UNIVERSITA' DEGLI STUDI DI PARMA

Dottorato di ricerca in Scienze Chimiche

XXIV Ciclo
Triennio 2009 - 2011

Modified peptide nucleic acids (PNAs) for nucleic acids detection and anti-miR activity

Coordinatore:
Chiar.mo Prof. Giovanni Predieri

Tutor:
Chiar.mo Prof. Roberto Corradini
Chiar.ma Prof. Rosangela Marchelli

Dottorando: Alex Manicardi

The mind is its own place, and in itself
Can make a heav'n of hell, a hell of heav'n
What matter where, if I be still the same,
And what I should be, all but less than he
Whom thunder hath made greater?

[John Milton – Paradise Lost – 1:254-258]

Table of contents

Outlook

The main target of the work presented in this Ph. D. Thesis is related to the development of new multifunctional PNA systems with both modified backbones or nucleobases, that can be used for targeting nucleic acids, in particular RNA, thus acting as new tools for different biotechnological fields, such as detection of target sequences of biological interest, inhibition of miR activity, and DNA/RNA template reactions.

In **chapter 1** a review of the literature about the challenges on nucleic acid recognition and in particular on RNA targeting (section I), and on the properties and applications of PNA and their chemical modifications (section II) is reported.

In **chapter 2** the synthesis of PNA containing a modified backbone or a modified nucleobase (Figure 2) is reports together with their recognition ability for the determination of SNPs. In the first part of the chapter the synthesis of "chiral box" PNAs containing an arginine side chain in the backbone is reported. These PNAs were covalently linked to microarrays, and an evaluation of their properties in detection of single nucleotide polymorphisms (SNPs) in DNA, in both solution and surface were investigated (as a function of structural parameter). In the second part of the chapter the synthesis of a modified monomer containing a thymine derivative bearing a pyrene moiety, and of PNAs containing this monomer in different positions is described. These compounds, upon hybridization with a complementary DNA sequence, were found to form excimers which allowed an easy detection by fluorescence.

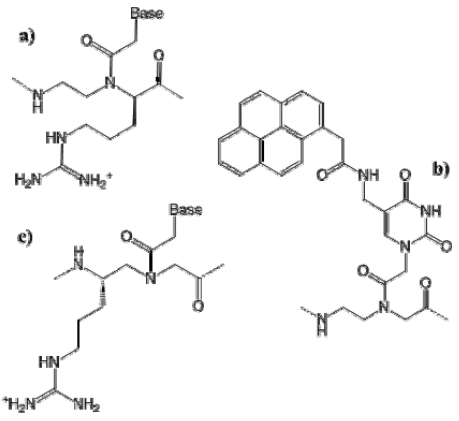


Figure 2: modifications introduced; a) 2D-Arg, b) Py-thymine, c) 5L-Arg

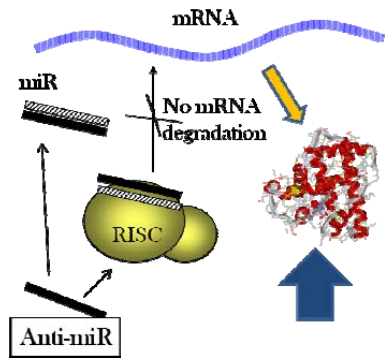


Figure 1: possible mechanism of miR inhibition.

In **chapter 3** the design and the synthesis of different PNAs as new potent anti-miR agents (Figure 1) are

described. PNAs containing a cell-penetrating peptide (CPP) motif conjugated or embedded in the backbone of the oligomer were synthesized. The stability of the complexes formed by these PNA with complementary DNA and RNA were found to be extremely high. The conjugated PNA were able to efficiently enter into cells and target miR210, a microRNA involved in erythroid differentiation, and inhibit its action on mRNA. In particular, the first example of anti-miR activity by a backbone-modified PNA is shown.

In **chapter 4** the synthesis of PNA-based systems able to interact and react with each other in the presence of a target sequence (DNA or RNA) is reported (Figure 4). In the first part the synthesis of a system bearing a pro-functional molecule that can be released only in the presence of a specific target sequence is reported; in the second part the design of different probes for the cellular detection of miR sequences is described; two probes, one with a rhodamine derivative bearing azide residues, and another with a phosphine unit were synthesized: these two probes were able to react in a template reaction which led to an increase in rhodamine fluorescence quantum yield, thus revealing the presence of the target DNA sequence.

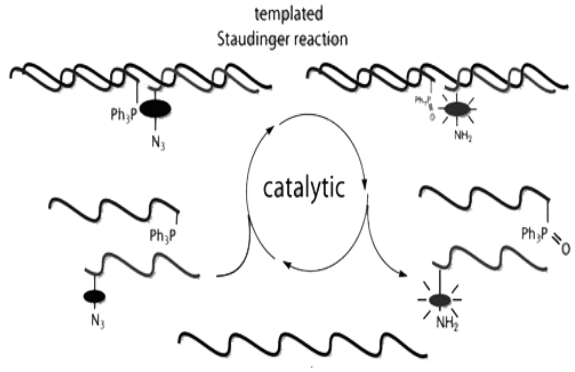


Figure 4: general mechanism of template reaction

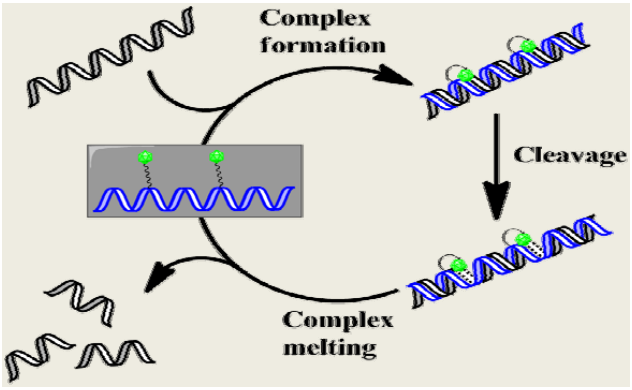


Figure 3: proposed mechanism of catalytic RNase

Combining the results obtained in the previous chapters, in **chapter 5** the synthesis of different multifunctional PNA systems bearing additional binding sites and a catalytic center with potential RNase activity (either on the backbone or on the nucleobase) is described (Figure 3).

.

Chapter 1 - Introduction

Challenges in Nucleic Acids Recognition: RNA

1.1 - General concepts

The role of storage, replication and use of all the information needed for the whole living being is played by nucleic acids. In this family two different types of molecules are presents, the deoxyribonucleic acids (DNA), which is involved in the conservation and transmission of the genetic information, and the ribonucleic acid (RNA), which is responsible for all processes related to the reading and the "interpretation" of this information. Both families of compound are polymers made by single units (nucleotides) constituted by a backbone of sugar units (deoxyribose and ribose respectively, Figure 1.1a, b) connected through ester bonds with a phosphate unit: the recognition occurs through four different nucleobases that are the core of all the processes of recognition on which the genetic information storage and expression are based on.

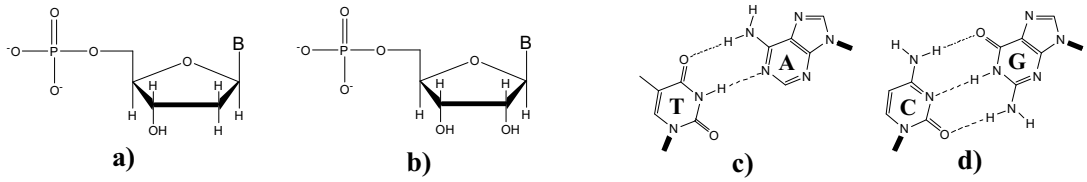


Figure 1.1: a) structure of deoxyribonucleotides; b) structure of nucleotides; c) H-bonding between thymine (T) and adenine (A); d) H-bonding between cytosine (C) and guanine (G).

The mutual recognition of adenine by thymine or uracil (Figure 1.1c) and of guanine by cytosine (Figure 1.1d) through H-bond was first discovered by Watson ad Crick[1] and it is the basis of the high fidelity in all the processes in which the nucleic acids are involved. Although this recognition pattern is the most important, there are other types of possible pairing via hydrogen bonds, among these one of the most important is the Hoogsteen base pairing, found in triple-stranded helices, in which in addition to the normal Watson-Crick base pairing a second pyrimidine base binds to the purine in the major groove.

Based on these characteristics, it is easy to predict and even to design the interactions between two different strands, so that this important issue is used in different field (medicine, biology,

biotechnology, …) to build artificial sequences able to control the transmission of the genetic information and to manipulate it in order to fine control the protein profile in cells, thus regulating the whole functions, from growth and replication, to induction or to block specific functions or even induce cellular death. It is also possible to use these artificial sequences as a sensing tool to characterize the variation in the genetic pool associated to a particular disease, in order to better understand the origin of the malfunction, since they allow the possibility of fast recognizing, by specific interactions, sequences related to that genetic disease. Nucleic acids and analogous molecules can therefore be very useful for the development of new tools for the early diagnosis of diseases, for the detection of pathogens, such as bacteria or viruses, and for the detection of DNA tracts of interest in forensic sciences and food analysis.

1.2 - RNA

Compared to DNA, the structure of RNA shows a great increase of complexity not only for the additional hydroxyl group in the 2' position of the ribonucleoside, but for its greater versatility in term of structure, conformation and reactivity. RNA has functions not only related to the expression of the encoded information, but it is also involved in the recognition and transport of other molecules and is endowed of different types of catalytic properties.

Differences between DNA and RNA structures are the presence of the uracil, a thymine analog lacking the C5 methyl group, in the nucleobases alphabet, and, most importantly, the presence of an hydroxyl group in the 2' position of the sugar ring, that confers to the structure different properties in term of conformation and hydration, and it provides a set of H-bond donor/acceptor groups along the strand[2].

From the conformational point of view the modification in the sugar backbone induce a more compact A-type structure in the double helix complex, compared to those of DNA in the B-form, with a wider and flatter minor groove and a narrower and deeper major groove; in this A-form the phosphate groups are placed in the major groove, confering a strong negative electrostatic potential, as is it for the minor groove in the B-form of DNA duplexes.

Another difference between DNA and RNA is the great variety of alternative base pairing that can be found in the latter class of molecules[3]. These non canonical recognition motifs (Figure 1.2) are found in many contexts: they are at the basis of the tertiary structure of RNA, they are responsible

of the degeneration in codon-anticodon recognition (in particular the wobble pairing, Figure 1.2a, b) and they are present in the core of some catalytic RNAs or in aptameric structures.

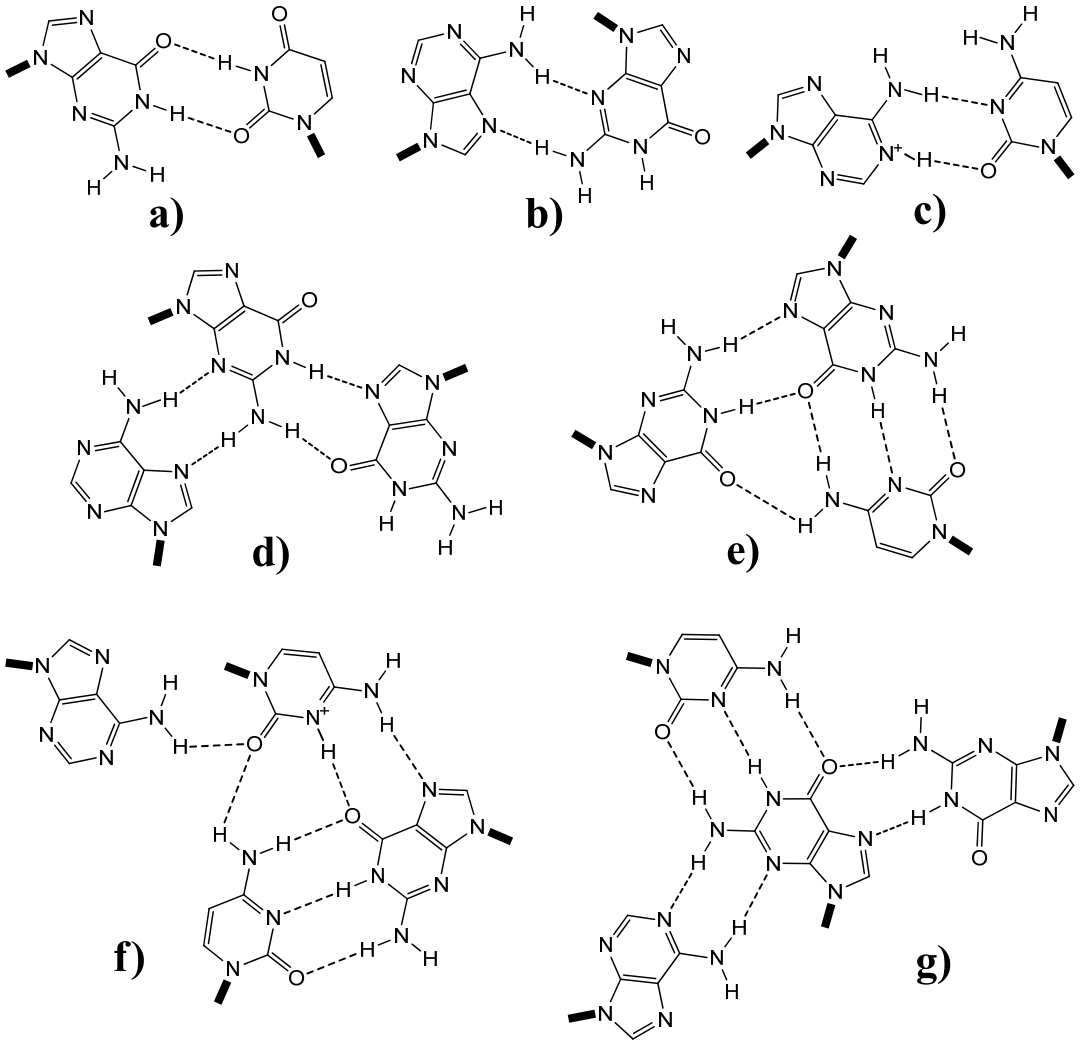


Figure 1.2: alternative base pairing in RNA: a) G-U wobble pair; b) G-A sheared pair; c) A^+-C pair; d) GGA triplex[4]; e) GGC triplex[5]; f) AC^+GC quadruplex[6]; g) GGCA quadruplex[7].

In contrast with DNA in which the tertiary structure is only related to its folding in the chromosomes, the tertiary structure of RNA is more important inasmuch as it is often related to the function that this nucleic acid play in cellular compartments. Tertiary structured RNA can be found in tRNA, rRNA, snRNA, some introns, ribozymes and in the interaction between sets of RNA duplexes. The tertiary structure can be stabilized by the presence of metal ions or by the donor/acceptor behavior of the –OH groups in the sugar scaffold, that can play an important role in the organization between different RNA duplexes (ribose zipper).

1.2.1 - Type of RNA

In the process of transcription the DNA acts as a template for the RNA synthesis by RNA polymerases, then the RNA transcripts are processed by different enzymes that can modify the primary structure of the strand or help the correct folding of the tertiary structure before the molecule can be transported to the target cellular compartment and perform its task. Different roles are played by these molecules inside the cells: messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), microRNAs (miRNAs or miRs), small interfering RNAs (siRNAs), long non-coding RNAs (lncRNAs), aptamers and ribozymes.

The main role played by RNA inside the cells is connected to the expression of the genetic information encoded in the DNA database; indeed three of these systems are involved in the complex mechanism that starts from the DNA transcription to the production of the target protein. After the splicing of the intronic sequence by the spliceosome, a ribonucleoprotein complex composed by five different snRNAs that are needed for the identification of the splicing sites, the mRNA is the tool used to physically transport the genetic information from the nucleus to the cytosol. The mature sequence can be recognized by the ribosome machinery in which the tRNA, RNAs showing a complex tertiary structure, each loaded with a specific amino acid, are non-covalently bound to the mRNA sequence in consecutive steps, from the 5' to the 3', using a system based on three nucleotide length code (codon on mRNA, anticodon on the tRNA) and transfer their amino acids to the growing protein chain. In this context the role for the rRNA is to act as a scaffold for recognition between mRNA and tRNA, and, together with the ribosomal protein, to help in the protein synthesis.

The role of the snRNA is not only related to the slicing of the transcribed mRNA to its mature form, but also to the regulation of transcription factors or RNA polymerase and to maintaining telomeres. The snoRNAs belong to this family of compounds but they are specifically located in the nucleolus and regulate the processes involved in the selective modification of the pre-rRNA, tRNA and snRNA[8].

Aptameric RNAs and ribozymes are two classes of RNA that are not specifically related in processes involved in gene expression, in fact the former specifically bind, with high affinity, to a particular ligand[9], whereas the latter are RNA with a specific tertiary structure that allows them to catalyze chemical reactions.

lncRNAs are mRNA-like non-protein-coding long RNA (more than 200 bases) that are involved in a lot of different processes, from chromatin remodeling to transcriptional regulation and in maintaining the integrity of subcellular compartment and other important biological functions[10]. The mechanisms involved in their regulation ability remain largely unknown, although they interact with transcription factors as co-activators or as decoy systems and with RNA-binding proteins, modifying their ability in the recognition of target sequences. MicroRNAs and siRNAs are short non-coding RNA (19÷25 bases) that are involved in the regulation of gene expression by incorporation is a multienzymatic complex that is responsible of the fate of the target mRNA sequence. Their biosynthesis and activity will be discussed in the next section.

In addition to the kinds of RNA described, in last years other classes were found to be involved in different other regulatory mechanisms, such as: Piwi-interacting small RNAs (piRNAs)[11], promoter-associated small RNAs (PASRs)[12], small RNAs positioned at the slice site (spliRNAs)[13],transcription initiation RNAs (tiRNAs)[14] and small interfering RNAs (siRNAs)[11].

1.2.2 - MicroRNAs

The microRNAs (miRNAs or miRs) are a family of small noncoding RNAs (19÷25 nucleotides) that are involved in the recognition of mRNA and in the inhibition of its translation (Figure 1.3).

There are two different mechanisms of production of miRNAs inside the cell: some are originated from the intronic excision during the mRNA maturation (intragenics miRNAs)[15], others come from the transcription of unique genes (intergenic miRNAs)[16].

In the former case the transcription of the miRNA is related to the factors that promote the transcription of the gene in which the sequence is embedded. During the processing of the mRNA for its maturation a microRNA-intron (Mirtron) is produced and it is processed by some enzyme to give rise to a precursor miRNA (pre-miRNA) that shows a stem-loop structure, 60÷100 nucleotides long. In the latter case, the transcription of the gene is promoted to a specific transcription factor that allows for the production of a long primary miRNA (pri-miRNA) with a peculiarly multiple stem-loop structure that is rapidly processed by a nuclear endonuclease (DROSHA) for the production of the pre-miRNAs.

In both cases the pre-miRNA produced in the nucleus is transported by exportin-5 to the cytosol in which it is processed by another endonuclease (DICER) to generate the mature miRNA (double

stranded unperfected matched structure) which is recognized by the RNA-induced silencing complex (RISC) incorporating one of the two strand. Once the complex is formed it is able to recognize the target mRNA sequence in the 3' untranslated region (3'UTR), and based on the degree of complementary to determine the fate of the target mRNA. Strong interactions between mRNA and RISC complex (perfect or near perfect matching) leads to degradation of the mRNA strand[17] while weak interactions (partial matching) leads only to inhibition of the protein synthesis without mRNA degradation. The central role in the selection of the silencing pathway is played by the "seed region" of the miRNA, which is located in the 2÷7 nucleotides region[18,19].

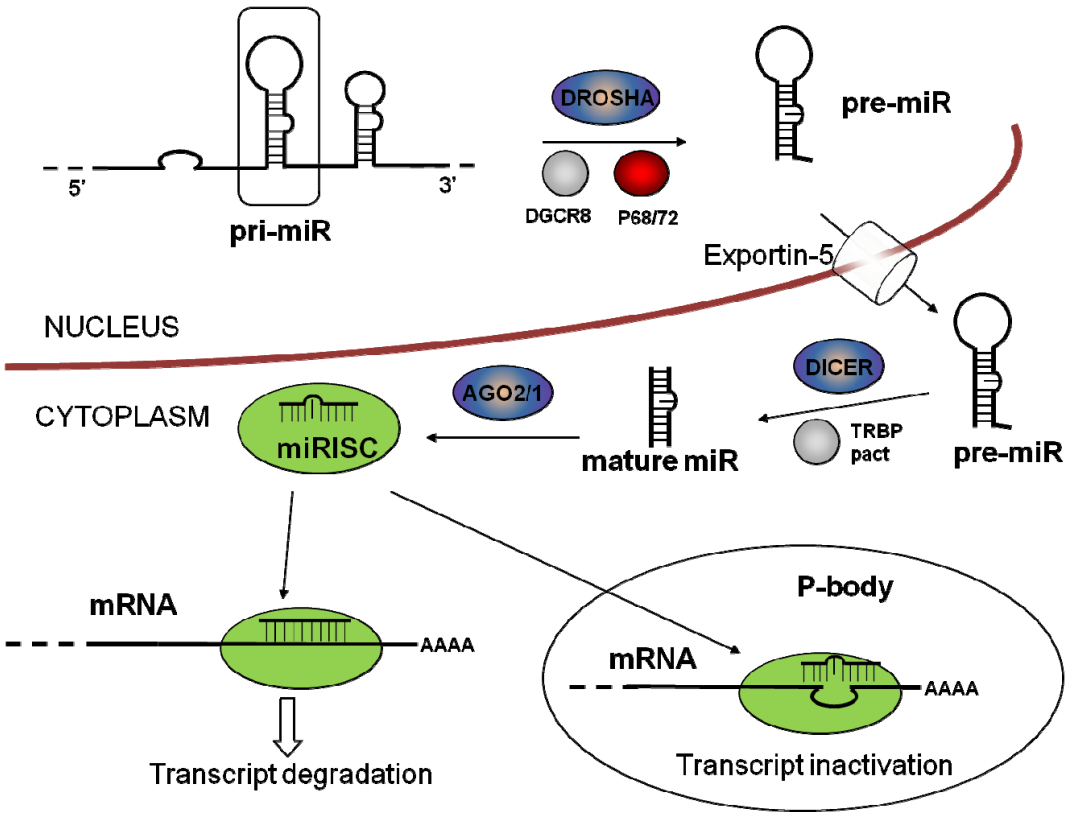


Figure 1.3: general mechanism of miR production in eukaryotic cells.

The number of know miRNAs in mammals is estimated to be in the order of 10^3, but might be much higher on the base of the knowledge coming from deep sequencing and on the evidence of cell-specific miRNAs[20]; moreover, since the 3'UTR of a single mRNA can be targeted by several miRNAs, and a given miRNA can control the silencing of multiple mRNA, the evaluation of the biological effects upon interaction with this complex machinery should be considered in order to

obtain a specific desired effect. From this point of view the target of multiple mRNA sequences should be considered as a very effective tool to obtain a strong biological effect[21].

It has been found that many alterations in miRNA expression or miRNA mutations correlate to the onset of diseases. A large-scale analysis and integration of the miRNA-disease associations at a system level offers a platform to dissect the mechanisms of miRNAs in diseases, useful for applications in diagnosis and prognosis[22]. Although the current miRNA-disease associations are far from complete, they have been listed in the Human MicroRNA Disease Database (HMDD)[23] since it has been shown that classification of tumor subtypes based on miRNA expression profiles is even more accurate than classification based on the expression profiles of protein-coding genes[24]. In cancer the role of miRNAs is dual since they can act both as oncogenes[25,26] (oncogenic miRNAs, oncomiRs), by targeting mRNA coding for tumor-suppression proteins, or as tumor suppressors[27,28], targeting mRNAs coding oncoproteins. Moreover, microRNAs have been definitely demonstrated to be involved in cancer metastasis (metastamiRs)[29,30], both promoting or blocking tumor development. By studies conducted in model systems it has be found that cancer-specific miRNA are present in the extracellular compartment, protected by exosome-like structure, suggesting the possibility that they can play an important role in the communication between cancer cells and the surrounding normal cells[31,32].

1.3 - Nucleic acid detection

Quantification, detection and imaging of nucleic acids are of broad interest and utility in biology and medicine. Detection methods are often used in the identification of pathogens[33,34] based on their DNA sequences or for the identification of single nucleotide polymorphisms (SNPs) for the discrimination of genetic variations often related to genetic diseases. From another point of view also the detection of the expression levels of various RNA sequences in the cellular medium can be correlated to the onset of diseases. Many methodologies have been reported during the years, some of them are already commercially available, while others are only at early stage of the development. In both cases the researchers play different roles: improving multiple analysis techniques, detecting small amounts of sequences, and simplifying the procedures involved in the analysis. Most of the existing methods base their mechanism of discrimination on the selective recognition of single stranded sequences, using sensing probes complementary to the region of interest to be detected[35]; for instance polymerase chain reaction (PCR) amplification, Southern and Northern blotting,

microarrays and fluorescence in situ hybridization (FISH) are based on this mechanism. But there are different systems which do not use Watson-Crick base-pairing in their mechanism of detection, some examples are polyamides able to bind to the minor groove of dsDNA with high sequence specificity[36] or engineered triplex-forming DNA and zinc finger DNA-binding proteins which show discrimination properties by major groove interactions[37,38].

The fidelity of the techniques is based on the specificity of recognition of the target sequences, while the sensitivity of the methods is related to the efficiency in the readout of the signal generated from a reporter group, that can be optical (fluorescent, colorimetric, turbidic,…), electrical (voltage, resistance, current) or mechanical (piezo-effect or bending).

These methods usually imply a different number of operations depending on the target and the type of analysis: for *in vivo* analysis only few step are required for obtaining a result, but only few qualitative information can be obtained, whereas from *in vitro* analysis multiple information can be obtained but usually require a larger number of steps. The development of new methods is therefore generally aimed to the discovery of new systems able to connect high throughput analysis with the reduction of the operative steps and with the increase of the sensitivity.

1.3.1 - Established methods

Many different methodologies have been developed during the years for the detection, quantification or imaging of nucleic acids in living cells or in treated sampled; however, a full discussion of these topics lie outside the goal of this introduction. We will propose a brief description of the milestones of this field.

Nucleic acid detection started with the invention of Southern blotting[39], a procedure in which DNA is first treated with restriction enzymes, then the fragments are separated by electrophoresis on agarose or polyacrilamide gels and “blotted” on a nitrocellulose or nylon sheet that undergoes a labeling process (radioactivity, fluorescence, …) with reporter sequences. In a similar way the Northern blotting[40] is used to detect RNA, in this case restriction enzymes are not present and the electrophoretic separation is run under denaturating conditions to avoid the formation of secondary structure that can modify the mobility of the sequences. These methods are considered the standard references in nucleic acid quantification.

Fluorescence in situ hybridization (FISH, Figure 1.4) is a cytogenetic technique generally used for

the *in vivo* imaging of specific DNA sequences of chromosomes by using a fluorescent probe[41]. Fluorescent probes are added to the medium containing the cells, and, after internalization, they bind to the target sequences, and finally, after a washing step to remove the excess of probe, the target sequence can be readily displayed in cells by fluorescence microscopy. This technique rises from previously developed *in situ* hybridization experiments based on the use of radiolabeled probes, allowing significant advances in resolution, speed and safety, and later paving the way for the development of simultaneous detection of multiple targets[42], quantitative analyses and live-cell imaging[43]. At first designed for the localization of DNA, it was rapidly applied also to the detection of RNA[44,45] and, with the help of software, FISH was applied to diagnostics.

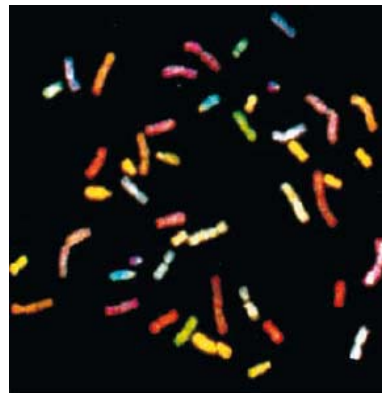


Figure 1.4: detection of 24 chromosomes by FISH (from Ref 42)

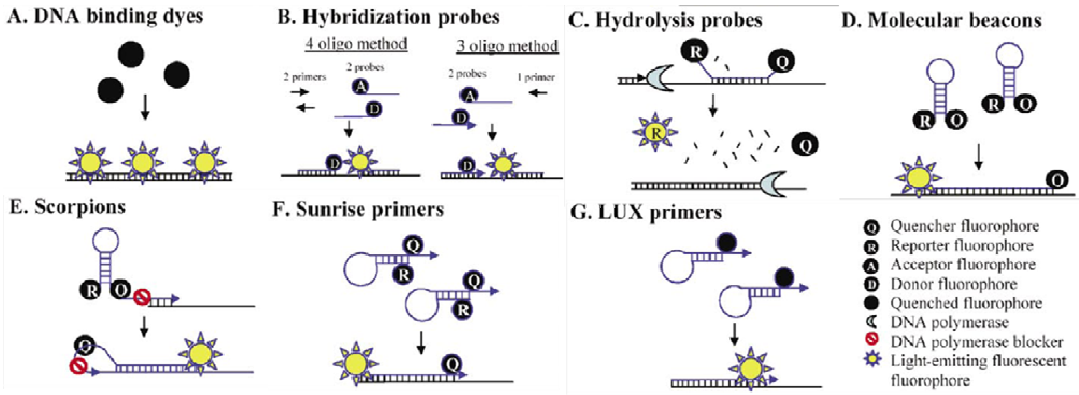


Figure 1.5: different approaches in Q-PCR fluorescence generation (from ref 47).

The use of PCR (and reverse transcriptase PCR for RNA) protocols[46] are the basis for most of the *in vitro* detection methodologies. In these techniques that allow the exponential amplification of a short sequence of a DNA target, or cDNA in case of RNA, the sequence is replicated by the action of a DNA polymerase that starts the replication from the tracts recognized by two different DNA primer sequences (each located on one of the two strands); the process in based on the iteration of temperature cycles consisting of denaturation, annealing and elongation steps. The PCR product can then be analysed by agarose gel electrophoresis with ethidium bromide staining. This technique leads only to a semi-quantitative analysis. However this problem has been overcome by the introduction of the quantitative (or real-time) PCR (Q-PCR), in which a fluorescent signal is produced step by step in relation to the amount of DNA present in solution, and correspondingly to

the starting amount of target[47,48]. A schematic representation of the different methods used for the generation of the fluorescent signal is shown in Figure 1.5.

Molecular beacons (MBs, Figure 1.6) are generally composed of an oligonucleotide containing a fluorophore at the 5'-end and a quencher at the 3'-end; however in recent years a different pattern of MBs without quencher has been proposed[49,50]. In MBs, the DNA strand, that contains two complementary antiparallel tracts at each ends which keeps the two terminal units in close proximity (7-10 nm), but once the target sequence has been recognized by the MB, the two extremities are drifted apart with the production of a "switch-on" fluorescence signal without the need to separate the probe-target hybrids from excess probes[51,52]. A broad variety of applications has been proposed using MBs, including SNP detection by carefully designing the sequence in order to have a proper duplex stability[53,54], by constructing self-reporting arrays[55] or *in vivo* analysis[56]. MBs are usually applied for the detection of a single sequence but combination of different systems with different fluorescent units can be applied in multiple target detection[57]. This approach can be followed by using different couples of excitation/emission wavelengths; however, the use of wavelength-shifting molecular beacons can open the possibility to use a single wavelength for the excitation of different molecular systems. In the first example proposed, the authors[58] used DABCYL as the quencher and luciferase/Texas red as harvester and emitter fluorophore.

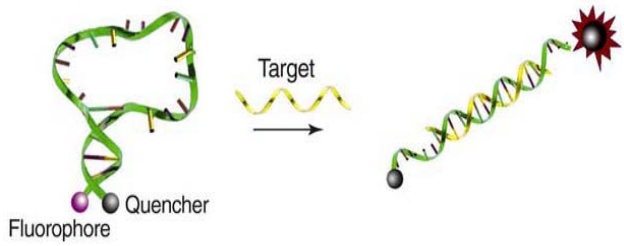


Figure 1.6: general mechanism of MB (from Ref 50)

DNA microarrays are now one of the most widely used tools in functional genomics[59]. These devices contain libraries of compounds automatically spotted, synthesized or printed on a surface (glass, silicon or plastic) with a specific 2D position for each given sequence. The basic principle of detection is again related to the ability of the nucleic acids in the sample solution (DNA or RNA) to selectively bind to a complementary sequence linked to the slide during a hybridization step that precede the detection step performed by laser excitation of specific dyes. Compared to the former techniques, microarrays have opened the possibility to simultaneously detect and quantify thousands of targets, but require functionalization of the sample with a fluorescent reporter group.

1.3.2 - Emerging technologies

On the basis of the concepts emerged from the above described methods, research has moved toward the development of new techniques based on new mechanisms of detection, label-free systems, reduction of the operator-dependent steps, miniaturization of the devices in order to reduce the amount of sample needed and the possibility to obtain a single cell evaluation or to avoid the use of PCR techniques that can produce artifacts[60,61].

Some examples of label-free analysis were proposed by using the target strand for template reactions between two different reactive probes in order to generate a fluorescent signal for the detection of different target, such as dsDNA structure or mRNA evaluation in cells[62,63].This technique is based on the increment of the effective concentration of the two reactive species after probe hybridization with the complementary strand. Another interesting approach was the development of forced intercalation probes (FIT probes)[64], in which a nucleobase was substituted with thiazole orange as a universal fluorescent nucleobase able to discriminate between perfect matching of single mismatched target by variation of the fluorescence emission of the probe in relation to the conformation of the double strand, with a successful application to allelic discrimination in Q-PCR[65].

Microfluidic technologies, also referred to as lab-on-a-chip or micro-total-analysis systems (μTAS), is the adaptation, miniaturization, integration, and automation of analytical laboratory procedures into a single device or "chip"[66,67]. These devices are designed to perform multiple tasks, from sample treatment until target detection, in a fully automated fashion from small amounts of sample, standardizing the procedure and with a fast displaying of the results in an easy-to-interpret readout, avoiding the operator-dependence of the results. Examples of this application are: i) the design of a credit card-sized chip that is capable of extracting and concentrating nucleic acids from milliliter aqueous samples and performing microliter chemical amplification, serial enzymatic reactions, metering, mixing, and nucleic acid hybridization[68]; ii) a microfluidic system that can perform cell lysis and/or DNA amplification[69,70]; iii) a microfluidic cartridge that could automate the fluidic handling steps required to carry out a microarray gene expression study of human leukemia[71]. Chips able to perform micro-Q-PCR were also developed[72].

For the quantification of DNA and RNA in amplification-free assays different approaches were followed during last years. In a first approach colorimetric variation of the scattering properties of gold nanoparticles[73] (AuNPs, Figure 1.7a), linked to different probes complementary to proximal

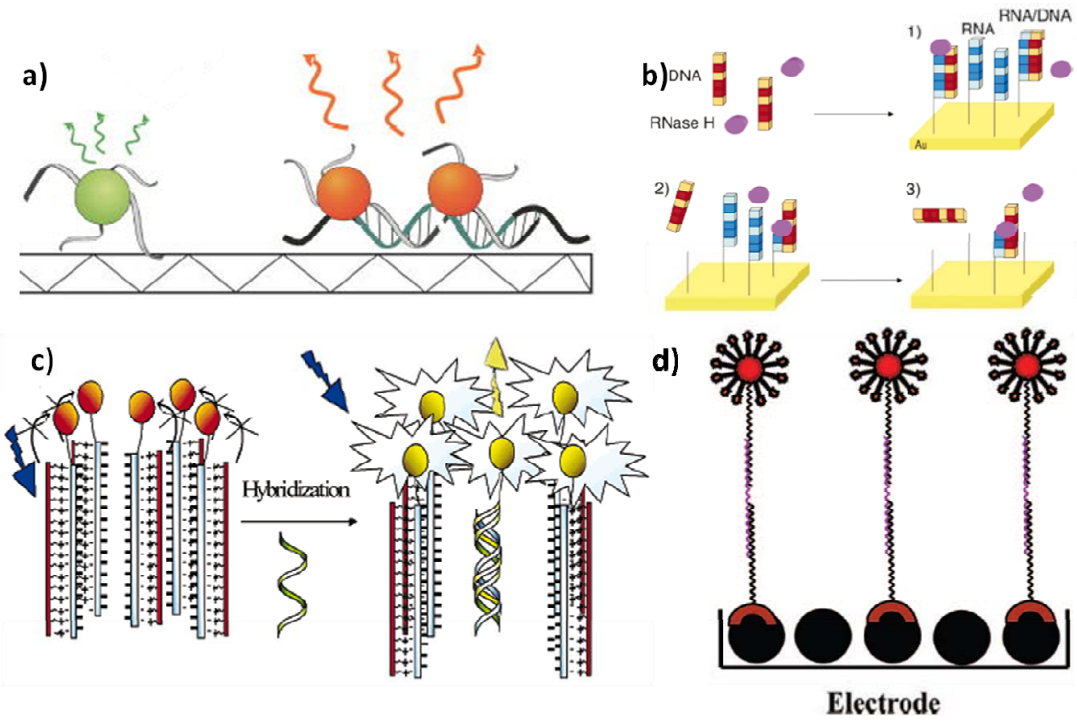


Figure 1.7: graphical representation of the mechanisms involved in PCR-free detection. a) scattering of AuNP; b) SPR variation after RNase degradation; c) enhanced fluorescence after conformational variation; d) electrochemiluminescence bar coding(from Ref 73-76).

regions in the target, was used for the detection of genomic DNA, at a 33fM concentration. Better results (1fM) were obtained using surface plasmon resonance (SPR) in combination with the activity of RNase H (Figure 1.7b)[74], in this case variations of the signal were induced by the degradation of a gold RNA microarray after the recognition with the target DNA. Further improvements (3zM and SNPs detection) were achieved taking advantages of the ability of a cationic polythiophene to induce FRET over a fluorescent marker of the sensing DNA probe after conformational changes induced by the recognition of the target DNA (Figure 1.7c)[75]. Electrochemiluminescence-based bar coding (Figure 1.7d)[76] was proposed for the detection of GMO by using a complex system involving the formation of a three stranded architecture in which the target DNA is recognized by two different probes, one attached to a AuNP bearing a ruthenium based fluorophore and another linked to a biotin moiety, used as anchoring point for the connection to streptavidin-coated magnetic beads, lying deposed on the electrode used to induce the electroluminescence. A schematic representation of the systems described is shown in Figure 1.7. Despite the great potential of these techniques, the continuous lowering of the target amounts of DNA to detect has opened the necessity to provide systems able to perform single molecule detection with significal confidence in the readout. This is often achieved bymicrofluidic devices,

since allowing the reporter molecules to flow through the detection volume either by pressure-driven or electrically-driven means could greatly enhance the encounter rate and hence reduce the analysis time dramatically[77].

1.3.3 - miRNA detection

miRNAs are a difficult target for the detection mainly for their short sequence, but also for their low expression and the extreme variability in their amount in cells, since they are tissue and developmental stage depending. Every method should be evaluated carefully before its application to avoid artifacts.

