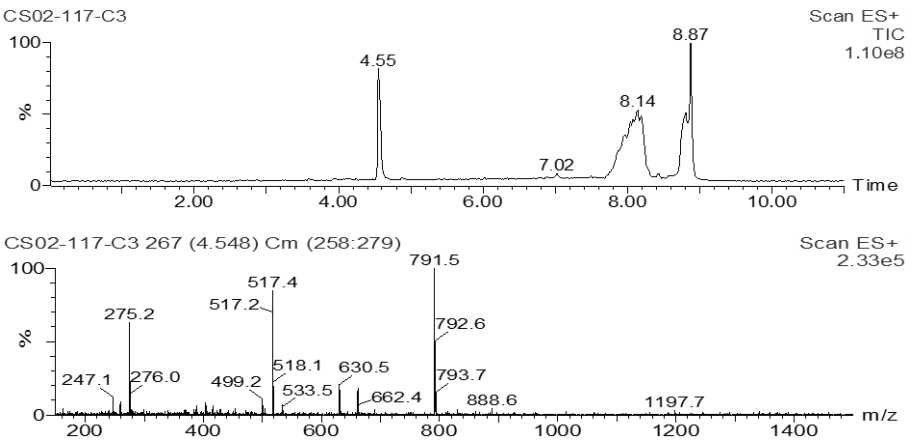



***Figure 2.A25.*** *UPLC-MS of CS02-117 C3: chromatogram and mass spectrum.*

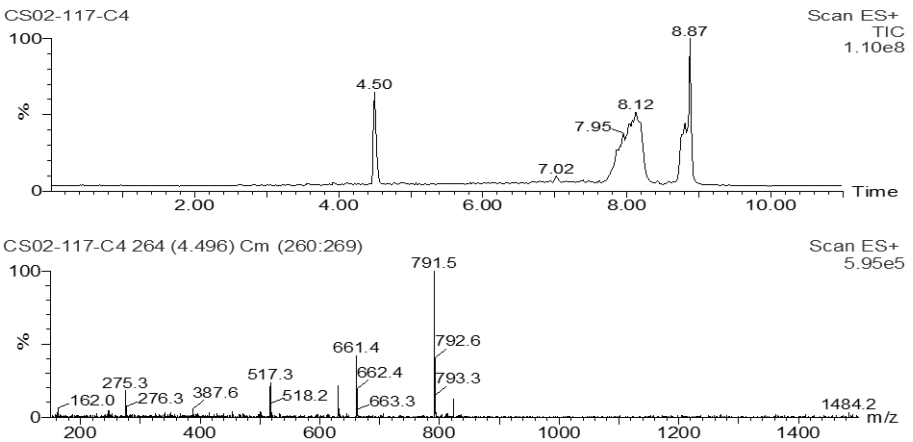



***Figure 2.A26.*** *UPLC-MS of CS02-117 C4: chromatogram and mass spectrum.*

## 2.6.3 Modified TRI-family peptides characterizations

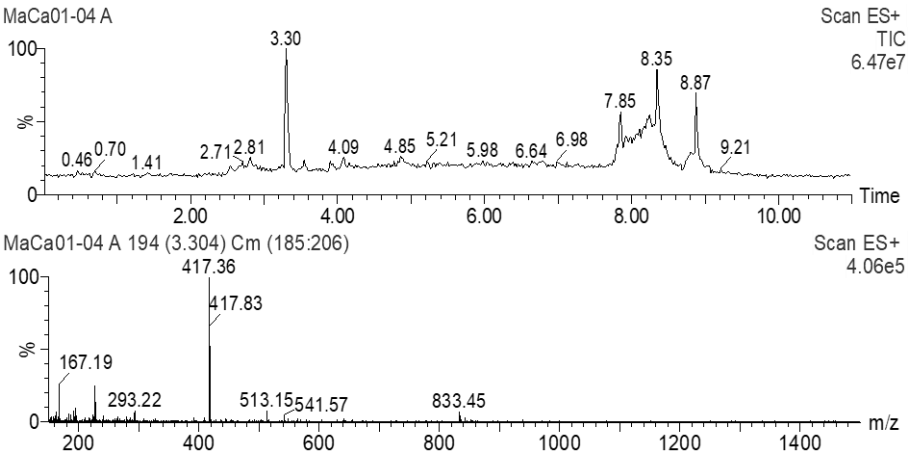



***Figure 2.A27.*** *UPLC-MS of MaCa01-04 A: chromatogram and mass spectrum.*

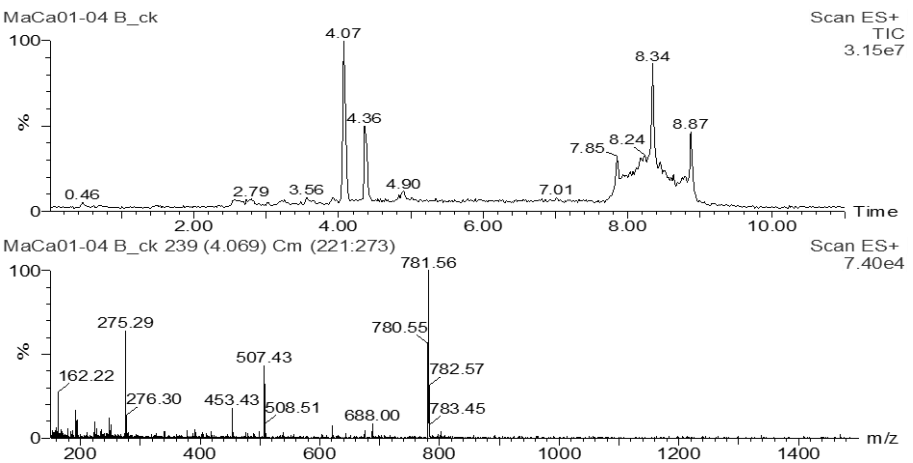



***Figure 2.A28.*** *UPLC-MS of MaCa01-04 B: chromatogram and mass spectrum.*

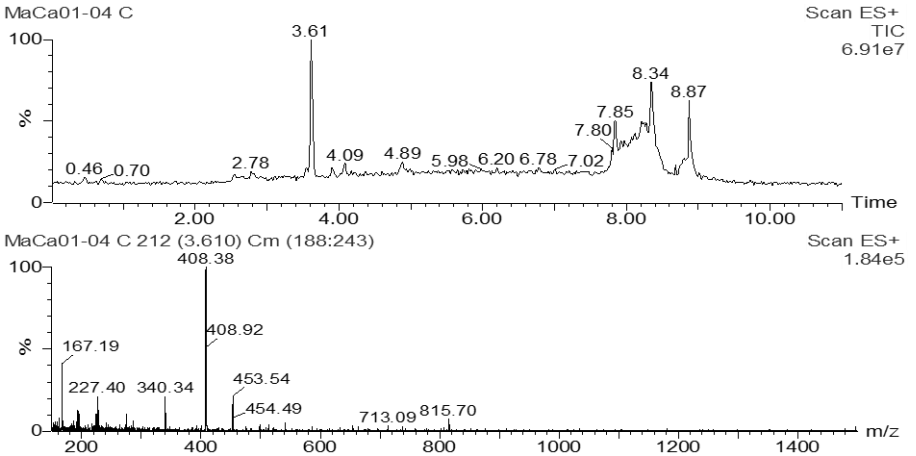



***Figure 2.A29.*** *UPLC-MS of MaCa01-04 C: chromatogram and mass spectrum.*

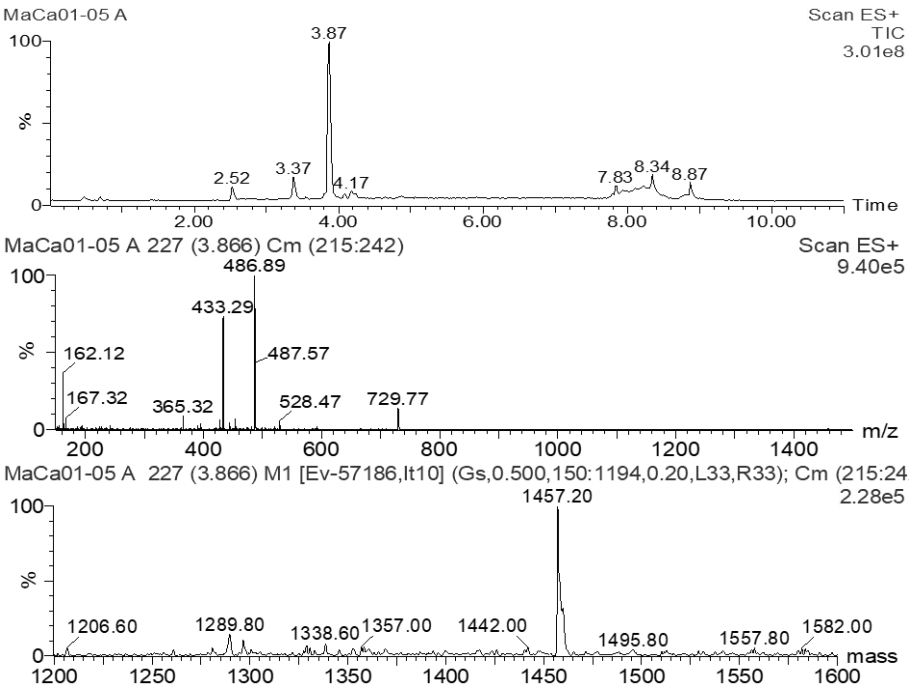



***Figure 2.A30.*** *UPLC-MS of MaCa01-05 A: chromatogram, mass spectrum and deconvoluted spectrum.*

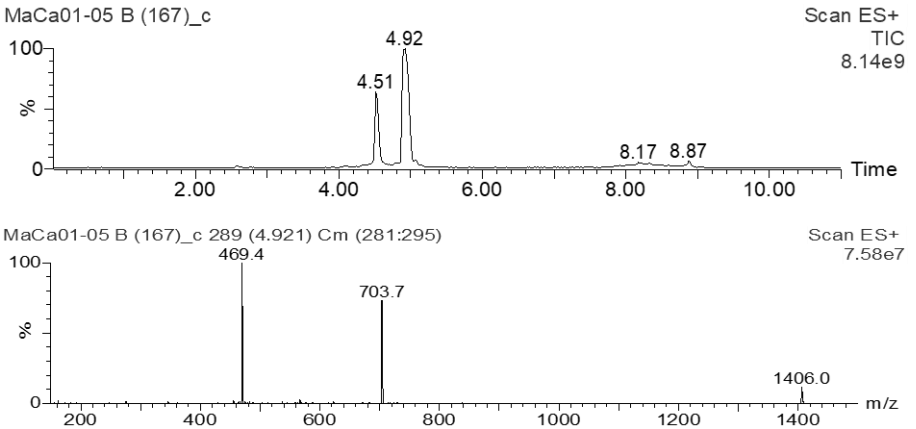



***Figure 2.A31.*** *UPLC-MS of MaCa01-05 B: chromatogram and mass spectrum.*

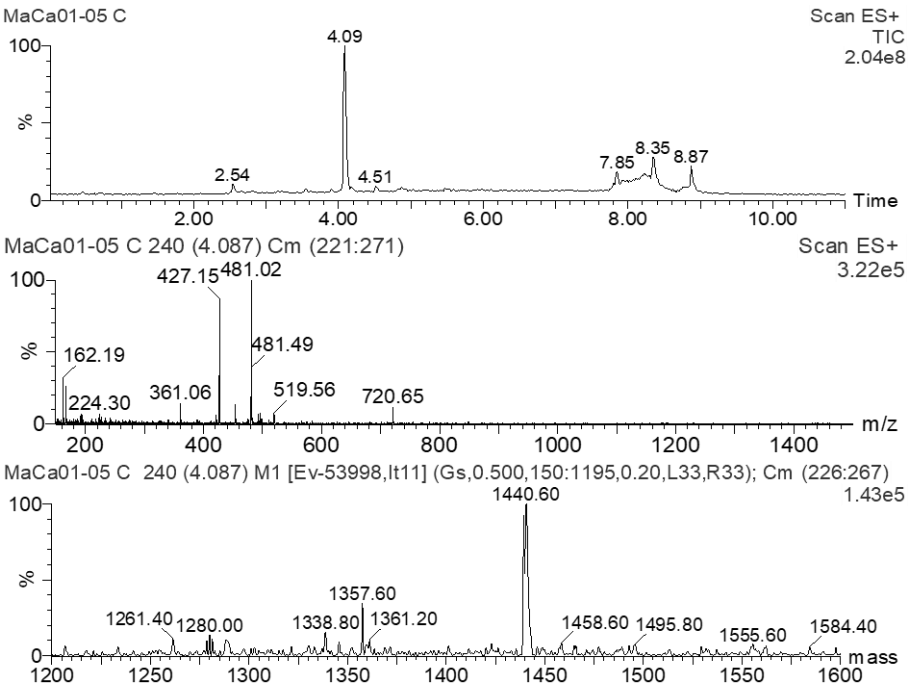



***Figure 2.A32.*** *UPLC-MS of MaCa01-05 C: chromatogram, mass spectrum and deconvoluted spectrum.*

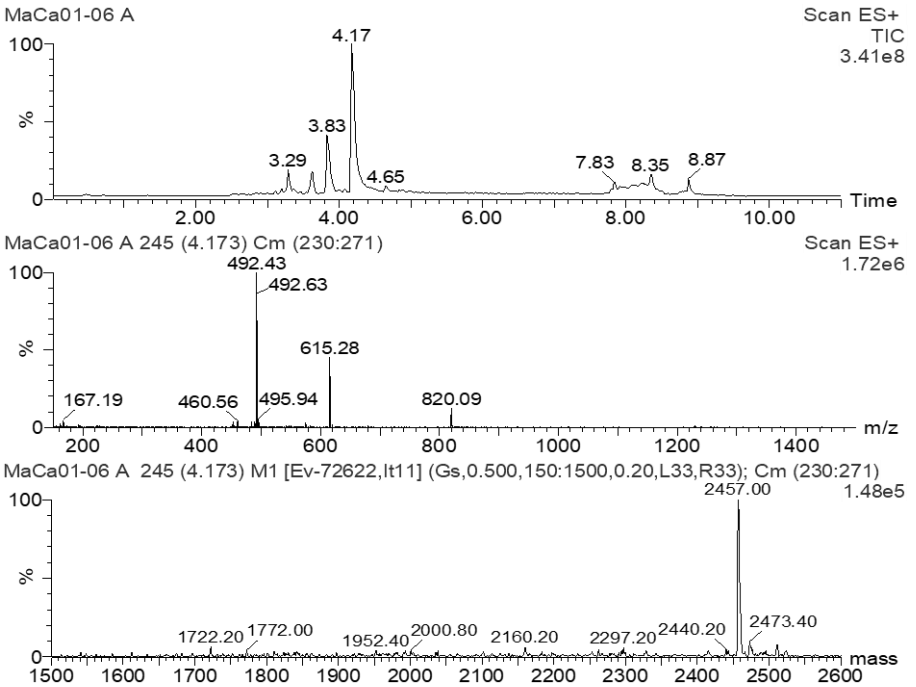



***Figure 2.A33.*** *UPLC-MS of MaCa01-06 A: chromatogram, mass spectrum and deconvoluted spectrum.*

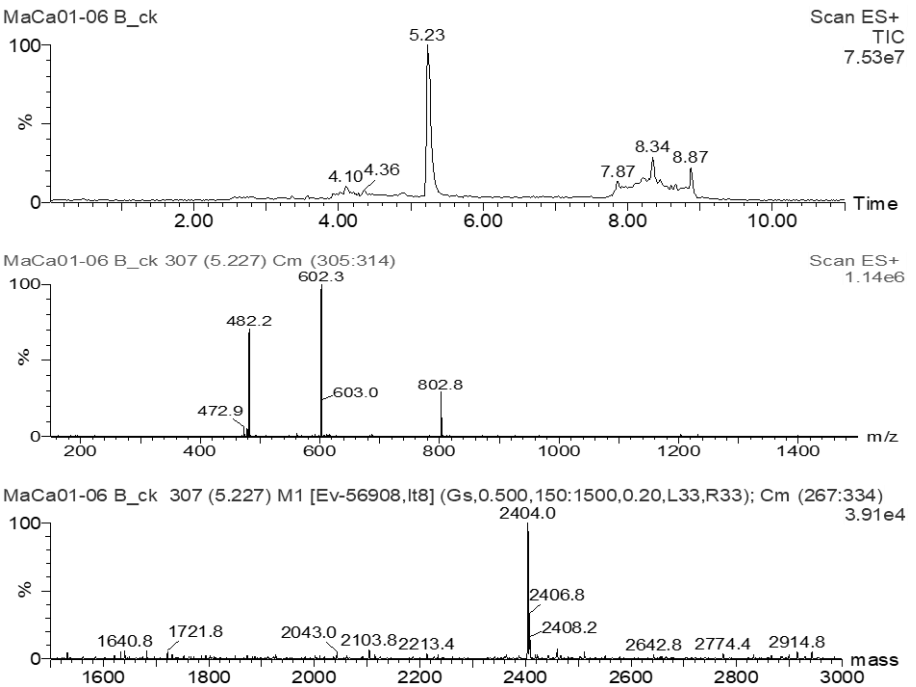



***Figure 2.A34.*** *UPLC-MS of MaCa01-06 B: chromatogram, mass spectrum and deconvoluted spectrum.*

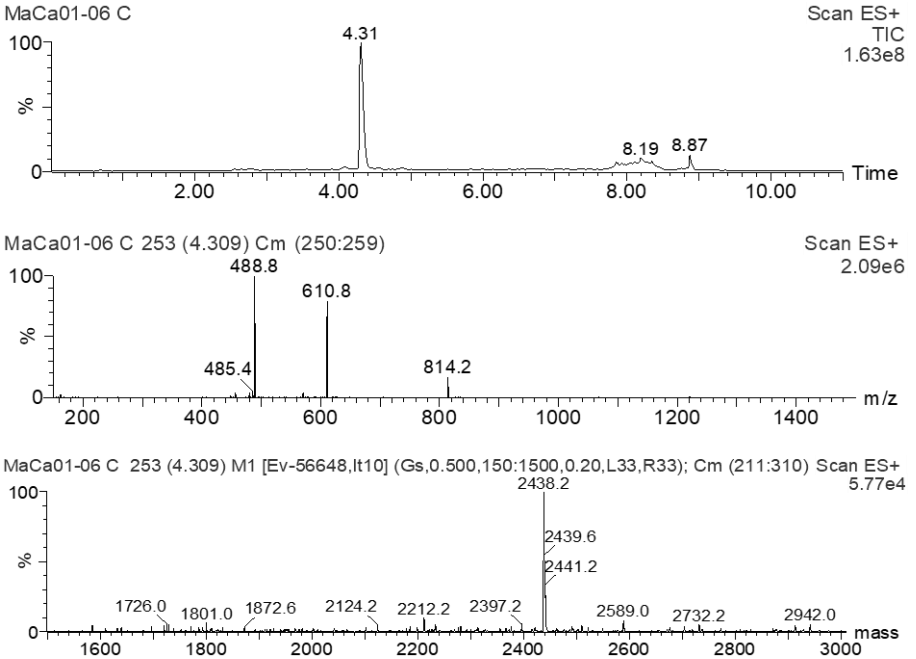



***Figure 2.A35.*** *UPLC-MS of MaCa01-06 C: chromatogram, mass spectrum and deconvoluted spectrum.*

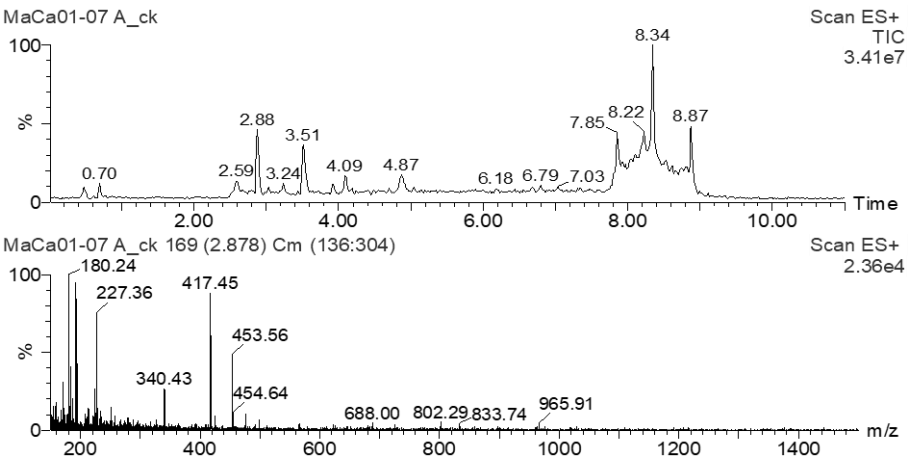



***Figure 2.A36.*** *UPLC-MS of MaCa01-06 B: chromatogram and mass spectrum.*

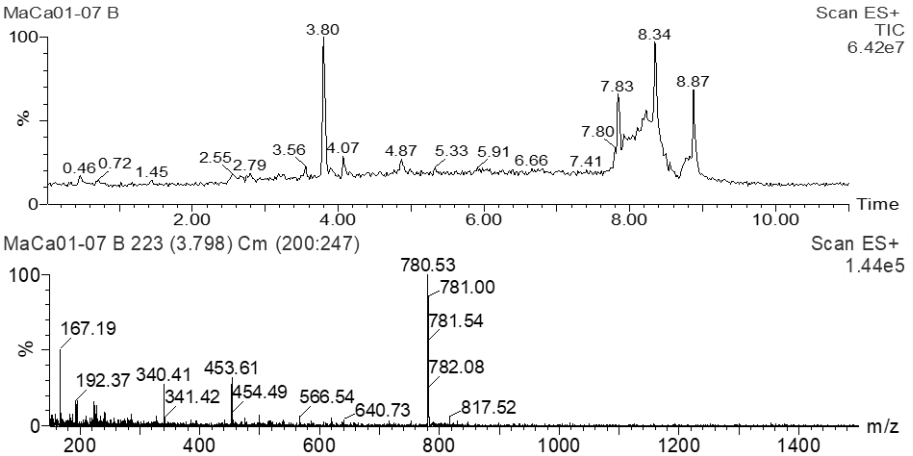



***Figure 2.A37.*** *UPLC-MS of MaCa01-07 B: chromatogram and mass spectrum.*

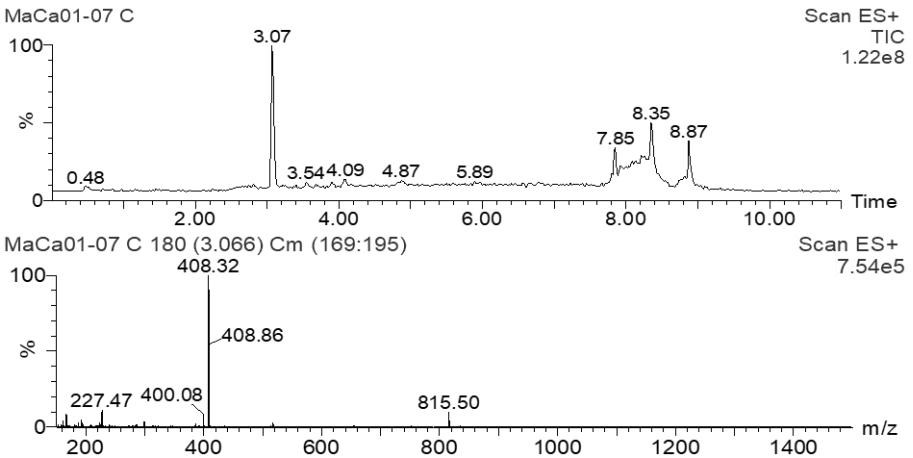



***Figure 2.A38.*** *UPLC-MS of MaCa01-07 C: chromatogram and mass spectrum.*

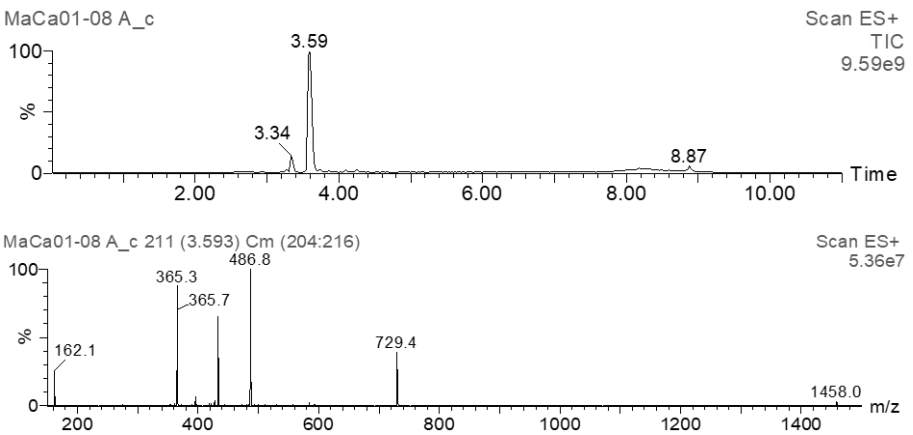



***Figure 2.A39.*** *UPLC-MS of MaCa01-08 A: chromatogram and mass spectrum.*

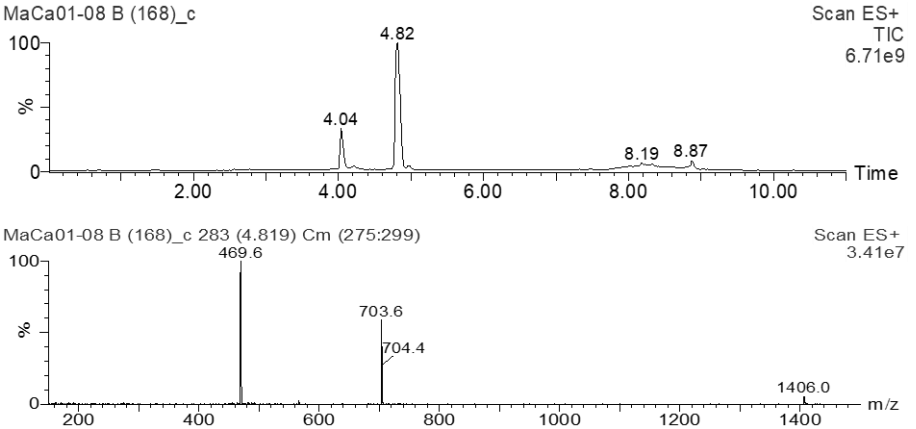



***Figure 2.A40.*** *UPLC-MS of MaCa01-08 B: chromatogram and mass spectrum.*

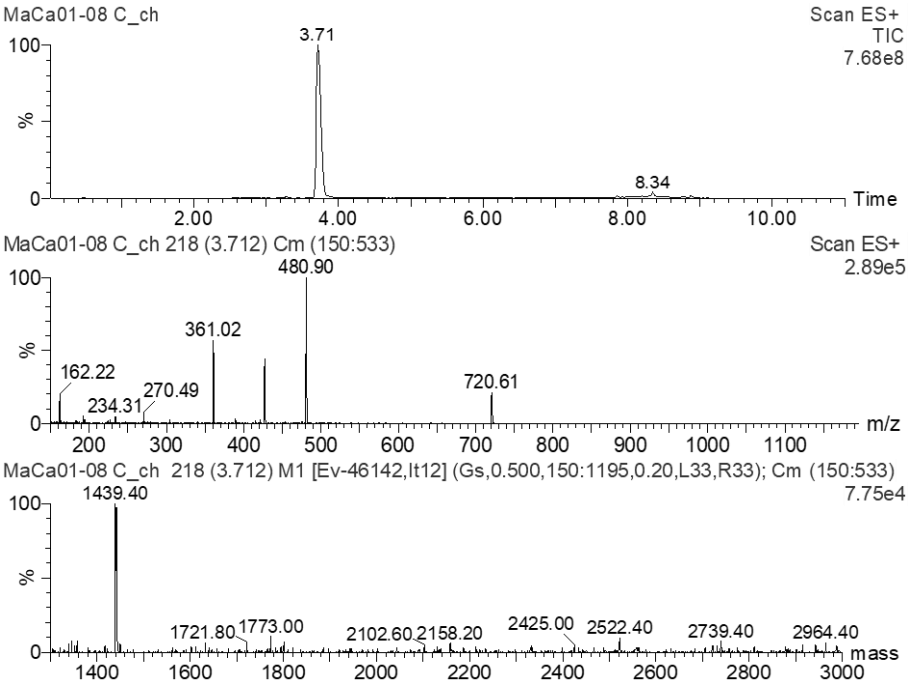



***Figure 2.A41.*** *UPLC-MS of MaCa01-08 C: chromatogram, mass spectrum and deconvoluted spectrum.*

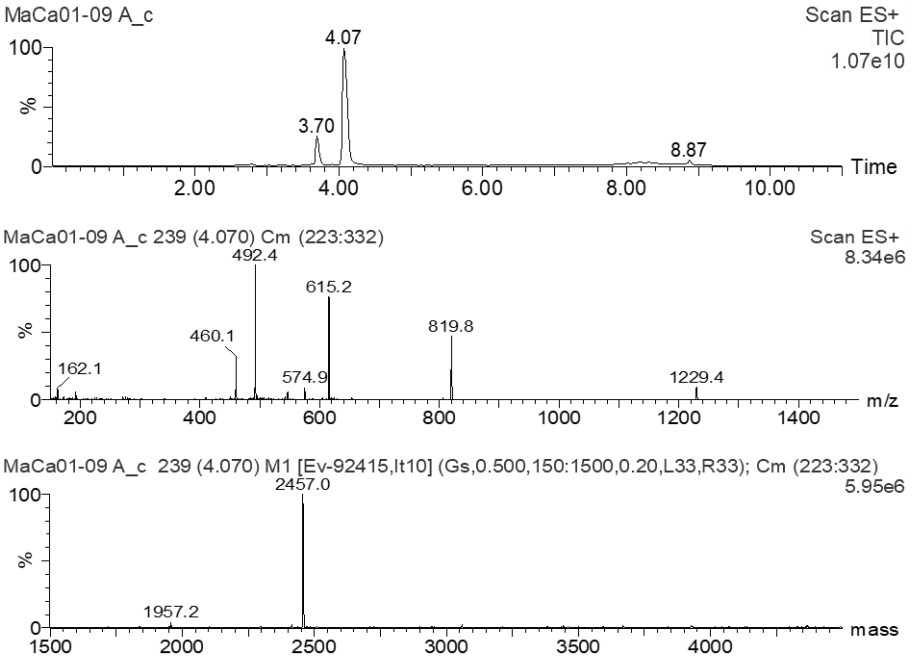



***Figure 2.A42.*** *UPLC-MS of MaCa01-09 A: chromatogram, mass spectrum and deconvoluted spectrum.*

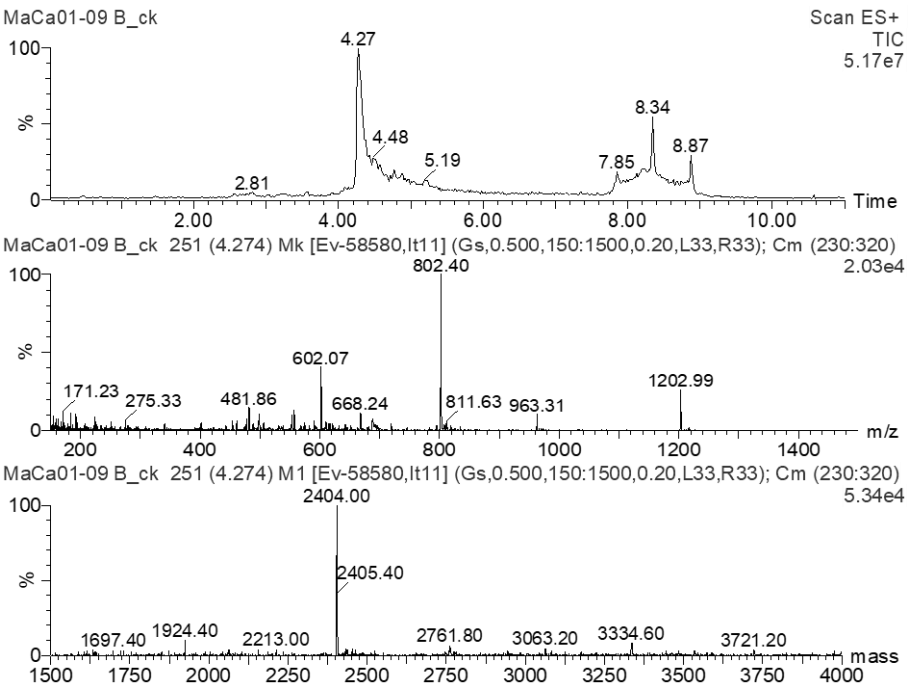



***Figure 2.A43.*** *UPLC-MS of MaCa01-09 B: chromatogram, mass spectrum and deconvoluted spectrum.*

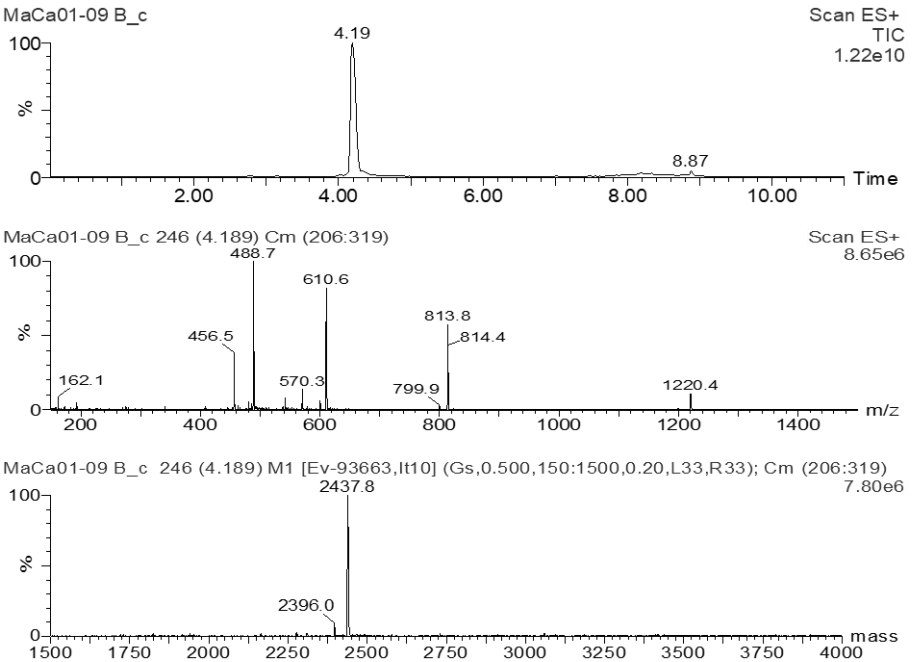



***Figure 2.A44.*** *UPLC-MS of MaCa01-09 C: chromatogram, mass spectrum and deconvoluted spectrum.*

## 2.6.4 PNA-peptide conjugates and PNAs characterizations

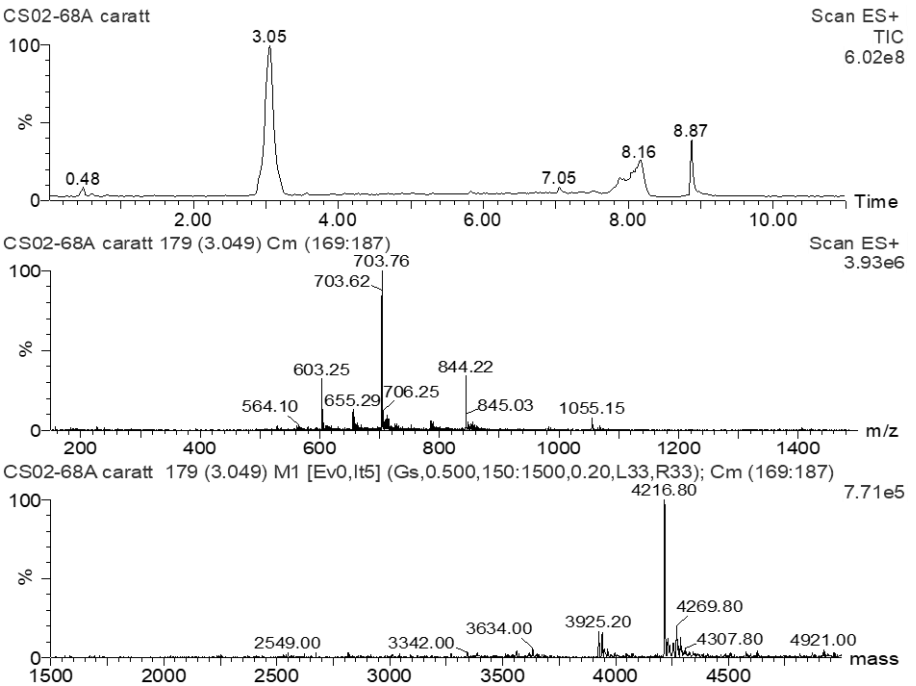



***Figure 2.A45.*** *UPLC-MS of CS02-68 A: chromatogram, mass spectrum and deconvoluted spectrum.*

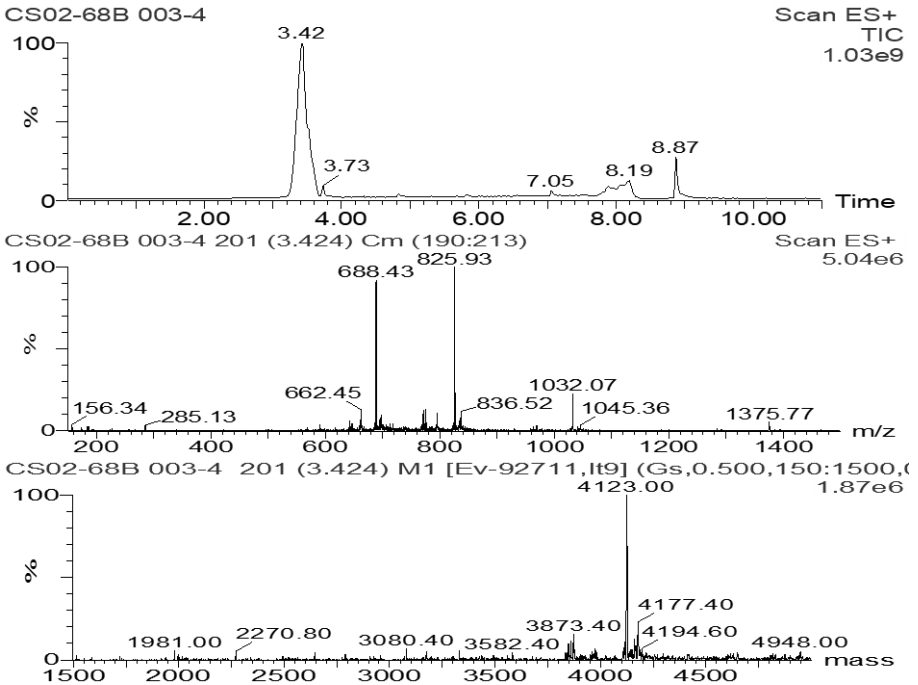



***Figure 2.A46.*** *UPLC-MS of CS02-68 B: chromatogram, mass spectrum and deconvoluted spectrum.*

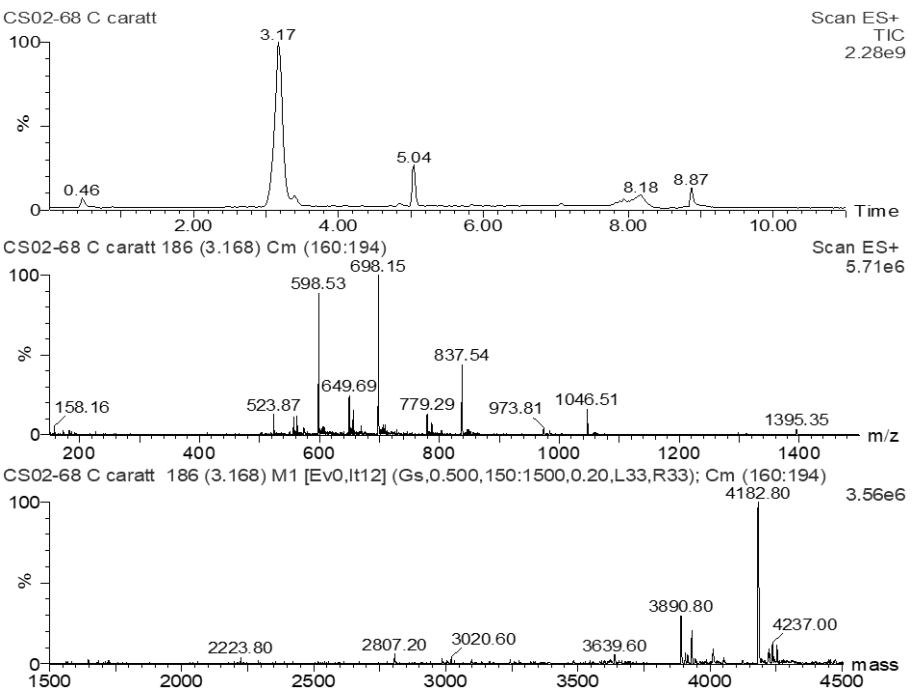



***Figure 2.A47.*** *UPLC-MS of CS02-68 C: chromatogram, mass spectrum and deconvoluted spectrum.*

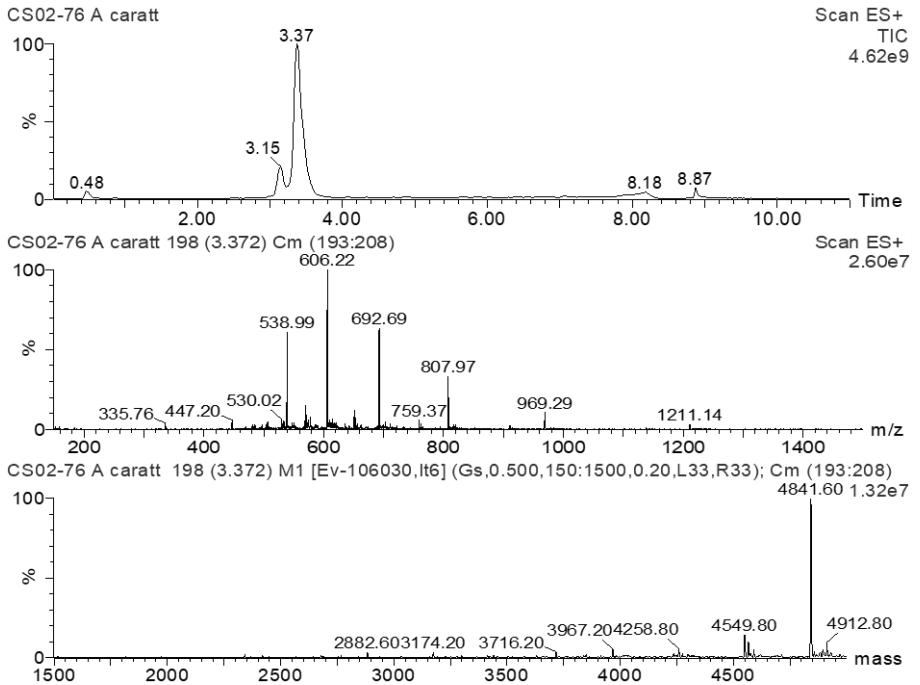



***Figure 2.A48.*** *UPLC-MS of CS02-76 A: chromatogram, mass spectrum and deconvoluted spectrum.*

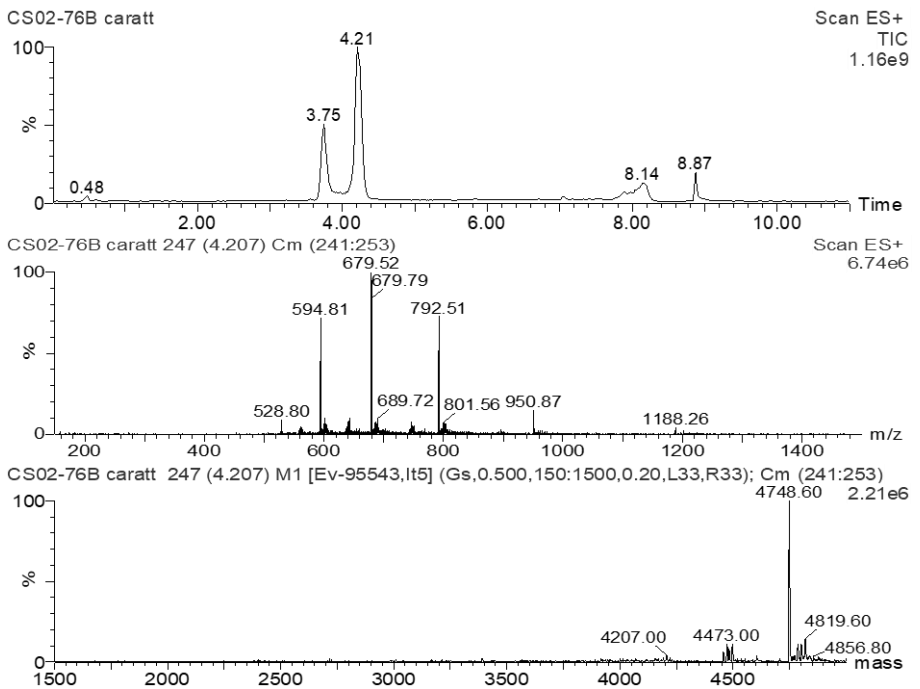



***Figure 2.A49.*** *UPLC-MS of CS02-76 B: chromatogram, mass spectrum and deconvoluted spectrum.*

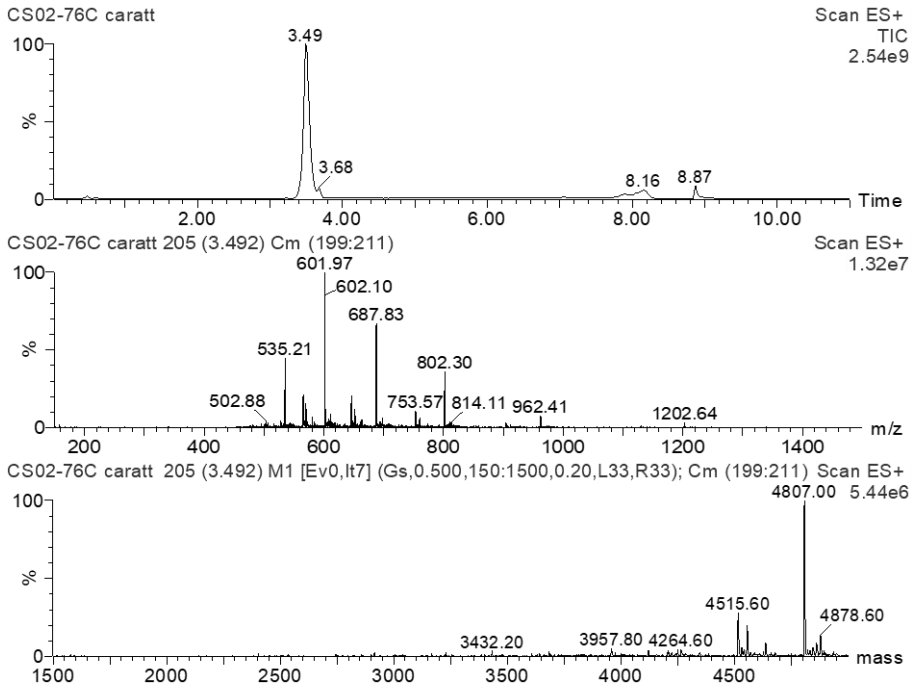



***Figure 2.A50.*** *UPLC-MS of CS02-76 C: chromatogram, mass spectrum and deconvoluted spectrum.*

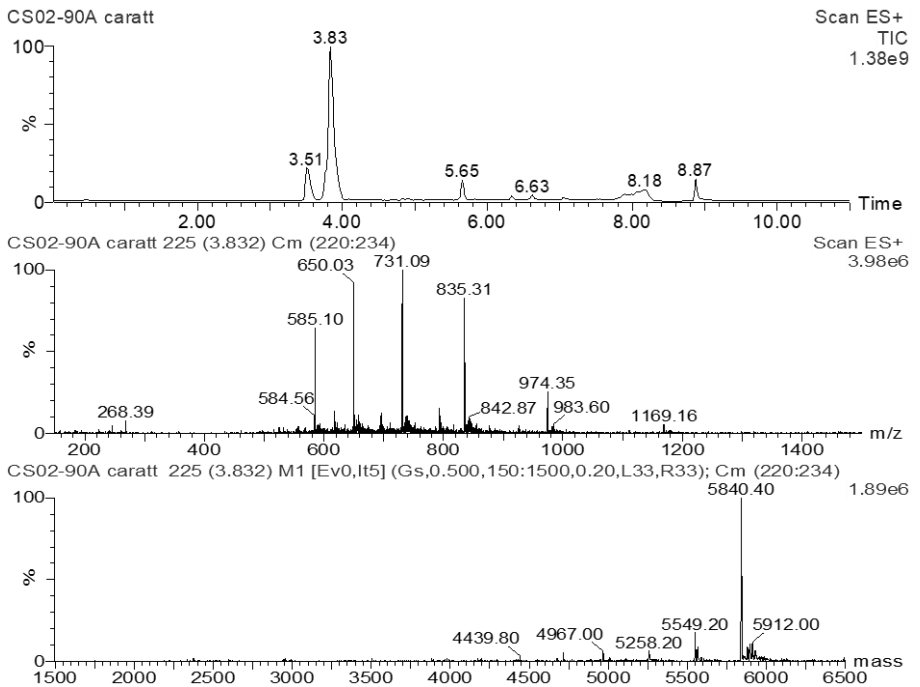



***Figure 2.A51.*** *UPLC-MS of CS02-90 A: chromatogram, mass spectrum and deconvoluted spectrum.*

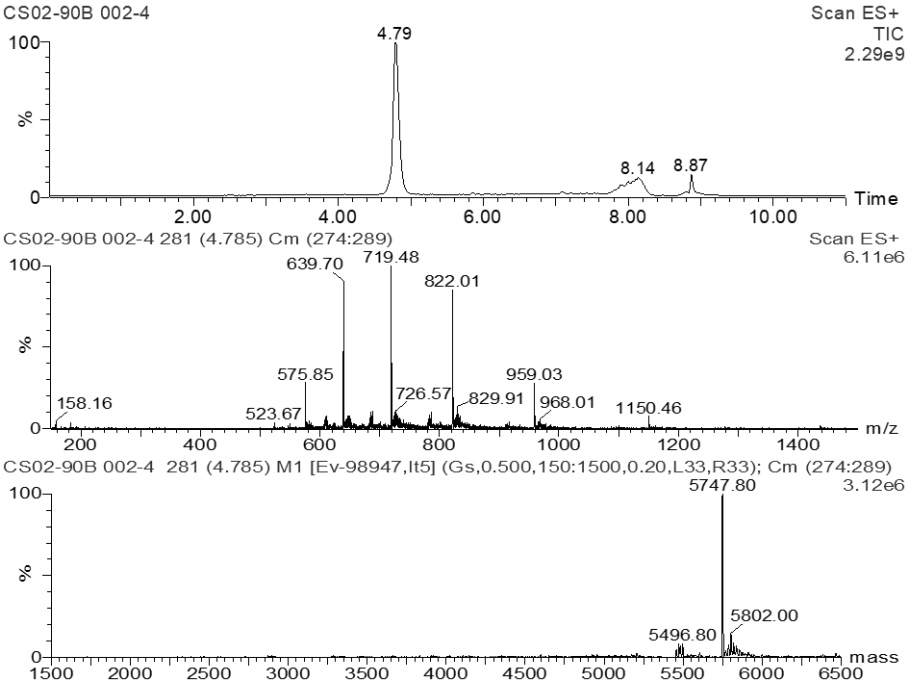



***Figure 2.A52.*** *UPLC-MS of CS02-90 B: chromatogram, mass spectrum and deconvoluted spectrum.*

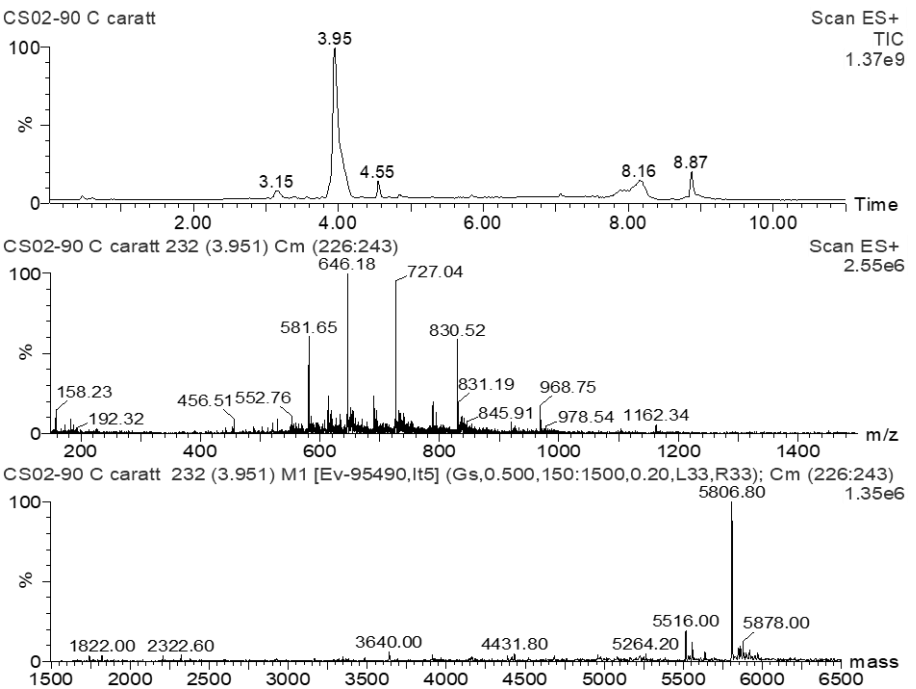



***Figure 2.A53.*** *UPLC-MS of CS02-90 C: chromatogram, mass spectrum and deconvoluted spectrum.*

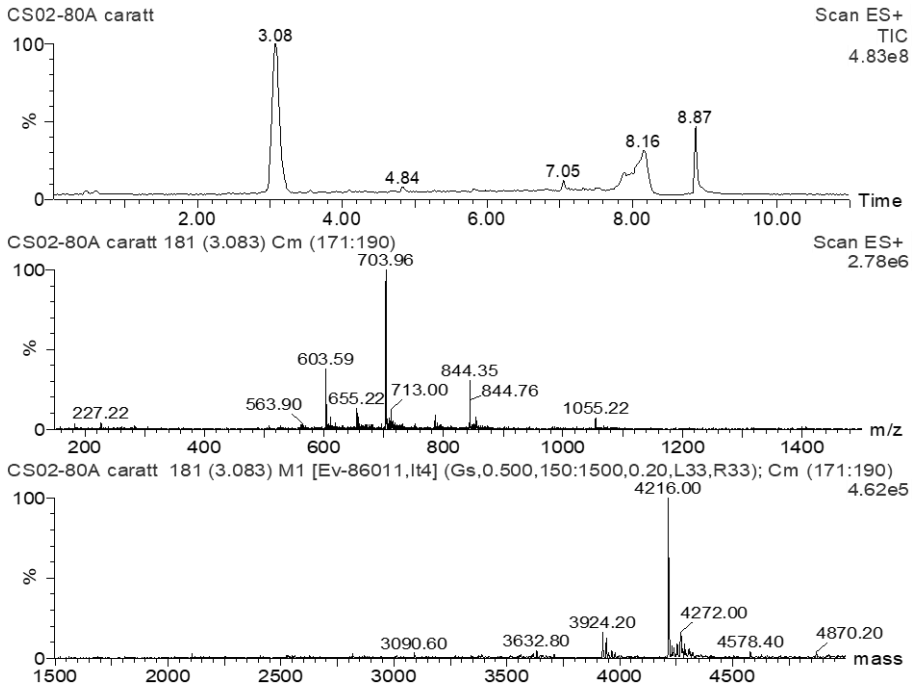



***Figure 2.A54.*** *UPLC-MS of CS02-80 A: chromatogram, mass spectrum and deconvoluted spectrum.*

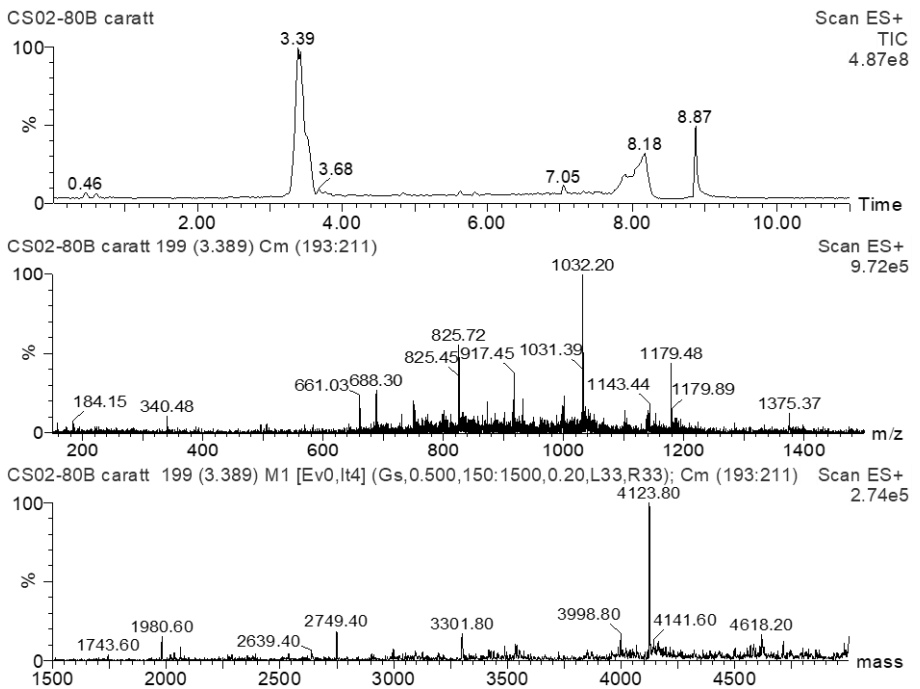



***Figure 2.A55.*** *UPLC-MS of CS02-80 B: chromatogram, mass spectrum and deconvoluted spectrum.*

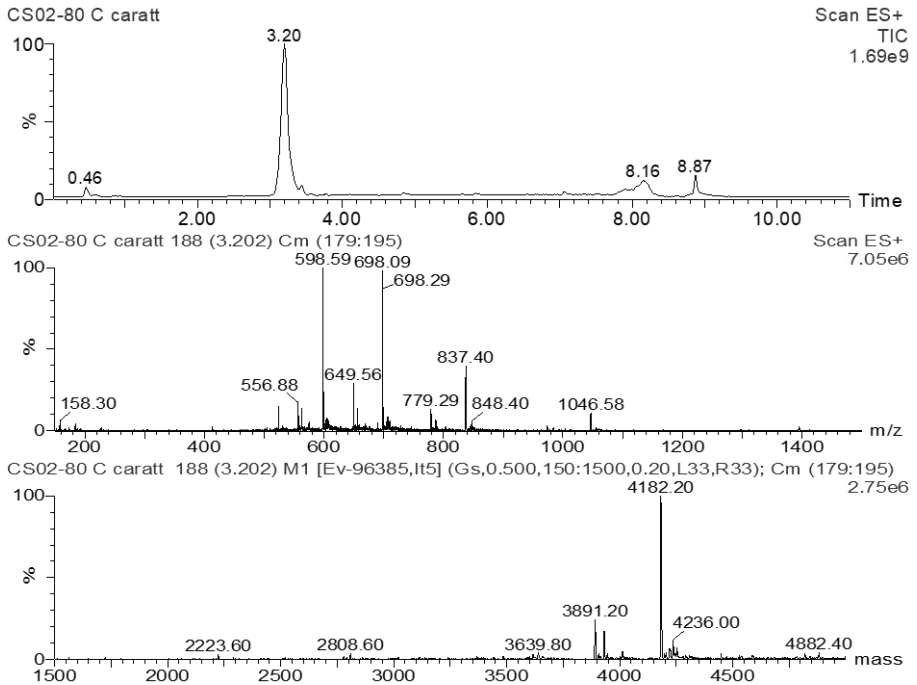



***Figure 2.A56.*** *UPLC-MS of CS02-80 C: chromatogram, mass spectrum and deconvoluted spectrum.*

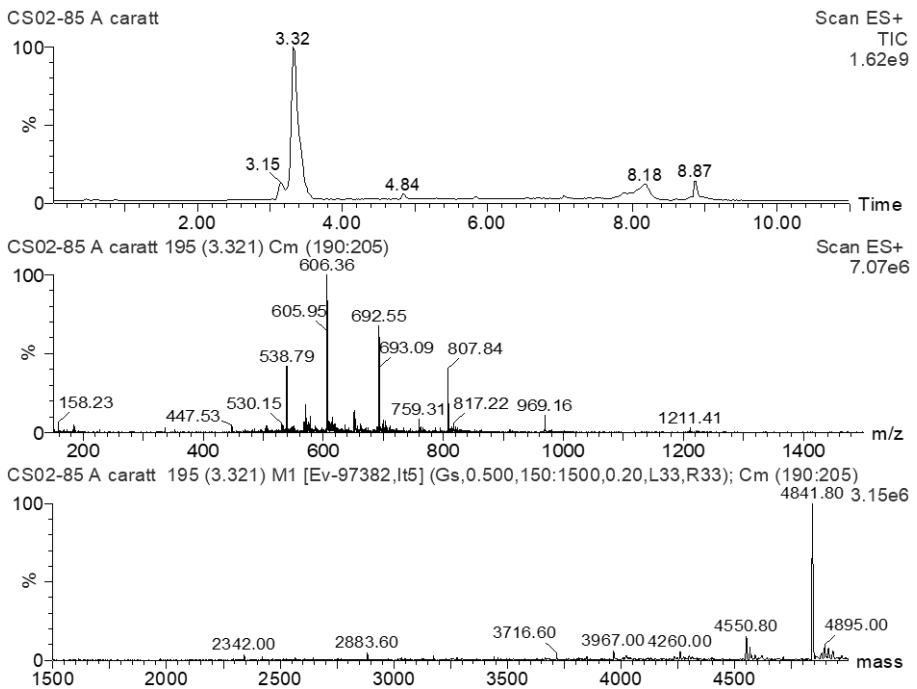



***Figure 2.A57.*** *UPLC-MS of CS02-85 A: chromatogram, mass spectrum and deconvoluted spectrum.*

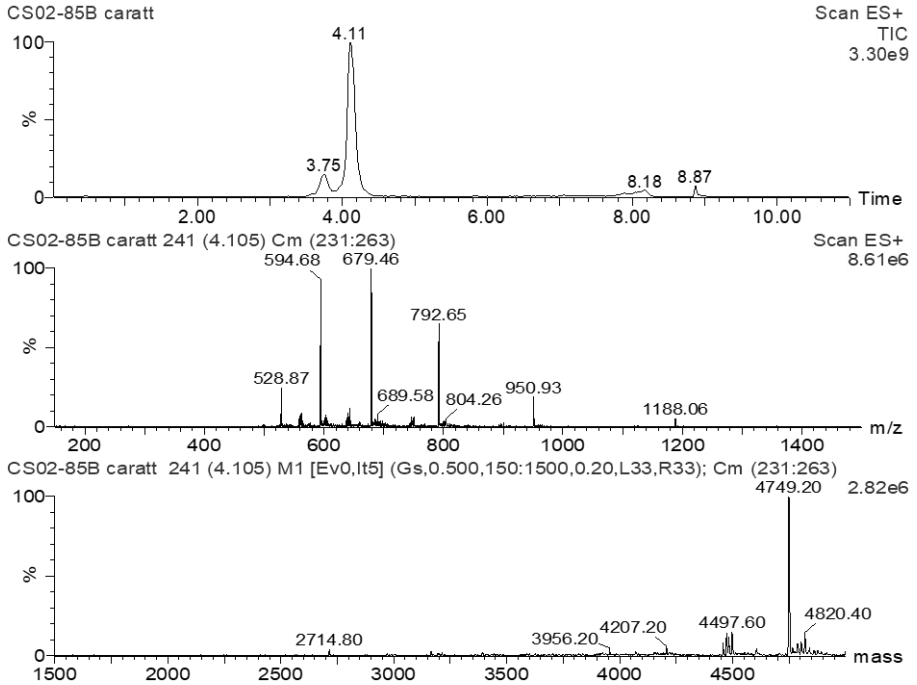



***Figure 2.A58.*** *UPLC-MS of CS02-85 B: chromatogram, mass spectrum and deconvoluted spectrum.*

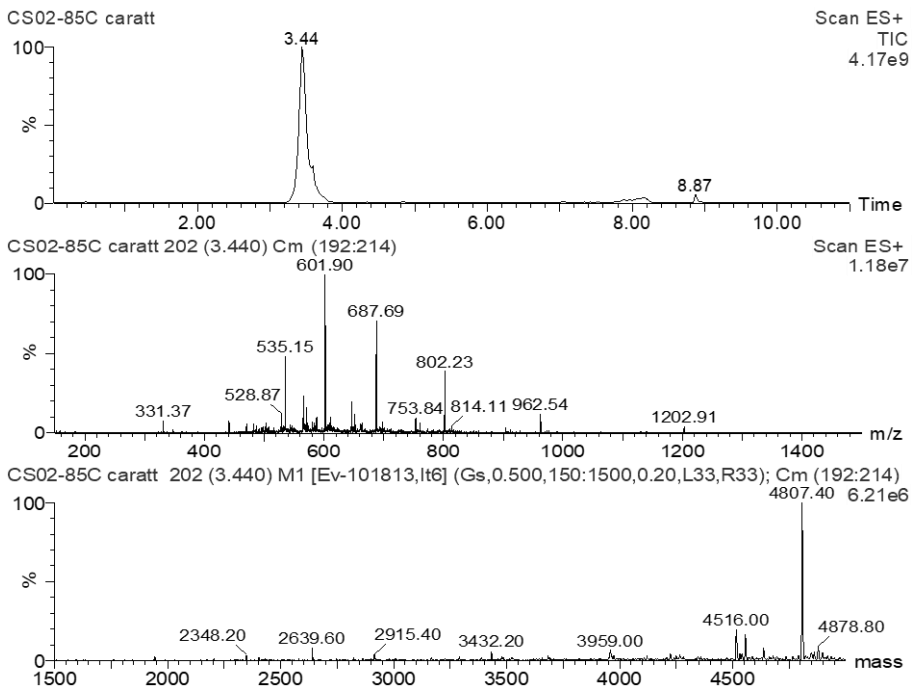



***Figure 2.A59.*** *UPLC-MS of CS02-85 C: chromatogram, mass spectrum and deconvoluted spectrum.*

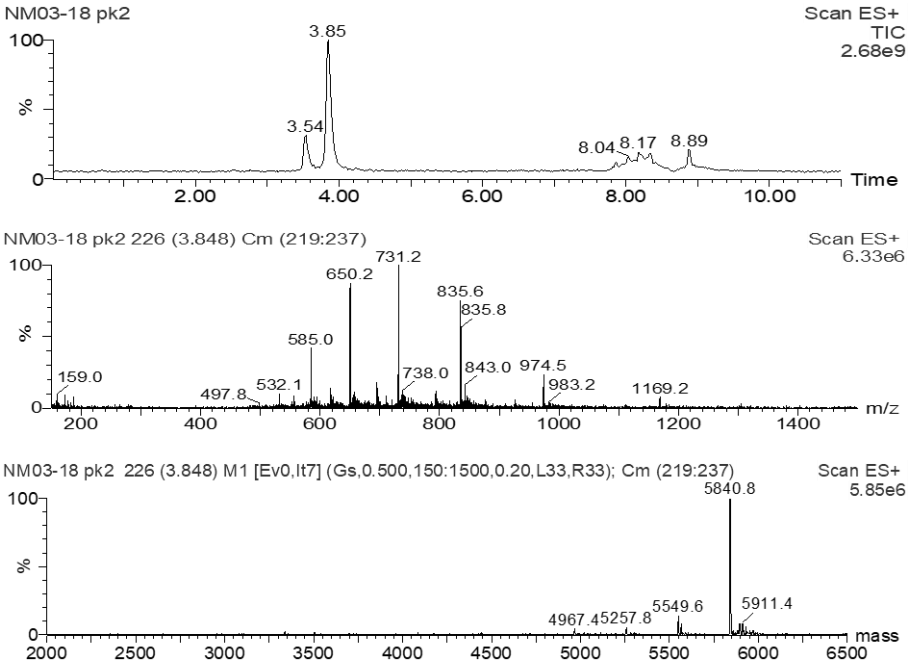



***Figure 2.A60.*** *UPLC-MS of NM03-18: chromatogram, mass spectrum and deconvoluted spectrum.*

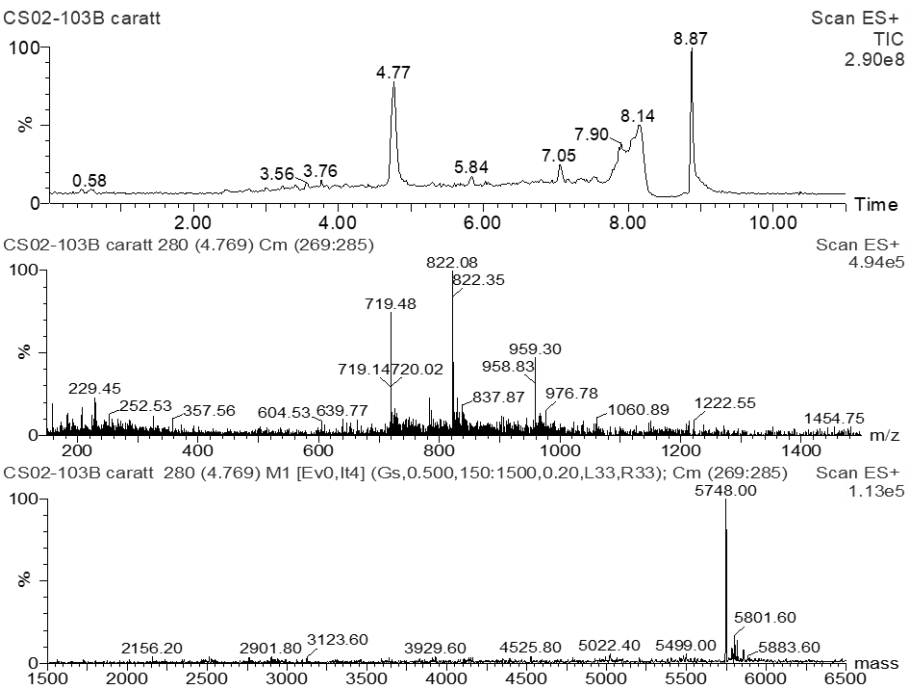



***Figure 2.A61.*** *UPLC-MS of CS02-103 B: chromatogram, mass spectrum and deconvoluted spectrum.*

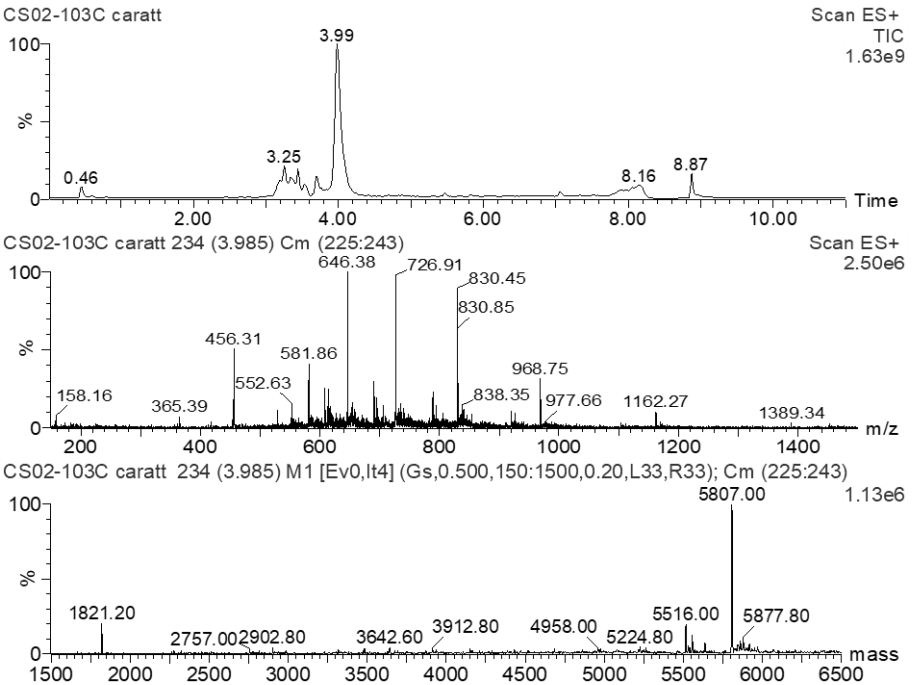



***Figure 2.A62.*** *UPLC-MS of CS02-103 C: chromatogram, mass spectrum and deconvoluted spectrum.*

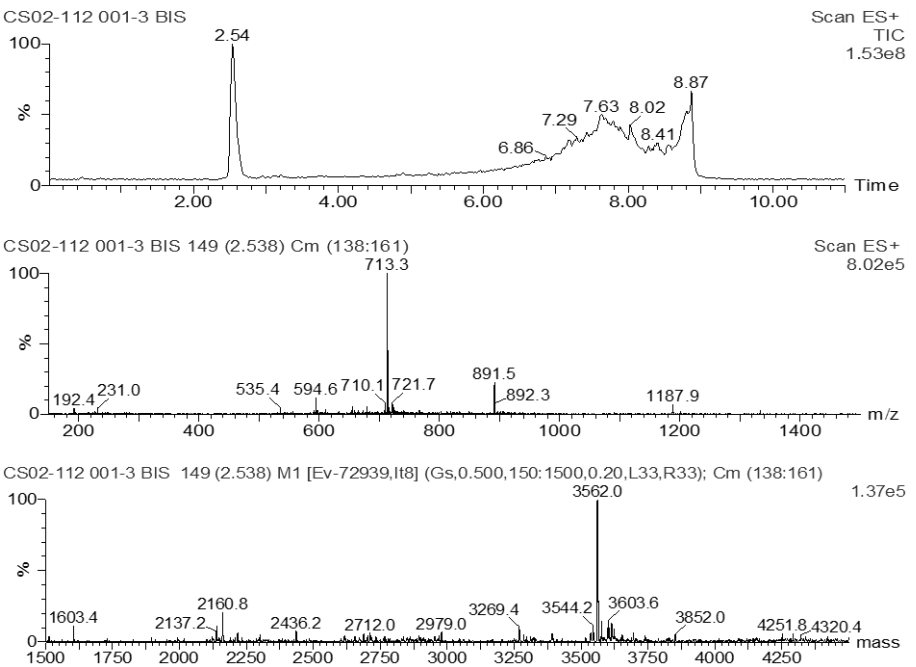



***Figure 2.A63.*** *UPLC-MS of PNA 1: chromatogram, mass spectrum and deconvoluted spectrum.*

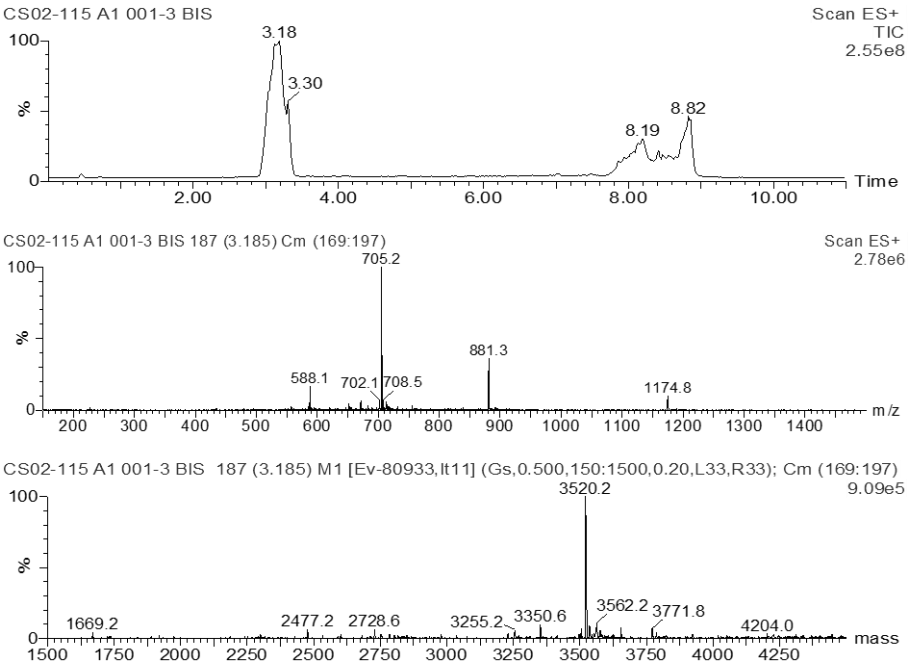



***Figure 2.A64.*** *UPLC-MS of PNA 2: chromatogram, mass spectrum and deconvoluted spectrum.*

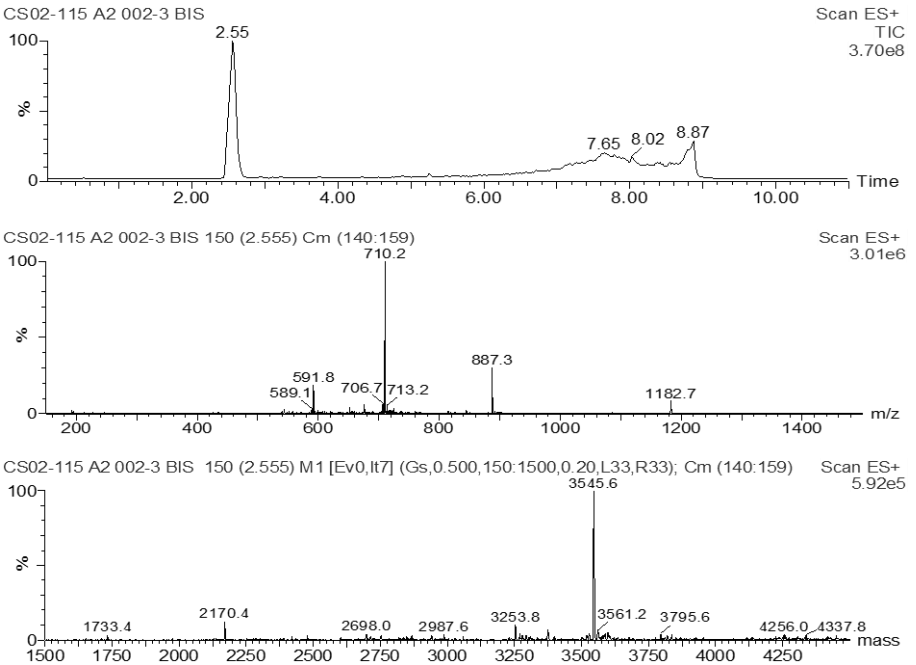



***Figure 2.A65.*** *UPLC-MS of PNA 3: chromatogram, mass spectrum and deconvoluted spectrum.*

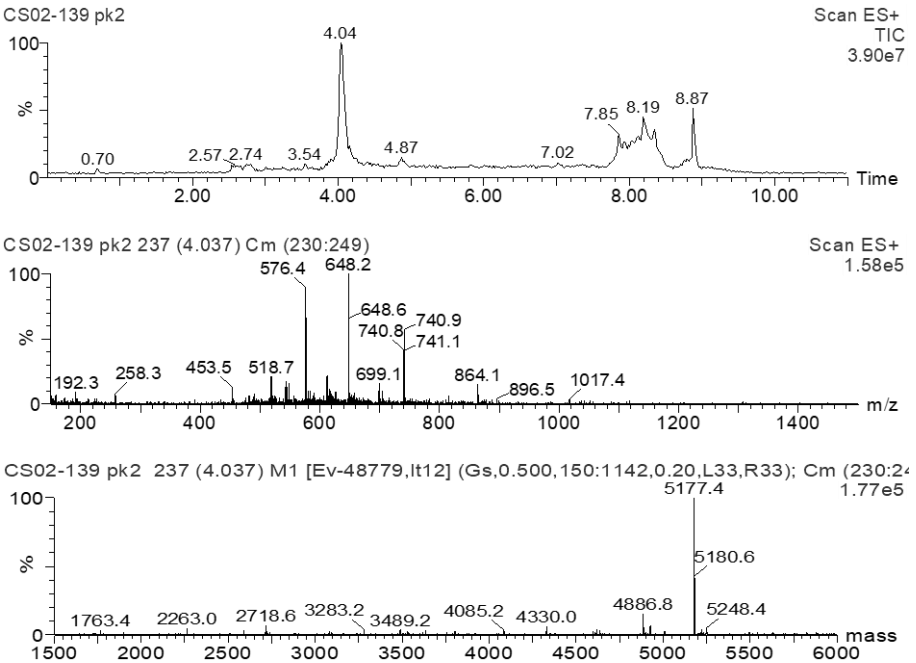



***Figure 2.A66.*** *UPLC-MS of CS02-139 pk2: chromatogram, mass spectrum and deconvoluted spectrum.*

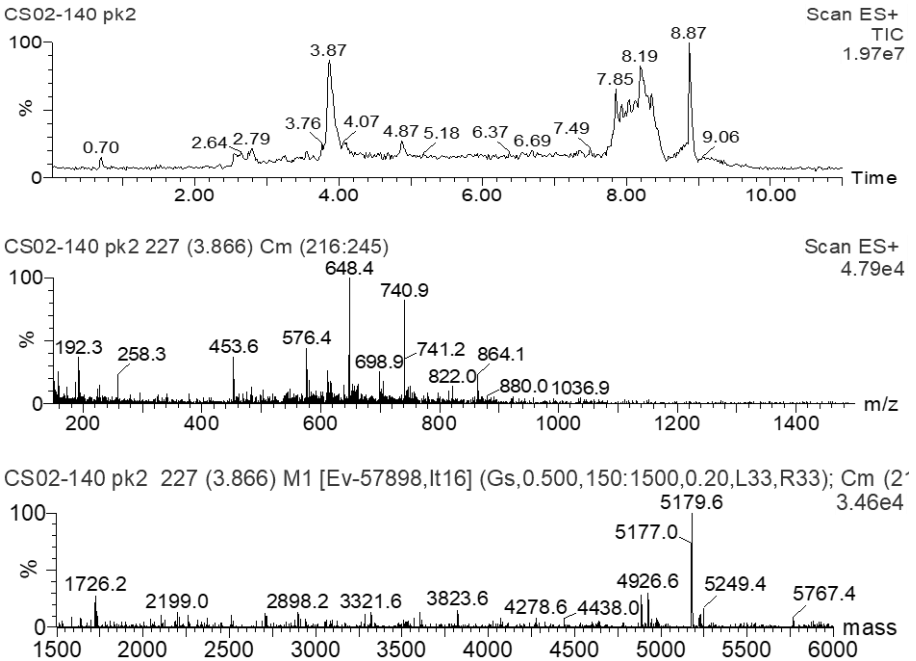



***Figure 2.A67.*** *UPLC-MS of CS02-140 pk2: chromatogram, mass spectrum and deconvoluted spectrum.*

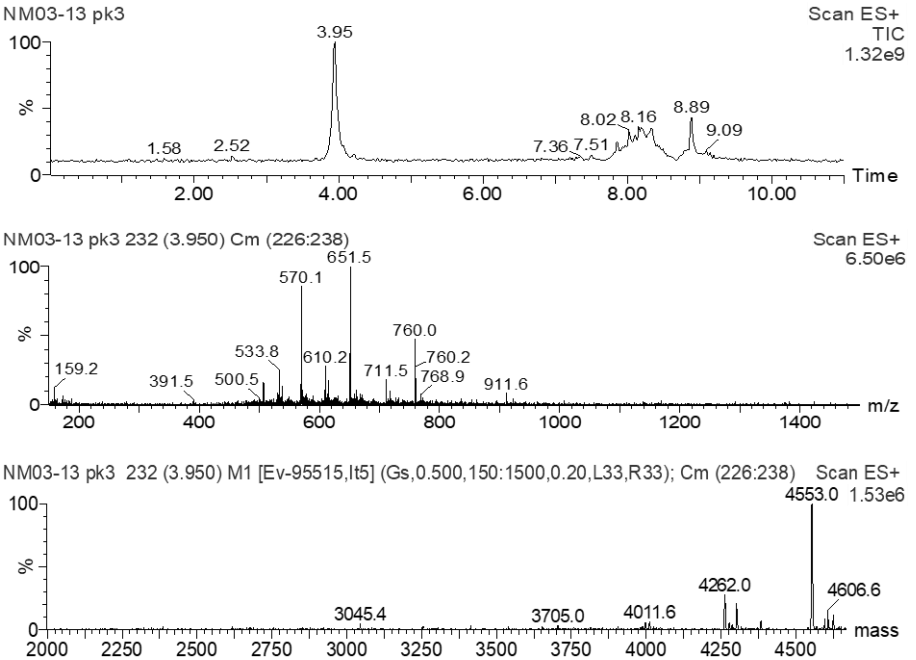



***Figure 2.A68.*** *UPLC-MS of NM03-13 pk3: chromatogram, mass spectrum and deconvoluted spectrum.*

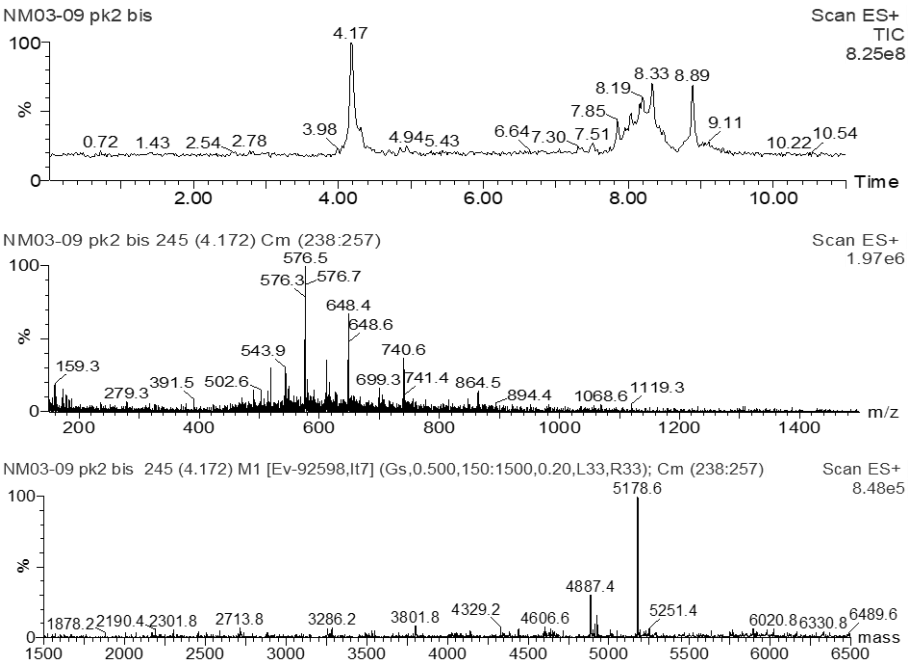



***Figure 2.A69.*** *UPLC-MS of NM03-09 pk2: chromatogram, mass spectrum and deconvoluted spectrum.*

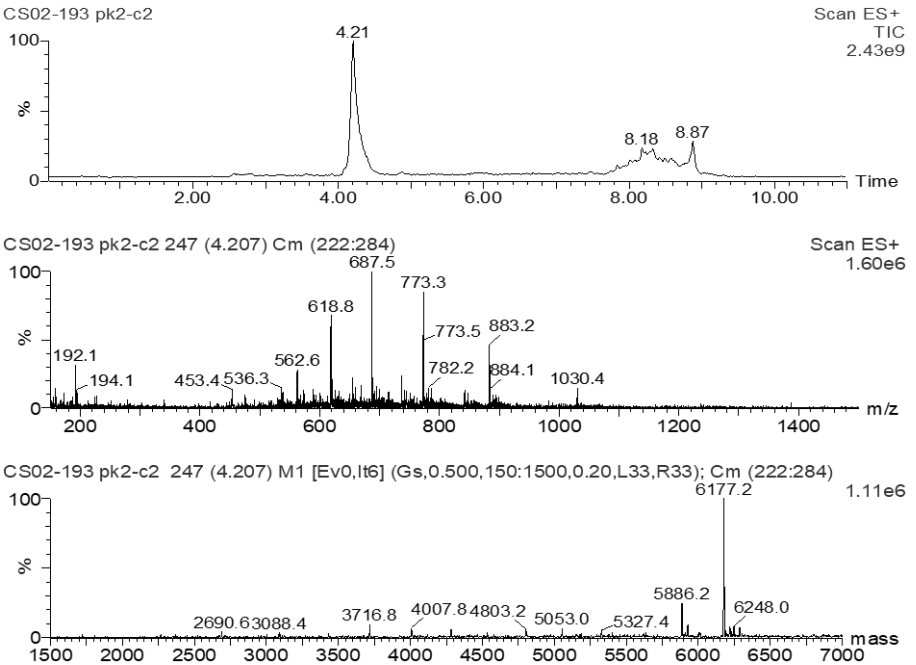



***Figure 2.A70.*** *UPLC-MS of CS02-193 pk2: chromatogram, mass spectrum and deconvoluted spectrum.*

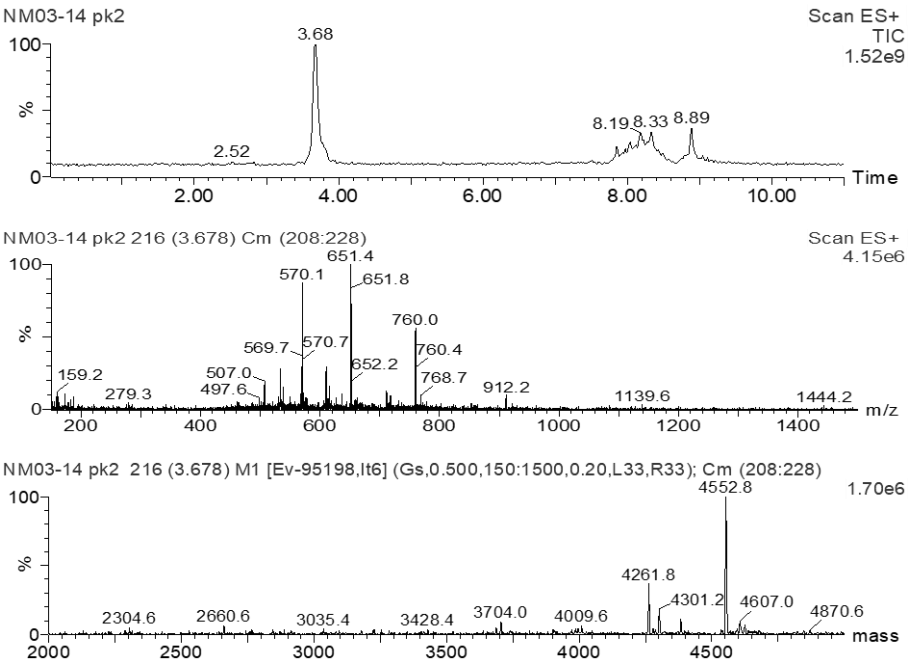



***Figure 2.A71.*** *UPLC-MS of NM03-14 pk2: chromatogram, mass spectrum and deconvoluted spectrum.*

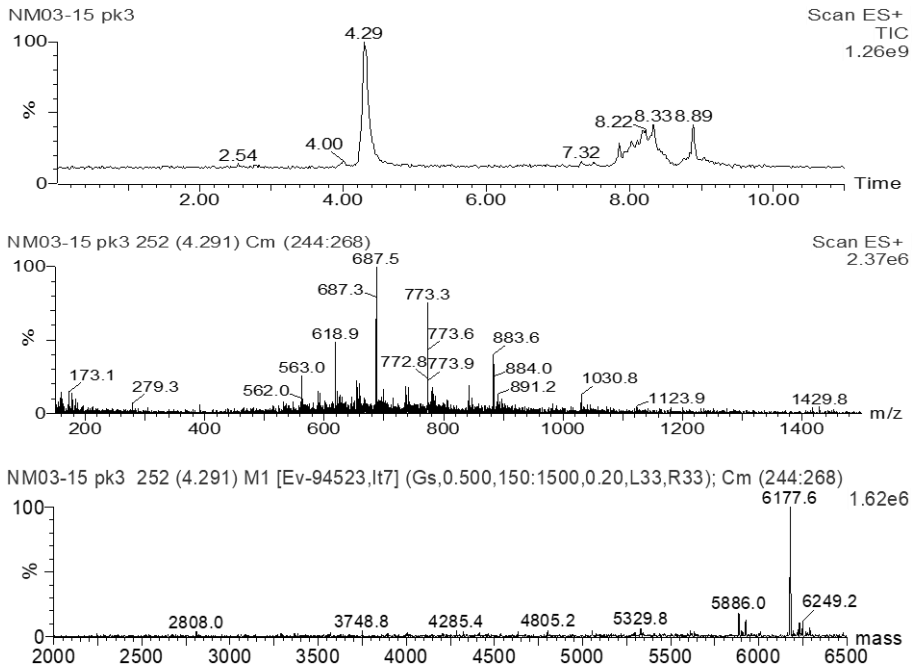



***Figure 2.A72.*** *UPLC-MS of NM03-15 pk3: chromatogram, mass spectrum and deconvoluted spectrum.*

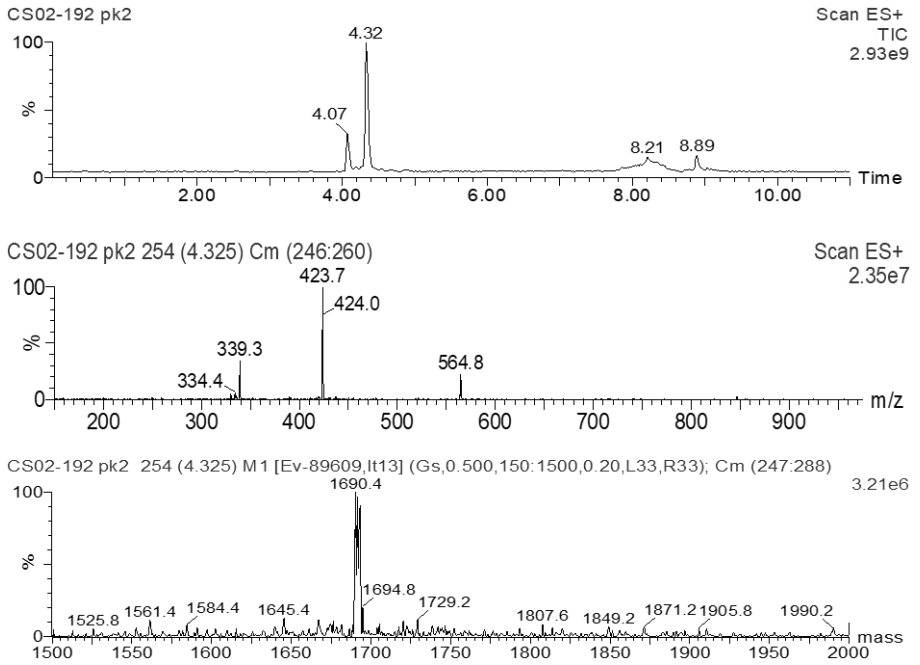



***Figure 2.A73.*** *UPLC-MS of CS02-192 pk2: chromatogram, mass spectrum and deconvoluted spectrum.*

## 2.7 References


1. M. Tegoni, F. Yu, M. Bersellini, J.E. Penner-Hahn, V.L. Pecoraro. Designing a functional type 2 copper center that has nitrite reductase activity within α-helical coiled coils. *Proc Natl Acad Sci U S A* **2012**, 109 (52), 21234–21239.
2. T. Trinh, D. Saliba, C. Liao, et al. "Printing" DNA Strand Patterns on Small Molecules with Control of Valency, Directionality, and Sequence. *Angewandte Chemie - International Edition* **2019**, 58 (10), 3042–3047.
3. L. PAULING, R.B. COREY. Compound Helical Configurations of Polypeptide Chains: Structure of Proteins of the α-Keratin Type. *Nature* **1953**, 171 (4341), 59–61.
4. F.H.C. Crick. The packing of α-helices: simple coiled-coils. *Acta Crystallogr* **1953**, 6 (8–9), 689–697.
5. L. Truebestein, T.A. Leonard. Coiled-coils: The long and short of it. *BioEssays* **2016**, 38 (9), 903–916.
6. D.N. Woolfson. Understanding a protein fold: The physics, chemistry, and biology of α-helical coiled coils. *Journal of Biological Chemistry* **2023**, 299 (4).
7. J.M. Mason, K.M. Arndt. Coiled Coil Domains: Stability, Specificity, and Biological Implications. *ChemBioChem* **2004**, 5 (2), 170–176.
8. A. Lupas. Coiled coils: new structures and new functions. *Trends Biochem Sci* **1996**, 21 (10), 375–382.
9. D.N. Woolfson. Coiled-Coil Design: Updated and Upgraded. In *Fibrous Proteins: Structures and Mechanisms*; Parry, D. A. D., Squire, J. M., Eds.; Springer International Publishing, Cham, **2017**; pp 35–61.
10. P.-S. Huang, S.E. Boyken, D. Baker. The coming of age of de novo protein design. *Nature* **2016**, 537 (7620), 320–327.
11. F. Yu, V.M. Cangelosi, M.L. Zastrow, et al. Protein Design: Toward Functional Metalloenzymes. *Chem Rev* **2014**, 114 (7), 3495–3578.
12. K. Madica, J.K. Lakshmi, S. Madhu, et al. Dimedone-Based Rigid Organic Scaffold for Organizing Symmetrical Helical Peptide Chains. *ChemistrySelect* **2019**, 4 (39), 11518–11522.
13. K.J. Stingley, B.A. Carpenter, K.M. Kean, M.L. Waters. Mismatched covalent and noncovalent templating leads to large coiled coil-templated macrocycles. *Chem Sci* **2023**, 14 (18), 4935–4944.
14. J.R. Litowski, R.S. Hodges. Designing heterodimeric two-stranded α-helical coiled-coils: the effect of chain length on protein folding, stability and specificity. *The Journal of Peptide Research* **2001**, 58 (6), 477–492.
15. D.N. Woolfson. The Design of Coiled-Coil Structures and Assemblies. In *Advances in Protein Chemistry*; Academic Press, **2005**; Vol. 70, pp 79–112.
16. L.K. Henchey, A.L. Jochim, P.S. Arora. Contemporary strategies for the stabilization of peptides in the α-helical conformation. *Curr Opin Chem Biol* **2008**, 12 (6), 692–697.
17. M.G. Wuo, A.B. Mahon, P.S. Arora. An Effective Strategy for Stabilizing Minimal Coiled Coil Mimetics. *J Am Chem Soc* **2015**, 137 (36), 11618–11621.
18. M.G. Wuo, S.H. Hong, A. Singh, P.S. Arora. Synthetic Control of Tertiary Helical Structures in Short Peptides. *J Am Chem Soc* **2018**, 140 (47), 16284–16290.

19. P.E. Dawson, S.B.H. Kent. Convenient total synthesis of a 4-helix template-assembled synthetic protein (TASP) molecule by chemoselective ligation. *J Am Chem Soc* **1993**, 115 (16), 7263–7266.
20. M. Mutter, P. Dumy, P. Garrouste, et al. Template assembled synthetic proteins (TASP) as functional mimetics of proteins. *Angewandte Chemie* **1996**, 35, 1482–1485.
21. O. Seitz. Templated chemistry for bioorganic synthesis and chemical biology. *Journal of Peptide Science* **2019**, 25 (7).
22. S. Madhu, K. Madica, R. Gonnade, G.J. Sanjayan. Synthesis of hexaphenylbenzene-based template assembled synthetic proteins. *Tetrahedron Lett* **2019**, 60 (33).
23. N.C. Seeman. Nucleic acid junctions and lattices. *J Theor Biol* **1982**, 99 (2), 237–247.
24. D. Saliba, T. Trinh, C. Lachance-Brais, et al. Asymmetric patterning drives the folding of a tripodal DNA nanotweezer. *Chem Sci* **2021**, 13, 74–80.
25. N. Farag, G. Ercolani, E. Del Grosso, F. Ricci. DNA Tile Self-Assembly Guided by Base Excision Repair Enzymes. *Angewandte Chemie - International Edition* **2022**.
26. T. Sabir, A. Toulmin, L. Ma, et al. Branchpoint expansion in a fully complementary three-way DNA junction. *J Am Chem Soc* **2012**, 134 (14), 6280–6285.
27. N. Farag, G. Ercolani, E. Del Grosso, F. Ricci. DNA Tile Self-Assembly Guided by Base Excision Repair Enzymes. *Angewandte Chemie - International Edition* **2022**.
28. D. Liu, M. Wang, Z. Deng, R. Walulu, C. Mao. Tensegrity: Construction of Rigid DNA Triangles with Flexible Four-Arm DNA Junctions. *J Am Chem Soc* **2004**, 126 (8), 2324–2325.
29. P.W.K. Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature* **2006**, 440 (7082), 297–302.
30. X. Wang, H. Jun, M. Bathe. Programming 2D Supramolecular Assemblies with Wireframe DNA Origami. *J Am Chem Soc* **2022**, 144 (10), 4403–4409.
31. T.G.W. Edwardson, K.M.M. Carneiro, C.K. McLaughlin, C.J. Serpell, H.F. Sleiman. Site-specific positioning of dendritic alkyl chains on DNA cages enables their geometry-dependent self-assembly. *Nat Chem* **2013**, 5 (10), 868–875.
32. T.G.W. Edwardson, K.M.M. Carneiro, C.J. Serpell, H.F. Sleiman. An efficient and modular route to sequence-defined polymers appended to DNA. *Angewandte Chemie - International Edition* **2014**, 53 (18), 4567–4571.
33. F.J. Rizzuto, M.D. Dore, M.G. Rafique, X. Luo, H.F. Sleiman. DNA Sequence and Length Dictate the Assembly of Nucleic Acid Block Copolymers. *J Am Chem Soc* **2022**, 144 (27), 12272–12279.
34. J.D. WATSON, F.H.C. CRICK. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, 171 (4356), 737–738.
35. N.C. Seeman, H.F. Sleiman. DNA nanotechnology. *Nat Rev Mater* **2017**, 3 (1), 17068.
36. T. Trinh, C. Liao, V. Toader, et al. DNA-imprinted polymer nanoparticles with monodispersity and prescribed DNA-strand patterns. *Nat Chem* **2018**, 10 (2), 184–192.
37. A. Buchberger, C.R. Simmons, N.E. Fahmi, R. Freeman, N. Stephanopoulos. Hierarchical Assembly of Nucleic Acid/Coiled-Coil Peptide Nanostructures. *J Am Chem Soc* **2020**, 142 (3), 1406–1416.

38. A. Buchberger, M. Al-Amin, C.R. Simmons, N. Stephanopoulos. Self-Assembly of Hybrid Peptide–DNA Nanostructures using Homotrimeric Coiled-Coil/Nucleic Acid Building Blocks. *ChemBioChem* **2023**, 24 (17), e202300223.
39. C. Lou, N.J. Christensen, M.C. Martos-Maldonado, et al. Folding Topology of a Short Coiled-Coil Peptide Structure Templated by an Oligonucleotide Triplex. *Chemistry – A European Journal* **2017**, 23 (39), 9297–9305.
40. C. Lou, J.T. Boesen, N.J. Christensen, et al. Self-Assembly of DNA–Peptide Supermolecules: Coiled-Coil Peptide Structures Templated by d-DNA and l-DNA Triplexes Exhibit Chirality-Independent but Orientation-Dependent Stabilizing Cooperativity. *Chemistry – A European Journal* **2020**, 26 (25), 5676–5684.
41. T.B.J. Pinter, K.J. Koebke, V.L. Pecoraro. Catalysis and Electron Transfer in De Novo Designed Helical Scaffolds. *Angewandte Chemie International Edition* **2020**, 59 (20), 7678–7699.
42. O. Iranzo, D. Ghosh, V.L. Pecoraro. Assessing the integrity of designed homomeric parallel three-stranded coiled coils in the presence of metal ions. *Inorg Chem* **2006**, 45 (25), 9959–9973.
43. M. Tegoni, F. Yu, M. Bersellini, J.E. Penner-Hahn, V.L. Pecoraro. Designing a functional type 2 copper center that has nitrite reductase activity within α-helical coiled coils. *Proceedings of the National Academy of Sciences* **2012**, 109 (52), 21234–21239.
44. A.G. Tebo, V.L. Pecoraro. Artificial metalloenzymes derived from three-helix bundles. *Curr Opin Chem Biol* **2015**, 25, 65–70.
45. J.S. Plegaria, M. Duca, C. Tard, et al. De Novo Design and Characterization of Copper Metallopeptides Inspired by Native Cupredoxins. *Inorg Chem* **2015**, 54 (19), 9470–9482.
46. J.S. Plegaria, C. Herrero, A. Quaranta, V.L. Pecoraro. Electron transfer activity of a de novo designed copper center in a three-helix bundle fold. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2016**, 1857 (5), 522–530.
47. M.C. Manning, R.W. Woody. Theoretical CD studies of polypeptide helices: Examination of important electronic and geometric factors. *Biopolymers* **1991**, 31 (5), 569–586.
48. C. Toniolo, A. Polese, F. Formaggio, M. Crisma, J. Kamphuis. Circular Dichroism Spectrum of a Peptide 310-Helix. *J Am Chem Soc* **1996**, 118 (11), 2744–2745.
49. A. Micsonai, F. Wien, L. Kernya, et al. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences* **2015**, 112 (24), E3095–E3103.
50. S.L. Beaucage, M.H. Caruthers. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett* **1981**, 22 (20), 1859–1862.
51. A. Manicardi, E. Cadoni, A. Madder. Visible-light triggered templated ligation on surface using furan-modified PNAs. *Chem Sci* **2020**, 11 (43), 11729–11739.

# *Chapter 3*

## *3.0 Heterotrimeric Coiled-Coil structures through PNA:PNA-Peptide interaction*

### 3.1 Abstract

In this Chapter is described the synthesis and optimization of a tri-functionalized rigid molecule to use it as a scaffold to establish a platform for the assembly of heterotrimeric coiled-coil structures and designed to be compatible with peptide solid phase synthesis.

Compared to the previous chapter, where the assembly was formed through DNA:PNA interactions, in this case the final complex will be formed through PNA:PNA-peptides interactions. Thanks to the three different functional groups (azide, carboxylic acid, and a Fmoc-protected amino function) featured on this rigid small molecule, it is possible to grow three different PNA sequences branched on this central core (*Figure 3.1*), in direction from $C_{term}$ to $N_{term}$.

At the same time, a study to understand the minimal PNA length, necessary to generate a PNA:PNA duplex able to operate as driving force for the formation of trimeric coiled-coil structures through PNA:PNA-peptides, was carried out eval of PNA:PNA complexes melting temperature.

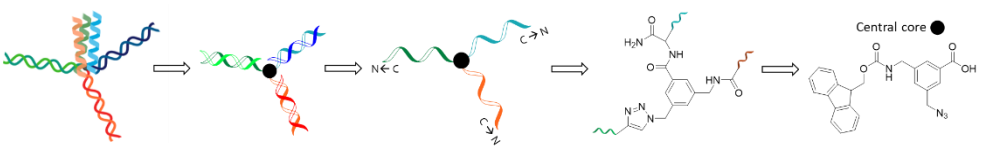



***Figure 3.1.*** *Schematic representation of a platform for the assembly of trimeric coiled-coil structure based on PNAs.*

## 3.2 Introduction

Since the use of a DNA template to generate heterotrimeric coiled-coil system requires the use of different strategies and chemistry (phopshoramidite synthesis for DNA and peptide synthesis for PNA), it is interesting to establish if a similar approach reported in *Chapter 2* can be performed using a trimeric Peptide Nucleic Acid (PNA)[1] as a template.

Unlike natural nucleic acids, which are susceptible to degradation by nucleases due to their phosphate backbone, the unnatural backbone PNA oligomers exhibit remarkable chemical and biochemical stability renders them impervious to nuclease and protease activity, as well as improves chemical stability.[2]

In many applications, PNA strands have been used for their capacity to bind DNA or RNA oligomers in either parallel or antiparallel orientations, forming very stable PNA:DNA or PNA:RNA duplexes.[3,4] However, PNAs are also capable of binding complementary PNA sequences both parallel and antiparallel, leading to the formation of stable duplexes that exhibit double helical structures akin to DNA or RNA duplexes. Consequently, this demonstrates that the sugar backbone found in nucleic acids (NA) is not a requisite feature for the creation of double helical architectures.[5] Complexes built upon nucleic acids are generally less stable compared to the corresponding PNA:NA or PNA:PNA duplexes. This heightened stability arises from the fact that the neutral backbone of PNA avoids electrostatic repulsion.[6] Overall, the thermal stability of these duplexes typically follows the order: antiparallel PNA:PNA > PNA:RNA > PNA:DNA ~ parallel PNA:PNA > DNA:DNA.

Over the years, investigations into the stability of duplexes formed by PNAs have primarily involved measuring their melting temperatures (*Tm*). An influential factor affecting the melting temperatures of both nucleic acid and PNA duplexes is the composition of the chains. Specifically, like in natural NA, a higher proportion of G-C base pairs demands more energy to dissociate the double helices due to the higher number of formed hydrogen bonds that can be formed.[7]

Peptide nucleic acid (PNA) oligomers are produced through solid-phase peptide synthesis (SPPS), a synthetic technique.[8] This method, coupled with the advancement and exploration of various protective groups[9], each with its unique reactivity, has played a crucial role in establishing PNAs as fundamental building blocks in the construction of PNA-based nanostructures. In addition, PNAs are excellent candidates for building nanostructures as they can be easily conjugated with small molecules, peptides, and proteins. They can tolerate modifications on both the backbone and the nucleobases, thereby adding further functionalities,

Guiding the organized assembly of PNAs, these nanostructures are formed spontaneously through noncovalent interactions, including hydrogen bonds, hydrophobic interactions, and van der Waals forces, resulting in precise structural organization and thermodynamic stability.[10] For these properties PNAs have been used in different works aiming to form self-assembled structures or nanostructures. For example, self-assembly of nanofiber structures based on gamma-modified peptide nucleic acid (γ-PNA) in organic solvents were developed.[11]

Supramolecular drugs can be obtained using PNA-mediated self-assembly. In fact, ligands for specific protein targets can be built by non-covalent self-assembly of complementary PNA-ligand conjugates, offering the possibility of multiple combinations.[12] Labelling with sensitive molecules or radioactive metal ions can be performed using these scaffolds. In fact, Gasser and coworkers have reported a complex formed by a PNA strand functionalized with a specific antibody and its corresponding radiolabeled complementary PNA sequence ($^{99m}$Tc-labeled PNA), demonstrating *in vivo* the ability of these two modified-PNAs to recognize each other.[13]

In addition, self-assembly of microparticles was reported by Ly and coworkers using complementary PNAs with appropriate stereochemistry.[14]

A very accurate prediction model is available for prediction of PNA:DNA melting temperature,[15] and so applications implying this type of structure can be tuned to the appropriate stability. However, a similar model is not available for PNA:PNA structures and few data are available for parallel PNA:PNA duplexes.

## 3.3 Results and discussion

### 3.3.1 Synthesis of a tri-functionalized rigid core molecule

To obtain three different PNA:PNA duplexes able to enforce the formation of the trimeric coiled-coil system through PNA:PNA-peptide interaction, a small tri-functionalized molecule was synthetized. This central core should ideally be: a) asymmetric, therefore with three orthogonal reactivities, to permit the growth of three different PNA sequences; b) fairly small to ensure the spatial proximity of the peptide domains; c) designed to generate equal distance between the three various functional groups to avoid steric hindrance to facilitate the attachment of chains; d) compatible with the solid phase peptide synthesis (SPPS); e) lacking of chiral centers. Therefore, we designed a benzene-core-based synthon **1** (*Figure 3.2*) which can be obtained as shown in the retrosynthesis reported in *Figure 3.2*.

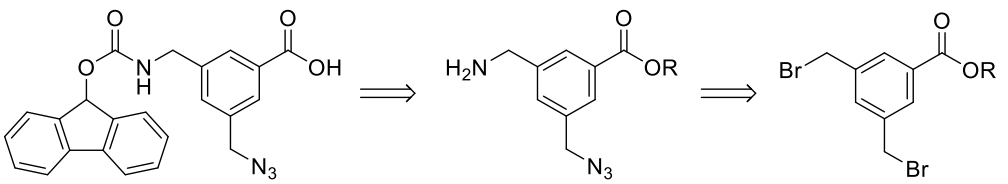



***Figure 3.2. R****etrosynthesis of the target 3-((9-fluorenylmethoxycarbonyl)amino methyl)-5-(azidomethyl)benzoic acid.*

It is possible to observe that the skeleton of the scaffold is a 1,3,5-substituted benzene derivative. The benzene core provides structural rigidity and the equal distance between the substituents was chosen to maximize the steric hindrance and, consequently, to ease the attachment of PNA sequences. This molecule was also selected for the lack of chiral centers.

Carboxylic acid, azide and Fmoc-protected amino function were the three functional groups chosen to obtain a molecule with orthogonal reactivities as a rigid center between three different PNA sequences, suitable for SPPS.

Fmoc protection of amine function, hydrolysis of ester, reduction of azide function and $S_N2$ nucleophilic substitution are the principal reactions of the retrosynthesis to achieve this desired central core. In specific, the Fmoc-protected target **1**, suitable for the solid phase protocols, can be obtained by protection of the amine group and by the hydrolysis of its carboxylic acid derivative, corresponding to molecule **2**. Compound **2** can be achieved through a selective mono-reduction (Staudinger reduction) after the installation through a nucleophilic substitution of two azido groups on a bis-bromo-containing starting material **3**, methyl 3,5-bis(bromomethyl) benzoate, which is commercially available.

The principal steps of the tri-functionalized rigid core molecule (**1**) synthetic protocol are $Sn_2$ nucleophilic substitution, Staudinger reduction of azido group, hydrolysis of ester, and Fmoc protection of amine function (*Figure 3.3*).

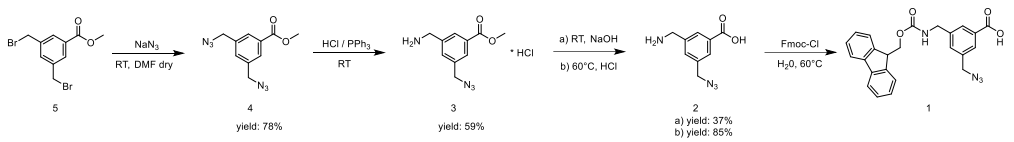



***Figure 3.3.*** *Tri-functionalized rigid core molecule (**1**) synthetic protocol.*

Synthetic steps to obtain compound **1**

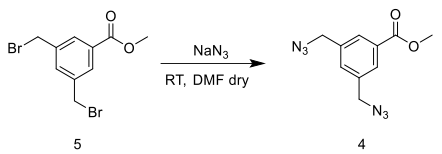



***Figure 3.4.*** *Synthesis of methyl 3,5-bis(azidomethyl)benzoate.*

Methyl 3,5-bis(azidomethyl)benzoate **4** can be obtained from the commercially available methyl 3,5-bis(bromomethyl)benzoate **5** via nucleophilic substitution in presence of sodium azide (*Figure 3.4*).

After optimization steps, this reaction was carried out in DMF dry with 2.1 equivalents of sodium azide, overnight at room temperature to ensure the

formation of compound **4** and to prevent the presence of monosubstituted intermediate. A pure product in fairly good yield (78%) was obtained.

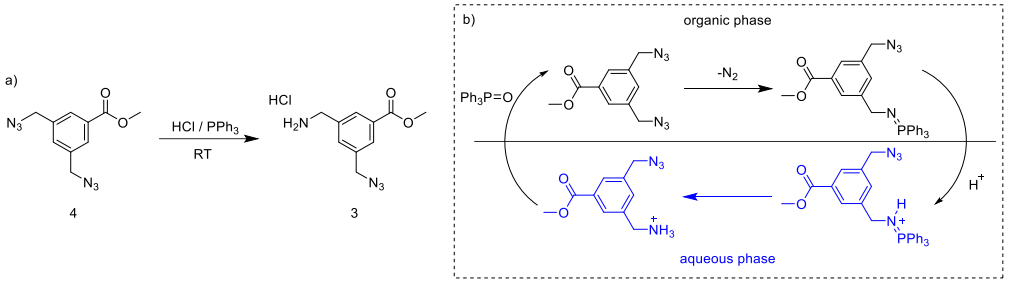



***Figure 3.5.*** *a) Synthesis of methyl 3-(aminomethyl)-5-(azidomethyl)benzoate. b) Mechanism of biphasic system.*

Methyl 3-(aminomethyl)-5-(azidomethyl)benzoate **3** can be obtained via Staudinger reduction of a single azido group present in compound **4** (*Figure 3.5a*). To achieve this, it is necessary to work in defect of triphenylphosphine and in a biphasic system containing an organic solvent and an acid aqueous solution (heterogeneous conditions, *Figure 3.5b*).

To optimize the drying of the aqueous phase and to check the formation of eventual degradation product (due to ester hydrolysis), the water phase was dried under reduced pressure, either with a rotavapor or freeze-dryer. The results (*Figure 3.6*) showed no difference between the two methods, at least in a small-scale (~100 mg).

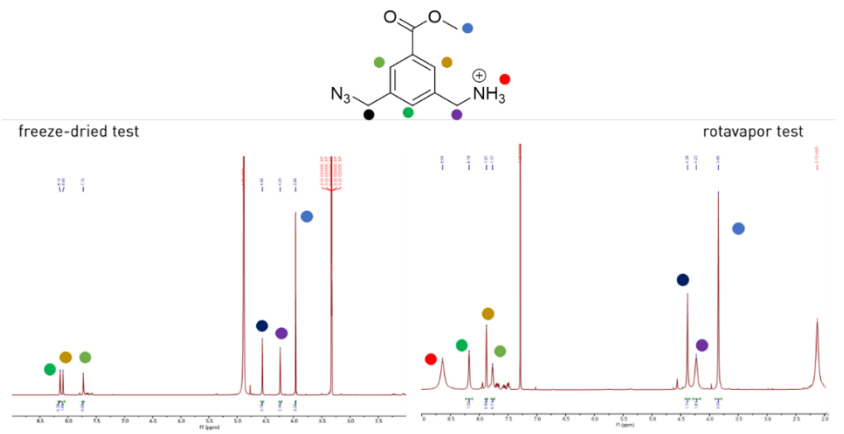



***Figure 3.6.*** *¹H-NMR spectra regarding two drying methods carried out during the synthesis of Compound* ***3****.*

However, on a bigger scale (~400 mg), using the rotavapor for evaporating the water phase led to partial ester hydrolysis, caused by the longer time required to dry the water phase (*Figure 3.7*). This side reaction, however, did not compromise the overall synthesis because the subsequent step is the ester hydrolysis (*Figure 3.8*).

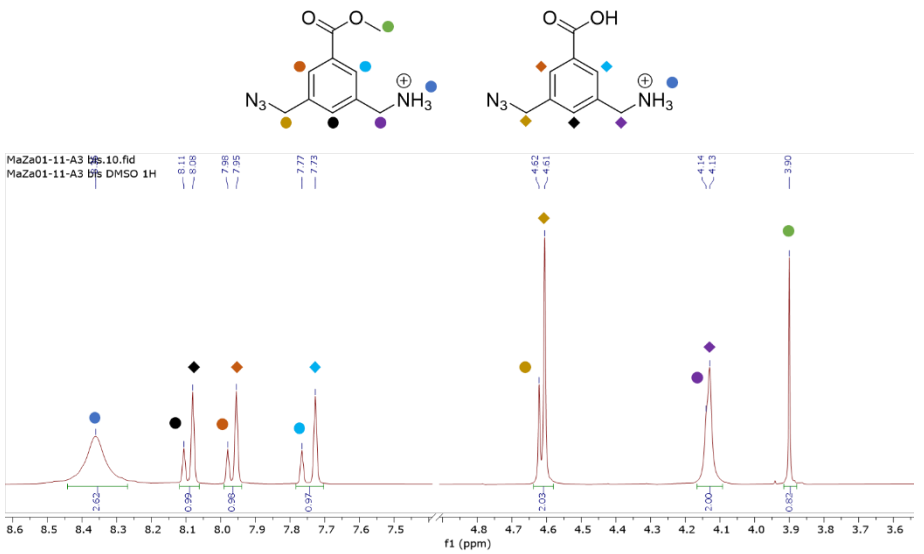
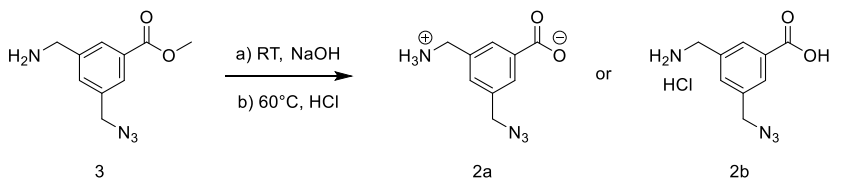



***Figure 3.7.*** *$^1$H-NMR (400 MHz, DMSO-d$^6$) spectrum of the mixture obtained in the rotavapor dried large-scale synthesis, showing partial ester hydrolysis of compound **3**.*




3 2a 2b

***Figure 3.8.*** *Synthesis of 3-(aminomethyl)-5-(azidomethyl)benzoic acid by ester hydrolysis of **3**.*

The synthesis of compound **2** was carried out by completing the hydrolysis of ester present in **3**. This reaction was performed both under basic and acidic conditions. When the reaction was performed with a NaOH 1 M solution in water/methanol at room temperature, the precipitation step of the zwitterion form (isoelectric point at 6.7, calculated by $pK_a$ of benzoic acid and benzylamine) was the crucial step. In addition, it is necessary not to exceed with the volume of the water to obtain the precipitation of the molecule in the zwitterion form.

The acid hydrolysis was studied as a one-pot reaction directly from compound **4** to **2**. A screening of concentrations of HCl solution and temperature was performed. The reaction time was kept overnight for all the experiments. The results were reported in *Table 3.1*.

***Table 3.1.*** *Results regarding the screening of concentrations of HCl solution and temperature performed for the acid hydrolysis of the ester **3**.*

| Test | HCl concentration | Temperature | Results | % |
|---|---|---|---|---|
| 1 | 1M | 50°C | Partial hydrolysis | 86 |
| 2 | 1M | 60°C | Partial hydrolysis | 98 |
| 3 | 2M | 60°C | Partial hydrolysis | 99 |
| 4 | 5M | 60°C | Fully hydrolysis | / |
| 5 | 12M | 60°C | Degradation | / |
| 6 | 12M | 100°C | Degradation | / |

It is possible to notice that the ester hydrolysis reaction improves, increasing the HCl concentration and the temperature up to 5 M, whereas too harsh conditions (those employing HCl 12 M) led to product degradation. In fact, by comparing the data, the tests with 1 M or 2 M HCl either at 50°C or 60°C has brought to a partial hydrolysis of compound **3**, while, in presence of 5 M HCl at 60°C, the compound **3** was fully hydrolyzed, as showed in $^{1}$H-NMR spectra (*Figure 3.9*). Recovery of the product led to a reaction yield of 85%. In the last two experiments of *Table 3.1* the reaction conditions were too harsh and leaded to degradation of compound **3**.

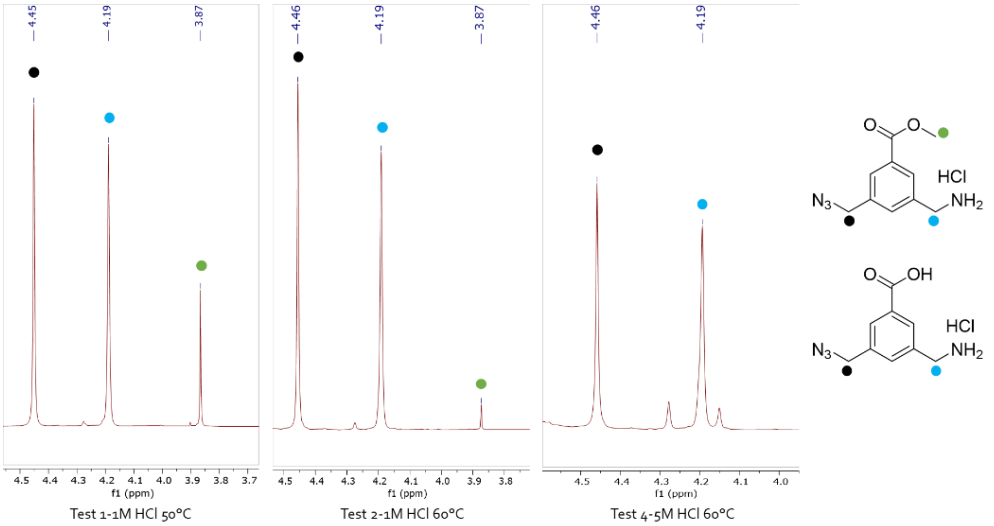



***Figure 3.9.*** *$^1$H-NMR (400MHz, $D_2O$ ) spectra of HCl concentration and temperature screening. The test numbers refer to those reported in Table 3.1*

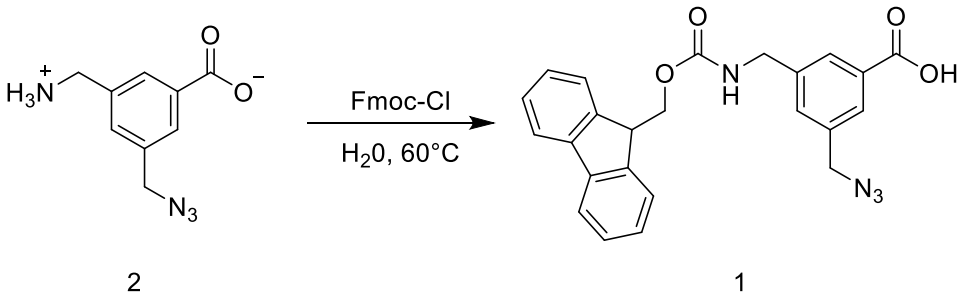



***Figure 3.10.*** *Synthesis of the target 3-((9-fluorenylmethoxycarbonyl)amino methyl)-5-(azidomethyl)benzoic acid (**1**).*

To obtain the desired tri-functionalized molecule, the protection of the amino moiety with Fmoc was the last synthetic step (*Figure 3.10*). In specific, the reaction was carried out in presence of Fmoc chloride as reagent.

Respect the typical Fmoc protective reaction, where an organic solvent is used, a different procedure adopted from the literature was employed.[16] It distinguishes itself from typical protective reactions by utilizing water as the solvent instead of an organic one.

Preliminary experiments to evaluate and optimize this procedure were carried out employing phenylalanine (Phe) to avoid the use of the costly compound **2** and given

that its structural resemblance to the target compound, sharing common features (benzene ring, a carboxylic acid moiety, and an amino functional group).

After these tests, the Fmoc protection of compound **2** was performed with 1.2 eq of Fmoc-Cl. However, the characterization shown the contamination of the desired product with Fmoc-Cl, but for limited time, it was not possible to optimize the purification of the compound **1**.

### 3.3.2 Study of PNA:PNA duplex stability

A library of PNA sequences of different lengths (11mer, 9mer, 7mer, and 5mer) and orientations (parallel and antiparallel) were designed and synthetized.

Using these series, the melting temperature of specific pairs of PNAs were analyzed with the aim of establishing the minimal number of base pairs for obtaining complete assembly of PNA:PNA duplex at room temperature.

The design (*Figure 3.11*) of the sequences was done starting from the three 11mer PNA sequences reported in *Chapter 2*, the corresponding complementary sequences in parallel orientation were individuated and designed to be used for the construction of templated assembly of peptides. In the final assembly, they should be installed in the trimeric small molecule to obtain the desired branched scaffold. Given that parallel PNA:PNA duplexes are less studied, to better understand the data, the synthesis of the respective PNA sequences in antiparallel orientation was also carried out.

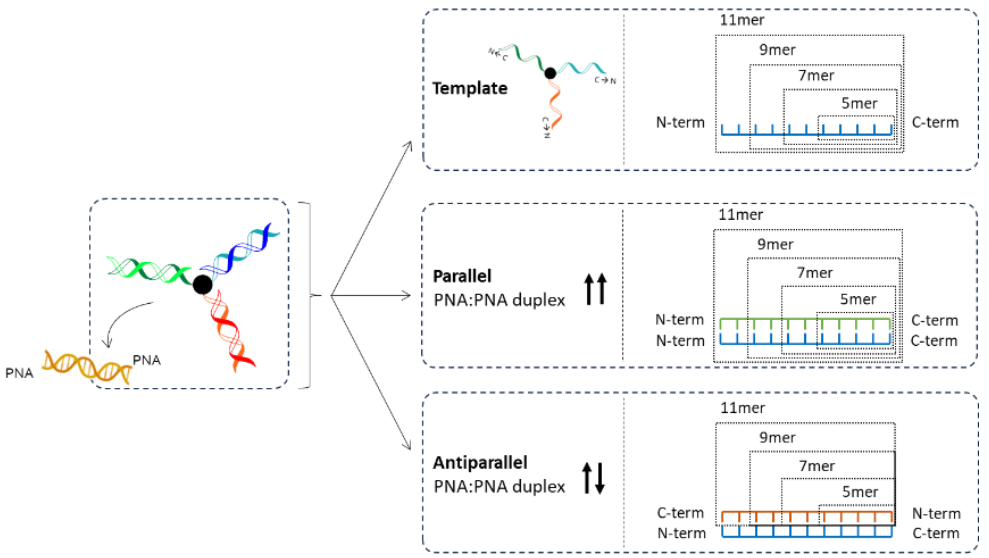



***Figure 3.11.*** *Design of PNA sequences for the construction of templated assembly of peptides.*

A summary of all the PNA sequences synthetized regarding the template, the related full-matched sequences in parallel and antiparallel orientation are listed in *Table 3.2.*

Nevertheless, all the melting temperature (*Tm*) reported subsequently are single experiments and, for limited time, it was not possible to perform replicates.

***Table 3.2.*** *PNA sequences of template (from Entry 1 to 12), PNA sequences complementary to PNAs of template in antiparallel orientation (from Entry 13 to 24), and PNA sequences complementary to PNAs of template in parallel orientation (from Entry 25 to 36).*

| Entry | PNA | Sequence | Yield (%) | | Length |
|---|---|---|---|---|---|
| 1 | 11mer A | Ac-CCCGTACTAGA-g-$NH_2$ | 12 | Template A | 11mer |
| 2 | 11mer B | Ac-ATGGCAGCTCA-g-$NH_2$ | 13 | Template B | 11mer |
| 3 | 11mer C | Ac-CACTGAGTCTA-g-$NH_2$ | 15 | Template C | 11mer |
| 4 | 9mer A | Ac-CGTACTAGA-g-$NH_2$ | 12 | Template A | 9mer |
| 5 | 9mer B | Ac-GGCAGCTCA-g-$NH_2$ | 5 | Template B | 9mer |
| 6 | 9mer C | Ac-CTGAGTCTA-g-$NH_2$ | 7 | Template C | 9mer |
| 7 | 7mer A | Ac-TACTAGA-g-$NH_2$ | 17 | Template A | 7mer |
| 8 | 7mer B | Ac-CAGCTCA-g-$NH_2$ | 8 | Template B | 7mer |
| 9 | 7mer C | Ac-GAGTCTA-g-$NH_2$ | 10 | Template C | 7mer |
| 10 | 5mer A | Ac-CTAGA-g-$NH_2$ | 22 | Template A | 5mer |
| 11 | 5mer B | Ac-GCTCA-g-$NH_2$ | 7 | Template B | 5mer |
| 12 | 5mer C | Ac-GTCTA-g-$NH_2$ | 12 | Template C | 5mer |
| 13 | 11mer $A_a$ | Ac-TCTAGTACGGG-g-$NH_2$ | 12 | Antiparallel A | 11mer |
| 14 | 11mer $B_a$ | Ac-TGAGCTGCCAT-g-$NH_2$ | 12 | Antiparallel B | 11mer |
| 15 | 11mer $C_a$ | Ac-TAGACTCAGTG-g-$NH_2$ | 20 | Antiparallel C | 11mer |
| 16 | 9mer $A_a$ | Ac-TCTAGTACG-g-$NH_2$ | 14 | Antiparallel A | 9mer |
| 17 | 9mer $B_a$ | Ac-TGAGCTGCC-g-$NH_2$ | 22 | Antiparallel B | 9mer |
| 18 | 9mer $C_a$ | Ac-TAGACTCAG-g-$NH_2$ | 28 | Antiparallel C | 9mer |
| 19 | 7mer $A_a$ | Ac-TCTAGTA-g-$NH_2$ | 10 | Antiparallel A | 7mer |
| 20 | 7mer $B_a$ | Ac-TGAGCTG-g-$NH_2$ | 25 | Antiparallel B | 7mer |
| 21 | 7mer $C_a$ | Ac-TAGACTC-g-$NH_2$ | 26 | Antiparallel C | 7mer |
| 22 | 5mer $A_a$ | Ac-TCTAG-g-$NH_2$ | 16 | Antiparallel A | 5mer |
| 23 | 5mer $B_a$ | Ac-TGAGC-g-$NH_2$ | 13 | Antiparallel B | 5mer |

| | | | | | |
|---|---|---|---|---|---|
| 24 | 5mer $C_a$ | Ac-TAGAC-g-$NH_2$ | 30 | Antiparallel C | 5mer |
| 25 | 11mer $A_p$ | Ac-GGGCATGATCT-g-$NH_2$ | 8 | Parallel A | 11mer |
| 26 | 11mer $B_p$ | Ac-TACCGTCGAGT-g-$NH_2$ | 8 | Parallel B | 11mer |
| 27 | 11mer $C_p$ | Ac-GTGACTCAGAT-g-$NH_2$ | 5 | Parallel C | 11mer |
| 28 | 9mer $A_p$ | Ac-GCATGATCT-g-$NH_2$ | 7 | Parallel A | 9mer |
| 29 | 9mer $B_p$ | Ac-CCGTCGAGT-g-$NH_2$ | 6 | Parallel B | 9mer |
| 30 | 9mer $C_p$ | Ac-GACTCAGAT-g-$NH_2$ | 9 | Parallel C | 9mer |
| 31 | 7mer $A_p$ | Ac-ATGATCT-g-$NH_2$ | 9 | Parallel A | 7mer |
| 32 | 7mer $B_p$ | Ac-GTCGAGT-g-$NH_2$ | 6 | Parallel B | 7mer |
| 33 | 7mer $C_p$ | Ac-CTCAGAT-g-$NH_2$ | 32 | Parallel C | 7mer |
| 34 | 5mer $A_p$ | Ac-GATCT-g-$NH_2$ | 15 | Parallel A | 5mer |
| 35 | 5mer $B_p$ | Ac-CGAGT-g-$NH_2$ | 9 | Parallel B | 5mer |
| 36 | 5mer $C_p$ | Ac-CAGAT-g-$NH_2$ | 40 | Parallel C | 5mer |
| Capital letters: PNA monomers, Lower case: amino acid | | | | | |

Given that the UV-Visible instrument was not available, the melting and annealing measurements were carried out at 5 μM concentration for each PNA in phosphate buffer saline (PBS, 10 mM phosphate and 100 mM NaCl, pH 7.4), using the HT channel (proportional to absorbance) of a CD spectropolarimeter, because these systems do not have a signal in the CD channel. Before the measurements, each sample was submitted to thermal annealing with a temperature ramp from 95 °C to room temperature.

Melting and annealing temperatures (*Tm* and *Ta*) of full-match and full-length antiparallel and parallel PNA:PNA duplexes were measured. To obtain a more comprehensive overview of PNA:PNA duplex stability, besides the two full-length orientations above mentioned, Tm hybrid duplexes of the 11mer with other three different lengths (9mer, 7mer and 5mer) were also measured.

The results are reported in *Table 3.3* for the antiparallel orientation and in *Table 3.4* for the parallel orientation. In both tables, the PNA sequences relative to the template are kept constant, varying thus the orientation of the corresponding PNA sequences. The melting and annealing temperature were calculated from the first

order derivative of a 10$^{th}$ order polynomial fitting function, using Matlab as software. An example is reported in *Figure 3.12*.

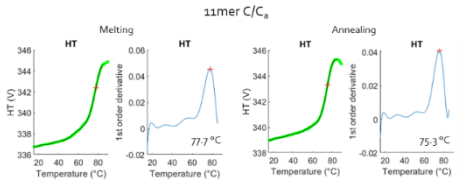



***Figure 3.12.*** *HT profile of melting and annealing analysis and values calculated from the first order derivative of a 10$^{th}$ order polynomial fitting function of 11mer C/Ca duplex.*

***Table 3.3.*** *Melting and annealing temperatures of PNA:PNA duplexes in antiparallel orientation.*

| Tm PNA:PNA duplex in antiparallel orientation | | | | |
|---|---|---|---|---|
| Serie A | | | | |
| | 5mer $A_a$ | 7mer $A_a$ | 9mer $A_a$ | 11mer $A_a$ |
| 5mer A | Tm: 24.1 °C<br>Ta: 20.1 °c | | | |
| 7mer A | | Tm: 36.4 °C<br>Ta: 36.4 °C | | |
| 9mer A | | | Tm: 62.2 °C<br>Ta: 60.1 °C | |
| 11mer A | | | | Tm: 76.7 °C<br>Ta: 75.4 °C |
| Serie B | | | | |
| | 5mer $B_p$ | 7mer $B_p$ | 9mer $B_p$ | 11mer $B_p$ |
| 5mer B | Tm: 35.0 °C<br>Ta: 35.2°C | | Tm: 48.1 °C | Tm: 50.2 °C |
| 7mer B | | Tm: 58.9 °C<br>Ta: 59.7°C | Tm: 61.7 °C | Tm: 63.5 °C |
| 9mer B | | | Tm: 78.1 °C<br>Ta: 77.0 °C | Tm: 85.2 °C |
| 11mer B | | | Tm: 79.5 °C | Tm: 85.1 °C<br>Ta: 83.5 °C |
| Serie C | | | | |
| | 5mer $C_a$ | 7mer $C_a$ | 9mer $C_a$ | 11mer $C_a$ |
| 5mer C | Tm: 21.9 °C<br>Ta: 22.7 °C | | | Tm: 59.9 °C |
| 7mer C | | Tm: 54.0 °C | | Tm: 62.8 °C |

| | | | | |
|---|---|---|---|---|
| | | Ta: 49.4 °C | | |
| 9mer C | | | Tm: 64.9 °C<br>Ta: 63.7 °C | Tm: 69.6 °C |
| 11mer C | | | | Tm: 77.7 °C<br>Ta: 75.3 °C |

***Table 3.4.*** *Melting and annealing temperatures of PNA:PNA duplexes in parallel orientation.*

| Tm PNA:PNA duplex in parallel orientation | | | | |
|---|---|---|---|---|
| Serie A′ | | | | |
| | 5mer $A_p$ | 7mer $A_p$ | 9mer $A_p$ | 11mer $A_p$ |
| 5mer A | Tm: 25.1 °C<br>Ta: 25.1 °c | | | |
| 7mer A | | Tm: 29.5 °C<br>Ta: 27.0 °C | | |
| 9mer A | | | Tm: 43.4 °C<br>Ta: 42.1 °C | |
| 11mer A | | | | Tm: 58.4 °C<br>Ta: 56.6 °C |
| Serie B′ | | | | |
| | 5mer $B_a$ | 7mer $B_a$ | 9mer $B_a$ | 11mer $B_a$ |
| 5mer B | Tm: 18.3 °C<br>Ta: 18.0 °C | Tm: 31.7 °C | Tm: 38.2 °C | Tm: 37.7 °C |
| 7mer B | | Tm: 37.4 °C<br>Ta: 35.5 °C | Tm: 43.9 °C | Tm: 40.6 °C |
| 9mer B | | Tm: 45.8 °C | Tm: 57.5 °C<br>Ta: 54.2 °C | Tm: 59.5 °C |
| 11mer B | | Tm: 45.3 °C | Tm: 57.3 °C | Tm: 62.7 °C<br>Ta: 59.7 °C |
| Serie C′ | | | | |
| | 5mer $C_p$ | 7mer $C_p$ | 9mer $C_p$ | 11mer $C_p$ |
| 5mer C | Tm: 20.1 °C<br>Ta: 20.0 °C | | | |
| 7mer C | | Tm: 43.0 °C<br>Ta: 35.7 °C | | |
| 9mer C | | | Tm: 68.8 °C<br>Ta: 67.2 °C | |
| 11mer C | | | | Tm: 70.8 °C<br>Ta: 65.8 °C |

From both tables, it is possible to observe, as expected, an increase of melting and annealing temperature of the PNA:PNA duplex going from 5mer to 11mer, in both the orientations. However, the melting and annealing temperatures are sequence dependent. Given the difference in number of hydrogen bonds in the G-C pairing respect the A-T, the first couple contributes to stabilize more the duplex.