Northern blotting is a valuable method for the discrimination of miRNAs against their precursor based on their sequence length. The final detection step with reporter probes has to be carefully designed, on account of the small difference in both length and sequence between different miRs and only the use of nucleotide mimics with better performances in term of sequence specificity (locked nucleic acids[78], LNAs, or PNAs) can really help in the correct evaluation[79]. Apart from the high quality of the results, this technique is time consuming and requires large amounts of starting RNA, due to its low sensitivity. miRNA microarrays were efficiently used for monitoring the whole miR expression in human and mouse tissues[80], despite the difficulty to simultaneously optimize the hybridization conditions for each sequence and the subsequent possibility to obtain false positive. As in the previous case the use of DNA mimics in the microarray device can enhance discrimination and detection[81]. Localization within the cells of a chosen miR was successfully achieved by the use of an in situ hybridization approach[82], with better results in combination with the use of LNA probes[83,84]; this method was used for the precise location of a single target but the extension to multiple targets is not easily achievable[85]. Recently a new approach that used MBs was applied to *in vivo* imaging[86], opening the possibility for the future detection of miR biogenesis and function in living animals.

Alongside with the application of the existing methods for the detection of this precious target new protocols were developed. An example is Q-PCR in which 3 different approaches were set up for miR evaluation, the main issue being the elongation of the primer sequence for an easier monitoring of the reaction and the introduction of an universal binding site: i) in a so called primer-extension method the elongation is made by a tailed primer[87]; ii) in a second method a stem-loop primer was used[88]; iii) in the third case a poly-adenine was enzymatically attached to the 3'-end[89]. In all cases

the approaches were applied to the simultaneous detection of different miR sequences, showing a wide detection range, the ability to discriminate between similar sequences and small amount of total RNA required.

Finally, new methods were also specifically developed for this class of molecules. An invader miR assay[90] was proposed for the evaluation of both miR and siRNA, based on a two steps reaction that involves the preliminary cleavage of a flapping part of the probe by a specific nuclease (cleavase). The cleaved fragment undergoes to a second recognition event on a secondary reaction template and to a second cleavage, induced on a FRET nucleotide for the separation of the fluorophore from the quencher with the subsequent generation of a fluorescence signal. In the rolling circle amplification detection[91] a padlock probe, a special linear probe, can be circularized by the action of a ligase after the formation of a perfectly matching miRNA abduct with perfect discrimination of fully matched or single mismatched sequences. In a complementary approach, using the radiolabeled splinted ligation[92], a ligase can connect a labeled RNA sequence to the miR sequence after the perfectly formation of a perfectly matched template with a bridge nucleotide. The last two procedures were coupled with an electrophoretic technique for the quantification of the target product.

1.4 - Gene therapy

The possibility to rapidly synthesize DNA, RNA or mimics with a predefined chosen sequence opened the possibility of a large number of applications. Beside the previously mentioned detection methods, therapeutics applications have been extensively studied[93]. A great variety of steric block methodologies were proposed for anti-gene (DNA block)[94] or anti-sense (RNA block)[95] targeting.

The first appearance of this approach in therapeutics arose in 1978 with the use of a synthetic oligonucleotide as antisense probe for blocking a specific gene expression in virus without alteration of other gene expression[96]. In the steric block approach, the formation of a stable duplex structure that inhibits the correct recognition of the target by the protein in the protein machinery is achieved, thus interfering with different events such as mRNA translation[97], nuclear splicing[98] or polyadenylation[99]. Together with this first approach, an RNase H induction can be often exploited to obtain strong RNA inhibition. This mechanism is based on the ubiquitous presence of this enzyme in the cellular media that selectively recognizes RNA/DNA adducts and usually degrades the RNA strand in the middle of the complex in a catalytic way rather than in a stoichiometric mode

(as required in the block mechanism).

For a good *in vivo* activity of the designed probes different aspects have to be taken in consideration: stability and selectivity of the interaction, biological stability, cellular uptake and toxicity. Above all, the complex stability is the most important and first requirement for obtaining any kind of effect: in fact, the probe must have sufficient length to form a stable complex in the cellular medium at 37°C, avoiding the possibility of formation of secondary structures that may hinder complex formation. Another important issue is the selectivity of the recognition of the target among all other sequences present in the cellular compartment, in order to avoid aspecific interaction with non complementary sequences, which may introduce the possibility of side effects. Moving from the *in vitro* optimization to the biological applications, the stability of the probes against the action of various enzymes should be considered, in order to avoid the degradation and to confer a long lasting effect to the treatment. Mandatory in cellular applications is internalization, a task difficult to obtain by the probe alone but easily achieved by conjugating the probes with specific carriers that allow cellular uptake and direct delivery into a specific type of cells or cellular compartments by receptor-mediated transport. Finally, for animal and human applications the systems must have no toxic effects.

1.4.1 - RNA interference

In early 1990s for the first time the unexpected result of gene suppression after the attempt to stimulate the overexpression of that gene by introduction of an artificial chimeric gene was described[100]. After this first appearance different studies were directed to the identification and description of this strange behavior that was found to be a general issue in plants[101] and in the worm *C. elegans*[102] and later on in many different organisms. This mechanism, named RNA interference (RNAi) was then showed to be a really powerful tool for gene suppression regardless of the use of either small dsRNA or small hairpin RNA (shRNA, generally introduced with vectors, for a long lasting effect[103]), and regardless of the kind of target to be suppressed. After introduction of this sequence in the cell, these RNAs are cleaved by the DICER complex to generate perfectly matched 21÷23nt long dsRNA with two nucleotides 3'-overhang each side, called small interfering RNA (siRNA), that are subsequently processed by the RISC complex which incorporates only one of the two strands (the antisense or guiding strand) and discard the other (the sense or passenger strand). The RNA activated enzymatic complex is finally able to recognize the target mRNA and to induce a

cleavage of the phosphodiester bond in the middle of the target sequence recognized[104]. This machinery was found to be easily triggered by all hybrids having similarity or which can induce the production of siRNA, with high tolerance to the introduction of chemical modifications of the passenger strand with a less extent for the guiding strand[105]. Studies on this mechanism permitted the discovery of the endogenous miRNA machinery.

Since process has the possibility to control the expression of every mRNA in cells, it has achieved great importance as a possible strategy for the treatment of different kinds of diseases, however, moving from cellular systems to animals or humans the stability inside the organism should be long enough to allow uptake inside the target cells [106]. Different approaches were proposed to address these problems[107]. Because of the necessity to maintain a possible recognition of the RNA by the DICER and RISC enzymatic machineries, drastic variations of the RNA structure are not allowed. Since it was found that the 2'-hydroxyl group of the ribose ring was not necessary for potent gene silencing[108], modifications at that site with beneficial effects on stability, and, in some case improving gene silencing were successfully used[109]. Other studies showed that the insertion of DNA nucleotides in the strand can lead to a reduction of both immune responses and off-target effects without loss of activity[110].

Beside the importance of the resistance to the biological conditions, the other main challenge for these types of drugs is to enter inside the cells to trigger the RNAi machinery. Naked siRNAs have been shown to have efficiently *in vivo* delivery[111,112], although in many cases the mechanism is still unknown[113]. For these reasons naked siRNAs were readily applied to clinical trials limited to some local administration (see Table 1.1 for a summary), and after few years excipient formulations (lipids, polymers, cell penetrating peptides, aptamers, ...) were also proposed for the clinical evaluation, due to their ability to readily overcome the stability and uptake challenges for a broader spectra of applications.

1.4.1 - miR as target for gene therapy

As previously described variation in miR content in cells can lead to the onset of diseases or cancer, being related to an over- or under-expression of a single or a related family of miR sequences.

The miRNA levels in cells can be easily increased by the introduction of short double stranded RNA (linear or hairpin) that activate the small interfering RNA mechanism or by the introduction of

Table 1.1: clinical trials with siRNA (form Ref 113).

Product owner	Technology originator	Clinical phase	Clinical outcome	Delivery technology	Administration	Indication	Target
Alnylam Pharma Inc.		Ph II (2007–09)	Well tolerated, antiviral activity	Naked siRNA	Nasal, inhalation	Respiratory syncytial infection	Respiratory syncytial virus
Alnylam Pharma Inc.	Tekmira	Ph I (2009–11)	Well tolerated, antitumor activity	SNALP	I.v. Infusion	Liver cancer	endothelial growth factor, Kinesin spindle
Alnylam Pharma Inc.	Tekmira	Ph I (2010–12)		SNALP	I.v. infusion	Amyloidosis	Transthyretin
Alnylam Pharma Inc. / Tekmira		Ph I (2009–10)	Terminated due to immune stimulation	SNALP	I.v. infusion	Lipid disorders	Apolipoprotein B
Tekmira / Alnylam Pharma Inc.		Ph I (2010–12)		SNALP	I.v. infusion	Cancer	Polo-like kinase-1
Merck, Allergan Inc.	Sirna Therapeutics	Ph II (2004–07)	Well tolerated, potentially unspecific clinical effects	Naked siRNA	Intravitreal	Wet AMD	Vascular endothelial growth factor
Opko Health Inc.	Acuity Pharmaceutics	Ph III (2007–09)	Well tolerated, clinical endpoint not met	Naked siRNA	Intravitreal	AMD	
Quark Pharma Inc. / Atugen AG / Pfizer		Ph II (2009–11)	Well tolerated, improved vision	Naked siRNA	Intravitreal	Wet AMD	Hypoxia-inducible gene RTP801
Quark Pharma Inc. / Silence Therapeutics	Atugen AG	Ph II (2008–12)		Naked siRNA	I.v. injection	Acute renal failure after kidney transplantation	p53
Quark Pharma Inc.		Ph I (2010–13)		Naked siRNA	Intravitreal	Eye diseases	Caspase 2
Alnylam / Calando		Ph I (2008–11)	Well tolerated, accumulation in tumor, reduction of mRNA level	Rondel	I.v. infusion	Cancer	M2 subunit of ribonucleotide reductase
Silence Therapeutics	Atugen AG	Ph I (2009–11)	Well tolerated, stabilization of disease	atuPLEX	I.v. infusion	Cancer	Protein kinase N3

the pre-miR precursor[114], while direct targeting of the miR for the reduction of free miR content (as putative disease effector) can be achieved by sequence specific binding of oligonucleotides, which can lead to multiple effects based on the ability of a single sequence to control the fate of different mRNA.

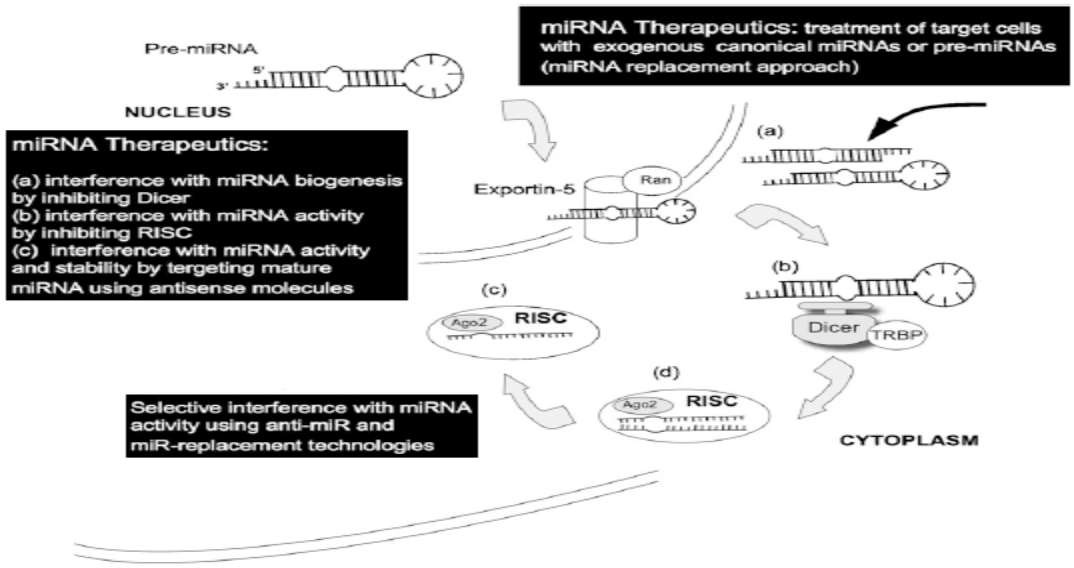


Figure 1.8: miRNA therapeutic approaches.

Chemically engineered oligonucleotides, termed “antagomirs”, are efficient and specific silencers of endogenous miRNAs in mice. Silencing of microRNAs *in vivo* with 'antagomirs' is a very interesting strategy, supporting studies on the involvement of miRs in gene expression and providing new tools for non-viral gene therapy[115]. miR-21 over-expression was shown to be associated to several solid tumors (lung, breast, colon, gastric and prostate carcinomas and endocrine pancreatic tumors) as well as to cholangiocarcinoma and glioblastoma, since it is correlated to inhibition of apoptosis[116,117]. Knock down of miR-21 by LNA oligonucleotides, associated with neural precursor cells (NPC) expressing a secretable variant of the cytotoxic agent tumor necrosis factor–related apoptosis inducing ligand (S-TRAIL), was shown to have high antitumor activity[118]. MiR-122 is another very important target: it is a well characterized liver-specific microRNA exhibiting particular therapeutic interest, since it is related to cholesterol levels in plasma, and it has been shown not only to facilitate Hepatitis C RNA replication[119], but also to be up-regulated in HIV-1 infected cells[120]. It has been demonstrated that intravenous administration of antagomirs against several miRs, including miR-122, resulted in a marked reduction of the corresponding miRNA levels in liver, lung, kidney, heart, intestine, fat, skin, bone marrow, muscle, ovaries and adrenals[121]. Silencing of endogenous miRNAs by this novel method is specific, efficient

and long-lasting; as an example, a five days administration of anti-miR122 induced a reduction of blood cholesterol levels, which was detected after three weeks, but lasted seven weeks[122]. These findings show that antagomirs are powerful tools to silence specific miRNAs *in vivo* and may represent a therapeutic strategy for silencing miRNAs in disease without strong evidence of toxicity. Interest on this target has been recently boosted by a report which showed increased resistance to chronic hepatitis C virus (HCV) in primates, achieved by targeting miR122 with LNA, with long-lasting suppression of HCV viremia, with no evidence of viral resistance or side effects in the treated animals[123]. Other highly expressed miRs (such as miR-142-3p) are very important tools in gene therapy protocols since specific targets can be inserted in gene constructs in order to suppress toxicity associated to viral vectors or to inhibit immune response against a transgene, however they are not easily used as targets for specific inhibition of a pathological state[124].

Despite the greater selectivity of complexation and inhibition of miR by LNA and LNA/DNA or LNA/OMe chimeras compared to the use of ordinary oligonucleotides it has been proved that their mechanism of down-regulation of target activity is only related to the sequestering of the target sequence without activation of the RNase H complex degradation explaining the role of the continuous administration required for obtaining good results[125].

Nucleic Acid Mimic: PNA

1.5 - Peptide nucleic acids (PNAs)

The peptide nucleic acids (PNAs) are a class of artificial nucleic acid mimics in which the sugar-phosphate backbone (ribose or deoxyribose) is substituted by a poly-amidic N-(2-aminoethyl)glycine backbone (Figure 1.9), while the nucleobases are linked by a carboxymethyl spacer[126].

As other nucleic acids mimics these derivatives obey to all pairing rules (both Watson-Crick and Hoogsteen), and, lacking the poly-anionic backbone, they present a higher affinity and selectivity in the recognition of antiparallel complementary strands[127] via formation of helical double o triple stranded structures. The stability of these structures is so high that can induce the dissociation of DNA:DNA duplexes to form DNA:PNA duplexes in a process called strand invasion[128].

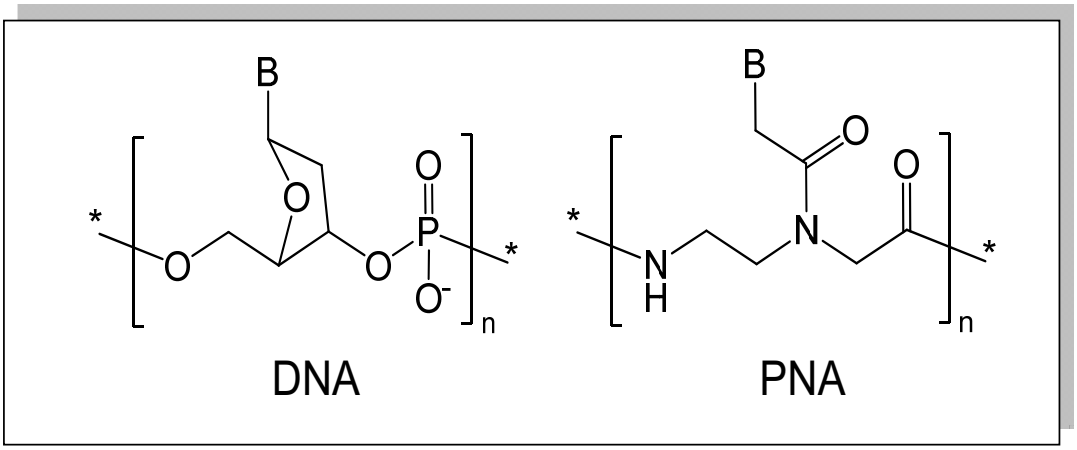


Figure 1.9: schematic structures of DNA and PNA.

On account of their artificial backbone these derivatives are not recognized by natural enzymes (both nucleases and peptidases) conferring high stability to these molecules in biological fluids[129]. On the other side their polyamidic nature confers great stability to a broad variety of chemicals and allows to obtain PNA strands using standard protocols of peptide solid-phase synthesis[130].

1.5.1 - PNA structure modification

The first PNA structure synthesized in 1991 was designed, using DNA as a model, to maintain unaltered the distance between two consecutive nucleobases and to have a flexible backbone that could allow an easy rearrangement around the more rigid DNA strand thus facilitating the duplex formation.

Despite the good performances showed by these derivatives in term of target recognition, the first model proposed presented several drawbacks[131]: low target specificity in the recognition of DNA and RNA, low water solubility, low cellular uptake and the tendency to show self-aggregation.

In the following decades more complex structures were proposed using the PNA scaffold as a starting point, and modification were introduced to overcome these problems. In this context the role of the synthetic organic chemistry was focused on introducing modifications of the standard backbone, on building new backbone structures, on the modification of the nucleobases and on the expansion of the synthetic strategies for building these polymers.

A great synthetic effort was dedicated to the modification of the backbone structure by introducing

functional groups by substitution of the aminoethyl or glycine synthons with chiral variants (Figure 1.10) or by substitution of the flexible scaffold with more rigid cyclic structures (Figure 1.11). A complete discussion of the backbone variations and of the relative consequences on the biological properties is reported in a recent review[132].

A first example of modification of the backbone was the introduction of a substituent on the C2 position (thanks to the availability of enantiopure α-amino acids, Figure 1.10a); the best results were obtained by introducing D-argine[133] or D-lysine[134,135] side chains, since it increased the water solubility and induced a preferred right handedness which favored the selectivity of complexation. After these first modifications, other structures bearing a chiral center in C5 position were proposed (Figure 1.10b)[136] or in both C2 and C5 position (Figure 1.10c)[137]. It was found that the configurations of the amino acids synthons which favored a right-handedness helicity in PNA, were also favoring DNA complexation.

In order to give rise to more rigid structures which could induce the preorganization of the PNA strand, thus favoring the complexation of natural DNA or RNA, a broad variety of structures containing cyclic moieties were proposed[138].

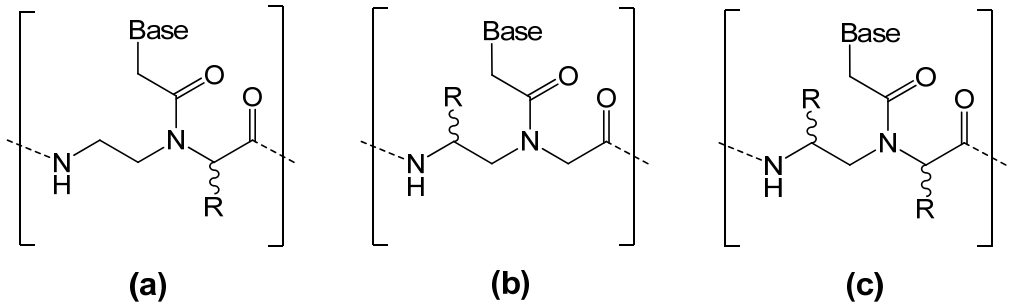


Figure 1.10: generic structure of modified chiral PNAs.

For example, Ganesh's and, later on, Appella's groups proposed substitution of the ethylenediamino moiety by cyclic diamine containing cyclopentane or cyclopropane rings (tcypPNA[139,140] and tcprPNA[141], Figure 1.11 a and b respectively). The insertion of these derivatives in chimera structures (containing both unmodified and modified PNA monomers) showed an increase in DNA-PNA duplex stability with an additive effect according to the number of modified monomers inserted.

Other research groups introduced rigid structure by substituting the glycine units with a proline derivative (POPNA, Figure 1.11c-d)[142] alternated with β-amino acids. In this case the first generation of compounds showed a preferential conformation that inhibited the complexation of the

target strands, whereas the introduction of a second ring in the ethylenediamino moiety (2-aminocyclopentanoic acid) drastically increased the stability and selectivity of the discrimination.

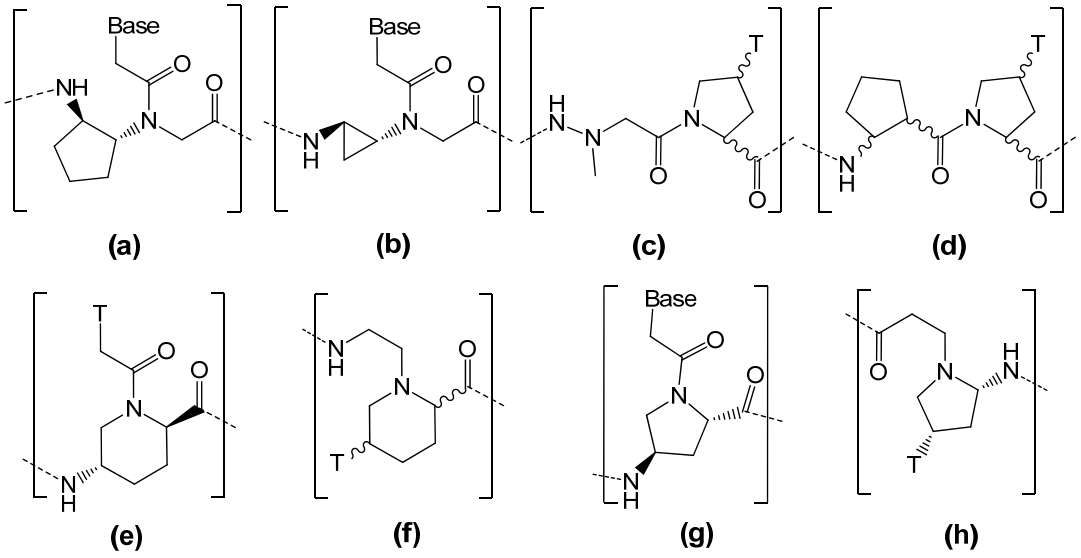


Figure 1.11: structure of modified PNA with increased structural rigidity.

In order to achieve discrimination between DNA and RNA, monomers based on pipecolic acid[143,144] (Figure 1.11e, f) or on a proline derivative[145] (Figure 1.11g) were reported; these structures were designed for an increased selectivity in the complexation of DNA strands whereas extended backbone based on pyrrolidine[146] (bepPNA, Figure 1.11h) and cyclohexyl PNA[143] were shown to selectively target RNA over DNA.

1.5.2 - Nucleobase modification

In many studies on natural nucleic acids or their mimics the role of the nucleobases was thoroughly explored, in order to design new systems with improved stability and selectivity in the recognition process or to introduce particular properties to the probes without altering the natural conformation of the complex formed[147].

A possible approach to the synthesis of modified nucleobases was the design of derivatives able to replace every natural nucleobase without altering the original stability of the complex formed (universal bases), so that they could be used in sensor systems for the detection of degenerate gene

sequences[148]. These bases do not present groups able to interact with the complementary strand via H-bond but only through π-π stacking.

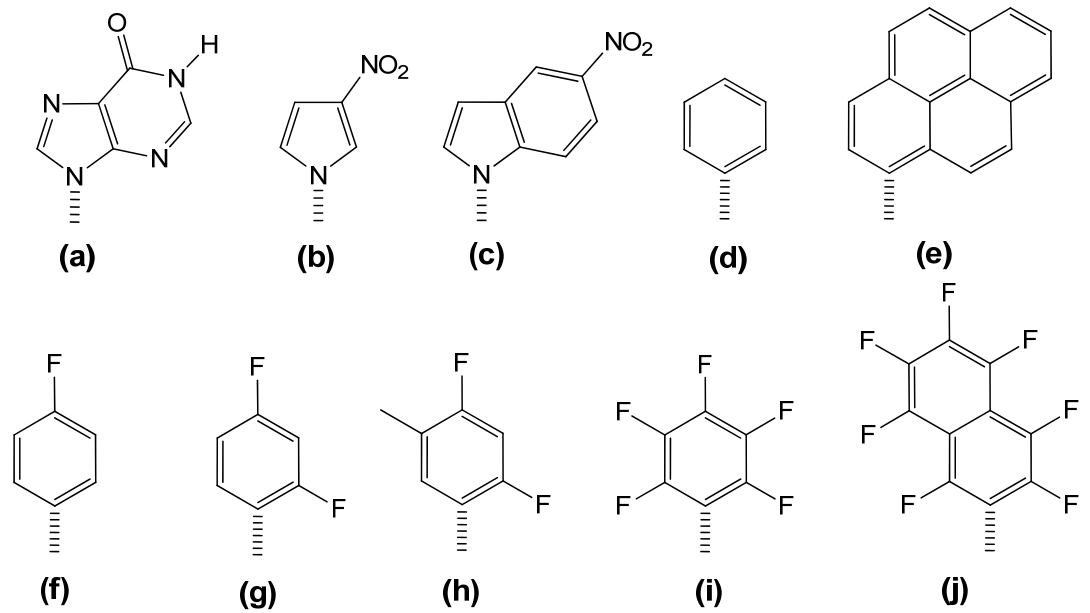


Figure 1.12: universal nucleobases; a) inosine, b) 3-nitropyrrole, c) 5-nitroindole, d) benzene, e) pyrene, f-j) fluoro derivatives.

Unfortunately, all systems proposed (some examples are reported in Figure 1.12) present a strong destabilization of the duplexes, even more than in the presence of a mismatch, with reduction of this effect by the increasing of the aromatic part. Furthermore, the formation of duplexes often showed a discrimination of the "complementary" base, and was affected by the neighboring environment. This effect of destabilization was found to be more effective with PNA than with other systems due to their higher sensitivity in mismatch discrimination.

A more successful approach was the increase the aromatic portion to stabilize the complex formed

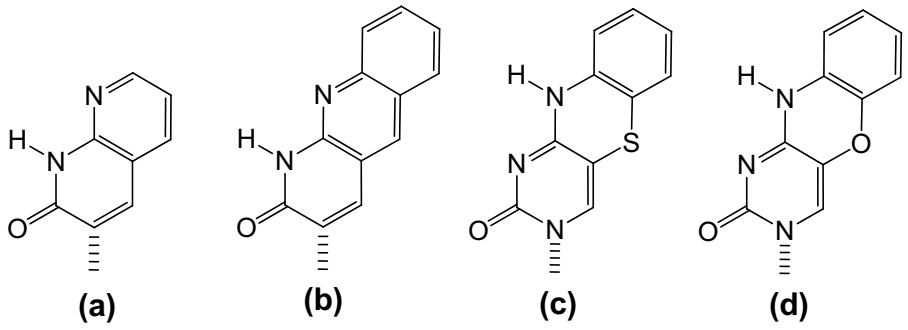


Figure 1.13: pyrimidine derivatives with extended aromatic system[149]; a-b) 1,8-naphthyridin-2-one derivatives, c) thiophenoxazine, d) phenoxazine.

was followed without the alteration of the H-bond donor/acceptor geometry, which allowed to obtain good results in term of complex stability. This can be done using extended bases as pyrimidine base analogs (Figure 1.13).

The systems proposed, with the additional moiety in the major grove, offered also the possibility of inducing additional stabilizing interactions, by placing different units in adjacent positions or in the proximity of purine nucleobases. Moreover, some of them (Figure 1.13c is an example) have fluorophoric properties than allow their use as reporter groups in the recognition of target sequences[150].

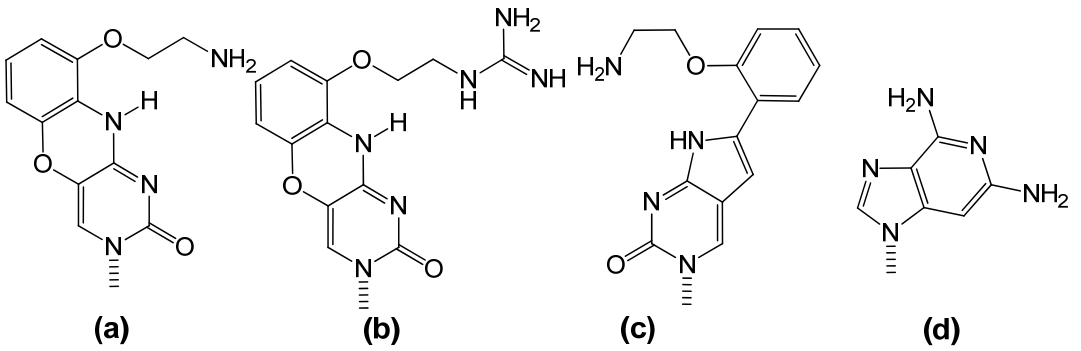


Figure 1.14: analogs with additional interactions; a-b) phenoxazine derivatives, c) phenylpyrrolocytosine derivative, d) 2,6-diaminopurine.

On the other hand, the insertion of additional sites able to interact with the complementary nucleobase through H-bond was a convenient strategy to increment the complex stability. G-clamp derivatives (Figure 1.14a and b) were designed on the phenoxazine scaffold (Figure 1.13d) to obtain an additional interaction with the N7 of the complementary guanine, able to increase the stability of the complex by 24°C in the melting temperature without losing in selectivity of recognition[151]. 2,6-diaminopurine derivatives (Figure 1.14d) were developed to reinforce the interaction with the complementary thymine by increment of 4÷6 °C/modification in the melting temperature[152], Their use was particularly important in combination with the use of 2-thiouracil in double strand invasion studies[153], due to the negative steric interaction between the two modified nucleobases, that avoided the formation of a stable PNA:PNA duplex, thus forming a pair of pseudo-complementary bases.

Particular interest was also placed in modified nucleobases showing fluorescence properties (Figure 1.15), for the synthesis of systems able to perform target detection without the necessity of labeling the target strand. In this field two different approaches were followed: modification able to maintain the recognition ability, or substitution of the original nucleobases with a fluorophore as a “universal nucleobase”.

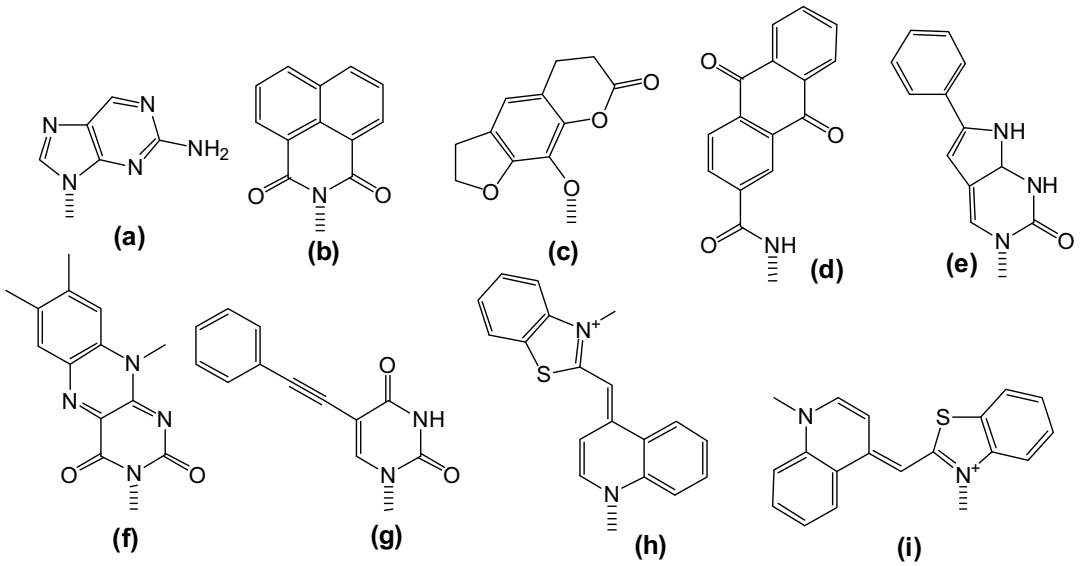


Figure 1.15: modified nucleobases with fluorescence properties; 2-aminopurine, b) naphtalenediimide, c) psoralen, d) anthracedione derivative, e) phenylpyrrolocytosine derivative, f) flavin, g) modified thymine, h-i) thiazole orange.

In the former case the groups introduced were designed to reside in the major groove with the aim of reducing the steric hindrance; the latter type of modifications was used to obtain forced intercalating probes (FITs) in which the interaction with the complementary strand induces a strong increment in the fluorescence of the reporter units; this effect allows to easily discriminate mismatches in adjacent positions. An important example was the use of the thiazole orange (Figure 1.15h, i) that stabilized the interaction by both π-π stacking and electrostatic interactions, and presented a strong increment in the fluorescence intensity (up to 20 fold) after hybridization with the perfect match counterpart, while mismatch induced a strong quenching[154].

1.5.3 - PNA in molecular biology and medicine

On account of their characteristics in DNA and RNA recognition, PNAs were chosen as important candidates for application in diagnostic and gene therapy.

PNAs were applied in different fields of detection[155]: microarray systems[156], FISH[157,158], PCR-clamping, chromatographic methods, southern or northern blotting[159,160], surface plasmon resonance[161,162] and beacon systems[163]. In microarrays technology, different PNA probes were covalently linked to a glass surface, as for DNA-microarrays, and, thanks to their higher selectivity and stability of their complexes in a broad operative condition range, they were efficiently used in SNP detection and species identification[164]. In the case of the PCR-clamping, the use of PNA

probes was developed to compete with the DNA primers in the recognition of the target region. In this contest the role of the PNA sequence is to bind to the target region for PCR primers and to stop the polymerization reaction. Once a mutation in this region is present, the destabilization in the PNA:DNA complex is greater than in the DNA:DNA case, so that the primers are able to recognize the binding site and start the reaction[165]. Chromatographic methods such as HPLC[166,167] and CE[168] were efficiently used for the discrimination of point mutation in DNA strands. The addition of PNA to the solution to be analyzed allow the formation of duplex structure that induced a difference in the retention times of perfectly matched sequences and single mismatched ones.

The development of PNA modification that reduced their drawback for the application to the biological compartment (low water solubility and low cellular uptake) opens the possibility to move the application of these systems from diagnostics to therapeutics. In this context PNA were efficiently used to complex different target in protein expression pathways (Figure 1.16).

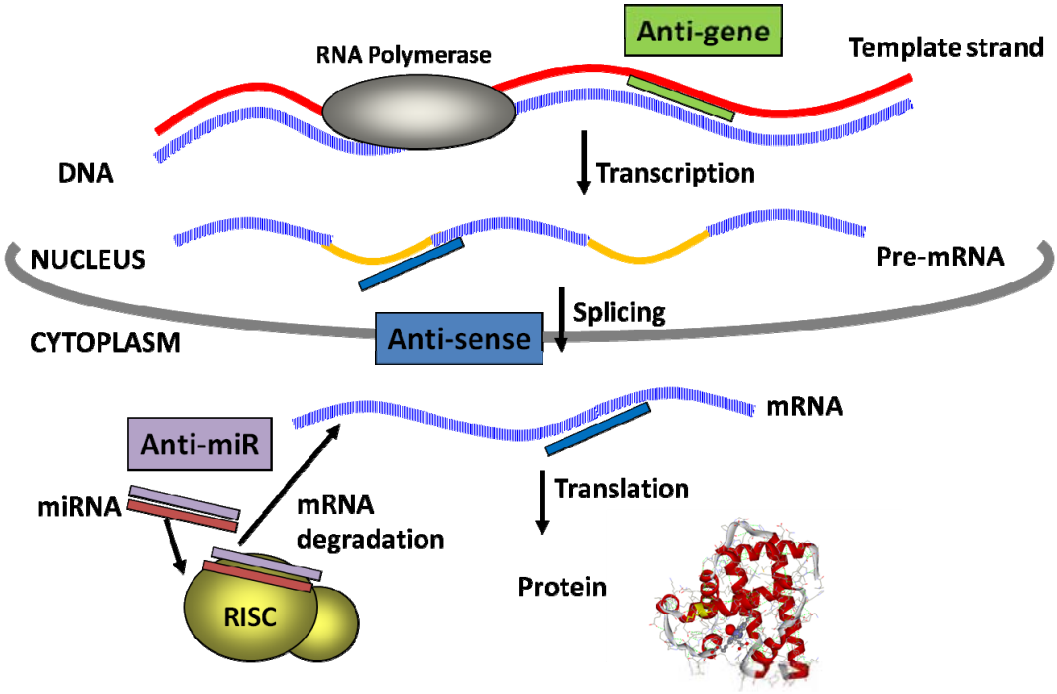


Figure 1.16: schematic representation of PNA action in genetic treatments.

One of the most important applications of PNA was their use as anti-gene drugs, in which the PNA strand was designed to bind to a target DNA sequence in correspondence of binding regions of the transcription factor, avoiding complexation of the strand by polymerases. To obtain an effective result, the probes must be designed to effectively penetrate inside the nucleus and then invade the double stranded DNA; these requirements are difficult to obtain in a general way without using

specific conditions or modification. Examples of the effectiveness of this kind of treatments were reported in targeting the progesterone gene promoter[169], the inhibition of the transcription of the HIV transfected genome[170], in the inhibition of MYCN expression in cells[171] on in animals[172] and in the inhibition of the DNA replication[173,174].

Another interesting target in cells is represented by RNA. The application of PNAs in anti-sense strategies appears to be a powerful method for gene silencing related to the fact that the probe has not to cross the cellular membrane since the mRNA hybridization takes place in the cytosol. On the other hand PNAs have to deal with the possible formation of secondary structure of the single stranded RNA and to the continuous production of the target in multiple copies. The former problem is easily overcome by the great stability of the PNA:RNA complex which is sufficiently strong to open eventual secondary structures, whereas the latter problem can be overcome by the continuous administration of the PNA for maintaining the effect. This protocol was often used in genetic knock down studies, in which the inhibitory effect of a protein in the cellular system was evaluated[175]. Other examples of application of the RNA targeting were the inhibition of the retrotrascription of the HIV genome[176] or the block of the bacterial rRNA for the inhibition of the translation process[177]

Finally, the recent discovery of the miRNA machinery of the genetic regulation has opened a new interesting cellular target for oligonucleotide mimics, although application of PNAs in this "anti-miR" field has still to be investigated and only few examples are reported in the literature[178,179].