In addition, the annealing measurements were carried out to evaluate if the PNA:PNA formation could be subject to hysteresis, which in turn could be a sign of kinetically slow annealing processes (such as those observed for triplexes). Comparing the melting and annealing data, no relevant hysteresis was noted. In fact, the values of melting and annealing temperature are similar for all the PNAs examined.

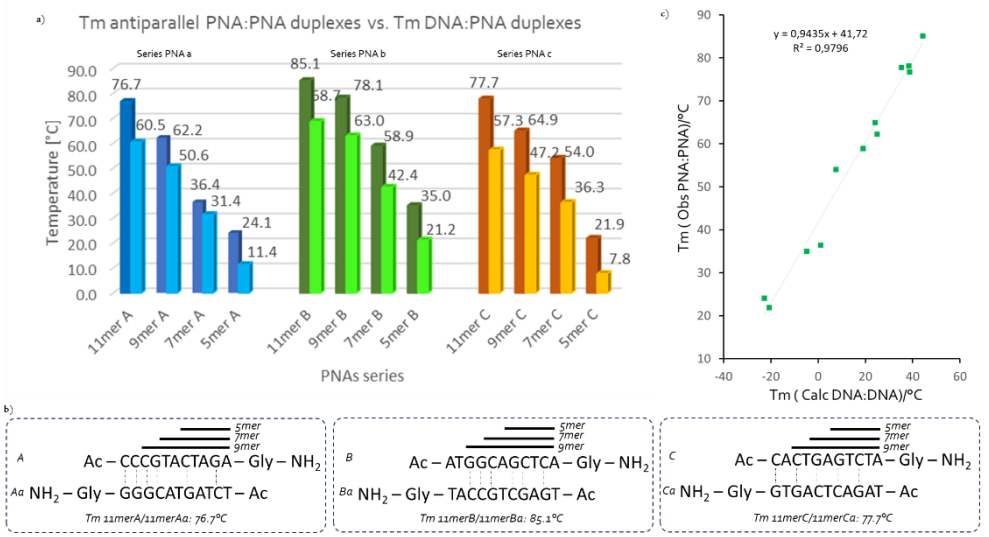



***Figure 3.13.*** *a) Schematic representation of Tm of PNA:PNA duplex in antiparallel orientation and comparison of these Tm with theoretical Tm values of DNA:PNA duplex. b) Graphic representation of G-C pairs in the 11mer PNAs. c) Correlation between the observed melting temperatures for full-match, full-length, antiparallel PNA:PNA duplexes and those calculated as average value for the DNA:DNA duplex (using the Santa Lucia model at 5 µM strand concentration, 50mM buffer and 0M Mg2+ , https://altogenlabs.com/resources/dna-tm-calculator/).*

Considering the antiparallel orientation and analyzing the data showed in *Figure 3.13a*, one can be seen a higher value of melting temperature in present of six G-C pairing distributed inside the 11mer sequence respect the same number of G-C at

the end of the sequence (11mer*A/Aa* respect *11merB/Ba*). In fact, 11mer*B/Ba* duplex present a Tm of 85.1 °C, while 11mer*A/Aa* of 76.7 °C. In addition, a homogeneous distribution of G-C pairing along the PNA sequences, as in *11merC/Ca*, also with a less G-C base pair, has given a denaturing temperature like *11merA/Aa*. A graphic representation was reported in *Figure 3.13b*.

Since the 9mer, 7mer and 5mer sequences are designed by the elimination of couple of nucleobases at the N-terminus respect the PNA sequences of template, the series *PNA a* shown a drop of Tm more pronounced because it lost, both in 11mer to 9mer and in 9mer to 7mer, a couple of G-C pairing. In fact, the Tm decrease approximately of 15-25°C, while losing two A-T pairing or one per type, a minor difference in temperature (about 10 °C) is observed like *series B* and *C*, except the 5mer PNAs. These last turn out to be not very stable at room temperature because they are too short and in fact the Tm value from 7mer to 5mer is roughly half, although the 5mer of *series B* has a slightly higher Tm value compared to the other two and this could be due to the presence of addition G-C base pairing.

A more precise model for nucleic acid duplex stability is the well-established Santa Lucia model, which considers the stabilizing effect of stacking interactions and thus is not only based on base composition, but also on nearest-neighbor pairs.[17] A prediction formula of PNA:DNA melting temperatures, which allows researchers to foresee the stability of the complex between PNA probes and their targets, have been proposed based on empirical data[15] but so far, no predictive model is available for PNA:PNA duplexes.

However, these previous results were compared with the theoretical Tm of corresponding PNA:DNA duplex, given that the PNA:PNA interaction are reported to be more stable than the corresponding DNA:PNA duplex. To obtain these data, that are reported in *Figure 3.12a*, an online platform called *PNA Bio* was used, and the sequences of template was selected as PNAs.

It seemed therefore useful to infer a correlation between the predicted melting temperature of DNA:DNA duplexes and that of the PNA:PNA duplexes of the same sequence. The results are reported in *Figure 3.13c,* and a linear correlation can be

noted between the experimental data of PNA:PNA duplexes and the theoretical data of DNA:DNA duplexes calculated using the Santa Lucia model. Thus, it is possible to use this model to have a rough estimate of the PNA:PNA stability when designing nanostructure containing this motif.

A similar correlation was found for predicted DNA:PNA (*Figure 3.A13* in Appendix). However, since the PNA chosen to calculate the PNA:DNA Tm can affect the stability of the latter, due to the asymmetry in Tm of PNA:DNA duplexes as evidenced in Sen and Nielsen, P.E. Bioph. J. , 2006, 90 ,1329–1337 , we chose to use the average of the predicted PNA:DNA Tm obtained for the two PNA strands.

Dealing with the PNA:PNA duplex in parallel orientation, the Tm of the three series is reported in *Figure 3.14a*. Contrary to the opposite orientation, the melting temperatures regarding *series Cp* are higher than the other two series, although it presents one G-C pairing in less, as can been seen also from the *Figure 3.14b*. However, also in this case, a homogeneous distribution of G-C pairing along the PNA sequences seems to need more energy to separate the two strands with respect to PNAs with G-C pairing more concentrated on one side of the PNA sequences.

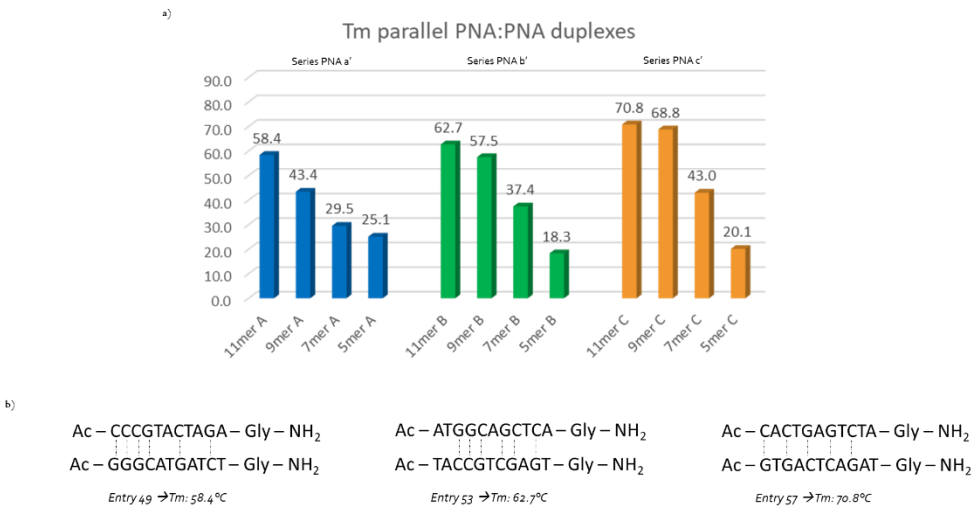



***Figure 3.14.*** *a) Schematic representation of Tm of PNA:PNA duplex in parallel orientation. B) Graphic representation of G-C pairs in the 11mer PNAs.*

A linear trend in the Tm of parallel PNA:PNA duplexes is difficult to identify, and in contrast to the antiparallel duplexes, these results do not seem to be roughly predictable based on the nearest-neighbor model of dsDNA stability. The reason for this can be attributed to the fact that the data was collected in a single measurement, the need to gather a larger dataset, and the presence of other parameters that differ from the standard ones.

In addition, several melting temperature experiments of duplex formed with the PNAs sequences at different lengths were carried out to better understand the effect of overhanging unpaired nucleobases both in parallel and antiparallel orientations. The results are reported in *Table 3.5, 3.6 and 3.7*.

***Table 3.5.*** *Melting temperature of PNA:PNA duplex in antiparallel orientation at different length correlated to PNA series B.*

| Serie B | | | | |
|---|---|---|---|---|
| | 9mer $B_a$ | ΔTm (Tm-Tm full-length)/°C | 11mer $B_a$ | ΔTm (Tm-Tm full-length)/°C |
| 5mer B | Tm: 48.1 °C | 13.1 | Tm: 50.2 °C | 15.2 |
| 7mer B | Tm: 61.7 °C | 2.8 | Tm: 63.5 °C | 4.6 |
| 9mer B | Tm: 78.1 °C | 0.0 | Tm: 85.2 °C | 7.1 |
| 11mer B | Tm: 79.5 °C | 1.4 | Tm: 85.1 °C | 0.0 |

***Table 3.6.*** *Melting temperature of PNA:PNA duplex in antiparallel orientation at different length correlated to PNA series C.*

| Serie C | | |
|---|---|---|
| | 11mer $C_a$ | ΔTm (Tm-Tm full-length)/°C |
| 5mer C | Tm: 59.9 °C | 38 |
| 7mer C | Tm: 62.8 °C | 8.8 |
| 9mer C | Tm: 69.6 °C | 5.3 |

***Table 3.7.*** *Melting temperature of PNA:PNA duplex in parallel orientation at different length correlated to PNA series b'.*

| Serie B' | | | | | | |
|---|---|---|---|---|---|---|
| | 7mer $B_p$ | ΔTm (Tm-Tm full-length)/°C | 9mer $B_p$ | ΔTm (Tm-Tm full-length)/°C | 11mer $B_p$ | ΔTm (Tm-Tm full-length)/°C |
| 5mer B | Tm: 31.7 °C | 13.4 | Tm: 38.2 °C | 19.9 | Tm: 37.7 °C | 19.4 |
| 7mer B | Tm: 37.4 °C | 0 | Tm: 43.9 °C | 6.5 | Tm: 40.6 °C | 3.2 |
| 9mer B | Tm: 45.8 °C | 8.4 | Tm: 57.5 °C | 0 | Tm: 59.5 °C | 2.0 |
| 11mer B | Tm: 45.3 °C | 7.9 | Tm: 57.3 °C | -0.2 | Tm: 62.7 °C | 0.0 |

As expected, by *Table 3.6*, it is possible to notice that generally the values of *Tm* decreasing from completely paired 11mer PNA:PNA duplex to 11mer:5mer as the number of nucleobases paired decrease is observed.

Similar observation can be drawn also for the results in the other two tables, although it is possible to notice that the flanking nucleobases from 9merB to 11merB for both antiparallel and parallel (corresponding to the same PNA sequences of template) allocate little contributions to duplex stability.

By analyzing the nucleobases type could be said that the presence of unpaired adenine and thymine does not contribute to duplex stabilization in both orientations. However, it is noteworthy from *Table 3.5-3.7* that the flanking nucleobases contribute to the stabilization of these complexes. When compared with Tm obtained with the full-length duplexes (*ΔTm*), those observed with overhanging PNA segments have always (expect one case) a higher stability.

This is due to the stacking interaction with the last bases of the flanking nucleobases, but other contributes could be present, such as hydrogen bonds or dipole-dipole

forces. . These can be accomplished by a fold-back of the unpaired segment, which is facilitated by PNA flexibility.

For the purpose of the design of an artificial assembly containing these PNA segments as templating bricks and considering the necessity to assembly PNA:PNA duplex completely at room temperature (therefore having a Tm of at least 45 °C), the above reported results seem to point to the conclusion that the minimum length of PNA:PNA duplex to be used is between 7 and 9 monomers. This would allow to avoid the loss of structural stability, while still maintaining the minimum length of the complex.

## 3.4 Conclusions and future directions

In this part of the work, the preliminary steps to obtain a PNA-based templated peptide nanostructure were performed.

The synthesis of a tri-functionalized rigid molecule was carried out by the optimization of the first three steps, starting from commercially available methyl 3,5-bis(bromomethyl) benzoate as starting material, though purification of the Fmoc-protected final product is still a challenge. This rigid molecule, containing a carboxylic acid, an azide, and a Fmoc-protected amino function can then be used to grow three different PNAs branched in direction from $C_{term}$ to $N_{term}$ thanks to the orthogonality of these three different groups.

On the other side, a library of PNA sequences at different composition, length and orientation was synthesized, allowing to perform melting and annealing measurements, and to study the stability of PNA:PNA duplex. This information was useful to establish the minimum length suitable to generate a three-fold PNA:PNA duplex as driving force for the assembly of hetero coiled-coil systems with short peptide domains. Interestingly, a general method to roughly predict the stability of antiparallel PNA:PNA duplex stability was obtained, whereas for the parallel complexes a trial-and-error approach is still necessary. From the results, it is possible to conclude that the minimum length to achieve a duplex based only on PNAs sufficiently stable at room temperature is between 7 and 9mer, with 9 being necessary for the less stable parallel duplexes. The duplex formed by 8-mer sequences can be also added to this study for fine-tuning, since this length for specific sequences could be a good balance between the structural stability and minimum sequence length to be used. In addition, the role of flanking segments on duplex stability was studied, revealing additional contributions to the PNA:PNA interactions. However, it might be interesting to systematically extend this study to other combinations of PNA:PNA duplexes with different length and composition of the flanking segments, to understand the rationale of the additional contribution to the PNA:PNA interactions.

Having chosen the correct PNA composition, using the central core synthesized, the synthesis of the templating unit for the heterotrimeric coiled-coil system can be performed in the following way (*Figure 3.15*): the solid support used to PNA synthesis could be loaded with an amino acid that present two different protecting groups, one in α-position and the other in the side chain (namely a doubly-protected lysine). Subsequently, the first PNA sequences could be grown, through cycles of deprotection/coupling reactions, starting from the amino group in α-position. After the first PNA segment synthesis, the second protecting group of the amino acid can be removed to permit the coupling of the carboxylic group of the trimeric molecule. A second PNA sequence can then be synthetized by deprotecting the Fmoc group and subsequent couplings with PNA monomers, until the designed PNA length is achieved. Lastly, the third PNA sequence can be linked off resin by exploiting the azido functionality through a click-chemistry (1,3-dipolar cycloaddition) reaction with an alkyne-modified oligomer (previously synthetized). Alternatively, the last PNA can also be built on the solid support, reducing the azido group to amine, and performing subsequent coupling reactions with PNA monomers. However, this last strategy might suffer from the high steric hindrance generated by the two PNA strands and would require a careful evaluation of "resin loading" A schematic representation of this scaffold formed by three different PNA sequences is reported in *Figure 3.15*.

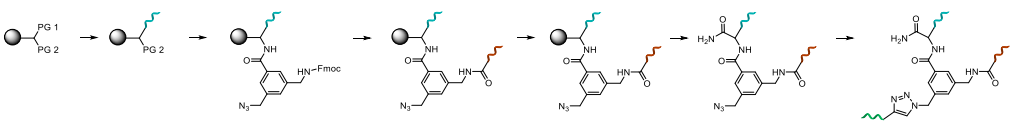



***Figure 3.15.*** *Schematic representation of trimeric scaffold based on PNAs synthesis.*

## 3.5 Experimental section

### 3.5.1 Reagent and instrumentations

The reagents were bought from BLD Pharmatech, Merck, abcr, Carlo Erba, VWR, Fluka, TCI Europe, Link Technologies, PolyOrg and used without additional purification. Only for the solid phase synthesis, the DMF was purged with nitrogen flow to prevent the formation of dimethylamine and was dried over 4Å molecular sieves.

NMR spectra were recorded, using $CDCl_3$, DMSO-d6 or $D_2O$, as solvents with a Brucker Avance 400 MHz instrument. δ values are expressed in ppm. TLC were performed on Supelco 56524-25EA silica gel on aluminum foil with indicator fluorescence at 254 nm.

All the PNA sequences are synthesized by standard solid-phase manual synthesis or by automatic synthesizer Biotage Syro I. Purification was performed by HPLC (Agilent Technologies 1260 Infinity I) using a SepaChrom Vydamas® (C18, 5 µm, 300 Å, 10 x 250 mm) column. Gradient used were as follows: a) Gradient 1 (for 5,7, and 9mer PNAs): 100% A for 1 min, then from 0% to 25% B for 16 min; b) Gradient 2 (for 11mer PNAs): 100% A for 1 min, then from 0% to 30% B for 19 min; for both gradients: flow: 4 mL/min and eluents: A: water + 0.1% trifluoroacetic acid; B: acetonitrile + 0.1% trifluoroacetic acid. Detector: UV set with wavelength: 260 nm.

After purification, the peptide nucleic acid sequences were characterized by UPLC-MS (Waters Acquity Ultra Performance LC) using the following instrumental set-up: Waters Acquity ultra-performance LC E07SQD086W, with Waters SQ detector and ESI-interface equipped with Acquity UPLC BEH 300 (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic condition: eluent A: water + 0.2% formic acid; eluent B: acetonitrile + 0.2% formic acid. Column temperature: 35 °C. Program: initial isocratic at 100% A (0.9 min), then linear gradient to 50% B (in 5.7 min). Final wash with 100% B for 1.2 min. Flow rate: 0.25 mL/min. Some PNAs, however, were characterized with Thermo LTQ Orbitrap XL detector and ESI-interface equipped with a Phenomenex Kinetex EVO (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic conditions: eluent

A: water + 0.1% formic acid; eluent B: acetonitrile + 0.1% formic acid. Column temperature: 35 °C. Program: initial isocratic at 98% A (3 min), then linear gradient to 50% B (in 20 min) and then to 95% (in 1 min). Final wash with 98% B for 8 min. Flow rate: 0.20 mL/min. UV wavelength: 260 nm.

The concentration of the PNAs was calculated using Evolution 260 Bio UV-Visible spectrophotometer with Peltier thermos tatting accessories (Thermo Fisher Scientific SPE8W), following the UV-absorbance at 260 nm and assuming an additive contribution of all bases.

The CD measurements were performed with J-1500 (Serial No. D067261638) with PM-539 as detector and PTC-510 as accessory for temperature monitoring. Experimental conditions to melting temperature: Tm/Ta setup; temperature interval: 15-90°C; data interval: 0.1°C; temperature was kept within +/- 0.1°C of the target temperature for 5 seconds; control sensor and monitor sensor: holder. Tm/Ta parameters: CD and HT channels; CD scale: 20 mdeg/0.05 dOD; FL scale: 200 mdeg/0.1 dOD; D.I.T.: 8 sec; bandwidth: 1.00 nm; monitor wavelength: 260 nm; start and end CD parameters: 260-200 nm; concentration: 5 umol/L to each probe; solvent: 10mM PBS solution; data pitch: 0.1 nm; scanning speed: 0.5°C/min. Melting and annealing measurements were performed to duplex PNA:PNA at the same lengths and only melting analysis were carried out to duplex PNA:PNA at different number of nucleobases.

### 3.5.2 Synthetic protocols of synthesis of a tri-functionalized rigid core molecule

Methyl 3,5-bis(azidomethyl)benzoate (Compound **4**)

In a round-bottom flask, **5** (519 mg, 1 eq) and sodium azide (224 mg, 2.1 eq) were mixed in DMF dry (1 mL) and left to react overnight at room temperature. After the complete conversion of the starting material (checked by TLC), the reaction mixture was diluted in ethyl acetate and extracted with a solution of $H_2O$/brine (1:1, 4x100 mL). The organic solution was dried over $Na_2SO_4$, and the solvent removed under reduced pressure. The product was a yellow oil (312 mg, 78%).

$^{1}$H-NMR (400 MHz, 25 °C, $CDCl_3$) δ(ppm): 7.97 (m, 2H, ArH), 7.48 (m, 1H, ArH), 4.44 (s, 4H, $CH_2$), 3.94 (s, 3H, $OCH_3$). $^{13}$C-NMR (100 MHz, 25 °C, $CDCl_3$) δ(ppm): 166.3 ($C_{quat}$), 136.9 ($C_{quat}$), 131.9 (CH), 131.5 ($C_{quat}$), 129.1 (CH), 54.2 ($C_2$), 52.5 ($CH_3$). UPLC-MS (ESI+): 5.41 min, m/z found 247.3 $[M+H]^+$ (m/z calculated for $[C_{10}H_{10}N_6O_2]$: 246.23). TLC: hexane/ethyl acetate (5:1), Rf: 0.32.

Spectroscopic data are consistent with those reported in the literature.[18]

Methyl 3-(aminomethyl)-5-(azidomethyl)benzoate hydrochloride (Compound **3**)

In a two-neck round-bottom flask, **4** (312 mg, 1 eq) was dissolved in a solution of hexane:diethyl ether (1:1, 30 mL) and subsequently 50ml of HCl 1M solution were added. The triphenylphosphine (324 mg, 0.98 eq) was dissolved in 35 mL of diethyl ether and after this solution was added dropwise using a dropping funnel, in 1 hour, followed by vigorous stirring for 3 hours. After that, the two phases of this biphasic solution were separated, and the aqueous phase was extracted with hexane (30 mL). The aqueous phase then was left to react overnight. After other three extractions with hexane, the aqueous phase was evaporated under reduced pressure or lyophilized by freeze-dried. The resulting product was isolated as sticky oil (168 mg, 59%).

$^{1}$H NMR (400 MHz, $CDCl_3$) δ(ppm): 8.64 (br, s, 3H, $NH_3^+$), 8.18 (s, 1H, ArH), 7.87 (s, 1H, ArH), 7.77 (s, 1H, ArH), 4.38 (s, 2H), 4.23 (s, 2H, $CH_2$), 3.85 (s, 3H, $OCH_3$). UPLC-MS (ESI+): 3.10 min m/z found 221.2 $[M+H]^+$ (m/z calculated for [C10H12N4O2]: 220.23). TLC: hexane/ethyl acetate (4:2), Rf: 0.08.

3-(Aminomethyl)-5-(azidomethyl)benzoic acid (Compound **2a**)

-Basic Hydrolysis

In a round-bottom flask, the compound *3* was dissolved in a $H_2O$:MeOH solution (1:1, 3ml) and then mixed with NaOH (54.4 mg, 3 eq). After stirring at room temperature overnight, the pH of the solution was lowered to 6.7 and the product was left to precipitate at 4°C overnight. Subsequently, it was centrifuged, and the white solid was treated with ethyl acetate (8 mL). The organic solution was evaporated under reduced pressure, and the product was thus obtained as a white solid (36.8 mg, 37%).

$^{1}$H NMR (400 MHz, DMSO-d$^{6}$ ) δ(ppm): 7.75 (s, 1H), 7.66 (s, 1H), 7.21 (s, 1H), 4.39 (s, 2H), 3.71 (s, 2H) $^{13}$C NMR (400 MHz, DMSO-d$^{6}$ ) δ(ppm): 174.6, 170.0, 142.8, 140.8, 134.6, 128.4, 128.9, 54.4, 45.6.

3-(Aminomethyl)-5-(azidomethyl)benzoic acid hydrochloride (Compound **2b**)

-Acidic Hydrolysis

In a round bottom-flask, ***3*** was mixed with HCl 1M solution and was left to stir overnight at 60°C. After that, the solvent was removed under reduced pressure and the product was obtained as a yellow oil (121 mg, 85%).

$^{1}$H NMR (400 MHz, $D_2O$) δ(ppm): 7.96 (s, 1H), 7.93 (s, 1H), 7.60 (s, 1H), 4.46 (s, 2H), 4.19 (s, 2H). UPLC-MS (ESI+): 2.98 min, m/z 207.2$[M+H]^+$ (m/z calculated for $[C_9H_{10}N_4O_2]$: 206.21).

3-((9-Fluorenylmethoxycarbonyl)amino methyl)-5-(azidomethyl)benzoic acid (Compound **1**)

In a vial, the compound **4** (16.7 mg, 1 eq) and Fmoc chloride (20 mg, 1.2 eq) was dissolved in a $H_2O$:EtOH solution (3:1, 1.5 mL) and the reaction was stirred at 60°C overnight. After, the pH of the solution was around 4-4.5 and the product was extracted with ethyl acetate (3x10 mL). After making anhydrous, the solvent was removed under reduced pressure. The product was a white solid.

## 3.5.3 Synthesis of peptide nucleic acids

All PNA sequences of template and those complementary in parallel orientation were synthetized with standard solid-phase manual synthesis, using 9-fluorenylmethoxycarbonyl (Fmoc) strategy while the PNA oligomers complementary to PNAs of template in antiparallel orientation were obtained, using the same strategy, but with automatic synthesis through automatic synthesizer Syro I. The PNA monomers employed to Fmoc strategy are Fmoc/Bhoc protected.

The H-Rink Amide Chemmatrix resin was loaded with Fmoc-Gly-OH to obtain 0.2 mmol/gr as theoretical loading. The principal steps to load the resin were: a) swelling

in DCM (1x1h), b) DMF dry wash, c) coupling (1x5h, activation for 10', using 1 eq of Fmoc-Gly-OH, 10 eq of DIC and 10 eq of DhBtOH in DMF dry as activation solution), d) DMF wash, e) capping (2x15', DMF:$Ac_2O$, 1:1), e) DMF wash and f) DCM wash. To measure the loading of the resin a UV/Visible technique was used, following the absorbance of divenzofulvene (DBF) formation after deprotection treatment (30', Piperidine:DMF, 1:4) at 290 nm. The loading of the resin was calculated from the difference between the absorbances at 290 nm and a 400 nm (baseline) which allowed to establish the amount of loaded Fmoc-Glycine, corrected with a value obtained by multiplication of a conversion factor with the weighed resin mass.

The synthesis of PNA synthesis of template and those complementary in parallel orientation was performed in a 10 µmol as a scale, after 5, 7 and 9 nucleobases an amount of resin corresponding to 2 µmol was set apart and the synthesis was continued using the remaining resin. PNA sequences complementary to PNAs of the template in antiparallel orientation were synthetized separately in a 5 µmol as a scale.

The principal steps and solutions for the standard solid-phase manual synthesis were: a) swelling in DCM (1x30'), b) deprotection (2x8', Piperidine:DMF, 1:4), c) DCM wash, d) DMF dry wash, e) Kaiser test (1', positive), f) coupling (1x30', activation for 2', using 5 eq of PNA monomer, 4.9 eq of HBTU as coupling reagent and 10 eq of DIPEA in DMF dry as activation solution), g) DMF wash, h) Kaiser test (1', negative), i) capping (2x1', $Ac_2O$:DIPEA:DMF dry, 5:6:89), l) DMF wash, m) 5% DIPEA wash (2x2', DIPEA:DMF, 5:95), n) DMF wash and o) DCM wash.

The solutions used for the Kaiser test were the following: K1) 1 g of ninhydrin in 10 mL absolute ethanol, K2) 8 g phenol in 2 mL of absolute ethanol, and K3) 0.2 mL aqueous solution 0.001 M of KCN diluited at 10 mL with pyridine. To perform this test 2-3 drops of each solution were added in a tube where few grains of the resin were previously included with a capillary. After that, the tube was put in an oil bath at 100°C for 1 minute. The test is positive if the resin becomes blue.

For the solid-phase automatic synthesis with Syro I the protocol does not allow the activation step, since the PNA monomer, HBTU and DIPEA are added in this order

to the reactor with the resin and mixed. The coupling step was performed two times for 40', using 5 eq of monomer, coupling reagent and 10 eq of DIPEA). In addition, the monomers Fmoc-PNA-A(Bhoc)-OH, Fmoc-PNA-B(Bhoc)-OH, and Fmoc-PNA-T-OH were dissolved in DMF dry to obtain a concentration of 0.1 M. The Fmoc-PNA-C(Bhoc)-OH were dissolved in NMP dry to avoid solubility problems. HBTU was dissolved in DMF dry at a concentration of 0.47 M; concentration of DIPEA in DMF dry was 0.40 M. Unlike the manual protocol, the 5% DIPEA washing was not performed.

After the synthesis 2 μmol of each PNA sequence were deprotected and acetylated, using the following protocol: a) deprotection (2x8', Piperidine:DMF, 1:4), b) DMF wash, c) capping (2x1', $Ac_2O$:DIPEA:DMF dry, 5:6:89), d) DMF wash, e) 5% DIPEA wash (2x2', DIPEA:DMF, 5:95), f) DMF wash, and g) DCM wash.

For the cleavage step a cocktail solution composed of TFA:m-cresol (9:1) was used and two cycles of 1h were performed to make sure a complete cleavage. After the first cycle the solution was filtered and collected in a falcon tube, and the resin was washed with only TFA to ensure a complete collection of the product, and the same thing was repeated at the end of the second cycle. The PNA oligomers were precipitated in diethyl ether in freezer for at least 2h. After removing diethyl ether and drying the product, the latter was dissolved in double distilled water.

All the PNAs were purified by HPLC and characterized by UPLC-MS and UV-Visible techniques. UV-Vis was used to calculate the concentration, using absorbance at 260 nm and 400 nm (as baseline). The molar absorptivity (ε) specific for each PNA was calculated considering 13700 $M^{-1}cm^{-1}$ for adenine monomer, 6600 $M^{-1}cm^{-1}$ for cytosine, 11700 $M^{-1}cm^{-1}$ for guanine and 8600 $M^{-1}cm^{-1}$ for thymine as molar absorptivity values of a single nucleobase. The results obtained in this way were used also to calculate the overall yield of the synthesis.

### 3.5.4 PNAs characterizations

Entry 1 (Figure 3.A14 in Appendix)
MaCa01-12 A4: Ac-CCCGTACTAGA-Gly-$NH_2$
yield: 12%; $t_r$: 2.61 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 3061.99 [M], observed (from ESI-MS deconvolution): 3061.2; ESI-MS: m/z observed (calcd): 1021.5 (1021.7) $[M+3H]^{3+}$, 766.3 (766.5) $[M+4H]^{4+}$, 613.3 (613.4) $[M+5H]^{5+}$, 511.3 (511.3) $[M+6H]^{6+}$, 438.6 (438.4) $[M+7H]^{7+}$.

Entry 2 (Figure 3.A15 in Appendix)
MaCa01-11 A4: Ac-ATGGCAGCTCA-Gly-$NH_2$
yield: 13%; $t_r$: 2.81 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 3102.02 [M], observed (from ESI-MS deconvolution): 3101.8 ESI-MS: m/z observed (calcd) [M]: 1035.2 (1035.0)$[M+3H]^{3+}$, 776.6 (776.5) $[M+4H]^{4+}$, 621.5 (621.4)$[M+5H]^{5+}$, 518.0 (518.0) $[M+6H]^{6+}$, 444.2 (444.1) $[M+6H]^{6+}$.

Entry 3 (Figure 3.A16 in Appendix)
MaCa01-10 A4: Ac-CACTGAGTCTA-Gly-$NH_2$
yield: 15%; $t_r$: 2.95 min; ε (260 nm): 110100 $M^{-1}cm^{-1}$; MW calculated: 3077.00 [M], observed (from ESI-MS deconvolution):3076.8 ESI-MS: m/z observed (calcd): 1026.9 (1026.7) $[M+3H]^{3+}$, 770.4 (770.3) $[M+4H]^{4+}$, 616.6 (616.4) $[M+5H]^{5+}$, 514.0 (513.8) $[M+6H]^{6+}$.

Entry 4 (Figure 3.A17 in Appendix)
MaCa01-12 A3: Ac-CGTACTAGA-Gly-$NH_2$
yield: 12%; $t_r$: 2.64 min; ε (260 nm): 94900 $M^{-1}cm^{-1}$; MW calculated: 2559.49 [M], observed (from ESI-MS deconvolution): 2559.0; ESI-MS: m/z observed (calcd): 854.1 (854.2) $[M+3H]^{3+}$, 640.8 (640.9) $[M+4H]^{4+}$, 512.8 (512.9) $[M+5H]^{5+}$, 427.8 (427.6) $[M+6H]^{6+}$.

Entry 5 (Figure 3.A18 in Appendix)
MaCa01-11 A3: Ac-GGCAGCTCA-Gly-$NH_2$
yield: 5%; $t_r$: 2.81 min; ε (260 nm): 90900 $M^{-1}cm^{-1}$; MW calculated: 2560.48 [M], observed (from ESI-MS deconvolution): 2560.4; ESI-MS: m/z observed (calcd): 854.5 (854.5) $[M+3H]^{3+}$, 641.2 (641.1) $[M+4H]^{4+}$, 513.2 (513.1) $[M+5H]^{5+}$, 427.8 (427.7)$[M+6H]^{6+}$.

Entry 6 (Figure 3.A19 in Appendix)
MaCa01-10 A3: Ac-CTGAGTCTA-Gly-$NH_2$
yield: 7%; $t_r$: 2.86 min; ε (260 nm): 89800 $M^{-1}cm^{-1}$; MW calculated: 2550.48 [M], observed (from ESI-MS deconvolution): 2550.8; ESI-MS: m/z observed (calcd): 851.3 (851.2) $[M+3H]^{3+}$, 638.7 (638.6) $[M+4H]^{4+}$, 511.1 (511.1)$[M+5H]^{5+}$.

Entry 7 (Figure 3.A20 in Appendix)
MaCa01-12 A2: Ac-TACTAGA-Gly-$NH_2$
yield: 17%; $t_r$: 2.90 min; ε (260 nm): 76600 $M^{-1}cm^{-1}$; MW calculated: 2016.97 [M], observed (from ESI-MS deconvolution): 2017.6; ESI-MS: m/z observed (calcd): 673.4 (673.3)$[M+3H]^{3+}$, 505.4 (505. 2) $[M+4H]^{4+}$, 404.5 (405.4) $[M+5H]^{5+}$.

Entry 8 (Figure 3.A21 in Appendix)
MaCa01-11 A2: Ac-CAGCTCA-Gly-$NH_2$
yield: 8%; $t_r$: 2.57 min; ε (260 nm): 67500 $M^{-1}cm^{-1}$; MW calculated: 1977.94 [M], observed (from ESI-MS deconvolution): 1977.6; ESI-MS: m/z observed (calcd): 990.3 (990.0) $[M+2H]^{2+}$, 660.5 (660.3) $[M+3H]^{3+}$, 495.6 (495.5) $[M+4H]^{4+}$, 396.7 (396.6)$[M+5H]^{5+}$.

Entry 9 (Figure 3.A22 in Appendix)
MaCa01-10 A2: Ac-GAGTCTA-Gly-$NH_2$
yield: 10%; $t_r$: 2.95 min; ε (260 nm): 74600 $M^{-1}cm^{-1}$; MW calculated: 2032.97 [M], observed (from ESI-MS deconvolution): 2032.6 ESI-MS: m/z observed (calcd): 1017.6 (1017.5) $[M+2H]^{2+}$, 678.8 (678.7) $[M+3H]^{3+}$, 509.3 (590. 2) $[M+4H]^{4+}$, 407.9 (407.6) $[M+5H]^{5+}$.

Entry 10 (Figure 3.A23 in Appendix)
MaCa01-12 A1: Ac-CTAGA-Gly-$NH_2$
yield: 7.2%; $t_r$: 2.55 min; ε (260 nm): 54300 $M^{-1}cm^{-1}$; ; MW calculated: 1475.44 [M], ESI-MS: m/z observed (calcd): 738.4 (738.7) $[M+2H]^{2+}$, 492.7 (492.8) $[M+3H]^{3+}$, 369.9 (369.9) $[M+4H]^{4+}$.

Entry 11 (Figure 3.A24 in Appendix)
MaCa01-11 A1: Ac-GCTCA-Gly-$NH_2$
yield: 7.2%; $t_r$: 2.57 min; ε (260 nm): 47200 $M^{-1}cm^{-1}$; MW calculated: 1451.42 [M]; ESI-MS: m/z observed (calcd): 726.6 (726.7) $[M+2H]^{2+}$, 484.8 (484.8) $[M+3H]^{3+}$, 364.0 (364.0)$[M+4H]^{4+}$.

Entry 12 (Figure 3.A25 in Appendix)
MaCa01-10 A1: Ac-GTCTA-Gly-$NH_2$
yield: 12%; $t_r$: 2.66 min; ε (260 nm): 49200 $M^{-1}cm^{-1}$; MW calculated: 1466.43 [M]; ESI-MS: m/z observed (calcd): 734.1 (734.2) $[M+2H]^{2+}$, 489.9 (489.8) $[M+3H]^{3+}$.

Entry 13 (Figure 3.A26 in Appendix)
MaCa01-01 A4: Ac-TCTAGTACGGG-Gly-$NH_2$
yield: 12%; $t_r$: 2.88 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 3133.02 [M], observed (from ESI-MS deconvolution): 3132.8; ESI-MS: m/z observed (calcd): 1566.4 (1567.5) $[M+2H]^{2+}$, 1045.4 (1045.3)$[M+3H]^{3+}$, 784.5 (784.3)$[M+4H]^{4+}$, 627.8 (627.6) $[M+5H]^{5+}$, 523.4 (523.2) $[M+6H]^{6+}$.

Entry 14 (Figure 3.A27 in Appendix)
MaCa01-02 A4: Ac-TGAGCTGCCAT-Gly-$NH_2$
yield: 12%; $t_r$: 2.96 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 3093.00 [M], observed (from ESI-MS deconvolution): 3092.6; ESI-MS: m/z observed (calcd): 1032.5 (1032.0) $[M+3H]^{3+}$, 774.3 (774.3) $[M+4H]^{4+}$, 619.9 (619.6) $[M+5H]^{5+}$, 516.5 (516.5) $[M+6H]^{6+}$.

Entry 15 (Figure 3.A28 in Appendix)
MaCa01-03 A4: Ac-TAGACTCAGTG-Gly-$NH_2$
yield: 20%; $t_r$: 2.91 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; MW calculated: 3117.02 [M], observed (from ESI-MS deconvolution): 3118.0; ESI-MS: m/z observed (calcd): 1040.2 (1040.0) $[M+3H]^{3+}$, 780.3 (780.3) $[M+4H]^{4+}$, 624.6 (624.4) $[M+5H]^{5+}$, 520.7 (520.5) $[M+6H]^{6+}$.

Entry 16 (Figure 3.A29 in Appendix)
MaCa01-01 A3: Ac-TCTAGTACG-Gly-$NH_2$
yield: 12%; $t_r$: 2.76 min; ε (260 nm): 89800 $M^{-1}cm^{-1}$; MW calculated: 2550.48 [M], observed (from ESI-MS deconvolution): 2550.6; ESI-MS: m/z observed (calcd) 1276.2 (1276.2) $[M+2H]^{2+}$, 851.4 (851.2) $[M+3H]^{3+}$, 638.7 (638.6) $[M+4H]^{4+}$, 511.2 (511.1) $[M+5H]^{5+}$.

Entry 17 (Figure 3.A30 in Appendix)
MaCa01-02 A3: Ac-TGAGCTGCC-Gly-$NH_2$
yield: 22%; $t_r$: 2.79 min; ε (260 nm): 85800 $M^{-1}cm^{-1}$; MW calculated: 2551.47 [M], observed (from ESI-MS deconvolution): 2550.80; ESI-MS: m/z observed (calcd): 851.4 (851.4) $[M+3H]^{3+}$, 638.9 (638.8) $[M+4H]^{4+}$, 511.3 (511.2) $[M+5H]^{5+}$, 426.3 (426.2)$[M+6H]^{6+}$.

Entry 18 (Figure 3.A31 in Appendix)
MaCa01-03 A3: Ac-TAGACTCAG-Gly-$NH_2$
yield: 28%; $t_r$: 2.86 min; ε (260 nm): 94900 $M^{-1}cm^{-1}$; MW calculated: 2559.49 [M], observed (from ESI-MS deconvolution): 2560.4; ESI-MS: m/z observed (calcd): 854.1 (854.2) $[M+3H]^{3+}$, 641.0 (640.9) $[M+4H]^{4+}$, 513.0 (512.9) $[M+5H]^{5+}$.

Entry 19 (Figure 3.A32 in Appendix)
MaCa01-01 A2: Ac-TCTAGTA-Gly-$NH_2$
yield: 10%; $t_r$: 2.81 min; ε (260 nm): 71500 $M^{-1}cm^{-1}$; MW calculated: 2007.96 [M], observed (from ESI-MS deconvolution): 2007.2; ESI-MS: m/z observed (calcd): 1005.1 (1005.0) $[M+2H]^{2+}$, 670.3 (670.3) $[M+3H]^{3+}$, 503.0 (503.0) $[M+4H]^{4+}$.

Entry 20 (Figure 3.A33 in Appendix)
MaCa01-02 A2: Ac-TGAGCTG-Gly-$NH_2$
yield: 25%; $t_r$: 7.89 min; ε (260 nm): 72600 $M^{-1}cm^{-1}$; MW calculated: 2048.97 [M], observed (from ESI-MS deconvolution): 2048.80; ESI-MS: m/z observed (calcd): 1025.4 (1025.4) $[M+2H]^{2+}$, 683.9 (683.9)$[M+3H]^{3+}$, 513.2 (513.2) $[M+4H]^{4+}$.

Entry 21 (Figure 3.A34 in Appendix)
MaCa01-03 A2: Ac-TAGACTC-Gly-$NH_2$
yield: 26%; $t_r$: 2.78 min; ε (260 nm): 69500 $M^{-1}cm^{-1}$; MW calculated: 1992.95 [M], observed (from ESI-MS deconvolution): 1992.6; ESI-MS: m/z observed (calcd): 997.7 (997.5) $[M+2H]^{2+}$, 665.3 (665.3) $[M+3H]^{3+}$, 499.3 (499.2) $[M+4H]^{4+}$.

Entry 22 (Figure 3.A35 in Appendix)
MaCa01-01 A1: Ac-TCTAG-Gly-$NH_2$
yield: 16%; $t_r$: 2.67 min; ε (260 nm): 49200 $M^{-1}cm^{-1}$; MW calculated: 1466.43 [M]; ESI-MS: m/z observed (calcd): 734.1 (734.2) $[M+2H]^{2+}$, 489.9 (489.8) $[M+3H]^{3+}$.

Entry 23 (Figure 3.A36 in Appendix)
MaCa01-02 A1: Ac-TGAGC-Gly-$NH_2$
yield: 13%; $t_r$: 2.61 min; ε (260 nm): 52300 $M^{-1}cm^{-1}$; MW calculated: 1491.44 [M]; ESI-MS: m/z observed (calcd): 746.6 (746.7) $[M+2H]^{2+}$, 498.2 (498.1) $[M+3H]^{3+}$, 374.0 (373.9) $[M+4H]^{4+}$.

Entry 24 (Figure 3.A37 in Appendix)
MaCa01-03 A1: Ac-TAGAC-Gly-$NH_2$
yield: 30%; $t_r$: 2.61 min; ε (260 nm): 54300 $M^{-1}cm^{-1}$; MW calculated: 1475.44 [M]; ESI-MS: m/z observed (calcd): 738.4 (738.7) $[M+2H]^{2+}$, 492.8 (492.8) $[M+3H]^{3+}$, 369.8 (369.8) $[M+4H]^{4+}$.

Entry 25 (Figure 3.A38 in Appendix)
MaCa01-16 A4: Ac-GGGCATGATCT-Gly-$NH_2$
yield: 8%; $t_r$: 2.90 min; ε (260 nm): 113200 $M^{-1}cm^{-1}$; MW calculated: 3133.02 [M], observed (from ESI-MS deconvolution): 3132.6; ESI-MS: m/z observed (calcd): 1566.2 (1567.5)$[M+2H]^{2+}$, 1045.3 (1045.3) $[M+3H]^{3+}$, 784.5 (784.3) $[M+4H]^{4+}$, 627.8 (627.6) $[M+5H]^{5+}$, 523.4 (523.2) $[M+6H]^{6+}$.

Entry 26 (Figure 3.A39 in Appendix)
MaCa01-15 A4: Ac-TACCGTCGAGT-Gly-$NH_2$
yield: 8%; $t_r$: 3.03 min; ε (260 nm): 108100 $M^{-1}cm^{-1}$; MW calculated: 3093.00 [M], observed (from ESI-MS deconvolution): 3092.6; ESI-MS: m/z observed (calcd): 1032.1 (1032.0) $[M+3H]^{3+}$, 774.3 (774.3) $[M+4H]^{4+}$, 619.7 (619.6) $[M+5H]^{5+}$, 516.4 (516.5) $[M+6H]^{6+}$.

Entry 27 (Figure 3.A40 in Appendix)
MaCa01-14 A4: Ac-GTGACTCAGAT-Gly-$NH_2$
yield: 5%; $t_r$: 2.96 min; ε (260 nm): 115200 $M^{-1}cm^{-1}$; ; MW calculated: 3117.02 [M], observed (from ESI-MS deconvolution): 3116.40; ESI-MS: m/z observed (calcd): 1040.1 (1040.0) $[M+3H]^{3+}$, 780.4 (780.2) $[M+4H]^{4+}$, 624.5 (624.4) $[M+5H]^{5+}$, 520.4 (520.5) $[M+6H]^{6+}$.

Entry 28 (Figure 3.A41 in Appendix)
MaCa01-16 A3: Ac-GCATGATCT-Gly-$NH_2$
yield: 7%; $t_r$: 2.79 min; ε (260 nm): 89800 $M^{-1}cm^{-1}$; MW calculated: 2550.48 [M], observed (from ESI-MS deconvolution): 2550.6; ESI-MS: m/z observed (calcd): 851.4 (851.2) $[M+3H]^{3+}$, 638.9 (638.6) $[M+4H]^{4+}$, 511.1 (511.1) $[M+5H]^{5+}$.

Entry 29 (Figure 3.A42 in Appendix)
MaCa01-15 A3: Ac-CCGTCGAGT-Gly-$NH_2$
yield: 6%; $t_r$: 2.95 min; ε (260 nm): 85800 $M^{-1}cm^{-1}$; MW calculated: 2551.47 [M], observed (from ESI-MS deconvolution): 2551.6; ESI-MS: m/z observed (calcd): 851.4 (851.5) $[M+3H]^{3+}$, 639.0 (638.9) $[M+4H]^{4+}$, 511.5 (511.3) $[M+5H]^{5+}$, 426.2 (426.2) $[M+6H]^{6+}$.

Entry 30 (Figure 3.A43 in Appendix)
MaCa01-14 A3: Ac-GACTCAGAT-Gly-$NH_2$
yield: 9%; $t_r$: 2.83 min; ε (260 nm): 94900 $M^{-1}cm^{-1}$; MW calculated: 2559.49 [M], observed (from ESI-MS deconvolution): 2559.00; ESI-MS: m/z observed (calcd): 854.4 (854.1) $[M+3H]^{3+}$, 640.9 (640.8) $[M+4H]^{4+}$, 512.9 (512.9) $[M+5H]^{5+}$, 427.6 (427.5) $[M+6H]^{6+}$.

Entry 31 (Figure 3.A44 in Appendix)
MaCa01-16 A2: Ac-ATGATCT-Gly-$NH_2$
yield: 9%; $t_r$: 2.78 min; ε (260 nm): 71500 $M^{-1}cm^{-1}$; MW calculated: 2007.96 [M], observed (from ESI-MS deconvolution): 2007.6; ESI-MS: m/z observed (calcd): 1005.2 (1005.0) $[M+2H]^{2+}$, 670.6 (670.3) $[M+3H]^{3+}$, 503.1 (503.0) $[M+4H]^{4+}$.

Entry 32 (Figure 3.A45 in Appendix)
MaCa01-15 A2: Ac-GTCGAGT-Gly-$NH_2$
yield: 6%; $t_r$: 2.78 min; ε (260 nm): 72600 $M^{-1}cm^{-1}$; MW calculated: 2048.97 [M], observed (from ESI-MS deconvolution): 2048.4; ESI-MS: m/z observed (calcd): 1025.5 (1025.5) $[M+2H]^{2+}$, 983.7 (684.0) $[M+3H]^{3+}$, 513.2 (513.2)$[M+4H]^{4+}$, 410.7 (410.8) $[M+5H]^{5+}$.

Entry 33 (Figure 3.A46 in Appendix)
MaCa01-14 A2: Ac-CTCAGAT-Gly-$NH_2$
yield: 32%; $t_r$: 2.74 min; ε (260 nm): 69500 $M^{-1}cm^{-1}$; MW calculated: 1992.95 [M], observed (from ESI-MS deconvolution): 1992.8; ESI-MS: m/z observed (calcd): 997.7 (997.5) $[M+2H]^{2+}$, 665.6 (665.3) $[M+3H]^{3+}$, 499.4 (499.2) $[M+4H]^{4+}$.

Entry 34 (Figure 3.A47 in Appendix)
MaCa01-16 A1: Ac-GATCT-Gly-$NH_2$
yield: 15%; $t_r$: 2.71 min; ε (260 nm): 49200 $M^{-1}cm^{-1}$; MW calculated: 1466.43 [M], observed (from ESI-MS Spectrum): 1468.0; ESI-MS: m/z observed (calcd): 734.1 (734.2) $[M+2H]^{2+}$, 489.9 (489.8) $[M+3H]^{3+}$

Entry 35 (Figure 3.A48 in Appendix)
MaCa01-15 A1: Ac-CGAGT-Gly-$NH_2$
yield: 9%; $t_r$: 2.62 min; ε (260 nm): 52300 $M^{-1}cm^{-1}$; MW calculated: 1491.44 [M]; ESI-MS: m/z observed (calcd): 746.7 (746.7) $[M+2H]^{2+}$, 498.4 (498.1) $[M+3H]^{3+}$, 374.0 (373.9)$[M+4H]^{4+}$.

Entry 36 (Figure 3.A49 in Appendix)
MaCa01-14 A1: Ac-CAGAT-Gly-$NH_2$
yield: 40%; $t_r$: 2.62 min; ε (260 nm): 54300 $M^{-1}cm^{-1}$; MW calculated: 1475.44 [M]; ESI-MS: m/z observed (calcd): 738.6 (738.7)$[M+2H]^{2+}$, 492.8 (492.8) $[M+3H]^{3+}$, 370.0 (369.9) $[M+4H]^{4+}$.

# 3.6 Appendix Chapter 3

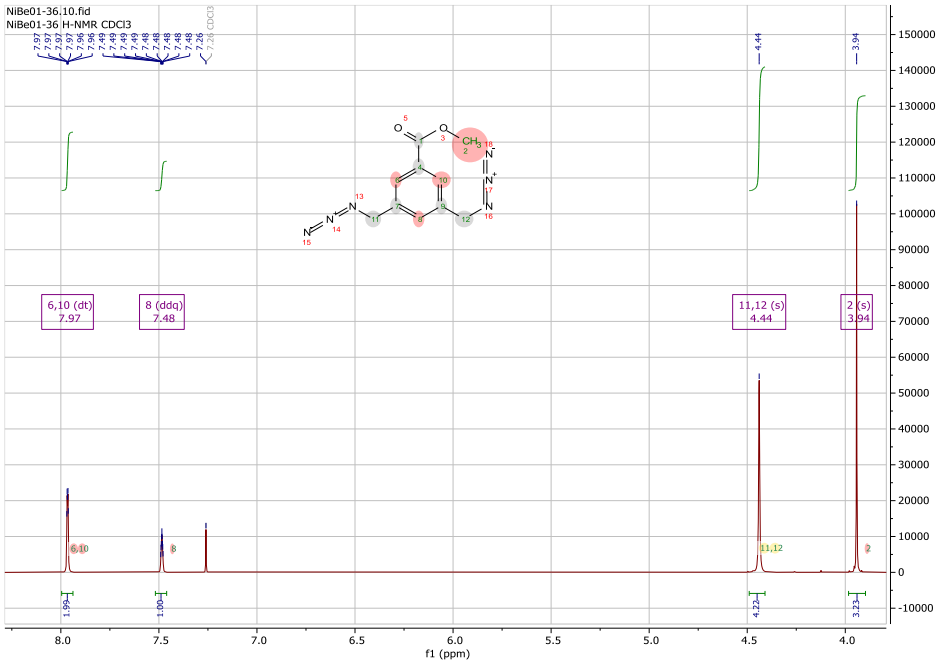



***Figure 3.A1.*** *$^1$H-NMR Spectrum (400MHz, $CDCl_3$) of compound* ***4****.*

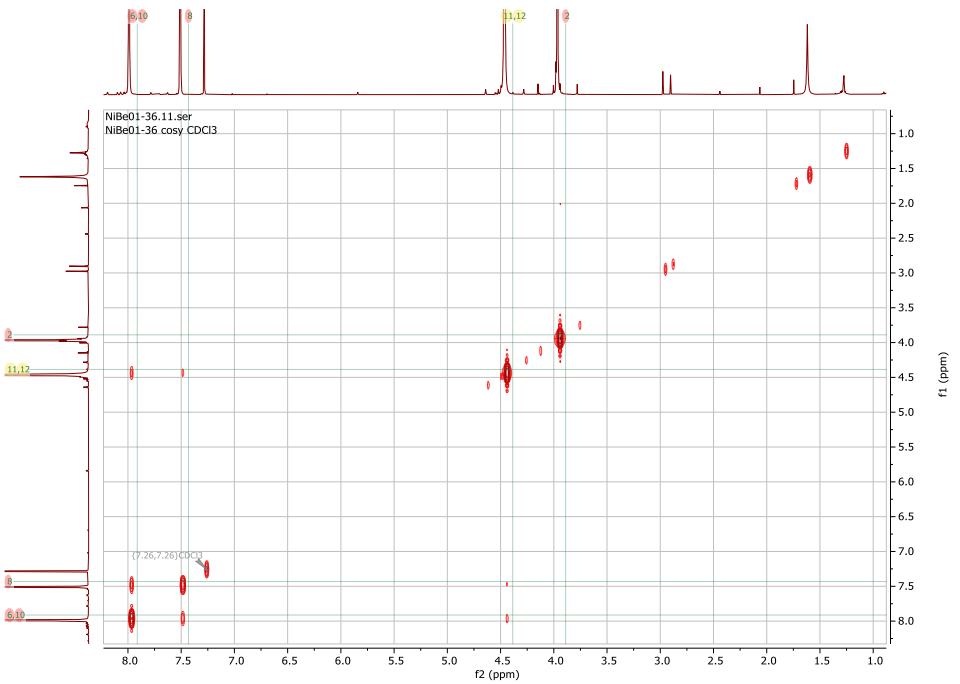



***Figure 3.A2.*** *COSY Spectrum (400MHz, $CDCl_3$) of compound* ***4****.*

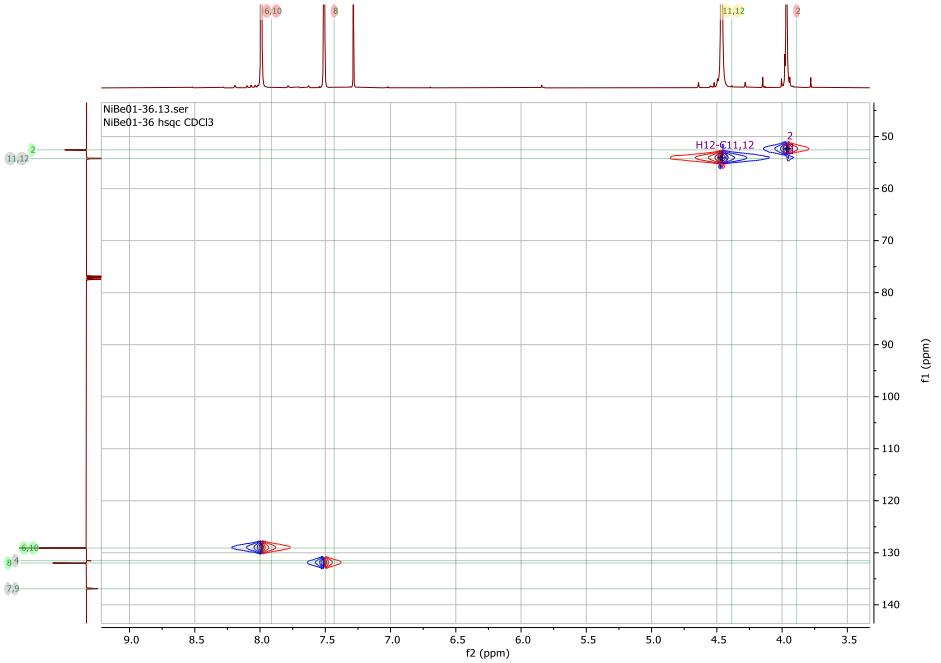



***Figure 3.A3.*** *HSQC Spectrum (400MHz, $CDCl_3$) of compound* ***4****.*

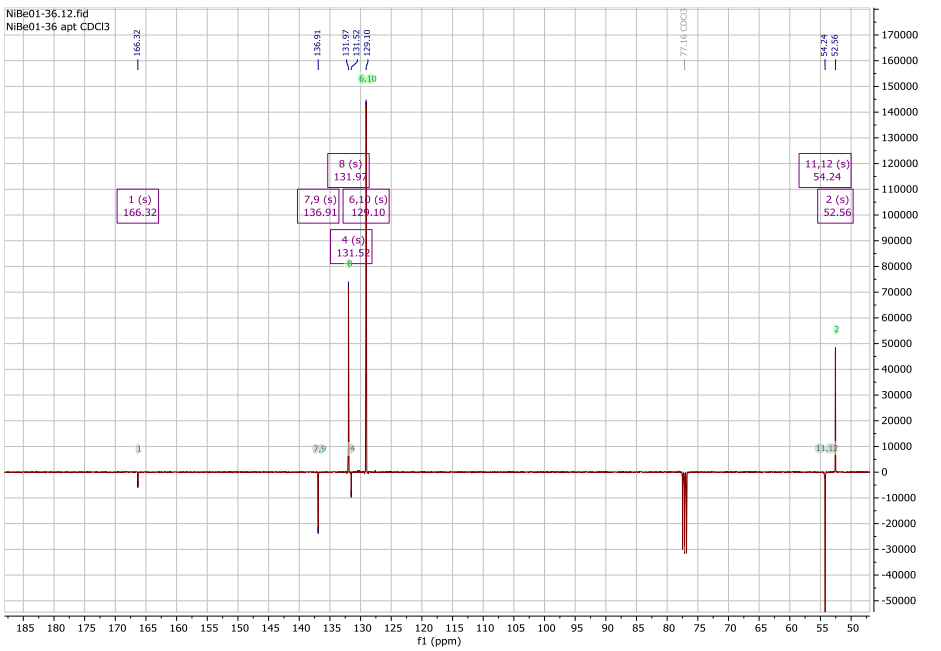



***Figure 3.A4.*** *$^{13}$C-NMR APT (400MHz, $CDCl_3$) of compound* ***4****.*

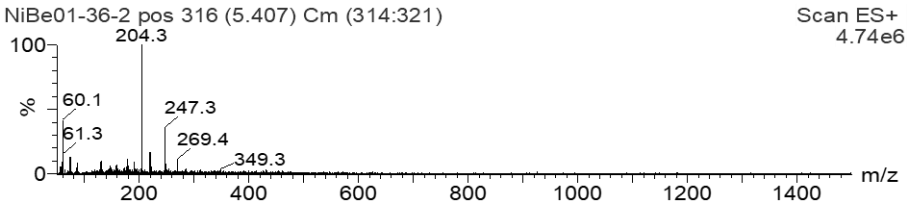



***Figure 3.A5.*** *UPLC-MS of Compound **4**.*

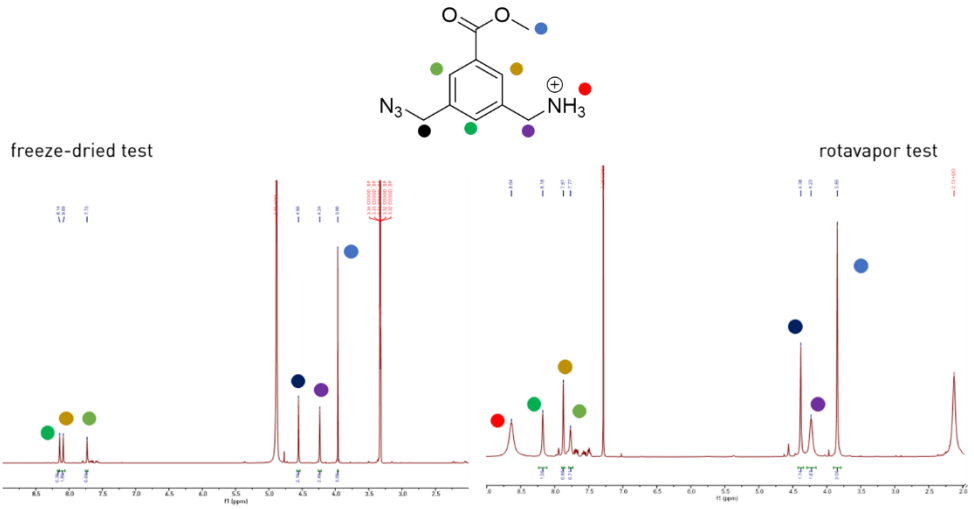



***Figure 3.A6.*** *$^1$H-NMR spectra regarding two drying methods carried out during the synthesis of Compound **3**.*

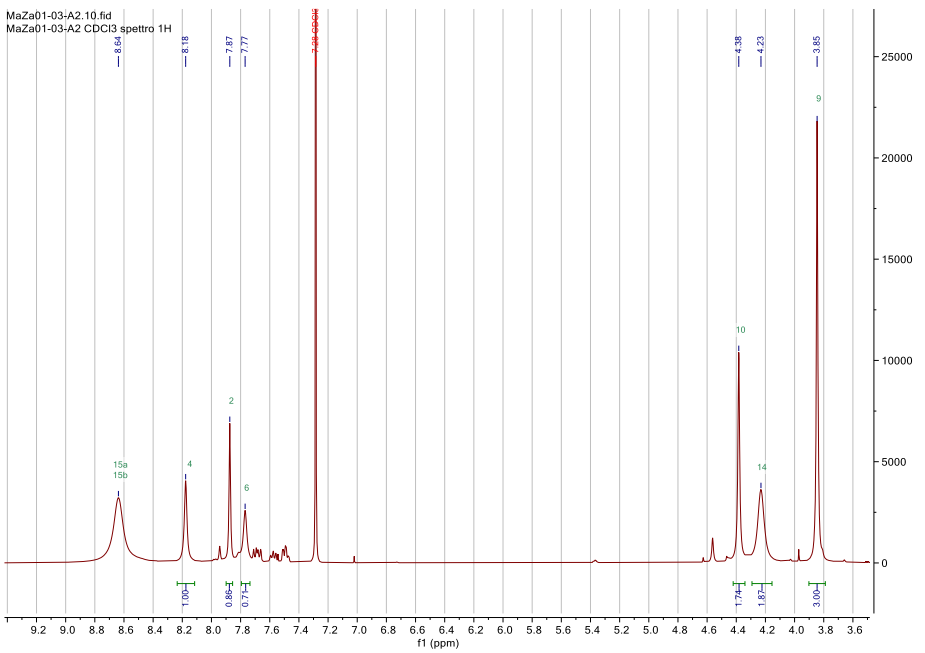



***Figure 3.A7.*** *$^1$H-NMR Spectrum (400MHz, $CDCl_3$) of compound **3**.*

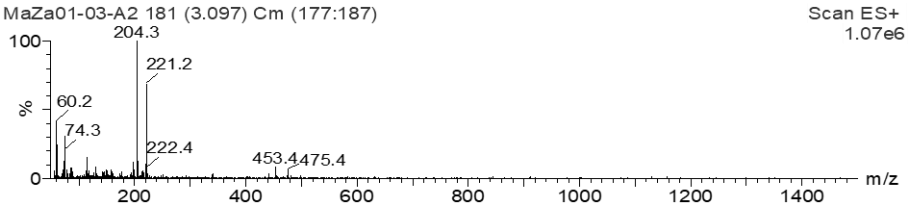



***Figure 3.A8.*** *UPLC-MS of Compound **3**.*

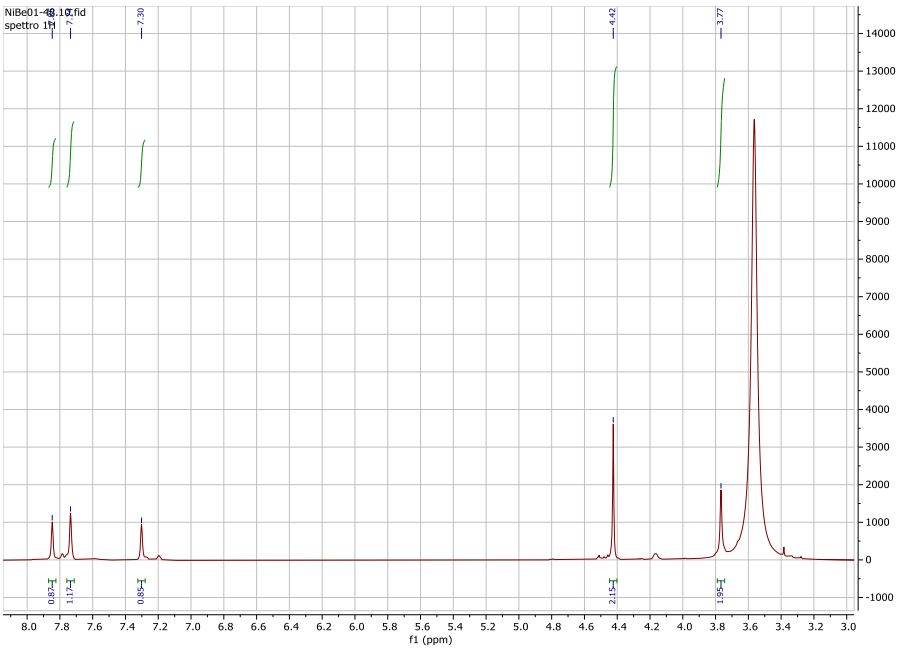



***Figure 3.A9.*** *$^{1}$H-NMR Spectrum (400MHz, DMSO) of Compound **2** via basic hydrolysis.*

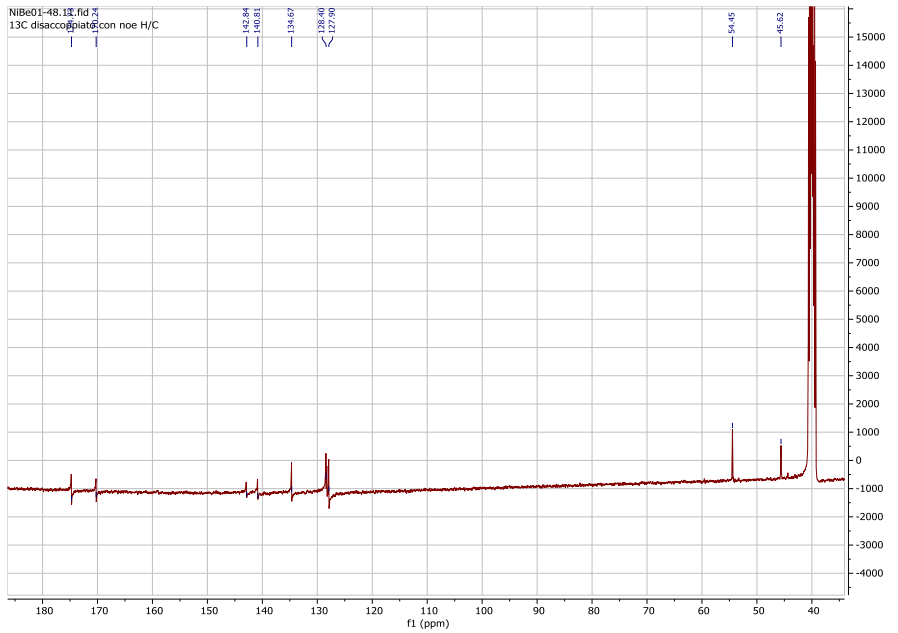



***Figure 3.A10.*** *$^{13}$C-NMR Spectrum (400MHz, DMSO) of compound **2** via basic hydrolysis.*

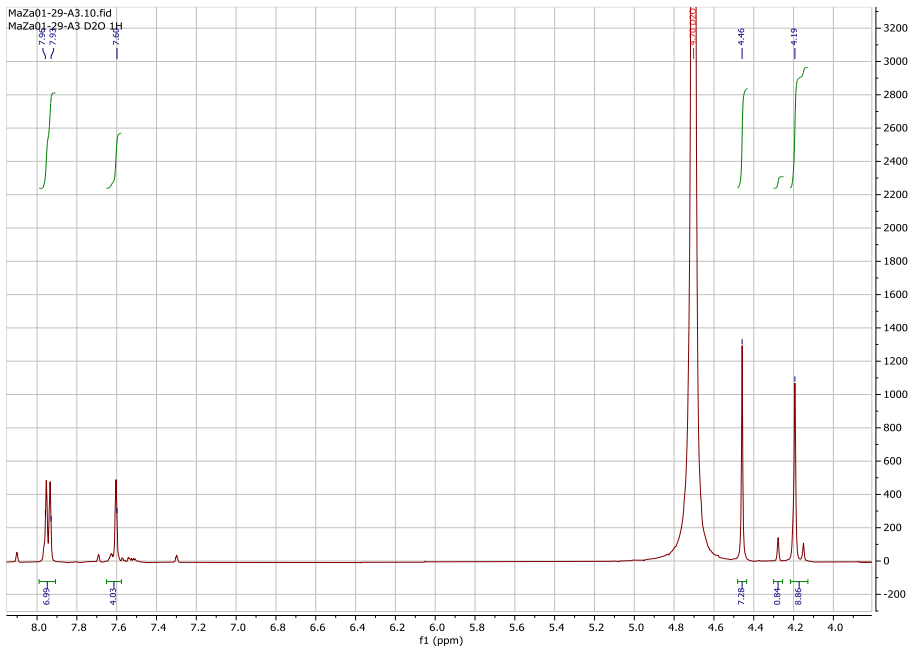



***Figure 3.A11.*** *$^{1}$H-NMR Spectrum (400MHz, $D_2O$) of compound **2** via acidic hydrolysis.*

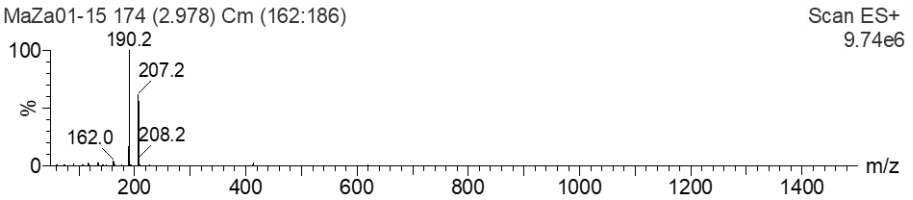



***Figure 3.A12.*** *UPLC-MS of Compound* ***2*** *via acidic hydrolysis.*

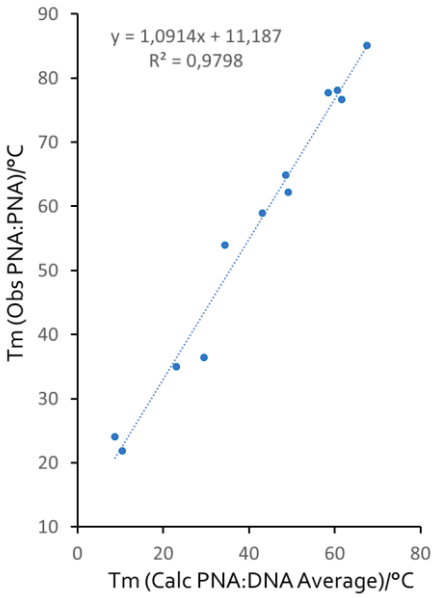



***Figure 3.A13.*** *Correlation between the observed melting temperatures for full-match, full-length, antiparallel PNA:PNA duplexes and those calculated as average value for PNA:DNA duplexes.*

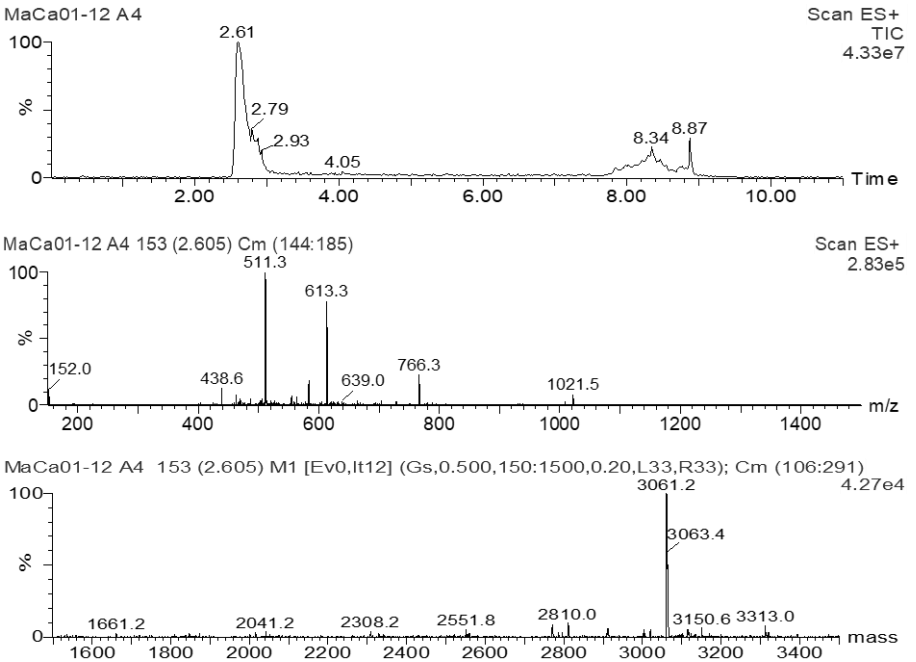



***Figure 3.A14.*** *UPLC-MS of MaCa01-12 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

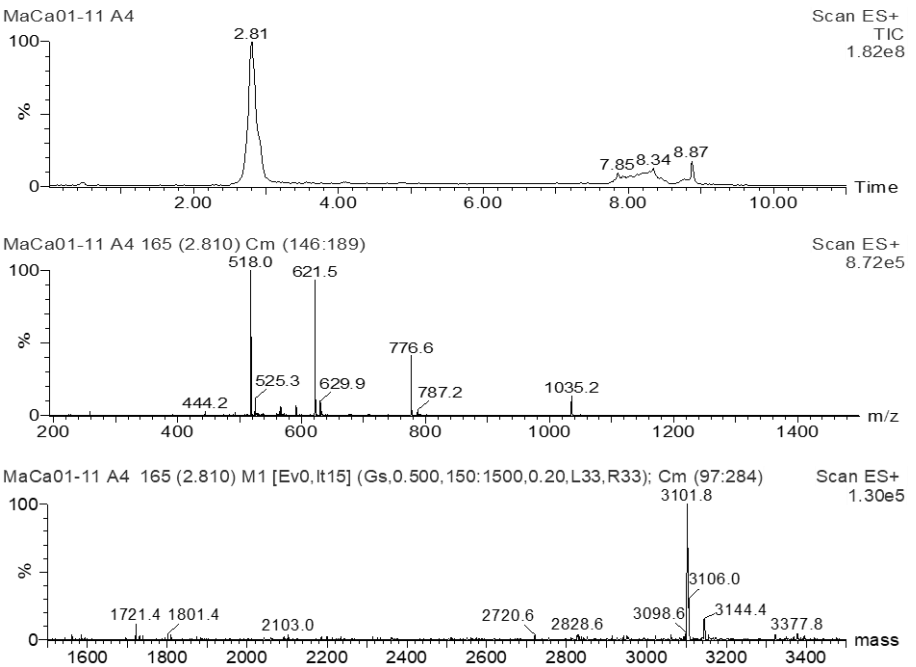



***Figure 3.A15.*** *UPLC-MS of MaCa01-11 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

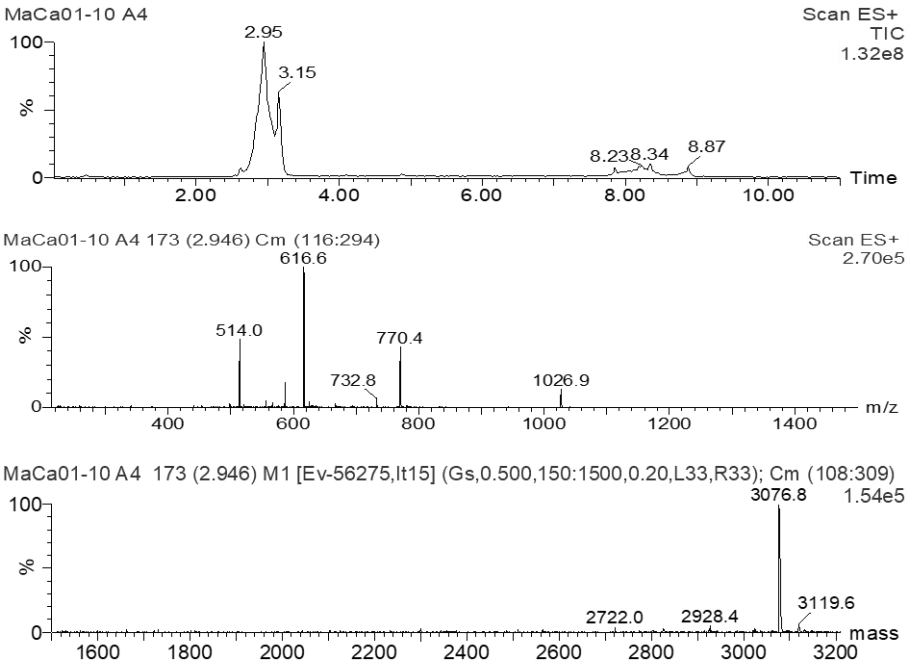



***Figure 3.A16.*** *UPLC-MS of MaCa01-10 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

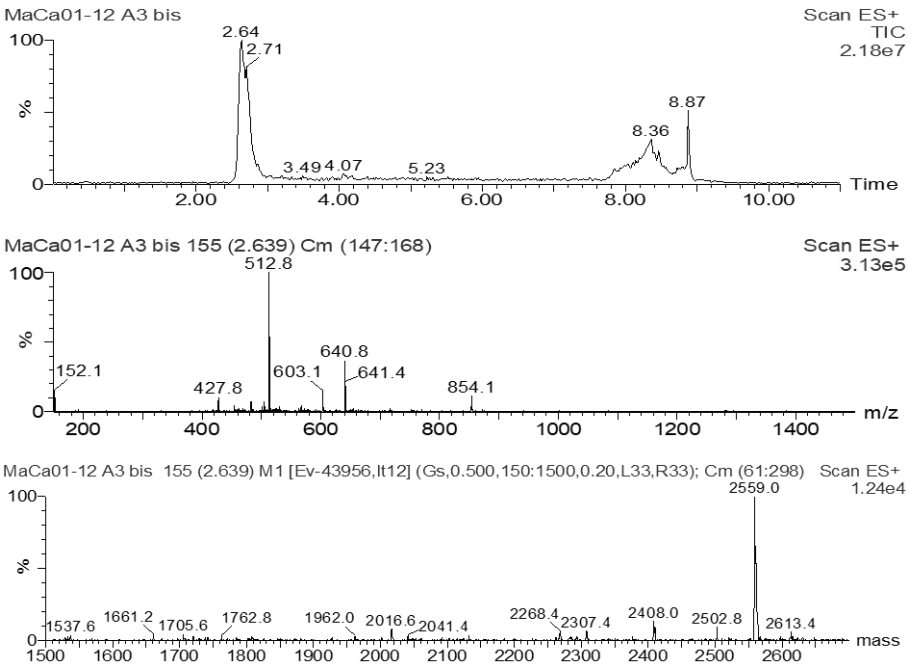



***Figure 3.A17.*** *UPLC-MS of MaCa01-12 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

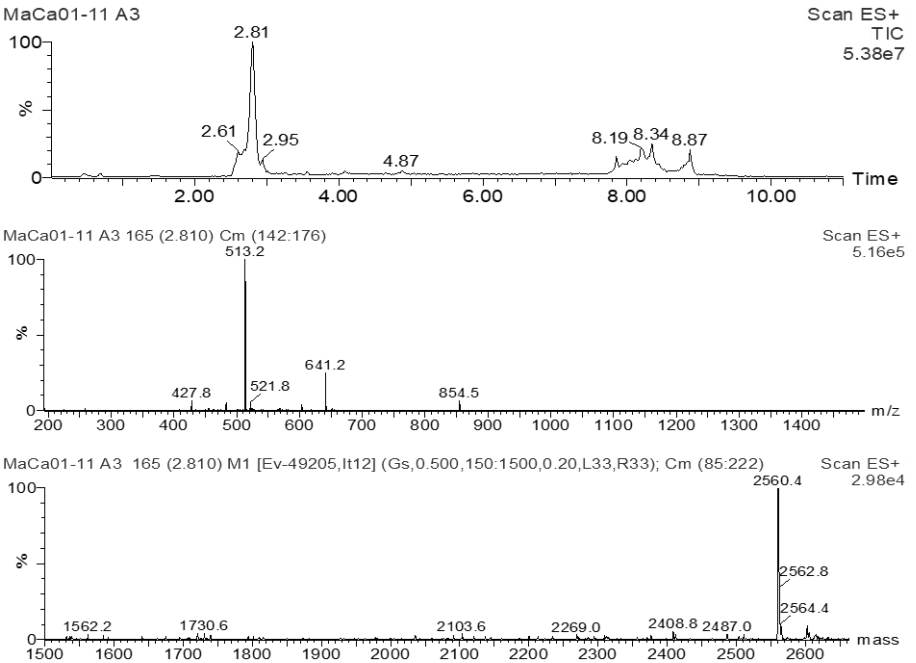



***Figure 3.A18.*** *UPLC-MS of MaCa01-11 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

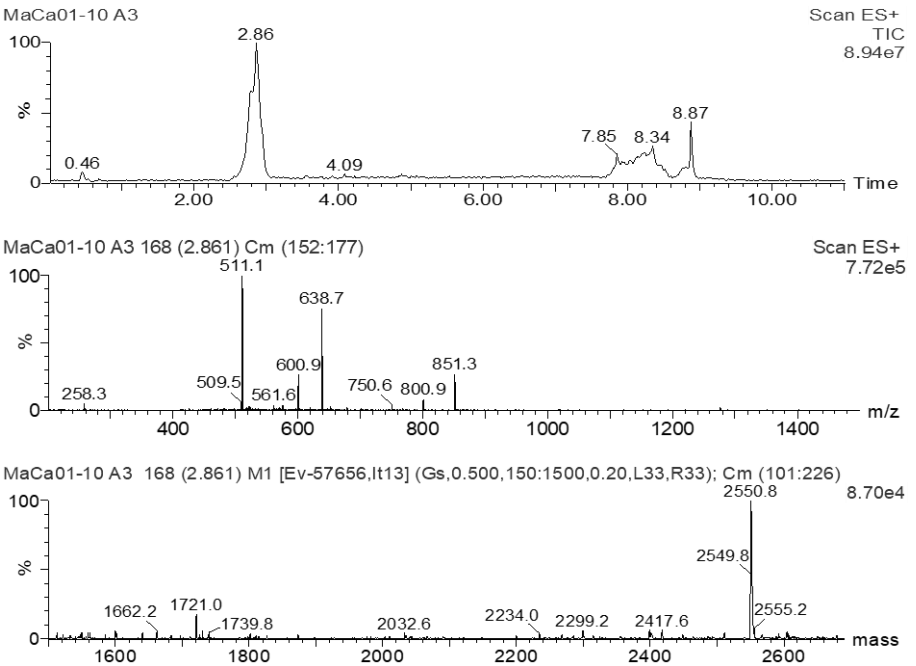



***Figure 3.A19.*** *UPLC-MS of MaCa01-10 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

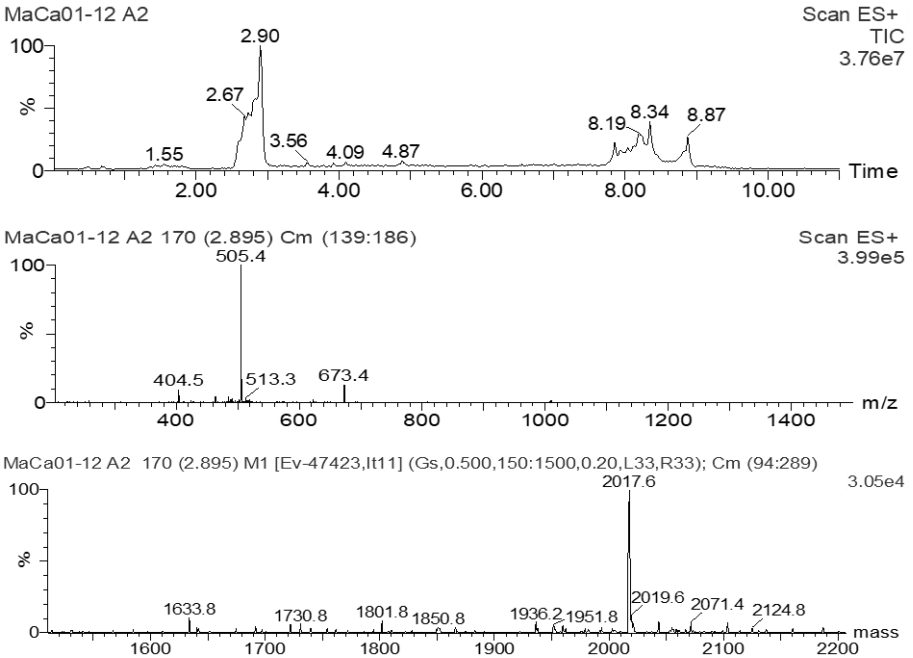



***Figure 3.A20.*** *UPLC-MS of MaCa01-12 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

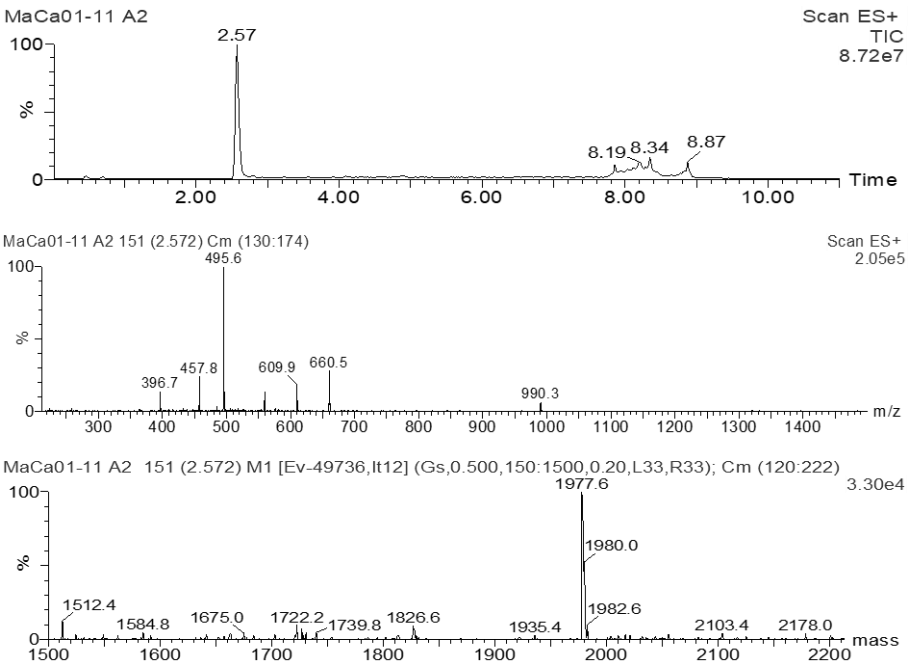



***Figure 3.A21.*** *UPLC-MS of MaCa01-11 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

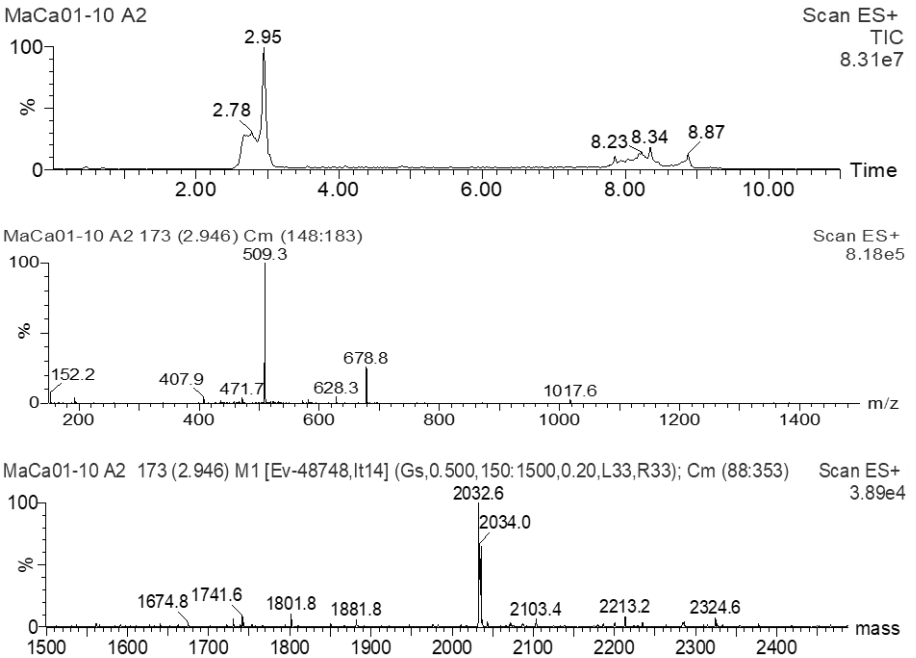



***Figure 3.A22.*** *UPLC-MS of MaCa01-10 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

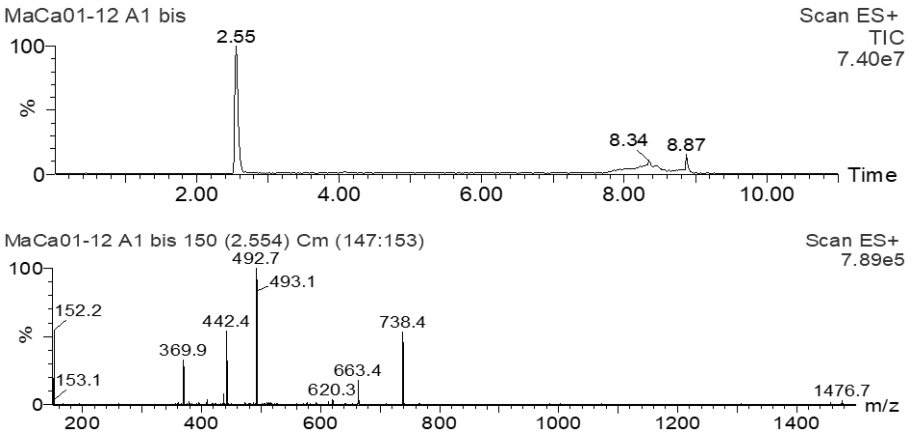



***Figure 3.A23.*** *UPLC-MS of MaCa01-12 A1: chromatogram and mass spectrum.*

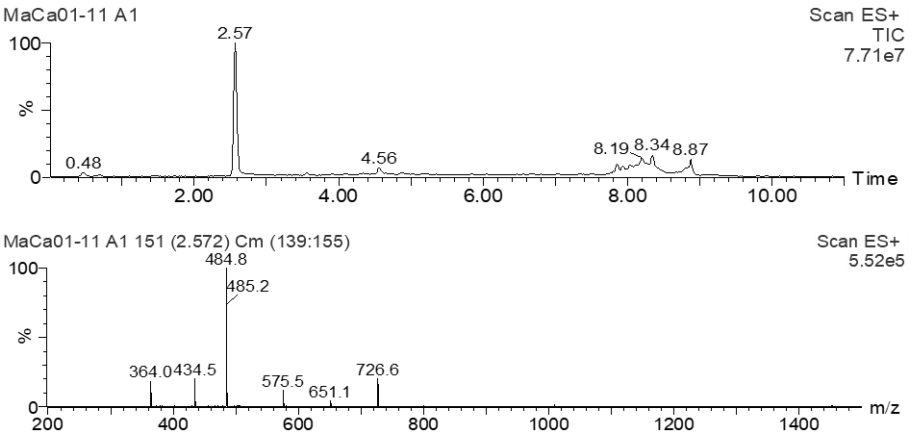



***Figure 3.A24.*** *UPLC-MS of MaCa01-11 A1: chromatogram and mass spectrum.*

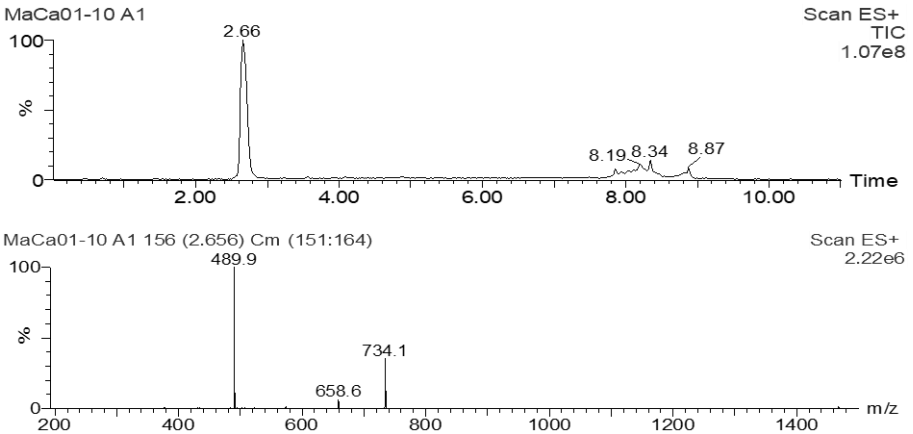



***Figure 3.A25.*** *UPLC-MS of MaCa01-10 A1: chromatogram and mass spectrum.*

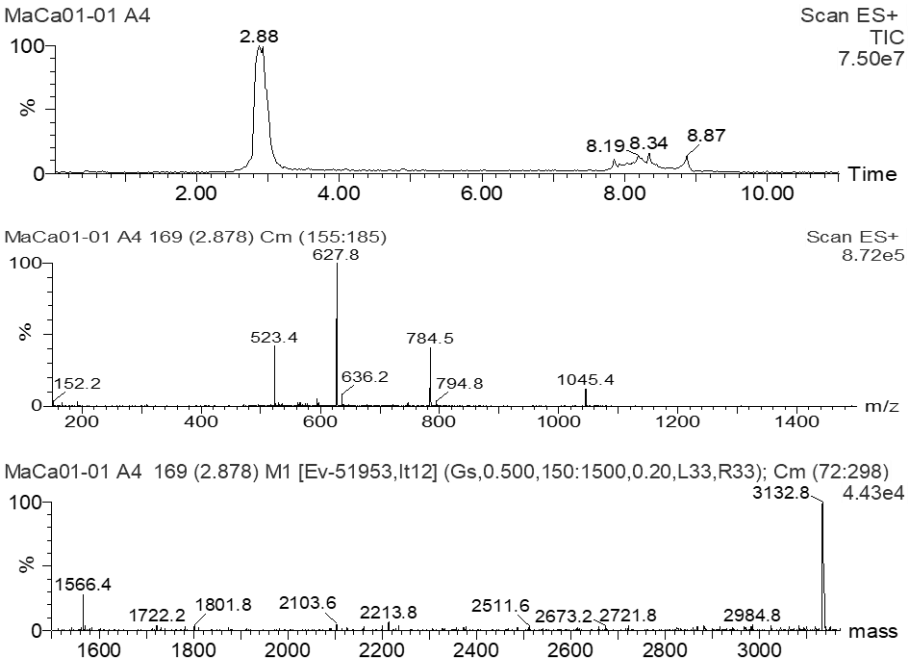



***Figure 3.A26.*** *UPLC-MS of MaCa01-01 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

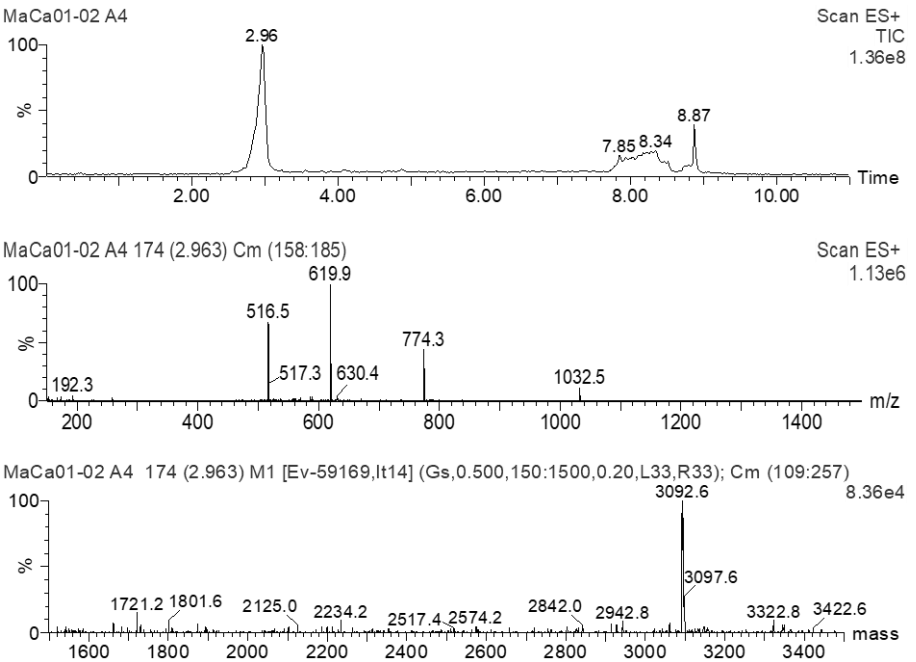



***Figure 3.A27.*** *UPLC-MS of MaCa01-02 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

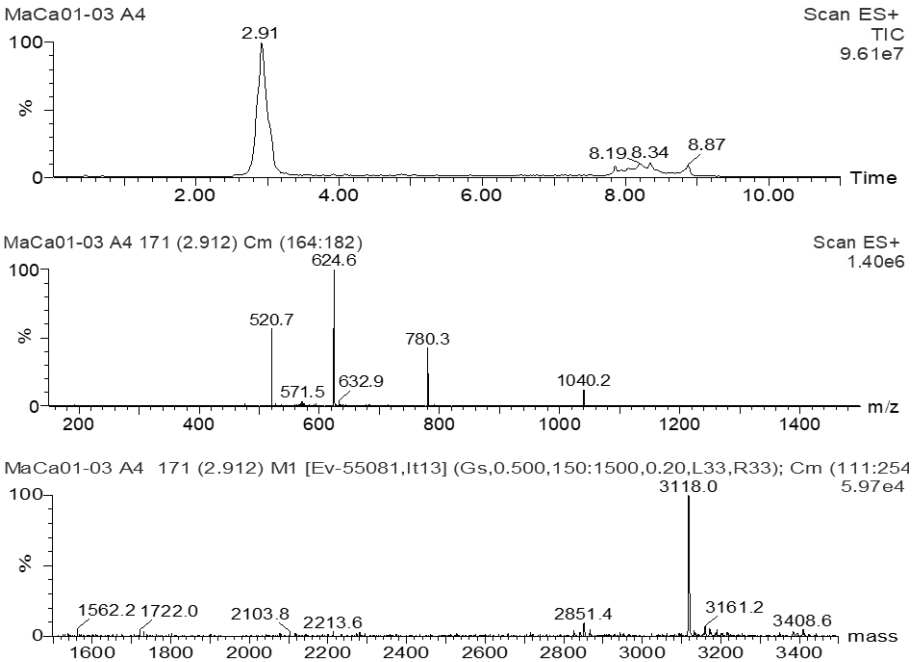



***Figure 3.A28.*** *UPLC-MS of MaCa01-03 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

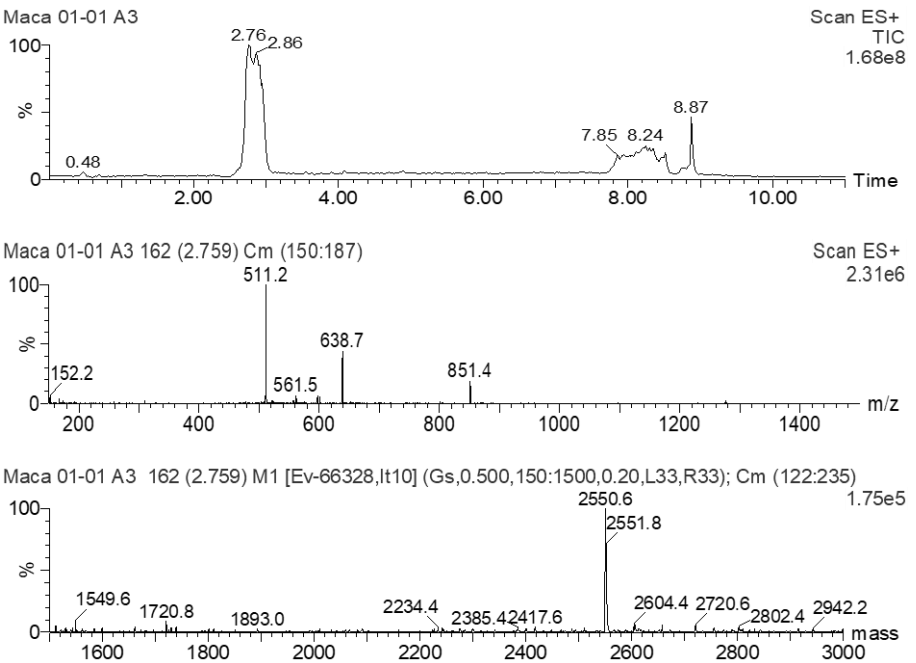



***Figure 3.A29.*** *UPLC-MS of MaCa01-01 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

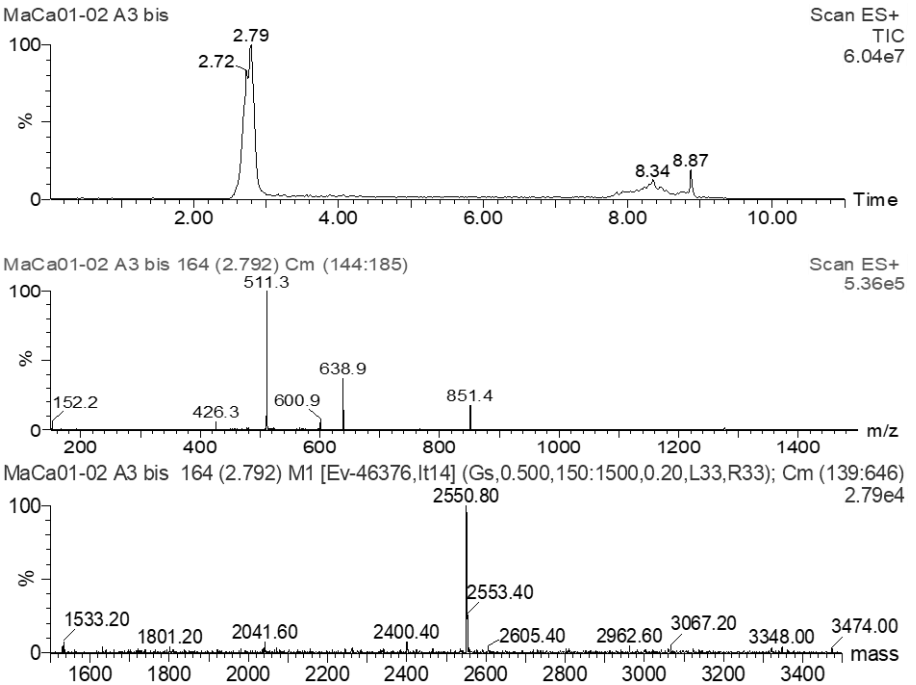



***Figure 3.A30.*** *UPLC-MS of MaCa01-02 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

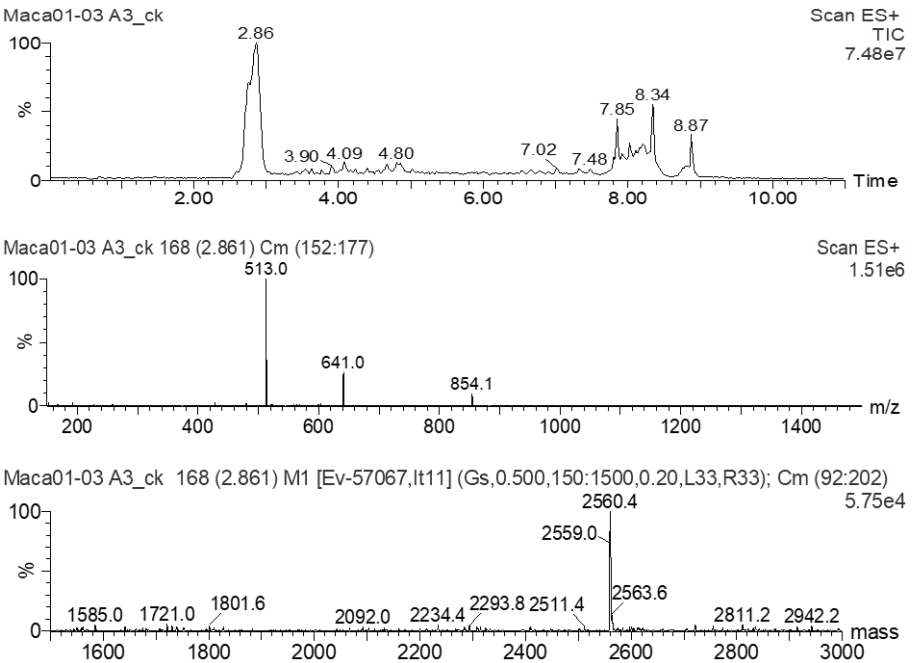



***Figure 3.A31.*** *UPLC-MS of MaCa01-03 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

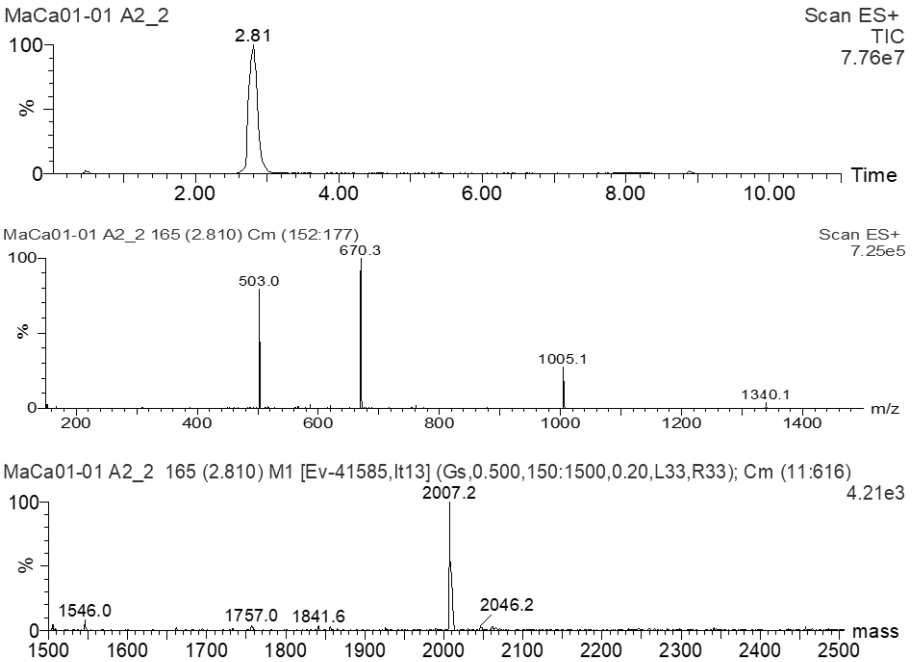



***Figure 3.A32.*** *UPLC-MS of MaCa01-01 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

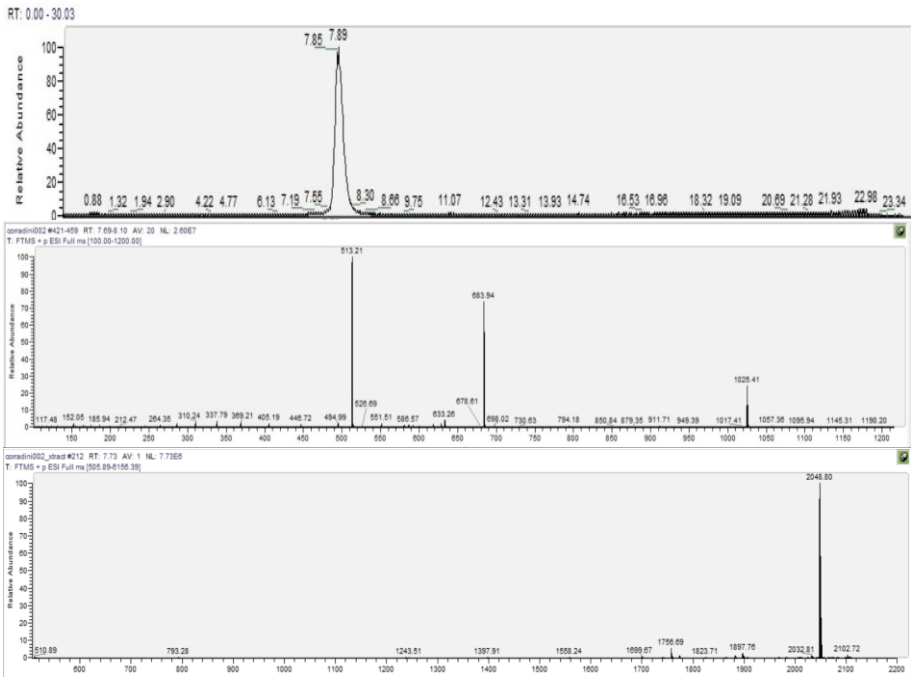



***Figure 3.A33.*** *Orbitrap of MaCa01-02 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

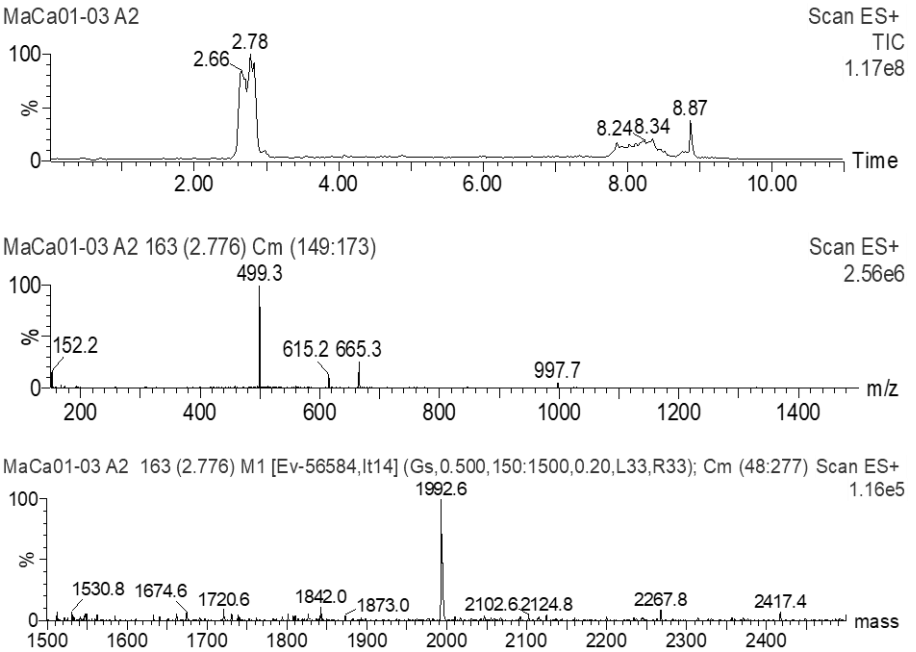



***Figure 3.A34.*** *UPLC-MS of MaCa01-03 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

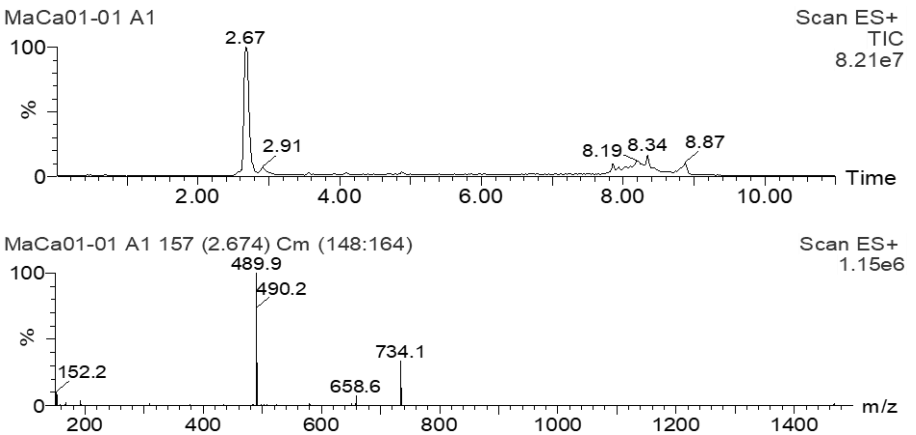



***Figure 3.A35.*** *UPLC-MS of MaCa01-01 A1: chromatogram and mass spectrum.*

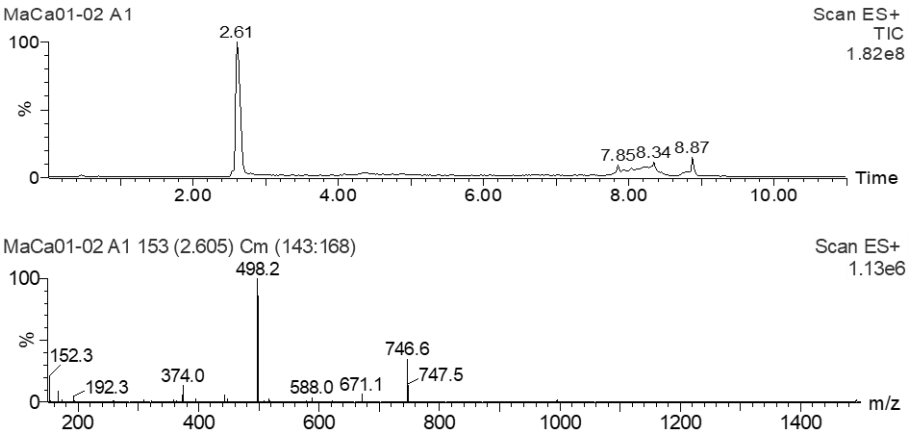



***Figure 3.A36.*** *UPLC-MS of MaCa01-02 A1: chromatogram and mass spectrum.*

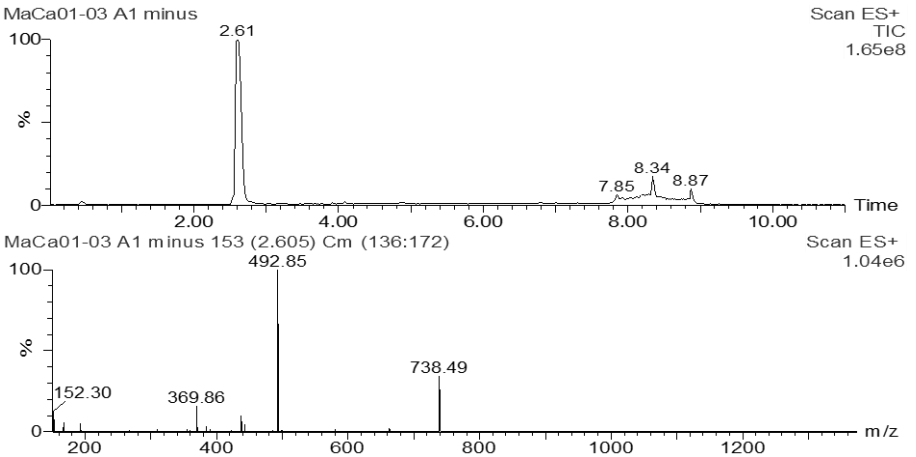



***Figure 3.A37.*** *UPLC-MS of MaCa01-03 A1: chromatogram and mass spectrum.*

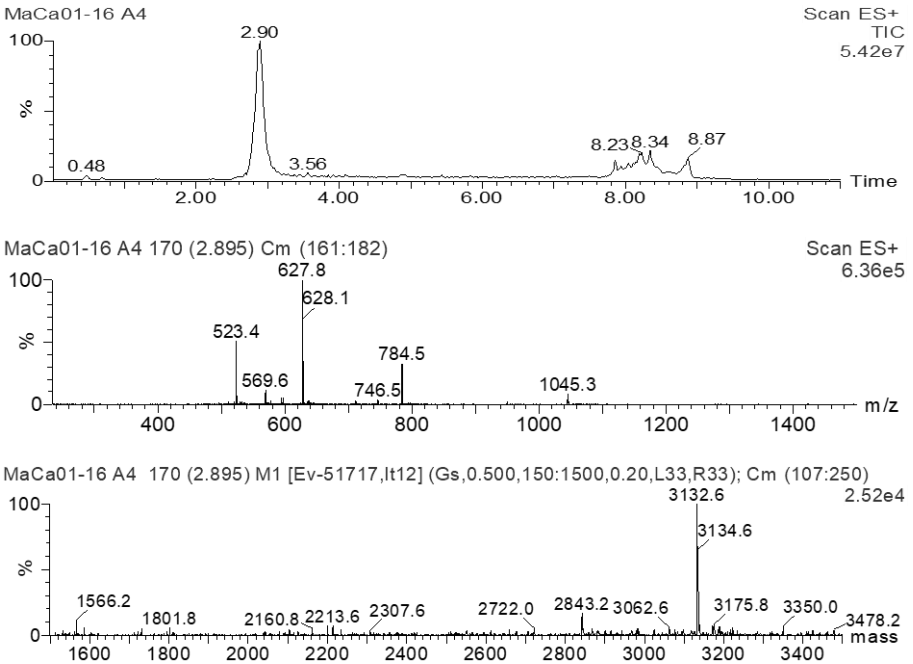



***Figure 3.A38.*** *UPLC-MS of MaCa01-16 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

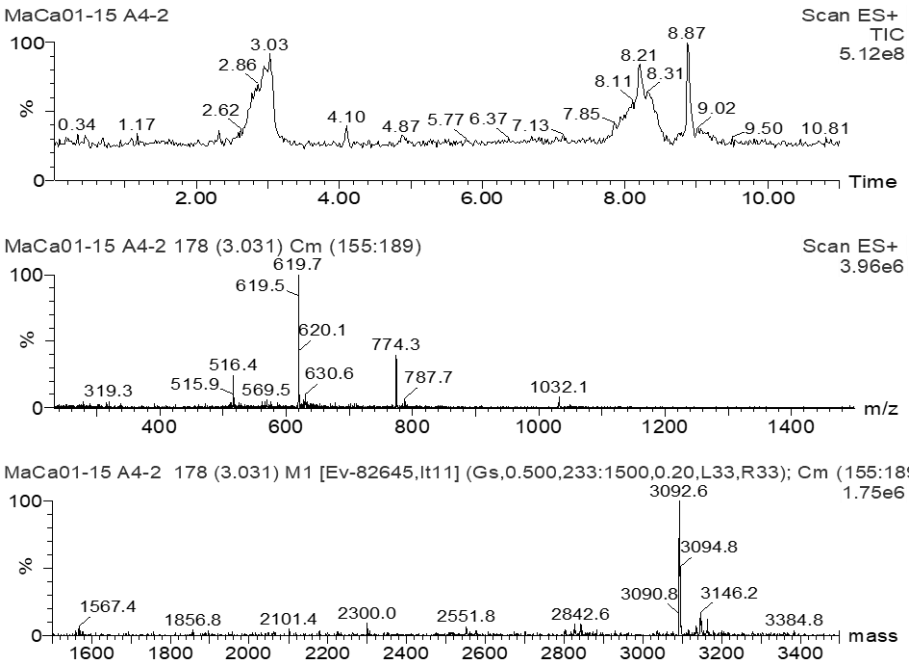



***Figure 3.A39.*** *UPLC-MS of MaCa01-15 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

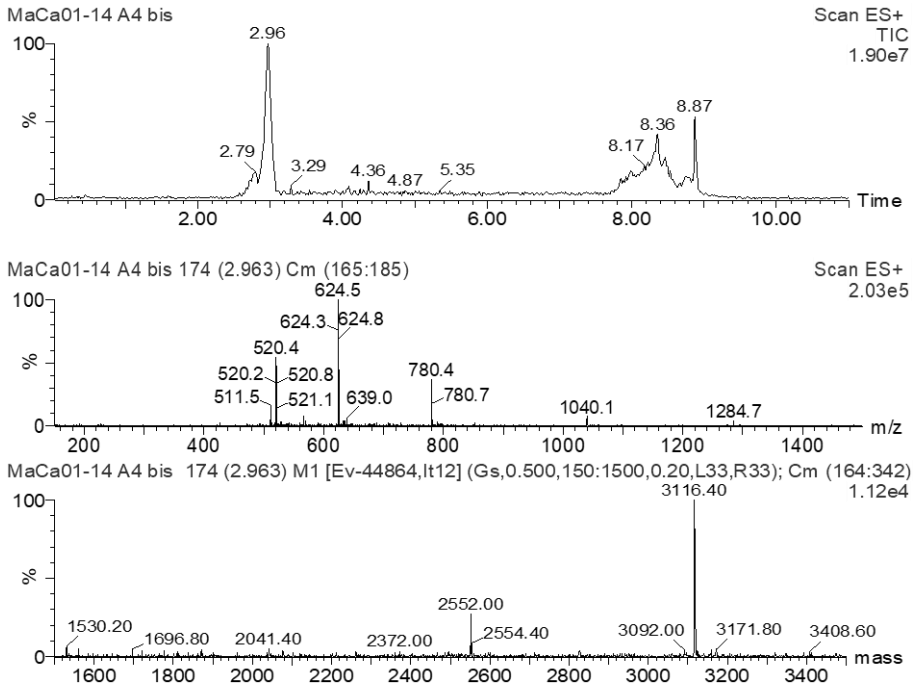



***Figure 3.A40.*** *UPLC-MS of MaCa01-14 A4: chromatogram, mass spectrum and deconvoluted spectrum.*

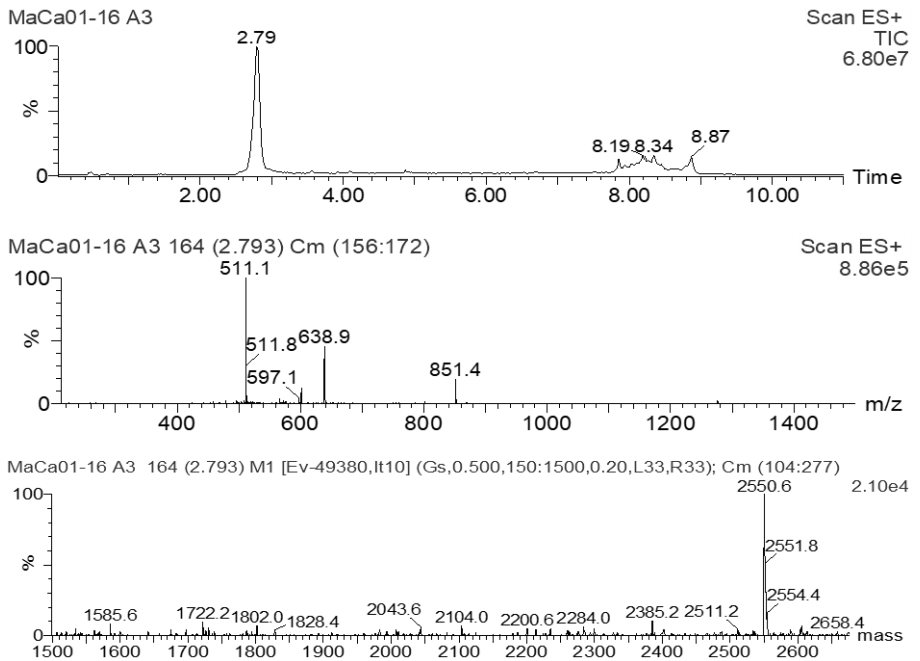



***Figure 3.A41.*** *UPLC-MS of MaCa01-16 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

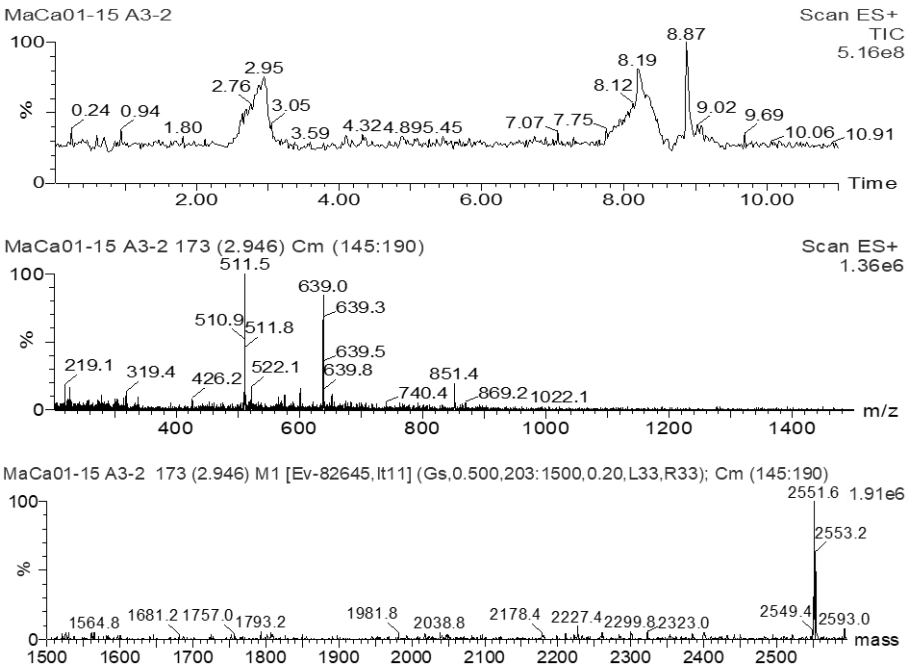



***Figure 3.A42.*** *UPLC-MS of MaCa01-15 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

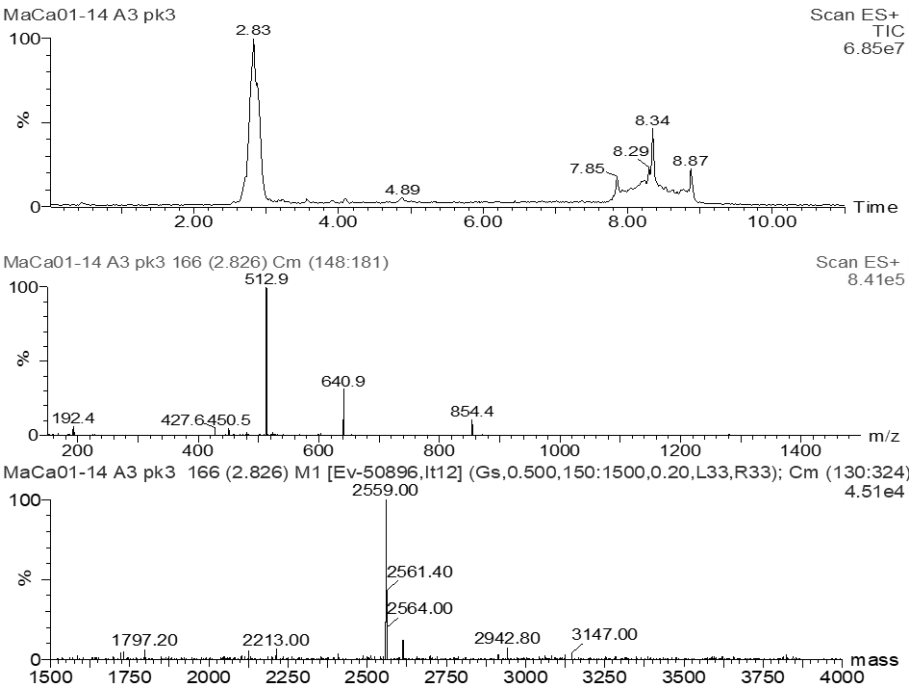



***Figure 3.A43.*** *UPLC-MS of MaCa01-14 A3: chromatogram, mass spectrum and deconvoluted spectrum.*

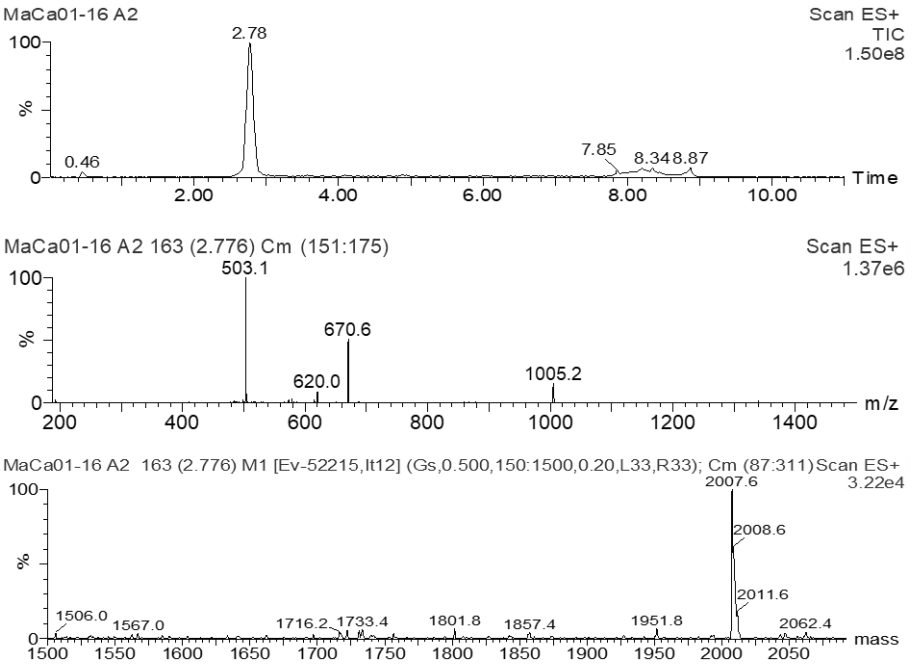



***Figure 3.A44.*** *UPLC-MS of MaCa01-16 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

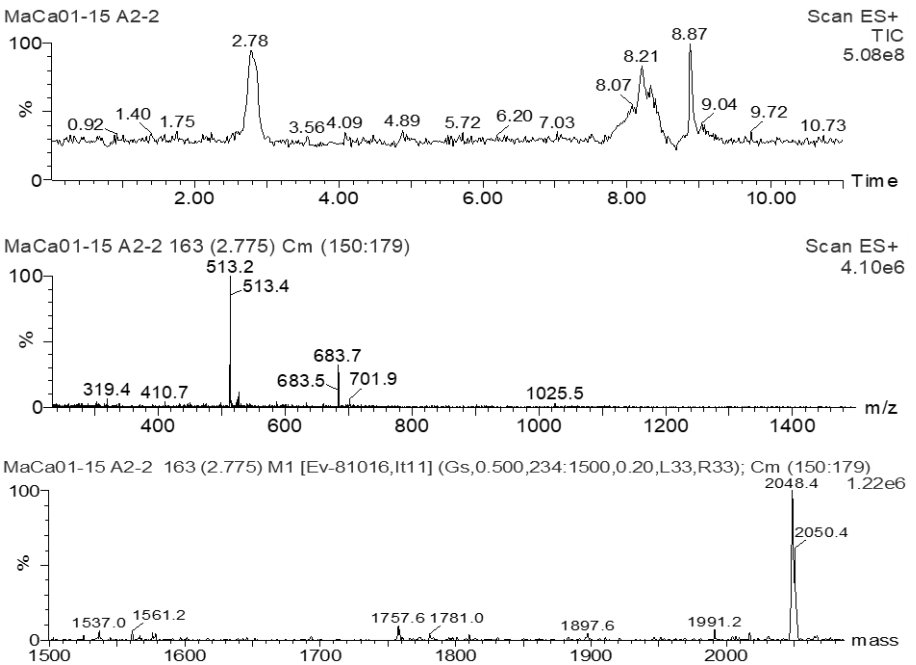



***Figure 3.A45.*** *UPLC-MS of MaCa01-15 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

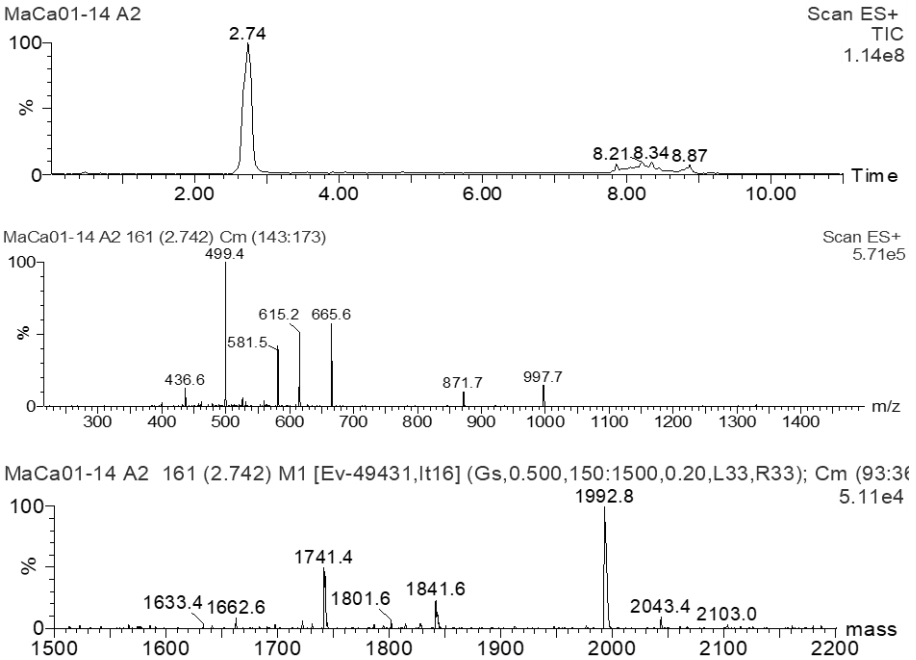



***Figure 3.A46.*** *UPLC-MS of MaCa01-14 A2: chromatogram, mass spectrum and deconvoluted spectrum.*

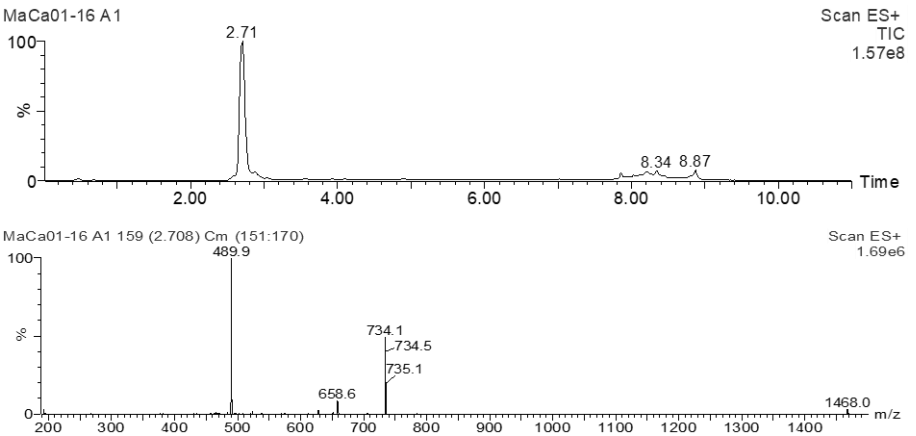



***Figure 3.A47.*** *UPLC-MS of MaCa01-16 A1: chromatogram and mass spectrum.*

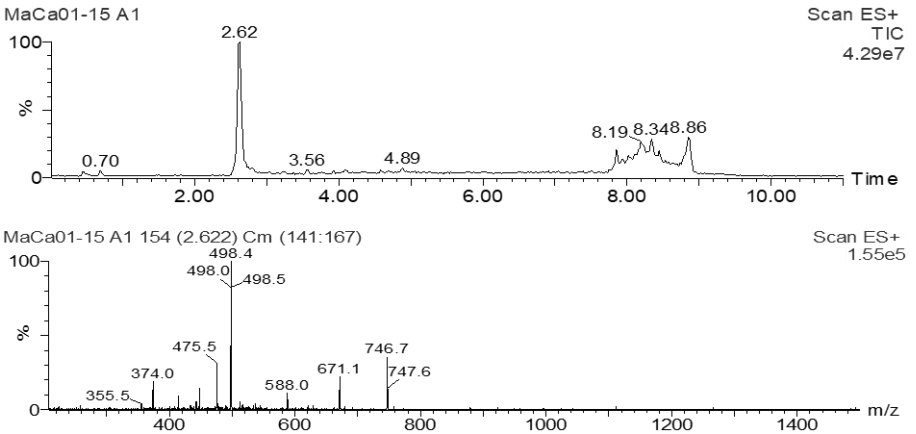



***Figure 3.A48.*** *UPLC-MS of MaCa01-15 A1: chromatogram and mass spectrum.*

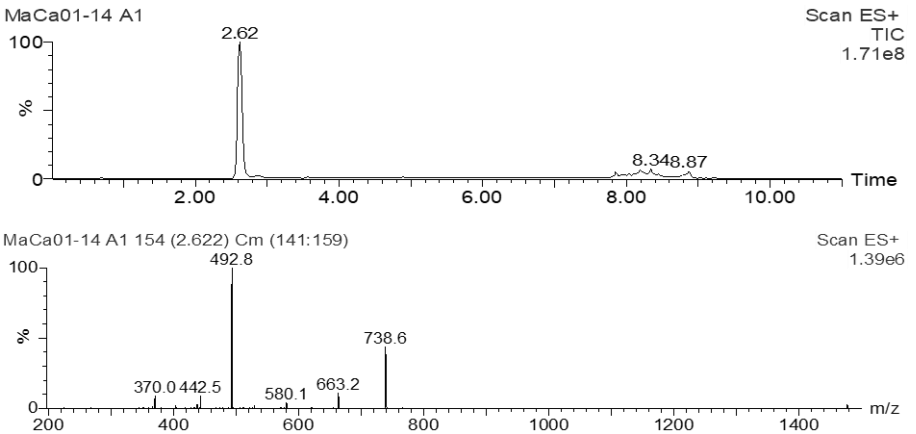
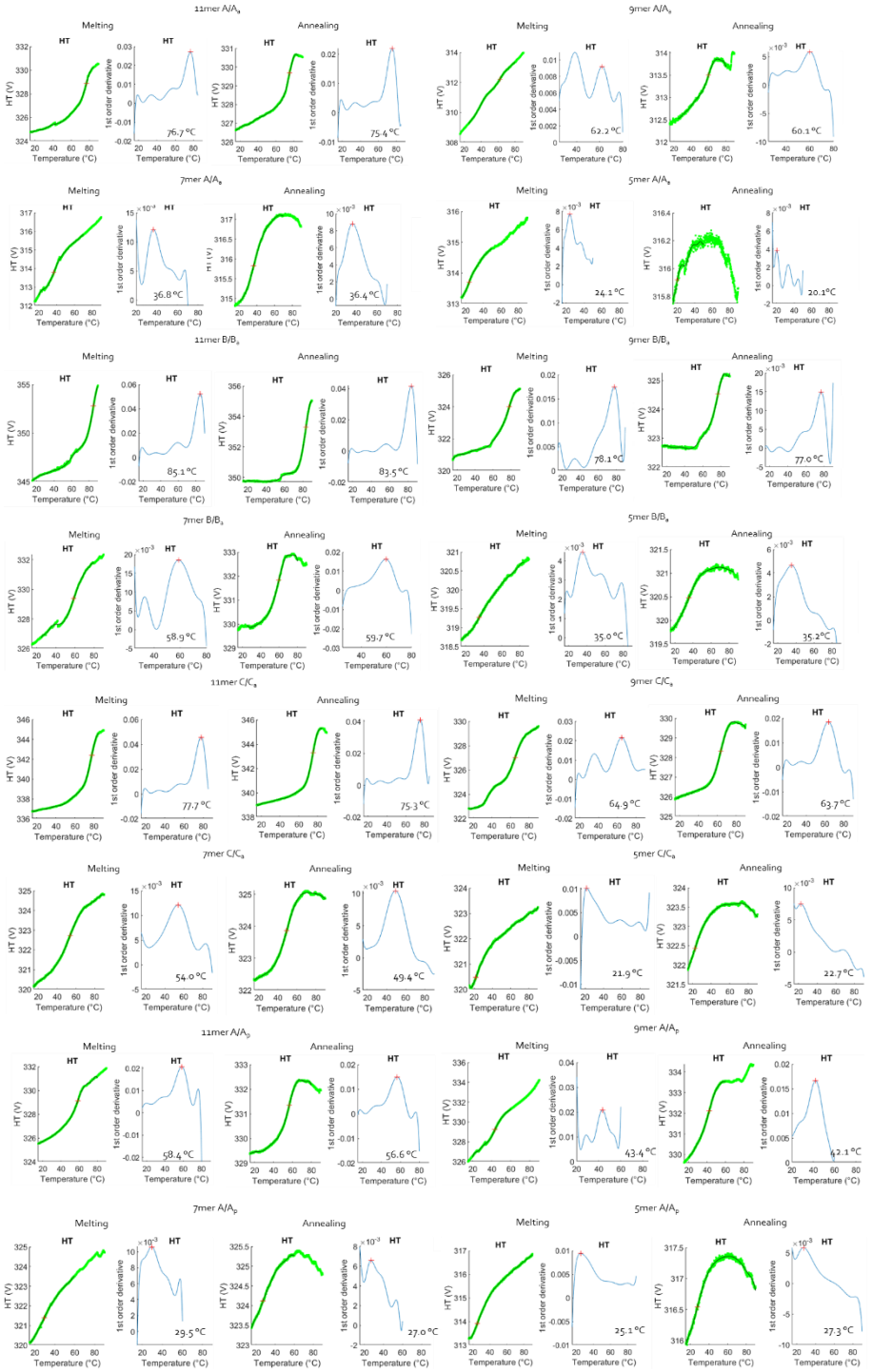