1.6 - References

[1] Watson J. D., Crick F. H. C. *Nature* **1953**, 171, 737-738.
[2] Blackburn G.M., Gait M. J., Loakes D., Williams D. M. *Nucleic Acids in Chemistry and Biology*. RSC Publishing, Cambridge, **2006**.
[3] Leontis N. B., Stombaugh J., Westhof E. *Nucleic Acids Res.* **2002**, 30, 3497-3531.
[4] Biou V., Yaremchuck A., Tukalo M., Cusak S. *Science* **1994**, 263, 1404-1410.
[5] Ban N., Nissen P., Hansen J., Moore P. B., Steitz T. A. *Science* **2000**, 289, 905-920.
[6] Su L.,Chen L., Egli M., Berger J. M., Rich A. *Nat. Struct. Biol.* **1999**, 3, 285-292.
[7] Baugh C., Grate D., Wilson C. *J. Mol. Biol.* **2000**, 301, 117-128.
[8] Smith C. M., Steiz J. A. *Cell* **1997**, 89, 669-672.
[9] Gold L., Polisky B., Uhlenbeck O. C., Yarus M. *Ann. Rev. Biochem.* **1995**, 64, 763-797.
[10] Chen L.-L., Carmichael G. G. *Current Opinion in Cell Biology* **2010**, 22, 357-364.
[11] Ghildiyal M., Zamore P. D. *Nat. Rev. Genet.* **2009**, 10 (2), 94-108.
[12] Kapranov P., Cheng J., Dike S., Nix D. A., Duttagupta R., Willingham A. T., Stadler P. F., Hertel J., Hackermüller J., Hofacker I. L., Bell I., Cheung E., Drenkow J., Dumais E., Patel S., Helt G., Ganesh M., Ghosh S., Piccolboni A., Sementchenko V., Tammana H., Gingeras T. R. *Science* **2007**, 316, 1484-1488.

[13] Taft R. J., Simons C., Nahkuri S., Oey H., Korbie D. J., Mercer T. R., Holst J., Ritchie W., Wong J. J., Rasko J. E., Rokhsar D. S., Degnan B. M., Mattick J. S. *Nat. Struct. Mol. Biol.* **2010**, 17(8), 1030-1034.

[14] Taft R. J., Kaplan C. D., Simons C., Mattick J. S. *Cell Cycle* **2009**, 8(15), 2332-2338.

[15] a) Ma N., Wang X., Qiao Y., Li F., Hui Y., Zou C., Jin J., Lv G., Peng Y., Wang L., Huang H., Zhou L., Zheng X., Gao X. *Mol Cell Endocrinol* **2011**, 333 (1), 96-101; b) Levy C., Khaled M., Iliopoulos D., Janas M. M., Schubert S., Pinner S., Chen P. H., Li S., Fletcher A. L., Yokoyama S., Scott K. L., Garraway L. A., Song J. S., Granter S. R., Turley S. J., Fisher D. E., Novina C. D. *Mol Cell* **2010,** 40 (5), 841-849; c) Hinske L. C., Galante P. A., Kuo W. P., Ohno-Machado L. *BMC Genomics* **2010**, 11, 533; d) Cao G., Huang B., Liu Z., Zhang J., Xu H., Xia W., Li J., Li S., Chen L., Ding H., Zhao Q., Fan M., Shen B., Shao N. *Biochem. Biophys. Res. Commun.* **2010**, 396 (4), 978-982.

[16] a) Ofir M., Hacohen D., Ginsberg D. *Mol. Cancer. Res.* **2011**, 9, 440-447; b) Woods K., Thomson J. M., Hammond S. M. *J. Biol. Chem.* **2007**, 282, 2130-2134; c) Rainer J., Ploner C., Jesacher S., Ploner A., Eduardoff M., Mansha M., Wasim M., Panzer-Grümayer R., Trajanoski Z., Niederegger H., Kofler R. *Leukemia* **2009**, 23 (4), 746-752; d) Berezikov E., Chung W. J., Willis J., Cuppen E., Lai E. C. Mol. Cell. **2007**, 28, 328-336.

[17] a) Choe J., Cho H., Lee H. C., Kim Y. K. *EMBO Rep.* **2010**, 11, 380-386; b) Bracken C. P., Szubert J. M., Mercer T. R., Dinger M. E., Thomson D. W., Mattick J. S., Michael M. Z., Goodall G. J. *Nucleic Acids Res.* **2011**, 39, 5658-5668; c) Huntzinger E., Izaurralde E. *Nat. Rev. Genet.* **2011**, 12, 99-110.

[18] Cuccato G., Polynikis A., Siciliano V., Graziano M., Di Bernardo M., Di Bernardo D. *BMC Syst. Biol.* **2011**, 5, 19.

[19] Yoda M., Kawamata T., Paroo Z., Ye X., Iwasaki S., Liu Q., Tomari Y. *Nat. Struct. Mol. Biol.* **2010**, 17, 17-23.

[20] Berezikov E., van Tetering G., Verheul M., van de Belt J., van Laake L., Vos J., Verloop R., van de Wetering M., Guryev V., Takada S., van Zonneveld A. J., Mano H., Plasterk R., Cuppen E. *Genome Res.* **2006**, 16 (10), 1289-1298.

[21] a) Hemida M. G., Ye X., Thair S., Yang D. *Mol. Diagn. Ther.* **2010**, 14, 271-282; b) Kota S. K., Balasubramanian S. *Drug. Discov. Today* **2010**, 15, 733-740; c) Bader A. G., Brown D., Winkler M. *Cancer. Res.* **2010**, 70, 7027-7030; d) Sibley C. R., Seow Y., Wood M. J. *Mol. Ther.* **2010**, 18, 466-476.

[22] Lu M., Zhang Q., Deng M., Miao J., Guo Y., Gao W., Cui Q. *PLoS One* **2008**, 3 (10), e3420.

[23] http://202.38.126.151/hmdd/mirna/md/

[24] Jay C., Nemunaitis J., Chen P., Fulgham P., Tong A. W. **DNA Cell Biol. 2007**, 26, 293-300.

[25] Björk J. K., Sandqvist A., Elsing A. N., Kotaja N., Sistonen L. *Development* **2010**, 137, 3177-3184.

[26] Jiang S., Zhang H. W., Lu M. H., He X. H., Li Y., Gu H., Liu M. F., Wang E. D. *Cancer Res.* **2010**, 70, 3119-3127.

[27] Kano M., Seki N., Kikkawa N., Fujimura L., Hoshino I., Akutsu Y., Chiyomaru T., Enokida H., Nakagawa M., Matsubara H. *Int. J. Cancer* **2010**, 127, 2804-2814.

[28] Nohata N., Sone Y., Hanazawa T., Fuse M., Kikkawa N., Yoshino H., Chiyomaru T., Kawakami K., Enokida H., Nakagawa M., Shozu M., Okamoto Y., Seki N. *Oncotarget* **2011**, 2, 29-42.

[29] Hurst D. R., Edmonds M. D., Welch D. R. *Cancer Res.* **2009**, 69, 7495-7498.

[30] Edmonds M. D., Hurst D. R., Welch D. R. *Cell Cycle* **2009**, 8, 2673-2675.

[31] Kosaka N., Iguchi H., Ochiya T. *Cancer Sci.* **2010**, 101, 2087-2092.

[32] Kosaka N., Iguchi H., Yoshioka Y., Takeshita F., Matsuki Y., Ochiya T. *J. Biol. Chem.* **2010**, 285, 17442-17445.

[33] Atkins S. D. , Clark I. M. *J. Appl. Genet.* **2004**, 45, 3-15.

[34] Lopez M. M., Bertolini E., Olmos A., Caruso P., Gorris M. T., Llop P., Penyalver R., Cambra M. *Int. Microbiol.* **2003**, 6, 233-243.

[35] Ghosh I., Stains C. I., Ooi A. T., Segal D. J. *Mol. BioSyst.* **2006**, 2, 551-560.

[36] Dervan P. B. *Bioorg. Med. Chem.* **2001**, 9, 2215-2235.

[37] Chan P. P., Glazer P. M. *J. Mol. Med.* **1997**, 75, 267-282.

[38] Segal D., Barbas C. F. *Curr. Opin. Biotechnol.* **2001**, 12, 632-637.

[39] Southern E. M. *J. Mol. Biol.* **1975**, 98, 503-517.

[40] Alwine J. C., Kemp D. J., Stark G. R. *Proc. Natl. Acad. Sci. USA* **1977**, 74(12), 5350-5354.

[41] Kumar S., Dagar S. S., Mohanty A. K., Sirohi S. K., Puniya M., Kuhad R. C., Sangu K. P. S., Griffith G. W., Puniya A. K. *Naturwissenschaften* **2011**, 98, 457-472.

[42] Macville M., Veldman T., Padilla-Nash H., Wangsa D., O'Brien P., Schrock E., Ried, T. *Histochem. Cell Biol.* **1997**, 108, 299-305.

[43] Levsky J. M., Singer R. H. *J. Cell Sci.* **2003**, 2833-2838.

[44] Femino, A. M. Fay F. S., Fogarty K., Singer R. H. **S***cience* **1998**, 280, 585-590.

[45] Levsky J. M., Shenoy S. M., Pezo R. C., Singer R. H. *Science* **2002**, 297, 836-840.

[46] Saiki R. K., Scharf S., Faloona F., Mullis K. B., Horn G. T., Erlich H. A., Arnheim N. *Science* **1985**, 230 (4732), 1350-1354.
[47] Wong M. L:, Medrano J. F. *BioTecniques* **2005***, 39, 75-85.*
[48] Gašparič M. B., Tengs T., La Paz J. L., Holst-Jensen A., Pla M., Esteve T., Žel J., Gruden K. *Anal. Bioanal. Chem.* **2010**, 369, 2023-2029.
[49] Venkatesan N., Seo Y. J., Kim B. H. *Chem. Soc. Rev.* **2008**, 37, 648-663.
[50] Plaxco K. W., Soh H. T. *Trends in Biotechnology* **2011**, 29 (1), 1-5
[51] Cady N. C., Strickland A. D., Batt C. A. *Mol. Cell Probe* **2007**, 21 (2), 116-124.
[52] Tyagi S., Kramer F. R. *Nat. Biotechnol.* **1996**, 14 (3), 303-308.
[53] Kostrikis L. G., Tyagi S., Mhlanga M. M., Ho D. D., Kramer F. R. *Science* **1998**, 279 (5354), 1228-1229.
[54] Nitin N., Santangelo P.J., Kim G., Nie S., Bao G. *Nucleic Acids Res*. **2004**, 32 (6), e58.
[55] Steemers F. J., Ferguson J. A:, Walt D. R., *Nat. Biotechnol.* **2000**, 18, 91-94.
[56] Sokol D. L., Zhang X., Lu P., GEwirtz A. M. *Proc. Natl. Sci. USA* **1998**, 95, 11538-11543.
[57] Tyagi S., Bratu D. P., Kramer F. R. *Nat. Biotechnol.* **1998**, 16 (1), 49-53.
[58] Tyagi S., Marras S. A., Kramer F. R. *Nat. Biotechnol.* **2000**, 18, 1191-1196.
[59] Southern E. M., Mir K., Shchepinov M. *Nat. Genet.* **1999**, 21 (1), 5-9.
[60] Reiss J., Krawczak M., Schloesser M., Wagner M., Cooper D. N. *Nucleic Acids Res.* **1990**, 18 (4), 973-978.
[61] Peccoud J., Jacob C. *Biophys. J.* **1996**, 71 (1), 101-108.
[62] Furukawa K., Abe H., Hibino K., Sako Y., Tsuneda S., Ito S. *Bioconjugate Chem.* **2009**, 20 (5), 1026-1036.
[63] Li H., Franzini R. M., Bruner C., Kool E. T. *ChemBioChem* **2010**, 11, 2132-2137.
[64] Köhler O., Jarikote D. V., Seitz O. *ChemBioChem* **2005**, 6, 69-77.
[65] Socher E., Jarikote D. V., Knoll A., Röglin L., Burmeister J., Seitz O. *Anal. Biochem.* **2008**, 375, 318-330.
[66] Ziober B. L., Mauk M. G., Falls E. M., Chen Z., Ziober A. F., Bau H. H. *Head & Neck* **2008**, 30 (1), 111-121.
[67] Bennett M. R., Hasty J. *Nat. Rev. Genetics* **2009**, 10 (9), 628-638.
[68] Anderson R. C., Su X., Bogdan G. J., Fenton J. *Nucleic Acids Res.* **2000**, 28 (12), e60.
[69] Lee C.-Y., Lee G.-B., Lin J.-L., Huang F.-C., Liao C.-S. *J. Micromech. Microeng.* **2005**, 15, 1215-1223.
[70] Zhang Y. H., Pinar O. *Analytica Chimica Acta* **2009**, 638 (2), 115-125.
[71] Liu R. H., Yang J., Lenigk R., Bonanno J., Grodzinski P. *Anal. Chem.* **2004**, 76, 1824-1831.
[72] Tsai N. C., Sue C. Y. *Biosens. Bioelectron.* **2006**, 22, 313-317.
[73] Storhoff J. J., Lucas A. D., Garimella V., Bao Y. P., Müller U. R. *Nat. Biotechnol.* **2004**, 22, 883-887.
[74] Goodrich T. T., Lee H. J., Corn R.M. *J. Am. Chem. Soc.* **2004**, 126, 4086-4087.
[75] Ho H. A., Doré K., Boissinot M., Bergeron M. G., Tanguay R. M., Boudreau D., Leclerc M. *J. Am. Chem. Soc.* **2005**, 127, 12673-12676.
[76] Zhu D., Tang Y., Xing D., Chen W. R. *Anal. Chem.* **2008,** 80, 3566-3571.
[77] Ranasinghe A. T., Brown T. *Chem. Commun.* **2011**, 47, 3717-3735.
[78] Varallyay E., Burgyan J., Havelda Z. *Methods* **2007**, 43, 140-145.
[79] Pall G. S., Codony-Servat C., Byrne J., Ritchie L., Hamilton A., *Nucleic Acid Res.* **2007**, 35, e60.
[80] Liu C. G., Calin G. A., Meloon B., Gamliel N., Sevignani C., Ferracin M., Dumitru C. D., Shimizu M., Zupo S., Dono M., Alder H., Bullrich F., Negrini M., Croce C.M. *Proc. Natl. Acad. Sci. USA* **2004**, 101, 9740-9774.
[81] Wang X. *Nucleic Acids Res.* **2006**, 34, 1646-1652.
[82] Wienholds E., Kloosterman W. P., Miska E., Alvarez-Saavedra E., Berezikov E., de Bruijn E., Horvitz H. R., Kauppinen S., Plasterk R. H. *Science* **2005**, 309, 310-311.
[83] Kloosterman W. P., Wienholds E., de Bruijn E., Kauppinen S., Plasterk R. H. *Nat. Methods* **2007**, 3, 27-29.
[84] Silahtaroglu A. N., Nolting D., Dyrskjøt L., Berezikov E., Møller M., Tommerup N., Kauppinen S. *NATURE PROTOCOLS* **2007**, 2 (10), 2520-2528.
[85] Berezikov E., Cuppen E., Plasterk R. H. *Nat. Genet.* **2006**, 38, S2–S7.
[86] Wang F., Niu G., Chen X. Cao F. *Eur. J. Med. Mol. Imaging* **2011**, 38, 1572-1579.
[87] Zeng Y., Cullen B. R. *RNA* **2003**, 9 (1), 112-123.
[88] Schmittgen T. D., Lee E. J., Jiang J., Sarkar A., Yang L., Elton T. S., Chen C. *Methods* **2008**, 44 (1), 31-38.
[89] Shi R., Chiang V. L. *Biotechniques* **2005**, 39 (4), 519-525.
[90] Allawi H. T., Dahlberg J. E., Olson S., Lund E., Olson M., Ma W. P., Takova T., Neri B. P., Lyamichev VI *RNA* **2004**, 10, 1153-1161.
[91] Jonstrup S. P., Koch J., Kjems J. *RNA* **2006**, 12, 1747-1752.

[92] Maroney P. A., Chamnongpol S., Souret F., Nilsen T. W. *Nat. Protoc.* **2008**, 3, 279-287.
[93] Opalinska J. B., Gewirtz A. M. *Nat. Rev. Drug Discovery* **2002**, 1, 503-514.
[94] Faria M., Wood C. D., Perrouault L., Nelson J. S., Winter A., White M. R. H., Hélène C., Giovannangeli C. *Proc. Natl. Acad. Sci. USA* **2000**, 97, 3862-3867.
[95] Scherer L. J., Rossi J. J. *Nat. Biotechnol.* **2003**, 21, 1457-1465.
[96] Zamecnik P. C., Stephenson M. L. *Proc. Natl. Acad. Sci. USA* **1978**, 78, 280-284.
[97] Faria M., Spiller D. G., Dubertret C., Nelson J. S., White M. R. H., Sherman D., Hélène C., Giovannangeli C. *Nat. Biotechnol.* **2001**, 19, 40-44.
[98] Mercatante D. R., Kole R. *Biochim. Biophys. Acta* **2002**, 29, 1293-1299.
[99] Vickers T. A., Wyatt J. R., Burckin T., Bennet C. F., Freier S. M. *Nucl. Acid Res.* **2001**, 29, 1293-1299.
[100] Napoli C., Lemieux C., Jorgensen R. *The Plant cell* **1990**, 2 (4), 279-289.
[101] Hamilton A. J., Baulcombe D. C. *Science* **1999**, 286, 950-952.
[102] Fire A., Xu S., Montgomery M. K., Kostas S. A., Driver S. E., Mello C. C. *Nature* **1998**, 39, 806-811.
[103] Amarzguioui M., Rossi J. J., Kim D. *FEBS Lett.* **2005**, 579, 5974-5981.
[104] Tijserman M., Plasterk R. H. A. *Cell* **2004**, 117, 1-4.
[105] Soutschek J., Akinc A., Bramlage B., Charisse K., Constien R., Donoghue M., Elbashir S. M., Geick A., Hadwiger P., Harborth J., John M., Kesavan V., Lavine G., Pandey R. K., Racie T., Rajeev K. G., Röhl I., Toudjarska I., Wang G., Wuschko S., Bumcrot D., Koteliansky V., Limmer S., Manoharan M., Vornlocher H. P. *Nature* **2004**, 432, 173-177.
[106] Czech M. P., Aouadi M., Tesz G. J. *Nat. Rev. Endocrinol.* **2011**, 7, 473-484.
[107] a) Watts J. K., Deleavey G. F., Damha M. J. *Drug Discov. Today* **2008**, 13, 842-855; b) Behlke M. A. *Oligonucleotides* **2008**, 18, 305-319; c) Chernolovskaya E. L., Zenkova M. A. *Curr. Opin. Mol. Ther.* **2010**, 12, 158-167.
[108] Chiu Y. L., Rana T. M. *RNA* **2003**, 9, 1034-1048.
[109] Allerson C. R., Sioufi N., Jarres R., Prakash T. P., Naik N., Berdeja A., Wanders L., Griffey R. H., Swayze E. E., Bhat B. *J. Med. Chem.* **2005**, 48, 901-904.
[110] Veedu R. N., Wengel J. *Chem. Biodivers.* **2010**, 7, 536-542.
[111] Zender L., Hutker S., Liedtke C., Tillmann H. L., Zender S., Mundt B., Waltemathe M., Gosling T., Flemming P., Malek N. P., Trautwein C., Manns M. P., Kuhnel F., Kubicka S., *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 7797-7802.
[112] Braasch D. A., Paroo Z., Constantinescu A., Ren G., Oz O. K., Mason R. P., Corey D. R. *Bioorg. Med. Chem. Lett.* **2004**, 14, 1139-1143.
[113] Bruno K. *Adv. Drug Deliv. Rev.* **2011**, 63 (13), 1210-1226.
[114] Seto A. G. The International Journal of Biochemistry & Cell Biology **2010**, 42, 1298-1305.
[115] Czech M. P. *New England Journal of Medicine* **2006***,* 354 (11), 1194-1195.
[116] Calin G. A., Croce C. M. *Nat. Rev. Cancer* **2006**, 6 (11), 857-866.
[117] Papagiannakopoulos T., Shapiro A., Kosik K. S. *Cancer Research* **2008**, 68 (19), 8164-8172.
[118] Corsten M.F., Miranda R., Kasmieh R., Krichevsky A. M., Weissleder R., Shah K. *Cancer Research* **2007**, 67 (19), 8994-9000.
[119] Jopling C. L., Yi M., Lancaster A.M., Lemon S.M., Sarnow P. *Science* **2005***,* 309 (5740), 1577-1581.
[120] Triboulet R., Mari B., Lin Y. L., Chable-Bessia C., Bennasser Y., Lebrigand K., Cardinaud B., Maurin T., Barbry P., Baillat V., Reynes J., Corbeau P., Jeang K. T., Benkirane M. *Science* **2007,** 315 (5818), 1579-82.
[121] Krützfeldt J., Rajewsky N., Braich R., Rajeev K. G., Tuschl T., Manoharan M., Stoffel M. *Nature* **2005**, 438 (7068), 685-689.
[122] Elmén J., Lindow M., Schütz S., Lawrence M., Petri A., Obad S., Lindholm M., Hedtjärn M., Hansen H. F., Berger U., Gullans S., Kearney P., Sarnow P., Straarup E. M., Kauppinen S. *Nature* **2008**, 452 (7189), 896-900.
[123] Lanford R.E., Hildebrandt-Eriksen E. S., Petri A., Persson R., Lindow M., Munk M. E., Kauppinen S., Ørum H. *Science* **2010**, 327 (5962), 198-201.
[124] Brown B. D., Naldini L. *Nat. Rev. Genetics* **2009**, 10 (8), 578-585.
[125] Torres A. G., Fabani M. M., Vigorito E., Gait M. J. *RNA* **2011**, 17, 933-943.
[126] Nielsen P.E., Egholm M., Berg R.H., Buchardt O. *Science* **1991**, 254, 1497-1500.
[127] Egholm M., Buchardt O., Christensen R., Behrens C., Freier S. M., Driver D. A., Berg R. H., Kim S. K., Norden B., Nielsen P. E., *Nature* **1993**, 365 (6446), 566-568.

[128] Uhlmann E., Peyman A., Breipohl G., Will D. W. *Angew. Chem. Int. Ed.* **1998**, 37, 2796-2823.
[129] Demidov V. A., Potaman V. N., Frank-Kamenetskii M. D., Egholm M., Buchardt O., Sonnichsen S. H., Nielsen P. E. *Biochem. Pharmscol.* **1994**, 48 (6), 1310-1313.
[130] Nielsen P. E.; Egholm M. Peptide Nucleic Acids. Protocols and Applications 2nd Ed. 2004, Horizon Scientific Press (Wymondham, UK)
[131] Howarth N.M. *Letters in Organic Chemistry* **2006**, 3, 495-503.
[132] R. Corradini, S. Sforza, T. Tedeschi, F. Totsingan, A. Manicardi, R. Marchelli. *Current topics in medicinal chemistry* **2011**, 11(12), 1535-1554.
[133] Zhou P., Wang M., Du L., Fisher G. W., Waggoner A. Ly D. H. *J. Am. Chem. Soc.* **2003**, 125, 6878-6879.
[134] Sforza S., Corradini R., Ghirardi S., Dossena A., Marchelli R. *Eur. J. Org. Chem.* **2000**, 2905-2913.
[135] Haaima G., Lohse A., Buchardt O., Nielsen P. E. *Angew. Chem.* **1996**, 35 (17), 1939-1942.
[136] a) Tedeschi T., Sforza S., Corradini R., Marchelli R. *Tetrehedron Lett.* **2005**, 46, 8395-8399; b) Falkiewicz B., Kołodziejczyk A. S., Liberek B., Wiśniewski K. *Tetrahedron* **2001**, 57, 7909-7917; c) Kosynkina L., Wang W., Liang T. C. *Tetrahedron Lett.* **1994**, 35, 5173-5176; d) Englung E. A., Appella D. H. *Org. Lett.* **2005**, 7 (16), 3465-3467; e) Zhou P., Wang M. M., Du L., Fisher G. W., Waggoner A., Ly D. H. *J. Am. Chem. Soc.* **2003**, 125 (23), 6878-6879.
[137] Sforza S., Tedeschi T., Ciavardelli D., Corradini R., Dossena A., Marchelli R. *Eur. J. Org. Chem.* **2003**, 1056-1063.
[138] Kumar V. A., Ganesh K. N. *Acc. Chem. Res.* **2005**, 38, 404-412.
[139] Myers M. C., Witschi M. A., Larionova N. V., Franck J. M., Haynes R. D., Hara T., Grajkowski A., Appella D. H. *Org. Lett.* **2003**, 5 (15), 2695-2698.
[140] Govindaraju T., Madhuri V., Kumar V. A., Ganesh K. N. *J. Org. Chem.* **2006**, 71, 14-21.
[141] Pokorski J. K., Myers M. C., Appella D. H. *Tetrahedron Lett.* **2005**, 46 (6), 915-917.
[142] Vilaivan T., Suparpprom C., Duanglaor P., Harnyuttanakorn P., Lowe G. *Tetrahedron Lett.* **2003**, 44 (8), 1663-1666.
[143] Shirude P. S., Kumar V. A., Ganesh K. N. *Tetrahedron Lett.* **2004**, 60, 9485-9491.
[144] Lonkar P. S., Kumar V. A. *J. Org. Chem.* **2005**, 70 (17), 6956-6959.
[145] Gagamani, B. P., Kumar, V. A., Ganesh K. N. *Tetrahedron* **1999**, 55, 177-192.
[146] Govindaraju T., Kumar V.A. *Chem. Commun.* **2005**, 495-497.
[147] Kool E.T. *Acc. Chem. Res.* **2002**, 35 (11), 936-943.
[148] Loakes D. *Nucleic Acid Res.* **2001**, 29 (12), 2437-2447.
[149] a) Eldrup A. B., Nielsen B. B., Haaima G., Rasmussen H., Kastrup J. S., Christensen C., Nielsen P. E. *Eur. J. Org. Chem.* **2001**, 9, 1781-1790; b) Wilds C.J., Maier M.A., Tereshko V., Manoharan M., Egli M. *Angew. Chem. Int. Ed.* **2002**, 41 (1), 115-117; c) Eldrup A. B., Christensen C., Haaima G., Nielsen P. E., *J. Am. Chem. Soc.,* **2002**, 124 (13), 3254-3262.
[150] Wilhelmsson L. M., Holmén A., Lincoln P., Nielsen P. M., Nordén B. *J. Am. Chem. Soc.* **2001**, 123 (10), 2434-2435.
[151] Rajeev K. G., Maier M. A:, Lasnik E. A., Manoharan M. *Org. Lett.* **2002**, 4, 4395-4398.
[152] Haaima G., Hansen H.F., Christensen L., Dahl O., Nielsen P.E. *Nucleic Acids Res.* **1997**, 25 (22), 4639-4643.
[153] Lohse J., Dahl O., Nielsen P.E. *Proc. Natl. Acad. Sci., USA* **1999**, 96, 11804-11808.
[154] Koehler O., Jarikote D.V., Seitz O. *ChemBioChem* **2005**, 6, 69-77.
[155] Nielsen P. E. *Curr. Opin. Biotech.* **2001**, 12, 16-20.
[156] Calabretta A., Tedeschi T., Di Cola G., Corradini R., Sforza S., Marchelli R. *Mol. Biosys.* **2009**, 5(11), 1323-1330.
[157] Taneja K. L., Chavez E. A., Coull J., Lansdorp P. M. *Genes Chromosomes Cancer* **2001**, 30 (1), 57-63.
[158] Stender H., Fiandaca M., Hyldig-Nielsen J. J., Coull J. *J. Microbiol. Met.* **2002**, 48 (1), 1-17.
[159] Perry-O'Keefe H., Yao X. W., Coull J. M., Fuchs M, Egholm M., *Proc. Natl. Acad. Sci. USA* **1996**, 93, 14670-14675.
[160] Tovar-Salazar A., Dhawan J., Lovejoy A., Liu Q. A., Gifford A. N. *Anal. Biochem.* **2007**, 360 (1), 92-98.
[161] D'Agata R., Corradini R., Grasso G., Marchelli R., Spoto G. *ChemBioChem* **2009**, 9(13), 2067-2070;
[162] D'Agata R., Corradini R., Ferretti C., Zanoli L., Gatti M., Marchelli R., Spoto G. *Biosensors and Bioelectronics* **2010**, 25, 2095-2100.
[163] Kuhn H., Demidov V. V., Coul J. M., Fiandaca M. J., Gildea B. D., Frank-Kamenetskii M. D. *J. Am. Chem. Soc.* **2002**, 124, 1097-1103.
[164] Lee J. W., Bang K.-H., Choi J.-J., Chung J.-W., Lee J.-H., Jo I.-H., Seo A.-Y., Kim Y.-C., Kim O.-T., Cha S.-W. *Genes & Genomics* **2010**, 32 (5), 463-468.
[165] Choi J.-J., Cho M., Oh M., Kim H., Kil M.-S., Park H. *Bull. Korean Chem. Soc.* **2010**, 31 (12), 3525.
[166] Huber C. G. *Journal of Chromatography A* **1998**, 806, 3-30.

[167] Germini A., Scaravelli E., Lesignoli F., Sforza S., Corradini R., Marchelli R.; *European Food Research and Technology* **2005**, 200(5-6), 619-624.

[168] Tedeschi T., Chiari M., Galaverna G., Sforza S., Cretich M., Corradini R., Marchelli R. *Electrophoresis* **2005**, 26, 4310-4316.

[169] Hu J., Corey D. R. *Biochemistry* **2007**, 46 (25), 7581-7589.

[170] Pesce C. D., Bolacchi F., Bongiovanni B., Cisotta F., Capozzi M., Diviacco S., Quadrifoglio F., Mango R., Novelli G., Mossa G., Esposito C., Ombres D., Rocchi G., Bergamini A. *Antiviral Res.* **2005**, 66 (1), 13-22.

[171] Tonelli R., Purgato S., Camerin C., Fronza R., Bologna F., Alboresi S., Franzoni M., Corradini R., Sforza S., Faccini A., Shohet J. M., Marchelli R. *Mol. Cancer Ther.* **2005**, 4, 779-786.

[172] Tonelli R., Mc Intyre A., Camerin C., Walters Z. S., Di Leo K., Selfe J., Purgato S. Missiaglia E., Tortori A., Astolfi A., Renshaw J., Taylor K. R:, Serravalle S., Bishop R., Nanni C., Valentijn L. J., Faccini A., Leuschner I., Formica S., Reis-Filho J. S., Ambrosini V., Thway K., Franzoni M., Summersgill B., Marchelli R., Hrelia P., Cantarelli-Forti G., Fansi S., Corradini R:, Pession A., Shipley J. *Clinical Cancer Res.*, in press, doi:10.1158/1078-0432.CCR-11-1981.

[173] Diviacco S., Rapozzi V., Xodo L., Helene C., Quadrifoglio F., Giovannangeli C. *The FASEB Journal*, **2001**, 15, 2660-2668.

[174] Taylor R. W., Chinnery P. F., Turnbull D. M., Lightowlers R. N. *Nature Genet*. **1997**, 15, 212-215.

[175] Nekohtiaeva N., Awasthi S. K., Nielsen P. E., Good L. *Molecular Therapy* **2005**, 10, 652-659.

[176] Lee R., Kaushik N., Modak M. J., Vinayak R., Pandey V. N. *Biochemistry* **1998**, 37(3), 900-910.

[177] Good L., Nielsen P. E. *PNAS* **1998**, 95(5), 2073-2076.

[178] Fabani M. M., Gait M. J. *RNA* **2008**, 14 (2), 336-346.

[179] Fabani M. M., Abreu-Goodger C., Williams D., Lyons P. A., Torres A. G., Smith K. G. C., Enright A. J., Gait M. J., Vigorito E. *Nucl. Acids Res.* **2010**, 38 (13), 4466-4475.

Chapter 2 - PNA-based systems for DNA single point mutations

2.1 - Introduction

Single nucleotide polymorphisms (SNPs) are one of the major sources of genetic variations in the genome[1]. Genetic diseases are often related to these type of mutations[2]; thus and the availability of methods to detect SNPs allows to perform accurate medical diagnosis and to discriminate between homozygous and heterozygous individuals. Methods able to perform SNP recognition beside medical applications can also be used in other different fields such as food analysis[3]

SNP recognition can be achieved through different techniques based on the use of specific probes: gel electrophoresis, capillary electrophoresis, affinity immobilization, PCR methods, HPLC analysis, microarrays and switching probes.

2.1.1 - Microarrays technology

The microarrays technology is a surface-based method for the discrimination of nucleic acids sequences[4,5], in which the recognition of a DNA sequence is performed by hybridization with an array of single stranded probes immobilized onto a surface (polymeric or glass) as show in Figure 2.1. The fabrication of this kind of devices is carried out by direct synthesis of oligonucleotides on the solid surface by photolithographic methods (for high density arrays), or by deposition of multiple probes on an activated surface (spotted arrays). These systems have several advantages in terms of extremely high efficiency, since they can be used for parallel high-throughput analyses (such as complete profiling of gene expression), usually with very small volumes of solutions for both deposition and analysis; moreover this format offers the possibility to denaturate the complexes formed, leaving the probes linked to the surface, thus allowing the recycle of the device.

The use of PNA in microarrays technology can meet some of the drawbacks present in DNA-microarrays, like the strong dependence of the duplex stability upon the ionic strength[6] and the

inability of DNA to efficiently invade the secondary structure of the target DNA[7]. Taken all together, these characteristics make PNA-based array technology an important tool for fast and multiple analyses of samples.

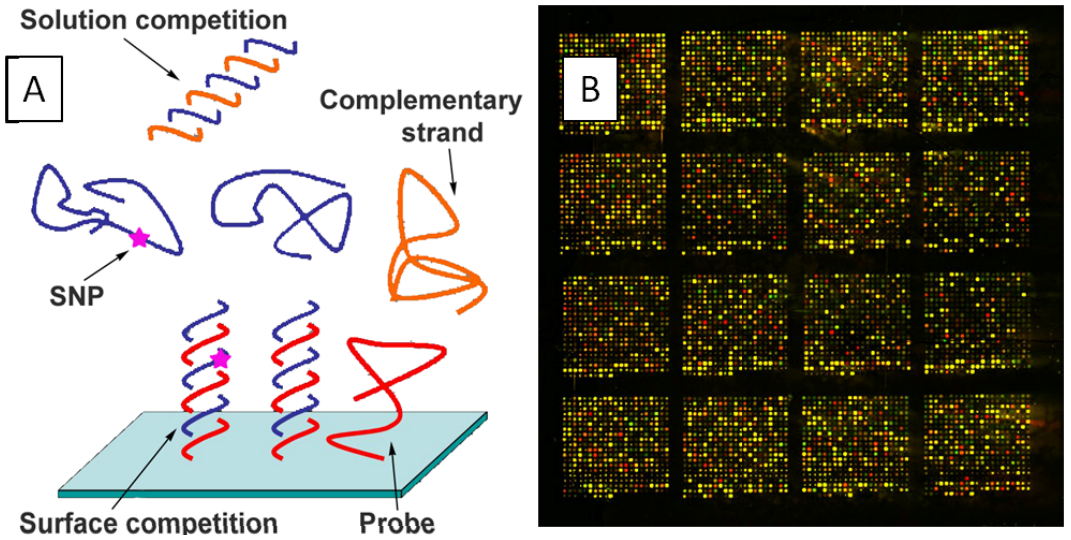


Figure 2.1: A) processes occuring during microarrays analysis; B) a typical microarray analysis.

2.1.2 - Pyrene based probes

Modified nucleotides containing pyrene derivatives have been used for nucleotide quantification, SNP detection[8,9] , increased duplex stability[10] or structural alteration studies[11,12]. For analytical purposes, pyrene-based probes were used in combination with different formats, such as beacon[13,14], fluorescence resonance energy transfer (FRET)[15] or base discriminating fluorescence (BDF)[16]. Depending on the application and on the technique used, pyrene units were linked to probes at different locations, that is in various positions of the sugar (or analog) backbone or appended to the nucleobases, and placed in different parts of the oligonucleotide chain, e.g. at one or at both ends of the sequence, as reported in Figure 2.2.

Many approaches were proposed, however, the most selective and efficient sensing mechanism is the excimer formation, thus several different geometries and combinations of the probes have been described, in order to obtain this signal upon interaction with the target sequence. A straightforward method was to use two pyrene conjugated probes targeting adjacent sequences (Figure 2.2b). In the presence of a complementary target, the two pyrene units get close and generate an excimer signal;

instead, in the presence of a mismatched target, the two pyrene units cannot come closer and there is no interaction between the two aromatic moieties. Another very efficient use of pyrene units is that introduced by "LNA invader" probes in which a dsLNA probe is used and both strands bind to target dsDNA, thus generating a fluorescent signal.[17]

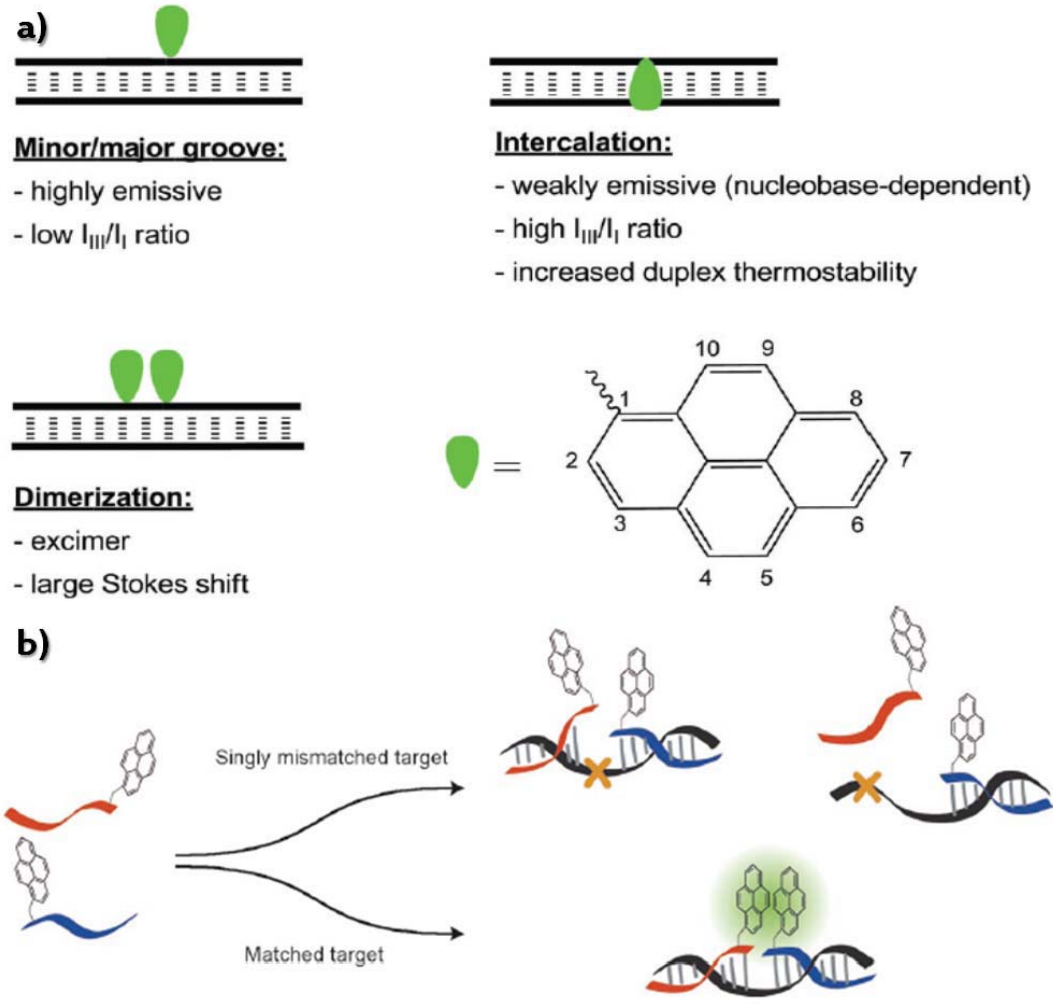


Figure 2.2: model systems used in combination with pyrene units; a) single probe systems, b) dual probe system.[18]

The introduction of a pyrene unit in PNA probes can be achieved through both backbone modification or nucleobases modification. In the literature many different approaches concerning its insertion on nucleobase have been reported (Figure 2.3), and in all cases the fluorophore is pointed to the major groove of the double-stranded structure. This choice is related to two factors: the first is related to stabilities issues: in fact, the nucleobase maintains its recognition ability and the modification is pointing inside the major groove reducing the destabilization induced by steric

hindrance; the second factor is related to the fact that the fluorescence of pyrene derivatives is strongly dependent on environment[19] and the complexation of PNA to perfectly matched or mismatched DNA lends to distortions that are extended to the pyrene location and its exposition to the solvent.