***Figure 3.A49.*** *UPLC-MS of MaCa01-14 A1: chromatogram and mass spectrum.*


11mer A/A$_a$
Melting
Annealing
76.7 °C
75.4 °C
9mer A/A$_a$
62.2 °C
60.1 °C
7mer A/A$_a$
36.8 °C
36.4 °C
5mer A/A$_a$
24.1 °C
20.1°C
11mer B/B$_a$
85.1 °C
83.5 °C
9mer B/B$_a$
78.1 °C
77.0 °C
7mer B/B$_a$
58.9 °C
59.7 °C
5mer B/B$_a$
35.0 °C
35.2°C
11mer C/C$_a$
77.7 °C
75.3 °C
9mer C/C$_a$
64.9 °C
63.7 °C
7mer C/C$_a$
54.0 °C
49.4 °C
5mer C/C$_a$
21.9 °C
22.7 °C
11mer A/A$_p$
58.4 °C
56.6 °C
9mer A/A$_p$
43.4 °C
42.1 °C
7mer A/A$_p$
29.5 °C
27.0 °C
5mer A/A$_p$
25.1 °C
27.3 °C
HT
HT (V)
1st order derivative
Temperature (°C)

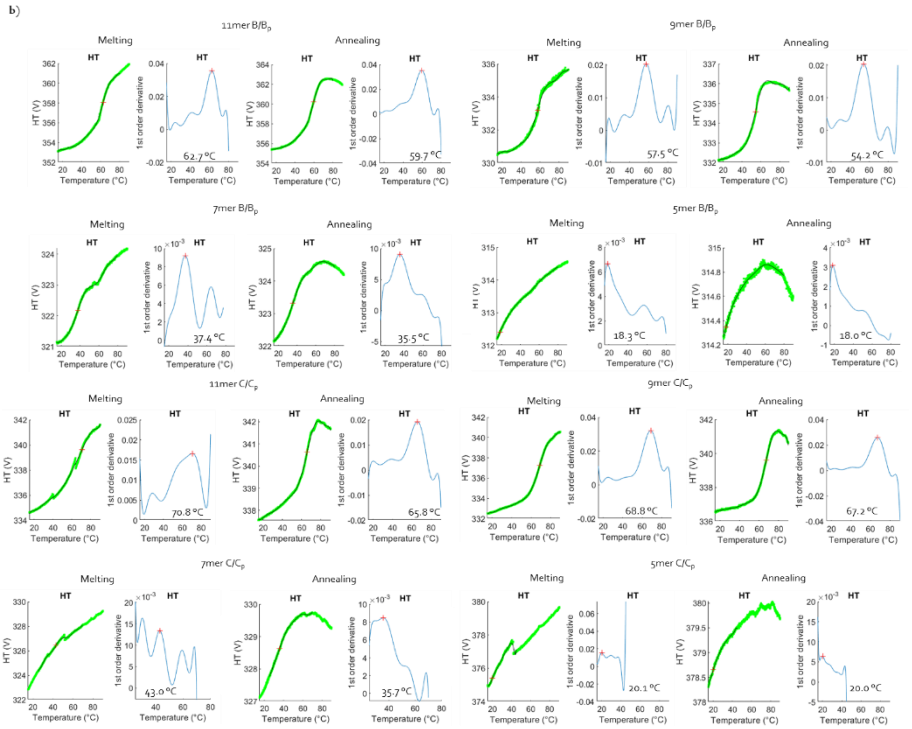



***Figure 3.A50.*** *HT profile of melting and annealing analysis and values calculated from the first order derivative of a 10th order polynomial fitting function of all the full-length PNA:PNA duplexes both parallel and antiparallel orientation.*

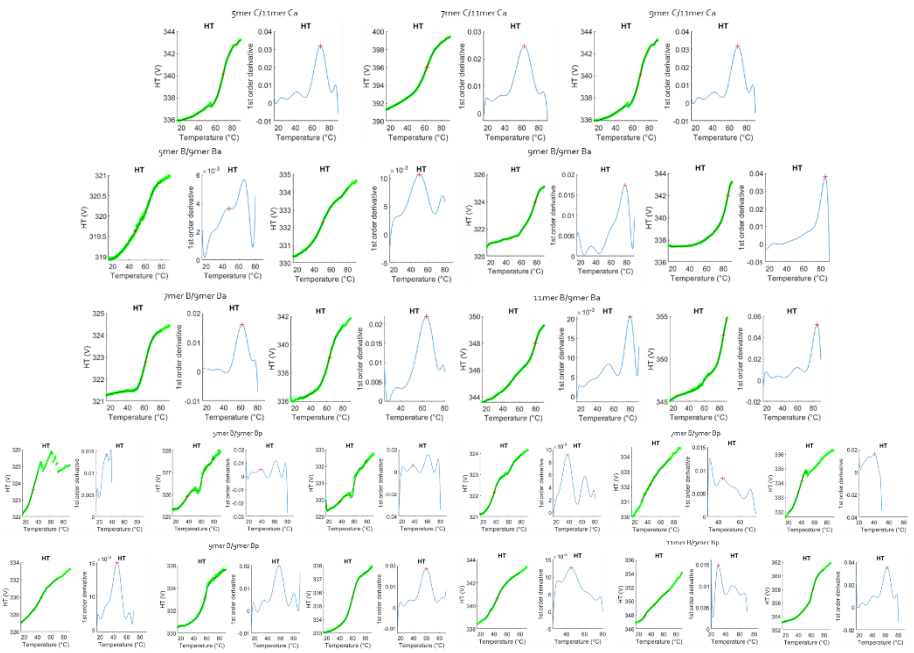



***Figure 3.A51.*** *HT profile of melting and annealing analysis and values calculated from the first order derivative of a 10th order polynomial fitting function of PNA:PNA duplexes reported in Table 3.5-3.7.*

## 3.7 References


1. P.E. Nielsen, M. Egholm, R.H. Berg, O. Buchardt. Sequence-Selective Recognition of DNA by Strand Displacement with a Thymine-Substituted Polyamide. *Science (1979)* **1991**, 254 (5037), 1497–1500.
2. V. V Demidov, N.E. Broude, I. V Lavrentieva-Smolina, H. Kuhn, M.D. Frank-Kamenetskii. An Artificial Primosome: Design, Function, and Applications. *ChemBioChem* **2001**, 2 (2), 133–139.
3. O. Buchardt, M. Egholm, R.H. Berg, P.E. Nielsen. Peptide nucleic acids and their potential applications in biotechnology. *Trends Biotechnol* **1993**, 11 (9), 384–386.
4. V. Menchise, G. De Simone, T. Tedeschi, et al. Insights into peptide nucleic acid (PNA) structural features: The crystal structure of a d-lysine-based chiral PNA–DNA duplex. *Proceedings of the National Academy of Sciences* **2003**, 100 (21), 12021–12026.
5. P. Wittung, P.E. Nielsen, O. Buchardt, M. Egholm, B. Norde´n. DNA-like double helix formed by peptide nucleic acid. *Nature* **1994**, 368 (6471), 561–563.
6. M. Jasiński, J. Miszkiewicz, M. Feig, J. Trylska. Thermal Stability of Peptide Nucleic Acid Complexes. *J Phys Chem B* **2019**, 123 (39), 8168–8177.
7. R. Corradini, T. Tedeschi, S. Sforza, R. Marchelli. Electronic Circular Dichroism of Peptide Nucleic Acids and their Analogues. In *Comprehensive Chiroptical Spectroscopy*; **2012**; pp 587–614.
8. R.B. Merrifield. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J Am Chem Soc* **1963**, 85 (14), 2149–2154.
9. G. Barany, R.B. Merrifield. A new amino protecting group removable by reduction. Chemistry of the dithiasuccinoyl (Dts) function. *J Am Chem Soc* **1977**, 99 (22), 7363–7365.
10. S. Sarkar. Recent advancements in bionanomaterial applications of peptide nucleic acid assemblies. *Biopolymers*. John Wiley and Sons Inc 2023.
11. S. Kumar, A. Pearse, Y. Liu, R.E. Taylor. Modular self-assembly of gamma-modified peptide nucleic acids in organic solvent mixtures. *Nat Commun* **2020**, 11 (1).
12. A. V. Dix, J.L. Conroy, K.M. George Rosenker, D.R. Sibley, D.H. Appella. PNA-based multivalent scaffolds activate the dopamine D2 receptor. *ACS Med Chem Lett* **2015**, 6 (4), 425–429.
13. A. Leonidova, C. Foerster, K. Zarschler, et al. In vivo demonstration of an active tumor pretargeting approach with peptide nucleic acid bioconjugates as complementary system. *Chem Sci* **2015**, 6 (10), 5601–5616.
14. I. Sacui, W.C. Hsieh, A. Manna, B. Sahu, D.H. Ly. Gamma Peptide Nucleic Acids: As Orthogonal Nucleic Acid Recognition Codes for Organizing Molecular Self-Assembly. *J Am Chem Soc* **2015**, 137 (26), 8603–8610.
15. U. Giesen, W. Kleider, C. Berding, et al. A formula for thermal stability (T m ) prediction of PNA/DNA duplexes; **1998**; Vol. 26.
16. M.B. Gawande, P.S. Branco. An efficient and expeditious Fmoc protection of amines and amino acids in aqueous media. *Green Chemistry* **2011**, 13 (12), 3355–3359.

17. J. SantaLucia, D. Hicks. The Thermodynamics of DNA Structural Motifs. *Annu Rev Biophys Biomol Struct* **2004,** 33 (1), 415–440.
18. Graphical Abstract: Eur. J. Org. Chem. 26/2009. *European J Org Chem* **2009,** 2009 (26), 4371–4376.

# *Chapter 4*

## *4.0 PNAs and Nuclease-recruiting PNAs as potential anti-COVID-19 approach*

### 4.1 Abstract

At the end of 2019, a novel viral strain of the coronavirus species, in specific the severe acute respiratory syndrome coronavirus 2 (Sars-CoV-2), has caused a pandemic situation starting from Wuhan (China) and causing acute respiratory illness. The principal goal of the work described in this chapter was the development of a new therapeutic approach based on PNAs. A series of PNAs were designed and synthetized to target different regions of viral genome of a model system based on a plasmid construct containing sequences coding for the spike (S) and nucleocapsid (N) proteins of the Sars-CoV-2 virus. At the same time, the synthesis of RNase L recruiter was carried out, adapting a synthetic protocol in the literature[1], and then this molecule was conjugated to some PNAs. Furthermore, as a proof of concept, these specific modified PNA with the RNase L recruiter were studied if it can recognize the target RNA, recruit the RNase L, and induce the selective target RNA cleavage, thus blocking more efficiently the virus replication. A schematic representation of RNA viral degradation with PNA-RNase L recruiter conjugate proposed is shown in *Figure 4.1*.

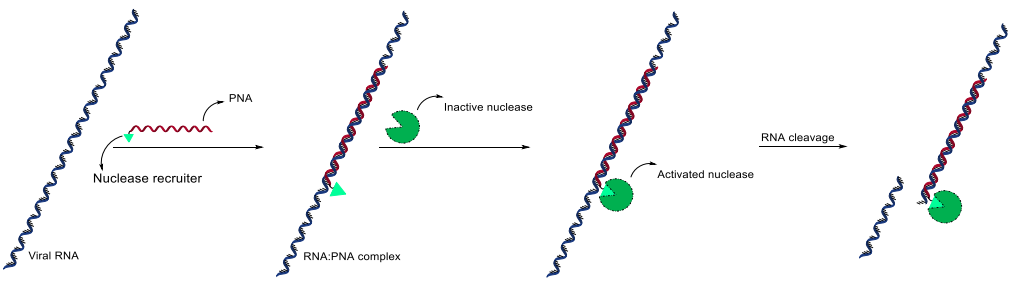



***Figure 4.1.*** *Schematic representation of RNA viral degradation in present of PNA with a RNase L recruiter able to induce the RNA cleavage by recruitment of nuclease.*

This project was financed by MIUR through the FISR-COVID program, and it was performed in collaboration with the laboratory of Prof. Roberto Gambari at University of Ferrara.

## 4.2 Introduction

The human population throughout its history has been forced to confront the emergence of pandemics caused by exposure to new viruses that, in some cases were extremely dangerous. The rise in the number of pathogenic viruses in recent years is mainly driven by factors such as population growth, climate changes, and economic globalization.

Coronaviruses (CoVs) are a family of RNA viruses that were first identified in the 1960s. Coronaviruses (order Nidovirales, family Coronaviridae, genus Coronavirus) represent a diverse group of large, lipid-coated, positive-stranded RNA viruses (the viral genome can be directly read by the cellular protein translation machinery to produce viral proteins) that cause respiratory and enteric diseases in humans and other animals. Each type of coronaviruses was classified into distinct groups and species based on typology, host range, antigenic relationships, and genomic organization. However, they can be classified in four principal genera: Alphacoronavirus, Betacoronavirus, Gammacoronavirus, and Deltacoronavirus.[2] In addition, they can infect both animals and humans, causing respiratory syndromes originally ranging from mild to moderate.[3] The emergence and epidemic spread of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) in 2002, Middle East Respiratory Syndrome Coronavirus (MERS-CoV) in 2012, and most notably, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in recent times have fundamentally altered the perception of the severity of this virus family.[4,5]

Severe acute respiratory syndrome (SARS) emerged at the end of 2002 in Guangdong, China, identifying its etiological agent as a new coronavirus (SARS-CoV) not previously present in humans. Some studies have shown that the spread of this virus was due to interspecies transmission from animals to humans, as there was a lack of serological evidence of prior infection in healthy individuals. However, its rapid spread has led to a significant global impact of the virus.[6] Instead, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is a recent pneumonia which

arose at the end of 2019 in the city of Wuhan (China), and subsequent quickly spread over the world, being highly transmissible. One of the major causes of death in COVID-19 patients is the cytokine storm syndrome, leading to acute respiratory distress syndrome and respiratory failure.[7]

Both genomes consist of a single-stranded positive-sense RNA formed by the typical structure of coronavirus (*Figure 4.2*), consisting in replicase 1a and 1b, spike protein (S), envelope protein (E), membrane protein (M), nucleocapsidic protein (N) and other non-structural proteins (nsps).

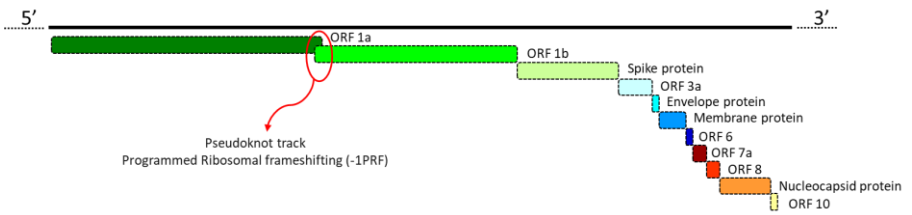



***Figure 4.2.*** *General structure of SARS-CoV.*

The genome of SARS-CoV was completely sequenced (GenBank accession number AY278741) and it is approximately 30 kilobases (kb) in length, resulting the largest among all RNA viruses.[8]

The genome of SARS-CoV-2, like SARS-CoV genome, falling in *betacoronavirus* genera, is a single-stranded positive-sense RNA of30 kb in length, (GenBank accession number: *MN908947.3*).

However, the first two-thirds of both genomes encompass the coding sequences for two extensive replicase polyproteins, denoted by open reading frames (ORF) 1a and ORF1b. An ORF is a sequence of nucleotides in a strand of nucleic acid identified by start codons (usually represented as *AUG* in RNA strands, indicating the translation start point of a protein) and stop codons. Analogous to other coronaviruses, there is a slight overlap between ORF1a and ORF1b. Due to the absence of independent translation initiation sites for ORF1b, the proteins encoded by this region are exclusively translated as a combined protein alongside ORF1a through the

mechanism of programmed -1 ribosomal frameshifting (-1 PRF). Because ORF1a is situated at the 5' end of the genomic RNA (gRNA), it is the first to be decoded.

The -1 PRF process involves the shifting of one base in the 5' direction by a fraction of elongating ribosomes, caused by a cis-acting RNA element within that overlapping region. Ribosomes, after the shift, continue translating the encoded proteins of ORF1b (with a one base shifted reading frame), facilitating the transition from stage 1 to stage 2 of the viral replication cycle.[5] However, the pseudoknot element can fold back on itself to create a complex, compact, and stable tridimensional secondary structure (a 3-stem architecture). An attenuator hairpin (AH), a slippery site (SS), and a pseudoknot frameshifting element (FSE) are therefore the critical motifs that enable the frameshifting process, all in the 5' → 3' order (*Figure 4.3*).

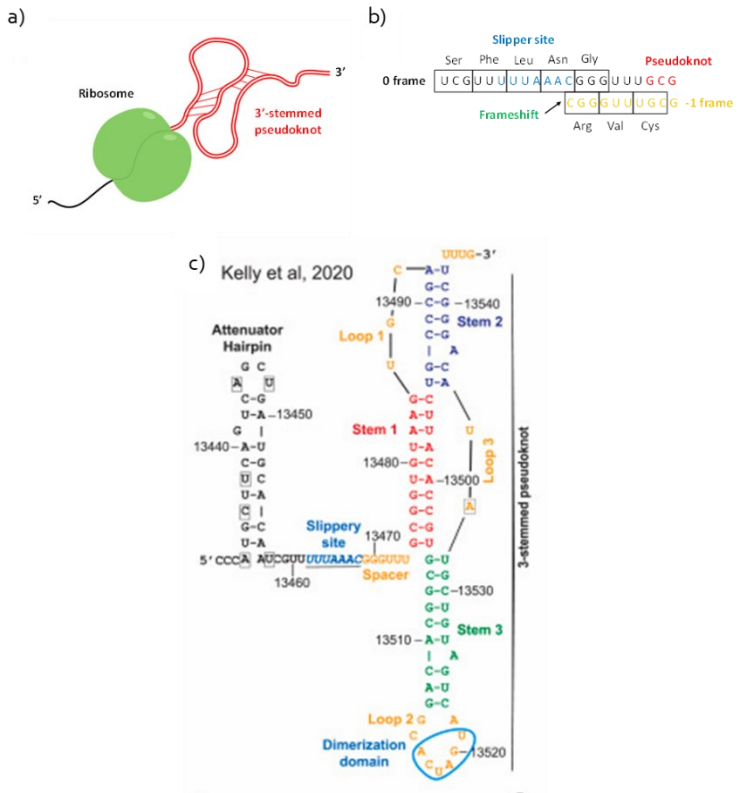



***Figure 4.3.*** *a) Schematic drawing of the pseudoknot structure inducing the frameshift, b) change in ORF in the slippery site and c) pseudoknot structure. Adapted with permission from Kelly et al. Copyright Elsevier © 2023.*[5]

Overall, the genome of SARS-CoV-2 is found to be quite like that of SARS-CoV, although some differences have been highlighted.[9]

The -1 PRF signal, both in terms of sequence and structure, remains conserved, potentially limiting the ability of SARS-CoV to generate drug-resistant mutants. In fact, this aspect makes it an appealing target for antiviral drug development.[10] Genetically, this -1 PRF signal exhibits a high degree of conservation in the analyses of SARS-CoV-2 sequence variations, resulting in an optimal target zone to delay or stop the virus replication in many variants. The –1 PRF was identified as a potential target for therapeutic intervention, given that small molecules can be used to modify –1 PRF rates, thus disrupting viral replication.[11,12] For example, 2-methylthiazol-4-ylmethyl)-[1,4]diazepane-1-carbonyl]amino benzoic acid ethyl ester (MDTB) is a small molecule capable of inhibiting the -1 PRF of SARS-CoV and SARS-CoV-2 and the viral replication.[13–15] Furthermore, as previously mentioned in the introduction (1.2.4.1 Antisense and antigene agent in gene therapeutic drugs), antisense peptide nucleic acids (PNAs) were engineered to specifically target the pseudoknot structure within the SARS-CoV frameshifting region. In fact, these PNAs were then evaluated for their effectiveness ability to inhibit the highly conserved -1 ribosomal frameshifting (-1 PRF) and SARS-CoV replication, leading to a substantial reduction of viral replication.[10]

Antiviral intervention can also target the –1 PRF attenuator hairpin. In fact, for example, Disney and colleagues found that screening small molecules binding to the SARS-CoV-2 attenuator hairpin element can prevent frameshifting process and reduce viral expression.[16] Within these identified small molecules, the compound named C5 was observed to have the ability to selectively bind and stabilize the attenuator hairpin structure of the FSE, thereby reducing its activity. Then, a specific ribonuclease targeting chimera (RIBOTAC) with a RNase L ligand[1] was functionalized with compound C5, resulting in C5-RIBOTAC, able not only to inhibit the frameshifting process, but also to degrade the SARS-CoV-2 viral RNA. The overall molecule proved to be catalytic, capable of effectively functioning at sub-

stoichiometric concentrations.[16] C5 compound and the C5-RIBOTAC complex are reported in *Figure 4.4*.

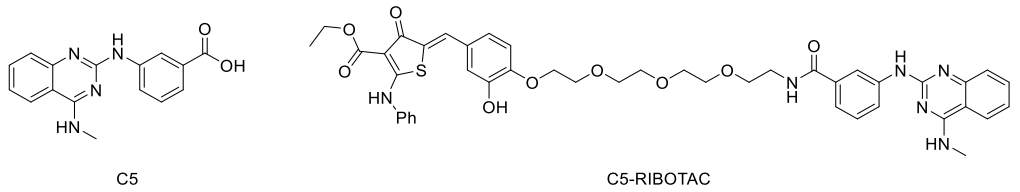



***Figure 4.4.*** *Structure of C5 compound and C5-RIBOTAC complex.*

The development of vaccines against the SARS-CoV-2 virus has had a significant impact on infection prevention and virus spread containment. However, in addition to vaccination, various therapeutic techniques such as immunotherapy and antiviral drugs have been, and continue to be, among the strategies used to prevent infection and successfully limit the virus' spread.[17] In fact, considering the emergence of new variants of SARS-CoV-2, it is essential that research can develop various differentiated therapeutics for the management of COVID-19.[18]

Hence, research continues to focus on the development of compounds targeting RNA, such as antisense, short interfering RNA (siRNA), and other approaches. One of the most promising approaches could be the use of peptide nucleic acids (PNAs) targeting viral RNA. This is due to the exceptional chemical and enzymatic stability of these compounds, their robust RNA affinity, and sequence-specific selectivity. PNAs have demonstrated their effectiveness as regulators of gene expression, especially in inhibiting aberrant splicing, acting as anti-miR agents, serving as innovative antimicrobial agents, and functioning as anti-gene drugs.[19,20]

Combination of PNA targeting viral RNA linked to small ligands able to recruit functional proteins, in particular RNAses, thus creating a ternary assembly in vivo, can produce a very potent and general strategy for limiting viral infections.

## 4.3 Result and discussion

### 4.3.1 Synthesis of RNase L recruiter and its variants

The synthesis of the RNase L recruiter described by Disney and collaborators, was carried out using the procedure reported in the literature [1], with the addition of a last synthetic step to introduce at the end of the spacer the functional group necessary for the coupling with to PNAs (*Figure 4.5*).

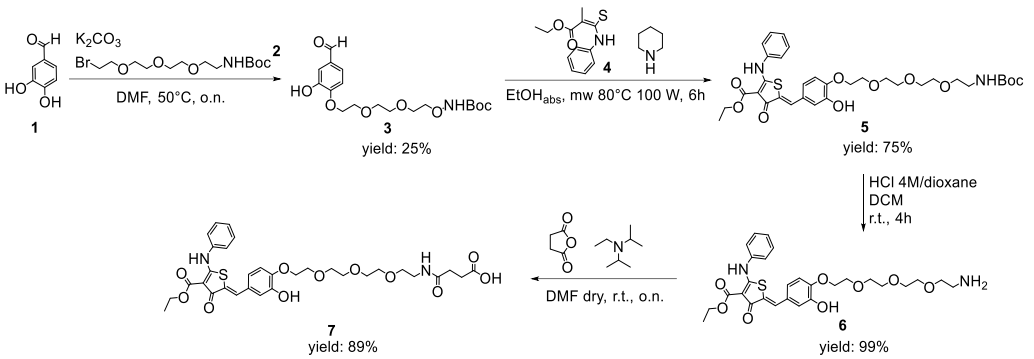



***Figure 4.5.*** *RNase L recruiter synthesis.*

To obtain the RNase L recruiter with a carboxylic acid terminal group, a protocol based on four reaction steps was used. Starting from the commercially available as 3,4-dihydroxybenzaldehyde and t-Boc-N-Amido-PEG4-bromide under $S_N2$ nucleophilic substitution conditions, in presence of potassium carbonate as a base, compound **3** was obtained. The $^1$H-NMR spectrum is reported in *Figure 4.A1* in Appendix.

Then, this modified aldehyde was reacted with a thiophene derivative (ethyl 4,5-dihydro-4-oxo-2-(phenylamino)-3-thiophenecarboxylate) to yield compound **4,** via a piperidine-catalyzed cross aldol condensation reaction. The $^1$H-NMR spectrum is reported in *Figure 4.A2* in Appendix. The Boc-deprotection of amino moiety was then performed as follows: compound **5** was dissolved in dichloromethane and a solution of HCl 4M in dioxane was added to obtain compound **6**. The $^1$H-NMR spectrum is reported in *Figure 4.A3* in Appendix. To obtain the carboxy-terminated compound necessary to perform the coupling to the amino moiety of a PNA oligomer, a

coupling between compound **6** and succinic anhydride with DIPEA was performed. In fact, under dry condition, to obtain compound **7**.

Subsequently, the synthesis of an alternative RNase L recruiter was designed and synthetized to have a spacer already presenting the carboxylic-terminal for the coupling with the PNAs (*Figure 4.6*).

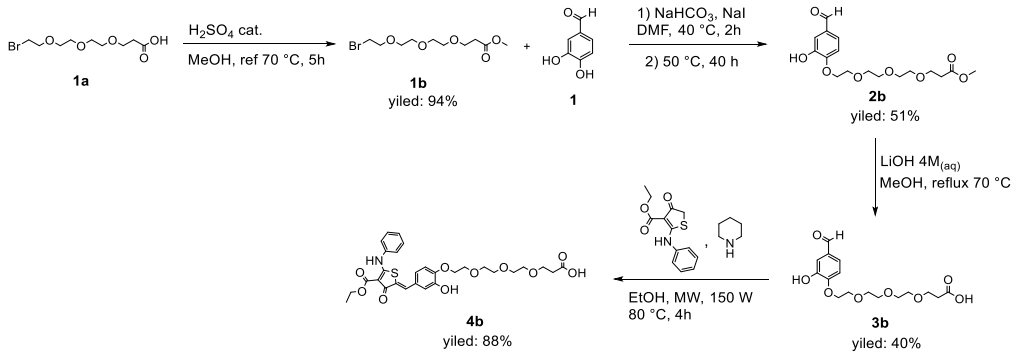



***Figure 4.6.*** *Alternative RNase L recruiter synthesis*

The synthesis of alternative RNase L recruiter was performed in four reaction steps of synthesis. Starting from $BrPEG_4COOH$ (reagent **1a**), commercially available, an esterification reaction with sulfuric acid as catalyst was done to have the carboxylic moiety protected (compound **1b**). The $^1$H-NMR spectrum of **1b** is reported in *Figure 4.A6* in Appendix. Then, **1b** was reacted with 3,4-dihydroxybenzaldehyde (reagent **1**), to obtain the compound **2b** by $S_N2$ nucleophilic substitution in the presence of sodium bicarbonate as base.

## 4.3.2 Design and synthesis of PNA and PNA-conjugate

The PNAs were designed to target different regions of mRNA encoding for spike (S) and nucleocapsid (N) proteins, both viral genome and plasmid construct, and a critical region between ORF1a and ORF1b of the complete viral genome including a pseudoknot able to induce a programmed ribosomal frameshifting (-1 PRF) in the open reading frame during the translation. In addition, two PNAs at different lengths

that do not target any viral genome and plasmid model system were designed as controls. The, all PNA sequences were also functionalized with the octa-arginine peptide moiety at their N-term to obtain an efficient cellular uptake in the target cells.

Before the synthesis, all the sequences of PNA targeting these RNA tract were analyzed with "PNA Tool" (http://pnabio.com/support/PNA_Tool.htm) to have an overview of the general properties of these PNAs and the results are showed in Appendix.

Furthermore, to understand all the possible interferents in the human genome, all the selected sequences were also analyzed for similarity in the NCBI database using BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch), but the analysis did not lead to the detection of problematic interferents.

Once the RNase L recruiter synthesis was completed, the coupling between this small molecule and the PNAs was performed on the solid support (*Figure 4.7*) to obtain a series of PNA conjugates that could be able to bind and degrade the viral genome.

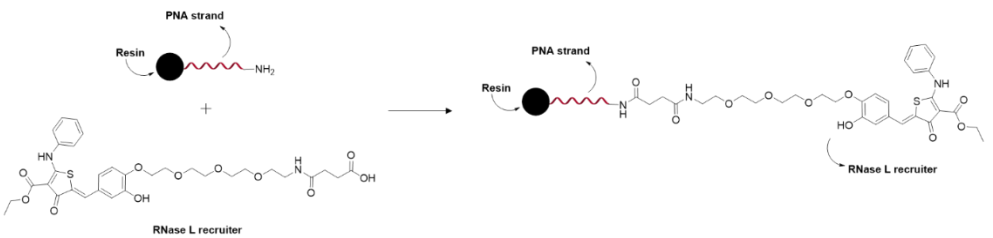



***Figure 4.7.*** *Coupling between the PNA sequence and the RNase L recruiter on solid support.*

Furthermore, to achieve the coupling between the PNAs and the chemical moiety able to recruit RNase L, some different tests with different coupling reagents, equivalents, temperatures, coupling, and activation time were carried out (*Table 4.1*). However, the best condition to perform the coupling appear to be *Entry 7* after the UPLC-MS analysis.

***Table 4.1.*** *Tests to optimize the coupling between the PNAs and the ribonuclease L recruiter and obtain the PNA conjugates.*

| Entry | eq RNase L recruiter | eq HBTU | eq DIPEA | Temperature | Time | Activation |
|---|---|---|---|---|---|---|
| 1 | 5 | 4.9 | 10 | r.t. | o.n. | 5′ |
| 2 | 5 | 4.9 | 10 | 40°C | o.n. | 5′ |
| 3 | 5 | 4.9 | 10 | 50°C | 6h | 5′ |
| 4 | 5 | 4.9 | 10 | 60°C | 5h | 5′ |
| Entry | eq RNase L recruiter | eq DIC | eq DhBtOH | Temperature | Time | Activation |
| 5 | 5 | 5 | 5 | r.t. | 5h | 2′ |
| 6 | 5 | 5 | 5 | r.t. | 2x5h | 2′ |
| 7 | 10 | 10 | 10 | r.t. | o.n. | 2′ |

A schematic representation of PNAs and PNAs-conjugates designed is shown in *Figure 4.8* and all the sequences synthetized targeting the S and N protein both in viral genome and plasmid model construct, the pseudoknot region and the two control was reported in *Table 4.2*.

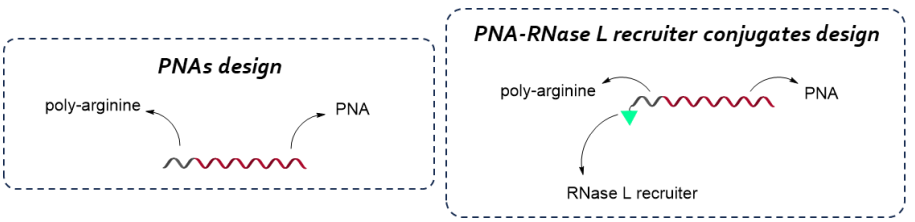



***Figure 4.8.*** *A schematic representation of PNAs and PNAs-conjugates design.*

***Table 4.2.*** *All PNAs and PNA-conjugates synthetized.*

| Entry | Name | Sequence | Yield (%) |
|---|---|---|---|
| 8 | S-P | r8-CAC CAG GAA CAC AAA CAT-g-$NH_2$ | 6 |
| 9 | S-St | r8-AAC AAG AAA AAC AAA CAT-g-$NH_2$ | 13 |
| 10 | S-3′ | r8-CAC TCC ATA ACA CTT-g-$NH_2$ | 8 |
| 11 | S-5′ | r8-AAA ACA AAC ATT GTT CGT-g-$NH_2$ | 12 |
| 12 | N-P | r8-TGG TCC ATT GTC ACT CAT-g-$NH_2$ | 5 |
| 13 | N-St | r8-GGG TCC ATT ATC AGA CAT-g-$NH_2$ | 5 |
| 14 | N-3′ | r8-TC TCC ATT CTG GTT-g-$NH_2$ | 6 |

| | | | |
|---|---|---|---|
| 15 | N-5′ | r8-ACA TTT TAG TTT GTT CGT-g-$NH_2$ | 11 |
| 16 | PK | r8-A AAG CCC TGT ATA CGA CAT-g-$NH_2$ | 7 |
| 17 | C-18mer | r8-TAT CCA GTC AAG ATC TAA-g-$NH_2$ | 9 |
| 18 | C-15mer | r8-ACA CTC TAC ATC ACT-g-$NH_2$ | 28 |
| 19 | S-P+L | L-r8-CAC CAG GAA CAC AAA CAT-g-$NH_2$ | 6 |
| 20 | S-St+L | L-r8-AAC AAG AAA AAC AAA CAT-g-$NH_2$ | 9 |
| 21 | S-3′+L | L-r8-CAC TCC ATA ACA CTT-g-$NH_2$ | 9 |
| 22 | S-5′+L | L-r8-AAA ACA AAC ATT GTT CGT-g-$NH_2$ | / |
| 23 | N-P+L | L-r8-TGG TCC ATT GTC ACT CAT-g-$NH_2$ | 3 |
| 24 | N-St+L | L-r8-GGG TCC ATT ATC AGA CAT-g-$NH_2$ | 2 |
| 25 | N-3′+L | L-r8-TC TCC ATT CTG GTT-g-$NH_2$ | 5 |
| 26 | N-5′+L | L-r8-ACA TTT TAG TTT GTT CGT-g-$NH_2$ | / |
| 27 | PK+L | L-r8-A AAG CCC TGT ATA CGA CAT-g-$NH_2$ | 3 |
| 28 | C-18mer+L | L-r8-TAT CCA GTC AAG ATC TAA-g-$NH_2$ | 14 |
| 29 | C-15mer+L | L-r8-ACA CTC TAC ATC ACT-g-$NH_2$ | 4 |
| Capital letters: PNA monomers, Lower case: amino acid | | | |

To understand the assigned codes: *S*, *N*, and *PK* indicate the spike protein, nucleocapsid protein, and pseudoknot regions, respectively, while *C* for control. *St*, *3′*, *5′* are relative to the position of target sequence with respect to protein translation start codon: *St* is relative to the tract containing the initial AUG, while *3′* and 5′ refer to sequences located in the 3′/5′ direction, respect to AUG. Instead, *P* represent the plasmid construct. *18mer* and *15mer* controls are denoted as *C-15mer* and *C-18mer*.

In addition, a second model was designed to test the role of the RNase L recruiter position in the PNA sequence, focusing on the pseudoknot PNA sequence (*Figure 4.9*). Fmoc-Lys(Mtt)-OH was used as first residue to allow the possibility to link the RNAse ligand to the C-term part of the PNA. The ligand was then linked to N or C terminus directly or after one or two (2-aminoethoxy)ethoxyacetyl spacers (AEEA,

or O). All the PNAs sequences synthetized with these two different designs were listed in *Table 4.3*.

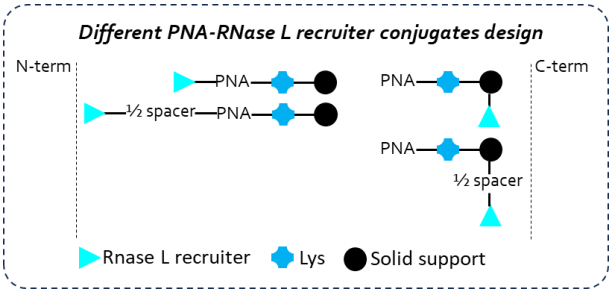



***Figure 4.9.*** *A schematic representation of different PNA-RNase L recruiter conjugates design.*

***Table 4.3.*** *PNAs series regarding the pseudoknot region with two different designs, from Entry 37 to 39 the RNase L recruiter (L) was positioned at the $N_{term}$ moiety instead from Entry 40 to 42 at the $C_{term}$. In addition, different number of AEEA spacer (O) was added.*

| Entry | Sigla | Sequence | Yield (%) |
|---|---|---|---|
| 30 | BeCa01-12 | L-A AAG CCC TGT ATA CGA CAT-k(Ac)-$NH_2$ | 1 |
| 31 | BeCa01-13 | L-O-A AAG CCC TGT ATA CGA CAT-k(Ac)-$NH_2$ | 1 |
| 32 | BeCa01-17 | L-O-O-A AAG CCC TGT ATA CGA CAT-k(Ac)-$NH_2$ | 5 |
| 33 | BeCa01-18 | Ac-A AAG CCC TGT ATA CGA CAT-k(L→)-$NH_2$ | 6 |
| 34 | BeCa01-19 | Ac-A AAG CCC TGT ATA CGA CAT-k(L-O→)-$NH_2$ | 4 |
| 35 | BeCa01-20 | Ac-A AAG CCC TGT ATA CGA CAT-k(L-O-O→)-$NH_2$ | 12 |
| Capital letters: PNA monomers, Lower case: amino acid, L: RNase L recruiter, O: spacer | | | |

For the second design was selected the pseudoknot PNA sequence given that, after a bio-informatic analysis, it appears to be conserved in the analyzed virus variants. In fact, a bio-informatic analysis was conducted to evaluate the conservation of synthesized sequences directed towards the viral genome in the predominant European variants from January to May and then from August to September 2023. The Global Initiative on Sharing All Influenza Data (GISAID, https://gisaid.org/) is the international online platform used for this bioinformatics analysis. This platform enables the global sharing of data related to the genome sequence of influenza and respiratory viruses, with a particular focus on the SARS-CoV-2 virus, identifying variants through any mutations present in their genome. However, GISAID employs

a naming/numbering system to identify variants, where the name includes the place of origin, sampling date, and sequence number. Several types of nomenclature of the variants were reported, like *World Health Organization* (WHO, https://covid19.who.int/) that identifies the variants using Greek alphabet letters, *Phylogenetic Assignment of Named Global Outbreak Lineages* (PANGO Lineage, https://cov-lineages.org/) where PANGO lineages determined through the PANGOLIN software are typically identified by a combination of letters and numbers, and instead Nextstrain (open-source bioinformatics platform, https://nextstrain.org/) uses a Year-Letter nomenclature.

Focusing on the epidemiological spread of the SARS-CoV-2 virus in the previously mentioned time periods, a picture extracted by GISAID platform showing how Omicron is the predominant European variant in the target temporal period of SARS-CoV-2 is depicted in the *Figure 4.10*.

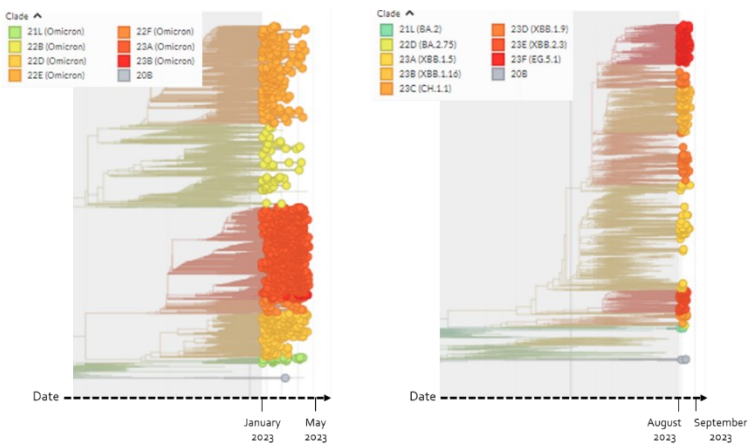



***Figure 4.10.*** *Focus of the epidemiological spread of the SARS-CoV-2 virus from January to May on the left spectrum and from August to September on the right picture.*

The main VOC (Variants of Concern) and VOI (Variants of Interest) in the same period are listed in the *Table 4.4*, that highlights again Omicron as the most widely spread European variant. In addition, from the first to the second screening, it can be affirmed that variants 23A, 22D, and 23F (market in blue) persist from January to

September. However, in the second time frame, the emergence of a new variant, denominated 21F (highlighted in green) can be observed.

***Table 4.4.*** *Principal VOC and VOI variants in the two period examinated.*

| WHO | NEXSTRAIN | PANGO Lineage | Distribution (%) |
|---|---|---|---|
| From January to May | | | |
| Omicron | 23A | XBB.1.5 | 81,2% |
| Omicron | 22D | BA.2.75 | 10,8% |
| Omicron | 22E | BQ.1 | 3,6% |
| Omicron | 21L | BA.2 | 1,1% |
| Omicron | 22F | XBB | 1.0% |
| Omicron | 22B | BA.5 | 0,8% |
| From August to September | | | |
| Omicron | 21F | EG.5 | 66,1% |
| Omicron | 23A | XBB.1.5 | 29,4 % |
| Omicron | 22F | XBB | 4,6 % |
| Omicron | 22D | BA.2.75 | 2,3 % |

The genome of these predominant variants, exported from NCBI (National Center for Biotechnology Information), was subsequently compared, using GENEDOC, with the original SARS-CoV-2 genome (MN908947.3) used for the initial design of the PNA sequences. This comparison was conducted with a focus on the target regions of the PNA sequences, specifically S-5', S-St, S-3' (S gene), N-5', N-St, N-3' (N gene), and PK. The results are showed in *Figure 4.11* for S and N genes, while in *Figure 4.12* for the pseudoknot sequence. From *Figure 4.11 and 12*, it is possible to observe that S-5', S-St and PK sequences appear to be completely conserved, in all the variants. On the other hand, the N-St sequence, which appeared to be conserved in the initial phase of bioinformatic analysis, shows a change of one nucleobase in the time range from August to September. However, compared to the reference genome, differences in terms of one nucleobase can be observed for the N-5' sequence in both analyzed periods, while the sequence N-3' is not conserved in the various variants of SARS-CoV-2 examined. In summary, the S-5', S-St, and PK sequences appear to be

conserved in the analyzed virus variants, leading to the conclusion that these regions could still be potential targets for the development of antiviral drugs.

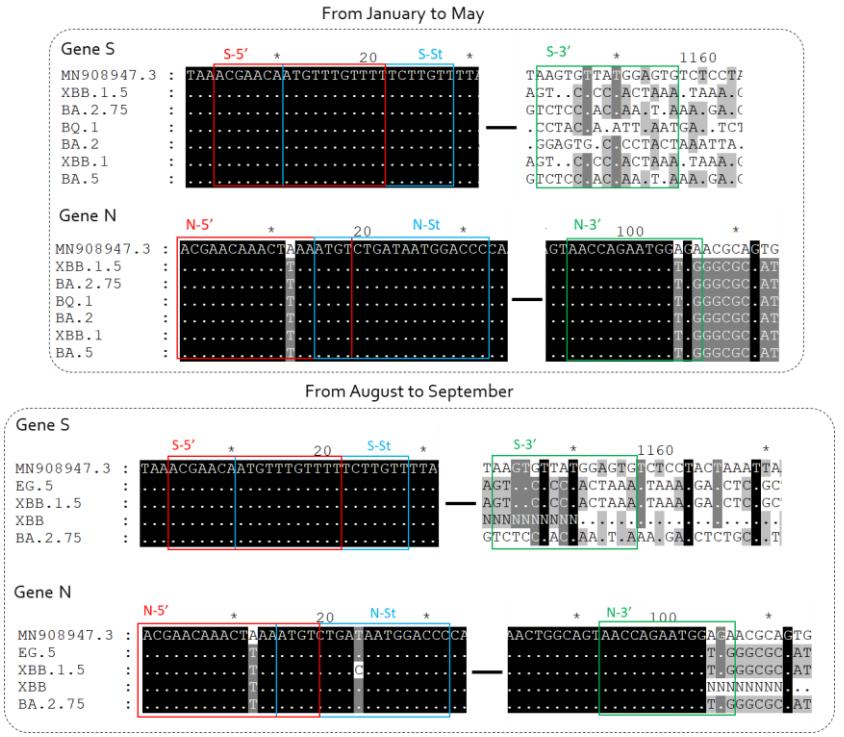



***Figure 4.11.*** *Comparison of the genome of the major variants from January to May (on the top) and from August to September (on the bottom) with respect to the SARS-CoV-2 genome initially used for the design of PNA sequences (reference genome), focusing on specific region of S- and N-gene.*

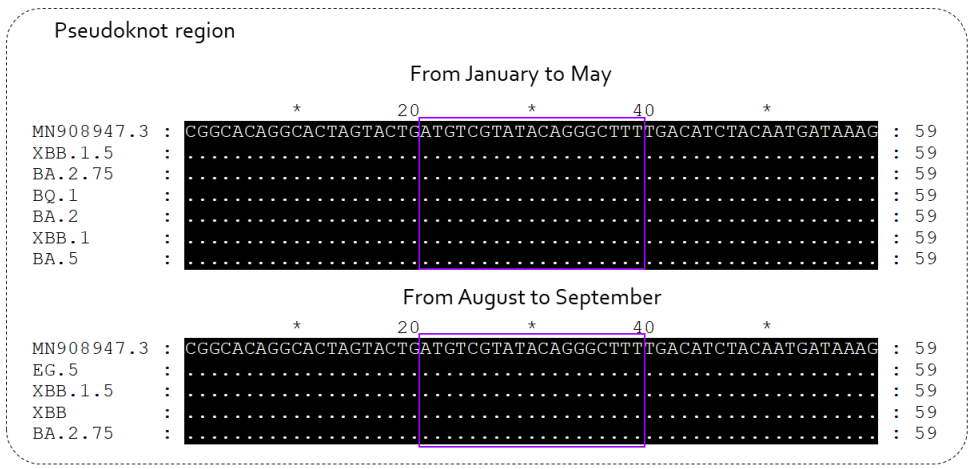



***Figure 4.12.*** *Comparison of the genome of the major variants from January to May (on the top) and from August to September (on the bottom) with respect to the SARS-CoV-2 genome initially used for the design of PNA sequences (reference genome), focusing on pseudoknot region.*

### 4.3.3 Biological experiments

The PNA of *Table 4.2* were provided to the group led by Prof. Roberto Gambari at the University of Ferrara, to test the antisense activity towards SARS-CoV-2 sequences. Some of the results obtained are reported hereinafter to show the potentiality of the PNA constructs.

A PCR inhibition experiment on a Plasmid model pUC57-2019-nCoV-PC: ORF1ab (cat. MC-0101079 from GeneScript, Piscataway, New Jersey, USA), containing the ORF1ab regions of the virus, was performed to assess the specific molecular interactions between the SARS-CoV-2 pseudoknot region and the corresponding PNA sequence. To carry out the PCR amplification two couple of PCR primers were used, one to amplify the SARS-CoV-2 pseudoknot region (F-RFS and R-RFS) and the other used as control primers (Fc and Rc), while regarding the PNA sequences, the pseudoknot PNA (PNA-PK) and the Control-18mer (uPNA) were employed (*Figure 4.13*).

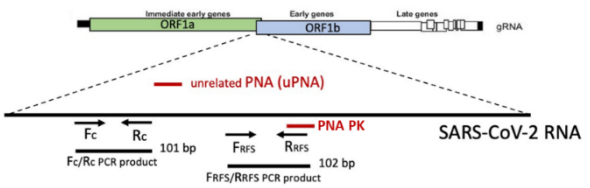



***Figure 4.13.*** *A schematic representation of SARS-Cov-2 RNA with the two PNA sequences (PNA-PK and PNA-Control 18mer) and couple of primers (F-RFS/R-RFS and Fc/Rc) used for the PCR experiments.*

Inhibition of PCR was observed in presence PNA-PK (*Figure 4.14A*) at concentrations of 10 nM or higher, observed by the rightward shift as threshold cycle ($C_T$) increases. $C_T$ is used to monitor the progress of amplification during a real-time PCR reaction, and it is the cycle at which the amount of amplified genetic material reaches a detectable level. Instead, the PCR amplification conducted in presence of control PNA at high concentrations (50-100 nM of uPNA) showed no inhibition (*Figure 4.14B*), while the results shown in *Figure 4.14C* highlight the inactivity of PNA-PK when Fc and Rc control primers were used for the amplification, except at higher

concentrations of PNA-PK. The dependence of Ct as a function of PNA concentrations used in the various tests is reported in *Figure 4.14D*.

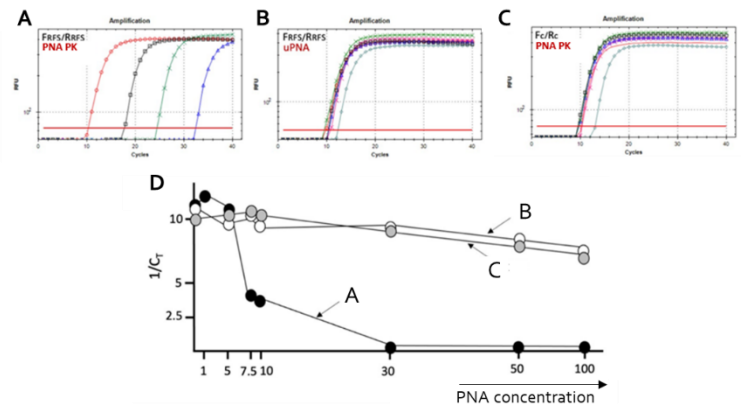



***Figure 4.14.*** *Data of PCR experiments carried out with the target SARS-CoV-2 region containing the Ribosomal Frame Shift Element (RFSE) and PNA-PK or Control-18 PNA (used as unrelated PNA) and two couple of primer: F-RFS/R-RFS and Fc/Rc, used to amplify the SARS-CoV-2 pseudoknot region and as a control, respectively. A) Specific molecular interaction between the PNA-PK and the target SARS-CoV-2 containing the pseudoknot region in presence of F-RFS/R-RFS as primers. B) Inactivity of Control-18 PNA. C) Inactivity of PNA-PK (less than the highest concentration) with control primers. D) Quantitative determination analysis (addicted to PNA concentration).*

SARS-CoV-2 manipulation was performed in the BSL-3 laboratory of the University of Ferrara, following the biosafety requirements. SARS-CoV-2 was isolated from a naso-pharyngeal swab retrieved from a patient with COVID-19 (Caucasian man of Italian origin, genome sequences available at GenBank, SARS-CoV-2-UNIBS-AP66: ERR4145453). Extraction of RNA, followed by RT-PCR showed inhibition of N-protein production, as well as of NFkB and NFkB-regulated pro-inflammatory genes (interleukins IL-1b, IL-6 and IL-8), upon treatment with the PNA-PK (results not shown, these are described in a paper submitted for publication).

Most importantly, both treatments with PNA-PK and PNA-PK-L inhibited the SARS-CoV-2 release from the infected cells (*Figure 4.15*).The inhibition of SARS-CoV-2 was also quantified once the functionalization of the cells with PNA-PK was performed, and the result are reported in *Figure 4.15A*, together data of with Calu-3 infected cells in absence of PNA treatment. The inhibition efficiency of the PNA-PK (*Figure 4.15B*)

was found to be around 75%, and that of PNA-PK with RNase L recruiter was slightly higher (83-85%).

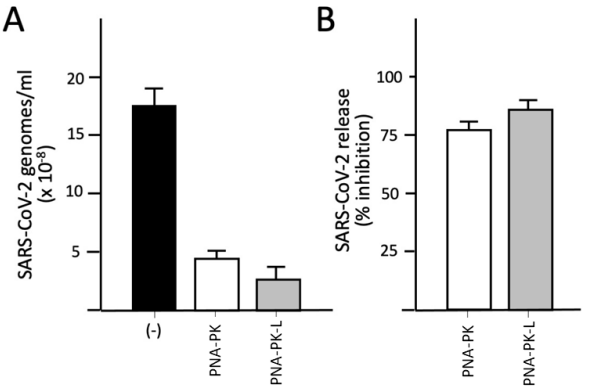



***Figure 4.15.*** *A) Release of SARS-CoV-2 in presence of PNA-PK and PNA-PK-L and in absence of PNA as a control. B) % of inhibition of SARS-Cov-2 release.*

### 4.3.4 Preliminary studies on selective RNA cleavage

Preliminary tests were performed to study if the RNase L recruiter modified PNAs can recognize the target RNA, recruit RNase L, and induce the selective RNA cleavage.

To perform these experiments the second model of PNA sequences (*Table 4.3, from Entry 30 to 35*), focusing on the pseudoknot region, labelled RNA sequences of 59 bp (*RNA-PK59*, 5'[FAM]-CGGCACAGGCACUAGUACUG**AUGUCGUAUACAGGGCUUU** UGACAUCUACAAUGAUAAAG-3') containing the pseudoknot sequence (highlighted in bold), and RNAse L were employed.

A solution of RNA-PK59 (1 μM) with a RNase L recruiter modified PNA (1 μM) was annealed by heating at 70 °C for 5 min and slowly cooling to room temperature in RNase L Buffer (25 mM Tris-HCl, 100 mM KCl, 10 mM $MgCl_2$, pH 7,4). After cooling, the solution was divided in two batch: one already ready for the next incubation step while in the second the RNase L and the RNase L supplement (10 mM MgCl2, fresh 7 mM β-mercaptoethanol, and 50 μM of ATP) were added (considering equal ratio of each component: PNA, RNA and RNseL). All the samples were transferred to a

thermocycler and incubated for two hours or overnight at 28 °C. The reaction is quenched by heating at 95 °C for 5 minutes. The samples were analyzed with denaturing 8 M urea 10% PAGE, using TAMg buffer (40 mM Tris-Base, 12.5 mM $Mg(AcO)_2$, 20 mM EDTA, pH: 8.0). The results, reported in *Figure 4.16*, showed that the RNA alone (*lane 1*) in both the experiments (two hours or overnight incubation) was partially degraded, probably to ageing of the stock solution sample. However, after 2 hours of incubation (*Figure 4.16a*) no substantial differences between the samples in the absence and presence of RNase L was observed. Instead, after overnight incubation (*Figure 4.16b*), a more intense band representative of an RNA sequence less than 50 base pairs in length (with a higher electrophoretic mobility than the *RNA-PK59* alone) was observed for the samples containing the RNase L. A band with lower electrophoretic mobility than the *RNA-PK59* was observed for the samples containing the PNAs reported in *Entry 32* and *33* for the first experiment (from *lane 7* to *10*), and only for the PNA *Entry 32* for the second experiment. This band is related to the formation of the RNA:PNA complex. However, since cleavage of RNA was also observed in the absence of RNase L (*Fig. 16, lane 11*), and with RNase L without the PNA (*Fig. 16, lane 2*) additional experiments with different conditions (e.g. non-degraded *RNA-PK59* probe) and different RNA tracts will be needed.

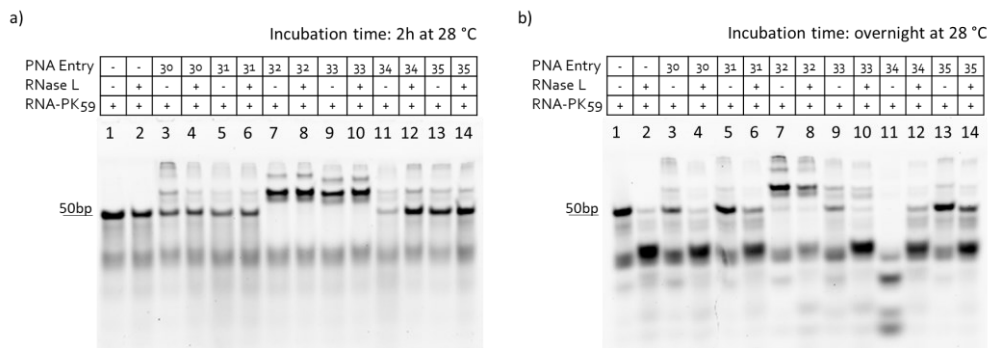



***Figure 4.16.*** *Denaturing PAGE analysis of preliminary experiments of RNA cleavage given by PNA-RNase L recruiter conjugate: a) 2h of incubation, b) overnight incubation.*

## 4.4 Conclusions and future directions

A series of PNA, potentially useful for SARS-CoV-2 treatment, was synthesized. The first biological data obtained are encouraging in terms of specificity and efficacy.

The entire library of PNA synthesized is now available for performing study of interaction/cleavage of the target RNA in, and, most importantly, in cellular systems modelling the COVID-19 infection. However, preliminary tests (based on polyacrylamide gel electrophoresis analysis) were performed to study if the modified PNA with the RNase L recruiter can recognize the target RNA, recruit the RNase L, and induce the selective target RNA cleavage. The preliminary results seem to point toward the direction of a RNA cleavage, but additional experiments with a non-degraded RNA-PK59 probe are required. Further studies are needed for the complete assessment of the best model for this type of application. So far, the possibility of assembly of RNA, PNA, RNase in a rational way seems to be a promising tool for the development of new, rationally designed treatments for precision medicine.

## 4.5 Experimental section

### 4.5.1 Reagents and instrumentations

The reagents were bought from BLD Pharmatech, Merck, abcr, Carlo Erba, VWR, Fluka, TCI Europe, Link Technologies, PolyOrg and used without additional purification. Only for the solid phase synthesis, the DMF was purged with nitrogen flux to prevent the formation of dimethylamine and was dried over 4Å molecular sieves.

NMR spectra were recorded with a Brucker Avance 400 and δ values are expressed in ppm concerning to $CDCl_3$, DMSO-d6 and $D_2O$. TLC are performed on Supelco 56524-25EA silica gel on aluminum foil with indicator fluorescence at 254 nm.

All the PNA sequences are synthesized by standard solid-phase manual synthesis. Purification was performed by HPLC (Agilent Technologies 1260 Infinity I) using a SepaChrom Vydamas® (C18, 5 µm, 300 Å, 10 x 250 mm) column. Gradient: 100% A for 5min, then from 0% to 50% B for 30 min. PNAs at 4 mL/min flow (A: water + 0.1% trifluoroacetic acid; B: acetonitrile + 0.1% trifluoroacetic acid). UV wavelength: 260 nm.

After purification, the peptide nucleic acid sequences were characterized by UPLC-MS (Waters Acquity Ultra Performance LC) using the following instrumental set-up: Waters Acquity ultra-performance LC E07SQD086W, with Waters SQ detector and ESI-interface equipped with Acquity UPLC BEH 300 (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic condition: eluent A: water + 0.2% formic acid; eluent B: acetonitrile + 0.2% formic acid. Column temperature: 35 °C. Program: initial isocratic at 100% A (0.9 min), then linear gradient to 50% B (in 5.7 min). Final wash with 100% B for 1.2 min. Flow rate: 0.25 mL/min. Some PNAs, however, were characterized with Thermo LTQ Orbitrap XL detector and ESI-interface equipped with a Phenomenex Kinetex EVO (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic conditions: eluent A: water + 0.1% formic acid; eluent B: acetonitrile + 0.1% formic acid. Column temperature: 35 °C. Program: initial isocratic at 98% A (3 min), then linear gradient

to 50% B (in 20 min) and after to 95% (in 1 min). Final wash with 98% B for 8 min. Flow rate: 0.20 mL/min. UV wavelength: 260 nm.

The concentration of the PNA was calculated using Evolution 260 Bio UV-Visible spectrophotometer with Peltier thermos tatting accessories (Thermo Fisher Scientific SPE8W), following the UV-absorbance at 260 nm and assuming an additive contribution of all bases and the RNase L recruiter.

For PAGE analysis: ChemiDoc MP Imaging System by Bio-Rad Laboratories S.R.L. for gel imaging.

### 4.5.2 Synthesis of RNase L recruiter

Compound **3**

A solution of 3,4-dihydroxybenzaldehyde **1** (50.5 mg, 0.37 mmol), BrPEG **2** (105 µL, 0.37 mmol) and $K_2CO_3$ (50 mg, 0.37 mmol) in DMF (1 mL) was allowed to react at 50 °C overnight. The solvent was removed under reduced pressure and the product was purified by silica gel column chromatography (Hex:AcOEt:MeOH = 4:6:0.5). The desired product **3** was a light-yellow oil (37.3 mg, 25%). Characterizations were consistent with previously reported data.[1]

$^1$H NMR (400 MHz, $CD_3OD$_SPE): δ 9.77 (s, 1H), 7.42 (dd, *J* = 8.3, 2.0 Hz, 1H), 7.33 (d, *J* = 2.0 Hz, 1H), 7.12 (d, *J* = 8.2 Hz, 1H), 4.32 – 4.25 (m, 2H), 3.97 – 3.88 (m, 2H), 3.79 – 3.55 (m, 8H), 3.49 (t, *J* = 5.6 Hz, 2H), 3.21 (t, *J* = 5.6 Hz, 2H), 1.43 (s, 9H).

Compound **5**

A solution of ethyl 4,5-dihydro-4-oxo-2-(phenylamino)-3-thiophenecarboxylate **4** (27.5 mg, 0.10 mmol), aldehyde **3** (35.8 mg, 0.09 mmol) and piperidine (8 µL, 0.08 mmol) in $EtOH_{abs}$ (2 mL) was heated at 80 °C (100 W) by microwave for 6h. The solvent was removed under reduced pressure and the product was purified by silica gel column chromatography (AcOEt:MeOH = 10:1). The product **5** was a light orange oil (57.1 mg, 75%). Characterizations were consistent with previously reported data.[1]

$^1$H NMR (400 MHz, $CDCl_3$): δ 11.49 (s, 1H), 7.73 (s, 1H), 7.49 (m, 2H), 7.42 – 7.35 (m, 3H), 7.19 (s, 1H), 7.12 (d, *J* = 2.2 Hz, 1H), 7.01 (dd, *J* = 8.4, 2.2 Hz, 1H), 6.89 (d, *J* = 8.4

Hz, 1H), 5.13 (s, 1H), 4.41 (q, *J* = 7.1 Hz, 2H), 4.23 – 4.17 (m, 2H), 3.88 – 3.82 (m, 2H), 3.76 – 3.58 (m, 8H), 3.53 (m, 2H), 3.34 – 3.26 (m, 2H), 1.44 (t, *J* = 7.1 Hz, 3H), 1.41 (s, 9H).

Compound **6**

A solution of **5** (57.1 mg, 87 nmol) and HCl 4M in 1,4-dioxane (0.65 mL, 2.6 mmol) in DCM (2 mL) was stirred at room temperature for 2 h. The solvent was removed under reduced pressure to give the desired product as a red oil (48.6 mg, 99%). Characterizations were consistent with previously reported data.[1]

$^1$H NMR (400 MHz, $CDCl_3$): δ 11.64 (s, 1H), 10.27 (s, 1H), 8.07 (brs, 3H), 7.65 (s, 1H), 7.57 – 7.44 (m, 5H), 7.32 (m, 1H), 6.88 (d, *J* = 8.0 Hz, 1H), 6.45 (d, *J* = 8.2 Hz, 1H), 4.42 (q, *J* = 7.1 Hz, 2H), 3.87 – 3.81 (m, 2H), 3.76 – 3.71 (m, 2H), 3.69 – 3.60 (m, 8H), 3.60 – 3.53 (m, 2H), 3.28 – 3.24 (m, 2H), 1.45 (t, *J* = 7.1 Hz, 3H).

Compound **7**

Under dry conditions, a solution of **6** (39.8 mg, 0.07 mmol), succinic anhydride (10.7 mg, 0.11 mmol) and DIPEA (25 μL, 0.14 mmol) in DMF dry (2 mL) was stirred at r.t. overnight. The solvent was removed under reduced pressure and the product was purified by silica gel chromatography (AcOEt:MeOH = 10:1 -> 0:1) to give **7** as a yellow solid (41.7 mg, 89%)

$^1$H NMR (400 MHz, $CDCl_3$): δ 11.46 (brs, 1H), 7.62 (s, 1H), 7.50 – 7.42 (m, 2H), 7.40 – 7.30 (m, 3H), 7.08 (s, 1H), 6.85 (d, *J* = 8.2 Hz, 1H), 6.66 (d, *J* = 8.5 Hz, 1H), 4.39 (q, *J* = 7.1 Hz, 2H), 4.01 (m, 2H), 3.78 (m, 2H), 3.68 – 3.54 (m, 8H), 3.51 (m, 2H), 3.35 (m, 2H), 2.43 (m, 4H), 1.42 (t, *J* = 7.1 Hz, 3H).

$^{13}$C NMR (101 MHz, $CDCl_3$): δ 182.3, 180.4, 176.0, 174.6, 166.9, 148.5, 147.7, 137.3, 131.9, 129.9, 127.8, 127.6, 125.2, 123.9, 121.8, 118.4, 112.6, 97.9, 69.9, 69.7, 69.3, 68.7, 67.1, 60.6, 38.3, 33.7, 33.0, 14.5.

UPLC-MS (ESI neg): Calcd for $C_{32}H_{37}N_2O_{11}S^-$ $[M-H]^-$ 657.22, found 657.1.

## 4.5.3 Synthesis of variant RNase L recruiter

Compound **1b**

In a round-bottom flask, $BrPEG_4COOH$ (260 mg, 0.91 mmol) **1a** was dissolved in MeOH (2 mL). After that, two drops of $H_2SO_4$ concentrated were added at the solution as catalyst. The solution was heated under reflux and left to react for 5 hours. The solvent was removed under reduced pressure and the product was extracted with a solution of H2O/DCM and washed with a solution of $H_2O$/saturated solution of NaCl. The solvent was removed again under reduced pressure and the desired product **1b** was a colorless oil (257 mg, 94%).

$^1$H NMR (400 MHz, $CDCl_3$) δ 3.81 (t, *J* = 6.3 Hz, 2H), 3.76 (t, *J* = 6.5 Hz, 2H), 3.68 (s, 3H), 3.67 – 3.60 (m, 8H), 3.47 (t, *J* = 6.3 Hz, 2H), 2.60 (t, *J* = 6.5 Hz, 2H).

UPLC-MS (ESI pos, MeOH): Calcd for $C_{10}H_{19}BrO_5$ $[M+H]^+$ 300.04, found 299.4 [100%, $^{79}M+H]^+$, 301.3 [100%, $^{81}M+H]^+$.

Compound **2b**

In a round-bottom flask, a solution of 3,4dihydroxybenzaldehyde **1** (59.3 mg, 0.43 mmol), $NaHCO_3$ (54.1 mg, 0.64 mmol), and NaI (19.3 mg, 0.1 3mmol) in DMF dry (1 mL) was stirred at 40°C for 2 hours under nitrogen condition. After that a solution of **1b** (257.0 mg, 0.86 mmol) in DMF dry (1 ml) was added. The reaction was left to react at 50°C for 40 hours. The reaction was turned off with $H_2O$ (20 mL) and the product was extracted with dichloromethane (3x10 mL). Then the organic phase was washed with $H_2O$/saturated solution of NaCl. The solvent was removed under reduced pressure and the product was purified by silica gel column chromatography (Hex:AcOEt:MeOH = 4:6:0.5). The desired product **2b** was a light-yellow oil (77.8 mg, 51%).

$^1$H NMR (400 MHz, $CDCl_3$) δ 9.87 (s, 1H), 7.46 (d, *J* = 2.0 Hz, 1H), 7.42 (dd, *J* = 8.2, 2.0 Hz, 1H), 7.02 (d, *J* = 8.2 Hz, 1H), 6.76 (s, 1H), 4.31 – 4.27 (m, 2H), 3.94 – 3.90 (m, 2H), 3.78 (t, *J* = 6.5 Hz, 2H), 3.77 – 3.73 (m, 2H), 3.73 – 3.71 (m, 2H), 3.70 (s, 3H), 3.67 (q, *J* = 1.6 Hz, 4H), 2.62 (t, *J* = 6.5 Hz, 2H).

$^{13}$C NMR (101 MHz, $CDCl_3$) δ 191.2, 172.1, 151.5, 147.4, 131.3, 123.7, 115.5, 112.6, 70.7, 70.5, 70.4, 70.4, 69.2, 68.8, 66.6, 51.7, 34.7.

UPLC-MS (ESI pos, MeOH): Calcd for $C_{17}H_{24}O_8$ $[M+H]^+$ 356.15, found 379.4 [100%, $M+Na]^+$.

Compound **3b**

In a round-bottom flask, compound **2b** (77.8 mg, 0.22 mmol) was dissolved in MeOH (2 ml) and after that a solution of 4M LiOH in water (55 µl x2, 0.22 mmol) was added. The solution was heated under reflux and left to react over weekend. The solvent was removed under reduced pressure and 1 M HCl solution (500µl) and water (2 mL) were added. After that the product was extracted with ethyl acetate (AcOEt) and the organic solution was removed under reduced pressure and the product was purified by silica gel column chromatography (Hex:AcOEt:MeOH = 2:8:1→0:10:1→0:10:2). The desired product **3b** was a pink resin (29.5 mg, 40%).

$^1$H NMR (400 MHz, $CDCl_3$) δ 9.81 (s, 1H), 7.41 (d, *J* = 2.0 Hz, 1H), 7.36 (dd, *J* = 8.2, 1.9 Hz, 1H), 6.94 (d, *J* = 8.2 Hz, 1H), 6.34 (s, 2H), 4.26 – 4.20 (m, 2H), 3.94 – 3.84 (m, 2H), 3.77 – 3.73 (m, 2H), 3.73 – 3.70 (m, 2H), 3.69 – 3.65 (m, 2H), 3.63 (m, 4H), 2.61 (t, *J* = 6.2 Hz, 2H).

$^{13}$C NMR (101 MHz, $CDCl_3$) δ 191.3, 175.5, 151.6, 147.2, 131.1, 123.8, 115.5, 112.3, 70.6, 70.3, 70.2, 69.2, 68.4, 66.6, 34.8.

UPLC-MS (ESI neg, MeOH): Calcd for $C_{16}H_{22}O_8$ $[M-H]^-$ 341.13, found 341.2 [100%, $M-H]^-$.

Compound **4b**

A solution of ethyl 4,5-dihydro-4-oxo-2-(phenylamino)-3-thiophenecarboxylate **4** (27.2 mg, 0.10 mmol), aldehyde **3b** (29.5 mg, 0.09 mmol) and piperidine (17 µL, 0.17 mmol) in $EtOH_{abs}$ (2 mL) was heated at 80 °C (100 W) by microwave for 4h. The solvent was removed under reduced pressure and the product was purified by silica gel column chromatography (AcOEt:MeOH = 9:1). The product **4b** was a light orange resin (57.1 mg, 75%).

$^{1}$H NMR (400 MHz, $CDCl_3$) δ 11.49 (s, 1H), 7.58 (s, 1H), 7.45 (s, 2H), 7.36 (d, *J* = 8.1 Hz, 3H), 6.95 (s, 1H), 6.74 (d, *J* = 8.1 Hz, 1H), 6.57 (s, 1H), 4.41 (q, *J* = 7.0 Hz, 2H), 4.02 (m, 2H), 3.80 (m, 2H), 3.65 (m, 8H), 3.55 (m, 2H), 2.41 (m, 2H), 1.44 (t, *J* = 7.0 Hz, 3H).
$^{13}$C NMR (101 MHz, $CDCl_3$) δ 182.3, 179.3, 175.9, 167.1, 148.7, 137.3, 132.3, 129.9, 127.5, 123.7, 121.3, 112.2, 98.0, 66.7, 60.6, 50.9, 37.8, 14.5.
UPLC-MS (ESI neg, MeOH): Calcd for $C_{29}H_{33}O_{10}S$ $[M-H]^-$ 586.18, found 586.2 [100%, $^{79}$M-H]$^-$.

### 4.5.4 Synthesis of peptide nucleic acid

All the PNA sequences designed were synthetized with standard solid-phase manual synthesis, using 9-fluorenylmethoxycarbonyl (Fmoc) strategy . The PNA monomers employed to Fmoc strategy are Fmoc/Bhoc protected.

Synthesis of the first set of PNAs (*Tables 4.2 and 4.3*)
The H-Rink Amide Chemmatrix resin was loaded with Fmoc-Gly-OH to obtain 0.2 mmol/gr as theoretical loading. The principal steps to load the resin were: a) swelling in DCM (1x1h), b) DMF dry wash, c) coupling (1x5h, activation for 10', using 1 eq of Fmoc-Gly-OH, 10 eq of DIC and 10 eq of DhBtOH in DMF dry as activation solution), d) DMF wash, e) capping (2x15', DMF:$Ac_2O$, 1:1), e) DMF wash and f) DCM wash. To measure the loading of the resin a UV/Visible technique was used, following the absorbance of divenzofulvene (DBF) formation after deprotection treatment (30', Piperidine:DMF, 1:4) at 290 nm. In fact, the values at 290 nm and 400 nm (to subtract the baseline contribution) were extracted from the measure. The loading of the resin was calculated dividing the value obtained from the difference between the absorbance at 290 nm and at 400 nm with a value obtained by multiplication of a conversion factor (1.65 g/mmol*mg) with the weighed resin mass.
The synthesis of PNA was performed in a 10 µmol as a scale, separating 5 µmol after attacking all the PNA monomer and continuing the synthesis in the remaining resin to obtain the product with the octa-arginine peptide moiety (using Fomc-Arg(Pbf)-OH). Subsequently, in 2 µmol was carried out the cleavage by the resin while in 1.5

μmol was performed the coupling with the RNase L recruiter and then the cleavage step.

The principal steps and solutions for the standard solid-phase manual synthesis were: a) swelling in DCM (1x30'), b) deprotection (2x8', Piperidine:DMF, 1:4), c) DCM wash, d) DMF dry wash, e) Kaiser test (1', positive), f) coupling (1x30', activation for 2', using 5 eq of PNA monomer, 4.9 eq of HBTU as coupling reagent and 10 eq of DIPEA in DMF dry as activation solution), g) DMF wash, h) Kaiser test (1', negative), i) capping (2x1', $Ac_2O$:DIPEA:DMF dry, 5:6:89), l) DMF wash, m) 5% DIPEA wash (2x2', DIPEA:DMF, 5:95), n) DMF wash and, o) DCM wash.

Synthesis of the second set of PNAs (*Table 4.5*)

The H-Rink Amide Chemmatrix resin was loaded with Fmoc-Lys(Mtt)-OH to obtain 0.2 mmol/gr as theoretical loading. The principal steps to load and to measure the loading of the resin are the same of before.

The synthesis of PNA was performed in a 20 μmol as a scale, separating 2.5 μmol after attacking all the PNA monomer in six reactors. In one design, three different peptide nucleic acid sequences were synthetized, where one present only the RNase L recruiter at the end of $N_{term}$ moiety in addition to the PNAs monomer, instead the other two have one or two AEEA spacer and the RNase L recruiter. In the other design, three different PNAs were synthetized using an orthogonal chemistry since the RNase L recruiter was attached on the lysine moiety at the $C_{term}$. One present only the RNase L recruiter at the end of $C_{term}$ moiety in addition to the PNAs monomer, instead the other two have one or two AEEA spacer and the RNase L recruiter. Subsequently, in 1.5 μmol was carried out the cleavage by the resin.

The principal steps and solutions for the standard solid-phase manual synthesis are the same of before with the addition of Mtt deprotection (0.5% BtOH•H2O in HFIP/DCM 1:1) and the coupling with the AEEA spacer (1x40', activation for 2', using 5 eq of AEEA spacer, 4.9 eq of HBTU as coupling reagent and 10 eq of DIPEA in DMF dry as activation solution).

For the coupling of the RNase L recruiter, after the tests previously discussed in *Results and Discussion part*, the following protocol was selected: a) deprotection (2x8', Piperidine:DMF, 1:4), b) DCM wash, c) DMF wash, d) coupling (o.n., activation for 2', using 10 eq of RNase L recruiter, 10 eq of DIC as coupling reagent and 10 eq of DhBtOH in DMF dry as activation solution), e) DMF wash, and f) DCM wash.
The solutions used for the Kaiser test were as those described in *Chapter 2*.
For the cleavage step a cocktail solution composed of TFA:m-cresol (9:1) was used and two cycles of 1h were performed to make sure a complete cleavage. After this first cycle the solution was filtered and collected in a falcon while the resin was washed with only TFA to ensure a complete collection of the product, and the same procedure was repeated at the end of the second cycle. The PNA oligomers were precipitated in diethyl ether in freezer for at least 2h. After removing diethyl ether, the product was dried and dissolved in double distilled water.
All the PNAs were purified by HPLC and characterized by UPLC-MS and UV-Visible techniques. In relation to this last technique, the values at 260 nm and 400 nm (to subtract the baseline contribution) were extracted by the measurement. In addition, the molar absorptivity (ε) specific for each PNA was calculated considering 13700 $M^{-1}cm^{-1}$ for adenine monomer, 6600 $M^{-1}cm^{-1}$ for cytosine, 11700 $M^{-1}cm^{-1}$ for guanine and 8600 $M^{-1}cm^{-1}$ for thymine as molar absorptivity values of a single nucleobase and 17465 $M^{-1}cm^{-1}$ as molar absorptivity value calculated for the RNase L recruiter. In presence of the absorbance, the molar absorptivity, and the solution dilution for the UV-Visible measurement, using a Lambert-Beer law, the concentration of each PNAs was calculated. Instead, the reaction yield was calculated using also the PNA solution volume.

### 4.5.5 Polyacrylamide gel electrophoresis (PAGE) experiments

A solution of RNA PK59 with a PNA sequence was annealed by heating at 70 °C for 5 min and slowly cooling to room temperature in RNase L Buffer (25 mM Tris-HCl, 100 mM KCl, 10 mM $MgCl_2$, pH 7,4). ). After cooling, the solution was divided in two batch: one already ready for the next incubation step while in the second the RNase

L and the RNase L supplement (10 mM $MgCl_2$, fresh 7 mM β-mercaptoethanol, and 50 μM of ATP) were added (considering equal ratio of each component: PNA, RNA and RNseL). All the samples are transferred to a thermocycler and incubated for two hours or overnight at 28 °C. The reaction is quenched by heating at 95 °C for 5 minutes. The samples are analyzed with denaturing urea (8 M) 10% PAGE (12 mL of 40% acrylamide/bis-acrylamide (19:1), 120 μL 10% of ammonium persulfate (APS), 12 μL of N, N, N′, N′-tetramethyl ethylene diamine (TEMED) and 4.8 mL of TaMg buffer (400 mM Tris-Base, 125 mM $Mg(AcO)_2$, 200 mM EDTA, pH: 8.0), using TAMg buffer (40 mM Tris-Base, 12.5 mM $Mg(AcO)_2$, 20 mM EDTA, pH: 8.0) as running buffer. A 70% v/v of formamide was added to loading the samples. The gels were run at 150V for 2 hours.

### 4.5.6 PNAs characterizations

Entry 8 (Figure 4.A13 in Appendix)
S-P (MF01-16): r8-CAC CAG GAA CAC AAA CAT-Gly-$NH_2$
yield: 6%; $t_r$: 2.50 min; ε (260 nm): 194900 $M^{-1}cm^{-1}$; MW calculated: 6157.33 [M], observed (from ESI-MS deconvoluted): 6157.8; ESI-MS: m/z observed (calcd): 1232.7 (1232.5) $[M+5H]^{5+}$, 1027.5 (1027.2) $[M+6H]^{6+}$, 880.6 (880.6) $[M+7H]^{7+}$, 770.8 (770.7) $[M+8H]^{8+}$, 685.4 (685.1) $[M+9H]^{9+}$, 616.8 (616.7) $[M+10H]^{10+}$,560.9 (560.8) $[M+11H]^{11+}$.

Entry 9 (Figure 4.A14 in Appendix)
S-St (MF01-17): r8-AAC AAG AAA AAC AAA CAT-Gly-$NH_2$
yield: 13%; $t_r$: 2.86 min; ε (260 nm): 218200 $M^{-1}cm^{-1}$; MW calculated: 6213.39 [M], observed (from ESI-MS deconvoluted):6213.8; ESI-MS: m/z observed (calcd): 1243.6 (1243.7)$[M+5H]^{5+}$, 1036.7 (1036.6) $[M+6H]^{6+}$, 888.8 (888.6) $[M+7H]^{7+}$, 777.8 (777.7) $[M+8H]^{8+}$, 691.5 (691.4) $[M+9H]^{9+}$, 622.4 (622.3) $[M+10H]^{10+}$, 565.9 (565.8) $[M+11H]^{11+}$, 518.7 (518.8) $[M+12H]^{12+}$.

Entry 10 (Figure 4.A15 in Appendix)
S-3′ (CS02-24): r8-CAC TCC ATA ACA CTT-Gly-$NH_2$
yield: 8%; $t_r$: 10.19 min; ε (260 nm): 142500 $M^{-1}cm^{-1}$; MW calculated: 5272.49 [M], observed (from ESI-MS deconvoluted): 5269.5; ESI-MS: m/z observed (calcd): 1318.9 (1319.1) $[M+4H]^{4+}$, 1055.3 (1055.5) $[M+5H]^{5+}$, 879.6 (879.7) $[M+6H]^{6+}$, 754.1 (754.2) $[M+7H]^{7+}$, 659.9 (660.1) $[M+8H]^{8+}$, 586.7 (586.8) $[M+9H]^{9+}$, 528.2 (528.2) $[M+10H]^{10+}$.

Entry 11 (Figure 4.A16 in Appendix)
S-5′ (SS01-15): r8-AAA ACA AAC ATT GTT CGT-Gly-$NH_2$
yield: 12%; $t_r$: 2.84 min; ε (260 nm): 195800 $M^{-1}cm^{-1}$; MW calculated: 6193.35 [M], observed (from ESI-MS deconvoluted): 6192.6; ESI-MS: m/z observed (calcd): 1239.4 (1239.7) $[M+5H]^{5+}$, 1033.2 (1033.2) $[M+6H]^{6+}$, 885.7 (885.8) $[M+7H]^{7+}$, 775.1 (775.2) $[M+8H]^{8+}$, 689.1 (689.2) $[M+9H]^{9+}$.

Entry 12 (Figure 4.A17 in Appendix)
N-P (CS02-14): r8-TGG TCC ATT GTC ACT CAT-Gly-$NH_2$
yield: 5%; $t_r$: 2.71 min; ε (260 nm): 169400 $M^{-1}cm^{-1}$; MW calculated: 6143.29 [M], observed (from ESI-MS deconvoluted): 6142.8; ESI-MS: m/z observed (calcd): 1229.6 (1229.7) $[M+5H]^{5+}$, 1024.9 (1024.9) $[M+6H]^{6+}$, 878.7 (878.6) $[M+7H]^{7+}$, 768.9 (768.9) $[M+8H]^{8+}$, 683.8 (683.6) $[M+9H]^{9+}$.

Entry 13 (Figure 4.A18 in Appendix)
N-St (CS02-15): r8-GGG TCC ATT ATC AGA CAT-Gly-$NH_2$
yield: 5%; $t_r$: 2.62 min; ε (260 nm): 184700 $M^{-1}cm^{-1}$; MW calculated: 6021.33 [M], observed (from ESI-MS deconvoluted): 6201.2; ESI-MS: m/z observed (calcd): 1241.0 (1241.3) $[M+5H]^{5+}$, 1034.6 (1034.6) $[M+6H]^{6+}$, 887.0 (886.9) $[M+7H]^{7+}$, 776.2 (776.2) $[M+8H]^{8+}$, 690.1 (690.0) $[M+9H]^{9+}$.

Entry 14 (Figure 4.A19 in Appendix)
N-3′ (MF01-18): r8-TC TCC ATT CTG GTT-Gly-$NH_2$
yield: 6%; $t_r$: 2.64 min; ε (260 nm): 123700 $M^{-1}cm^{-1}$; MW calculated: 5050.23 [M], observed (from ESI-MS deconvoluted): 5048.4; ESI-MS: m/z observed (calcd): 1263.4 (1263.6) $[M+4H]^{4+}$, 1011.1 (1011.1) $[M+5H]^{5+}$, 842.5 (842.7) $[M+6H]^{6+}$, 722.2 (722.5) $[M+7H]^{7+}$, 632.3 (632.3) $[M+8H]^{8+}$.

Entry 15 (Figure 4.A20 in Appendix)
N-5′ (SS01-19): r8-ACA TTT TAG TTT GTT CGT-Gly-$NH_2$
yield: 11%; $t_r$: 2.74 min; ε (260 nm): 175400 $M^{-1}cm^{-1}$; MW calculated: 6188.32 [M], observed (from ESI-MS deconvoluted): 6187.6; ESI-MS: m/z observed (calcd): 1238.8 (1238.7) $[M+5H]^{5+}$, 1032.6 (1032.4) $[M+6H]^{6+}$, 885.0 (885.0) $[M+7H]^{7+}$, 774.6 (774.5) $[M+8H]^{8+}$, 688.0 (688.6) $[M+9H]^{9+}$.

Entry 16 (Figure 4.A21 in Appendix)
PK (MF01-22): r8-A AAG CCC TGT ATA CGA CAT-Gly-$NH_2$
yield: 7%; $t_r$: 2.72 min; ε (260 nm): 198400 $M^{-1}cm^{-1}$; MW calculated: 6445.59 [M], observed (from ESI-MS deconvoluted): 6444.6; ESI-MS: m/z observed (calcd): 1289.8 (1290.1) $[M+5H]^{5+}$, 1075.2 (1075.3) $[M+6H]^{6+}$, 921.9 (921.8) $[M+7H]^{7+}$, 806.7 (806.7) $[M+8H]^{8+}$, 717.1 (717.2) $[M+9H]^{9+}$, 645.6 (654.6) $[M+10H]^{10+}$.

Entry 17 (Figure 4.A22 in Appendix)
C-18mer (SS01-21): r8-TAT CCA GTC AAG ATC TAA-Gly-$NH_2$
yield: 9%; $t_r$: 2.61 min; ε (260 nm): 188700 $M^{-1}cm^{-1}$; MW calculated: 6169.33 [M], observed (from ESI-MS deconvoluted): 6169.0; ESI-MS: m/z observed (calcd): 1029.0 (1029.2) $[M+6H]^{6+}$, 882.3 (882.3) $[M+7H]^{7+}$, 772.1 (772.2) $[M+8H]^{8+}$, 686.3 (686.5) $[M+9H]^{9+}$.

Entry 18 (Figure 4.A23 in Appendix)
C-15mer (MF01-29): r8-ACA CTC TAC ATC ACT-Gly-$NH_2$
yield: 28%; $t_r$: 10.09 min; ε (260 nm): 142500 $M^{-1}cm^{-1}$; MW calculated: 5272.49 [M], observed (from ESI-MS deconvoluted): 5269.5; ESI-MS: m/z observed (calcd): 1318.9 (1319.1) $[M+4H]^{4+}$, 1055.3 (1055.5) $[M+5H]^{5+}$, 879.6 (879.7) $[M+6H]^{6+}$, 754.1 (754.2) $[M+7H]^{7+}$, 659.9 (660.1) $[M+8H]^{8+}$, 586.7 (586.8) $[M+9H]^{9+}$, 528.2 (528.2) $[M+10H]^{10+}$, 480.2 (480.3) $[M+11H]^{11+}$.

Entry 19 (Figure 4.A24 in Appendix)
S-P+L (MF01-58): RNase L recruiter-r8-CAC CAG GAA CAC AAA CAT-Gly-$NH_2$
yield: 6%; $t_r$: 3.46 min; ε (260 nm): 212365 $M^{-1}cm^{-1}$; MW calculated: 6799.04 [M], observed (from ESI-MS deconvoluted):6798.4; ESI-MS: m/z observed (calcd): 972.6 (972.3) $[M+7H]^{7+}$, 850.8 (850.9) $[M+8H]^{8+}$, 756.3 (756.4) $[M+9H]^{9+}$, 680.9 (680.9) $[M+10H]^{10+}$, 619.0 (619.1) $[M+11H]^{11+}$.

Entry 20 (Figure 4.A25 in Appendix)
S-St+L (MF01-61): RNase L recruiter-r8-AAC AAG AAA AAC AAA CAT-Gly-$NH_2$
yield: 9%; $t_r$: 3.41 min; ε (260 nm): 235665 $M^{-1}cm^{-1}$; MW calculated: 6855.10 [M], observed (from ESI-MS deconvoluted): 6855.6; ESI-MS: m/z observed (calcd): 1372.8 (1372.0) $[M+5H]^{5+}$, 1144.0 (1143.5) $[M+6H]^{6+}$, 980.5 (980.3) $[M+7H]^{7+}$, 858.0 (857.9) $[M+8H]^{8+}$, 762.8 (762.7) $[M+9H]^{9+}$, 686.8 (686.5) $[M+10H]^{10+}$, 624.3 (624.2) $[M+11H]^{11+}$, 572.6 (572.3) $[M+12H]^{12+}$, 528.4 (528.3) $[M+13H]^{13+}$.

Entry 21 (Figure 4.A26 in Appendix)
S-3'+L (MF01-62): RNase L recruiter-r8-CAC TCC ATA ACA CTT-Gly-$NH_2$
yield: 9%; $t_r$: 3.46 min; ε (260 nm): 159965 $M^{-1}cm^{-1}$; MW calculated: 5914.20 [M], observed (from ESI-MS deconvoluted): 5914.6; ESI-MS: m/z observed (calcd): 1184.1 (1183.8) $[M+5H]^{5+}$, 987.0 (986.7) $[M+6H]^{6+}$, 845.9 (845.9) $[M+7H]^{7+}$, 740.3 (740.3) $[M+8H]^{8+}$, 658.3 (658.1) $[M+9H]^{9+}$, 592.5 (592.4) $[M+10H]^{10+}$, 538.4 (538.6) $[M+11H]^{11+}$.

Entry 23 (Figure 4.A27 in Appendix)
N-P+L (MF01-57): RNase L recruiter-r8-TGG TCC ATT GTC ACT CAT-Gly-$NH_2$
yield: 3%; $t_r$: 3.58 min; ε (260 nm): 186865 $M^{-1}cm^{-1}$; MW calculated: 6785.00 [M], observed (from ESI-MS deconvoluted): 6784.6; ESI-MS: m/z observed (calcd): 1357.6 (1358.0) $[M+5H]^{5+}$, 1131.5 (1131.8) $[M+6H]^{6+}$, 970.3 (970.3) $[M+7H]^{7+}$, 849.0 (849.1) $[M+8H]^{8+}$, 754.9 (754.9) $[M+9H]^{9+}$.

Entry 24 (Figure 4.A28 in Appendix)
N-St+L (CS02-31): RNase L recruiter-r8-GGG TCC ATT ATC AGA CAT-Gly-$NH_2$
yield: 2%; $t_r$: 15.65 min; ε (260 nm): 202165 $M^{-1}cm^{-1}$; MW calculated: 6843.05 [M], observed (from ESI-MS deconvoluted): 6838.0; ESI-MS: m/z observed (calcd): 1141.3 (1141.5) $[M+6H]^{6+}$, 978.3 (978.6) $[M+7H]^{7+}$, 856.1 (856.4) $[M+8H]^{8+}$, 761.1 (761.39 $[M+9H]^{9+}$, 685.1 (685.3) $[M+10H]^{10+}$.

Entry 25 (Figure 4.A29 in Appendix)
N-3'+L (CS02-32): RNase L recruiter-r8-TC TCC ATT CTG GTT-Gly-$NH_2$
yield: 5%; $t_r$: 15.72 min; ε (260 nm): 141165 $M^{-1}cm^{-1}$; MW calculated: 5691.94 [M], observed (from ESI-MS deconvoluted): 5687.5; ESI-MS: m/z observed (calcd): 949.4 (949.7) $[M+6H]^{6+}$, 813.9 (814.1) $[M+7H]^{7+}$, 712.3 (712.5) $[M+8H]^{8+}$, 633.3 (633.4) $[M+9H]^{9+}$.

Entry 27 (Figure 4.A30 in Appendix)
PK+L (MF01-69): RNase L recruiter-r8-A AAG CCC TGT ATA CGA CAT-Gly-$NH_2$
yield: 3%; $t_r$: 3.49 min; ε (260 nm): 215865 $M^{-1}cm^{-1}$; MW calculated: 7087.30 [M], observed (from ESI-MS deconvoluted): 7088.2; ESI-MS: m/z observed (calcd): 1418.6 (1418.5) $[M+5H]^{5+}$, 1182.5 (1182.2) $[M+6H]^{6+}$, 1013.6 (1013.5) $[M+7H]^{7+}$, 886.9 (886.9) $[M+8H]^{8+}$, 788.5 (788.5) $[M+9H]^{9+}$, 709.8 (709.7) $[M+10H]^{10+}$, 645.2 (645.3) $[M+11H]^{11+}$.

Entry 28 (Figure 4.A31 in Appendix)
C-18mer+L (CS02-70): RNase L recruiter-r8-TAT CCA GTC AAG ATC TAA-Gly-$NH_2$
yield: 14%; $t_r$: 3.47 min; ε (260 nm): 206165 $M^{-1}cm^{-1}$; MW calculated: 6811.04 [M], observed (from ESI-MS deconvoluted):6808.8; ESI-MS: m/z observed (calcd): 1362.8 (1363.2) $[M+5H]^{5+}$, 1135.9 (1136.2) $[M+6H]^{6+}$, 973.7 (974.0) $[M+7H]^{7+}$, 852.1 (852.4) $[M+8H]^{8+}$, 757.5 (757.8) $[M+9H]^{9+}$, 681.9 (682.1) $[M+10H]^{10+}$, 620.1 (620.1) $[M+11H]^{11+}$.

Entry 29 (Figure 4.A32 in Appendix)
C-15mer+L (CS02-69): RNase L recruiter-r8-ACA CTC TAC ATC ACT-Gly-$NH_2$
yield: 4%; $t_r$: 3.51 min; ε (260 nm): 159965 $M^{-1}cm^{-1}$; MW calculated: 5914.20 [M], observed (from ESI-MS deconvoluted): 5912.8; ESI-MS: m/z observed (calcd): 986.4 (986.7) $[M+6H]^{6+}$, 845.6 (845.9) $[M+7H]^{7+}$, 740.0 (740.3) $[M+8H]^{8+}$, 657.9 (658.1) $[M+9H]^{9+}$, 592.0 (592.4) $[M+10H]^{10+}$.

Entry 37 (Figure 4.A33 in Appendix)
BeCa01-12: RNase L recruiter-A AAG CCC TGT ATA CGA CAT-Lys(Ac)-$NH_2$
yield: 1%; $t_r$: 4.26 min; ε (260 nm): 215865 $M^{-1}cm^{-1}$; MW calculated: 5950.94 [M], observed (from ESI-MS deconvoluted):5951.0; ESI-MS: m/z observed (calcd): 1488.5 (1488.7) $[M+4H]^{4+}$, 1190.9 (1191.2) $[M+5H]^{5+}$, 992.6 (992.8) $[M+6H]^{6+}$, 851.0 (851.1) $[M+7H]^{7+}$, 744.9 (744.9) $[M+8H]^{8+}$, 662.2 (662.2) $[M+9H]^{9+}$, 596.2 (596.1) $[M+10H]^{10+}$.

Entry 38 (Figure 4.A34 in Appendix)
BeCa01-13: RNase L recruiter-AEEA-A AAG CCC TGT ATA CGA CAT-Lys(Ac)-$NH_2$
yield: 1%; $t_r$: 4.38 min; ε (260 nm): 215865 $M^{-1}cm^{-1}$; MW calculated: 6096.10 [M], observed (from ESI-MS deconvoluted): 6096.0; ESI-MS: m/z observed (calcd): 1220.0 (1220.2) $[M+5H]^{5+}$, 1017.0 (1017.0) $[M+6H]^{6+}$, 871.8 (871.9) $[M+7H]^{7+}$, 762.9 (763.0) $[M+8H]^{8+}$, 678.4 (678.3) $[M+9H]^{9+}$, 610.5 (610.6) $[M+10H]^{10+}$.

Entry 39 (Figure 4.A35 in Appendix)
BeCa01-17: RNase L recruiter-AEEA-AEEA-A AAG CCC TGT ATA CGA CAT-Lys(Ac)-$NH_2$
yield: 5%; $t_r$: 4.43 min; ε (260 nm): 215865 $M^{-1}cm^{-1}$; MW calculated: 6241.26 [M], observed (from ESI-MS deconvoluted): 6241.0; ESI-MS: m/z observed (calcd): 1249.2 (1249.3) $[M+5H]^{5+}$, 1041.2 (1041.2) $[M+6H]^{6+}$, 892.5 (892.6) $[M+7H]^{7+}$, 781.2 (781.2) $[M+8H]^{8+}$, 694.6 (694.5) $[M+9H]^{9+}$, 625.1 (625.1) $[M+10H]^{10+}$, 568.5 (568.4) $[M+11H]^{11+}$, 391.5 (391.1) $[M+16H]^{16+}$.

Entry 40 (Figure 4.A36 in Appendix)
BeCa01-18: Ac-A AAG CCC TGT ATA CGA CAT-Lys(RNase L recruiter→)-$NH_2$
yield: 6%; $t_r$: 4.33 min; ε (260 nm): 215865 $M^{-1}cm^{-1}$; MW calculated: 5950.94 [M], observed (from ESI-MS deconvoluted):5951.0; ESI-MS: m/z observed (calcd): 1488.4 (1488.7) $[M+4H]^{4+}$, 1191.1 (1191.2) $[M+5H]^{5+}$, 992.9 (992.8) $[M+6H]^{6+}$, 851.2 (851.1) $[M+7H]^{7+}$, 744.9 (744.9) $[M+8H]^{8+}$, 662.2 (662.2) $[M+9H]^{9+}$, 596.2 (596.1) $[M+10H]^{10+}$, 541.9 (542.0) $[M+11H]^{11+}$.

Entry 41 (Figure 4.A37 in Appendix)
BeCa01-19: Ac-A AAG CCC TGT ATA CGA CAT-Lys(RNase L recruiter-AEEA→)-$NH_2$
yield: 4%; $t_r$: 4.31 min; ε (260 nm): 215865 $M^{-1}cm^{-1}$; MW calculated: 6096.10 [M], observed (from ESI-MS deconvoluted): 6096.0; ESI-MS: m/z observed (calcd): 1220.0 (1220.2) $[M+5H]^{5+}$, 1017.0 (1017.0) $[M+6H]^{6+}$, 871.8 (871.9) $[M+7H]^{7+}$, 763.0 (763.0) $[M+8H]^{8+}$, 678.3 (678.3) $[M+9H]^{9+}$, 610.5 (610.6) $[M+10H]^{10+}$, 555.0 (555.2) $[M+11H]^{11+}$.

Entry 42 (Figure 4.A38 in Appendix)
BeCa01-20: Ac-A AAG CCC TGT ATA CGA CAT-Lys(RNase L recruiter-AEEA-AEEA→)-$NH_2$
yield: 12%; $t_r$: 4.46 min; ε (260 nm): 215865 $M^{-1}cm^{-1}$; MW calculated: 6241.26 [M], observed (from ESI-MS deconvoluted): 6241.0; ESI-MS: m/z observed (calcd): 1249.0 (1249.3) $[M+5H]^{5+}$, 1041.1 (1041.2) $[M+6H]^{6+}$, 892.4 (892.6) $[M+7H]^{7+}$, 781.1 (781.2) $[M+8H]^{8+}$, 694.4 (694.5) $[M+9H]^{9+}$, 625.0 (625.1) $[M+10H]^{10+}$, 568.4 (568.4) $[M+11H]^{11+}$.

The biological experiments and the elaboration of the data were performed by laboratory of Prof. Roberto Gambari from University of Ferrara. The experimental part of these is reported in a paper submitted for publication.