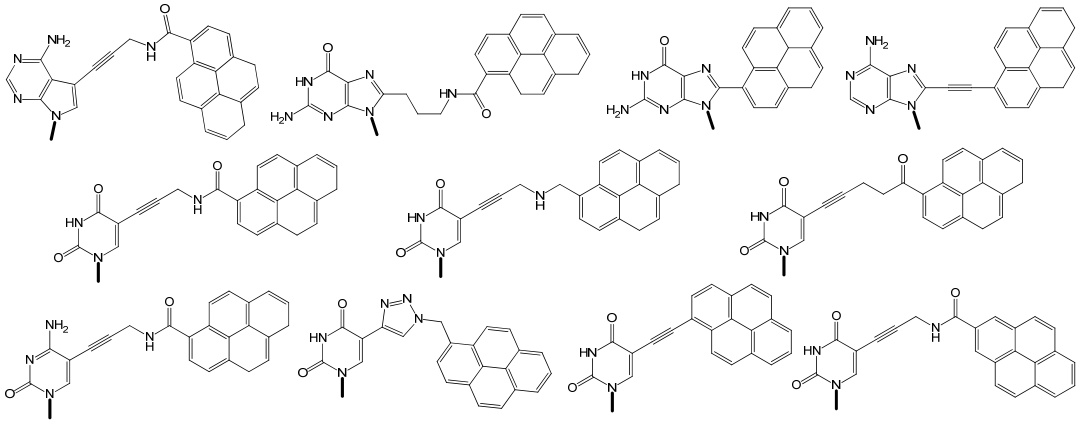


Figure 2.3: modified nucleobases containing a pyrene unit (from ref 8-16).

2.2 - Surface SNPs detection by "chiral-box" PNAs

The insertion of a single C2- or C5- modified chiral monomer in the middle of a PNA strand resulted in the modulation of the preferred PNA handedness, which, in turn, affected the preference for the right-handed complementary DNA giving rise to a very stable PNA-DNA duplex[20]. In particular, the use of C2-modified monomers derived from D-amino acids and C5-modified monomers derived from L-amino acids were shown to induce a preferential right-handedness, with the latter being a stronger inducer than the former. As a result, C2,C5-modified PNAs were found to bind very effectively to DNA when the stereochemistry was 2D,5L (or 2R,5S)[21].

The use of chiral PNAs for the detection of single point mutation was demonstrated to be very effective in several analytical techniques. By introducing a series of three adjacent C2-modified chiral monomers, termed "chiral-box" PNAs, very high direction control (antiparallel vs. parallel binding) and exceptionally high sequence selectivity were achieved[22]. In this work we compare the performances of modified PNAs bearing a "chiral-box" composed of C2- or C5-modified monomers in terms of binding affinity and mismatch recognition both in solution and in microarray technology.

2.2.1 - Design and synthesis of chiral PNAs

Since positively charged side chains were shown to improve binding ability in C2-modified PNAs, compared to neutral or negatively charged ones, the "chiral box" model was previously built using lysine synthons. However, since the amino group of lysine can interfere with in the binding of the PNA terminal amino group in the microarray spotting process, arginine units were rather chosen. The "chiral box" model was extended to C5-modified monomers. The stereochemistry was chosen on the basis of previously reported studies, using L-Arg for the C5-modifed monomers, and D-Arg for the C2-modified monomers.

The model system used has a DNA target sequence of interest contained a SNP related to vegetal varieties; and the "chiral-box" motif in PNA was centred on the base corresponding to the SNP to be recognized (Figure 2.4).

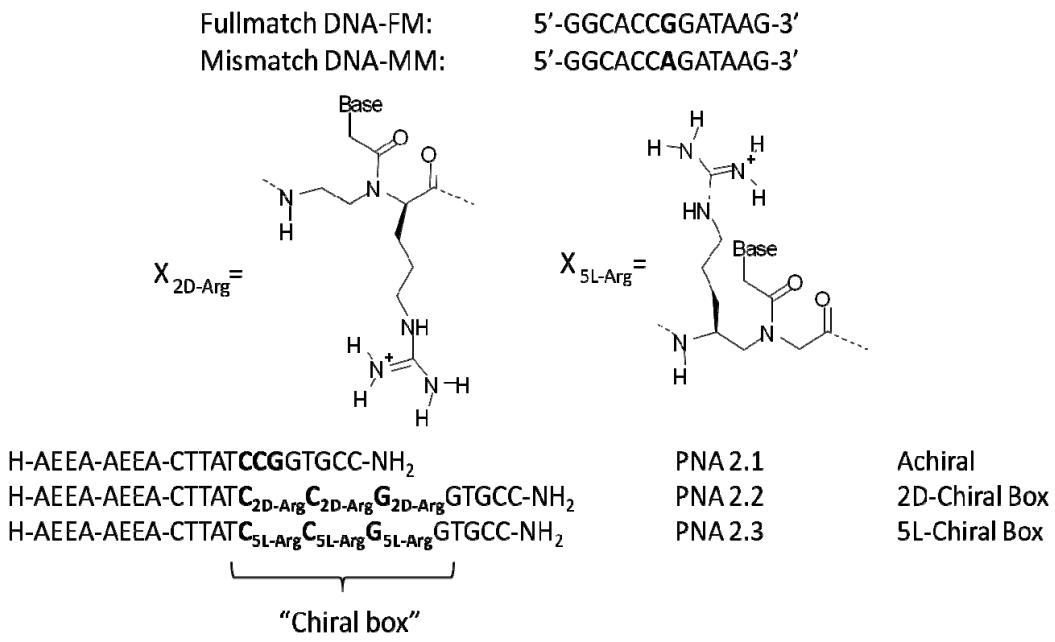
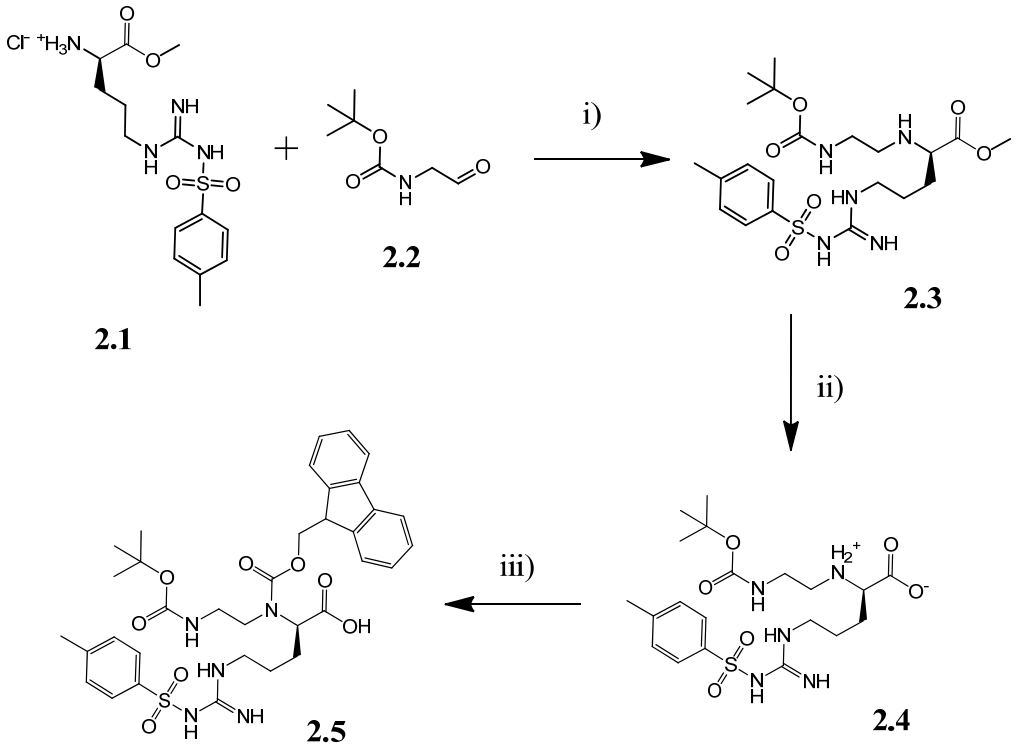


Figure 2.4: sequences used for this study. AEEA refers to the 2-(2-aminoetoxy)etoxyacetyl spacer.

C-2 substituted chiral PNA monomers are sensitive to racemization[22], and should therefore be synthesized using a special protocol called "submonomeric approach" in which only the submonomer **2.5** lacking of the carboxymethyl base is linked to the peptide chain, followed by deprotection of the α-amino group and coupling with the desired carboxymethylnucleobase. The synthesis of the C2-modified D-Arg submonomer is depicted in Scheme 2.1. The tosyl protected

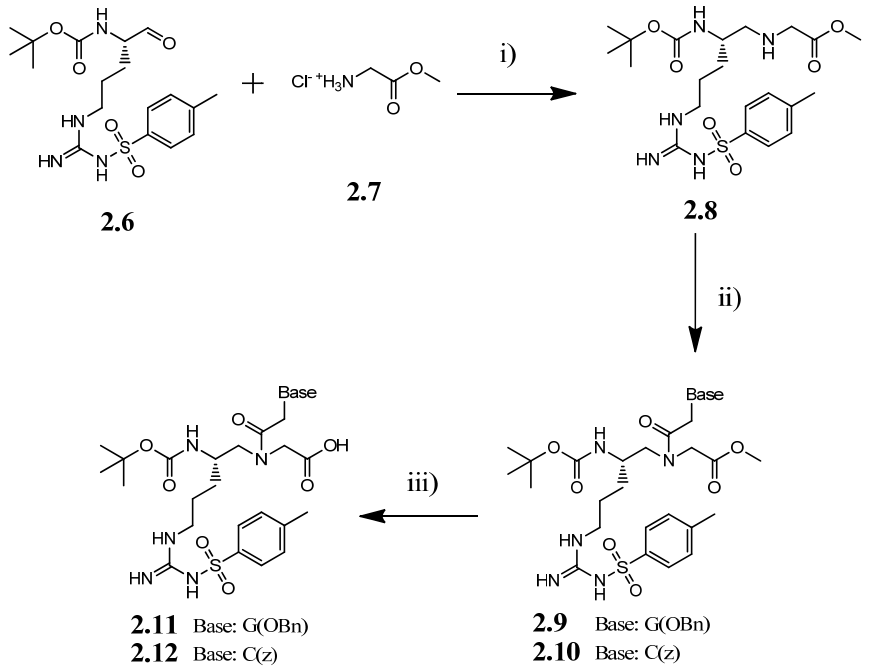
Arg methyl ester **2.1** was coupled with Boc-2-aminoacetaldehyde **2.2** under reductive amination conditions to obtain the C-protected backbone **2.3**. Deprotection of the methyl ester gave the backbone **2.4,** and subsequent protection of the α-amino group with Fmoc-chloride in the presence of a temporary protecting silylating agent (BTSA) led to the final C2-D-Arg submonomer **2.5**.



Scheme 2.1: Synthesis of the chiral 2D-Arg PNA submonomer 2.5: i) $NaBH_3CN$, CH_3COOH in MeOH, 68%; ii) NaOH in water, 97%; iii) BTSA, DIPEA, Fmoc-Cl in DCM, 82%.

C5-substituted PNA monomers do not undergo racemisation during solid-phase synthesis and were therefore synthesized entirely in solution. Thus, we designed two C5-substituted Boc-PNA monomers by using L-$Arg_{(Tos)}$ synthons, following the synthetic scheme previously described for other amino acids[23]: synthesis of α-substituted Boc-aminoacetaldehyde **2.6** (through reduction of the corresponding Weinreb amide), which was subsequently used in reductive amination with glycine methyl ester to yield the backbone methyl ester **2.8**, followed by reaction with the carboxymethylbase to yield the C-protected monomers **2.9** and **2.10**. Hydrolysis of these compounds by 1M $Ba(OH)_2$ in a THF/H_2O solution yielded the chiral monomers **2.11** and **2.12**, as reported in Scheme 2.2**.**

Each PNA oligomer was synthesized using a different approach: the achiral **PNA 2.1** was synthesized on an automatic ABI 433A Synthesizer using Fmoc-based solid phase synthesis, the



Scheme 2.2: Synthesis of the 5L-Arg chiral PNA monomers 2.11 and 2.12: i) $NaBH_3CN$, CH_3COOH in MeOH, 50%; ii) $CMG_{(OBn)}$ or $CMC_{(Z)}$, EDC·HCl, DhBtOH in DMF, 79-95%; iii) $Ba(OH)_2$ in THF/H_2O = 1: 1, 72-92%.

synthesis of 2D-chiral-box **PNA 2.2** was carried out using the submonomeric synthesis previously described for other chiral-box PNAs[23,24]; finally the 5L-chiral-box **PNA 2.3** was synthesized using Boc-based solid phase synthesis with the two preformed chiral monomers and commercially available Boc-PNA monomers. The crude products were purified by RP-HPLC and characterized by HPLC-MS, as described in the Experimental part, obtaining satisfactory yields (12÷16%) and purity, as required for the evaluation of the properties of recognition in both solution and surface.

2.2.2 - Recognition properties in solution

UV-melting temperatures of the DNA:PNA duplexes were used to evaluate the affinity and sequence selectivity of the PNA in solution. The UV melting curves of the two chiral PNAs are reported in Figure 2.5, and the corresponding melting temperatures are reported in Table 2.1 and compared with those of the achiral **PNA 2.1**. The 2D-chiral-box **PNA 2.2**: DNA duplex was found to have a neat melting transition, with only a slightly lower melting temperature than the achiral PNA:DNA, similarly to what reported for lysine-based 2D-chiral-box PNAs with different sequences[22,25]. On the contrary, the 5L-chiral-box **PNA 2.3** was found to have a less steep increase

in the absorbance upon melting, with a larger transition range. Although under the present conditions it was not possible to determine enthalpy/entropy contributions, it is evident that the thermodynamic properties should be significantly different from the 2D **PNA 2.2** case. The melting temperature for 5L-chiral-box **PNA 2.3** was significantly higher than in the achiral PNA:DNA duplex, thus confirming that this modification is the most suited for obtaining PNA with high affinity for complementary DNA. This behaviour is in line with the results of Ly and co-workers, who have used this type of modification for performing strand invasion of duplex DNA[26,27].

Table 2.1: UV and CD melting temperatures of the duplexes of the PNA with full matched and mismatched DNA duplexes. (ΔT= Tm FM – Tm MM)

PNA	DNA	UV T_m	UV ΔT_m	CD T_m	CD ΔT_m
PNA 2.1	FM	67 °C	22 °C	69 °C	22 °C
	MM	45 °C		47 °C	
PNA 2.2	FM	66 °C	26 °C	67 °C	27 °C
	MM	40 °C		40 °C	
PNA 2.3	FM	75 °C	20 °C	75 °C	19 °C
	MM	55 °C		56 °C	

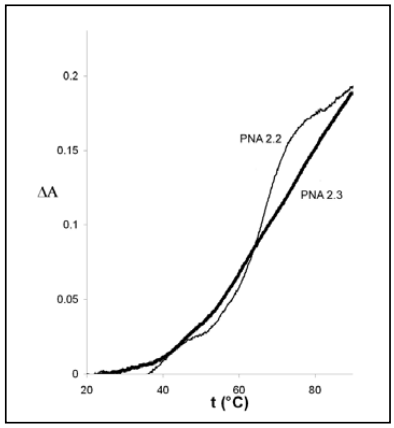


Figure 2.5: UV Melting curves (260 nm) for the DNA-FM with PNA 2.2 (thin line) and PNA 2.3 (thick line), in PBS buffer, pH 7; concentration of each strand was 5 μM. Data are expressed as variation in absorbance (ΔA). Cell length was 1 cm.

The recognition of a single mismatch (A instead of G) corresponding to the SNP of interest was also evaluated by the decrease of the melting temperatures. In this case, the **PNA 2.2**: DNA was found to have the lowest melting temperature, with highest mismatch selectivity, since the difference between the FM and the MM DNA complexes (ΔTm) was 26 °C, higher than that observed for the achiral **PNA 2.1** (22°C), while **PNA 2.3** gave a stable duplex also in the presence of the mismatch, with ΔTm of 20°C. Therefore, while **PNA 2.3** was superior in terms of binding, **PNA 2.2** was found to be the best model for the recognition of a single mismatch.

The circular dichroism spectra of the two chiral PNA and of the PNA:DNA duplexes are reported in Figure 2.6. A certain degree of preorganization was present for both, as reported in similar cases for chiral PNA[28]. In particular, single strand 5L-PNAs was found to be quite preorganized, with a CD spectrum reminiscent of a PNA arranged in a P-helix form[28].

The PNA:DNA duplexes gave CD spectra consistent with that of achiral PNA, with a maximum at 265nm and a minimum at 244nm, suggesting that the highly constrained PNA did not induce a

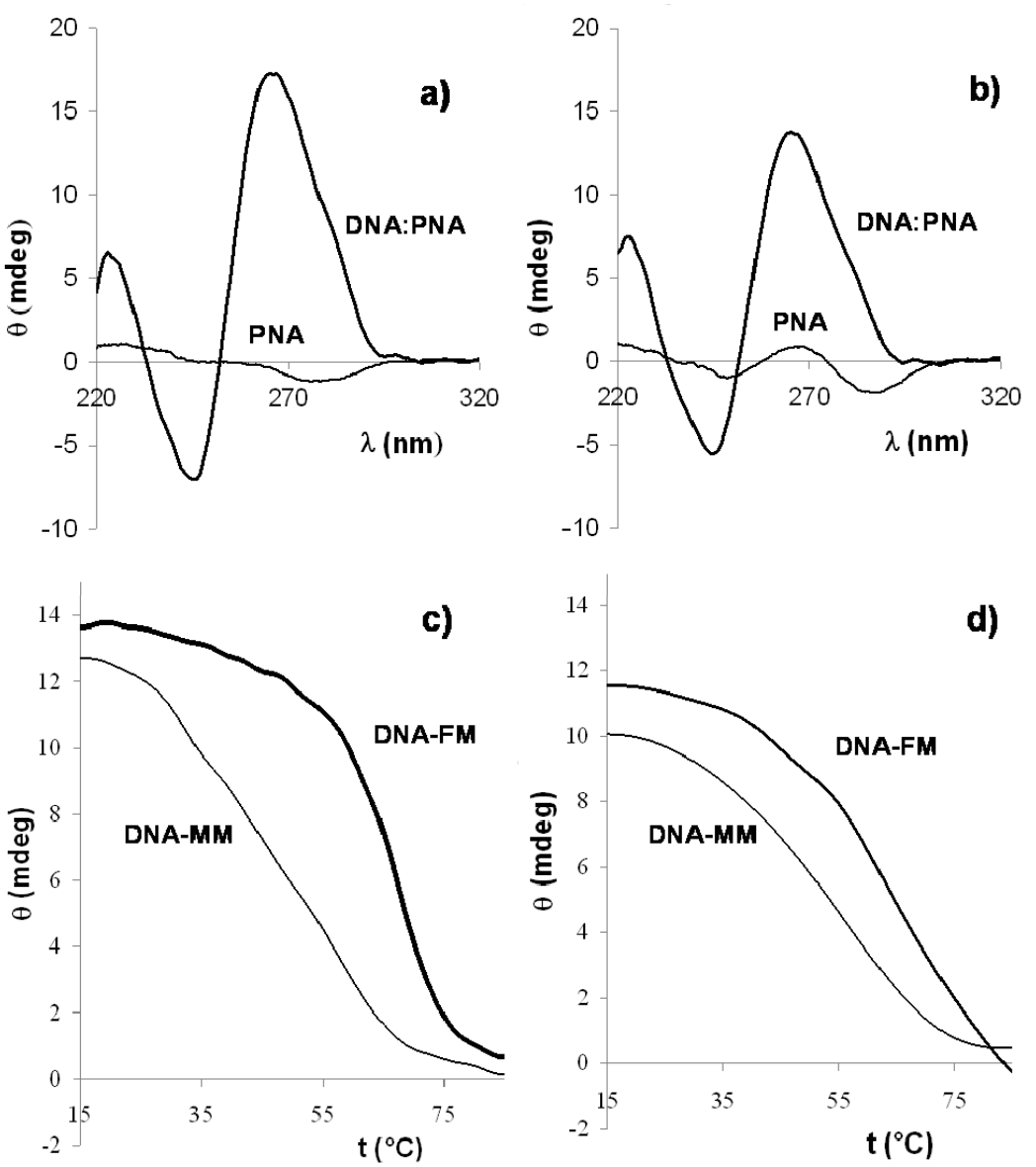


Figure 2.6: CD spectra of a) PNA 2.2 and b) PNA 2.3 alone (thin line) and of the duplex with full complementary DNA (thick line); c, d) melting curves at 260nm for the full-match and single mismatch (G-A) DNA for PNA 2.2 (c), and PNA 2.3 (d). All measurements were done in PBS buffer, pH 7; concentration of each strand was 5 μM. Cell length was 1 cm.

distorted conformation of the duplex, in agreement with the effect of 2D and 5L stereochemistry of the chiral monomers, which are both compatible with right-handed helices[20].

The CD melting curves of the duplexes recorder at 260nm were found to be consistent with those obtained by UV absorption (Figure 2.6,c-d), with a significant difference in the recognition of the single mismatch, which, also in this case, was more pronounced for **PNA 2.2** (Table 2.1).

The higher recognition properties of the 2D chiral-box model could be due to the different position of the substituent in the PNA backbone, which is attached to a carbon atom between two rigid amidic groups for the 2D-derivative, whereas in the case of the 5L derivative, the substituent is located on the aminoethyl moiety, which allows a higher conformational freedom. The presence of a mismatch in the central base of the 2D chiral-box could induce a conformational change which then affect the proper conformation of the monomer and of adjacent residues. If these are also highly constrained, the overall conformation of the "chiral-box" segment would be highly distorted and give rise to less stable PNA:DNA duplex. This effect is less pronounced in the case of 5L monomers, due to the possibility for the flanking monomers to better adapt to the distorted conformation, therefore preserving the positive electrostatic interactions which stabilize the duplex.

2.2.3 - DNA recognition by microarrays

The recognition properties of the two PNAs were also evaluated on a solid support, using the microarray technology which was previously used for the analysis of multiple sequences, including point mutations[29].

PNAs were spotted on the surface of activated slides using solutions of two different concentrations

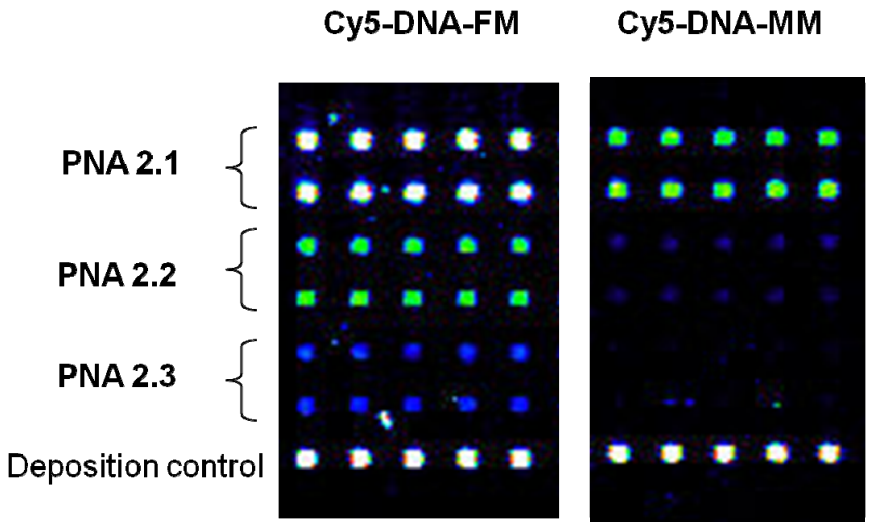


Figure 2.7: Microarray analysis of Cy-5 labeled DNA oligonucleotides: full-match DNA-FM (left side) and single-mismatch DNA-MM sequences (right side). Hybridization was performed in SSC buffer (0.1 μM of DNA in 2× (SSC) solution and 0.1% SDS). Images were obtained using a ScanArray™ Express Microarray Scanner, with λ_{ex}= 646nm and λ_{em}= 664nm.

(30 and 50 μM), both producing similar results (Figure 2.7), suggesting that the concentration was sufficient in both cases to saturate the active sites of the slide. A set of the achiral, 2D chiral-box and 5L chiral-box PNAs was spotted in duplicate on the same slide, together with a Cy5-labelled oligonucleotide as a control of deposition. Using a multi-chamber silicone gasket, one set of spots was then hybridized with the full-match oligonucleotide, and the other with the mismatched one, each bearing Cy5-labelling.

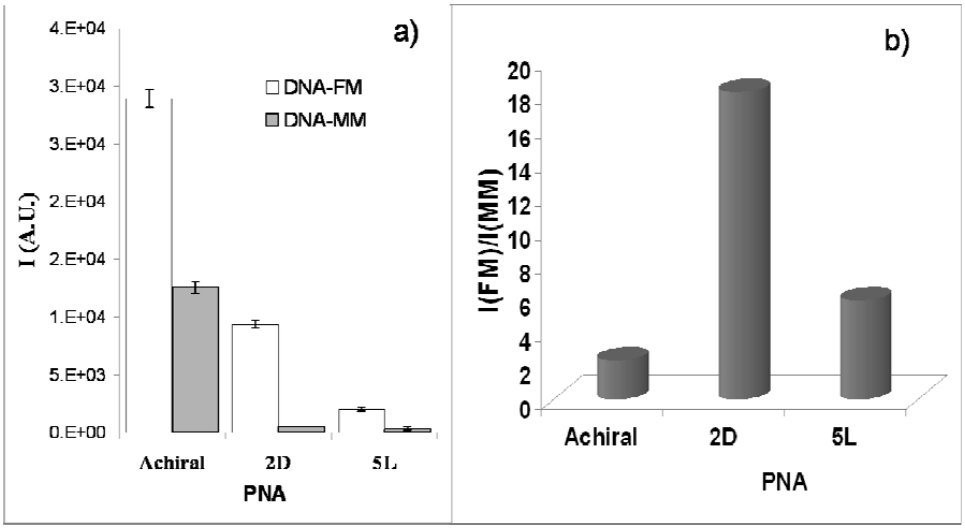


Figure 2.8: a) Quantitative analysis of microarray signals obtained with achiral PNA 2.1, PNA 2.2 and PNA 2.3 with full match (DNA-FM, white bars) and mismatched DNA (DNA-MM, grey bars); vertical bars indicate standard deviations. b) Selectivity ($I_{full\ match}$/ $I_{mismatch}$) observed in the microarray hybridization for the various PNA.

The results, shown in Figure 2.7, show that under the conditions used, the achiral **PNA 2.1** gave rise to the more intense signal with the full-match DNA. **PNA 2.2** gave a stronger signal than **PNA 2.3**, which showed very weak hybridization. Therefore, the stability of the PNA:DNA duplexes under these conditions were in a different order than those observed in solution, suggesting that electrostatic interactions which stabilize the duplex in solution can be affected by the matrix effect of the activated slide surface and by the additives normally used for hybridization. The hybridization with the mismatched oligonucleotide gave results in line with the sequence selectivity observed in solution, confirming the superiority of **PNA 2.2** in the discrimination of a single point mutation. Quantitative evaluation of the sensitivity and selectivity of the three different systems is show in Figure 2.8.

2.2.4 - Origin of differences in selectivity

It was previously established that stereoselectivity in the DNA binding ability was determined mainly by the intramolecular steric interactions between the PNA side chains and the DNA backbone, as evaluated by the structural features of PNA:DNA duplexes[21,30]. However, the different behavior of the 2D- and 5L-chiral-box described previously requires further discussion. Using the crystal structure of a 2D chiral box PNA-DNA duplex as a scaffold, and introducing arginine side chains either with the 2D or 5L stereocenter, the schematic model of the PNA-DNA interaction represented in Figure 2.9 can be proposed.

It can be noted that there is a major difference between the two types of modifications: while the side chains in the 2D model are pointing toward the major groove, in the 5L model they are directed toward the minor groove. This is in line with more favorable electrostatic interactions with the phosphate groups, due to a shorter distance between charges, which generates higher stability of the PNA-DNA duplex. The higher sequence selectivity observed in the case of 2D chiral box can be due to the position of the side chain, which is linked to the α-carbon of the more rigid glycine moiety of the PNA backbone, whereas the side chain in the 5L derivative is placed in the more flexible aminoethyl group.

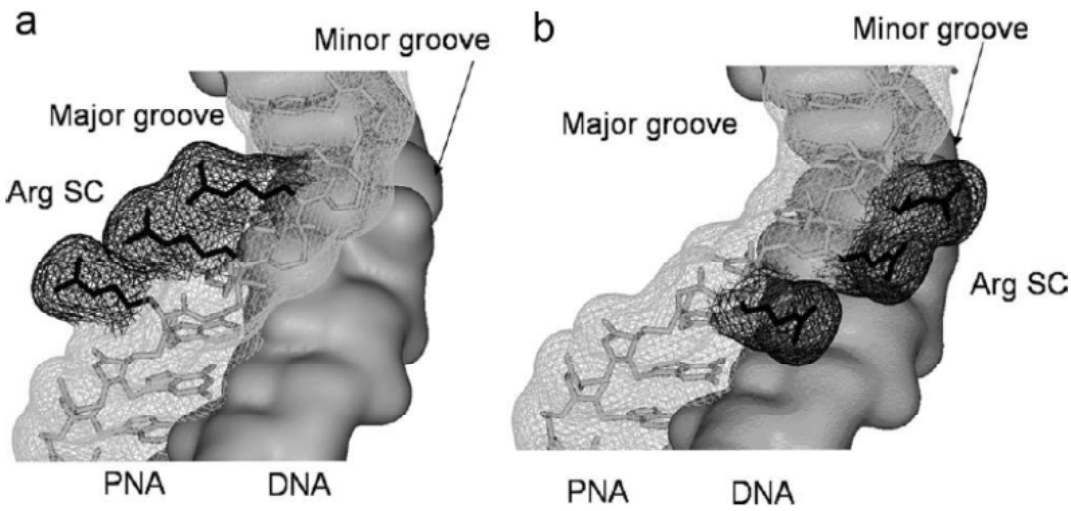


Figure 2.9: Model system for the 2D (a) and 5L (b) chiral box Arginine-based PNA, showing the disposition of the Arginine side chains (Arg SC) relative to PNA:DNA major and minor grooves (based on the crystal structure reported in Ref. 24).

Distortion of the former will likely generate a conformation in which the side chains collide with each other, whereas in the case of the 5L derivative, they can be rearranged and the eventual

repulsive interactions are compensated by the electrostatic interactions with the negative potential of the minor groove.

Since the side chains in the 5L model are closer to the DNA backbone, however, eventual surface-PNA interactions on this side will strongly hamper the interaction with DNA, while the same effect occurring on the major groove side would be better tolerated, thus leading to a better performance on the microarray system.

2.2.5 - Conclusions

Chirality of PNAs can affect DNA recognition in a positive way by increasing signal sensitivity and sequence selectivity [25]. In particular, the introduction of a stereogenic center derived from D-amino acid in the C-2 carbon or from L-amino acid at the C-5 carbon of the PNA backbone has turned out to be very effective in inducing higher affinity and increased selectivity. This effect was magnified when several chiral monomers were used to form C-2 substituted chiral-box PNAs, which were used in combination with several techniques for the specific detection of point mutations and single nucleotide polymorphisms.

Here we compared the effect of 2D- and 5L- chiral modified PNAs in the chiral-box model to increase selectivity in solution and on solid surfaces. The results showed that **PNA 2.3** was superior in terms of binding affinity in solution, while **PNA 2.2** model was superior in performances when recognition of single nucleotide polymorphisms were considered both in solution and in the microarray format. The latter information can be very precious for the design of new devices for the recognition of point mutations or single nucleotide polymorphisms in biomedical or food analysis.

These results have been the subject of a scientific publication in the Journal Chirality[31].

2.3 - Pyrene-based-PNA switching probes

Systems bearing one or more pyrene units linked along a PNA chain strand have spectrofluorimetric properties which are modulated by the conformation of the PNA and by the complexation with

DNA/RNA. In particular, two pyrene units can interact through π-π stacking to generate an excimeric complex with a red shift emission compared to that of the pyrene alone.

This property can be utilized for the detection of DNA strands in solution. In this work we exploited the potentiality of base modification with an azide group for generating pyrene based fluorescent probes. As target sequence we have chosen the DNA tract corresponding to the W1282X single point mutation of the CF gene, related to cystic fibrosis disease. One or more pyrene residues were introduced in PNA, and the different abilities of these probes were compared in terms of sensitivity and selectivity of discrimination between the wild type gene and the mutated gene.

2.3.1 - Base-modified monomers and PNA synthesis

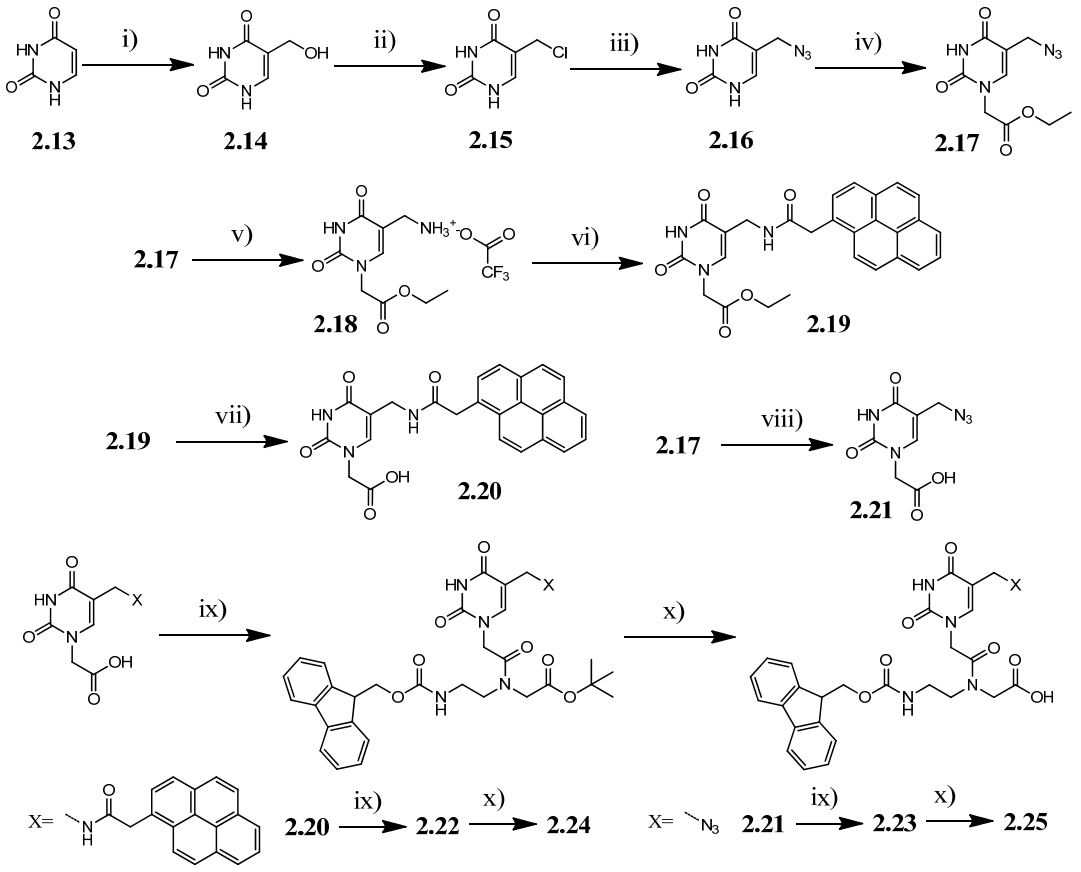
Two different approaches were followed for the introduction of the pyrene units in the PNA strand. In the former case (pyrene insertion by on-resin modification) a modified uracil bearing an azide group was introduced in the PNA strand during solid-phase synthesis. The azide group was chosen since it is equivalent to a masked amine that can be deprotected using orthogonal Staudinger reaction. In the latter case (pyrene insertion by preformed monomer) a modified monomer, in which the pyrene unit was already linked to the modified uracil was used for solid-phase synthesis.

The former approach provides a modification that can be used in both Boc- or Fmoc-based chemistry and that can be efficiently used to build a library of derivatives avoiding the need of synthesizing a large number of different monomers, so it can also be used to obtain many different modifications with less synthetic effort. On the contrary, the latter approach is more suitable for automated synthesis.

The syntheses of both monomers were devised in order to use a common pathway of reactions (Scheme 2.3) via the formation of 5-azidomethyluracil. Uracil, was hydroxymethylated at 5-carbon, following a literature protocol,[32] to obtain 5-(hydroxymethyl)uracil **2.14**, that can be rapidly converted into the corresponding azide **2.16** after activation of the leaving group by substitution of the hydroxyl group with chloride under acidic conditions. The N-1 position was then alkylated with ethyl 2-bromoacetate to produce **2.17,** bearing the linker for connecting the nucleobase to the PNA backbone.

For the strategy of pyrene insertion during solid-phase synthesis, **2.17** was hydrolyzed to the corresponding acid **2.21**, which was subsequently linked to the Fmoc-protected PNA backbone

using EDC/DhBtOH as activating mixture, and hydrolyzed under acid conditions to obtain the final monomer **2.25**.