# 4.6 Appendix Chapter 4

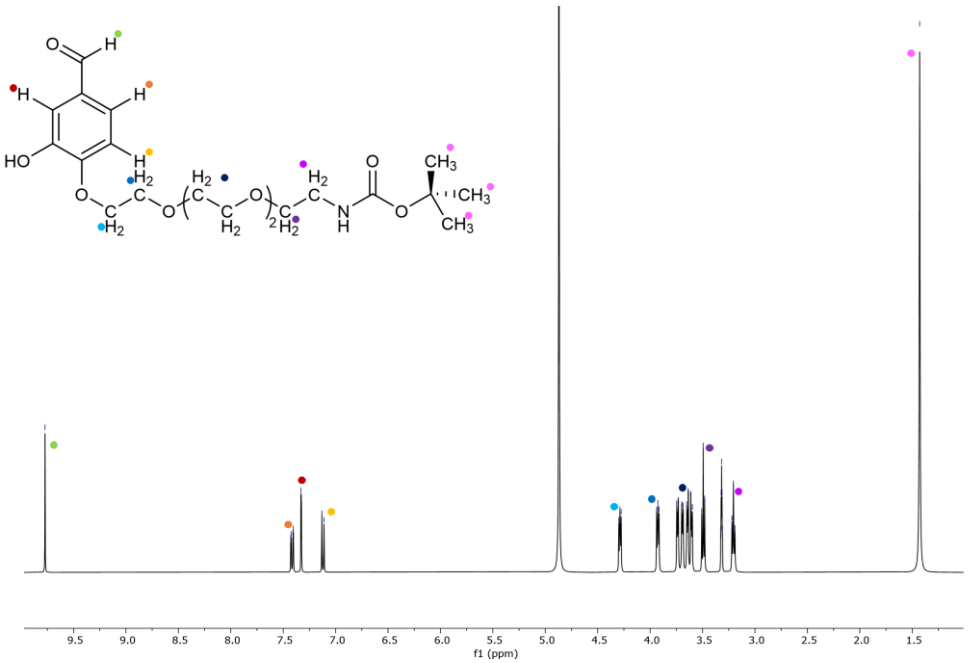



***Figure 4.A1.*** *$^1$H-NMR Spectrum (400 MHz, $CD_3OD$_SPE, 25 °C) of compound 3.*

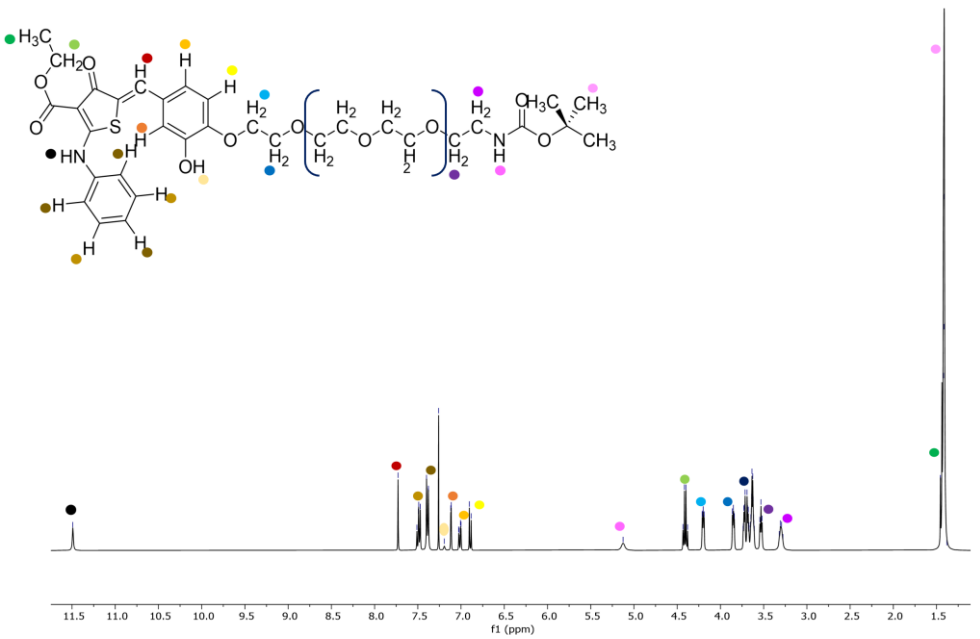



***Figure 4.A2.*** *$^1$H-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 5.*

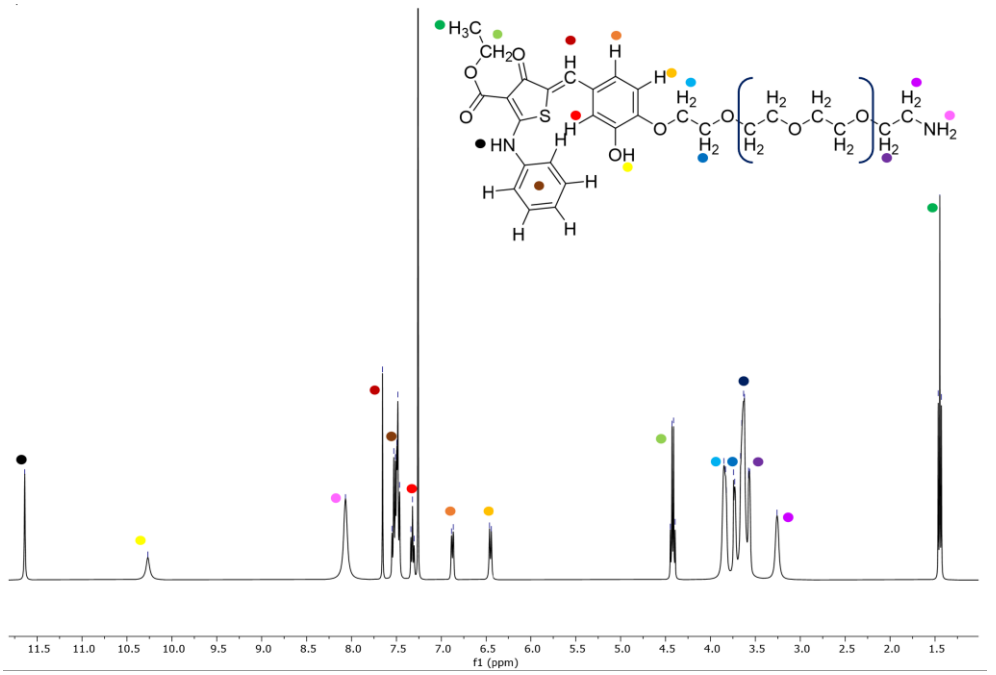



***Figure 4.A3.*** *$^1$H-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 6.*

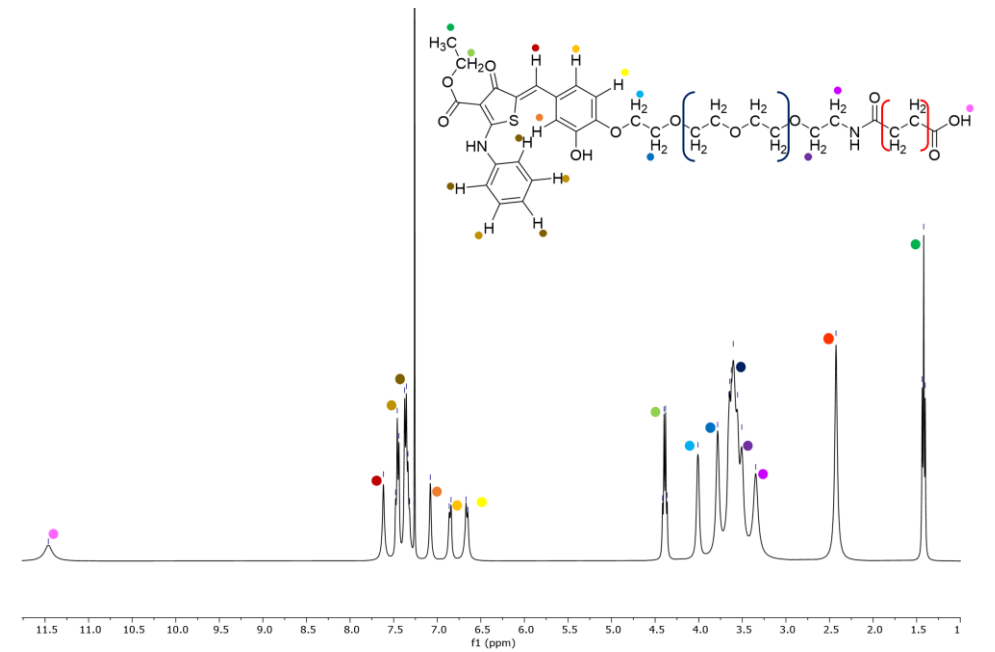



***Figure 4.A4.*** *$^1$H-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 7.*

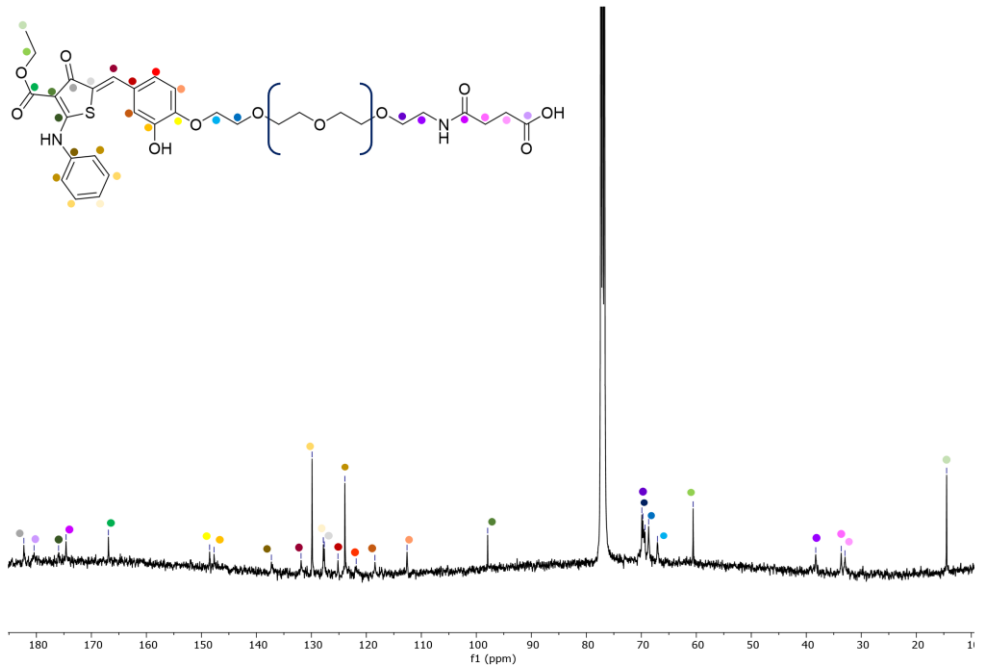



***Figure 4.A5.*** *$^{13}$C-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 7.*

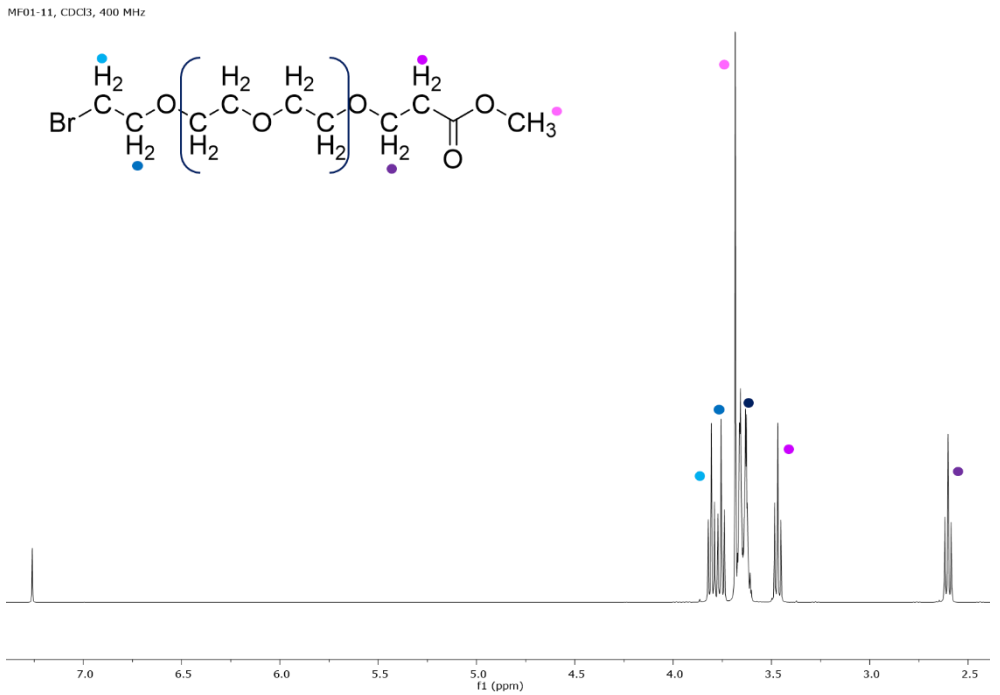



***Figure 4.A6.*** *$^{1}$H-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 1b.*

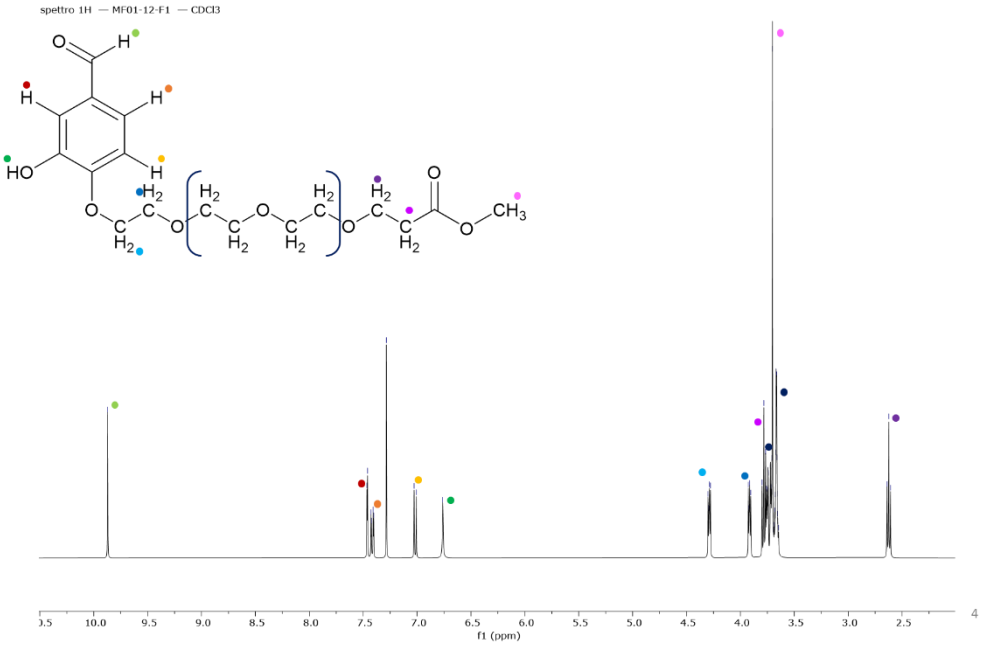



***Figure 4.A7.*** *$^{1}$H-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 2b.*

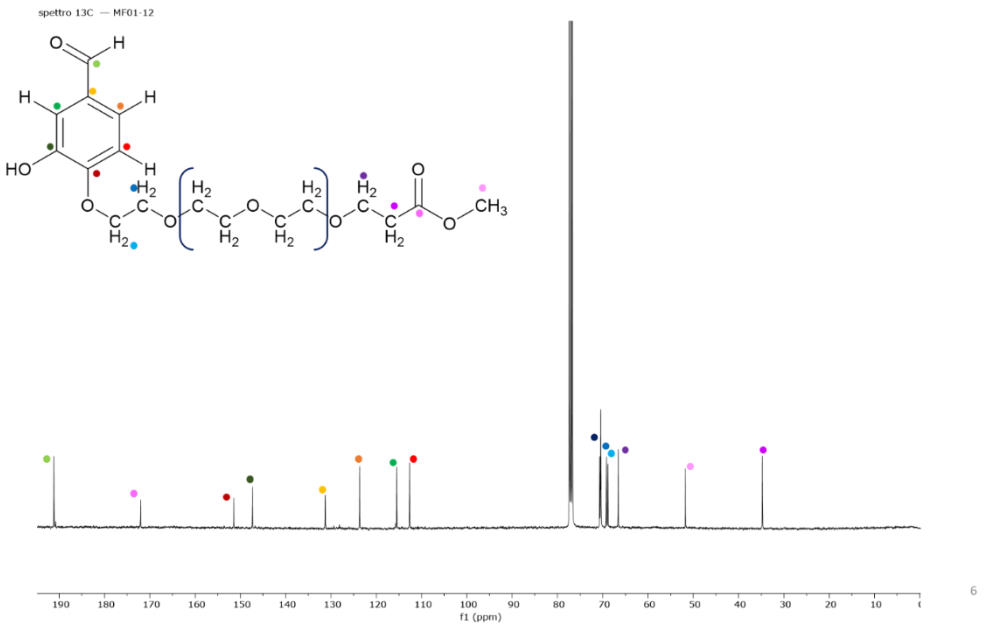



***Figure 4.A8.*** *$^{13}$C-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 2b.*

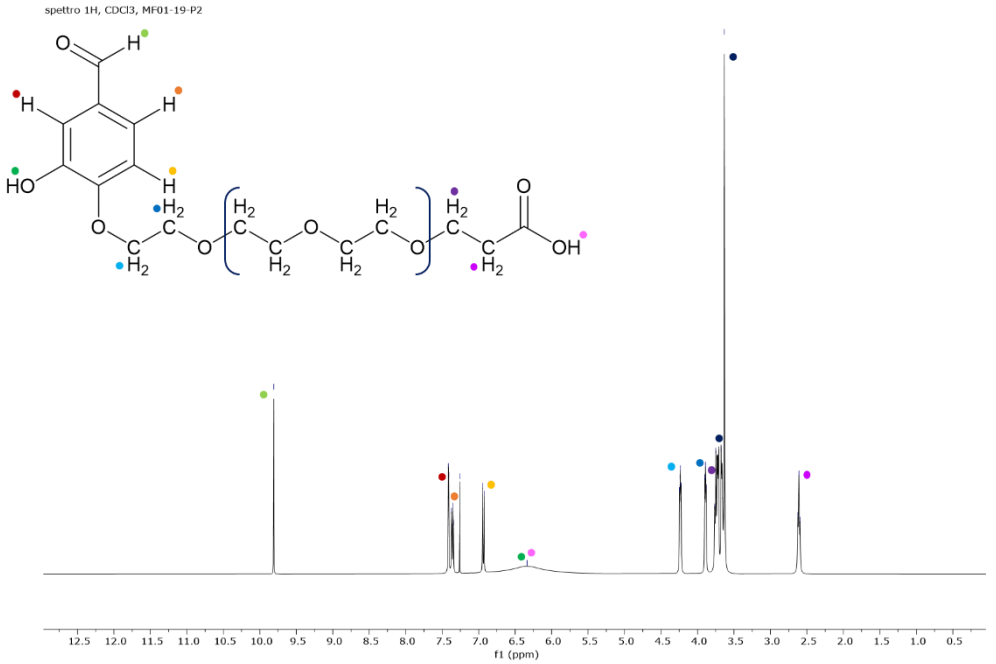



***Figure 4.A9.*** *$^{1}$H-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 3b.*

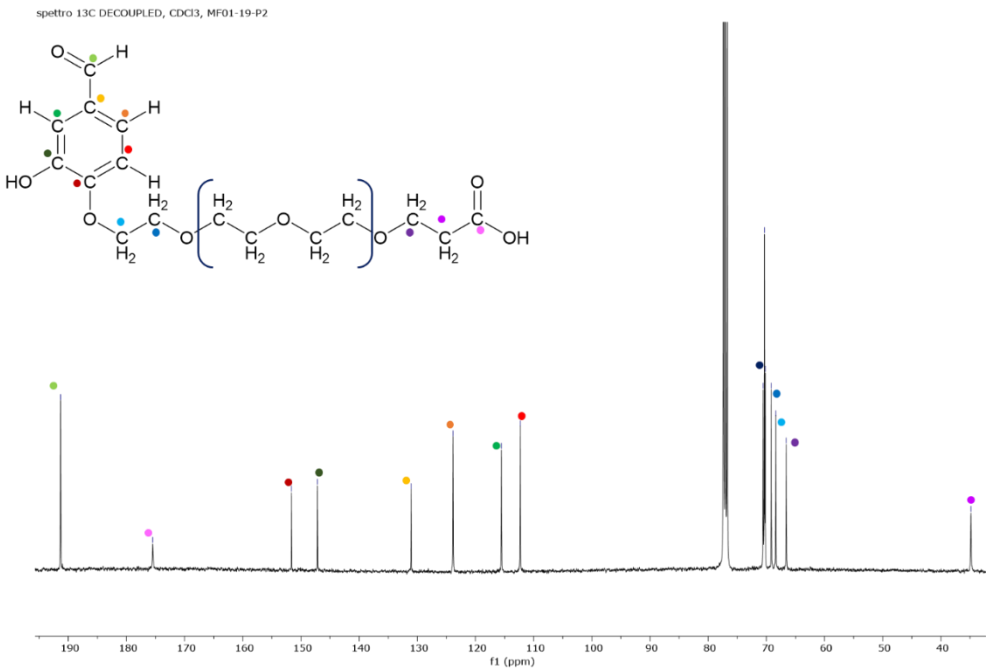



***Figure 4.A10.*** *$^{13}$C-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 3b.*

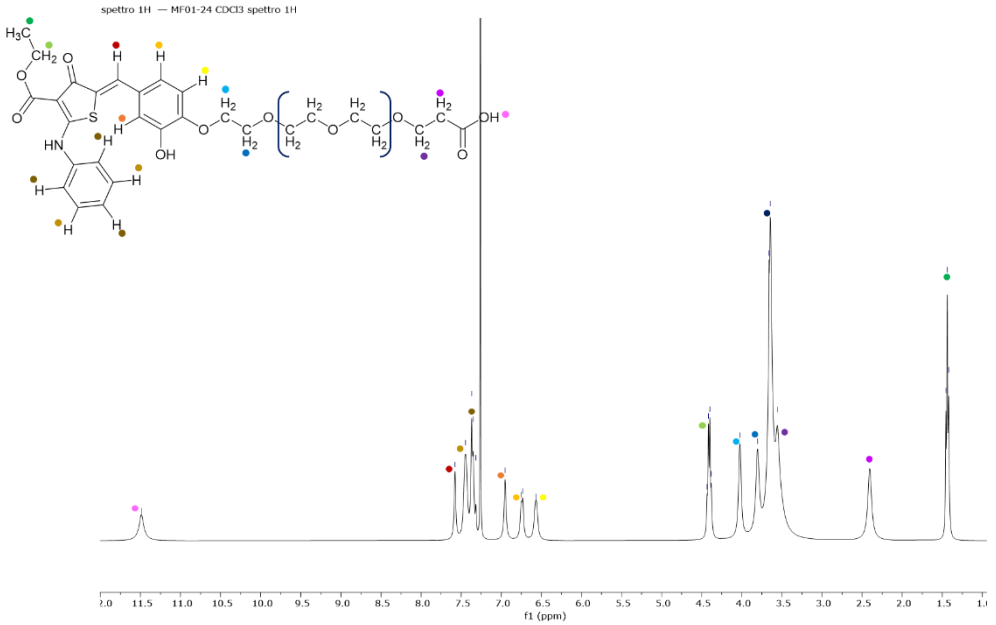



***Figure 4.A11.*** *$^{1}$H-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 4b.*

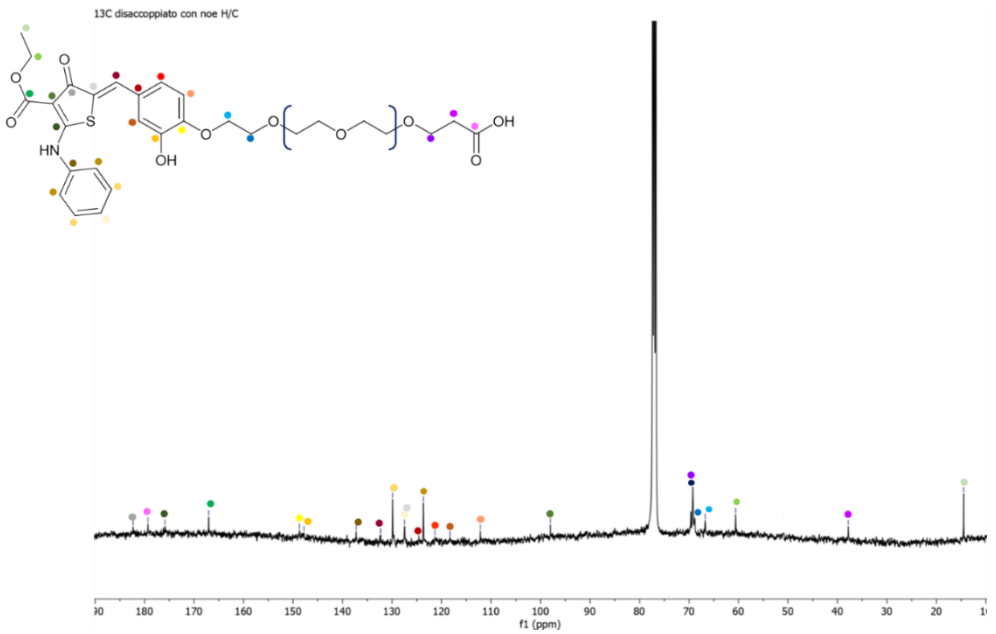



***Figure 4.A12.*** *$^{13}$C-NMR Spectrum (400 MHz, $CDCl_3$, 25 °C) of compound 3b.*

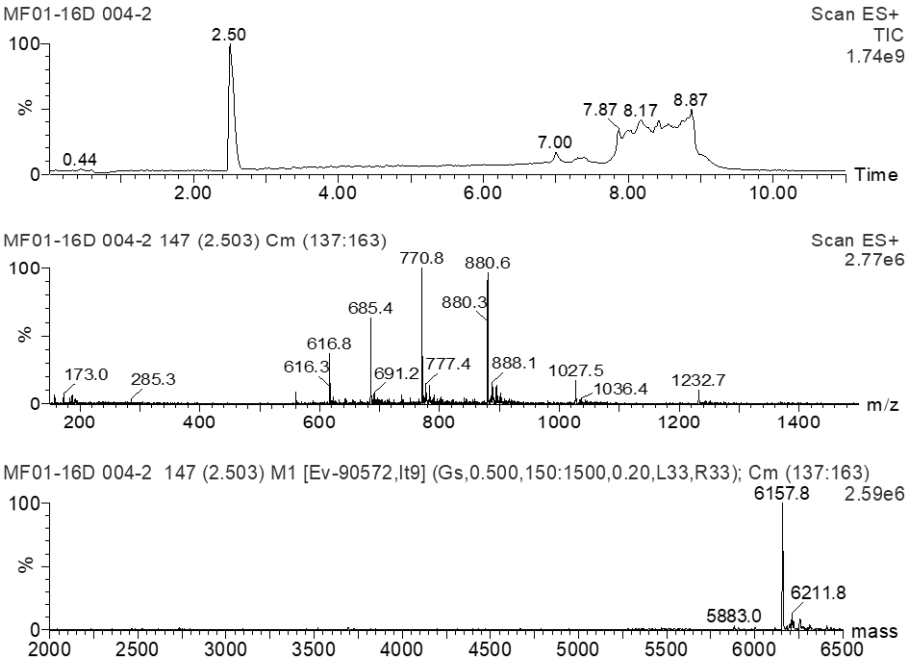



***Figure 4.A13.*** *UPLC-MS of S-P (MF01-16): chromatogram, mass spectrum and deconvoluted spectrum.*

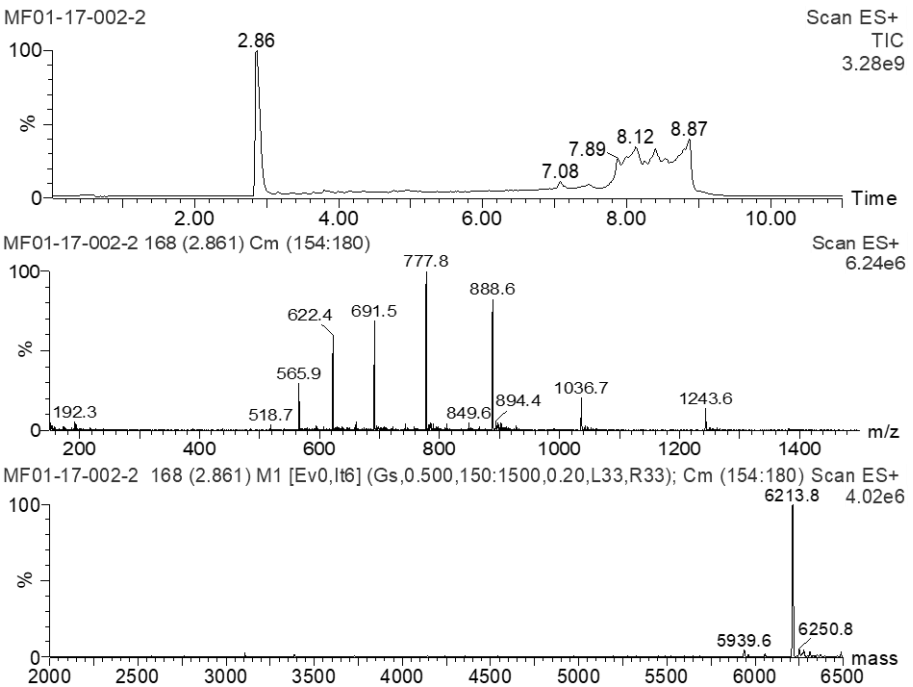



***Figure 4.A14.*** *UPLC-MS of S-St' (MF01-17): chromatogram, mass spectrum and deconvoluted spectrum.*

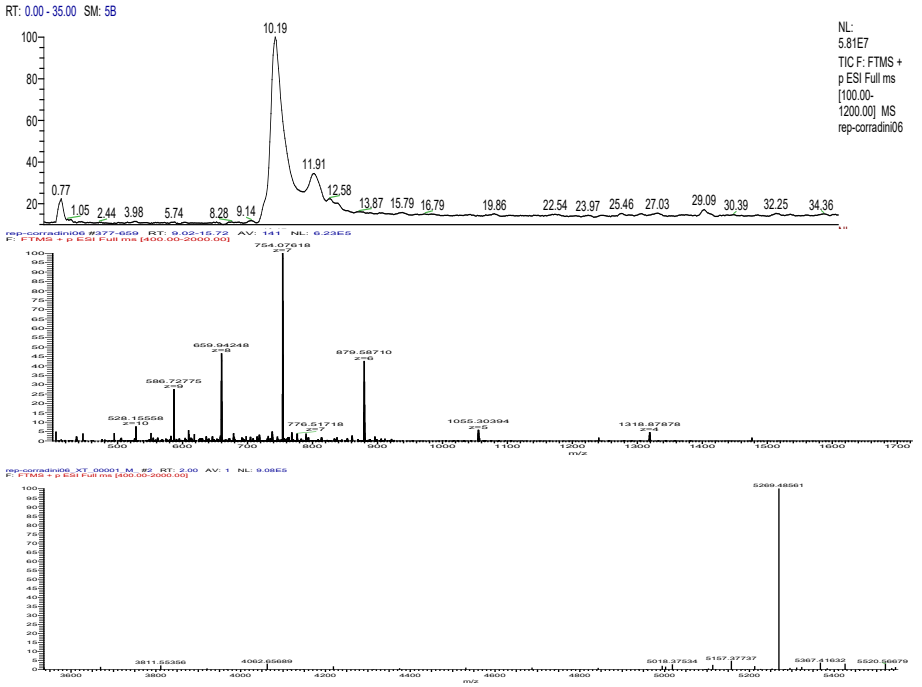



***Figure 4.A15.*** *Orbitrap of S-3' (CS02-24): chromatogram, mass spectrum and deconvoluted spectrum.*

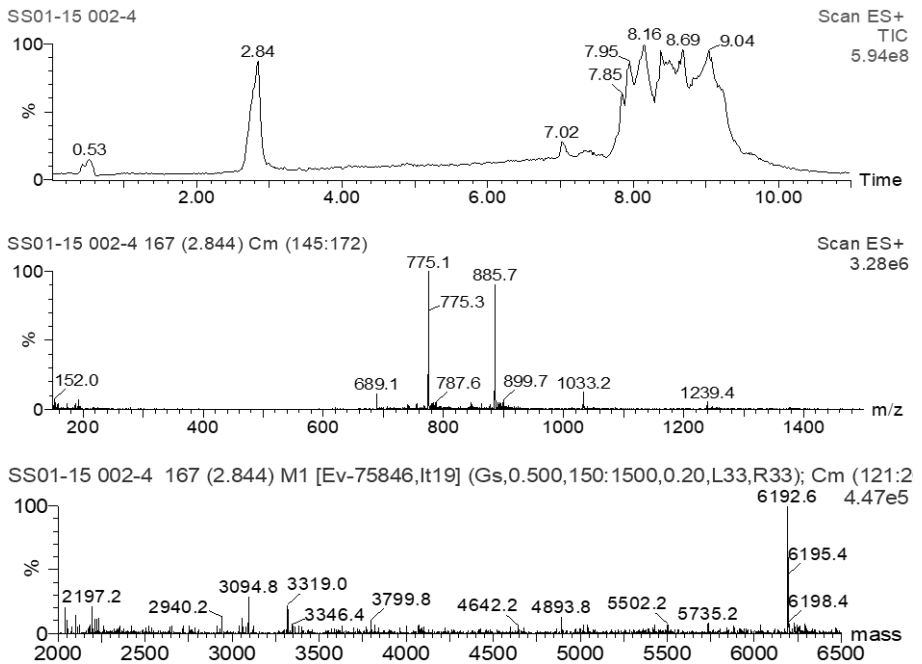



***Figure 4.A16.*** *UPLC-MS of S-5' (SS01-15): chromatogram, mass spectrum and deconvoluted spectrum.*

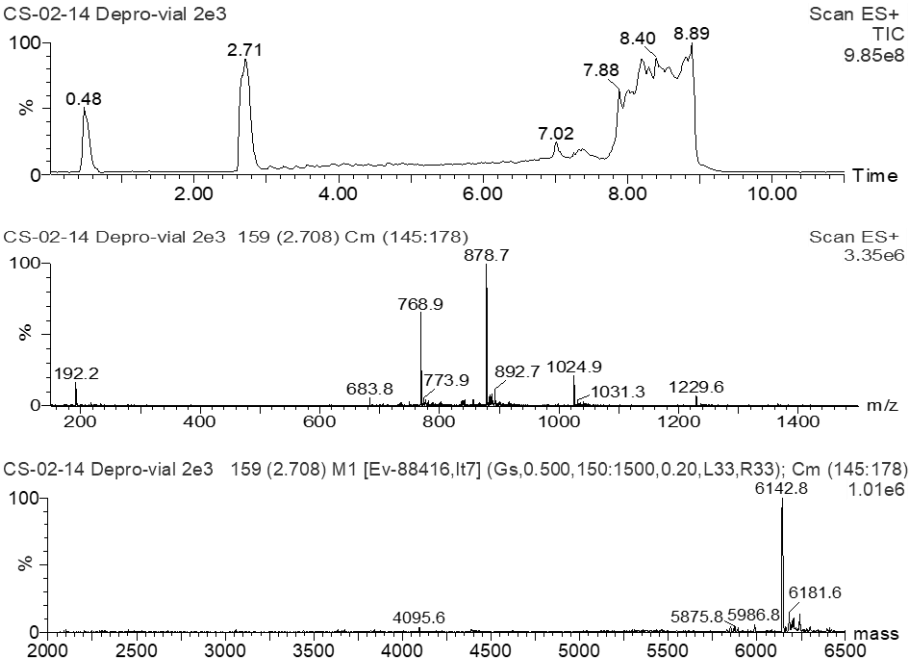



***Figure 4.A17.*** *UPLC-MS of N-P (CS02-14): chromatogram, mass spectrum and deconvoluted spectrum.*

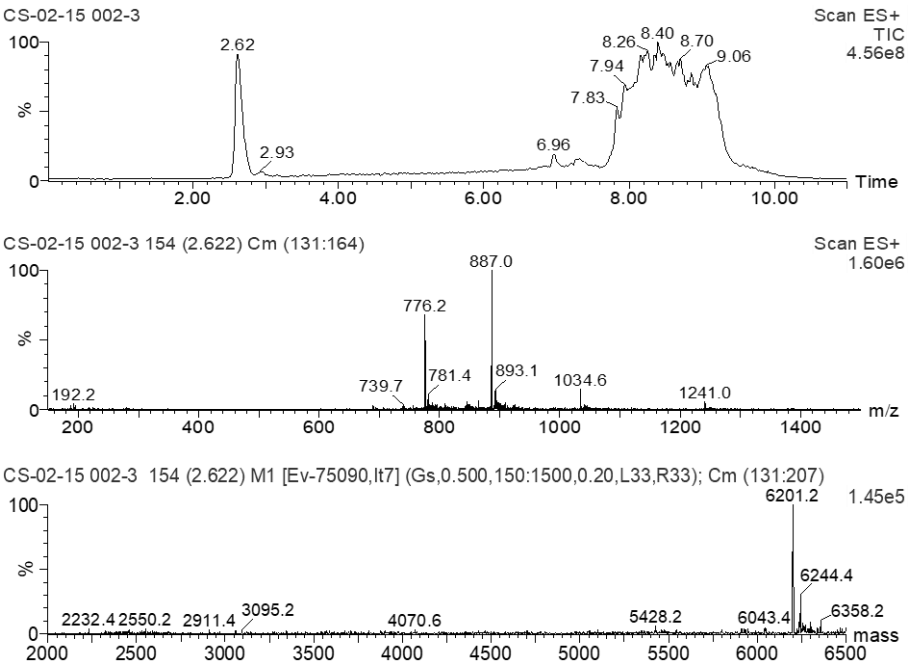



***Figure 4.A18.*** *UPLC-MS of N-St (CS02-15): chromatogram, mass spectrum and deconvoluted spectrum.*

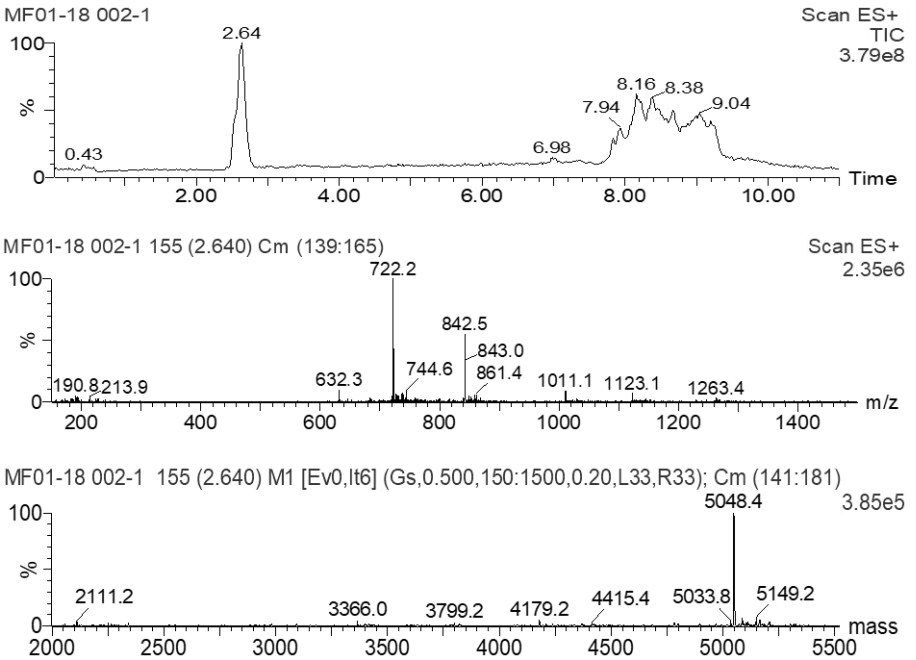



***Figure 4.A19.*** *UPLC-MS of N-3′ (MF01-18): chromatogram, mass spectrum and deconvoluted spectrum.*

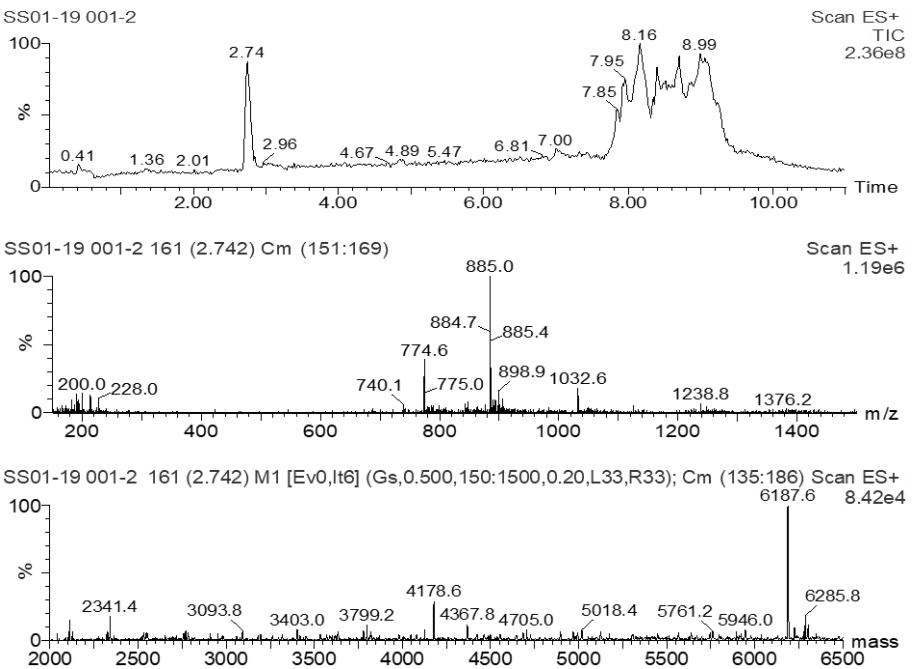



***Figure 4.A20.*** *UPLC-MS of N-5′ (SS01-19): chromatogram, mass spectrum and deconvoluted spectrum.*

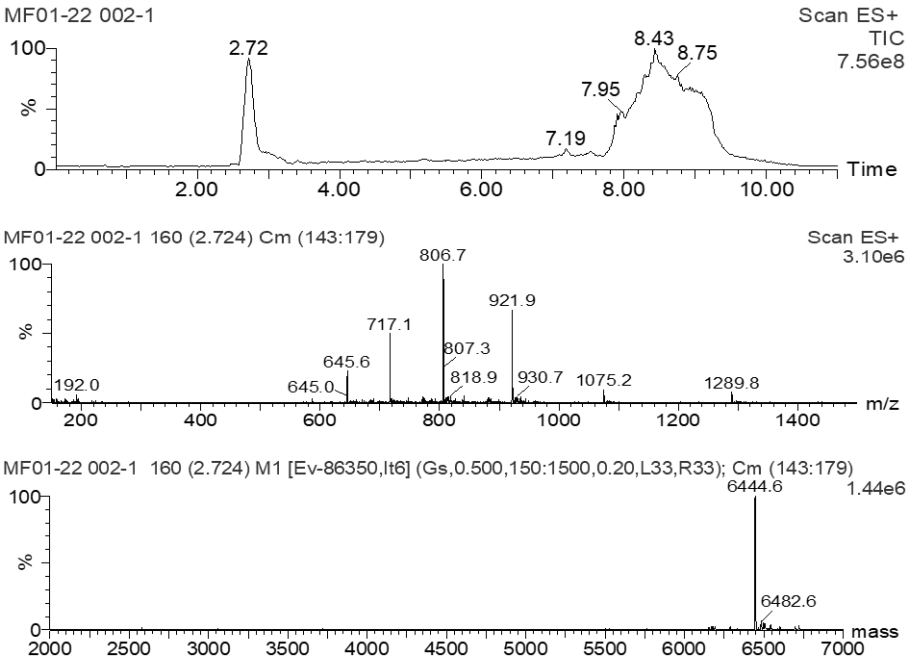



***Figure 4.A21.*** *UPLC-MS of PK (MF01-22): chromatogram, mass spectrum and deconvoluted spectrum.*

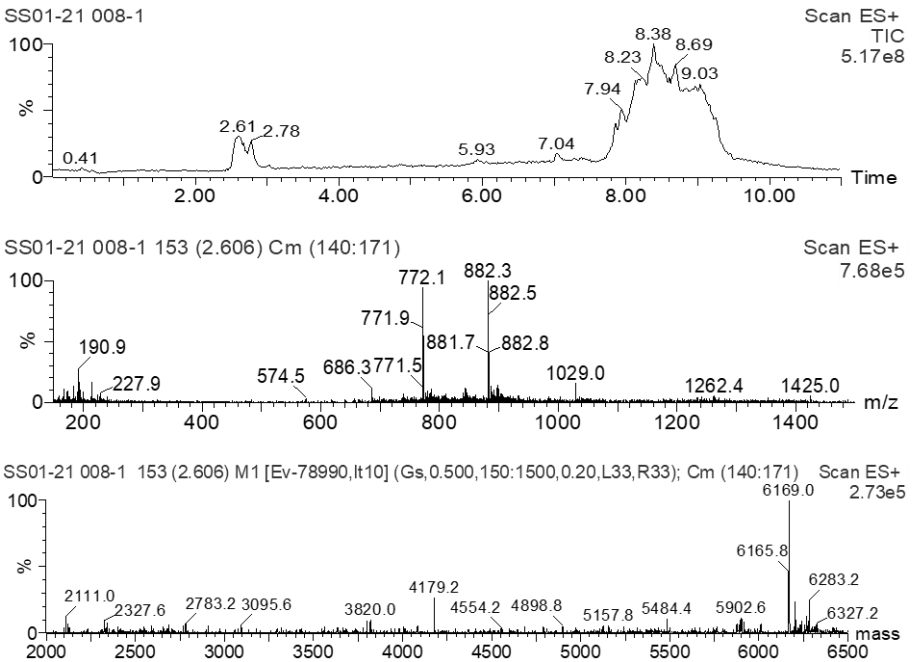



***Figure 4.A22.*** *UPLC-MS of C-18mer (SS01-21): chromatogram, mass spectrum and deconvoluted spectrum.*

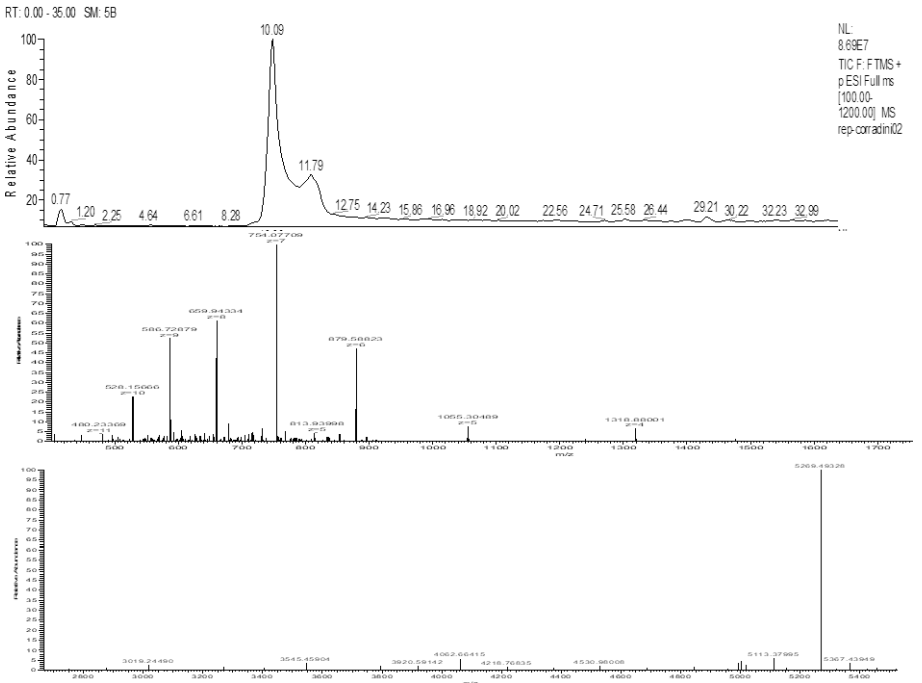



***Figure 4.A23.*** *Orbitrap of C-15mer (MF01-29): chromatogram, mass spectrum and deconvoluted spectrum.*

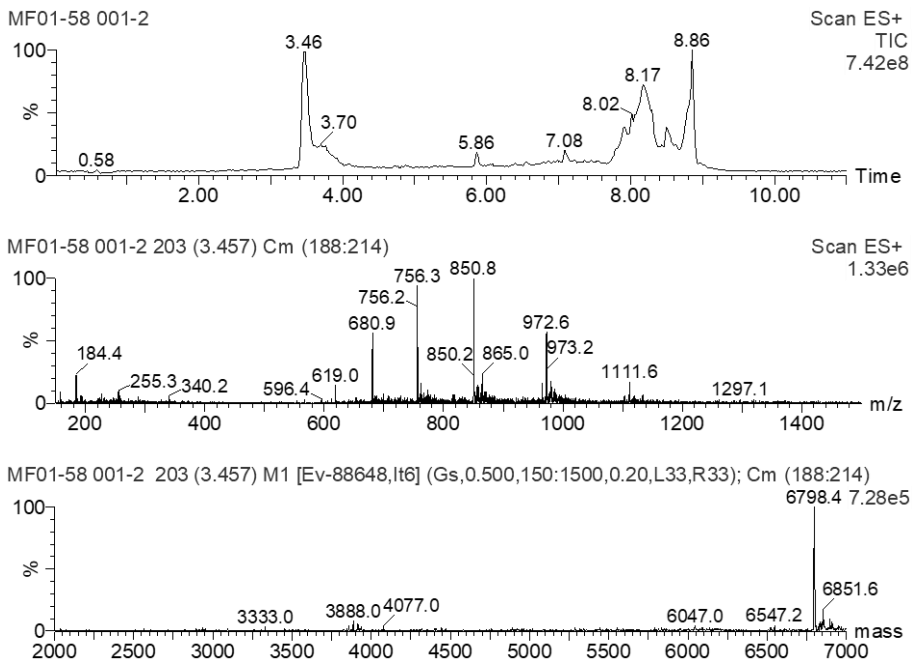



***Figure 4.A24.*** *UPLC-MS of S-P+L (MF01-58): chromatogram, mass spectrum and deconvoluted spectrum.*

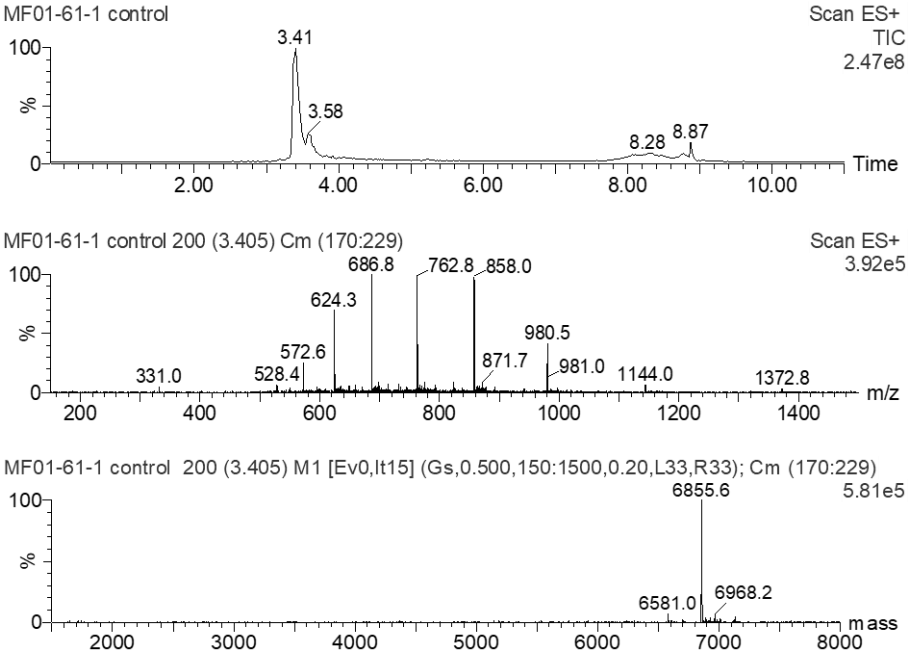



***Figure 4.A25.*** *UPLC-MS of S-St+L (MF01-61): chromatogram, mass spectrum and deconvoluted spectrum.*

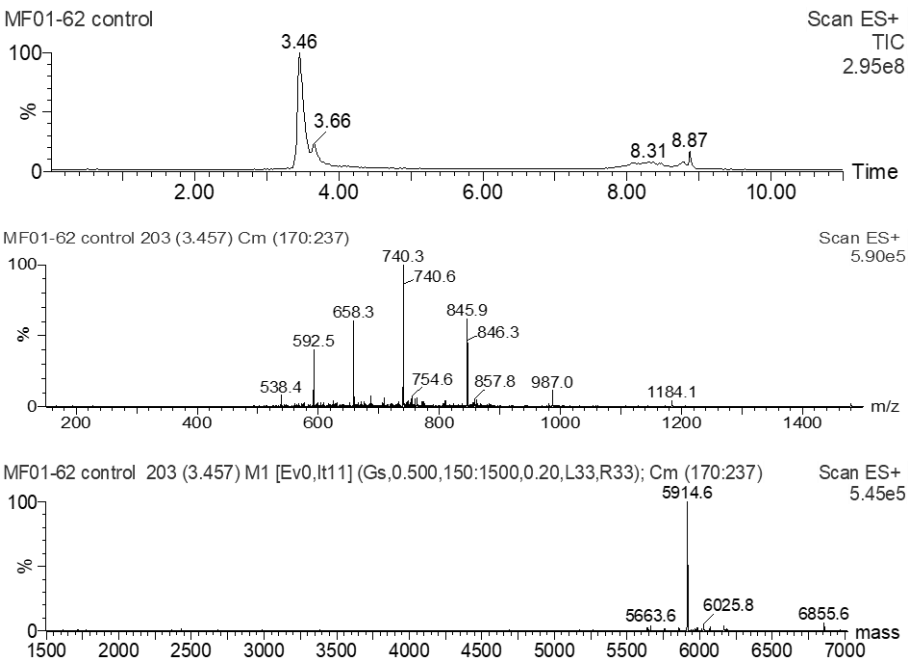



***Figure 4.A26.*** *UPLC-MS of S-3'+L (MF01-62): chromatogram, mass spectrum and deconvoluted spectrum.*

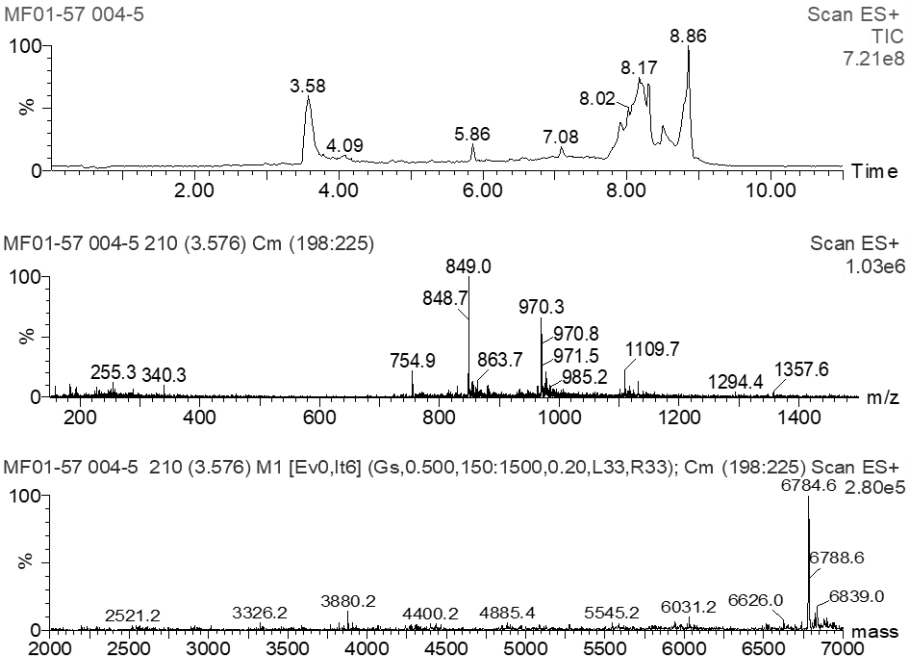



***Figure 4.A27.*** *UPLC-MS of N-P+L (MF01-57): chromatogram, mass spectrum and deconvoluted spectrum.*

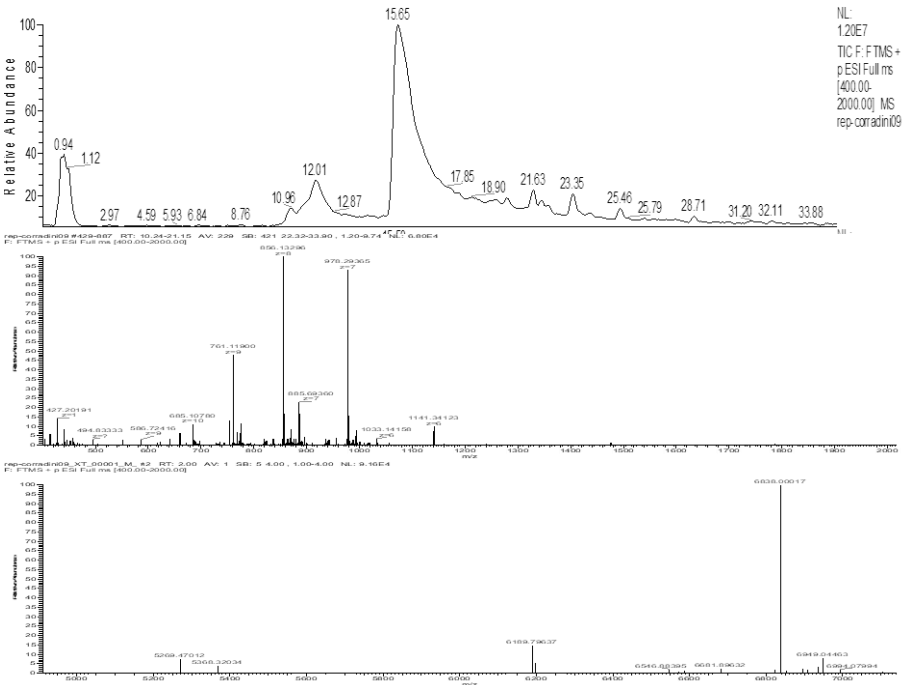



***Figure 4.A28.*** *Orbitrap of N-St+L (CS02-31): chromatogram, mass spectrum and deconvoluted spectrum.*

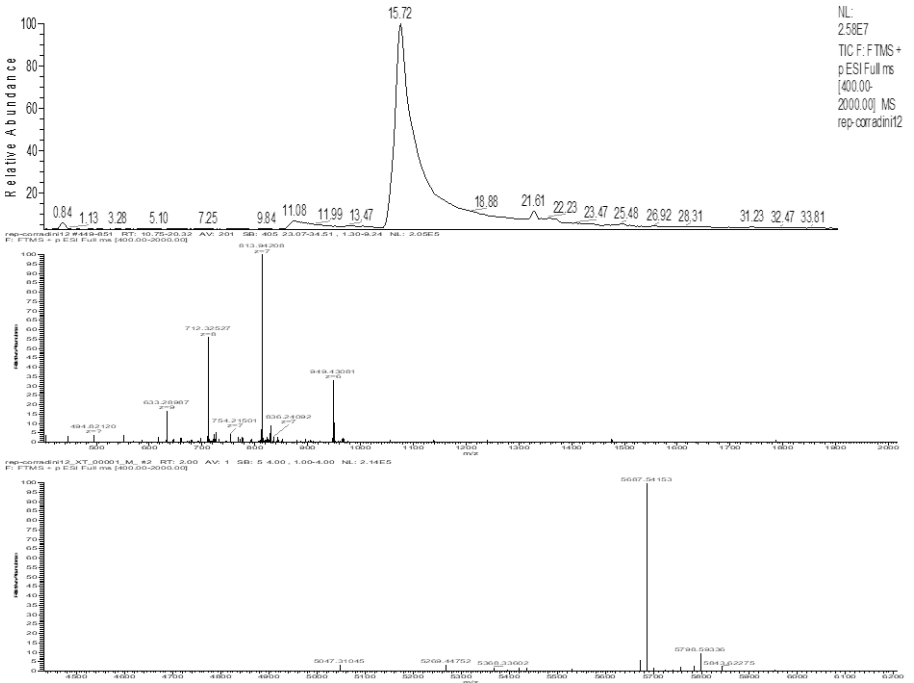



***Figure 4.A29.*** *Orbitrap of N-3'+L (CS02-32): chromatogram, mass spectrum and deconvoluted spectrum.*

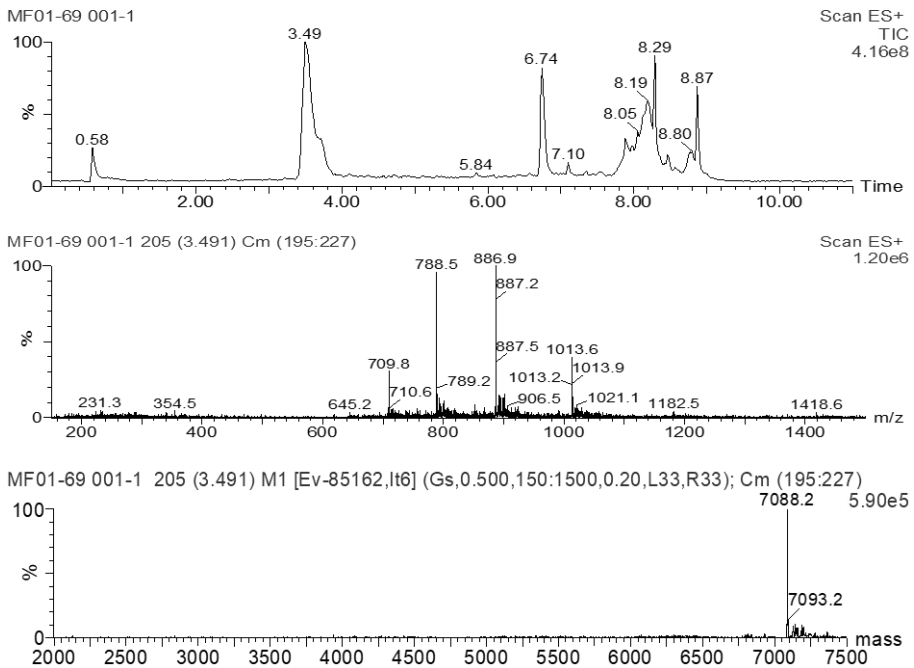



***Figure 4.A30.*** *UPLC-MS of PK+L (MF01-69): chromatogram, mass spectrum and deconvoluted spectrum.*

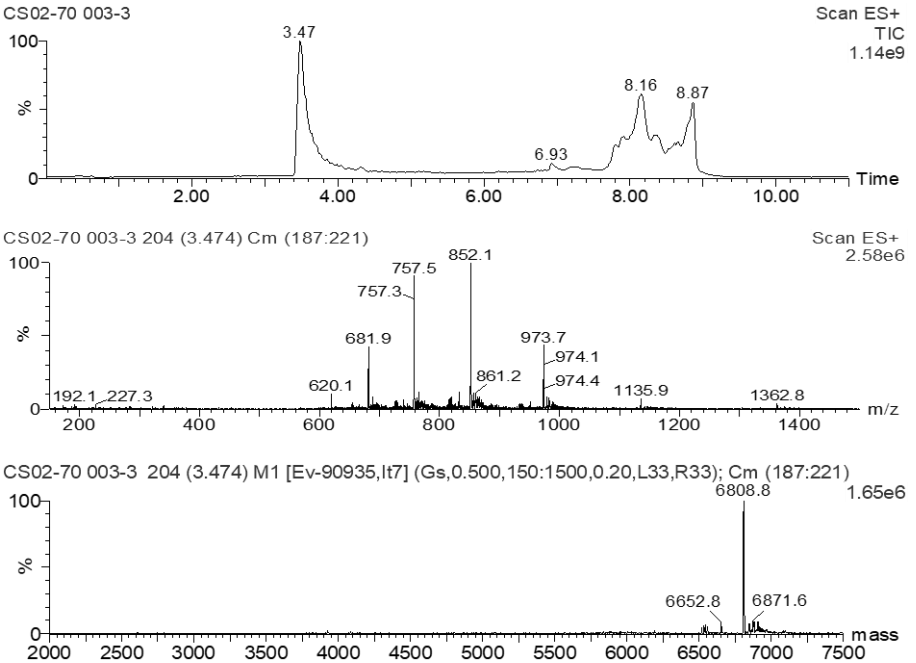



***Figure 4.A31.*** *UPLC-MS of C-18mer+L (CS02-70): chromatogram, mass spectrum and deconvoluted spectrum.*

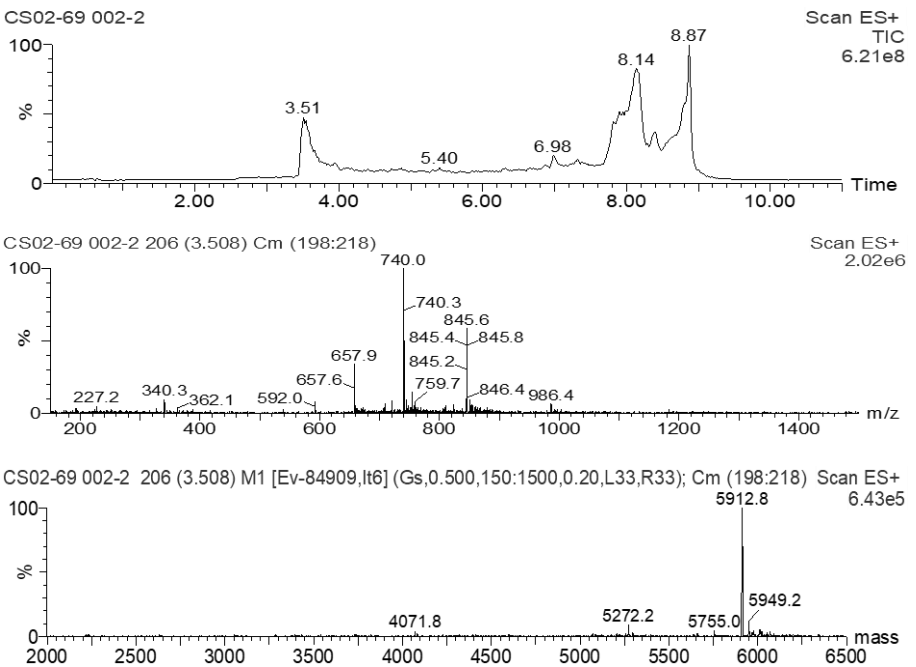



***Figure 4.A32.*** *UPLC-MS of C-15mer+L (CS02-69): chromatogram, mass spectrum and deconvoluted spectrum.*

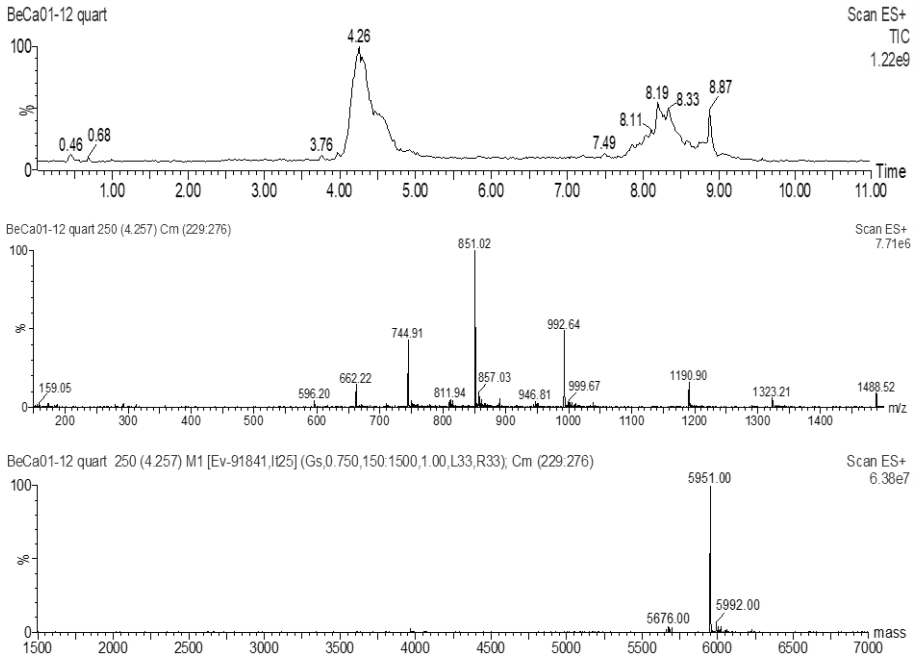



***Figure 4.A33.*** *UPLC-MS of BeCa01-12: chromatogram, mass spectrum and deconvoluted spectrum.*

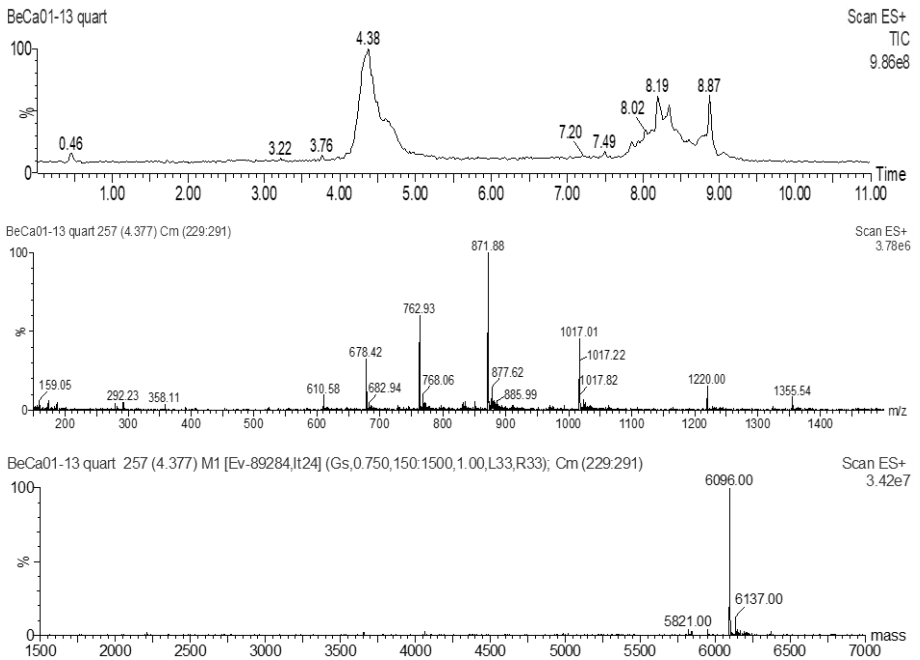



***Figure 4.A34.*** *UPLC-MS of BeCa01-13: chromatogram, mass spectrum and deconvoluted spectrum.*

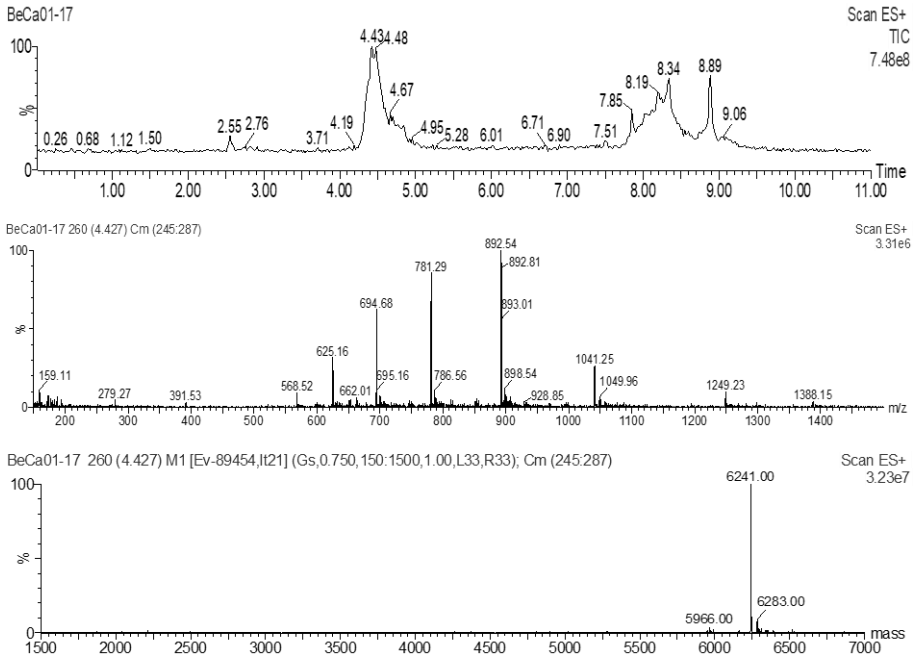



***Figure 4.A35.*** *UPLC-MS of BeCa01-17: chromatogram, mass spectrum and deconvoluted spectrum.*

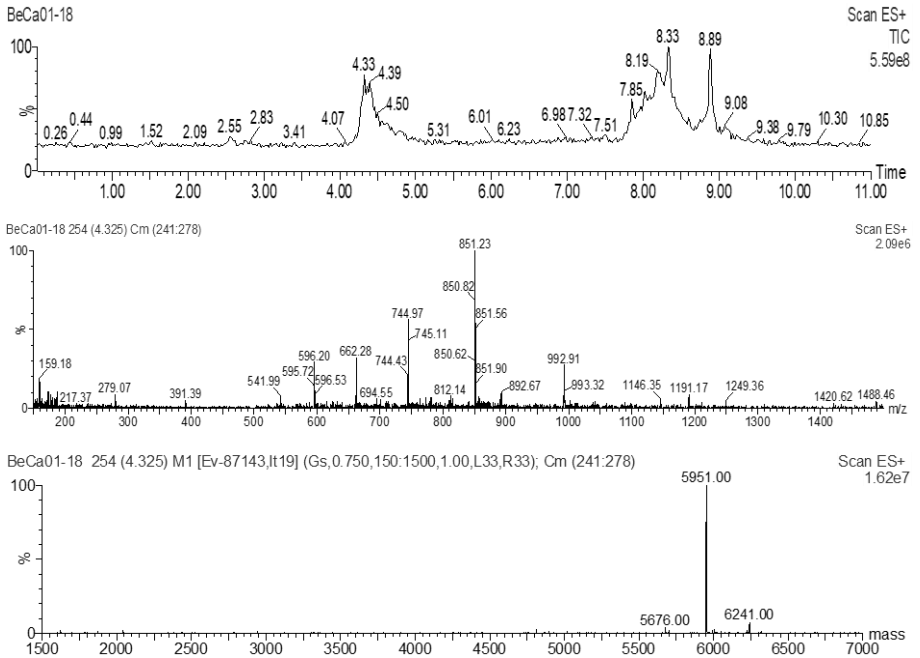



***Figure 4.A36.*** *UPLC-MS of BeCa01-18: chromatogram, mass spectrum and deconvoluted spectrum.*

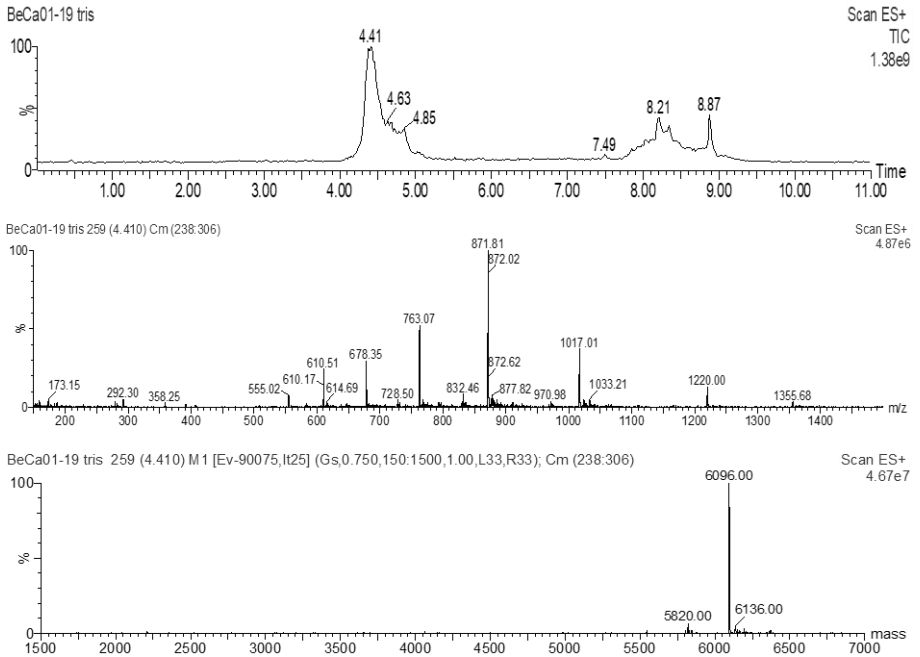



***Figure 4.A37.*** *UPLC-MS of BeCa01-19: chromatogram, mass spectrum and deconvoluted spectrum.*

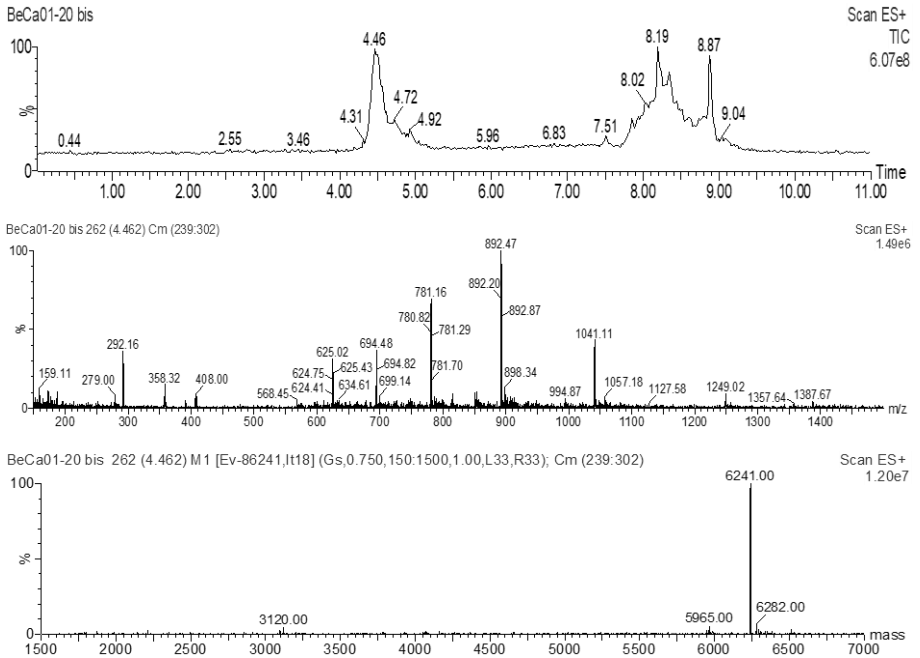



***Figure 4.A38.*** *UPLC-MS of BeCa01-20: chromatogram, mass spectrum and deconvoluted spectrum.*

Results of the analysis performed with "PNA Tool":

PNA S-P: Sequences: CAC CAG GAA CAC AAA CAT-Gly-$NH_2$; Tm at 4 µM: 80.4°C; Base composition: A=9 (50.0%), T=1 (5.6%), G=2 (11.1%), C=6 (33.3%); Purines %: 61.1%; GC Content: 44.4%; Purine Stretch: 5 (AGGAA).

PNA S-St: Sequences: AAC AAG AAA AAC AAA CAT-Gly-$NH_2$; Tm at 4 µM: 78.6°C; Base composition: A=13 (72.2%), T=1 (5.6%), G=1 (5.6%), C=3 (16.7%); Purines %: 77.8%; GC Content: 22.2%; Purine Stretch: 8 (AAGGAAAAA).

PNA S-3′: Sequences: CAC TCC ATA ACA CTT-Gly-$NH_2$; Tm at 4 µM: 64.7°C; Base composition: A=5 (33.3%), T=4 (26.7%), G=0 (0.0%), C=6 (40.0%); Purines %: 33.3%; GC Content: 40.0%; Purine Stretch: 8 (AA).

PNA S-5′: Sequences: AAA ACA AAC ATT GTT CGT-Gly-$NH_2$; Tm at 4 µM: 75.2°C; Base composition: A=8 (44.4%), T=5 (27.8%), G=2 (11.1%), C=3 (16.7%); Purines %: 55.6%; GC Content: 27.8%; Purine Stretch: 4 (AAAA); Self complementary: AA**AACAA**ACA**TTGTT**CGT, AAA**AC**A**AACA**T**TGTT**C**GT**.

PNA N-P: Sequences: TGG TCC ATT GTC ACT CAT-Gly-$NH_2$; Tm at 4 µM: 73.8°C; Base composition: A=3 (16.7%), T= 7 (38.9%), G=3 (16.7%), C=5 (27.8%); Purines %: 33.3%; GC Content: 44.4%; Purine Stretch: 2 (GG).

PNA N-St: Sequences: GGG TCC ATT ATC AGA CAT-Gly-$NH_2$; Tm at 4 µM: 75.9°C; Base composition: A=5 (27.8%), T= 5 (27.8%), G=4 (22.2%), C=4 (22.2%); Purines %: 50.0%; GC Content: 44.4%; Purine Stretch: 3 (GGG).

PNA N-3′: Sequences: TC TCC ATT CTG GTT-Gly-$NH_2$; Tm at 4 µM: 62.2°C; Base composition: A=1 (7.1%), T= 7 (50.0%), G=2 (14.3%), C=4 (28.6%); Purines %: 21.4%; GC Content: 42.9%; Purine Stretch: 2 (GG); Self complementary: TCT**CCA**TTC**TGG**TT.

PNA 5′: Sequences: ACA TTT TAG TTT GTT CGT-Gly-$NH_2$; Tm at 4 µM: 68.3°C; Base composition: A=3 (16.7%), T=10 (55.6%), G=3 (16.7%), C=2 (11.1%); Purines %: 33.3%; GC Content: 27.8%; Purine Stretch: 2 (AG); Self complementary: **ACA**TTT**TA**GTT**TGT**TCGT.

PNA PK: Sequences: A AAG CCC TGT ATA CGA CAT-Gly-$NH_2$; Tm at 4 µM: 79.1°C; Base composition: A=7 (36.8%), T=4 (21.1%), G=3 (15.8%), C=5 (26.3%); Purines %:

52.6%; GC Content: 42.1%; Purine Stretch: 4 (AAAG); Self complementary: AAAGCCCT**GTATAC**GACAT.

PNA C-18mer: Sequences: TAT CCA GTC AAG ATC TAA-Gly-$NH_2$; Tm at 4 µM: 71.9°C; Base composition: A=7 (38.9%), T=5 (27.8%), G=2 (11.1%), C=4 (22.2%); Purines %: 50.0%; GC Content: 33.3%; Purine Stretch: 4 (AAGA); Self complementary: T**ATC**CAGTCAA**GAT**CTAA, TATCCAGTCA**AGATCT**AA.

PNA C-15mer: Sequences: ACA CTC TAC ATC ACT-Gly-$NH_2$; Tm at 4 µM: 65.5°C; Base composition: A=5 (33.3%), T=4 (26.7%), G=0 (0.0%), C=6 (40.0%); Purines %: 33.3%; GC Content: 40.0%; Purine Stretch: 1 (A).

## 4.7 References


1. M.G. Costales, H. Aikawa, Y. Li, et al. Small-molecule targeted recruitment of a nuclease to cleave an oncogenic RNA in a mouse model of metastatic cancer. *Proceedings of the National Academy of Sciences* **2020**, 117 (5), 2406–2411.
2. F. Li. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annual Review of Virology*. Annual Reviews Inc. September 29, 2016, pp 237–261.
3. A.R. Fehr, S. Perlman. Coronaviruses: An overview of their replication and pathogenesis. In *Coronaviruses: Methods and Protocols*; Springer New York, **2015**; pp 1–23.
4. C.I. Paules, H.D. Marston, A.S. Fauci. Coronavirus Infections—More Than Just the Common Cold. *JAMA* **2020**, 323 (8), 707–708.
5. J.A. Kelly, M.T. Woodside, J.D. Dinman. Programmed –1 Ribosomal Frameshifting in coronaviruses: A therapeutic target. *Virology* **2021**, 554, 75–82.
6. J.S.M. Peiris, Y. Guan, K.Y. Yuen. Severe acute respiratory syndrome. *Nature Medicine*. 2004, pp S88–S97.
7. B. Hu, H. Guo, P. Zhou, Z.L. Shi. Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*. Nature Research March 1, 2021, pp 141–154.
8. P.A. Rota, M.S. Oberste, S.S. Monroe, et al. Characterization of a Novel Coronavirus Associated with Severe Acute Respiratory Syndrome.
9. A. Wu, Y. Peng, B. Huang, et al. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host and Microbe*. Cell Press March 11, 2020, pp 325–328.
10. D.G. Ahn, W. Lee, J.K. Choi, et al. Interference of ribosomal frameshifting by antisense peptide nucleic acids suppresses SARS coronavirus replication. *Antiviral Res* **2011**, 91 (1), 1–10.
11. J.D. Dinman, M.J. Ruiz-Echevarria, K. Czaplinski, S.W. Peltz. Peptidyl-transferase inhibitors have antiviral properties by altering programmed –1 ribosomal frameshifting efficiencies: Development of model systems. *Proceedings of the National Academy of Sciences* **1997**, 94 (13), 6606–6611.
12. T. Goss Kinzy, J.W. Harger, A. Carr-Schmid, et al. New Targets for Antivirals: The Ribosomal A-Site and the Factors That Interact with It. *Virology* **2002**, 300 (1), 60–70.
13. S.-J. Park, Y.-G. Kim, H.-J. Park. Identification of RNA Pseudoknot-Binding Ligand That Inhibits the –1 Ribosomal Frameshifting of SARS-Coronavirus by Structure-Based Virtual Screening. *J Am Chem Soc* **2011**, 133 (26), 10094–10100.
14. J.A. Kelly, A.N. Olson, K. Neupane, et al. Structural and functional conservation of the programmed –1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2). *Journal of Biological Chemistry* **2020**, 295 (31), 10741–10748.
15. P.R. Bhatt, A. Scaiola, G. Loughran, et al. Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome.
16. H.S. Haniff, Y. Tong, X. Liu, et al. Targeting the SARS-COV-2 RNA genome with small molecule binders and ribonuclease targeting chimera (RiboTAC) degraders. *ACS Cent Sci* **2020**, 6 (10), 1713–1721.

17. D. Miteva, M. Kitanova, H. Batselova, et al. The End or a New Era of Development of SARS-CoV-2 Virus: Genetic Variants Responsible for Severe COVID-19 and Clinical Efficacy of the Most Commonly Used Vaccines in Clinical Practice. *Vaccines*. Multidisciplinary Digital Publishing Institute (MDPI) July 1, 2023.
18. B.A. Babalola, O.S. Akinsuyi, E.O. Folajimi, et al. Exploring the future of SARS-CoV-2 treatment after the first two years of the pandemic: A comparative study of alternative therapeutics. *Biomedicine and Pharmacotherapy*. Elsevier Masson s.r.l. September 1, 2023.
19. P.E. Nielsen. Peptide Nucleic Acids (PNA) in Chemical Biology and Drug Discovery. *Chem Biodivers* **2010,** 7 (4), 786–804.
20. J. Saarbach, P.M. Sabale, N. Winssinger. Peptide nucleic acid (PNA) and its applications in chemical biology, diagnostics, and therapeutics. *Curr Opin Chem Biol* **2019,** 52, 112–124.

# Chapter 5

## *5.0 Design, assembly, and characterization of hybrid DNA/PNA nanospheres*

### 5.1 Abstract

In this chapter, I reported the work done in a project previously started in our laboratory by Dr. Stefano Volpi. This project was carried out in the framework of a European RISE project (Nano Oligomed) in collaboration with the group of Prof. Hanadi Sleiman of Mc Gill University.

Deepening the study regarding the design, assembly, and characterization of hybrid DNA:PNA nanostructures, the aim of this project was the production of new DNA:PNA hybrid nanostructures. Given that, the addition of PNAs in this assembly allows to change the DNA nanosphere under different aspects: a) it stabilizes the DNA from enzymatic cleavage; b) the PNA can rigidify the terminal part of the nanospheres by forming a PNA:DNA duplex; c) PNA conjugated with different peptide or ligand units can potentially be used to *decorate* the self-assembly, thus changing their fate in biological systems. The production of DNA micelles decorated with complementary PNA strands can be obtained by a self-assembly process (*Figure 5.1*) and the use of PNA-peptide conjugates can potentially modify their properties.

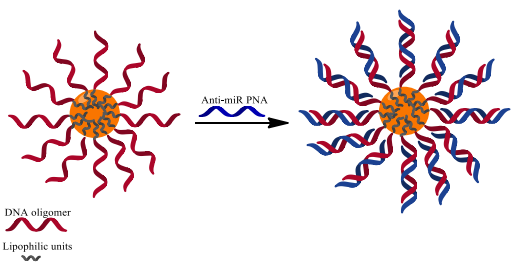



***Figure 5.1.*** *Structure of hybrid PNA:DNA micelles.*

My work was aimed at studying the behavior of these PNA:DNA nanostructure as a function of the overall charge introduced by the PNA-conjugates.

## 5.2 Introduction

DNA nanotechnology capitalizes on the reliable and controllable self-assembly of DNA oligonucleotides, enabling the construction of nanoscale structures with well-defined shapes and dimensions.[1,2] These systems fall into three primary categories, determined by the methodologies employed for constructing their architectures: DNA origami, tile-based DNA assembly, and DNA minimal assembly.

In the last thirty years, extensive research has been conducted on DNA origami and tile-based assembly techniques. These approaches depend on numerous single-stranded DNA sequences to craft intricate nanostructures, which can encompass openable boxes, molecular walkers, and logic gates.[3,4] It is important to acknowledge that designing and synthesizing the essential building blocks for these methods can be both expensive and time-consuming. Conversely, the minimal assembly strategy represents a newer approach that uses fewer DNA strands to produce more straightforward and stylized structures in a cost-efficient and less labor-intensive manner. Despite its simplicity, this method retains the ability to generate precisely defined 3D structures, such as prisms, cubes, and wireframe fibers. These structures can demonstrate dynamic behaviors when subjected to specific external stimuli.[2,5] Simplified structures can also be achieved by substituting the conventional Watson-Crick base-pairing with alternative self-assembly methods. As an illustration, amphiphilic DNA oligomers containing polyhexaethylenephosphate components can create well-dispersed micelles primarily through hydrophobic interactions in the presence of an aqueous buffer.[6] This phenomenon can be utilized in conjunction with DNA base-pairing to elevate the intricacy of previously assembled structures by introducing nonpolar polymers at specific sites, leading to hierarchical assemblies.[7] A particularly fascinating feature of minimal DNA structures is the retention of substantial single-stranded sections following assembly. This characteristic offers the potential to tailor and enhance these structures with complementary DNA derivatives through post-synthetic modifications. Such modifications have facilitated non-covalent bonding with

crucial biomolecules,[8,9] integration of antisense oligonucleotides,[10–14] encapsulation of gold nanoparticles,[15–17] incorporation of anticancer drugs,[18] and the creation of more complex structures.[7,19] As a result, these structures present promising platforms for applications in bioimaging, biosensing, and drug delivery.

DNA micelles have proven to be effective structures for delivering antisense oligonucleotides, especially when combined with appropriate transfection agents.[10,12,13] Recent research by the Sleiman group[6], which focused mainly on DNA micelles, indicates that these systems exhibit a capacity to penetrate cell membranes to some extent, making them more viable for *in-vivo* applications. As a result, DNA minimal structures are now being considered as promising options for creating a novel class of multivalent, biocompatible platforms designed for transporting PNAs inside cells, ultimately boosting their antisense capabilities.

Furthermore, the attachment of PNA derivatives has been found to enhance the overall robustness of the resulting hybrid nanostructures, as DNA/PNA duplexes are recognized for their remarkable stability in biologically relevant environments.[20] While DNA minimal assemblies exhibit considerable inherent resilience, surpassing that of their individual components, their resistance to nucleases remains somewhat constrained.[21] Complementary PNAs are expected to function as protective agents against this category of enzymes.

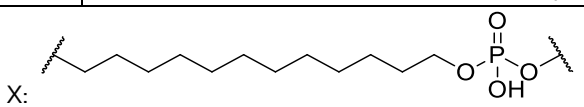
## 5.3 Result and discussion

The DNA derivatives to produce new DNA:PNA hybrid nanostructures was performed by Dr. Stefano Volpi, during a secondment period in the Sleiman's group, by linking lipophilic oligomers with DNA segments, and the optimization of conditions for the formation of hybrid PNA:DNA nanostructures is reported in a paper in preparation. The PNA used in this study were chosen among those able to target microRNA sequences, most notably tumor-related miR 21 and miR-221, with the aim of testing a possible use of the nanosphere for PNA delivery. In fact, these systems can penetrate the cell membranes independently and are thus promising to create a novel class of multivalent, biocompatible platforms drug-delivery.

The effect of modification of the overall charge was tested using an octa-arginine conjugated PNA.

The DNA and PNA strands used for the formation of the nanospheres are shown in *Table 5.1*.

***Table 5.1.*** *DNA and PNA strands.*

| Entry | Sigla | Sequence |
|---|---|---|
| 1 | $(C_{12})_{12}$-21 | XXXXXXXXXXXX TTTTT TTA TCA GAC TGA TGT TGA |
| 2 | $(C_{12})_{12}$-21-3' | TTA TCA GAC TGA TGT TGA TTTTT XXXXXXXXXXXX |
| 3 | PNA-21-Fl | Fl-TCA ACA TCA GTC TGA TAA-Gly-$NH_2$ |
| 4 | PNA-21-R8-Fl | Fl-R8-TCA ACA TCA GTC TGA TAA-Gly-$NH_2$ |
| 5 | X:  | |

The PNAs sequences were synthesized by solid-phase peptide synthesis (SPPS) and labelled with 5-(6)-carboxyfluorescein (Fl) to effectively demonstrate the hybridization with target DNA strands. These compounds were 18-mer oligonucleotides with sequence complementary to miR-21-5p. This microRNA is commonly implicated in gene expression alterations and is overexpressed in several tumors.[22] The design of amphiphilic DNA derivatives drew inspiration from literature.[6] Their polar component comprised a 23-mer DNA strand, which was

divided into a pentathymine spacer and an 18-mer segment with a sequence resembling miR21 to ensure full complementarity with the PNA oligomers. The nonpolar portion of these derivatives consisted of twelve consecutive hexaethylenephosphate moieties (X).

From these building blocks, DNA:PNA nanospheres were formed by mixing the amphiphilic DNA ($(C_{12})_{12}$-21) with different quantity of PNAs (PNA-21-Fl or PNA-21-R8-Fl) in water solution. The preparations of these structures were performed at room temperature or by annealing at 90 °C and cooling to 4 °C. The formation of the spherical micellar nanostructures was followed by native Agarose Gel Electrophoresis analysis (1.5% AGE) or by Dynamic Light Scattering (DLS) characterization.

### 5.3.1 Native Agarose Gel Electrophoresis analysis

The addition of uncharged PNA-21-Fl to $(C_{12})_{12}$-21 DNA, as shown in *Figure 5.2a*, allows the formation of a hybrid structure of PNA:DNA, identified as single band with lower electrophoretic mobility than $(C_{12})_{12}$-21 DNA. The incorporation of PNA into the nanospheres can be demonstrated by the presence of fluorescein emission. The same band could be stained using DNA-specific Gel-Red dye.

This band is formed already in the presence of 0.5 eq of PNA with respect of the target DNA. AFM measurements of these systems, carried out at Mc Gill University, showed the nanometric dimensions of these particles ( *Figure 5.2b*).

Thermal annealing allowed to obtain a more defined band, and hence a more homogeneous structure.

Furthermore, the stability of the PNA-capped DNA nanospheres to serum enzymes was found to be much higher than that of the DNA nanospheres alone (results not shown), confirming that the PNA part could protect the nanosphere from degradation.

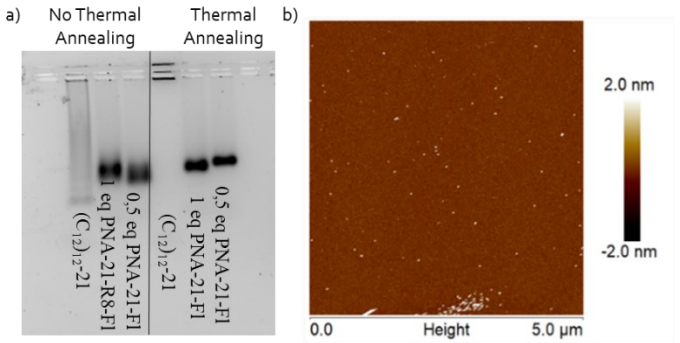



***Figure 5.2.*** *a) 2.5% Native agarose gel electrophoresis analysis of nanospheres formed by $(C_{12})_{12}$-21 as DNA and 1 eq or 0.5 eq of PNA-21-Fl, b) AFM image a 5 mL x 200 nM sample containing an equimolar amount of (C12)12-21-short and PNA-21-Fl which was not submitted to thermal annealing.*

The addition of PNA-21-R8-Fl gave rise to more complex behavior, leading to the disruption of the assembly with formation of larger aggregates ( *Figure 5.3*, bands in the gel pits).

In a first series of test, a screening of different AGE percentage was performed, using samples prepared by different equivalents of PNA-21-R8-Fl and 5 µM of DNA.

The results, reported in *Figure 5.3*, showed that the band relative to the nanospheres is not visible for all the three agarose percentages tested, in specific 2.5%, 1.5%, 0.5% agarose (denoted as 2.5%, 1.5%, and 0.5% AGE), in the presence of the cationic PNA-21-R8-Fl.

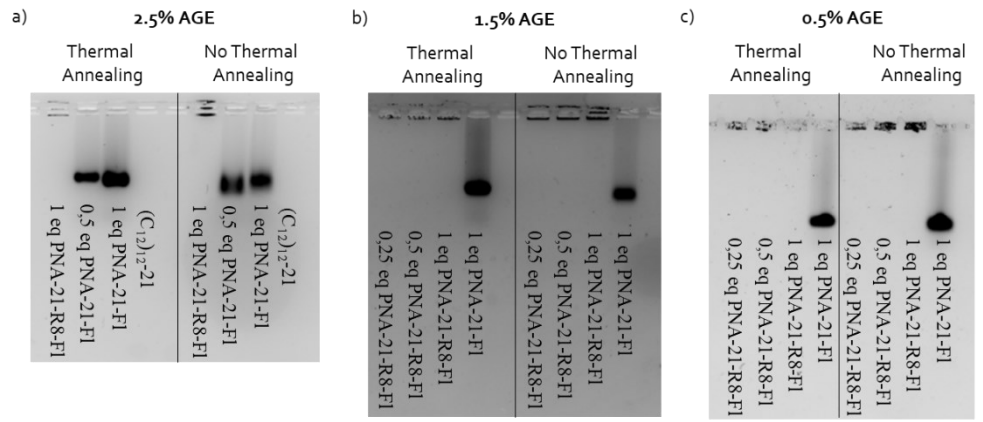



***Figure 5.3.*** *a) 2.5%, b) 1.5%, and c) 0.5% Native agarose gel electrophoresis analysis of nanospheres formed by $(C_{12})_{12}$-21 as DNA and different equivalents of PNA-21-R8-Fl.*

Since the presence of positively-charged groups was found to disrupt the nanosphere structure, we addressed the problem preparing several samples at different percentage of PNA-21-Fl and PNA-21-R8-Fl, constant overall concentration (5 μM), and DNA at 5 μM concentration. The results, reported in *Figure 5.4*, showed that the band relative to the nanospheres is not visible in the presence of 100% cationic PNA-21-R8-Fl, but the samples with a mixture of neutral PNA-21-Fl and a fraction of PNA-21-R8-Fl also showed this band. By this screening, it is also possible to observe that increasing the concentration of PNA with eight arginine, the nanospheres are still formed until a 50:50 respectively of PNA-21-Fl and PNA-21-R8-Fl.

Additionally, the results highlight the importance of subjecting the samples to a thermal treatment (annealing at 90°C and cooling to 4°C) to achieve better formation of these PNA:DNA nanostructures also when using cationic PNAs.

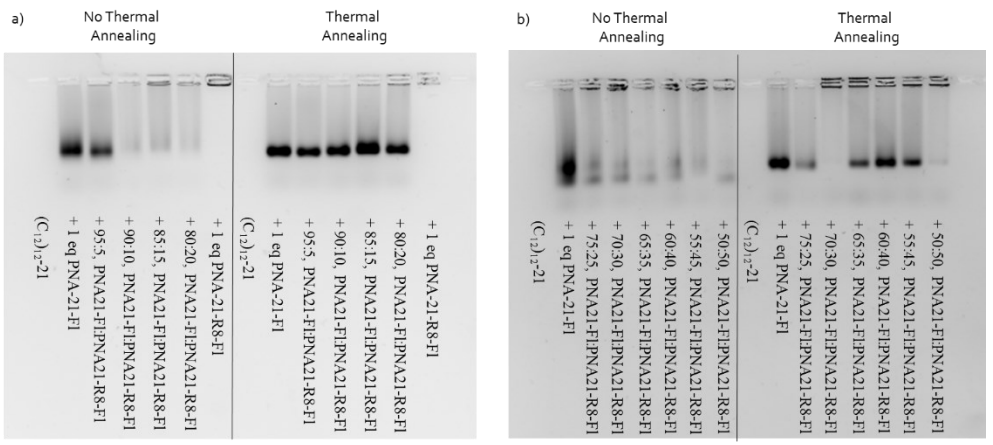



***Figure 5.4.*** *1.5% Native agarose gel electrophoresis analysis of micelles formed by $(C_{12})_{12}$-21 as DNA and different percentage of PNA-21-Fl and PNA-21-R8-Fl, with or without thermal annealing.*

Similar results were obtained with the $(C_{12})_{12}$-21-3' as DNA and the respectively PNA ( *Figure 5.5*).

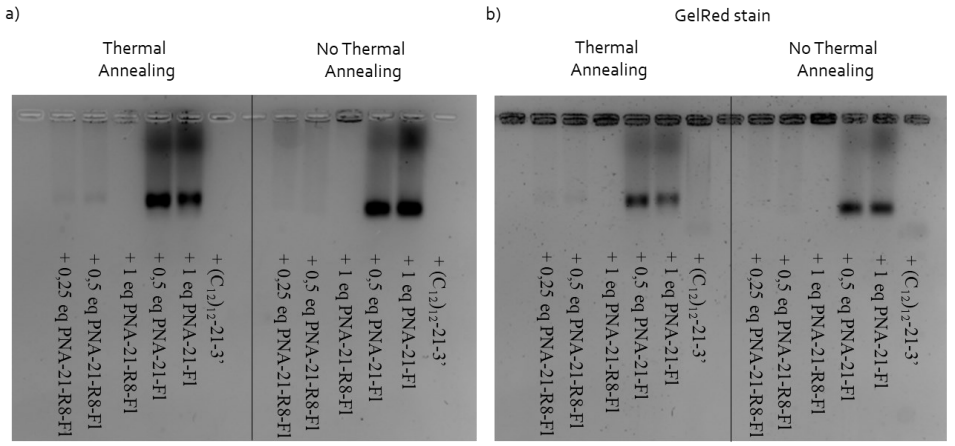



***Figure 5.5.*** *1.5% Native agarose gel electrophoresis analysis of micelles formed by (C12)12-21-3' as DNA and PNA-21-Fl or PNA-21-R8-Fl with or without thermal annealing. a) Fluorescein channel; b) GelRed stain.*

## 5.3.2 Dynamic Light Scattering (DLS) characterization

The same samples were characterized by DLS, to evaluate the dimension of the particles and gain further evidence on the presence of higher aggregates. Without thermal annealing, the DLS experiments (*Figure 5.6*) gave a hydrodynamic radius (Z-average) of 168 nm for the solution containing $(C_{12})_{12}$-21 alone, 75 nm for DNA/1 eq of PNA-21-Fl and 1197 nm for DNA/1 eq of PNA-21-R8-Fl, thus confirming that the presence of the cationic PNA could destabilize the charge balance of these systems, leading to larger type of aggregates, consistently with the agarose gel electrophoretic data.

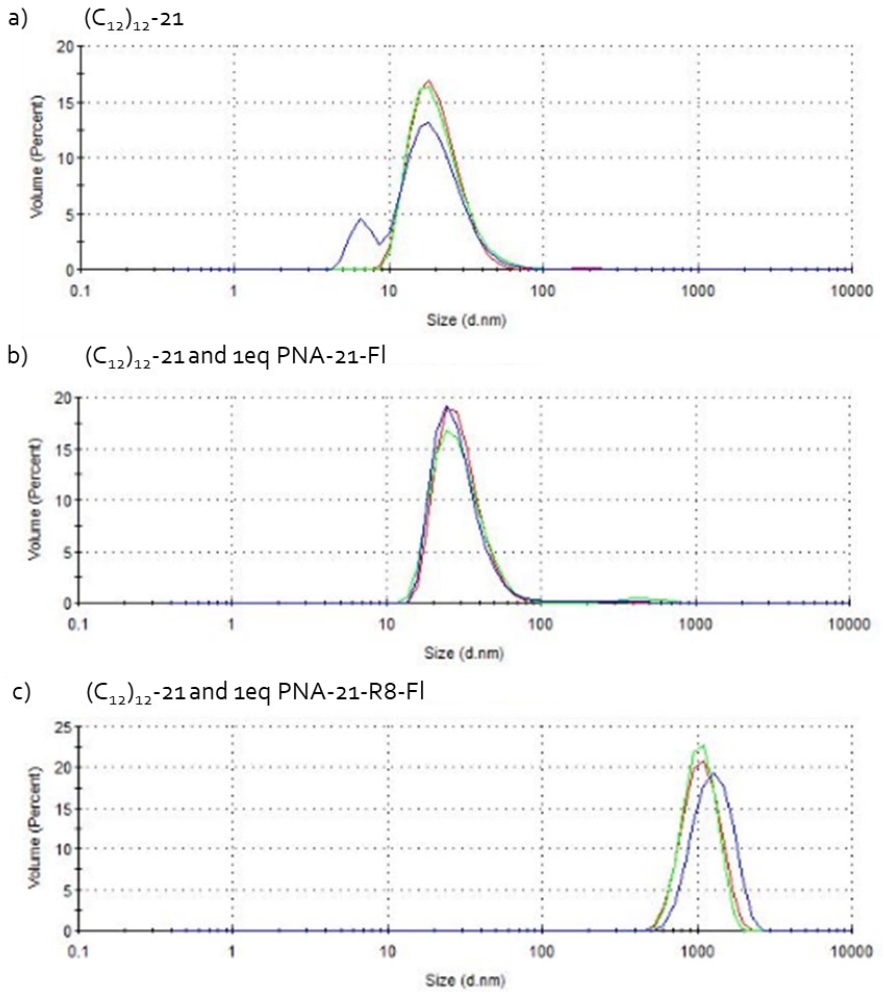



***Figure 5.6.*** *Size distribution by volume of a) $(C_{12})_{12}$-21, b) $(C_{12})_{12}$-21 and 1 eq of PNA-21-Fl, c) $(C_{12})_{12}$-21 and 1 eq of PNA-21-R8-Fl. Different colors refer to different samples measured (measurements in triplicate).*

## 5.4 Conclusions and future directions

Although PNA-loaded DNA nanospheres are interesting nano-objects in which the PNA can be used as protecting agent against enzymatic cleavage of DNA, or as a further cargo to be delivered to cells, the addition of charges on the PNA chain can have dramatic effects on the nanosphere stability, as revealed by the case reported here of polyarginine-conjugated PNA and inferred from both electrophoresis and DLS measurements.