Scheme 2.3: synthesis of the two monomers used for the introduction of the pyrene unit on PNA: i) CH_2O, Et_3N in H_2O, 81%; ii) HCl 37%, 80%; iii) NaN_3 in DMF, 90%; iv) $BrCH_2COOEt$, K_2CO_3 in DMF, 41%; v) PPh_3, H_2O in THF, 71%;vi) 1-pyreneacetic acid, HBTU, DIPEA in DMF, 66%; vii) NaOH in H_2O/MeOH 1:1, 91%; viii) NaOH in H_2O/MeOH 1:1, 96%; ix) EDC·HCl, DhBtOH, DIPEA, tert-butyl 2-((2-Fmoc-aminoethyl)amino)acetate hydrochloride in DMF, 68-96%; x) TFA in DCM, 86-86%.

For the insertion of pyrene on the monomer, the azide was first reduced under Staudinger conditions to the corresponding amine **2.18,** on which 1-pyreneacetic acid was linked using HBTU/DIPEA as condensing agent. The ester **2.19** was then hydrolyzed to the acid **2.20**, and linked to the backbone with the procedure described above the monomer; by ester hydrolysis, the PNA monomer **2.24**, containing the pyrene moiety, was obtained. Compared with the former synthetic pathway, the latter strategy was affected by a reduced solubility of intermediates induced by the pyrene moiety, which decreased the yields of the coupling steps.

Using either of these procedures, a series of PNAs containing one to three pyrene units at different positions along the chain, were synthesized, using the sequence complementary to the W1282X mutated form of CF gene tract (Table 2.2). All PNAs were synthesized using standard Fmoc-based

manual solid phase protocol, using monomer **2.25** for **PNA 2.4**, containing a single pyrene unit, and monomer **2.24** for all the other PNAs bearing multiple pyrene units. The on-resin modification was carried out after PNA synthesis on the Fmoc protected sequence, by reducing the azide function with a mixture H_2O/THF/$P(CH_3)_3$ 1M in THF 3:2:1 (10 minutes, 2 times) and subsequent coupling of 1-pyreneacetic acid (10 equivalents) with HBTU/DIPEA as activating agent(Figure 2.10).

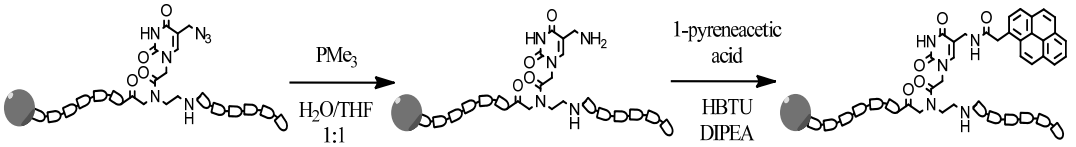


Figure 2.10: scheme of solid phase modification of monomer 2.25.

The crude products were purified by RP-HPLC and characterized by HPLC-MS, as described in the Experimental part. The quantification of the PNA was performed after evaluation of the molar extinction factor of **2.20**. Low yields obtained after purification are only related to the poor solubility of this derivatives in water.

Table 2.2: PNA sequences used for this study. Underlined letters are for modified monomers.

Name	PNA sequence	PNA synthesis	Yield
PNA 2.4	H-TCCTTCACT-Gly-NH_2	On-resin modification	10%
PNA 2.5	H-TCCTTCACT-Gly-NH_2	Preformed monomer	9%
PNA 2.6	H-TCCTTCACT-Gly-NH_2	Preformed monomer	10%
PNA 2.7	H-TCCTTCACT-Gly-NH_2	Preformed monomer	11%
PNA 2.8	H-TCCTTCACT-Gly-NH_2	Preformed monomer	15%
PNA 2.9	H-TCCTTCACT-Gly-NH_2	Preformed monomer	11%

2.3.2 - Thermal stability of PNA:DNA complexes

The introduction of a modification in a PNA strand can lead to different effects, from electronic to steric, which affect the interactions between the PNA and the complementary DNA strand. Substitution at the C-5 position of the uracil ring allows positioning of the substituent in the direction of the major groove of the double helix, thus reducing the destabilization induced by steric factors; moreover the large aromatic portion introduced with the pyrene ring can interact with the flanking bases of the strand through π-π stacking interactions, thus stabilizing the complex formed.

For the evaluation of the sum of all these effects, we measured the melting temperatures of the complexes formed between the PNAs and a full matched (DNA FM, 5'-AGT GAA GGA-3') or a single mismatched (DNA MM, 5'-AGT GGA GGA-3') DNA, evaluating both melting and annealing processes. In Table 2.3 the T_m of transition between 18°C and 50°C are reported, showing a complex behavior of the different PNAs in relation to the number and position of the pyrene units along the PNA strand.

Table 2.3: UV melting temperature of PNA:DNA complexes. DNA FM: 5'-AGTGAAGGA-3'; DNA MM: 5'-AGTGGAGGA-3'; all measurement were done in PBS pH 7 at 1μM strand concentration except for unmodified PNA measurements that were conducted at 5μM strands concentration. *literature data.

PNA	Tm DNA FM (°C)		Tm DNA MM (°C)	
	Melting	Annealing	Melting	Annealing
Unmodified*	34	-	24	-
2.4	26	22	20	< 18
2.5	39	31	19	< 18
2.6	24	< 18	< 18	< 18
2.7	33 (broad)	32	22	< 18
2.8	28	25	22	< 18
2.9	Not detectable	Not detectable	Not detectable	Not detectable

Comparing **PNA 2.4** (with a single pyrene unit) with the unmodified PNA, we can see that the introduction of the aromatic system led to destabilization of both complexes with a reduction in the selectivity of discrimination of SNPs. Introducing a second pyrene unit, the behavior of the PNAs in term of stability and selectivity is strongly related to the relative position of the different modification. Rationalizing, the presence of an additional N-terminal modification confers a higher stability to the FM complex compared to that at C-terminus (comparison between **PNA 2.5** and **PNA 2.8**). Moreover, the introduction of the modification in correspondence of mismatch position lead to a higher selectivity between FM and MM complexes (comparison between **PNA 2.5** and **PNA 2.7**). The stabilization of the complex by the terminal modification can be related to a sort of terminal clamping due to π-π stacking between the electron rich pyrene unit and the electron poor nucleobases; on the other hand the increased selectivity induced by the central modification can be related to steric factors or reduced stabilization induced by the distortion of the complex in correspondence of the mismatch. Finally, steric hindrance, or a distortion induced by a possible stacking interaction of the two consecutive modifications can be the reason why **PNA 2.6** displays an even lower stability (compared to **PNA 2.5).** Adding one more pyrene unit (**PNA 2.9**), no melting was detectable in this region suggesting a too strong interaction not explainable only with

the effects previously described for the PNAs bearing only two units, a possible melting transition was observed at Tm higher than 50°C, but it was hidden by another transition of the probe itself. For these reasons **PNA 2.5**, which presents both central and N-terminal modifications, appear to be the best candidate for the discrimination between the two sequences.

From the data reported in Table 2.3 it is also possible to note the presence of a significant hysteresis between the profile obtained during melting and annealing processes can be noticed (from data in Table 2.3 or shown in Figure 2.11 for **PNA 2.5**) suggesting the presence of kinetic effects typical of the formation of triplex structures PNA:DNA:PNA. Stabilities were also evaluated following the decrease in excimer signal detected by fluorescence studies (described below) in which fluorescence spectra were recorded every 5°C starting from 10°C to 50°C, obtaining similar results to those obtained by UV measurement (as show in Figure 2.11 for **PNA 2.5**, and in Figure 2.13 for the other PNAs).

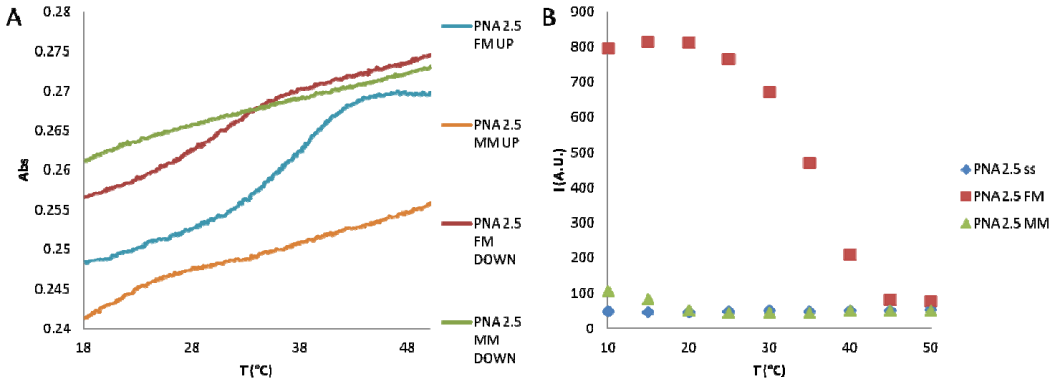


Figure 2.11: melting curves for the full-match and single mismatch DNA (G-A) for PNA 2.5; A) UV analysis at 260nm; B) Fluorescence analysis at 347nm excitation and 470nm emission as a function of temperature. All measurements were done in PBS buffer, pH 7; concentration of each strand was 1 μM. Cell length was 1cm for UV analysis.

2.3.3 - Preliminary fluorescence studies

Beside the discriminating properties induced by incorporation of pyrene moieties in PNA described above, we evaluated the fluorescence properties of these probes upon hybridization with DNA, since it is well know that the fluorescence of this fluorophore is strongly related to the environment[19], so that distortion of the complex formed between DNA and PNA can result in variation of the fluorescence properties (quantum yield and lambda switch). In a first assay we

compared the fluorescence of the probes alone and in presence of perfectly matched and singly mismatched DNA.

Due to the low melting temperature determined by UV analysis, all spectra were recorded every 5°C, ranging from 10°C to 50°C, to obtain a signal response in function of both sequence and temperature.

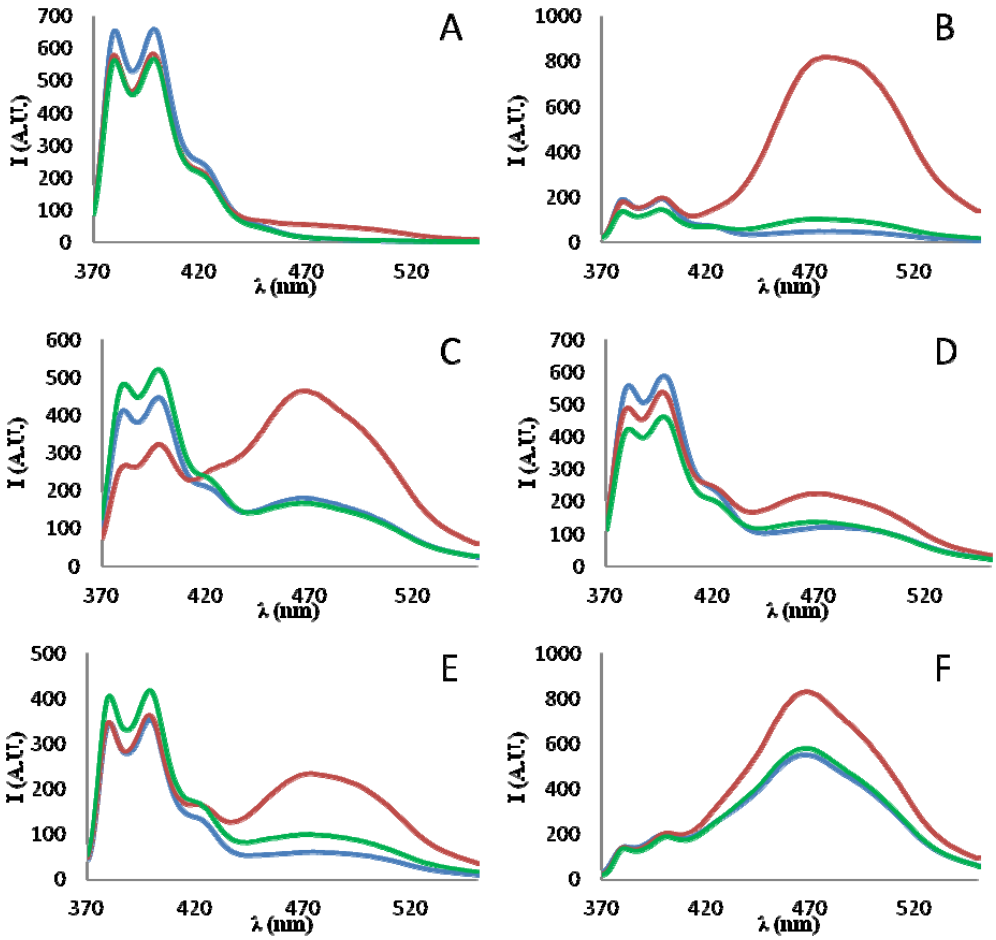


Figure 2.12: fluorescence studies for λ excitation= 347nm. A) PNA 2.4; B) PNA 2.5; C) PNA 2.6; D) PNA 2.7; E) PNA 2.8; F) PNA 2.9; blue lines: ssPNA ; red lines: PNA:DNA FM; green lines: PNA:DNA MM. All spectra were recorded at 1μM strand concentration in PBS pH 7.0 at 10°C. All spectra were obtained as mean of 3 different scans.

As shows in Figure 2.12 there is no a clear correlation between the fluorescence in the typical region of the pyrene emission (370÷430nm) and the position of the pyrene along the strand or the degree of complementarities of the PNA:DNA complex. Looking at the spectra, it is possible to see in some cases an additional band at 470nm, due to the excimer formation between two pyrene units.

As this could be hypothesized for **PNA 2.6**, in which the two aromatic systems are located on consecutive nucleobases, the presence of this band was not expected for the other PNA sequences, in particular for **PNA 2.4** which contains only one pyrene unit in the sequence. Looking at this band, the formation of the excimeric complex is strongly favored only in the presence of a perfectly matched DNA (apart **PNA 2.9** that show a strong excimer band even in the absence of DNA), whereas the presence of single mismatched DNA induce only a slight- or no increase of the signal

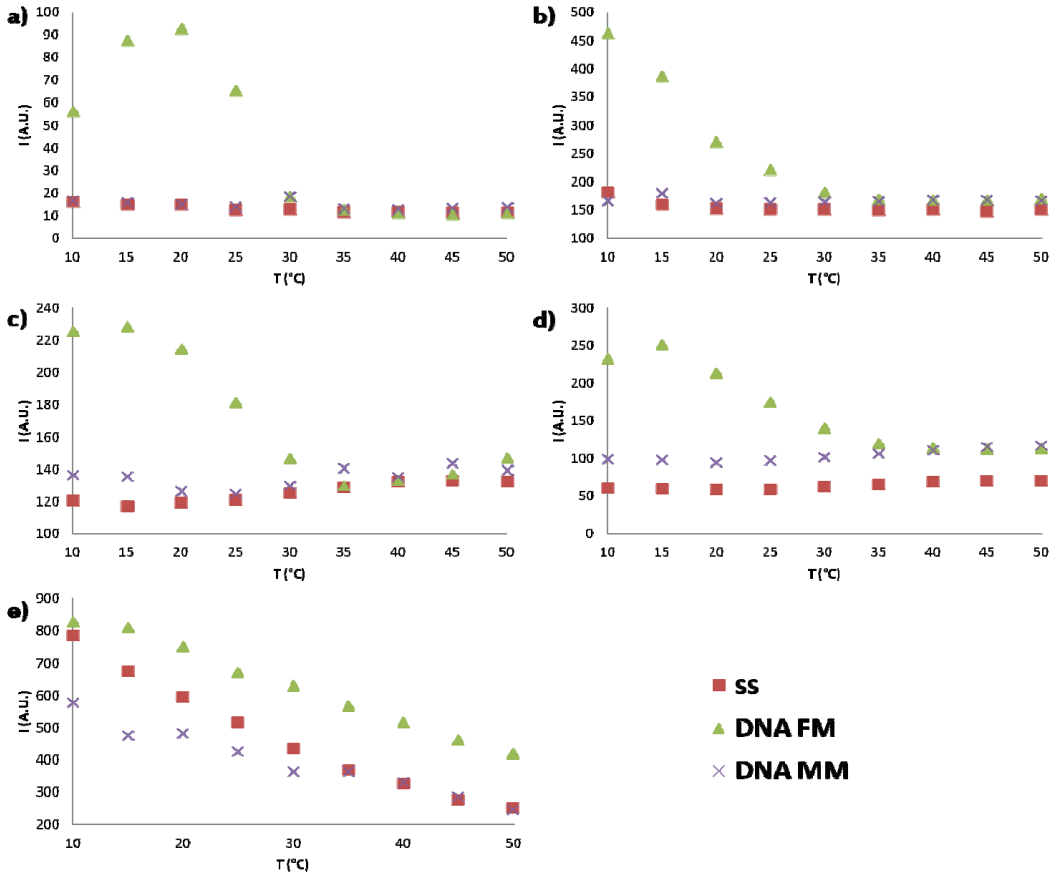


Figure 2.13: fluorescence analysis at 347nm excitation and 470nm emission as a function of temperature; a) PNA 2.4, b) PNA 2.6, c) PNA 2.7, d) PNA 2.8, e) PNA 2.9. All measurements were done in PBS buffer, pH 7; concentration of each strand was 1 µM.

2.3.4 - Titration and discrimination of PNA:DNA complexes

The behavior showed in the preliminary fluorescence studies together with the hysteresis determined by UV analysis, suggested the possible formation of more complex structures than the

simple PNA:DNA double helix. For this reason, we evaluated the effective stoichiometry of the complex, titrating **PNA 2.5** and **PNA 2.6** (the two systems with the higher differences in the excimer signal) with FM or MM DNA.

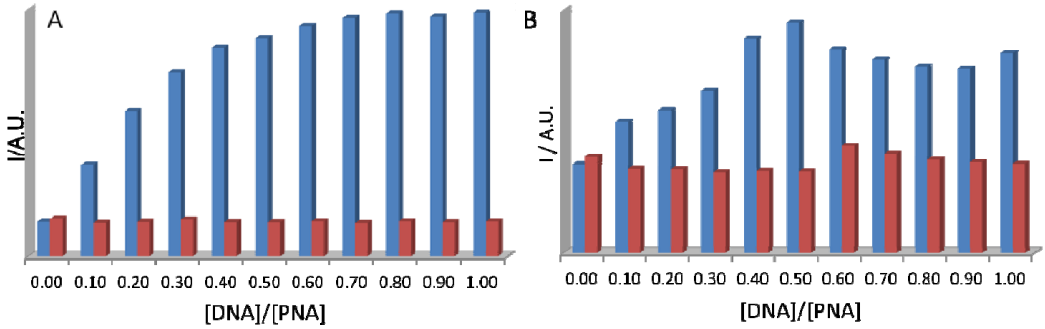


Figure 2.14: fluorescence titration with DNA FM (blue bars) and DNA MM (red bars); A) PNA 2.5, B) PNA 2.6. All measurements were done in PBS pH 7.0 at 1μM PNA strand concentration. λ excitation: 347nm; λ emission: 470nm.

As shown in Figure 2.14A for **PNA 2.5** in the presence of DNA FM the excimer signal strongly increase reaching a molar ratio DNA/PNA of 1:2, then the signal have only a modest variation upon the addition of more DNA aliquots, whereas the addition of DNA MM did not induce any variation of the excimer signal. This result suggests the formation of a PNA:DNA:PNA triplex structure, in fully agreement with all the other information obtained from former analysis. Moreover, looking at the melting profile in Figure 2.11 for this PNA, the behavior of the fluorescence signal perfectly fit with the partial melting of the **PNA 2.5**:DNA FM complex and the full melting of the **PNA 2.5**:DNA MM complex. The same behavior was obtained for **PNA 2.6**.

One step further, we also tested the performances of this two system in the recognition of the target DNA in presence of the single mismatched DNA, and the ability to recognize and emit fluorescence in presence of a longer DNA strand (19mer, DNA FM long: 5'-GCA ACA GTG AAG GAA AGC C-3', DNA MM long: 5'-GCA ACA GTG GAG GAA AGC C-3', Figure 2.15).

Concerning the ability of PNAs to recognize the presence of the target DNA in the presence of an interfering one (Figure 2.14B for **PNA 2.5** and **PNA 2.6**), the total absence of signal variations from the ssPNA after hybridization with DNA MM and the strong variation of signal after hybridization with DNA FM make these probes very interesting for detection of the point mutation by fluorimetric experiments. Adding half equivalent of DNA FM to the solution of PNA and DNA MM 1μM the signal raised and it was almost identical to the signal of the FM complex only (data not shown). This suggests a good discrimination ability for both PNA systems tested in the

recognition of the target DNA, even in the presence of a strong interfering sequence, allowing its application to the discrimination of point mutations for heterozygous subjects.

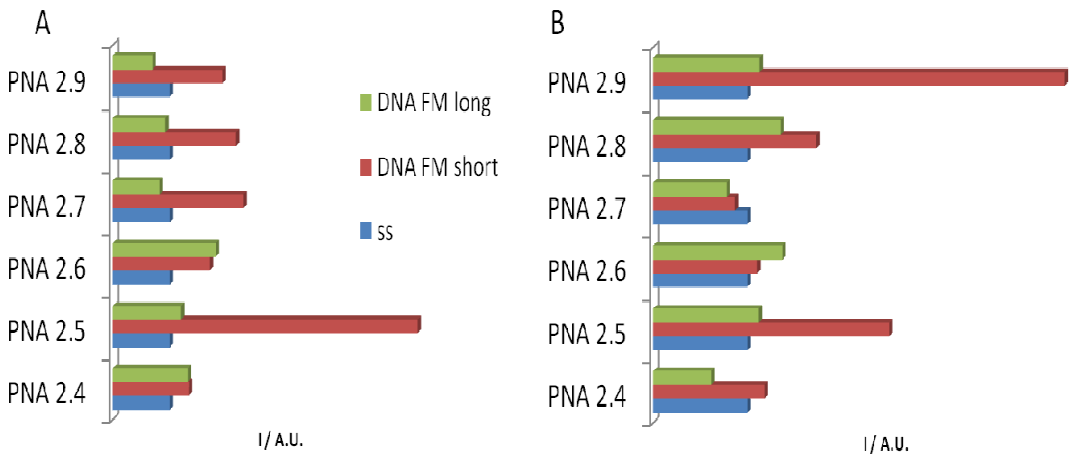


Figure 2.15: fluorescence evaluation upon complexation with different DNA sequences. A) 25°C ; B) 37°C. All data were obtained from PBS solution pH 7.0, 1μM PNA, 0.5μM DNA, λ excitation: 347nm, λ emission: 470nm. Signals were normalized at single strand excimer emission.

From the data obtained in the evaluation of fluorescence after hybridization with long sequences, it is possible to notice a drop in the fluorescence signal in all cases except for **PNA 2.6**. The drop can be induced by the impossibility to form the triplex structure that can be induced by the secondary structure of this longer DNA or by the negative steric repulsion between terminal pyrene and flanking nucleobases in the DNA strand, in fact the only probe able to maintain a signal (and the triplex structure) is **PNA 2.6**, which does not present modified units at the edge of the sequence, avoiding this possible destabilization. Comparing the data obtained at 25°C with those obtained at 37°C, it is possible to see that this PNA in the presence of long DNA is still able to have excimer fluorescence significantly different from the single strand, whereas this does not occur in the presence of a short DNA strand. Increasing the temperature, also **PNA 2.8** showed a raise in the excimer signal, but no explanation about this phenomenon were found from previous data.

Taken together, these fluorescence data suggest that the π-π interaction between different pyrene units shifts the equilibrium of complexation to the formation of the PNA_2:DNA triplex structure (Figure 2.16). For **PNA 2.5** the triplex structure allows the pyrene units to have a correct geometry for a good π-π stacking, while for **PNA 2.6** the position of the four pyrene units in close proximity leads to an unfavorable interaction due to steric hindrance, not allowing them to adopt the correct disposition. Terminal pyrenes can further interact with terminal nucleobase increasing the stability of the complex[33], this explaining the complete loss of signal by increasing the DNA strand length.

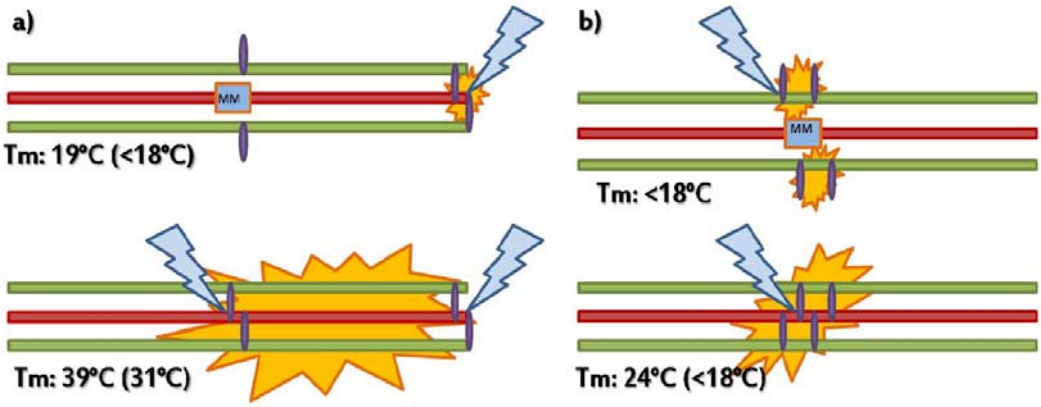


Figure 2.16: proposed triplex formation excimer induction; a) PNA 2.5, b) PNA 2.6. MM complexes in the upper row; FM complexes in lower row.

2.3.5 - Conclusions

PNA-based pyrene switching probe systems directed against the W1282X single point mutation, related to cystic fibrosys (CF), were designed, synthesized and characterized. The properties of the different PNAs in term of DNA complex stability and selectivity through discrimination of the SNP were compared and their responses in term of fluorescence signal upon hybridization were also evaluated.

By titration analyses it was demonstrated that the recognition of the target occurs through triplex formation induced by the additional stabilization provided by the π-π stacking interaction between pyrene units. In fact this additional interaction between two different strands can help for the stabilization of the structure, shifting the equilibrium from duplex structure to the more complex triplex structure. From the UV melting analyses it can be inferred that N-terminal rather than C-terminal modification induces the formation of more stable complexes. This can be explained by the presence of an adenine residue at the 3' side of the strand that can distort the system not allowing the perfect disposition of the aromatic systems, thus preventing a strong interaction, which translates in a lower complex stability.

Among all the models tested, **PNA 2.5** seemed to be the best system in term of sensitivity and selectivity for the discrimination of a single point mutation placed in the middle of the sequence. However, this PNA does not show any of its properties when the target DNA is elongated to more realistically simulate DNA in cellular systems (both extracted or PCR-amplified). Luckily, **PNA**

2.6, despite showing less pronounced discriminating properties, maintained its recognition ability even when the target DNA was elongated, suggesting its possible application to more complex systems.

2.4 - Experimental Section

General. Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, Link, ASM and used without further purification. All reactions were carried out under a nitrogen atmosphere with dry solvents under anhydrous conditions, unless otherwise noted. Anhydrous solvents were obtained by distillation or anhydrification with molecular sieves. Reactions were monitored by TLC carried out on 0.25mm E. Merck silica-gel plates (60F-254) by using UV light as visualizing agent and ninhydrin solution and heat as developing agents. E. Merck silica gel (60, particle size 0.040÷0.063 mm) was used for flash-column chromatography. NMR spectra were recorded on Bruker Avance 400 instrument and calibrated by using residual undeuterated solvent as an internal reference. The following abbreviations were used to explain the multiplicities: s=singlet, d=doublet, t=triplet, q=quartet, m=multiplet, and br=broad. IR spectra were measured using a FT-IR Thermo Nicolet 5700, in transmission mode using KBr or NaCl. HPLC-UV-MS were recorded by using a Waters Alliance 2695 HPLC with Micromass Quattro microAPI spectrometer, a Waters 996 PDA and equipped with a Phenomenex Jupiter column (250x4.6mm, 5µ, C18, 300Å) (method A, 5 minutes in H_2O 0.2% FA, then linear gradient to 50% MeCN 0.2% FA in 30 minutes at a flow rate of 1ml/min). PNA oligomers were purified with RP-HPLC using a XTerra Prep RP_{18} column (7.8x300mm, 10µm) (method B, linear gradient from H_2O 0.1% TFA to 50% MeCN 0.1 % TFA in 30 minutes at a flow rate of 4.0ml/min).

Boc-glycinal, Boc-Gly-H, D-Arg(Tos)-OMe·HCl and Boc-L-Arg(Tos)-H were synthesized as described previously[20,34].

Boc-PNA(2D-Arg$_{(Tos)}$)-OMe (2.3). Boc-Glycinal (0.76g, 4.81mmol) and Arginine methyl ester hydrochloride (1.52g, 4.01mmol) were dissolved in 30ml MeOH with addition of DIPEA (662µl, 4.01mmol). The solution was stirred for 30 minutes at r.t., then cooled to 0° C with an ice bath and $NaBH_3CN$ (0.30mg, 4.81mmol) and acetic acid (275µl, 4.81mmol) were added. The reaction was allowed to stir for 30 minutes at 0°C, then at room temperature for 2.5h and was monitored via TLC. The solvent was evaporated and the residue was taken up with 200ml AcOEt and washed with saturated $NaHCO_3$ (2x200ml) and brine (200ml). The organic layer was dried over Na_2SO_4, filtered and evaporated to afford an oil that was purified via flash chromatography (from AcOEt to AcOEt/MeOH 9:1) to afford **2.3** as a white foamy solid (1.32g, 68%); **TLC (AcOEt/MeOH 9:1) Rf:** 0.31; **^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 7.64 (2H, d, J= 9Hz), 7.29 (2H, d, J= 6Hz), 7.03 (1H, br s), 6.72 (1H, br s), 6.56 (1H, br s), 3.61 (3H, s), 3.34 (1H, s),

2.86÷3.23 (6H, m), 2.35 (3H, s), 1.61÷1.26 (13H, m); **^{13}C NMR (DMSO-d^6, 300MHz) δ(ppm):** 174.1, 156.7, 156.6, 141.7, 141.0, 129.1, 125.9, 79.6, 60.6, 51.9, 47.9, 40.6, 29.6, 29.2, 28.3, 25.5, 21.4; **MS (ESI, MeOH):** *m/z* calcd for $C_{21}H_{35}N_5O_6S$ [M]: 485.23080, found: 486.5 $[M+H]^+$, 508.5$[M+Na]^+$, 524.4 $[M+K]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3421.4 (s), 3350.6 (s), 2976.5 (m), 1734.8 (s), 1701.0 (s), 1623.8 (m), 1576.8 (m), 1550.4 (s), 1366.5 (m), 1254.2 (s), 1168.6 (s).

Boc-PNA(2D-Arg$_{(Tos)}$)-OH (2.4). To a stirred solution of **2.3** (2.60 g, 5.37mmol) in THF (50ml) was added a solution of NaOH(2.14 g, 53.7mmol) in water (50ml). The reaction mixture was stirred for 1h. The THF was then evaporated and the pH of the solution was lowered to 7.1 with a solution of HCl to induce the precipitation of the zwitterionic form. The solution was cooled at 4°C for 2 hours, then filtered over Buchner to afford **2.4** as a white solid (2.44g, 97%); **TLC (AcOEt/MeOH 9:1) Rf:** 0.00; **^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 7.97 (1H, br s), 7.64 (2H, d, J= 9Hz), 7.29 (2H, d, J= 6Hz), 6.98 (1H, br s), 3.29÷3.05 (5H, m), 2.79÷2.69 (3H, m), 2.34 (3H, s), 1.73÷1.26 (13H, m); **^{13}C NMR (DMSO-d^6, 75MHz) δ(ppm):** 170.8, 156.8, 155.5, 141.6, 140.8, 128.9, 125.4, 77.9, 61.1, 45.9, 37.0, 28.0, 27.2, 24.9, 20.7; **MS (ESI, MeOH):** *m/z* calcd for $C_{20}H_{33}N_5O_6S$ [M]: 471.21515, found: 472.4 $[M+H]^+$, 490.4 $[M+Na]^+$, 510.4 $[M+K]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3421.4 (s), 3356.8 (s), 2975.9 (w), 1710.1 (s), 1689.5 (s), 1609.6 (s), 1577.7 (s), 1539.7 (m), 1390.4 (w), 1261.3 (s), 1167.4 (s).

Boc-PNA(2D-Arg$_{(Tos)}$)-Fmoc-OH (2.5). In a round bottom flask **2.4** (2.42 g, 5.00mmol) was dispersed in 170ml DCM, then BTSA (3.7ml, 15mmol) and DIPEA (1.8ml, 11mmol) were added. The reaction mixture was stirred at room temperature for 30 minutes until the solution became nearly clear, then Fmoc-Cl (1.55g, 5.99mmol) was added at 0°C. The mixture was stirred for 10 minute at 0°C then for 2h at room temperature. The reaction was quenched with addition of 25ml MeOH, the solvent evaporated and the residue was taken up with 150ml DCM, washed with satured $KHSO_4$ (2x150ml) and brine (150ml). The organic layer was dried over Na_2SO_4, filtered and evaporated to afford an oil that was purified via flash chromatography (from DCM to DCM/MeOH 9:1) to afford **2.5** as a pale yellow solid (2.85g, 82%); **TLC (DCM/MeOH 9:1) Rf:** 0.31; **MP (°C):** decompose without melt at 142°C;**^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 7.92÷7.82 (2H, m), 7.70÷7.59 (4H, m), 7.18÷7.45 (6H, m), 4.16÷4.33 (3H, m), 3.38 (1H, br s), 3.14÷2.87 (2H, m), 2.30 (3H, s), 1.42÷1.23 (13H, m); **^{13}C NMR (DMSO-d^6, 75MHz) δ(ppm):** 175.0, 156.7, 155.9, 155.5, 143.8, 141.7, 140.8, 140.5, 128.8, 127.5, 127.1, 126.9, 125.4, 125.0, 77.3, 66.4, 60.8, 48.3, 46.6, 45.2, 29.9, 28.1, 26.6, 20.7; **MS (ESI, MeOH):** *m/z* calcd for $C_{35}H_{43}N_5O_8S$ [M]: 693.28323, found: 694.5 $[M+H]^+$, 716.4 $[M+Na]^+$, 732.4 $[M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 692.2749 $[C_{35}H_{42}N_5O_8S]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3425.6 (s), 3346.8 (m), 2974.8 (w), 1689.5 (m), 1684.9 (m), 1617.4 (s), 1550.5 (s), 1367.1 (m), 1252.4 (m), 1165.8 (s).

Boc-PNA(5L-Arg$_{(Tos)}$)-OMe (2.8). In around bottom flask Boc-Arg$_{(Tos)}$-H (687mg, 1.66mmol) and Gly-OMe·HCl (251mg, 2.00mmol) were dissolved in 20ml MeOH. The reaction mixture was cooled to 0° C with an ice bath and $NaBH_3CN$ (132mg, 2.00mmol) and acetic acid (144 µl, 2.00mmol) were added to the stirred solution. The reaction was allowed to stir for 30 minutes at 0°C, then at room temperature for 4h. The

solvent was evaporated and the residue was taken up with 200ml EtOAc and washed with saturated $NaHCO_3$ (3x100ml) and brine (100ml). The organic layer was dried over Na_2SO_4, filtered and evaporated to afford an oil that was purified via flash chromatography (from EtOAc to AcOEt/MeOH 9:1) to afford **2.8** as a white foamy solid(407mg, 50%); **TLC (AcOEt/MeOH 94:6) Rf:** 0.17; **MP (°C):** 58.6÷61.8°C; **1H NMR ($CDCl_3$, 300MHz) δ(ppm):** 7.75 (2H, d, J= 8.2Hz), 7.22 (2H, d, J= 8.2Hz), 6.45 (3H, br s), 5.00 (1H, d, J= 8,4Hz), 3.72 (3H, s), 3.70÷3.55 (1H, m), 3.44 (1H, d, J= 17,5Hz), 3.35 (1H, d, J= 17,5Hz), 3.18 (2H, br s), 2.70÷2.50 (2H, m), 2.38 (3H, s), 1.78 (1H, br s), 1.61÷1.43 (4H, m), 1.41 (9H, s); **^{13}C NMR ($CDCl_3$, 75MHz) δ(ppm):** 172.8, 156.8, 156.2, 141.8, 140.6, 129.1, 125.7, 79.0, 53.5, 52.9, 51.6, 50.3, 40.7, 30.1, 28.2, 25.6, 21.2; **MS (ESI, MeOH):** *m/z* calcd for $C_{21}H_{35}N_5O_6S$ [M]: 485.23080, found: 486.7 $[M+H]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 486.2375 $[C_{21}H_{36}N_5O_6S]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3431.1 (m), 3342.1 (m), 3158.2 (w), 2976.3 (m), 2951.6 (m), 2866.2 (w), 1743.0 (s), 1695.8 (s), 1623.32 (s), 1576.4 (s), 1549.7 (s), 1456.8 (m), 1436.9 (m), 1366.2 (m), 1253.6 (m), 1168.8 (s), 1132.3 (s).

Boc-PNA(5L-Arg$_{(Tos)}$)-G$_{(OBn)}$-OMe (2.9). In a round bottom flask 2-(O-Bn-Guanin-9-yl)acetic acid (501.3mg, 1.68mmol) was dissolved in 6ml DMF at 0° C, together with DHBtOH (273.2mg, 1.68mmol) and DIPEA (416 µl, 2.51mmol). The EDC·HCl (318.2mg, 1.66mmol) was added and the solution was stirred for 10 minutes at 0° C, for 20 minutes at room temperature, then a solution of **2.8** (406.7mg, 0.84mmol) in 6ml DMF was added to the mixture. The solution was stirred overnight and the DMF was then evaporated. The residue was taken up with 200ml AcOEt and washed with saturated $NaHCO_3$ (2x200ml), saturated $KHSO_4$ (2x200ml) and brine (200ml). The organic layer was dried over Na_2SO_4, the solvent evaporated under reduced pressure and the residue was purified via flash chromatography (AcOEt/MeOH 95:5) to afford **2.9** as a pale yellow foamy solid (511mg, 79%); **TLC (AcOEt/MeOH 9:1) Rf:** 0.21; **^{1}H NMR ($CDCl_3$, 300MHz, major rotamer) δ(ppm):** 7.70÷7.60 (3H, m), 7.39 (2H, d, J= 6.0Hz), 7.30÷7.20 (3H, m), 7.13 (2H, d, J= 7.8Hz), 6.50÷6.20 (5H, br m), 5.73 (1H, br s), 5.45 (2H, s), 4.97 (1H, d, J= 15.9Hz), 4.84 (1H, d, J= 15.9Hz), 4.30÷4.20 (2H, m) 3.90÷3.70 (1H, m), 3.61 (3H, s), 3.40÷3.00 (4H, m), 2.30 (3H, s), 1.40÷1.15 (13H, m); **^{13}C NMR ($CDCl_3$, 300MHz, major rotamer) δ(ppm):** 169.5, 167.4, 160.9, 159.5, 156.1 142.1, 142.0, 140.6, 140.4, 136.3, 129.2, 128.4, 128.1, 125.9, 114.5, 79.9, 68.1, 52.9, 52.7, 52.4, 51.5, 50.2, 48.8, 40.6, 28.8, 28.4, 25.7, 21.4; **MS (ESI, MeOH):** *m/z* calcd for $C_{36}H_{46}N_{10}O_8S$ [M]: 766.32208, found: 767.8 $[M+H]^+$, 789.7 $[M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 767.3292 $[C_{36}H_{47}N_{10}O_8S]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3351.7 (m), 3156.8 (w), 2975.6 (m), 2952.4 (m), 1745.5 (m), 1665.7 (m), 1615.5 (s), 1584.4 (s), 1551.5 (s), 1456.6 (m), 1365.9 (m), 1256.3 (s), 1132.3 (m).

Boc-PNA(5L-Arg$_{(Tos)}$)-C$_{(Cbz)}$-OMe (2.10). In a round bottom flask 2-(N-Cbz-cytosin-1-yl)acetic acid (502.9mg, 1.65mmol) was dissolved in 10ml DMF at 0° C together with DHBtOH (268.8mg, 1.65mmol) and DIPEA (423 µl, 2.56mmol). EDC·HCl (314.4mg, 1.64mmol) was added and the solution was stirred for 10 minutes at 0° C, for 20 minutes at room temperature, then **2.8** (408.2mg, 0.84mmol) was added to the mixture. The solution was stirred overnight and the DMF was then evaporated. The residue was taken up with 100ml AcOEt and washed with saturated $KHSO_4$ (2x50ml), saturated $NaHCO_3$ (2x50ml) and brine

(50ml). The organic layer was dried over Na_2SO_4, the solvent evaporated under reduced pressure and the residue was purified via flash chromatography (AcOEt/MeOH 95:5) to afford **2.10** as a pale yellow foamy solid (615mg, 95%); **TLC (AcOEt/MeOH 1:1) Rf:** 0.27; **MP (°C):** decompose without melt at 146°C; **^{1}H NMR ($CDCl_3$, 300MHz, major rotamer) δ(ppm):** 10.81 (1H, br s), 7.69 (2H, d, J= 8.1Hz), 7.55 (1H, br s), 7.35÷7.10 (8H, m), 6.55 (2H, br s), 5.56 (1H, br s), 5.16 (2H, s), 4.90÷4.45 (2H, m), 4.40÷3.85 (2H, m), 4.00÷3.70 (2H, m), 3.59 (3H, s), 3.41 (1H, br s), 3.17 (2H, br s), 2.32 (3H, s), 1.80÷1.30 (13H, m); **^{13}C NMR ($CDCl_3$, 75MHz, major rotamer) δ(ppm):** 169.5, 167.4, 163.3, 156.9, 156.2, 156.0, 152.5, 150.2, 141.8, 140.9, 135.0, 129.1, 128.6, 128.3, 128.2, 125.9, 95.6, 79.7, 77.2, 67.7, 52.8, 52.2, 50.9, 49.2, 48.9, 40.2, 28.8, 28.3, 25.6, 21.4.; **MS (ESI, MeOH):** *m/z* calcd for $C_{35}H_{46}N_8O_{10}S$ [M]: 770.30576, found: 771.8 $[M+H]^+$, 793.8 $[M+Na]^+$, 809.9 $[M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 771.5744 $[C_{35}H_{47}N_8O_{10}S]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3429.0 (m), 3338.6 (m), 3140.0 (w), 2976.0 (m), 1750.0 (s), 1664.9 (s), 1630.1 (s), 1561.9 (s), 1453.8 (m), 1367.6 (s), 1213.8 (s), 1132.3 (m).

Boc-PNA(5L-Arg$_{(Tos)}$)-G$_{(OBn)}$-OH (2.11). To a stirred solution of **2.9** (511mg, 0.67mmol) in 40ml THF was added a solution of $Ba(OH)_2 \cdot 8H_2O$ (315mg, 1.00mmol) in 40ml water. The reaction mixture was stirred for 15 minutes and checked via TLC. The THF was then evaporated and the pH of the solution was lowered to 4.5 with HCl. The solution was cooled at 4°C for 2 hours, then filtered over Buchner and dried under vacuum to afford **2.11** as a white solid (461mg ,92%); **TLC (AcOEt/MeOH 9:1) Rf:** 0.00; **MP (°C):** decompose without melt at 205°C;**^{1}H NMR (DMSO-d^6, 300MHz, major rotamer) δ(ppm):** 7.67 (1H, s), 7.62 (2H, d, J= 8.1Hz), 7.50 (2H, d, J= 7.8Hz), 7.45÷7.30 (3H, m), 7.26 (2H, d, J= 7.8Hz), 6.99 (3H, br s), 6.73 (1H, br d, J= 7.5Hz), 6.44 (2H , br s), 5.50 (2H, s), 4.84 (2H, s), 3.95÷3.55 (2H, m), 3.50÷2.90 (5H, m), 2.31 (3H, s), 1.50÷1.15 (13H, m); **^{13}C NMR (DMSO-d^6, 300MHz, major rotamer) δ(ppm):** 172.6, 167.5, 159.9, 159.6, 156.8, 154.9, 141.7, 140.9, 140.7, 140.4 136.7, 129.0, 128.3, 127.9, 125.5, 113.0, 77.5, 66.7, 52.9, 51.7, 49.5, 48.5, 43.5, 40.6, 35.7, 28.7, 28.2, 25.3, 20.8; **MS (ESI, MeOH):** *m/z* calcd for $C_{35}H_{45}N_9O_8S$ [M]: 751.31118, found: 753.8 $[M+H]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 753.3157 $[C_{34}H_{45}N_{10}O_8S]^+$, 751.2981 $[C_{34}H_{43}N_{10}O_8S]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3348.8 (m), 3219.5 (m), 2976.2 (w), 1721.4 (m), 1685.4 (s), 1615.8 (s), 1588.0 (s), 1456.1 (m), 1366.1 (m), 1255.3 (m), 1166.6 (m).

Boc-PNA(5L-Arg$_{(Tos)}$)-C$_{(Cbz)}$-OH (2.12). To a stirred solution of **2.10** (594mg, 0.77mmol) in 40ml THF was added a solution of $Ba(OH)_2 \cdot 8H_2O$ (364mg, 1.15mmol) 40ml in H_2O. The reaction mixture was stirred for 30 minutes and checked via TLC. The THF was then evaporated and the pH of the solution was lowered to 4.5 with HCl. The solution was cooled at 4°C for 2 hours, then filtered over Buchner and dried under vacuum to afford **2.12** as a white solid (419mg ,72%); **TLC (AcOEt/MeOH 9:1) Rf:** 0.00; **MP (°C):** decompose without melt at 191°C; **^{1}H NMR (DMSO-d^6, 300MHz, major rotamer) δ(ppm):** 10.81 (1H, br s), 7.81 (1H, d, J= 7.2Hz), 7.64 (2H, d, J= 8.1Hz), 7.45÷7.25 (7H, m), 7.02 (1H, d, J= 7.2Hz), 6.90÷6.60 (3H, br m), 5.19 (2H, s), 4.90 (1H, d, J= 15.9Hz), 4.77 (1H, d, J= 16.2Hz), 4.03 (1H, d, J= 17.1Hz), 3.87 (1H, d, J= 17.4Hz), 3.70÷2.90 (5H, m), 2.34 (3H, s), 1.60÷1.30 (13H, m); **^{13}C NMR (DMSO-d^6, 75MHz, major rotamer) δ(ppm):** 170.3, 167.1, 163.0, 156.5, 155.5, 154.9, 153.1, 150.5, 141.6, 141.0, 135.9, 129.0, 128.4,

128.1, 127.9, 125.5, 93.8, 77.9, 66.4, 51.4, 49.4, 48.5, 47.8, 40.2, 28.6, 28.1, 25.4, 20.8; **MS (ESI, MeOH):** *m/z* calcd for $C_{34}H_{44}N_8O_{10}S$ [M]: 756.29011, found: 757.8 $[M+H]^+$, 755.6 $[M-H]^-$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 757.3019 $[C_{34}H_{45}N_8O_{10}S]^+$, 779.2831 $[C_{34}H_{44}N_8O_{10}NaS]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3447.0 (m), 3350.3 (m), 2977.8 (w), 2933.5 (w), 1748.3 (m), 1653.4 (s), 1635.5 (s), 1558.7 (s), 1457.1 (m), 1368.6 (m), 1214.7 (s), 1132.3 (m).

5-hydroxymethyluracil (2.14). In a Shlenk tube uracil (10g, 89.5mmol) was dispersed together with paraformaldehyde (8.5g, 283mmol) in 200ml H_2O then tryethylamine (17.5ml, 125.9mmol) was added and the reaction mixture was heated overnight at 60°C. The excess formaldehyde was then purged with nitrogen in a 15% solution of sodium hypochlorite and the reaction solution was concentrated under vacuum to afford an oil that was taken up with 25ml H_2O; the product was precipitated adding 25ml ethanol, collected through sintered glass funnel and washed with cold ethanol. Subsequent aliquots of product were collected by repeating the concentration-precipitation process to afford **2.14** as white solid (10.3g, 81%); **TLC (MeOH) Rf:** 0.68, **MP (°C):** decompose without melt at 215°C; **^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 11.05 (1H, s), 10.71 (1H, s), 7.24 (1H, s), 4.85 (1H, br s), 4.10 (2H, s); **^{13}C NMR (DMSO-d^6, 75MHz) δ(ppm):** 163.8, 151.3, 138.2, 112.7, 55.8; **MS (ESI, MeOH):** *m/z* calcd for $C_5H_6N_2O_3$ [M]: 142.0378, found: 165.0 $[M+Na]^+$, 307.2 $[2M+Na]^+$, 141.0 $[M-H]^-$, 283.2 $[2M-H]^-$, 305.0 $[2M+Na-H]^-$; **Elemental composition:** calcd %C 42.26, %H 4.26, %N 19.41, found %C 42.24, %H 4.35, %N 19.33; **FT-IR (KBr) ν(cm^{-1}):** 3369 (m), 3189 (m), 3038 (m), 2865 (m), 1704 (s), 1672 (s).

5-chloromethyluracil (2.15). In a round bottom flask **2.14** (6.76g, 47.6mmol) was dissolved in HCl 37% (25ml). After 4h a precipitate was formed and the product was collected by filtration through a sintered glass funnel and dried overnight over P_2O_5 to afford **2.15** as white solid (6.09g, 80%). **TLC (MeOH) Rf:** 0.60, **MP (°C):** decompose without melt at 239°C; **^{1}H NMR (DMSO-d^6, 75MHz) δ(ppm):** 11.27 (1H, s), 11.04 (1H, d, J= 5.0Hz), 7.73 (1H, d, J= 6.0Hz), 4.40 (2H, s); **^{13}C NMR (DMSO-d^6, 300MHz) δ(ppm):** 162.8, 151.0, 142.4, 108.8, 39.8; **MS (ESI, MeOH):** *m/z* calcd for $C_5H_5ClN_2O_2$ [M]: 160.0040, found: 198.9 $[M+K]^+$, 159.0 $[M-H]^-$; **Elemental composition:** calcd %C 37.40, %H 3.14, %N 17.45, found %C 37.48, %H 3.20, %N 16.94; **FT-IR (KBr) ν(cm^{-1}):** 3305 (s), 3125 (m), 3044 (m), 2829 (m), 1759 (s), 1697 (s), 1656 (s), 1178 (s), 738 (m).

5-azidomethyluracil (2.16). In round bottom flask sodium azide (2.65g, 40.78mmol) was dispersed in 25ml dry DMF and cooled to 0°C with an ice bath. **2.15** (5.88g, 36.63mmol) dissolved in 80ml dry DMF was added dropwise over 30 minutes. After 2h the orange solution was quenched by adding 1.0ml HCl 37% and the resulting white suspension was concentrated under vacuum (T= 40°C). The crude product was sonicated with 25ml DCM and subsequently with 30ml cold water. **2.16** was collected through sintered glass funnel as white solid (5.01g, 90%). **TLC (MeOH) Rf:** 0.79, **MP (°C):** decompose without melting at 188°C; **^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 11.27(1H, s), 11.17 (1H, s), 7.65 (1H, s), 4.02 (2H, s); **^{13}C NMR (DMSO-d^6, 75MHz) δ(ppm):** 163.8, 151.0, 141.9, 106.4, 46.3; **MS (ESI, MeOH):** *m/z* calcd for $C_5H_5N_5O_2$ [M]:

167.0443, found: 190.9 $[M+Na]^+$, 166.0 $[M-H]^-$, 333.1 $[2M-H]^-$; **Elemental composition:** calcd %C 35.93, %H 3.02, %N 41.90, found %C 35.53, %H 2.95, %N 41.51; **FT-IR (KBr) ν(cm^{-1}):** 3219 (s), 3079 (m), 3026 (m), 2825 (m), 2132 (m), 1757 (s), 1689 (s), 1667 (s).

Ethyl 2-(5-azidomethyluracil-1-yl)acetate (2.17). In a round bottom flask **2.16** (1.02g, 6.07mmol) and K_2CO_3 (0.84g, 6.07) were dispersed in 10ml dry DMF and cooled to 0°C with an ice bath. Ethyl 2-bromoacetate (0.674ml, 6.07mmol) was diluted with 1ml dry DMF and added dropwise over 1h. The reaction was then left to warm to r.t. and left to stir overnight. The solvent was evaporated under vacuum and the resulting oil was partitioned between AcOEt (40ml) and water (40ml) and transferred in a separatory funnel, the aqueous phase was then extracted with AcOEt (2x40ml). The combined organic fractions were dried together with Na_2SO_4 and concentrated under vacuum. Flash chromatography (EtOAc:hexane 8:2) yielded **2.17** as white solid (0.63g, 41%). **TLC (AcOEt) Rf:** 0.46, **MP (°C):** 128.5-129.4; **^{1}H NMR ($CDCl_3$, 300MHz) δ(ppm):** 8.83 (1H, s), 7.17 (1H, s), 4.48 (2H, s), 4.26 (2H, q, J= 7.1Hz), 4.18 (2H, s), 1.31 (3H, t, J= 7.1Hz); **^{13}C NMR ($CDCl_3$, 75MHz) δ(ppm):** 167.1, 162.4, 150.2, 142.4, 110.0, 62.4, 48.8, 46.9, 14.0; **MS (ESI, MeOH):** *m/z* calcd for $C_9H_{11}N_5O_4$ [M]: 253.0811, found: 276.2 $[M+Na]^+$, 292.1 $[M+K]^+$, 252.2 $[M-H]^-$; **Elemental composition:** calcd %C 42.69, %H 4.38, %N 27.66, found %C 42.82, %H 4.28, %N 27.37; **FT-IR (KBr) ν(cm^{-1}):** 3160 (m), 3036 (s), 2841 (m), 2109 (s), 2082 (s), 1739 (s), 1701(s), 1647 (s).

Ethyl 2-(5-aminomethyluracil-1-yl)acetate trifluoroacetate salt (2.18). In a round bottom flask **2.17** (1.10g, 4.33mmol) was solubilized in 20ml THF, then water (0.156ml, 8.67mmol) was added and the reaction mixture was cooled to 0°C with an ice bath before the addition of triphenilphosphine (1.70g, 6.50mmol). The reaction was left to warm to r.t. and to react overnight. The solvent was then evaporated under reduced pressure, the crude was taken up with the minimum amount of TFA then precipitated with Et_2O, collected by centrifugation and washed with DCM (3 times) to yield **2.18** as white solid (0.70g, 71%). **TLC (AcOEt/MeOH/$NH_{3(aq)}$ 5:4:1) Rf:** 0.32, **MP (°C):** 113.8-115.2; **^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 11.84 (1H, s), 7.99 (3H, br s), 7.88 (1H, s), 4.56 (2H, s), 4.16 (2H, q, J= 7.1Hz), 3.65 (2H, d, J= 5.0Hz), 1.21 (3H, t, J= 7.1Hz); **^{13}C NMR (DMSO-d^6, 75MHz) δ(ppm):** 167.7, 163.2, 158.2 (q, J= 33Hz), 150.4, 146.1, 116.3 (q, J= 294Hz), 105.7, 61.1, 48.8, 35.1, 13.8; **MS (ESI, MeOH):** *m/z* calcd for $C_9H_{13}N_3O_4$ [M]: 227.0906, found: 228.2 $[M+H]^+$, 455.4 $[2M+H]^+$; **Elemental composition:** calcd %C 38.72, %H 4.14, %N 12.31, found %C36.85, %H 4.40, %N 11.28; **FT-IR (KBr) ν(cm^{-1}):** 3566 (m), 3446 (m), 3200 (m), 2998 (m), 2829 (m), 1728 (s), 1713 (s), 1667(s).

Ethyl 2-(5-(2-(pyren-1-yl)acetamido)methyluracil-1-yl)acetate (2.19). In a round bottom flask 1-pyreneacetic acid (305mg, 1.17mmol) and HBTU (434mg, 1.15mmol) were dissolved in 5ml DMF, the solution was then cooled to 0°C with an ice bath and DIPEA was added(485µl, 2.94mmol). The mixture was left to react for 30 minutes at 0°C, then to warm to r.t.. After 30 minutes **2.18** (200mg, 0.59mmol) was added and the mixture was left to react overnight. **2.19** was then collected as a beige solid over büchner after precipitation from the reaction mixture with 20ml AcOEt and 20ml H_2O. A second aliquot of product was

collected by precipitation with 20ml hexane from the organic layer washed with saturated $KHSO_4$ (2x50ml), saturated $NaHCO_3$ (2x50ml) and brine (50ml) (183mg, 66%); **TLC (AcOEt/MeOH/$NH_{3\,(aq)}$ 5:4:1) Rf:** 0.33, **MP (°C):** decompose without melting at 226°C; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 11.57 (1H, s), 8.53 (1H, t, J= 5.5Hz), 8.41 (1H, d, J= 9.3Hz), 8.29 (2H, d, J= 7.5Hz), 8.24 (1H, d, J= 7.8Hz), 8.20 (1H, d, J= 9.3Hz), 8.15 (2H, s), 8.07 (1H, d, J= 7.6Hz), 8.02 (1H, d, J= 7.9Hz), 7.53 (1H, s), 4.46 (2H, s), 4.24 (2H, s), 4.10 (2H, q, J= 7.1Hz), 3.92 (2H, d, J= 5.4Hz), 1.17 (3H, t, J= 7.1Hz); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 170.7, 168.4, 163.9, 151.2, 143.5, 131.3, 131.4, 130.8, 130.2, 129.5, 129.1, 127.8, 127.7, 127.3, 126.6, 125.6, 125.4, 125.2, 124.6, 124.4, 110.7, 61.6, 49.2, 40.4, 35.7, 14.4; **MS (ESI, MeOH):** *m/z* calcd for $C_{27}H_{23}N_3O_5$ [M]: 469.16377, found: 470.4 $[M+H]^+$, 492.3 $[M+Na]^+$, 508.3 $[M+K]^+$, 468.4 $[M-H]^-$, 504.4 $[M+Cl]^-$; **HRMS (LTQ-Orbitrap, MeOH):** *m/z* found: 470.17075 $[C_{27}H_{24}O_5N_3]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3041 (m), 1670 (s), 1684 (s), 1472 (w).

2-(5-(2-(pyren-1-yl)acetamido)methyluracil-1-yl)acetic acid (2.20). In a round bottom flask **2.19** (248mg, 0.53mmol) was dissolved in 10ml MeOH, 5ml NaOH 1M were added and the mixture and left to react overnight. The organic solvent was then evaporated under vacuum, the pH was lowered to 3 with HCl 37% and **2.20** was collect over buckner as a red solid (212mg, 91%). **TLC (AcOEt) Rf:** 0.00, **MP (°C):** decompose without melting at 165°C; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 13.15 (1H, br s), 11.52 (1H, s), 8.51 (1H, br t, J= 5.4Hz), 8.40 (1H, d, J= 9.3Hz), 8.29 (2H, d, J= 7.6Hz), 8.24 (1H, d, J= 7.9Hz), 8.20 (1H, d, J= 9.3Hz), 8.15 (2H, s), 8.08 (1H, d, J= 7.6Hz), 8.02 (1H, d, J= 7.9Hz), 7.56 (1H, s), 4.39 (2H, s), 4.25 (2H, s), 3.92 (2H, d, J= 5.2Hz); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 170.7, 169.9, 164.0, 151.2, 143.7, 131.4, 131.3, 130.8, 130.2, 129.5, 129.1, 127.8, 127.7, 127.3, 126.6, 125.5, 125.4, 125.2, 124.6, 124.4, 110.5, 49.3, 49.2, 35.8 ; **MS (ESI, MeOH):** *m/z* calcd for $C_{25}H_{19}N_3O_5$ [M]: 441.13247, found: 442.2 $[M+H]^+$, 464.3 $[M+Na]^+$, 480.2 $[M+K]^+$, 905.6 $[2M+Na]^+$, 440.3 $[M-H]^-$, 881.6 $[2M-H]^-$, 903.6 $[2M+Na-2H]^-$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 440.12518 $[C_{25}H_{18}N_3O_5]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3041 (m), 1700 (s), 1684 (s), 1472 (w).

2-(5-azidomethyluracil-1-yl)acetic acid (2.21). In a round bottom flask **2.17** (1.15g, 4.54mmol) was dispersed in 10ml MeOH, 10ml NaOH 1M were added and the mixture and left to react. After 1h the solution became clear and the organic solvent was then evaporated under vacuum, the pH was lowered to 3 with HCl 37% and **2.21** was collect over buckner as a white solid (0.98g, 96%). **TLC (AcOEt/MeOH) Rf:** 0.31, **MP (°C):** decompose without melting at 175°C; **^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 13.21 (1H, br s), 11.65 (1H, s), 7.83 (1H, s), 4.43 (2H, s), 4.06 (2H, s); **^{13}C NMR (DMSO-d^6, 75MHz) δ(ppm):** 169.3, 163.3, 150.6, 145.3, 107.3, 48.6, 46.5; **MS (ESI, MeOH):** *m/z* calcd for $C_7H_7N_5O_4$ [M]: 225.04980, found: 226.4 $[M+H]^+$, 248.3 $[M+Na]^+$, 264.2 $[M+K]^+$, 473.2 $[2M+Na]^+$, 224.3 $[M-H]^-$, 449.3 $[2M-H]^-$; **Elemental composition:** calcd %C 37.34, %H 3.13, %N 31.10, found %C 37.45, %H 3.32, %N 30.24; **FT-IR (KBr) ν(cm^{-1}):** 3412(w), 3002(br), 2825(m), 2108(s), 1751(s), 1700(s), 1691(s), 1482(s).

Tert-buthyl 2-(N-(2-Fmoc-aminoethyl)-2-(5-(2-(pyren-1-yl)acetamido)methyluracil-1-yl)acetamido)acetate (2.22). In a round bottom flask **2.20** (200mg, 0.452mmol) was solubilized in 2ml dry DMF at 0°C together with EDC·HCl (104mg, 0.543mmol), DhBtOH (74mg, 0.453mmol) and DIPEA (179µl, 1.086mmol) and left to react for 10 minutes before warming to r.t.. After 20 minutes tert-butyl 2-((2-Fmoc-aminoethyl)amino)acetate hydrochloride (104mg, 0.543mmol) was added and the reaction mixture was stirred for further 4h. The reaction was then diluted with 200ml AcOEt and washed with saturated $KHSO_4$ (2x200ml), saturated $NaHCO_3$ (2x200ml) and brine (200ml). The organic fraction was dried over Na_2SO_4 and concentrated under vacuum. Flash chromatography (from AcOEt to AcOEt/MeOH 9:1) yielded **2.22** as pale brown solid (252mg, 68%). **TLC (AcOEt/MeOH 7:3) Rf:** 0.71; **MP (°C):** d; **^{1}H NMR ($CDCl_3$, 400MHz, major rotamer) δ(ppm):** 9.67(1H, s), 8.10÷8.03 (3H, m), 8.01÷7.95 (2H, m), 7.93÷7.85 (3H, m), 7.83÷7.75 (1H, m), 7.72 (2H, d, J= 7.3Hz), 7.59 (2H, d, J= 7.3Hz), 7.39÷7.32 (2H, m), 7.31÷7.21 (2H, m), 7.01 (1H, s), 6.67 (1H, t, J= 5.8Hz), 6.04 (1H, t, J= 5.7Hz), 4.42 (2H, d, J= 6.7Hz), 4.25÷4.05 (5H, m), 3.90÷3.81 (4H, m), 3.44 (2H, br s), 3.26 (2H, br s), 1.44 (9H, s); **^{13}C NMR ($CDCl_3$, 100MHz, major rotamer) δ(ppm):** 28.0, 36.3, 39.3, 41.5, 47.9, 48.9, 49.8, 66.8, 120.0, 123.1, 124.5, 125.0, 125.2, 126.0, 127.1, 127.2, 127.3, 127.7, 128.0, 128.5, 128.8, 129.5, 130.7, 131.1, 143.9, 150.5, 163.9, 167.7, 168.5, 171.2, 171.7; **MS (ESI, MeOH):** *m/z* calcd for $C_{48}H_{45}N_5O_8$ [M]: 819.32681, found: 842.5 $[M+Na]^+$, 858.5 $[M+K]^+$, 1663.0 $[2M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 820.33386 $[C_{48}H_{46}O_8N_5]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3414 (m), 3043 (w), 1676 (s), 1522 (m).

Tert-buthyl 2-(N-(2-Fmoc-aminoethyl)-2-(5-azidomethyluracil-1-yl)acetamido)acetate (2.23). In a round bottom flask **2.21** (251mg, 1.12mmol) was solubilized in 3ml dry DMF at 0°C together with EDC·HCl (213mg, 1.12mmol), DhBtOH (181mg, 1.12mmol) and DIPEA (290µl, 1.68mmol) and left to react for 5 minutes before the addition of tert-butyl 2-((2-Fmoc-aminoethyl)amino)acetate hydrochloride (240mg, 0.56mmol) was added and the reaction mixture was stirred for 15 minutes at 0°C and further 2.5h at room temperature. The reaction was then diluted with 200ml AcOEt and washed with saturated $KHSO_4$ (2x200ml), saturated $NaHCO_3$ (2x200ml) and brine (200ml). The organic fraction was dried over Na_2SO_4 and concentrated under vacuum. Flash chromatography (from AcOEt/hexane 1:1 to AcOEt/hexane 3:1) yielded **2.23** as white solid (323mg, 96%). **TLC (AcOEt) Rf:** 0.52; **MP (°C):** 83.1÷86.0°C; **^{1}H NMR ($CDCl_3$, 400MHz, major rotamer) δ(ppm):** 9.65 (1H, s), 7.77 (2H, d, J= 7.5Hz), 7.63 (2H, d, J= 7.4Hz), 7.41 (2H, t, J= 7.4Hz), 7.32 (2H, t, J= 7.4Hz), 7.11, 6.08 (1H, br t), 4.50 (2H, s), 4.46 (2H, t, J= 7.6Hz), 4.23 (1H, t, J= 5.9Hz, 3H), 4.09 (2H, s), 4.07 (2H, s), 3.63÷3.48 (2H, m), 3.45÷3.33 (2H, m), 1.48 (9H, s); **^{13}C NMR ($CDCl_3$, 100MHz, major rotamer) δ(ppm):** 168.7, 167.0, 163.2, 156.8, 150.8, 144.0, 143.8, 141.3, 127.8, 127.1, 125.1, 120.0, 109.5, 82.8, 66.8, 51.1, 49.8, 48.8, 48.0, 47.0, 39.2, 28.0; **MS (ESI, MeOH):** *m/z* calcd for $C_{30}H_{33}N_7O_7$ [M]: 603.24415, found: 626.3 $[M+Na]^+$, 642.5 $[M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 604.2510 $[C_{30}H_{33}O_7N_7]^+$.

2-(N-(2-Fmoc-aminoethyl)-2-(5-(2-(pyren-1-yl)acetamido)methyluracil-1-yl)acetamido)acetic acid (2.24). In a round bottom flask **2.22** (241mg, 0.2941mmol) was solubilized in 6ml DCM at 0°C, then TFA

(4ml) was added and the reaction mixture was left to react. After 5 minutes the reaction mixture was left to warm to r.t. and to reach for other 2h. The solvent was then co-evaporated with MeOH and $CHCl_3$ under reduced pressure. The resulting oil was dispersed in 20ml H_2O and filtered through Buchner to yield **2.24** as pale brown solid (193mg, 86%). **TLC (AcOEt/MeOH 1:1) Rf:** 0.49; **MP (°C):** decompose without melting at 166°C; **^{1}H NMR (DMSO-d^6, 400MHz, major rotamer) δ(ppm):** 12.82 (1H, br s), 11.48 (1H, s), 8.52 (1H, br s), 8.41 (1H, s), 8.30÷8.25 (2H, m), 8.23 (2H, d, J= 9.0Hz), 8.14 (1H, s), 8.09÷7.98 (2H, m), 7.88 (2H, d, J= 7.5Hz), 7.67 (2H, d, J= 7.1Hz), 7.40 (6H, t, J= 7.5Hz), 4.66 (2H, br s), 4.30÷4.37 (3H, m), 4.29÷4.18 (4H, m), 3.92 (2H, br s), 3.44÷3.32 (2H, m), 3.29÷3.10 (2H, m); **^{13}C NMR (DMSO-d^6, 100MHz, major rotamer) δ(ppm):** 47.2, 47.3, 48.2, 48.6, 49.5, 65.4, 66.0, 110.4, 120.6, 124.4, 124.6, 125.2, 125.4, 125.5, 125.6, 126.6, 127.3, 127.5, 127.7, 127.9, 128.1, 129.2, 129.5, 130.2, 130.8, 131.3, 131.4, 141.2, 144.3, 151.2, 151.8, 164.4, 167.9, 171.3; **MS (ESI, MeOH):** *m/z* calcd for $C_{44}H_{37}N_5O_8$ [M]: 763.26421, found: 786.5 $[M+Na]^+$, 802.3 $[M+K]^+$, 762.5 $[M-H]^-$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 762.25629 $[C_{44}H_{36}N_5O_8]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3409 (m), 1684 (s), 1559 (m), 1472 (m).

2-(N-(2-Fmoc-aminoethyl)-2-(5-azidomethyluracil-1-yl)acetamido)acetic acid (2.25). In a round bottom flask **2.23** (311mg, 0.51mmol) was solubilized in 6ml DCM at 0°C, then TFA (4ml) was added and the reaction mixture was left to react. After 30 minutes the reaction mixture was left to warm to r.t. and to reach for other 1h. The solvent was then co-evaporated with MeOH and $CHCl_3$ under reduced pressure. The resulting oil was dispersed in 20ml H_2O and filtered through Buchner to yield **2.25** as white solid (241mg, 86%). **TLC (AcOEt/MeOH 1:1) Rf:** 0.41; **MP (°C):** decompose without melting at 110°C; **^{1}H NMR (DMSO-d^6, 400MHz, major rotamer) δ(ppm):** 12.81 (1H, br s), 11.60 (1H, s), 7.90 (2H, d, J= 7.5Hz), 7.77÷7.62 (3H, m), 7.42 (2H, t, J= 7.3Hz), 7.33 (2H, t, J= 7.4Hz), 4.73 (2H, s), 4.32 (2H, dd, J= 18.3, 6.8Hz), 4.23 (1H, t, J= 7.0Hz), 4.05 (2H, s), 4.00 (2H, s), 3.48 – 3.17 (4H, m); **^{13}C NMR (DMSO-d^6, 100MHz, major rotamer) δ(ppm):** 170.8, 167.5, 163.9, 156.8, 151.2, 146.3, 144.3, 141.2, 128.1, 127.5, 125.6, 120.6, 107.7, 66.0, 49.7, 48.6, 48.4, 48.2, 47.2, 38.4; **MS (ESI, MeOH):** *m/z* calcd for $C_{26}H_{25}N_7O_7$ [M]: 547.18155, found: 570.5 $[M+Na]^+$, 586.5 $[M+K]^+$, 546.4 $[M-H]^-$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 546.17335 $[C_{26}H_{24}N_7O_7]^-$.

Synthesis and characterization of PNAs. The synthesis of **PNA 2.1** was performed on an ABI 433A peptide synthesizer with a software modified to run the PNA synthetic steps, using Fmoc-based chemistry and standard protocols with HBTU/DIPEA coupling and a Rink amide loaded with Fmoc-PNA-C(Bhoc)-OH as first monomer (0.2mmol/g). **PNA 2.2** was synthesized using standard Boc-based manual peptide solid-phase synthesis for the introduction of achiral monomers and with submonomeric strategy using **2.5** on a MBHA resin loaded with Boc-PNA-C(Cbz)-OH as first monomer (0.2mmol/g). **PNA 2.3** was synthesized with standard manual Boc-based solid-phase synthesis using the preformed monomers **2.11** and **2.12**, in addition to commercially available Boc-PNA monomers on a MBHA resin loaded with Boc-PNA-C(Cbz)-OH as first monomer (0.2mmol/g).The synthesis of **PNA 2.4** was performed with standard Fmoc-based manual synthesis protocol using **2.25** in addition to standard monomers, on a Rink amide resin loaded with

Fmoc-Gly-OH as first monomer (0.2mmol/g). The pyrene unit was introduced by on-resin reduction of the azide function with a mixture H_2O/THF/$P(CH_3)_3$ 1M in THF 3:2:1 (10 minutes, 2 times) and subsequent coupling of 1-pyreneacetic acid (10 equivalents) with DIC/DhBtOH as activating agent. The syntheses of the PNAs bearing multiple pyrene units (**PNA 2.5**, **PNA 2.6**, **PNA 2.7**, **PNA 2.8** and **PNA 2.9**) were performed with standard Fmoc-based manual synthesis protocol using **2.24** in addition to standard monomers, on a Rink amide resin loaded with Fmoc-Gly-OH as first monomer (0.2mmol/g). PNAs purification were performed by RP-HPLC with UV detection at 260nm (gradient B). The purity and identity of the purified PNAs were determined by HPLC-UV-MS (gradient A). **PNA 2.1:** 12%; t_r: 15.9min; ESI-MS: *m/z* calcd 3775.5266 [M]: 945.4 $[MH_4]^{4+}$, 756.2 $[MH_5]^{5+}$, 630.4 $[MH_6]^{6+}$, 540.7 $[MH_7]^{7+}$; **PNA 2.2:** 13%; t_r: 14.3min; ESI-MS: *m/z* calcd 4072.7655 [M]: 1019.5 $[MH_4]^{4+}$, 816.0 $[MH_5]^{5+}$, 680.1 $[MH_6]^{6+}$, 583.1 $[MH_7]^{7+}$, 510.3 $[MH_8]^{8+}$; **PNA 2.3:** 16%; t_r: 14.2min; ESI-MS: *m/z* calcd 4072.7655 [M]: 1019.5 $[MH_4]^{4+}$, 815.7 $[MH_5]^{5+}$, 680.0 $[MH_6]^{6+}$, 583.0 $[MH_7]^{7+}$, 510.4 $[MH_8]^{8+}$; **PNA 2.4:** 10%; t_r: 18.4min; ESI-MS: *m/z* calcd 2675.0585 [M]: 1338.6 $[MH_2]^{2+}$, 892.7 $[MH_3]^{3+}$, 669.5 $[MH_4]^{4+}$, 535.8 $[MH_5]^{5+}$; **PNA 2.5:** 9%; t_r: 24.7min; ESI-MS: *m/z* calcd 2932.1426 [M]: 1466.9 $[MH_2]^{2+}$, 978.2 $[MH_3]^{3+}$, 733.9 $[MH_4]^{4+}$, 587.2 $[MH_5]^{5+}$; **PNA 2.6:** 10%; t_r: 24.6min; ESI-MS: *m/z* calcd 2932.1426 [M]: 1466.9 $[MH_2]^{2+}$, 978.2 $[MH_3]^{3+}$, 733.9 $[MH_4]^{4+}$, 587.3 $[MH_5]^{5+}$; **PNA 2.7:** 11%; t_r: 23.7min; ESI-MS: *m/z* calcd 2932.1426 [M]: 1466.8 $[MH_2]^{2+}$, 978.2 $[MH_3]^{3+}$, 733.9 $[MH_4]^{4+}$, 587.3 $[MH_5]^{5+}$; **PNA 2.8:** 15%; t_r: 24.0min; ESI-MS: *m/z* calcd 2932.1426 [M]: 1467.1 $[MH_2]^{2+}$, 978.2 $[MH_3]^{3+}$, 733.8 $[MH_4]^{4+}$, 587.3 $[MH_5]^{5+}$; **PNA 2.9:** 11%; t_r: 27.4min; ESI-MS: *m/z* calcd 3189.2266 [M]: 1064.1 $[MH_3]^{3+}$, 798.2 $[MH_4]^{4+}$, 638.7 $[MH_5]^{5+}$. Yields % reported for each PNA are those of purified products, calculated by UV-Vis analysis.

Estimation of the extinction coefficient for 2.20. Stock solution of **2.20** was prepared in H_2O (1.17mg in 2.0ml, 1.32mM). From this solution an intermediate 1:10 dilution was prepared, which was used to prepare the final solution with a 1:100, 1.5:100, 2:100, 3:100, 5:100, 7.5:100 dilution. Each solution was prepared in triplicate and the absorbance of the solutions was evaluated within the range 200÷500nm at a scan speed of 100nm/min.

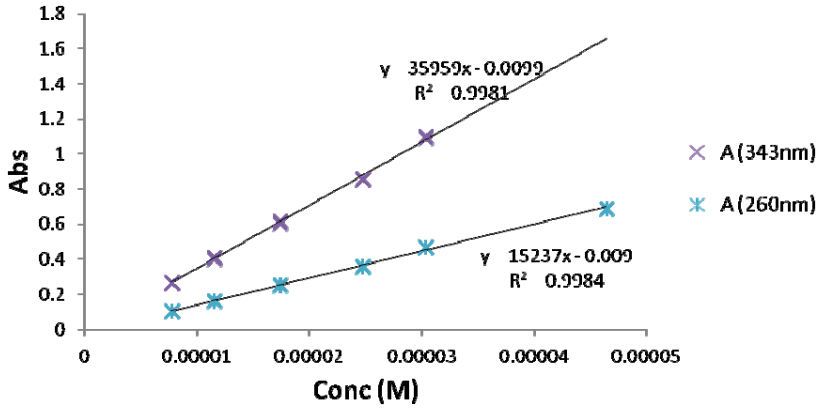


Figure 2.17: linear regression for the evaluation of the extinction coefficient at 260nm and 343nm.