Indeed, upon binding of cationic PNAs, the balance between attractive forces and repulsive electrostatic interactions that is necessary to maintain the spherical and determines the size of the particle is disrupted. Furthermore, the repulsion between nanoparticles is lowered, thus favoring the formation of larger aggregates. The same does not occur with the neutral PNA, which could be bound to the DNA nanospheres with only a reduction of the particle size.

The presence of cationic PNAs on the DNA nanospheres can only be limited to a fraction of all the binding sites available, with a limit of 50% found in the present case. However, it would be interesting to follow the behavior and biological fate of these nanoparticles as a function of the overall particle charge in the available 0-50% range.

## 5.5 Experimental section

### 5.5.1 Reagent and instrumentation

The reagents were bought from BLD Pharmatech, Merck, abcr, Carlo Erba, VWR, Fluka, TCI Europe, Link Technologies, PolyOrg and used without additional purification. Only for the solid phase synthesis, the DMF was purged with nitrogen flux to prevent the formation of dimethylamine and was dried over 4Å molecular sieves.

The two PNA sequences are synthesized by standard solid-phase manual synthesis or by automatic synthesizer Biotage Syro I. Purification was performed by HPLC (Agilent Technologies 1260 Infinity I) using a SepaChrom Vydamas® (C18, 5 µm, 300 Å, 10 x 250 mm) column. Gradient: 100% A for 1 min, then from 0% to 50% B for 31 min at 4 mL/min flow (A: water + 0.1% trifluoroacetic acid; B: acetonitrile + 0.1% trifluoroacetic acid). Detector UV set with wavelength: 260 nm.

After purification, the peptide nucleic acid sequences were characterized by UPLC-MS (Waters Acquity Ultra Performance LC) using the following instrumental set-up: Waters Acquity ultra-performance LC E07SQD086W, with Waters SQ detector and ESI-interface equipped with Acquity UPLC BEH 300 (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic condition: eluent A: water + 0.2% formic acid; eluent B: acetonitrile + 0.2% formic acid. Column temperature: 35 °C. Program: initial isocratic at 100% A (0.9 min), then linear gradient to 50% B (in 5.7 min). Final wash with 100% B for 1.2 min. Flow rate: 0.25 mL/min. Some PNAs, however, were characterized with Thermo LTQ Orbitrap XL detector and ESI-interface equipped with a Phenomenex Kinetex EVO (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic conditions: eluent A: water + 0.1% formic acid; eluent B: acetonitrile + 0.1% formic acid. Column temperature: 35 °C. Program: initial isocratic at 98% A (3 min), then linear gradient to 50% B (in 20 min) and then to 95% (in 1 min). Final wash with 98% B for 8 min. Flow rate: 0.20 mL/min. UV wavelength: 260 nm.

The concentration of the PNAs was calculated using Evolution 260 Bio UV-Visible spectrophotometer with Peltier thermos tatting accessories (Thermo Fisher

Scientific SPE8W), following the UV-absorbance at 260 nm and assuming an additive contribution of all bases.

For Native Agarose Gel Electrophoresis analysis: microwave for preparation of agarose gel solution and ChemiDoc MP Imaging System by Bio-Rad Laboratories S.R.L. for gel imaging.

### 5.5.2 Synthesis of peptide nucleic acids

The two PNAs were synthetized with standard solid-phase manual synthesis, using tert-butyloxycarbonyl (Boc) strategy. The PNA monomers employed to Boc strategy are Boc/Cbz (benzyloxycarbonyl) protected.

The HMBHA Rink Amide resin was loaded with Fmoc-Gly-OH to obtain 0.2 mmol/gr as theoretical loading. The principal steps to load the resin were: a) swelling in DCM (1x1h), b) DMF dry wash, c) coupling (1x5h, activation for 10', using 1 eq of Fmoc-Gly-OH, 10 eq of DIC and 10 eq of DhBtOH in DMF dry as activation solution), d) DMF wash, e) capping (2x15', DMF:$Ac_2O$, 1:1), e) DMF wash and f) DCM wash. To measure the loading of the resin a UV/Visible technique was used, following the absorbance of divenzofulvene (DBF) formation after deprotection treatment (30', Piperidine:DMF, 1:4) at 290 nm. In fact, the values at 290 nm and 400 nm (to subtract the baseline contribution) were extracted from the measure. The loading of the resin was calculated dividing the value obtained from the difference between the absorbance at 290 nm and at 400 nm with a value obtained by multiplication of a conversion factor (1.65 g/mmol*mg) with the weighed resin mass.

The principal steps and solutions for the standard solid-phase manual synthesis, using Boc strategy, were: a) swelling in DCM (1x30'), b) deprotection (2x4', TFA/m-cresol, 95:5), c) DCM wash, d) DMF dry wash, e) Kaiser test (1', positive), f) coupling (1x30', activation for 2', using 5 eq of PNA monomer, 4.9 eq of HBTU as coupling reagent and 10 eq of DIPEA in DMF dry as activation solution), g) DMF wash, h) Kaiser test (1', negative), i) capping (2x1', $Ac_2O$:py:NMP dry, 1:25:25), l) DMF wash, m) 5% Piperidine wash (2x2', Piperidine:DMF, 1:9), n) DMF wash and o) DCM wash. The solution formed by 5-(6)-carboxyfluorescein (5 eq), HBTU (4.9 eq), and DIPEA

(10 eq) was used to perform the coupling with the fluorescein. Instead for the polyarginine moiety, a solution composed by Fmoc-Arg(Pbf)-OH (5 eq), HBTU (4.9 eq), and DIPEA (10 eq) was employed.

The solutions and procedures used for the Kaiser test were as described in *Chapter 2*.

For the cleavage step a cocktail solution composed of TFA:TFMA:m-cresol:thioanisole (6:2:1:1) was used and two cycles of 1h were performed to make sure a complete cleavage. After the first cycle the solution was filtered and collect in a falcon. Then, the resin was wash with only TFA to ensure a complete collection of the product, and the same thing was repeated at the end of the second cycle. The PNA oligomers were precipitated in diethyl ether in freezer for at least 2h. After removing diethyl ether and take drying the product, the latter was dissolved in double distilled water.

The two PNAs were purified by HPLC and characterized by UPLC-MS and UV-Visible techniques, as described in *Chapter 2*.

### 5.5.3 Native Agarose gel electrophoresis experiments

2.5%, 1.5%, and 0.5% native agarose gel electrophoresis analysis were carried out to study the formation of nanospheres. TAMg buffer was used, and in specific TAMg buffer was formed by 45 mM Tris and 12.5 mM $MgCl_2 \bullet 6H_2O$ with a pH of 8.0 adjusted with glacial acetic acid. All the samples were prepared in water and submitted to thermal annealing with a temperature ramp from 90°C to 4°C in 1h.

### *5.5.4 DLS experiments*

DLS experiments were carried out with DNA and PNA concentrations of 5 μM, both with and without the octa-arginine moiety, in a 10 mM Tris-Acetate-Magnesium (TAMg) buffer at pH 8.0, and at a temperature of 25°C.

## 5.5.3 PNAs characterizations

Entry 1 ( Figure 5.A1 in Appendix)
PNA-21-Fl: Fl-TCA ACA TCA GTC TGA TAA-Gly
yield: 22%; $t_r$: 3.52 min; ε (260 nm): 227065 $M^{-1}cm^{-1}$; MW calculated: 5278.12 [M], observed (from ESI-MS deconvolution): 5278.0; ESI-MS: m/z observed (calcd): 1320.7 (1320.5) $[M+4H]^{4+}$, 1056.8 (1056.6) $[M+5H]^{5+}$, 880.8 (880.7) $[M+6H]^{6+}$, 755.2 (755.0) $[M+7H]^{7+}$, 660.9 (660.7) $[M+8H]^{8+}$, 587.6 (587.4) $[M+9H]^{9+}$.

Entry 2 ( Figure 5.A2 in Appendix)
PNA-21-R8-Fl: Fl-R8-TCA ACA TCA GTC TGA TAA-Gly
yield: 8%; $t_r$: 2.67 min; ε (260 nm): 227065 $M^{-1}cm^{-1}$; MW calculated: 6527.64 [M], observed (from ESI-MS deconvolution): 6527.0; ESI-MS: m/z observed (calcd): 1088.9 (1089.1) $[M+6H]^{6+}$, 933.3 (933.5) $[M+7H]^{7+}$, 816.8 (816.9) $[M+8H]^{8+}$, 726.2 (726.3) $[M+9H]^{9+}$.

## 5.6 Appendix Chapter 5

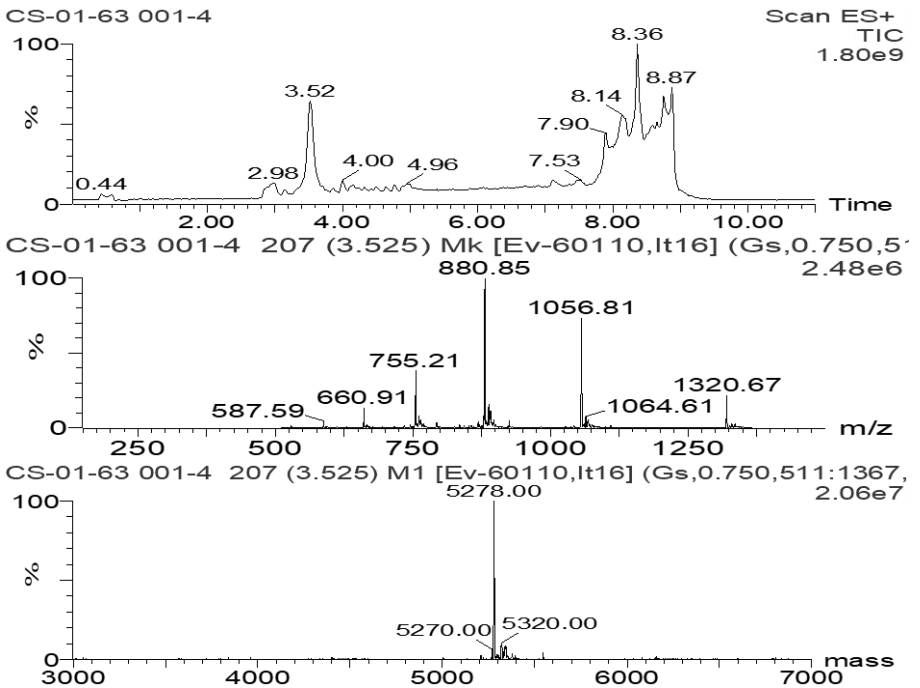



***Figure 5.A1.*** *UPLC-MS of PNA-21-Fl: chromatogram, mass spectrum and deconvoluted spectrum.*

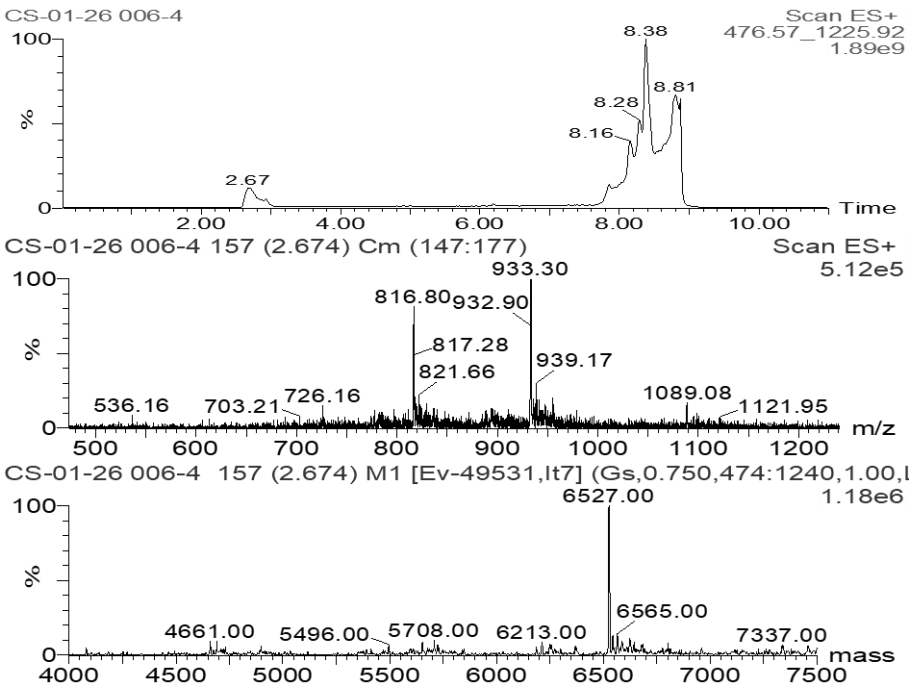



***Figure 5.A2.*** *UPLC-MS of PNA-21-R8-Fl: chromatogram, mass spectrum and deconvoluted spectrum.*

## 5.7 References


1. J.D. WATSON, F.H.C. CRICK. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, 171 (4356), 737–738.
2. N.C. Seeman, H.F. Sleiman. DNA nanotechnology. *Nat Rev Mater* **2017**, 3 (1), 17068.
3. F. Hong, F. Zhang, Y. Liu, H. Yan. DNA Origami: Scaffolds for Creating Higher Order Structures. *Chem Rev* **2017**, 117 (20), 12584–12640.
4. D.Y. Zhang, G. Seelig. Dynamic DNA nanotechnology using strand-displacement reactions. *Nat Chem* **2011**, 3 (2), 103–113.
5. P. Chidchob, H.F. Sleiman. Recent advances in DNA nanotechnology. *Curr Opin Chem Biol* **2018**, 46, 63–70.
6. T.G.W. Edwardson, K.M.M. Carneiro, C.J. Serpell, H.F. Sleiman. An efficient and modular route to sequence-defined polymers appended to DNA. *Angewandte Chemie - International Edition* **2014**, 53 (18), 4567–4571.
7. C.J. Serpell, T.G.W. Edwardson, P. Chidchob, K.M.M. Carneiro, H.F. Sleiman. Precision polymers and 3D DNA nanostructures: Emergent assemblies from new parameter space. *J Am Chem Soc* **2014**, 136 (44), 15767–15774.
8. P. Chidchob, D. Offenbartl-Stiegert, D. McCarthy, et al. Spatial Presentation of Cholesterol Units on a DNA Cube as a Determinant of Membrane Protein-Mimicking Functions. *J Am Chem Soc* **2019**, 141 (2), 1100–1108.
9. A. Lacroix, T.G.W. Edwardson, M.A. Hancock, M.D. Dore, H.F. Sleiman. Development of DNA Nanostructures for High-Affinity Binding to Human Serum Albumin. *J Am Chem Soc* **2017**, 139 (21), 7355–7362.
10. M.D. Dore, J.J. Fakhoury, A. Lacroix, H.F. Sleiman. Templated synthesis of spherical RNA nanoparticles with gene silencing activity. *Chemical Communications* **2018**, 54 (80), 11296–11299.
11. J.W. Conway, C. Madwar, T.G. Edwardson, et al. Dynamic Behavior of DNA Cages Anchored on Spherically Supported Lipid Bilayers. *J Am Chem Soc* **2014**, 136 (37), 12987–12997.
12. K.E. Bujold, J. Fakhoury, T.G.W. Edwardson, et al. Sequence-responsive unzipping DNA cubes with tunable cellular uptake profiles. *Chem Sci* **2014**, 5 (6), 2449–2455.
13. J.J. Fakhoury, T.G. Edwardson, J.W. Conway, et al. Antisense precision polymer micelles require less poly(ethylenimine) for efficient gene knockdown. *Nanoscale* **2015**, 7 (48), 20625–20634.
14. H.H. Fakih, J.J. Fakhoury, D. Bousmail, H.F. Sleiman. Minimalist Design of a Stimuli-Responsive Spherical Nucleic Acid for Conditional Delivery of Oligonucleotide Therapeutics. *ACS Appl Mater Interfaces* **2019**, 11 (15), 13912–13920.
15. T. Trinh, C. Liao, V. Toader, et al. DNA-imprinted polymer nanoparticles with monodispersity and prescribed DNA-strand patterns. *Nat Chem* **2018**, 10 (2), 184–192.
16. X. Luo, P. Chidchob, J.F. Rahbani, H.F. Sleiman. Encapsulation of Gold Nanoparticles into DNA Minimal Cages for 3D-Anisotropic Functionalization and Assembly. *Small* **2018**, 14 (5), 1702660.

17. T.G.W. Edwardson, K.L. Lau, D. Bousmail, C.J. Serpell, H.F. Sleiman. Transfer of molecular recognition information from DNA nanostructures to gold nanoparticles. *Nat Chem* **2016**, 8 (2), 162–170.
18. D. Bousmail, L. Amrein, J.J. Fakhoury, et al. Precision spherical nucleic acids for delivery of anticancer drugs. *Chem Sci* **2017**, 8 (9), 6218–6229.
19. T.G.W. Edwardson, K.M.M. Carneiro, C.K. McLaughlin, C.J. Serpell, H.F. Sleiman. Site-specific positioning of dendritic alkyl chains on DNA cages enables their geometry-dependent self-assembly. *Nat Chem* **2013**, 5 (10), 868–875.
20. P.E. Nielsen. Peptide Nucleic Acids (PNA) in Chemical Biology and Drug Discovery. *Chem Biodivers* **2010**, 7 (4), 786–804.
21. J.W. Conway, C.K. McLaughlin, K.J. Castor, H. Sleiman. DNA nanostructure serum stability: greater than the sum of its parts. *Chemical Communications* **2013**, 49 (12), 1172–1174.
22. R. Marchelli, R. Corradini, A. Manicardi, et al. Gene Modulation by Peptide Nucleic Acids (PNAs) Targeting microRNAs (miRs). In *Targets in Gene Therapy*; InTech, **2011**.

# Chapter 6

## 6.0 DNA detection with Hollow Core optical fiber

### 6.1 Abstract

In this chapter the use of PNA as sensing elements for fiber optic biosensors is reported, the work carried out, in collaboration with Prof. Annamaria Cucinotta's laboratory (Department of Engineering and Architecture), was aimed to develop a platform for lab-in-fiber optical biosensors for DNA detection. The principal aim of this project is to functionalize the inner surface of one hollow-core tube lattice fiber (HC-TLF) with some specific PNA probes capable to recognize a target DNA. The fiber's inner surface was derivatized with a PNA probe or with biotin. The signal of this recognition event was then amplified using a sandwich-type construct ,forming a bio-layer on the cladding to allow optical transduction of the binding events. Upon infiltration in the fiber of solutions of the analytes, thanks to an increase of the inner surface thickness upon capture of the biological target, a red-shift in the fiber transmission spectrum was observed. In this way, the optical fiber acted both as a microfluidic devise and as a sensing element itself. A general scheme of the working principle of this platform is shown in *Figure 6.1*.

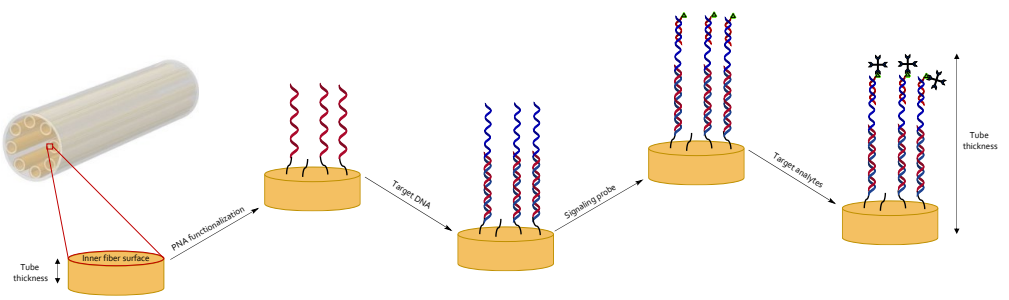



***Figure 6.1.*** *Schematic representation of working principle to DNA sensing with Hollow Core Fibers.*

## 6.2 Introduction

Biosensors are analytical systems that integrate a biological component with a physicochemical detector to detect, quantify, and analyze specific biological elements. Biorecognition element, transducer, and receiver are the three principal components of a biosensor apparatus. Furthermore, biosensors can be classified both on signal transduction methods (such as optical, electrochemical, and mass-based) and based on bioreceptors (like enzyme, antibody, DNA, and others). Some of their benefits encompass exceptional specificity, sensitivity, compact dimensions, and cost-effectiveness.[1]

One of the frequently documented categories of biosensors is that of optical biosensors, sophisticated analytical tools that combine a biorecognition sensing element with an optical transducer system based on light-guiding technologies. In these devices, the modifications in the optical properties of light due to the interaction of the sensor surface with specific analytes are employed to identify biological or chemical analytes, given that these interactions lead to adjustments in the output spectrum, or the electrical signal captured by the photodetector. Its fundamental purpose is to generate a signal that correlates with the concentration of a targeted analyte. Label-free and label-based are the two categories of optical biosensing. Within the label-free mode, the detected signal originates directly from the interaction between the analyzed material and the transducer. Conversely, label-based sensing entails the utilization of a label, and the optical signal is subsequently generated through colorimetric, fluorescent, or luminescent methods.[2,3] Comparing the two categories, label-free optical biosensors are promising tool as they do not require sophisticated molecule labeling processes.[4,5] Selective, rapid, and extremely sensitive measurements can be carried out with optical biosensors, and in fact this type of biosensors have found extensive use in various fields, like medicine, diagnostic, food quality control, environmental safety, and others.[3]

Over the years, optical sensors have been integrated with photonic and fluidic technologies, and one group of optical biosensors that incorporate both

characteristics is biosensors based on Photonic Crystal Fiber (PCF), incrementing the development of lab-in-fiber technologies.[6–8] PCFs are a type of optical fiber that exhibit photonic crystal properties, confining light within a hollow core. The inner surface of these fibers is composed of a micro-structured cladding consisting of air holes running along the fiber's length.[9,10] Solution of biological compound can be infiltered inside this air holes, forming bio-layers on the inner surface of the fiber.[11] Solid-core fibers and Hollow-core fibers are the two classes of PCFs.[12,13] The first class of fibers can guide light both in a high-refractive-index core and a low-refractive-index core. In contrast, hollow-core PCFs can confine light within a low-refractive-index core through two different transmission mechanisms, which are Photonic Band-gap (PBG) or inhibited coupling (IC). The former approach focused on a distinct range of frequencies or wavelengths where light propagation encounters significant prohibition within a complex periodic structure. Conversely, the latter approach relies on the recurrent pattern of high and low transmission bands within the fiber's transmission spectra, determined by the micro-structured cladding's thickness.[14,15] However, the inner surface of a fiber can be functionalized with a bio-receptor such as a PNA probe, ligand, or antibody, generating a biological layer capable of increasing the surface thickness, consequently resulting in a shift in the fiber's transmission spectrum without the use of additional transducers. Furthermore, the fiber can function as a label-free sensor for various kinds of bio-analytes, like DNA and proteins, if the thickness variation is significant enough.[16,17]

For example, HC-TLFs, utilizing the inhibited coupling waveguiding principle, have been suggested for the detection of streptavidin analytes through the formation of biotin-streptavidin bio-layer. In fact, the fiber's inner surface was first functionalized with biotin, which is a small molecule that, having a carboxylic group, can be attached to the sensor's surface through a silanization process, and then a solution of streptavidin analytes was fluxed inside the fiber. The streptavidin-biotin system exhibits a very high binding constant (dissociation constant of about $K_d \approx 10^{-14}$), and

following the infiltration of streptavidin, a redshift in the fiber's transmission spectrum was observed, comparable to the size of the analyte itself.[17]

Nevertheless, PNAs are well-suited for sensing applications due to their strong affinity and selectivity for complementary natural nucleic acid (DNA or RNA), as well as their resistance to chemical and enzymatic degradation.[18,19] Optical fibers modified with PNA have demonstrated the capability to detect complementary DNA with high sensitivity and specificity.[20–22]

## 6.3 Result and discussion

The work regarding this platform for lab-in-fiber optical biosensors for DNA was focused on increasing the thickness of the fiber's inner surface with the formation of a bio-layer through a chemical functionalization protocol.

All the chemical solutions for both derivatization reactions and analysis were infiltered into the channel of the fiber to get in contact with the inner surface, in an *optofluidic* format. The experimental setup was designed and realized using a poly(tetrafluoroethylene) (PTFE) tubing reservoir (100 µL), connected to one moiety of the HC-TLF through a PEEK (polyetheretherketone) ferrule and a PTFE adapter, in presence of a 2 atm nitrogen pressure. To ensure the correct liquid flow, the formation of drops at the opposite extremity of the fiber was visually monitored using filter paper.

Monitoring of the reaction outcome was performed using optical measurements in Prof. Annamaria Cucinotta's laboratory in the Department of Engineering and Architecture (UNIPR) and, also the infiltration and the optical analysis was carried out by Foroogh Khozeymeh and Federico Melli.

The characteristics of the piece of HC-TLF used as a biosensor in this work is reported in *Figure 6.2*. In specific, the length of the piece of HC-TLF used and the cross-section are schematically reported in *Figure 6.2a* and *6.2b*, respectively. However, the fiber features a hollow core with a radius ($R_{co}$), encircled by eight silica tubes characterized by a radius ($r_t$), a thickness ($t$), and a refractive index ($n_d$). Instead, a schematic representation of the bio-layer formed in the fiber's inner surface is showed in *Figure 6.2c*, while in *Figure 6.2d* is reported the chemical bio-layer created with also bioreceptors and streptavidin analytes.

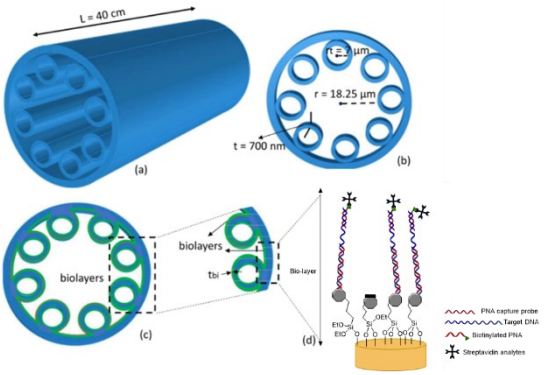



***Figure 6.2.*** *Characteristics of the HC-TLF piece used as biosensor: a) Length; b) fiber cross-section and related structural parameters; c) representation of bio-layer formed inside the fiber; d) chemical details of bio-layers, bio-receptors, and streptavidin analytes.*

The chemical functionalization and detection protocol was obtained using PNA and modified PNA probes as reported in *Figure 6.3*.

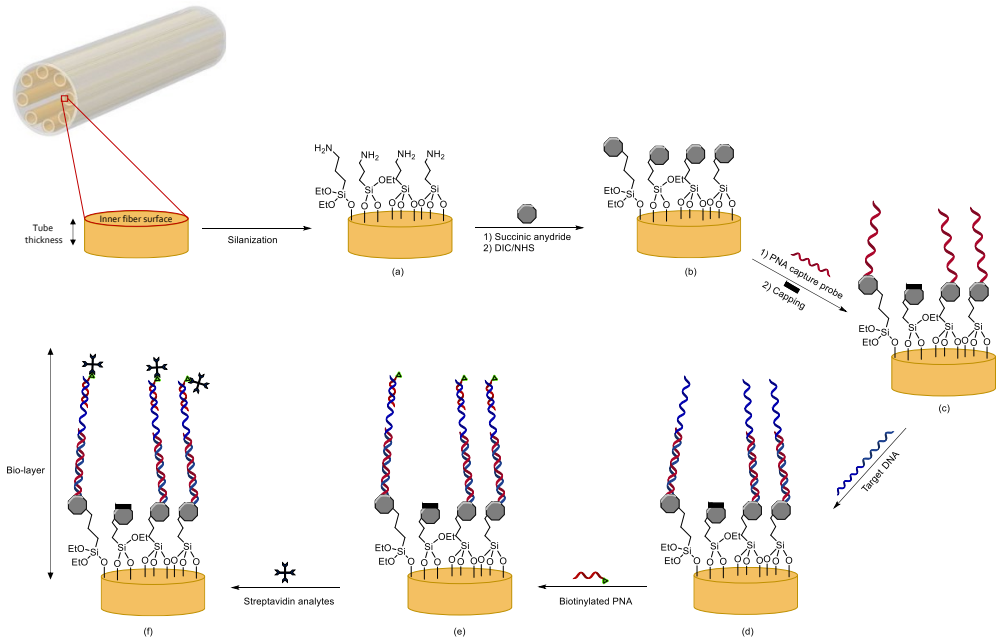



***Figure 6.3.*** *Schematic representation of chemical functionalization process of one Hollow-core tube lattice fiber (a-c) and of the detection of target DNA (d-f): a) silanization; b) reaction with succinic anhydride and formation of active ester; c) PNA capture probe immobilization with amide bond and capping of unreacted active sites; d) flowing the target DNA; e) Biotinylated PNA infiltration as signaling probe, and f) streptavidin solution infiltration to increase the bio-layer thickness and to amplify optical DNA detection.*

The sequences of PNA capture probe, target DNA, and biotinylated PNA used in the present work are reported in *Table 6.1*.

***Table 6.1.*** *Characteristics of PNA capture probe, target DNA and Biotinylated PNA used in the chemical functionalization process.*

| Entry | Oligo | Sigla | Sequence |
|---|---|---|---|
| 1 | PNA capture probe | PNA Soy RR | $N_{term}$ H-O-TGC TAG AGT CAG CTT-$NH_2$ $C_{term}$ |
| 2 | Target DNA | Modified 50-mer SOY RR | 5'- ACC CTA ATC ATT TCA TTT GGA GAGGAC ACG CTG ACA AGC TGA CTC TAG CA -3' |
| 3 | Biotinylated PNA | PNA 8 mer-Biotin | $N_{term}$ Biotin-O-TGG GAT TA-Gly-$NH_2$ $C_{term}$ |

First an acidic solution (MeOH:HCl, 1:1) was fluxed in the HC-TLF to clean the inner surface by possible organic compounds residues and to activate the silica surface. After that, the silanization of the fiber's inner surface, a solution of (3-aminopropyl)triethoxysilane (APTES, 0.1 %) in absolute ethanol, was flushed. Thanks to this organofunctional alkoxysilane molecule, a positively charged layer on the cladding, given by the presence of the protonated terminal amino groups, was obtained. Once the surface was silanized, to obtain terminal carboxylic groups, a solution of succinic anhydride (0.25 M in DMF) was infiltered. Subsequently, a solution of N,N'-diisopropylcarbodiimide (DIC) and N-hydroxysuccinimide (NHS) (0.25 M for both solutions in DMF) was fluxed through the fiber to obtain an active ester on the cladding needed to the formation of the amide bond with the free amino moiety of the PNA probe. To perform the coupling, a solution of the PNA capture probe (30 µM concentration in DMF) with an excess of DIPEA (50 µM concentration) was infiltered.

Up to this point, the derivatization process of the fiber was the same as a former work.[17] However, the protocol was modified to optimize the DNA detection process and increase the precision of the chemical binding. In fact, a solution of ethanolamine in TRIS base (0.3% ETA, pH:9) was flushed to the HC-TLF to block the unreacted active sites.

Subsequently, to obtain the PNA:DNA duplex and the capture of the DNA, a solution of complementary DNA (5 μM concentration) in PBS as a buffer was fluxed in the inner fiber surface. After that, a biotinylated PNA (signaling probe) in DMF as a solvent (30 μM concentration) was infiltered to create a *sandwich* complex. In this case, the specificity of the DNA capture was ensured thanks to the high sequence selectivity of the PNA probe. At the end, a streptavidin analytes solution (1 mg/ml concentration) was injected inside the fiber. In this way, a biotin-streptavidin complex can be formed, which is formed through a very specific and very strong interaction. After each step of infiltrations, doubly-distilled water and ethanol were infiltered to clean the fiber by excess solutions and then it was emptied by flowing the nitrogen in it before all the optical measurements.

Using this protocol, a bio-layer able to increase the thickness of the cladding around the hollow core and to guarantee the optical DNA detection with a signal magnification and a change in the transmission spectrum was formed.

Optical measurements were performed using the following set up: a supercontinuum white light source emitting diffraction-limited light within the 450–2400 nm wavelength range served as the light source. In the initial step, laser output light is directed into the 40 cm long HC-TLF through a specific lens with the correct focal length and a 3-axis NanoMax Flexure stage, allowing for precise and continuous micrometric adjustments in the X, Y, and Z axes. This facilitates the coupling of input light into the HC-TLF input. Then, the output light is directed through an optical objective lens and onto a CCM1–PBS252/M–30 mm Cage Cube-Mounted beam splitter cube, suitable for wavelengths in the 620–1000 nm range. One of the beams is directed to a compact Zelux Camera sensitive to visible wavelengths. Connecting this Zelux Camera with a laptop, the output beam is monitored in real-time. The other beam, after passing through the final objective, is directed to an optical spectrum analyzer (OSA) (AQ-6315A/-6315B) with a 0.5 nm resolution bandwidth, used to measure the output transmission spectrum within the 350–1750 nm wavelength range.

The experiment employed an HC-TLF cross-section, featuring a core with a radius ($R_{co}$) of 18 µm, encased by eight silica tubes with an outer radius ($r_t$) of 7 µm and a thickness (t) of 700 nm. The HC-TLF fiber transmission spectrum before derivatization is reported in *Figure 6.4*. Three high loss regions were noted, one in the 500-540 nm range, the second between 729-805 nm and the last from 1277 nm to 1645 nm, due to the optical fiber allows the transmission of light in these specific range of wavelength.

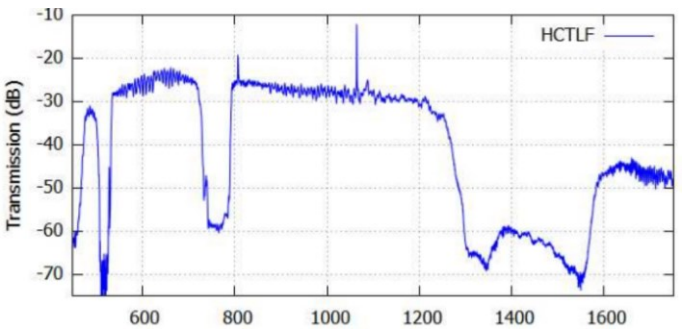



***Figure 6.4*** *Transmission spectrum of HC-TLF used without any infiltration steps.*

From the optical measurements it was observed that in presence of a target DNA fluxed through the inner surface of the HC-TLF in the *optofluidic* process, a bio-layer on the core cladding was formed causing a red-shift on the fiber's transmission spectrum.

However, focus on the comparison between the transmission spectra, before and after the streptavidin-biotin complex formation, in specific after the biotinylated PNA used as signaling probe (*step e*) and subsequently after the injection of streptavidin analytes solution (*step f*), a red shift of the transitions in all the three regions previously mentioned was encountered (*Figure 6.5*).

By the same figure, considering a zoom at -35dB as a reference signal level, it is possible to observe a wavelength red shift from 1.83 to 3.81 nm for the first loss region, from 2.84 nm to 5.76 nm for the second, and a range from 9.82 nm to 17.92

nm concerning the third region. The optical measurements were carried in triplicate to have the possibility to analyze the significance of these shifts.

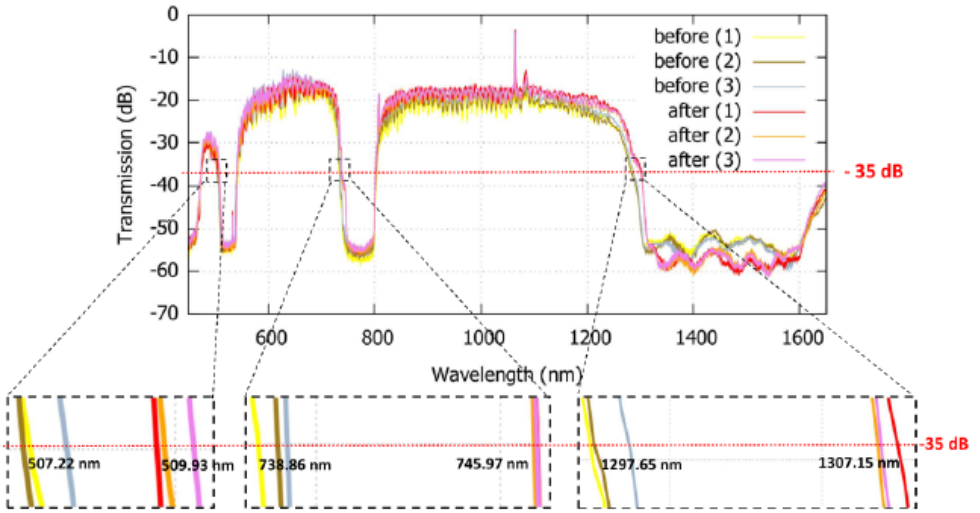



***Figure 6.5.*** *On the top, experimental transmission spectra of HC-TLF before and after the infiltration with streptavidin analytes solution. To test the assurance of the results, the optical measurements was performed three times: (1), (2), and (3). In addition, from the expanded images at -35 dB as reference signal level, the red shift of the transitions in the three regions observed can better be seen.*

Considering the same three regions of transmission spectra above showed, moreover, it is possible to measure the optical red-shift between the measurements performed before and after the streptavidin analytes infiltration (*Figure 6.6*). However, in addition at the -35 dB signal taken above as a reference signal, other three reference signal levels were analyzed, like -40 dB, -45 dB and -50 dB. These measurements have been taken at three distinct wavelengths: 505 nm, 735 nm, and 1290 nm. Observing this changing towards the red zone of the transmission spectrum, the DNA and streptavidin molecules detection was confirmed because the values are similar with the protein's size detection.

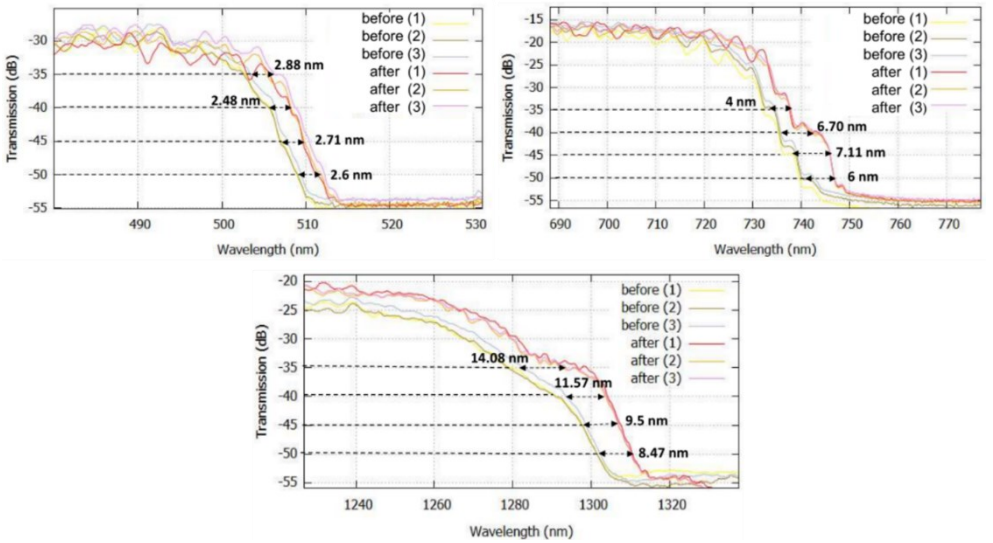



***Figure 6.6.*** *Expanded images at -35 dB, -40dB, -45 dB and -50 dB as references signal levels in different regions of the transmission spectra and relative red-shift value.*

In addition, the observed shifts in the second and third bands were substantial (5.76 nm and 17.92 nm, respectively). According to the 3-sigma criterion, even though we did not carry out a calibration curve to establish the limit of detection (LOD), it can be inferred that the LOD is likely lower than the concentration employed in the test (5 M).

Furthermore, to confirm that the PNA capture probe hybridized with the target DNA and if this bond was specific or a-specific, some qualitative fluorescence measurements with microarray as instrument were carried out. The probes used to perform these measurements were reported in *Table 6.2*, DNA complementary and scramble were labelled with fluorophore, in specific a cyanine 3 dye (Cy3).

***Table 6.2.*** *Characteristics of PNA capture probe, complementary and scramble DNA used for the qualitative fluorescence measurements with microarray.*

| Entry | Oligo | Sigla | Sequence |
|---|---|---|---|
| 4 | PNA capture probe | PNA Soy RR | $N_{term}$ H-O-TGC TAG AGT CAG CTT-$NH_2$ $C_{term}$ |
| 5 | Complementary DNA | 50-mer SOY RR | 5′- [CY3] AGG AAG TTC ATT TCA TTT GGA GAG GAC ACG CTG ACA AGC TGA CTC TAG CA -3′ |
| 6 | Control DNA | 50-mer scramble | 5′- [CY3] TGT AGT CCA TAC CAT TGA GTT ACC AAC CTT ATA GAT AAC CTG TAC ATT AC -3′ |

HC-TLF fiber, a Kagomè lattice (KL) fibers were used for this test. These fibers were derivatized with complementary or scramble DNA functionalized with Cy3 and with the PNA capture probe. In specific, three different fibers were subjected to a derivatization process, using the protocol discussed above, in one fiber only the injection of PNA capture probe in DMF dry (30 μM concentration) was performed, and for the other two a solution of complementary or scramble DNA solution in PBS (both 5 μM concentration) was fluxed after the PNA. These fibers were analyzed using a microarray reader with laser excitation and fluorescence read-out (*Figure 6.7*). It is possible to observe that only the fiber functionalized with the complementary DNA showed a high signal intensity, confirming the presence and the specificity of PNA/complementary DNA hybridization.

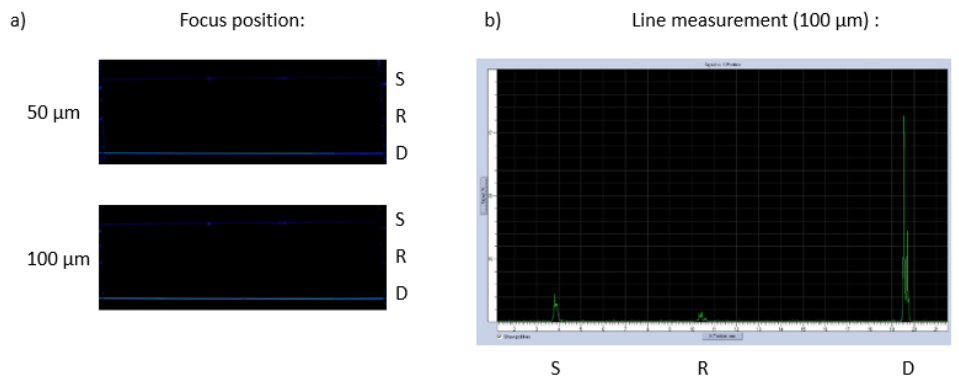



***Figure 6.7.*** *Images obtained with Microarray reader (ScanArray Express) using KL fibers functionalized with PNA capture probe and infiltrated with complementary or scramble DNA. Three different fibers were compared: S) Kagomè fiber with PNA deposited infiltrated with scramble DNA, R) Kagomè fiber after the deposition of PNA capture probe, and D) Kagomè fiber with PNA deposited infiltrated with complementary DNA. Figure 6.7a refers to the three results at two focus positions, 50 μm and 100 μm, instead the Figure 6.7b relates their signal intensity at 100 μm.*

In a second set of experiments, to confirm that the complementary DNA solution remains inside the fiber after the water wash in presence of the PNA capture probe, three fibers were analyzed. One fiber was derivatized with the PNA capture probe in DMF dry (30 μM concentration) and used to capture the complementary DNA in PBS (5 μM concentration). Another one was only infiltered with the DNA (without the PNA deposition step). The third one was only derivatized by the PNA capture probe solution. The microarray reader results (*Figure 6.8*) showed a high signal intensity for both the fibers functionalized with the complementary DNA before the water wash. However, after the water wash, the sample without the PNA infiltration shows a decrease in signal intensity, this means that the complementary DNA is left inside the fiber only in the presence of the PNA probe, confirming the formation of PNA:DNA duplex.

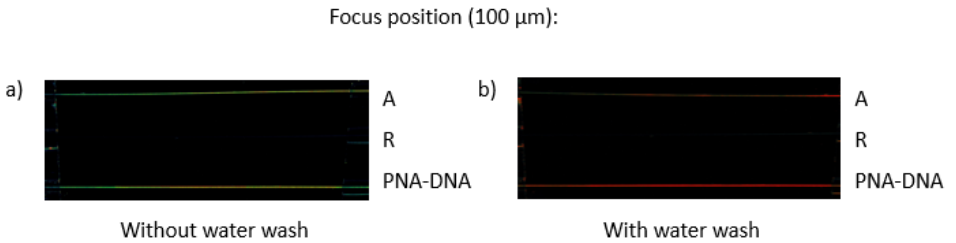



***Figure 6.8.*** *Image obtained with Microarray reader (ScanArray Express) using KL fibers: A) Kagomè fiber with complementary DNA solution in absence of PNA, R) Kagomè fiber after the infiltration of PNA capture probe, and PNA-DNA) Kagomè fiber in presence of PNA probe and complementary DNA. Figure 6.8a and 6.8b display the results before and after the water wash, respectively and after the water wash a decrease of signal intensity was observed in the sample without the PNA functionalization.*

## 6.4 Conclusions

A platform for lab-in-fiber optical biosensors for DNA detection was developed using a Hollow-core tube lattice fiber (HC-TLF). In fact, derivatizing the inner surface of one fiber in a "optofluidic" way with a specific chemical functionalization protocol the detection of DNA molecule and streptavidin analytes was carried out. This chemical process allowed the formation of a bio-layer in the internal cladding of the fiber tubes able to be revealed, thanks to the specific thickness generated with chemical solutions infiltered, by optical measurements. From the optical analysis, a red wavelength shift was observed in the transmission spectra.

In this type of experiments, the assembly of the sandwich architecture was necessary to obtain a significant signal, similarly to the nanoparticle enhancement observed in previous work.[20] The fluorescence measurements carried out on similar fiber using a microarray reader confirm that the specificity of PNA:DNA hybridization is retained in these systems.

## 6.5 Experimental section

### 6.5.1 Reagent and instrumentations

The reagents were bought from Merck, abcr, Carlo Erba, VWR, Fluka, TCI Europe, Janssen, Link Technologies, PolyOrg, Eurofins Genomics and used without additional purification. Only for the solid phase synthesis, the DMF was purged with nitrogen flux to prevent the formation of dimethylamine and was dried over 4Å molecular sieves.

Both the PNAs used in the chemical functionalization protocol (PNA capture probe and biotinylated PNA) were already present in the laboratory and only the characterization with the UPLC-MS (Waters Acquity Ultra Performance LC) and the concentration with Evolution 260 Bio UV-Visible spectrophotometer (Thermo Fisher Scientific SPE8W) were carried out.

For the characterization the following instrumental set-up were used: Waters Acquity ultra-performance LC E07SQD086W, with Waters SQ detector and ESI-interface equipped with Acquity UPLC BEH 300 (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic condition: eluent A: water + 0.2% formic acid; eluent B: acetonitrile + 0.2% formic acid. Column temperature: 35 °C. Program: initial isocratic at 100% A (0.9 min), then linear gradient to 50% B (in 5.7 min). Final wash with 100% B for 1.2 min. Flow rate: 0.25 mL/min. Some PNAs, however, were characterized with Thermo LTQ Orbitrap XL detector and ESI-interface equipped with a Phenomenex Kinetex EVO (50 × 2.1 mm, 1.7 µm, C18, 100 Å). Chromatographic conditions: eluent A: water + 0.1% formic acid; eluent B: acetonitrile + 0.1% formic acid. Column temperature: 35 °C. Program: initial isocratic at 98% A (3 min), then linear gradient to 50% B (in 20 min) and after to 95% (in 1 min). Final wash with 98% B for 8 min. Flow rate: 0.20 mL/min. UV wavelength: 260 nm. The qualitative fluorescence measurements were performed with ScanArray $G_x$ Microarray Scanner (Perkin Elmer).

The concentration of the two PNAs was calculated following the UV-absorbance at 260 nm and assuming an additive contribution of all nucleobases (see Chapter 2).

### 6.5.4 Chemical Functionalization Protocol

The derivatization of the inner surface of the fiber was performed with specific chemical solutions, which was flushed through the HC-TLF with a nitrogen overpressure (2 atm), as described below.

An acidic solution (MeOH:HCl, 1:1) to clean and active the fiber was infiltered and then a solution of (3-aminopropyl) triethoxysilane 0.1% (APTES) in absolute ethanol was used to silanize the glass. After ethanol and water washes, a solution of succinic anhydride 0.25 M in DMF dry was flushed to obtain the terminal acidic moiety. To perform the coupling with the PNA, the N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC) and N-Hydroxysuccinimide (NHS) at 0.25 M for both in DMF dry were infiltered Subsequently, the PNA capture probe (30 µM concentration) in DMF dry was flushed inside the fiber with an excess of DIPEA (50 µM concentration), and after that a water wash was performed. To quench the excessive activated esters, an ethanolamine in TRIS base (0.3% ETA, pH:9) solution was used, followed by TRIS base and water washes. Then, a solution 5 µM of target DNA in phosphate buffered saline (PBS) was injected and later water washes was performed. After that, the signaling probe solution, that is 30 µM of biotinylated PNA in DMF dry, was flushed inside the fiber. Finally, after a water washes, a 1 mg/ml of a streptavidin analytes solution was infiltered. Furthermore, 1ml of each solution was prepared and the duration of infiltration was one hour in one direction, while the biotinylated PNA and the streptavidin solution was flushed in both the direction (from the left to right and vice-versa). In addition, before to perform the optical measurements, the inner surface of the fiber was dried under nitrogen flux for 30 minutes.

For the microarray experiments some Kagomè lattice (KL) fibers were used and derivatized with the same chemical functionalization protocol discussed above until the DNA solution step.

The solution infiltrations, the measurements and elaboration of the data were performed by Foroogh Khozeymeh and Federico Melli.

## 6.5.3 PNAs characterizations

Entry 1 (Figure 6.A1 in Appendix)
Check of the PNA capture probe by UPLC-MS: H-O-TGC TAG AGT CAG CTT-$NH_2$
$t_r$: 3.10 min; ε (260 nm): 150700 $M^{-1}cm^{-1}$; MW calculated: 4238.13 [M], observed (from ESI-MS deconvoluted): 4240.4; ESI-MS: m/z observed (calcd): 1414.6 (1413.7) $[M+3H]^{3+}$, 1061.1 (1060.5) $[M+4H]^{4+}$, 848.8 (848.6) $[M+5H]^{5+}$, 707.4 (707.4) $[M+6H]^{6+}$, 606.5 (606.4) $[M+7H]^{7+}$.

Entry 3 (Figure 6.A2 in Appendix)
Check of the Biotinylated PNA probe by UPLC-MS: Biotin-O-TGG GAT TA-Gly-$NH_2$
$t_r$: 3.24 min; ε (260 nm): 88300 $M^{-1}cm^{-1}$; MW calculated: 2668.66 [M], observed (from ESI-MS deconvoluted): 2668.6; ESI-MS: m/z observed (calcd): 1335.3 (1335.3) $[M+2H]^{2+}$, 890.6 (890.6) $[M+3H]^{3+}$, 668.1 (668.2) $[M+4H]^{4+}$, 534.5 (534.7) $[M+5H]^{5+}$.

# 6.6 Appendix Chapter 6

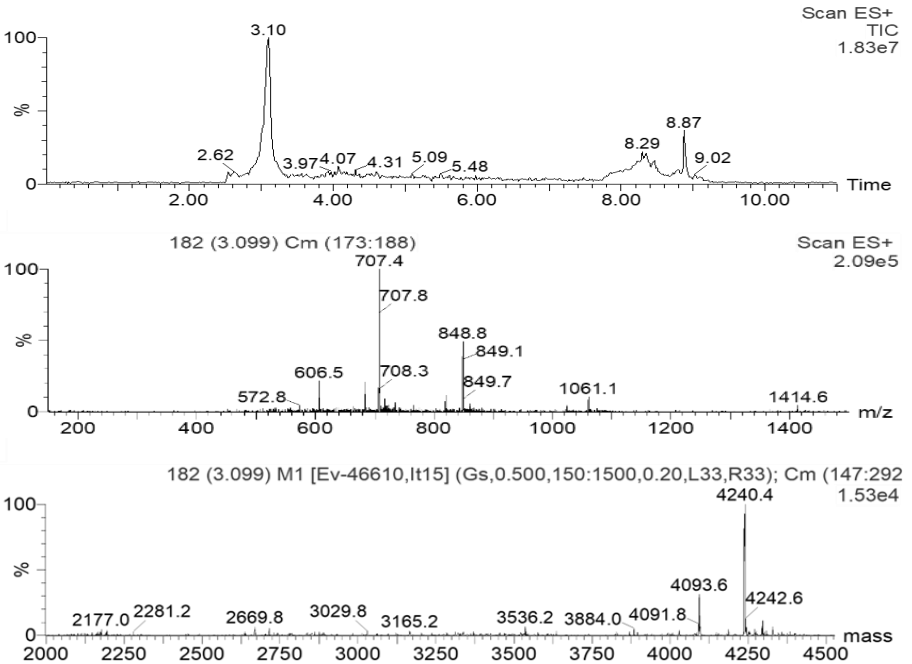



***Figure 6.A1.*** *UPLC-MS of PNA capture probe: chromatogram, mass spectrum and deconvoluted spectrum.*

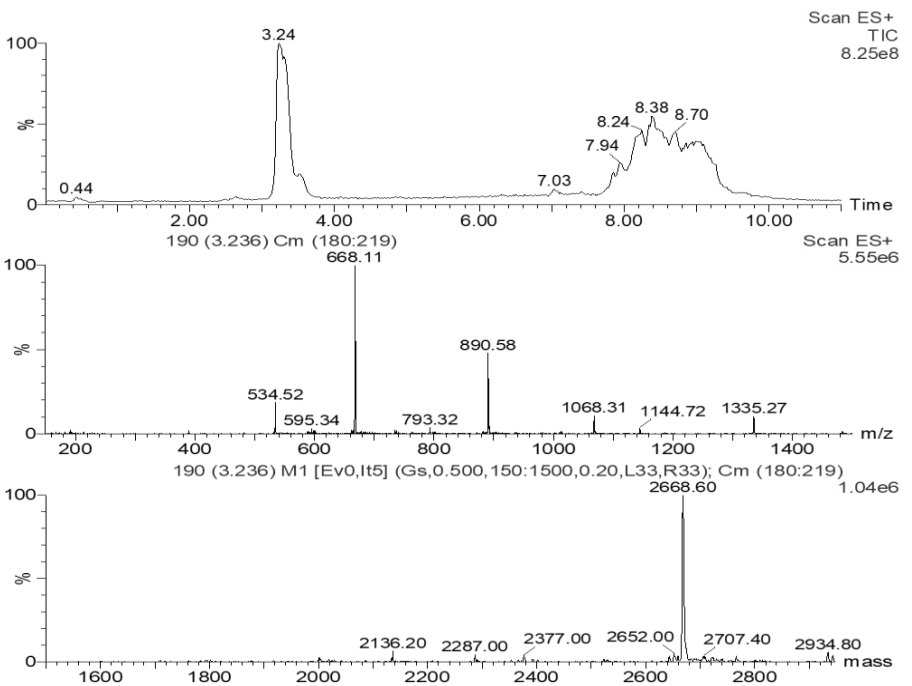



***Figure 6.A2.*** *UPLC-MS of Biotinylated PNA: chromatogram, mass spectrum and deconvoluted spectrum.*

## 6.7 References Chapter 6


1. C. Karunakaran, R. Rajkumar, K. Bhargava. Introduction to Biosensors. In *Biosensors and Bioelectronics*; Elsevier Inc., **2015**; pp 1–68.
2. P. Damborský, J. Švitel, J. Katrlík. Optical biosensors. *Essays Biochem* **2016**, 60 (1), 91–100.
3. A.K. Singh, S. Mittal, M. Das, A. Saharia, M. Tiwari. Optical biosensors: a decade in review. *Alexandria Engineering Journal*. Elsevier B.V. March 15, 2023, pp 673–691.
4. G. Zanchetta, R. Lanfranco, F. Giavazzi, T. Bellini, M. Buscaglia. Emerging applications of label-free optical biosensors. *Nanophotonics*. Walter de Gruyter GmbH 2017, pp 627–645.
5. P. Chamorro-Posada. Asymmetric Concentric Microring Resonator Label-Free Biosensors. *Photonics* **2022**, 9 (1).
6. A. Ricciardi, A. Crescitelli, P. Vaiano, et al. Lab-on-fiber technology: A new vision for chemical and biological sensing. *Analyst*. Royal Society of Chemistry December 21, 2015, pp 8068–8079.
7. J.Z. Zhang, X. Yang, C. Shi, R. Newhouse, C. Gu. Hollow-core photonic crystal fibers for surface-enhanced raman scattering probes. *International Journal of Optics*. 2011.
8. R. Nasirifar, M. Danaie, A. Dideban. Hollow-core graded index optical fiber refractive index sensor based on surface plasmon resonance. *Opt Quantum Electron* **2020**, 52 (7), 341.
9. F. (Federica) Poli, A. (Annamaria) Cucinotta, Stefano. Selleri. Photonic crystal fibers : properties and applications; Springer, **2007**.
10. P.St.J. Russell. Photonic-Crystal Fibers. *J. Lightwave Technol.* **2006**, 24 (12), 4729–4749.
11. E. Coscelli, M. Sozzi, F. Poli, et al. Toward a highly specific DNA biosensor: PNA-modified suspended-core photonic crystal fibers. *IEEE Journal on Selected Topics in Quantum Electronics* **2010**, 16 (4), 967–972.
12. F. Luan, A.K. George, T.D. Hedley, et al. All-solid photonic bandgap fiber. *Opt. Lett.* **2004**, 29 (20), 2369–2371.
13. T.T. Larsen, A. Bjarklev, D.S. Hermann, J. Broeng. Optical devices based on liquid crystal photonic bandgap fibres. *Opt. Express* **2003**, 11 (20), 2589–2596.
14. F. Couny, F. Benabid, P.J. Roberts, P.S. Light, M.G. Raymer. Generation and Photonic Guidance of Multi-Octave Optical-Frequency Combs. *Science (1979)* **2007**, 318 (5853), 1118–1121.
15. L. Vincetti, V. Setti. Waveguiding mechanism in tube lattice fibers. *Opt. Express* **2010**, 18 (22), 23133–23146.
16. F. Giovanardi, A. Cucinotta, L. Vincetti. Inhibited coupling guiding hollow fibers for label-free DNA detection. *Opt Express* **2017**, 25 (21), 26215–26220.
17. F. Giovanardi, A. Cucinotta, A. Rozzi, et al. Hollow Core Inhibited Coupling Fibers for Biological Optical Sensing. *Journal of Lightwave Technology* **2019**, 37 (11), 2598–2604.

18. R. D'Agata, R. Corradini, G. Grasso, R. Marchelli, G. Spoto . Ultrasensitive Detection of DNA by PNA and Nanoparticle-Enhanced Surface Plasmon Resonance Imaging. *ChemBioChem* **2008**, 9 (13), 2067–2070.
19. A. Bertucci, A. Manicardi, R. Corradini. Advanced Molecular Probes for Sequence-Specific DNA Recognition. In *Detection of Non-Amplified Genomic DNA*; Spoto, G., Corradini, R., Eds.; Springer Netherlands, Dordrecht, **2012**; pp 89–124.
20. A. Bertucci, A. Manicardi, A. Candiani, et al. Detection of unamplified genomic DNA by a PNA-based microstructured optical fiber (MOF) Bragg-grating optofluidic system. *Biosens Bioelectron* **2015**, 63, 248–254.
21. A. Candiani, M. Sozzi, A. Cucinotta, et al. Optical fiber ring cavity sensor for label-free DNA detection. *IEEE Journal on Selected Topics in Quantum Electronics* **2012**, 18 (3), 1176–1183.
22. A. Candiani, A. Bertucci, S. Giannetti, et al. Label-free DNA biosensor based on a peptide nucleic acid-functionalized microstructured optical fiber-Bragg grating. *J Biomed Opt* **2013**, 18 (5), 057004.

# Conclusion remarks

This PhD thesis focused the attention on PNAs and their employment as building blocks to mediate the self-assembly of nanostructures (*Chapter 2, 3 and 5*), as probe in RNA-targeting therapy (*Chapter 4*), and in DNA-detection (*Chapter 6*).

In relation to *Chapter 2*, leveraging our laboratory's proficiency in PNAs and the DNA expertise of Prof. Hanadi Sleiman's group, the assembly of new nanostructures was carried out thanks to the capability of PNAs to bind specifically DNA oligomers and their ability to form very stable DNA:PNA complexes. Given that the PNAs were functionalized, in orthogonal way, with specific peptide sequences, it was possible to force the formation of heterotrimeric coiled-coil structures in presence of short peptide moieties. Moreover, the peptide design was selected to generate minimal coiled-coil structures with metal binding abilities mimicking plastocyanin. As control, circular dichroism (CD) studies performed with peptide sequences containing or not the copper binding site and with PNA-peptide sequences to understand the minimal length necessary at which self-assembly of coiled-coil structure, without a rigid platform that ensure the spatial proximity of peptide domains, was occurring. The programmed assembly of heterotrimeric peptide-PNA:DNA was confirmed by polyacrylamide gel electrophoresis (PAGE) and CD analyses, the latter consistent with the formation of coiled-coil structures. Binding of copper(II) ion was also demonstrated via fluorescence quenching experiments. The results obtained with the trimeric template and the three PNA-11mer peptide conjugates in one orientation (orientation B), revealed interactions with Cu(II), but the data available did not allow to establish if the actual binding site are those designed for the mini-metalloprotein model. For this reason, other type of experiments and comparison should be performed to understand this important feature. If successful, this first proof-of-concept work can lead to a strategy for the construction of mini-proteins of various functions, such as mini-enzymes or electron transporter units, using a rational design.

Since the use of a DNA template requires different strategies and chemistries (phopshoramidite synthesis for DNA and peptide synthesis for PNA), it was interesting to establish if the same approach can be performed employing a trimeric PNA as a template. Therefore, in *Chapter 3* the research was focused on: a) the synthesis and the optimization of a tri-functionalized rigid molecule designed to be used as a scaffold for the synthesis of the trimeric PNA-template, and b) the study of the length of PNA:PNA segments necessary to obtain a sufficient thermal stability to ensure the formation of the final templated assembly. Regarding the first point, the synthesis of a tri-functionalized rigid molecule was carried out by the optimization of the synthetic protocol, starting from commercially available starting material, which can be Fmoc-protected to obtain the proper branching unit. This rigid molecule would then allow to grow three different PNAs branched in direction from $C_{term}$ to $N_{term}$ thanks to the orthogonality of these three different groups. For the second point, a library of PNA sequences having different composition, length, and orientation was synthesized, allowing to perform melting and annealing measurements, and study the stability of PNA:PNA duplexes, especially in parallel orientation. This information was useful to establish the minimum length suitable to generate a three-fold PNA:PNA duplex as driving force for the assembly of minimal hetero coiled-coil systems. Interestingly, a general method to roughly predict the stability of antiparallel PNA:PNA duplex stability was obtained, whereas for the parallel complexes a trial-and-error approach is still necessary. This part therefore provides a new model for designing nano-structures containing PNA:PNA duplexes. The study of hybrid DNA:PNA nanospheres (*Chapter 5*) allowed to understand the effect of their decoration with charged PNA (polyarginine-conjugate PNA) on the overall stability of these structures. PAGE and DLS measurements showed that the addition of charges on the PNA chain can have a dramatic effect on the nanosphere stability, given that the equilibrium between attractive forces and repulsive electrostatic interactions, crucial for maintaining the spherical shape and determining the size of the particle, is perturbed.

Taking into consideration the potentially use of PNA as probe in RNA-targeting therapy (*Chapter 4*), a series of PNAs were designed and synthetized to target different regions of SARS-CoV-2. At the same time, the syntheses of RNase L recruiters were carried out and some of the anti-viral PNAs were functionalized with these RNase L recruiters. The first biological data obtained are encouraging in terms of specificity and efficacy. Moreover, the preliminary results aimed at investigating whether the modified PNA with the RNase L recruiter can effectively identify the target RNA, engage the RNase L, and prompt selective cleavage of the target RNA suggest a tendency toward RNA cleavage. However, comprehensive studies are necessary for a thorough evaluation of the optimal model for this specific application.

PNA-mediated self-assemblies can also be very useful in sensing technologies, though this is a well-established principle, still new applications can be designed on its basis. As example, in *Chapter 6* was reported a platform for lab-in-fiber optical biosensors for DNA detection using a hollow-core tube lattice fiber (HC-TLF). The inner surface of HC-TLF was functionalized using flow chemistry with PNA probes capable to recognize a target DNA, allowing subsequent self-assembly of a sandwich-type complex with a second biotinylated probe and streptavidin. The formation of a bio-layer in the internal cladding of the fiber tubes could then be revealed by measuring of the optical properties of the fiber itself: a red wavelength shift was observed in the transmission spectra. Though conceptually simple, the use of these sophisticated optical fibers can ultimately lead to devices of great potential, since they combine the wave-guiding properties of optical fibers with the highly specific properties of the PNA probes. Sensing "optofluidic" devices of this type, though still in their infancy, could be of extreme interest in the medical field, eventually leading to advance in-vivo monitoring techniques.

# Publications

*Publications related to this PhD work:*

1. F. Khozeymeh, F. Melli, S. Capodaglio, R. Corradini, F. Benabid, L. Vincetti, A. Cucinotta, Hollow-Core Fiber-Based Biosensor: A Platform for Lab-in-Fiber Optical Biosensors for DNA Detection Sensor, 2022, Vol. 22, 5144.
2. F. Zhao, M. Frandsen, S. Capodaglio, H.F. Sleiman, DNA-Mediated Peptide Assembly into Protein Mimics, Journal of the American Chemical Society, 2024.

*Other publications:*

1. G. Pastore, S. Gabrielli, R. Giacomantonio, G. Lupidi, S. Capodaglio, F. Stella, E. Leone, T. Compagnucci, E. Marcantoni, An efficient synthesis of bio-based Poly(urethane-acrylate) by $SiO_2$-Supported $CeCl_3 \cdot 7H_2O$–NaI as recyclable Catalyst, Results in Material, 2022, Vol. 15, 100294. Publication on Master's degree.
2. S. Fortunati, C. Giliberti, M. Giannetto, A. Bertucci, S. Capodaglio, E. Ricciardi, P. Giacomini, V. Bianchi, A. Boni, I. De Munari, R. Corradini, M. Careri, A highly sensitive electrochemical magneto-genosensing assay for the specific detection of a single nucleotide variation in the KRAS oncogene in human plasma, Biosensors and Bioelectronics: X, 2023, Vol. 15, 100404.

# Acknowledgements

At the end of these three intense years, it seems inevitable to dedicate these few lines to thank the people who have accompanied me, some in one way and some in another, along this journey of professional and personal growth.

I would like to thank Prof. Roberto Corradini, the supervisor of this PhD thesis, for giving me the opportunity, by hosting me in his laboratory, to undertake this path of professional and personal growth. I would also like to thank Dr. Alex Manicardi, the co-supervisor of this PhD thesis, for the project ideas, the teachings, the help, and the patience.

I also want to thank all the people that I have met in Lab50, from trainees to post-doc. Thank you for everything you have taught me, for your patience and assistance, for every shared moment, and for the conversations and laughter.

I would also like to thank Prof. Hanadi Sleiman for welcoming me into her laboratory, for the meetings that have allowed me to grow, and for the professional and personal attention and care during my brief stay in her lab. A special thanks goes to the entire Sleiman group, but particularly to Fangzhou Zhao for his help, patience, collaboration, and desire to do. Consequently, I would also like to thank the Research and Innovation Staff Exchange (RISE) program for giving me the opportunity to live this work experience at the McGill University.

I also wish to thank all the collaborations, both internal and external at UNIPR, that have contributed to the creation and development of some projects mentioned: thanks to Prof. Matteo Tegoni, thanks to Prof. Roberto Gambari and his entire research group, and thanks to Prof. Annamaria Cucinotta and her entire research group. In addition, I would also like to thank the MIUR that financed the COVID project through the FISR-COVID program.

I also want to thank all the people I had the opportunity to meet in this department during my Ph.D. and during my secondment in Montreal. Thank you for your words of support, for the help, for the collaboration, for the conversations, for letting me

share, for the hugs, for the aperitifs, for the dinners, and for making each day different from the previous one.

I would also like to thank all the technicians at UNIPR, but especially Marco and Paola, for your dedication, support, and the patience you have shown with all my requests.

I would like to thank all my friends, especially those friends from 'always' (only distant in terms of kilometers) and the fencing friends (for allowing me to cancel out the outside world with every single match). A special thanks goes to both of you, Gullo and Bene, and thanking you is still not enough for everything you have done for me.

Last but not least, special thanks go to my parents, my brother, Miriam, and all my family for supporting and tolerating me, and for giving me the freedom to choose, to make mistakes, and to grow.

To you, Irene, I dedicate this small work for our stolen time:

"Quanto manca alla vetta? Tu sali e non pensarci"_ Friedrich Nietzsche