UV and CD measurements. Stock solutions of PNA and DNA synthetic oligonucleotides (Thermo-Fisher Scientific, HPLC-grade) were prepared in double-distilled water, and the PNA concentration was calculated by UV absorbance with the following extinction coefficients (ε_{260} [$M^{-1}cm^{-1}$]) for the nucleobases: T 8600, T* 15237, C 6600, A 13700, and G 11700. For DNA the data provided by the producer were used. From these, solutions containing single stranded PNA and DNA, PNA–DNA duplexes were prepared. Measurement conditions: [PNA]=[DNA]= 5μM (for backbone-modified PNAs) or 1μM (for base modified PNAs) in PBS (100mM NaCl, 10mM $NaH_2PO_4 \cdot H_2O$, 0.1mM EDTA, pH 7.0). All the samples were first incubated at 90°C for 5 min, then slowly cooled to room temperature. Thermal denaturation profiles (A_{260} versus T) of the hybrids were measured with a UV/Vis Lambda Bio 20 spectrophotometer equipped with a Peltier temperature programmer PTP6 interfaced to a personal computer. For the temperature range 18÷90°C, A_{260} values were recorded at 0.1°C increments, with a temperature ramp of 1°C/min. A melting curve was recorded for each duplex. The melting temperature (Tm) was determined from the maximum of the first derivative of the melting curves. CD measurements were carried out on a Jasco J715 spectropolarimeter equipped with a Peltier PTC 348 temperature controller unit, using solutions prepared as described above for the UV measurements. A scan speed of 50 nm/min was used and 4 accumulations for each spectrum. The spectra were corrected by subtracting the signal obtained with the corresponding buffer. CD meting curves were obtained in a similar way to UV.

Array preparation. ‘‘Genorama1 SAL Slides’’ (Asper Biotech) were used as solid support to which the amino-terminal group of the PNA probes was covalently linked. The deposition of the probes was carried out using a Perkin Elmer SpotArray 24 Microarray Printing System. The manufacturer’s instructions for the deposition protocol were slightly changed to comply with the special requirement of the chemical structures of PNAs: in particular a 100mM carbonate buffer (pH 9.0) containing 0.001% sodium dodecyl sulfate (SDS) was used as deposition buffer. Moreover, after every deposition, the pin printing system was washed in three solutions containing 100mM carbonate buffer (pH 9.0) and SDS at different concentrations (0.2%, 1%, 0.001%), to avoid carry-over of the probes in subsequent depositions. The probes were coupled to the surface by leaving the slides in a humid chamber (relative humidity: 75%) at room temperature for 12 h, and the remaining reactive sites were blocked by immersion in a 1% aqueous solution of NH_3 at room temperature for 15 min under gently shaking. The slides were slowly shaken for 15 min with 0.1% SDS buffer pre-warmed at 40°C and then washed twice with doubly distilled water at room temperature. Each slide was spin-dried twice in a falcon tube at 1200rpm for 5min, to remove the remaining washing solution. Slides were then ready to undergo the hybridization protocol or could be stored under dried atmosphere for future use. Since a fluorescent control probe ($(AC)_{11}$-Cy5) was deposited to check the efficiency of the deposition step, all the previously described operations were carried out away from direct light in order to prevent degradation of the Cy5 fluorophore.

Hybridization on microarray. The Cy5-labeled DNA-FM and DNA-MM oligonucleotides were used at a final concentration of 0.1μM in 2x saline/sodium citrate (SSC) solution and 0.1% SDS buffer. Hybridization

was performed by loading 65µl of each solution to multiwall chambers (Arrayit Corporation, AHC1 x 16), using a silicone gasket (Arrayit Corporation, GAHC1 x 16), to avoid cross contamination, and leaving the slides under slow shaking for 2.5h at 40°C. After the hybridization step, the slides were washed under slow shaking for 5 min at 40°C with a 2x SSC, 0.1% SDS buffer pre-warmed at 40°C, followed by treatment for 1min with 0.2x SSC and for 1min with 0.1x SSC at room temperature. The slides were then spin-dried twice at 1200rpm for 5min. All post-hybridization steps were performed in a dark environment to prevent degradation of the Cy5 fluorophore used to label the oligonucleotides.

Array image acquisition. The fluorescent signal deriving from the hybridization was acquired using a Perkin Elmer ScanArray Express Microarray Scanner at λ_{ex}= 646nm and λ_{em}= 664nm. To correctly compare the hybridization data, all the images reported were acquired with laser power= 100 and photomultiplier gain= 70. Images were analyzed using the ScanArray program with integration of 170 µm diameter circular area entirely containing the fluorescent spots.

Evaluation of the Tm with fluorescence. From the stock solution described above were prepared both single stranded and double stranded solution at 1µM concentration in PBS with 1.25%DMF to avoid precipitation. All the samples were first incubated at 90°C for 5 min, then slowly cooled to 10°C. Fluorescence spectra were recorded every 5°C allowing the sample to equilibrate 5min to the new temperature before the new record. Fluorescence spectra were recorded on a Perkin Elmer LS55 Luminescence Spectrometer equipped with a LAUDA ECOline RE104 temperature control system, exciting at 347nm (slit: 5.0nm), scanning from 370nm to 550nm, a scan speed of 200nm/min was used and 3 accumulation for each spectrum. All measurements were normalized using a reference standard solution containing 20nM 1-pyreneacetic acid in PBS.

Fluorescence titration of PNA 2.5 and PNA 2.6. From the stock solution described above were prepared single stranded PNA solutions (1µM in PBS) and single stranded DNA solutions (10µM in PBS). PNA solution were first incubated at the experimental temperature then spectra were recorded after addition of portions of DNA (10% of the PNA amount each), allowing an equilibration time of 8min. Fluorescence emission spectra were recorded with an excitation wavelength of 347nm (slit: 5.0nm), scanning from 370nm to 550nm, scan speed of 200nm/min, with 3 spectra accumulation for each solution. All measurements were normalized using a reference standard solution containing 20nM 1-pyreneacetic acid in PBS. **PNA 2.5**: T= 25°C, emission slit= 6nm; **PNA 2.6**: T= 10°C, emission slit= 10nm.

Fluorescence analysis at variable temperature. From the stock solution described above single stranded PNA solutions (1µM in PBS) and triple stranded PNA:DNA:PNA solutions (1µM PNA and 0.5µM DNA in PBS) were prepared. All the samples were first incubated at 90°C for 5 min, then slowly cooled to temperature of analysis. Fluorescence emission spectra were recorded with an excitation wavelength of 347nm (slit excitation: 5.0nm), scanning from 370nm to 550nm (slit emission: 10.0nm), a scan speed of

200nm/min was used with 3 spectra accumulation for each solution. All measurements were normalized using a reference standard solution containing 20nM 1-pyreneacetic acid in PBS.

2.5 - References

[1] Brookes A. J. *Gene*, **1999**, 234, 177.

[2] a) Ozaki, K., Ohnishi, Y., Iida, A., Sekine A., Yamada R., Tsunoda, T., Sato H., Sato H., Hori M., Nakamura Y., Tanaka T. *NATURE GENETICS* **2002** , 32 (4), 650-654 ; b) Kontola K., Aalto-Sălă K., Kuusi T., Hămălăinen L., Syvănen A. *Clin. Chem.* **1990**, 36 (12), 2087-2097; c) Ng P. C., Henikoff S. *GENOME RESEARCH* **2002**, 12 (13), 436-446.

[3] a) Blais B. W., Phillippe L. M., Vary N. *Biotechnol. Lett.* **2002**, 24, 1407-1411; b) Rudi K., Rud I., Holck A. *Nucleic Acids Res* **2003**, 31, e62/1-e62/8, c) Germini A., Mezzelani A., Lesignoli F., Corradini R., Marchelli R., Bordoni R., Consolandi C., De Bellis G. *J. Agric. Food. Chem.* **2004**, 52, 4535-4540, d) Germini A., Rossi S., Zanetti A., Corradini R., Fogher C., Marchelli R. *J. Agric. Food. Chem.* **2005**, 53, 3958-3962, e) Totsingan F., Tedeschi T., Sforza S., Corradini R., Marchelli R. *Chirality* **2009**, 21, 245-253.

[4] Kian-Kon J., Wen-Tso L. *Anal. Bioanal. Chem*. **2006**, 386, 427-434.

[5] Sassolas A., Leca-Bouvier B. D., Blum L. J. *Chem. Rev.* **2008**, 108, 109-139.

[6] Southern E. M. *Electrophoresis* **1995**, 16, 1539-1542.

[7] Weiler J., Gausepohl H., Hauser N., Jensen O. N., Hoheisel J. D. *Nucleic Acids Res.* **1997**, 25, 2792-2799.

[8] Okamoto A., Ichiba T., Saito I. *J. Am. Chem. Soc.* **2004**, 126 (27), 8364-8365.

[9] Hwang G. T., Seo Y. J., Kim B. H. *J. Am. Chem. Soc.* **2004**, 126 (21), 6528-6529.

[10] Pasternak A., Kierzek E., Pasternak K., Fratczak A., Turner D. H., Kierek R. *Biochemistry* **2008**, 47, 1249-1258.

[11] Okamoto A., Ochi Y., Saito I. *Chem. Commun.* **2005**, 9, 1128-1130.

[12] Seo Y. J., Kim B. H. *Chem. Commun.* **2006**, 2, 150-152.

[13] Conlon P., Yang C. Y., Wu Y., Chen Y., Martinez K., Kim Y., Stevens N., Marti A. A., Jockusch S., Turro N. J., Tan W. *J. Am. Chem. Soc.* **2008**, 130, 336-342.

[14] Yamana K., Ohshita Y., Fukunaga Y., Nakamuna M., Maruyama A. *Bioorg. Med. Chem.* **2008**, 16, 78-83.

[15] a) Marti A. A., Li X., Jockusch S., Li Z., Raveendra B., Kalachikov S., Russo J. J., Morozova I., Puthanveettil S. V. Ju J., Turro N. J. *Nucleic Acid Research* **2006**, 34 (10), 3161-3168; b) Bichenkova E. V., Gbaj A., Walsh L., Savage H. E., Rogert C., Sardarian A. R., Etchells L. L. Douglas K. T. *Org. Biomol. Chem.* **2007**, 5, 1039-1051; c) Umemoto T., Hrdlicka P. J., Babu B. R., Wengel J. *ChemBioChem* **2007**, 8, 2240-2248.

[16] a) Seo Y. J., Hwang G. T., Kim B. H. *Nucleic Acids Symposium Series No. 49*, 135-136; b) Wanninger-Weiß C., Valis L., Wagenknecht H.-A. *Bioorg. Med. Chem.* **2008**, 16, 100-106; c) Østergaard M. E., Guenther D. C., Kumar P., Baral B., Deobald L., Paszczynski A. J., Sharma P. K., Hrdlicka P. J. *Chem. Commun.* **2010**, 46, 4929-4931; d) Tanaka K., Okamoto A. *Bioprg. Med. Chem.* **2008**, 16, 400-404

[17] Hrdlicka P. J., Kumar T. S., Wengel J. *Chem. Commun.* **2005**, 34, 4279-4281.

[18] Østergaard M. E., Hrdlicka P. J. *Chem. Soc. Rev.* **2011**, 40, 5771-5788.

[19] Kalyanasundaram K., Thomas J. K. *J. Phys. Chem.* **1977**, 81, 2176-2180.

[20] Tedeschi T., Sforza S., Dossena A., Corradini R., Marchelli R. *Chirality* **2005**, 17, S196-S204.

[21] Sforza S., Tedeschi T., Corradini R., Marchelli R. *Eur J Org Chem* **2007**, 35, 5879-5885

[22] a) Tedeschi T., Chiari M., Galaverna G., Sforza S., Cretich M., Corradini R., Marchelli R. *Electrophoresis* **2005**, 26, 4310-4316; b) Corradini R., Feriotto G., Sforza S., Marchelli R., Gambari R. *J. Mol. Rec.* **2004**, 17, 76–84; c) Tedeschi T., Corradini R., Marchelli R., Pushl A., Nielsen P.E. *Tetrahedron Asymm* **2002**, 13, 1629-1636.

[23] Sforza S., Tedeschi T., Corradini R., Ciavardelli D., Dossena A., Marchelli R. *Eur J Org Chem* **2003**, 6, 1056-1063.

[24] Menchise V., De Simone G., Tedeschi T., Corradini R., Sforza S., Marchelli R., Capasso D., Saviano M., Pedone C. *Proc Natl Acad Sci USA* **2003**, 100, 12021-12026.

[25] Sforza S., Corradini R., Ghirardi S., Dossena A., Marchelli R. *Eur J Org Chem* **2000**, 17, 2905-2913.

[26] He G., Rapireddy S., Bahal R., Sahu B., Ly D. H. *J. Am. Chem. Soc.* **2009**, 131, 12088-12090.

[27] Rapireddy S., He G., Roy S., Armitage B. A., Ly D. H. *J. Am. Chem. Soc.* **2007**, 129, 15596-15600.

[28] Dragulescu-Andrasi A., Rapireddy S., Frezza B.M., Gayathri C., Gil R.R., Ly D.H. *J Am Chem Soc* **2006**, 128, 10258-10267.
[29] Calabretta A., Tedeschi T., Di Cola G., Corradini R., Sforza S., Marchelli R. *Mol Bio-Syst* **2009**, 5, 1323-1330.
[30] Corradini R., Tedeschi T., Sforza S., Totsingan F., Marchelli R. *Curr Top Med Chem* **2007**, 7, 681-694.
[31] Manicardi A., Calabretta A., Bencivenni M., Tedeschi T., Sforza S., Corradini R., Marchelli R.. *Chirality* **2010**, 22 (1E), 161-172.
[32] Hudson R. H. E., Liu X., Wojciechowski F. *Can. J. Chem.* **2007**, 85 (4), 302-312.
[33] Seo Y. J., Ryu J. H., Kim B. H. *Org. Lett.* **2005**, 7, 4931-4933.
[34] Calabretta A., Tedeschi T., Di Cola G., Corradini R., Sforza S., Marchelli R. *Mol. Bio-Syst.* **2009**, 5 (11), 1323-1330.

Chapter 3 - Modulation of miR activity with PNAs

3.1 - Introduction

Micro-RNA (miR) are small (19-23 bp) noncoding double-stranded RNA which modulate gene expression through a mechanism similar to RNA interference (RNAi): incorporation of one strand into the microRNA-induced Silencing Complex (miRISC) and suppression of mRNA expression by degradation or complexation of target sequence depending on complementarity level[1,2]. This miR/mRNA interaction leads to regulation of a great number of biological processes such as differentiation, cell cycle and apoptosis[3,4]. Down- or up-regulation of these molecules is often related to diseases. On one side, the correct miR level can be achieved by introducing short dsRNA to induce the RNAi mechanism, on the other side inhibition of miR activity by specific molecules able to bind to them by base pairing (antago-miR or anti-miR) has been shown to be of great interest in drug development, since they allow to restore the expression of miR-targeted genes[5]. Ideally, the anti-miR activity should be accompanied by good cellular uptake, high chemical and enzymatic stability, and very high affinity for the target RNA.

PNAs are ideal candidates for targeting miRs since they form a very stable PNA:RNA duplex which can efficiently disrupt the dsRNA duplex. However, very few works have been reported so far concerning the use of PNAs as anti-miR agents, all showing best performances among a series of oligonucleotide analogues[6]. The major limit in the use of PNAs for alteration of gene expression is their low cellular uptake[7]. In order to solve this problem several approaches have been proposed, by conjugating the PNA with carrier molecules or by chemical modification[8]. One of these strategies is to increase the cellular delivery by linking them to cell penetrating peptides (CPP), exploiting the similarity between peptide and PNA solid-phase synthesis. Thus, PNA conjugated with penetratin, transportan and other peptidic carriers have been extensively used in antisense research. Conjugation with short cationic peptides such as tetra-lysine or the SV40 nuclear localization signal (NLS) was shown to greatly improve the cellular permeability of PNAs[9], and favour dsDNA invasion[10]. Introduction of positively charged amino acid side chains has been shown to favour cellular uptake in several cellular lines, [11,12] and to enhance DNA affinity, also favouring dsDNA invasion, if used in combination with other modifications such as modified nucleobases[13].

A very efficient delivery system, developed by Wender and co-workers, is based on arginine oligomers, which have been demonstrated to be efficient carriers for cellular delivery[14] , more efficient than other cationic amino acid-containing peptides such as poly-Lys or poly-His derivatives[15]. Oligoarginines were subsequently reported to be internalized into mammalian cells predominantly through an endocytotic pathway, which is mediated through binding to cell surface heparan sulfates linked to proteoglycans[16], followed by back-transport in the Golgi complex and endoplasmic reticulum[17]. A modelling study on oligoarginine peptides suggested that not all the arginine side chains are able to interact with the cellular membrane, and a systematic study showed that longer spacing between the arginine side-chains could improve the uptake of these carrier peptides, with activity increasing by increasing the spacer length, and with optimal values obtained for the 6-aminocaproic acid[14]. In spite of these interesting properties as carriers, antisense oligoarginine-PNA conjugates were reported to have decreased activity when compared to those conjugated to other cationic peptides, which was attributed to their rapid degradation by peptidases[18].

It has been reported that a homothymine PNA decamer having all the C2-carbons substituted with guanidinium side chains derived from L-arginine (GPNA) is able to cross the cellular membrane with same efficiency as TAT transduction domain (amino acids 49-57), in both cases with specific localization on the endoplasmatic reticulum[12]. An uptake mechanism neither related to endocytotic nor receptor mediated phenomena was proposed, similarly to poly-arginine or arginine-rich peptides; in fact the cationic side chains of these peptides are able to form ionic complexes with the phospholipidic bilayer, which is able to translocate the target molecules inside the cytosol, on account of a difference between the electrostatic potential of the two faces of the cellular membrane[14]. A similar behavior was showed by GPNA based on L-Arg in the C-5 position[19].

Substitution at both C2 and C5 carbon atoms of the PNA backbone with amino acid side chains leads to ambivalent structures having properties of DNA or RNA mimic on one side and peptide mimic on the other side, thus allowing recognition by specific receptors, as shown very recently by a short PNA mimicking the function of a nuclear localization peptide (NLS)[20]. Thus, this strategy can be used to obtain PNAs with both peptide properties and RNA binding ability; furthermore, the use of chiral PNAs allows to achieve higher sequence selectivity for even single mismatch recognition[21].

However, a systematic study on the cellular uptake and biological properties of modified PNAs as a function of the backbone substitution and distribution of the side chains along the PNA chain has not yet been carried out; thus a general model on how substituents should be placed is not yet

available.

3.2 - Results and discussion

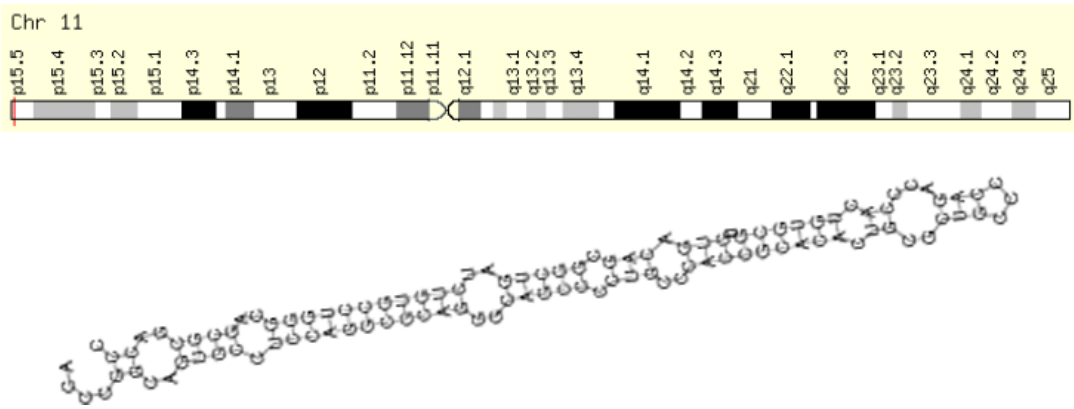


Figure 3.1: genomic locus of miR210 and predicted secondary structure of stem-loop. Estimation by: http://rna.tbi.univie.ac.at/.

miR-210 is a microRNA produced by transcription of a DNA tract on chromosome 11 (Figure 3.1), and is an important target which has been associated with hypoxia and is modulated during erythroid differentiation in leukemic cell lines[22,23].

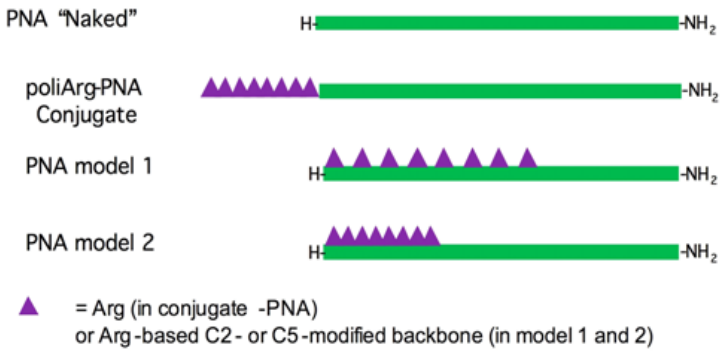


Figure 3.2: structural models of the PNAs used in this study.

In this chapter the activity of a series of PNAs bearing a series of arginine residues was tested as anti-miR-210 on K562 cells, with the aim of developing an optimized model for cellular uptake, biostability and anti-miR activity. Four different models were chosen to achieve this goal (Figure 3.2): an unmodified PNA ("naked"), used as a control system, a PNA bearing a CPP peptide (R_8 tail), and two different PNAs bearing modified arginine-based chiral monomers in which the side

chain are on consecutive monomers (Model 1) or spaced with an unmodified monomer (Model 2). We will present the comparison between these three different model systems, comparing their ability to bind DNA and RNA strands, cellular uptake and anti-miR biological effects.

3.2.1 - PNAs design and synthesis

For the design of a drug able to bind to a molecular target, such as miR inside the cells, it is important to have a tested model known to be able to selectively bind the target sequence without interaction with other RNAs present in the cytosol together with the desired target; PNAs are, therefore, good candidates for this application, since they have high sequence selectivity, though single-base mismatches are less discriminated by longer PNA. It is important also to take in to account that the target miR is present inside the cell in different forms and that it will interact with other RNA sequences in a more or less stable way, thus competing with the binding of PNA. Based on these factors, a 18mer PNA sequence was chosen as a good compromise to have good selectivity in the recognition of the target strand, good stability, and fair yields in solid-phase synthesis.

Since the mature miR-210 is a 22mer, we had the possibility to chose which region to target to reduce the possibility to have multiple targets inside the cytosol. We then performed a BLAST search for 3 different PNA sequences on the GeneBank RNA database to determine which sequence presented the smallest number of possible interfering interactions with undesired mRNA targets, thus theoretically giving rise to more selective miR binding. As shown in Figure 3.3, comparison of the mRNA with the largest number of base pairing for each PNA indicated that the PNA matching the 5' side of the miR sequence should present less unspecific mRNA binding (13 base pairs).

Mature mir-210: 5'-CUG UGC GUG UGA CAG CGG CUG A-3'

PNA 3'-side: H-T CAG CCG CTG TCA CAC GC-NH_2
NM 000138.3 3-16(14) FBN1

PNA 5'-side: H- CCG CTG TCA CAC GCA CAG-NH_2
NM 005928.1 1-13(13) MFGE8

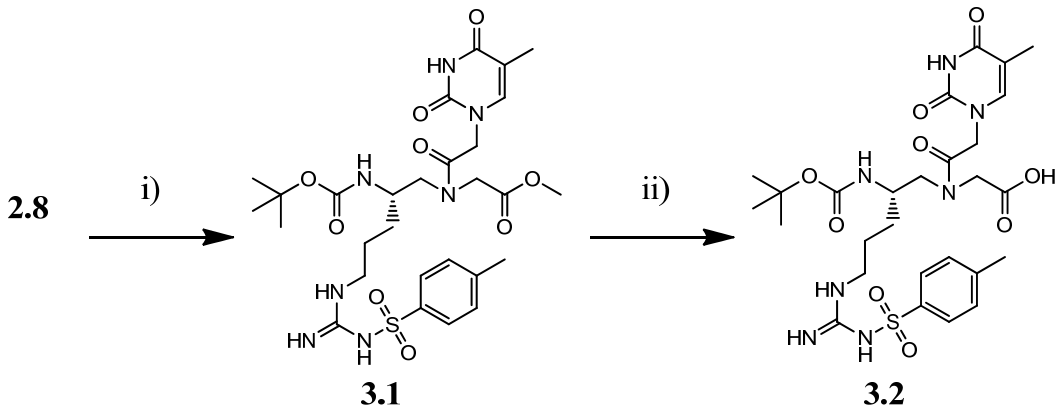
Central PNA: H-AG CCG CTG TCA CAC GCA C-NH_2
NM 000138.3 5-18(14) FBN1

Figure 3.3: mature miR-210 sequence and PNA probes blasted on database. Below each sequence is shown the top scored interferent.

Based on these results and on models proposed in Figure **3.2**, six different PNA probes were synthesized (Table 3.1), all sequences were also synthesized with a fluorescent (5(6)-carboxyfluoresceine) unit for the evaluation of cellular uptake. Moreover, they are expected to possibly bind to both pre-miR-210 and mature miR-210.

Table 3.1: PNA sequences synthesized in this study. Underlined letters are for modified monomer.

Name	Model name	Sequence	Modification
R_8-pep	CPP peptide	Fl-AEEA-RRRRRRRR-NH_2	None
PNA 3.1	PNA "naked"	H-CCGCTGTCACACGCACAG-NH_2	None
PNA 3.2	poliArg-PNA	H-RRRRRRRR-CCGCTGTCACACGCACAG-NH_2	None
PNA 3.3	Model 1	H-C<u>C</u>G<u>C</u>T<u>G</u>T<u>C</u>A<u>C</u>A<u>C</u>G<u>C</u>A<u>C</u>AG-NH_2	5L-Arg
PNA 3.4	Model 2	H-<u>CCGCTGTC</u>ACACGCACAG-NH_2	5L-Arg
PNA 3.5	Model 1	H-C<u>C</u>G<u>C</u>T<u>G</u>T<u>C</u>A<u>C</u>A<u>C</u>G<u>C</u>A<u>C</u>AG-NH_2	2D-Arg
PNA 3.6	Model 2	H-<u>CCGCTGTC</u>ACACGCACAG-NH_2	2D-Arg



Scheme 3.1: synthesis of the 5L-Arg-T-PNA monomer 3.2: i) CMT, EDC·HCl, DHBtOH in DMF, 76%; ii) $Ba(OH)_2$ in THF/H_2O = 1: 1, 62%.

For the synthesis of achiral **PNA 3.1**, **PNA 3.2** and of 5L-modified **PNA 3.3**, **PNA 3.4** a standard Boc-based manual solid phase peptide synthesis was used with standard monomers or 5L-Arg monomers synthesized as reported in chapter 2 (**2.11**, **2.12** and **3.2**); instead for the synthesis of the 2D-modified **PNA 3.5** and **PNA 3.6** the sub-monomeric approach described in section 2.2.1 was followed. The synthesis of the monomer **3.2** (Scheme 3.1) followed a scheme similar to those of previously described 5L-Arg monomers.

3.2.1 - PNA:DNA and PNA:RNA duplex stability

The long PNA sequence and the additional electrostatic contribution provided by the guanidinium groups of the arginine side chain (**PNA 3.3-6**) induce an exceptionally high melting temperature in water which exceeds 90°C. The differences between different structures can thus best be appreciated using strongly denaturing conditions, such as in a buffer containing 5M urea (Figure 3.4).

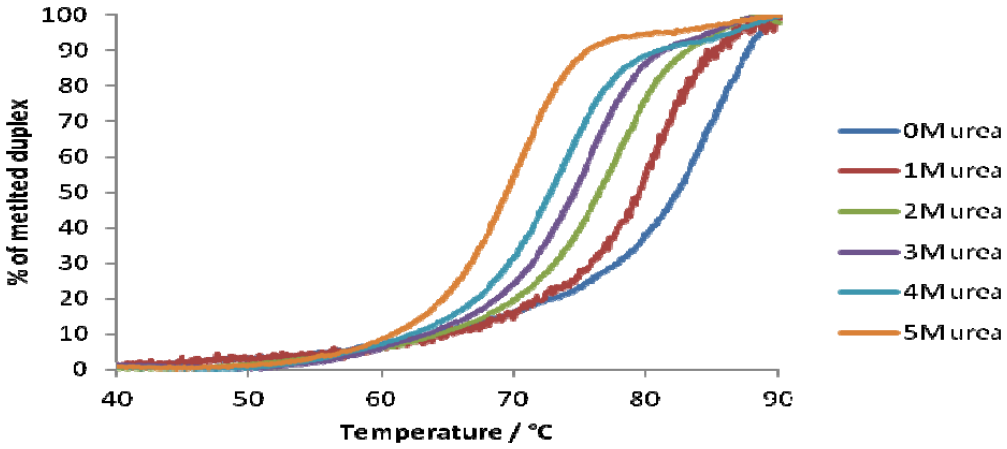


Figure 3.4: melting profile of PNA 3.3:DNA FM duplexes at different urea concentration. All measurements were done in PBS buffer, pH 7; concentration of each strand was 5 μM.

Most notably, the T_m observed even under strong denaturing conditions are still very high if compared to room temperature (Table 3.2), being always in the range 70-88°C. Thus all these compounds are able to strongly bind both DNA and RNA even under this severely unfavourable environment. The highest temperatures were observed for **PNA 3.2**, which is conjugated with the arginine peptide, in agreement with previous results obtained with PNA conjugated with the cationic SV40 NLS peptide[24]. The melting temperatures of the backbone modified PNA:DNA duplex under denaturing conditions follow a general trend observed for this type of derivatives in other studies[21], with the 5L-modified (**PNA 3.3** and **PNA 3.4**) or the 2D-modified PNA (**PNA 3.5** and **PNA 3.6**) showing higher T_m than unmodified **PNA 3.1**. This effect can be attributed to electrostatic contributions to DNA binding, which are present for both 5L- and 2D-Arg-PNA. 5L-modified PNA (**PNA 3.3** and **PNA 3.4**) showed higher T_m than 2D modified (**PNA 3.5** and **PNA 3.6**), in line with previously reported data, supporting the model of a preferred preorganization of the 5L-modified backbone[25], which is more suitable to form the right handed conformation of the

PNA:DNA duplex. However, under these strongly denaturing conditions, the differences in melting temperatures as a function of backbone modification and stereochemistry were less pronounced than expected from previous studies.

Table 3.2: melting temperatures (T_m) for PNA:DNA and PNA:RNA complexes with full match (FM) or with a single mismatch (MM); DNA FM: 5'-CTG TGC GTG TGA CAG CGG-3', DNA MM: 5'-CTG TGC GTG TTA CAG CGG-3', RNA FM: 5'-CUG UGC GUG UGA CAG CGG-3', RNA MM: 5'-CUG UGC GUG UUA CAG CGG-3'.

PNA	DNA FM in water	DNA FM in 5M urea	DNA MM in 5M urea	RNA FM in water	RNA FM in 5M urea	RNA MM in 5M urea
PNA 3.1	85 °C	70 °C	58 °C	88 °C	77 °C	62 °C
PNA 3.2	> 90 °C	80 °C	68 °C	> 90 °C	88 °C	73 °C
PNA 3.3	> 90 °C	77 °C	65 °C	> 90 °C	76 °C	60 °C (broad)
PNA 3.4	> 90 °C	77 °C	56 °C	> 90 °C	86 °C	75 °C
PNA 3.5	> 90 °C	74 °C	64 °C	> 90 °C	78 °C	62 °C
PNA 3.6	> 90 °C	77 °C	69 °C	> 90 °C	75 °C	65 °C

The melting temperatures observed with RNA, as expected from literature[26], were higher than those obtained with DNA for unmodified **PNA 3.1**, for the peptide-conjugated **PNA 3.2**, and for the backbone modified **PNA 3.4** and **PNA 3.5**, whereas they were slightly lower in the case of **PNA 3.3** and **PNA 3.6**, suggesting that the conformational differences between the PNA:DNA structure and the PNA:RNA structure[27] can have a strong impact when associated with strongly constrained PNA structure.

Sequence selectivity is a very important issue in view of the use of PNA in cellular systems, since partial hybridization could occur with other unrelated miR, or, most importantly, with mRNA targets, which, according to our analysis, could lead to a maximum pairing of 13 bases. We analysed the sequence selectivity of our PNAs using a stricter model (DNA2 and RNA2) bearing a single base mismatch (G→T or G→U) located in the middle of the target sequence. In all cases the sequence selectivity was confirmed by significant lowering of the melting temperature. The ranges of ΔTm due to the presence of this mismatch were 8-21°C for DNA, with the maximum drop in the case of **PNA 3.4** and the minimum observed in the case of **PNA 3.6**. For the latter it should be considered that the mismatched base was facing the last chiral monomer, thus preventing the cooperative effect observed in the case of 2D-modifed "chiral box" PNA. For RNA, the ΔTm due to the presence of the mismatch were less disperse, going from 10°C to 16°C, with best results obtained for **PNA 3.2** and **PNA 3.3**, and worst for **PNA 3.6**. Thus, also in the case of mismatch

recognition the effects observed for DNA and RNA turned out to be different; this prediction of both binding and selectivity of PNA for RNA cannot be inferred from DNA data.

3.2.2 - Circular dichroism studies

The PNA:DNA antiparallel duplex formed by an achiral PNA shows a complex spectrum, evidencing several transitions. Common features of PNA:DNA spectra are an intense maximum in the range 260-265 nm and a minimum at 240-245 nm, which is much more variable both in sign and intensity. Both bands are slightly shifted towards shorter wavelengths if compared with DNA:DNA duplexes of the same sequence. In Figure 3.1 the spectra of the **PNA 3.1-3.6** with antiparallel full-match DNA FM are reported, showing very similar features. Thus, although electrostatic interactions are differently distributed along the PNA chain, the overall conformation of the duplex is not significantly distorted from that of the achiral, unmodified **PNA 3.1**. A similar behaviour was observed in the presence of antiparallel RNA FM, corresponding to the target region of miR 210, but in this case more pronounced differences were observed; the most significant differences from the spectrum of **PNA 3.1**:RNA duplex were observed in the case of **PNA 3.3**, corresponding to 5L-modified PNA with model 1 structure, and to a lesser extent for **PNA 3.6**, which is 2D-modified PNA following the model 2 scheme. Thus, this type of PNA seems to induce a preorganization which prevents the binding to RNA in the same manner as unmodified PNA. This difference can explain the difference in stability observed previously.

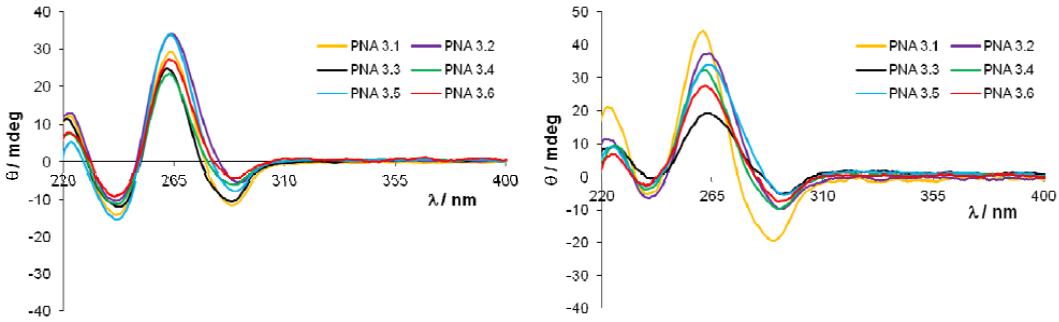


Figure 3.5: CD spectra of PNA:DNA (left) and PNA:RNA (right) duplexes with unmodified and modified PNA. All measurements were done in PBS buffer, pH 7; concentration of each strand was 5 μM.

Same results were obtained when the PNA:DNA or PNA:RNA duplexes were analysed in a solution containing 5M urea, confirming that the denaturant is not able to induce dissociation of these duplexes at room temperature, and, more important, not even to distort the overall structure of

the duplexes. All these data are in agreement with the extremely high thermal stability observed for these duplexes, as described in the previous paragraph.

3.2.3 - Cellular uptake

The K562 cell line has proved to be an excellent experimental system for evaluating the anti-miR-210 activity, since the expression of this microRNA is increased during mithramycin-induced erythroid differentiation[23]. From the technical point of view, erythroid differentiation can be easily visualized via a benzidine test, and treatment with anti-miR 210 agents, including PNAs targeting miR-210, prevents the differentiation process, leading to a different phenotype in respect to MTH-treated cells.

The first step necessary for obtaining anti-miR activity in cells is the effective cellular uptake. We studied the uptake of PNAs in this cellular model, and compared the results obtained in the presence of increasing concentrations of fluorescein-labelled PNAs for 24 and 48 hours by fluorescence-activated cell sorting analysis (FACS). The results are shown in the left side of Figure 3.6. Since different PNA structures could be affected by proteases present in serum, leading to different cellular uptake, these experiments were conducted in the presence or in the absence of serum. In agreement with the model proposed, in the absence of serum low internalization of **PNA 3.1-Fl** was obtained (red line); on the contrary, the R_8-conjugated **PNA 3.2-Fl** (blue line) is efficiently internalized. Interestingly, all the C-2 and C-5 backbone modified **PNA 3.3-3.6-Fl**, independently from the model proposed, exhibit an even higher efficiency of internalization than **PNA 3.2-Fl**. In the presence of serum, no decrease of the uptake of **PNA 3.5-Fl** and **PNA 3.6-Fl** was observed, and only a minor decrease of uptake of **PNA 3.3-Fl** and **PNA 3.4-Fl** were detectable. The most remarkable difference is evident for **PNA 3.2-Fl**, which exhibits a decreased uptake by the target cells in the presence of serum. This effect might be due to several factors, including aptameric activity of this PNA, leading to binding to serum factors, or to degradation of the peptide part of the molecule by serum proteases, which instead are unable to degrade the unnatural PNA backbone[28]. The embedding of the peptide features in the PNA backbone leads to very high stability to these agents, though preserving the recognition elements of the arginine peptide.

On the basis of the FACS analysis we cannot exclude the possibility that the fluorescence signal is at least partially due to cell-surface interactions caused by the positively charged arginine tails, which might strongly interact with negatively charged protein components on the cell surface. In

order to determine the distribution of PNAs within the target cells, analysis was performed using a fluorescence microscope. A representative example of this analysis is presented in the right side of Figure 3.6, which shows that the fluorescence of the **PNA 3.2** is mainly cytoplasmatic. Similar results were obtained with all the other PNAs. Low efficient uptake was also found with fluorescence microscopy using **PNA 3.1**. These data suggest that a real uptake occurs, rather than a simple interaction of PNAs with the cellular membrane either in FACS analysis and in microscopy.

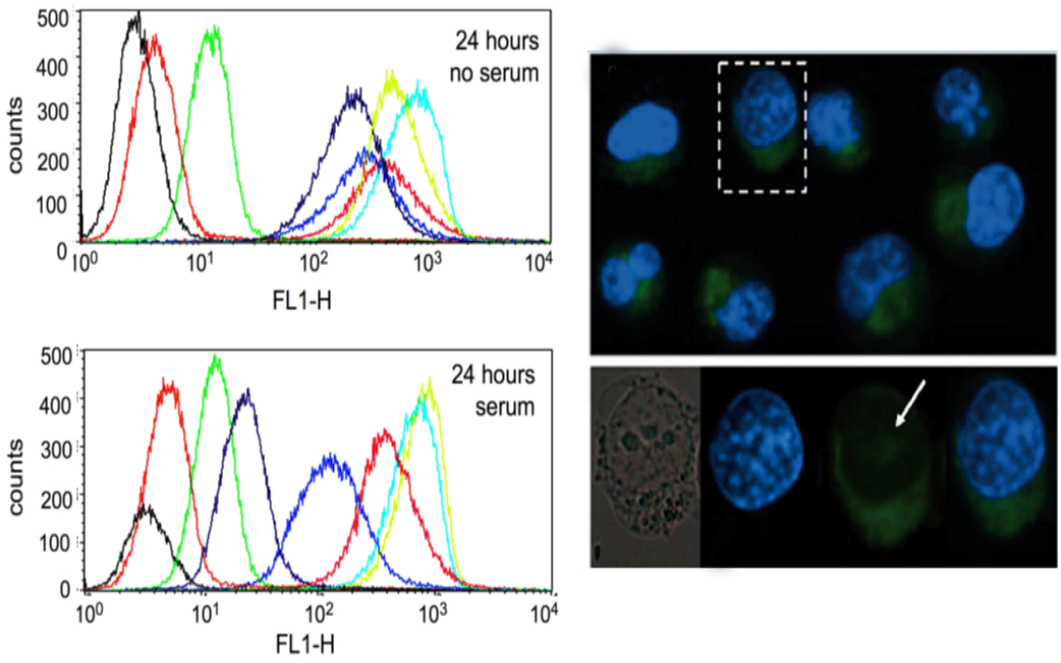


Figure 3.6: left side: FACS analysis showing the uptake after 24 incubation of K562 cells, in the absence (upper panel) or in the presence of serum (lower panel), of fluorescein-labeled probes: R_8-pep (green lines), PNA 3.1-Fl (light red lines), PNA 3.2-Fl (dark blue lines), PNA 3.3-Fl (light blue lines), PNA 3.4-Fl (dark red lines), PNA 3.5-Fl (cyan lines) and PNA 3.6-Fl (yellow lines). Right side intracellular distribution of PNA: in this experiment, cells were cultured for 24 hours with 2 μM of PNA 3.2-Fl and then analyzed using a fluorescence microscope; the pictures in the lower part of the panel are the respectively: optical microscope image, Hoechst 33258 staining (selective for nuclei), fluorescence staining and merged image.

Furthermore, in order to prove the *in vivo* stability of the systems, similar analyses were conducted pre-incubating the PNAs with trypsin before the interaction with K562 cells: the results shown in Figure 3.7 demonstrated, as expected, that trypsin treatment affects the cellular uptake of **PNA 3.2-Fl**, whereas it does not affect the uptake by backbone modified arginine-PNA. Thus the embedding into the PNA backbone of arginine residues allows to obtain efficient cellular uptake without the risk of degradation even by a basic amino acid-specific peptidase, allowing a long-lasting effect. On the contrary, the carrier peptide is, to a certain degree, an Achille's heel of the **PNA 3.2-Fl** molecule, leading to possible degradation by proteases, an effect to be considered in view of possible developments of PNAs as drugs based on anti-miR activity.

The use of backbone modified PNAs can therefore be viewed as a strategy to deliver PNAs in an efficient and less environmental sensitive way, as demonstrated by Ly and co-workers in several papers[12,29].

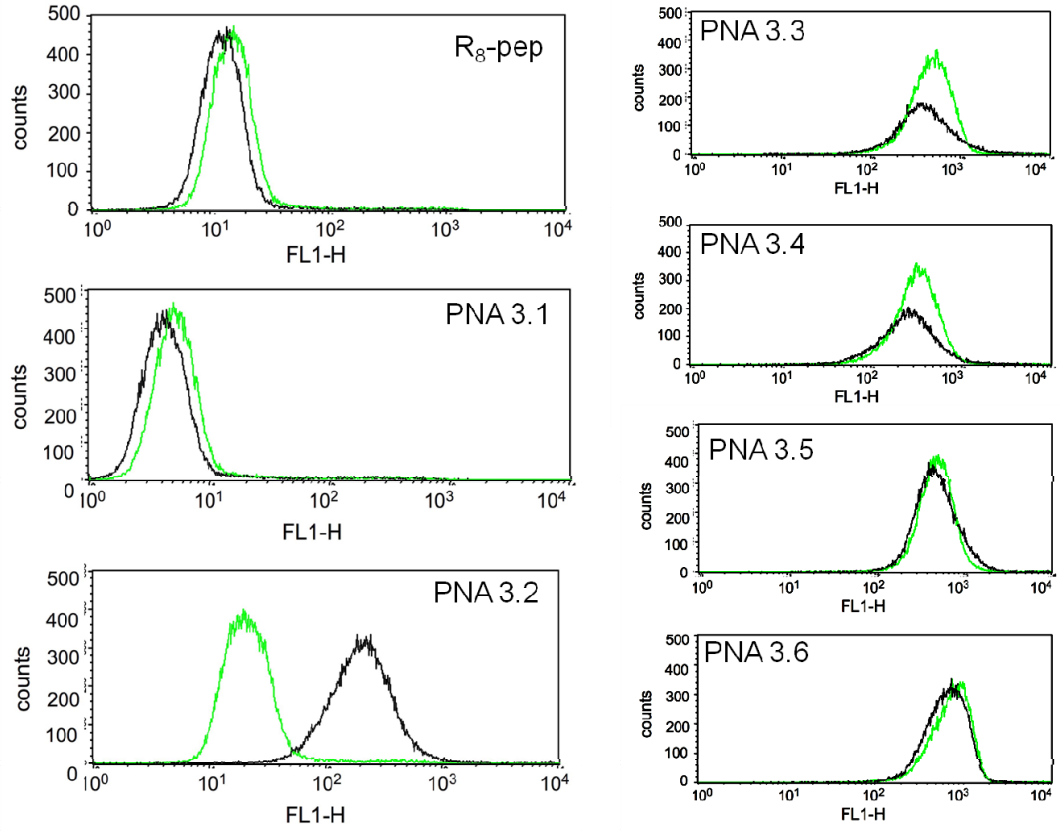


Figure 3.7: effects of pre-incubation with (green lines) or without (black lines) trypsin on the cellular uptake.

3.2.4 - Biological effects of PNAs

Preliminary tests on the biological effects were conducted comparing the anti-miR210 activity of achiral PNAs and CPP. These experiment performed on mithramycin (MTH) induced K562 cells with an increased miR210 level[23], evaluating the different targets involved in the biological cycle of this miR, such as: miR210 level, *raptor* mRNA content (downregulated by miR210), γ-globin mRNA content (downregulated by *raptor*) and erythroid differentiation (related to γ-globin content). A representation of the relationship among all the targets is shown in Figure 3.8.

In a first set of experiments, the anti-miR activity of the "naked" PNA **3.1** and of the R_8-conjugated PNA **3.2** were compared, while the R8-peptide (R_8 pep) was used as a control, in order to assess the eventual contribution of the peptide itself to some biological effect. As show in Figure 3.9A the miR210 level was significantly reduced only in presence of **PNA 3.2** (irrespective of the MTH treatment) with a minor inhibition in the case of **PNA 3.1**; and in Figure 3.9B that this effect is show to be restricted to miR210 with only minor changes in some miR, demonstrating the specificity of the PNA treatment. Evaluating the effect of the treatment on the other molecular target (Figure 3.9-C-E) it can be noted that, as expected, the raptor mRNA content was increased while γ-globin mRNA and diffentiated cells were reduced.

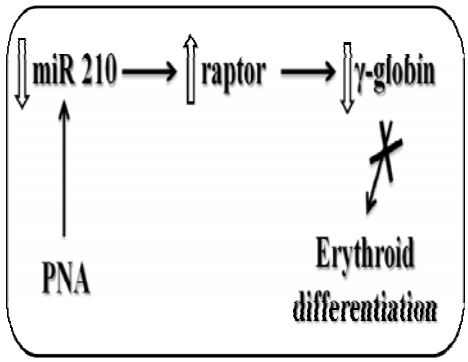


Figure 3.8: miR 210 regulation pathway.

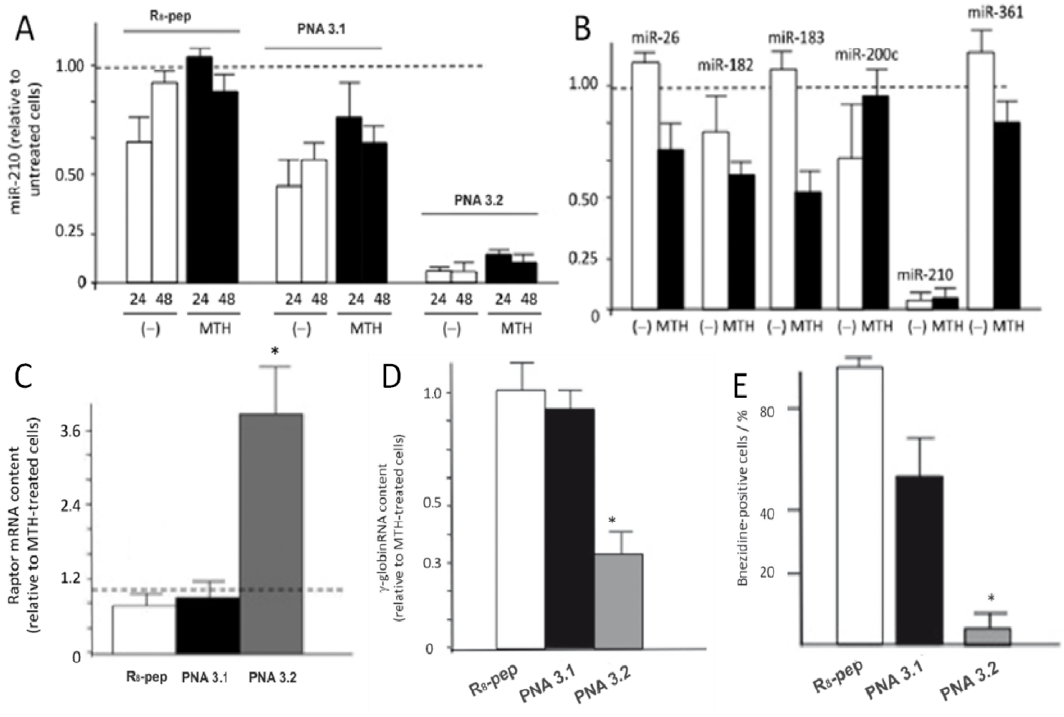


Figure 3.9: A) Effects of treatment with 2 μM PNAs on miR-210 content in cells cultured for 24 and 48 h in the absence (white bars) and in the presence(black bars) of MTH; B) Effects of treating K562 cells for 48 h with PNA 3.2 on the content of miR-26, miR-182, miR-183, miR-200c, miR-210, miR-361, and miR-221. C) Effects of 2 μM PNAs on raptor mRNA analyzed after 48 h; D) Effects of 500 nM PNAs on the percentage of benzidine-positive cells after 5 days treatment with 15 nM MTH; E) Effects of 2 μM PNAs on γ-globin content in K562 cell cultures for 3 days with 15 nM MTH; all RT-qPCR analyses are reported with respect to MTH treated cells; all data represent the average ±SD (n=3, *$p<0.05$).

Thus, the results achieved after 2/3 days with a single administration are in agreement with those obtained with a commercially available anti-miR210, administered every day in the presence of a transfecting agent (data not shown). Then we compared the activity of those unmodified PNAs with the modified ones. Thus, treating K562 cells with a 2μM solution of PNAs with or without MTH induction, after 24 hours, RNA was isolated and miR-210 assayed by RT-qPCR analysis.

Figure 3.10 reports the data obtained, showing that in the absence of MTH treatment miR bioavailability was reduced with the modified PNA as well as with the peptide-conjugated PNA, whereas it was unchanged in the case of **PNA 3.1**, as previously described. Stronger inhibition was observed for **PNA 3.2**, **PNA3.3**, **PNA 3.4** and **PNA 3.6** to a similar extent, whereas less efficient inhibition was obtained with **PNA 3.5**. This reduced miR210 amount might be related to inhibition of pre-miR maturation or to unavailability of miR to RT-PCR analysis, due to strong complexation with the PNA. In both cases, the bioavailability of miR for targeting other genes should be reduced.

Therefore, although all the PNAs considered have very high melting temperatures with complementary RNA, and thus bind quantitatively to complementary RNA *in vitro*, in the mithramycin-treated cells, which produce high levels of miR-210, the inhibition becomes more competitive. Thus, the anti-miR activity is modulated by the type of substitution on the monomer (5L-Arg better than 2D-Arg) and by the distribution of the modified monomers along the chain (model 2 better than model 1). The most efficient model turned out to be the use of C5-substituted monomers with a compact arrangement (model 2), and to a less extent the unmodified PNA bearing the polyarginine chain.

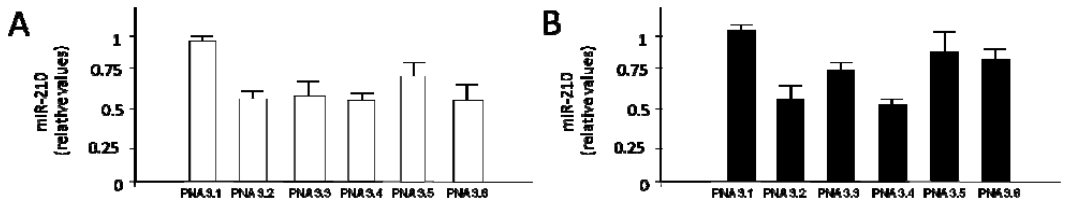


Figure 3.10: Effects of treatment with PNAs on miR-210 content in cells cultured for 24 hours in the absence (A) or in the presence (B) of 15 nM MTH. Data are reported as arbitrary units relative to cells treated with PNA 3.1 (white boxes) or PNA 3.1 and MTH (black boxes) (average ± SD; n = 3; *, $p < 0.05$).

On account of the higher effect on miR-210 exhibited by **PNA 3.2** and **PNA 3.4**, these two probes were further employed to determine the effects on the other targets. We have first compared the activity of **PNA 3.1**, **PNA 3.2** and **PNA 3.4** on γ-globin gene expression and MTH-induced differentiation. As expected from previous experiments, the inhibition of miR210 induces overexpression of the *raptor* gene, leading in turn to underexpression of the mithramycin-induced γ-

globin gene. Fully in agreement with the strong inhibition of miR210, both **PNA 3.2** and **PNA 3.4** inhibit γ-globin gene expression in mithramycin treated K562 cells to a similar extent, whereas **PNA 3.1** did not show appreciable effects, as determined by RT-qPCR analysis (Figure 3.11).

In order to evaluate the effect on mithramycin induced cellular differentiation, K562 cells were treated with MTH in the presence of 2 μM **R_8-pep**, **PNA 3.1**, **PNA 3.2** and **PNA 3.4**. The proportion of benzidine-positive cells was determined after 5 days of induction. The quantitative analysis is reported in Figure 3.11B, while in Figure 3.11C a representative experiment indicating that, unlike **PNA 3.1**, lacking conjugation to the R_8 peptide, both **PNA 3.2** and **PNA 3.4** are able to inhibit erythroid differentiation. In this case, significant differences are observed between the effects of the two PNAs, despite they show comparable miR-210 inhibition and γ-globin mRNA reduction, this last result indicates that the conjugation of arginine peptide and the incorporation of the arginine side chains into the PNA structure do not have the same effect. This effect cannot be attributed neither to differences in the cellular systems used, or to interference with the miR-210 pathway, rather it should be generated by the activity of the R_8-PNA conjugate in other different

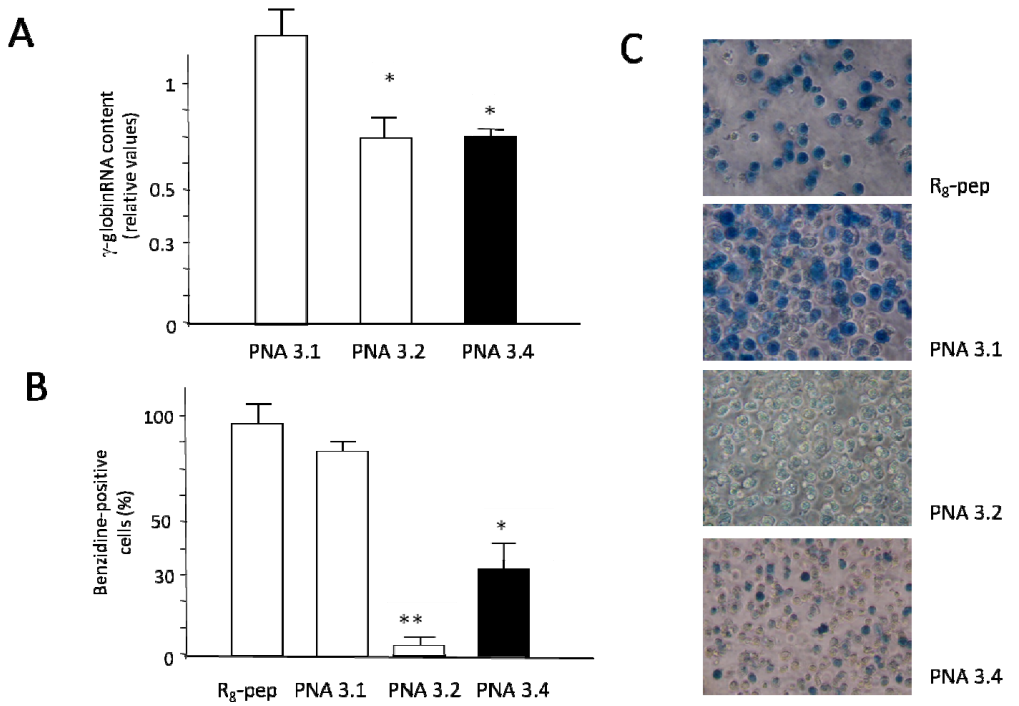


Figure 3.11: A) Effects of 2 μM PNA 3.1, PNA 3.2 and PNA 3.4 molecules on γ-globin content in K562 cells cultures for 72 hours with 15 nM MTH. Data are reported as arbitrary units relative to cells treated with R_8-pep and MTH (average ± SD; n = 3; *, p <0.05); B) and C) Effects of treatment of K562 cells for 72 hours with R_8-pep, PNA 3.1, PNA 3.2, PNA 3.4 molecules on K562 cells induced to erythroid differentiation by 15 nM MTH (B: summary of three independent experiments; C: a representative experiment showing staining of the cells with benzidine).

pathways, probably mediated by the conjugation with polyarginine peptide. Though this effect in this case reinforces the anti-differentiating activity of **PNA 3.2**, its consequences are not easily predictable, whereas for the chiral **PNA 3.4** the effect observed is a consequence only of the rationally designed anti-miR activity.

3.3 - Conclusions

With this study we have demonstrated for the first time an anti-miR activity of the backbone modified PNAs. Overall, the combined good cellular uptake, higher biostability, good anti-miR activity, and reduction of differentiated cells and of γ-globin production, make these molecules ideal candidates for the next-generation of drugs able to perform gene modulation. The experimental system (inhibition of miR210 in K562 cells) has been chosen in order to better assess the performances of the different types of PNA, modification as anti-miR. However the present results represent a systematic study involving both C2- and C5-modified PNA and with different spacing between modification and can be extended to other interesting targets, most notably tumor-associated miR (onco-miR). The results show that high affinity and excellent cellular uptake can be obtained using all the types of arginine-derived PNAs, although differences in anti-miR activity are still present as a function of the type of modification and of the model used, with contiguous stretch of modified monomers being the preferred model.

Furthermore, our results might be relevant also in the field of applied biology, since miR-210 has been demonstrated to be linked to hypoxia and up-regulated together with several hypoxia-related genes[30], and is hypothesized to be involved in cancer. Since miR-210 is a marker of the growth of advanced tumors under hypoxic conditions and correlates in general with poor prognosis[31,32], it has been proposed as a possible therapeutic target. For example, Radojicic *et al.*, who analyzed forty-nine primary triple-negative breast cancer cases, along with 34 matched tumor-associated normal samples, reported miR-210 as overexpressed in the triple-negative primary breast cancers, suggesting a role in breast cancer tumor progression [33]. Therefore, PNAs against miR-210 might be useful on one hand for understanding the function of miR-210 in solid tumors, on the other to alter miR-210 expression in tumors in which this microRNA has a clear role.

The results of these studies have been published in the following papers: Gene modulation by Peptide Nucleic Acids (PNAs) targeting microRNA (miR)[34], A novel approach for modification of

gene expression and drug development[35], miRNA therapeutics: delivery and biological activity of peptide nucleic acids targeting miRNAs[36], Modulation of the Biological Activity of microRNA-210 with Peptide Nucleic Acids (PNAs)[37], and submitted to ChemBioChem.

3.4 - Experimental section

General. Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, Link, ASM and used without further purification. All reactions were carried out under a nitrogen atmosphere with dry solvents under anhydrous conditions, unless otherwise noted. Anhydrous solvents were obtained by distillation or anhydrification with molecular sieves. Reactions were monitored by TLC carried out on 0.25mm E. Merck silica-gel plates (60F-254) by using UV light as visualizing agent and ninhydrin solution and heat as developing agents. E. Merck silica gel (60, particle size 0.040÷0.063 mm) was used for flash-column chromatography. NMR spectra were recorded on Bruker Avance 400 instrument and calibrated by using residual undeuterated solvent as an internal reference. The following abbreviations were used to explain the multiplicities: s=singlet, d=doublet, t=triplet, q=quartet, m=multiplet, and br=broad. IR spectra were measured using a FT-IR Thermo Nicolet 5700, in transmission mode using KBr or NaCl. HPLC-UV-MS were recorded by using a Waters Alliance 2695 HPLC with Micromass Quattro microAPI spectrometer, a Waters 996 PDA and equipped with a Phenomenex Jupiter column (250x4.6mm, 5µ, C18, 300Å) (method A, 5 minutes in H_2O 0.2% FA, then linear gradient to 50% MeCN 0.2% FA in 30 minutes at a flow rate of 1ml/min). PNA oligomers were purified with RP-HPLC using a XTerra Prep RP_{18} column (7.8x300mm, 10µm) (method B, linear gradient from H_2O 0.1% TFA to 50% MeCN 0.1 % TFA in 30 minutes at a flow rate of 4.0ml/min).

Boc-PNA(5L-Arg$_{(Tos)}$)-T-OMe (3.1). In a round bottom flask 2-(thymin-1-yl)acetic acid (190mg, 1.03mmol) was dissolved in 6ml DMF at 0° C together with DHBtOH (169mg, 1.03mmol) and DIPEA (270µl, 1.55mmol). The EDC·HCl (199mg, 1.03mmol) was added and the solution was stirred for 10 minutes at 0° C, for 20 minutes at room temperature, then **2.8** (251mg, 0.52mmol) was added to the mixture. The solution was stirred overnight and the DMF was then evaporated under reduced pressure. The residue was taken up with 50ml AcOEt and washed with saturated $KHSO_4$ (2x25ml), saturated $NaHCO_3$ (2x25ml) and brine (25ml). The organic layer was dried over Na_2SO_4, the solvent evaporated under reduced pressure and the residue was purified via flash chromatography (from AcOEt to AcOEt/MeOH 95:5) to afford **3.1** as a pale yellow foamy solid (258mg, 76%); **TLC (AcOEt/MeOH 95:5) Rf:** 0.26, **MP (°C):** decompose without melt at 145°C; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 11.31 (1H, s), 7.63 (2H, d, J= 7.4Hz), 7.29 (2H, d, J= 7.4Hz), 7.21 (1H, s), 7.05 (1H, br s), 6.83 (2H, d, J= 7.4Hz), 6.79 (1H, br s), 6.58 (1H, br s), 4.75 (1H, d, J= 17.6Hz), 4.62 (1H, d, J= 17.6Hz), 4.09 (1H, d, J= 17.6Hz), 3.98 (1H, d, J= 18Hz), 3.62 (3H, s), 3.57÷3.45 (1H, br m), 3.40÷3.20 (2H, br m), 3.10÷2.95 (2H, br m), 2.35 (3H, s), 1.76 (3H, s), 1.55÷1.20 (4H, m), 1.36 (9H, s); **^{13}C NMR (DMSO, 100MHz) δ(ppm):** 169.9, 167.9, 164.8, 157.1, 156.1, 151.4, 142.4, 142.2, 141.5, 129.5, 126.1, 108.6, 78.5, 52.2, 51.9, 48.9, 48.3, 48.1, 29.8, 29.2, 28.7, 21.3, 12.41 *one signal estimated under DMSO peak; **MS (ESI, MeOH):** *m/z* calcd for $C_{28}H_{41}N_7O_9S$ [M]: 651.26865, found: 652.3 $[M+H]^+$, 674.3 $[M+Na]^+$, 690.3 $[M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 652.27661 $[C_{28}H_{42}N_7O_9S]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3444.3 (m), 3342.8 (m), 3142.8 (w), 2978.4 (m), 958.0 (m), 2929.4 (m), 1733.7 (s), 1699.9 (s), 1684.0 (s), 1558.4 (s), 1472.5 (m), 1367.9 (m), 1253.4 (s), 1132.5 (m).

Boc-PNA(5L-Arg$_{(Tos)}$)-T-OH (3.2). To a stirred solution of **3.1** (240mg, 0.37mmol) in 20ml THF was added a solution of $Ba(OH)_2 \cdot 8H_2O$ (175.1mg, 0.55mmol) in 20ml water. The reaction mixture was stirred for 10 minutes. The THF was then evaporated and the pH of the solution was lowered to 4.5 with a diluted solution of HCl to induce the precipitation of the product. The solution was cooled at 4°C for 2 hours, then filtered over Buchner to afford **3.2** as white solid (145mg, 62%) ; **TLC (AcOEt/MeOH 9:1) Rf:** 0.00, **MP (°C):** decompose without melt at 179°C; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 12.81 (1H, s), 11.31 (1H, s), 7.64 (2H, d, J= 7.6Hz), 7.29 (2H, d, J= 7.6Hz), 7.22 (1H, s), 7.05 (1H, br s), 6.93 (1H, br s), 6.84 (2H, d, J= 7.6Hz), 6.64 (1H, br s), 4.75 (1H, d, J= 17.6Hz), 4.59 (1H, d, J= 17.8Hz), 4.01 (1H, d, J= 17.6Hz), 3.88 (1H, d, J= 17.6Hz), 3.67 (1H, br m), 3.55÷2.90 (4H, br m), 2.35 (3H, s), 1.76 (3H, s), 1.50÷1.15 (4H, m), 1.37 (9H, s); **^{13}C NMR (DMSO, 100MHz) δ(ppm):** 171.4, 170.9, 168.2, 167.7, 164.8, 157.1, 156.1, 151.5, 142.3, 141.5, 129.5, 126.1, 108.6, 78.4, 51.8, 49.0, 48.9, 48.4, 48.13, 29.5, 29.2, 28.7, 21.3, 12.4 *one signal estimated under DMSO peak; **MS (ESI, MeOH):** *m/z* calcd for $C_{27}H_{39}N_7O_9S$ [M]: 637.25300, found: 638.3 $[M+H]^+$, 660.3 $[M+Na]^+$, 676.2 $[M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 636.24564 $[C_{27}H_{38}N_7O_9S]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3437.4 (m), 3346.0 (m), 3063.9 (w), 2977.8 (m), 1683.7 (s), 1549.6 (s), 1455.8 (s), 1367.8 (m), 1251.9 (s), 1132.1 (m).

Synthesis and characterization of PNAs. The synthesis of the Fl-R_8 peptide was performed using standard automated Fmoc-based chemistry with HBTU/DIPEA coupling on a Rink amide resin loaded with Fmoc-Arg(Pmc)-OH as first monomer (0.2mmol/g). The achiral PNAs (**PNA 3.1**, **PNA 3.2**, **PNA 3.1-Fl** and **PNA 3.2-Fl**) were synthesized with standard manual Boc-based chemistry using commercially available monomers with HBTU/DIPEA coupling. The 5L-modified PNAs (**PNA 3.3**, **PNA 3.4**, **PNA 3.3-Fl** and **PNA 3.4-Fl**) were synthesized with standard manual Boc-based chemistry using **2.12**, **2.11** and **3.2** in addition to normal monomers. The 2D-modified PNAs (**PNA 3.5**, **PNA 3.6**, **PNA 3.5-Fl** and **PNA 3.6-Fl**) were synthesized with submonomeric approach using **2.5** in addition to normal monomer. All the PNAs were synthesized at the 5 mmol scale using MBHA resin loaded with Boc-PNA-G(Z)-OH as first monomer (0.2mmol/g). The carboxyfluorescein unit was introduced after the 2-(2-(Fmoc-amino)ethoxy)ethoxyacetic acid (AEEA) spacer using DIC/DhBtOH protocol after Fmoc deprotection with a 20% piperidine solution in DMF. PNAs purification was performed by RP-HPLC with UV detection at 260nm. The purity and identity of the purified PNAs were determined by HPLC-UV-MS. **R_8 peptide:** 15%; ESI-MS: *m/z* calcd 1768.9576 [M], found: 590.8 $[MH_3]^{3+}$, 443.4 $[MH_4]^{4+}$, 355.0 [M H$_5$]$^+$; **PNA 3.1**: 39%, t_r: 19.8min; ESI-MS: *m/z* calcd 4822.9285 [M]: 1207.3 $[MH_4]^{4+}$, 966.1 $[MH_5]^{5+}$, 805.1 $[MH_6]^{6+}$, 690.3 $[MH_7]^{7+}$, 604.1 $[MH_8]^{8+}$; **PNA 3.2**: 18%, t_r: 14.6min; ESI-MS: *m/z* calcd 6071.7374 [M]: 1013.5 $[MH_6]^{6+}$, 868.8 $[MH_7]^{7+}$, 760.4 $[MH_8]^{8+}$, 675.9 $[MH_9]^{9+}$, 608.6 $[MH_{10}]^{10+}$, 553.3 $[MH_{11}]^{11+}$; **PNA 3.3**: 19%; t_r: 15.9min; ESI-MS: *m/z* calcd 5615.5657 [M]: 1124.6 $[MH_5]^+$, 937.4 $[MH_6]^+$, 803.7 $[MH_7]^+$, 703.3 $[MH_8]^+$; **PNA 3.4**: 11%; t_r: 16.1min; ESI-MS: *m/z* calcd 5615.5657 [M]: 1124.9 $[MH_5]^{5+}$, 937.5 $[MH_6]^{6+}$, 803.6 $[MH_7]^{7+}$, 703.4 $[MH_8]^{8+}$, 625.4 $[MH_9]^{9+}$; **PNA 3.5**: 3.6%; t_r: 13.4min; ESI-MS: *m/z* calcd 5615.5657 [M]: 937.1 $[MH_6]^{6+}$, 803.4 $[MH_7]^{7+}$, 703.2 $[MH_8]^{8+}$, 625.3 $[MH_9]^{9+}$; **PNA 3.6**: 3.2%; t_r: 13.4min; ESI-MS: *m/z* calcd 5615.5657 [M]: 1124.9 $[MH_5]^{5+}$, 937.5 $[MH_6]^{6+}$, 803.6 $[MH_7]^{7+}$, 703.4 $[MH_8]^{8+}$, 625.4 $[MH_9]^{9+}$; **PNA 3.1-Fl**: 18%, t_r: 19.8min; ESI-MS: *m/z* calcd 5326.0501 [M]: 1333.1 $[MH_4]^{4+}$, 1066.9 $[MH_5]^{5+}$, 889.1 $[MH_6]^{6+}$, 762.2 $[MH_7]^{7+}$, 667.1 $[MH_8]^{8+}$, 593.1 $[MH_9]^{9+}$; **PNA 3.2-Fl**: 20%, t_r: 16.1min; ESI-MS: *m/z* calcd 6574.8590 [M]: 1097.3 $[MH_6]^{6+}$, 940.7 $[MH_7]^{7+}$, 823.3 $[MH_8]^{8+}$, 732.0 $[MH_9]^{9+}$, 658.8 $[MH_{10}]^{10+}$, 599.0 $[MH_{11}]^{11+}$, 549.3 $[MH_{12}]^{12+}$; **PNA 3.3-Fl**: 17%; t_r: 12.5min; ESI-MS: *m/z* calcd 6118.6873 [M]: 1224.7 $[MH_5]^{5+}$, 1020.9 $[MH_6]^{6+}$, 875.3 $[MH_7]^{7+}$, 766.0 $[MH_8]^{8+}$, 681.0 $[MH_9]^{9+}$, 613.2 $[MH_{10}]^{10+}$; **PNA 3.4-Fl**: 17%; t_r: 12.5min; ESI-MS: *m/z* calcd 6118.6873 [M]: 1224.8 $[MH_5]^{5+}$, 1021.0 $[MH_6]^{6+}$, 875.1 $[MH_7]^{7+}$, 765.9 $[MH_8]^{8+}$, 680.9 $[MH_9]^{9+}$, 613.0 $[MH_{10}]^{10+}$; **PNA 3.5-Fl**: 6.8%; t_r: 12.9min; ESI-MS: *m/z* calcd 6118.6873 [M]: 1021.1 $[MH_6]^{6+}$, 875.8 $[MH_7]^{7+}$, 766.3 $[MH_8]^{8+}$, 681.2 $[MH_9]^{9+}$, 613.2 $[MH_{10}]^{10+}$; **PNA 3.6-Fl**: 2.8%; t_r: 12.6min; ESI-MS: *m/z* calcd 6118.6873 [M]: 1020.8 $[MH_6]^{6+}$, 875.1 $[MH_7]^{7+}$, 766.2 $[MH_8]^{8+}$, 681.0 $[MH_9]^{9+}$. Yields % reported for each PNA are those of purified products, calculated by UV-Vis analysis.

CD and UV measurements. Stock solutions of PNA, and of complementary DNA or RNA synthetic oligonucleotides (Thermo-Fisher Scientific, HPLC-grade) were prepared in double-distilled water, and the PNA concentration was calculated by UV absorbance with the following extinction coefficients (ε_{260} [$M^{-1}cm^{-1}$]) for the nucleobases: T 8600, C 6600, A 13700, and G 11700. For DNA and RNA the data provided by the producer were used. From these, solutions containing single stranded PNA, DNA, RNA, or PNA–DNA PNA–RNA duplexes were prepared. Measurement conditions: [PNA]=[DNA] or [RNA]= 5μM in PBS (100mM NaCl, 10mM $NaH_2PO_4 \cdot H_2O$, 0.1mM EDTA, pH 7.0) or in the same buffer containing from 1M to 5M urea as denaturating agent. All the samples were first incubated at 90°C for 5 min, then slowly cooled to room temperature. Thermal denaturation profiles (A_{260} versus T) of the hybrids were measured with a UV/Vis Lambda Bio 20 spectrophotometer equipped with a Peltier temperature programmer PTP6 interfaced to a personal computer. For the temperature range 18÷90°C, A_{260} values were recorded at 0.1°C increments. A melting curve was recorded for each duplex. The melting temperature (Tm) was determined from the maximum of the first derivative of the melting curves. CD measurements were carried out on a Jasco J715 spectropolarimeter equipped with a Peltier PTC 348 temperature controller unit, using solutions prepared as described above for the UV measurements. A scan speed of 50 nm/min was used, and four accumulations for each spectrum. The spectra were corrected by subtracting the signal obtained with the corresponding buffer.

Uptake and cellular studies. FACS analysis, fluorescence microscopy, miR210 , raptor and g-globin expression evaluation by RT-PCR, and differentiation studies were carried out at the Biochemistry and Molecular Biology Department at the University of Ferrara, by the group of Prof. Roberto Gambari, and experimental details are reported in the publications[37].

3.5 - References

[1] Alvarez-Garcia I., Miska E. A., *Development* **2005**, 132, 4653-4662.

[2] He L., Hammon G. J. *Nat. Rev. Genet.* **2004**, 5, 522-531.

[3] Subramania S., Steer C. J. *J. Cell. Physiology* **2010**,223, 89-98.

[4] Wang Y. M., Blelloch R. *Cancer Res.* **2010**, 69, 4093-4096.

[5] Czech M. P. *New. Engl. J. Med.* **2006**, 354, 1194-1195.

[6] a) Oh S. Y., Ju Y. S., Park H. *Mol. Cells* **2009**, 28, 341-345; b) Fabani M. M., Gait M. J. *RNA* **2008**, 14, 336-346; c) Torres A. G., Threlfall R. N., Gait M. J. *Artificial DNA: PNA & XNA* **2011**, 2(3), 1-8.

[7] Rasmussen F. W., Bendifallah N., Zachar V., Shiraishi T., Fink T., Ebbesen P., Nielsen P. E., Koppelhus U., *Oligonucleotides* **2006**, 16, 43-57.

[8] Goun E. A:, Pillow T. H., Jones L. R., Rothbard J. B., Wender P. A. *ChemBioChem* **2006**,7, 1497-1515.

[9] Sazani P., Gemignani F., Kang S. H., Maier M. A., Manoharan M., Persmark M.., Bortner D., Kole R. *NATURE BIOTECHNOLOGY* **2002**, 20 (12), 1228-1233.

[10] Faccini A., Tortori A., Tedeschi T., Sforza S., Tonelli R., Pession A., Corradini R., Marchelli R. *Chirality* **2008**, *20*, **494-500.**

[11] Rapireddy S., He G., Roy S., Armitage B. A., Ly D. H. *J. Am. Chem. Soc.* **2007**, 129, 15596-15600.

[12] Dragulescu-Andrasi A., Zhou P., He G., Ly D. H. *Chem. Commun.* **2005,** 244-246.

[13] a) Ishizuka T., Yoshida J., Yamamoto Y., Sumaoka J., Tedeschi T., **Corradini** R.**,** Sforza S., Komiyama M. *Nucleic Acid Res.* **2008***, 36,* 1464-1471; b) Ishizuka T., Tedeschi T., Corradini R., Komiyama M., Sforza S., Marchelli R. *ChemBioChem* **2009**, 10, 2607-2612; c) Chenna V., Rapireddy S., Sahu B., Ausin C., Pedroso E., Ly D. H. *ChemBioChem* **2008**, *9*, 2388-2391.

[14] Rothbard J. B., Kreider E.,VanDeusen C. L., Wright L., Wylie B. L., Wender P. A. *J. Med. Chem.* **2002**, 45, 3612-3618.

[15] Mitchell D. J., Kim D. T., Steinman L., Fathman C. G., Rothbard J. B. *J. Pept. Res.* **2000**, 56, 318-325.

[16] Fuchs S. M., Raines R. T. *Biochemistry* **2004**, 43, 2438-2444.

[17] Fischer R., Koehler K., Fotin-Mleczek M., Brock R. A. *J. Biol. Chem.* **2004**, 279, 12625-12635.

[18] Albertshofer K., Siwkowski A. M., Wancewicz E. V., Esau C. C., Watanabe T., Nishihara K. J., Kinberger G. A., Malik L., Eldrup A. B., Manoharan M., Geary R. S., Monia B. P., Swayze E. E., Griffey R. H., Bennett C. F., Maier M. A. *J. Med. Chem.* **2005**, 48, 6741-6749.

[19] Sahu B., Chenna V., Lathrop K. L., Thomas S.M., Zon G., Livak K. J., Ly D. H. *J.Org. Chem.,* **2009**, 74 (4), 1509-1516.

[20] Sforza S., Tedeschi T., Calabretta A., Corradini R., Camerin C., Tonelli R., Pession A., Marchelli R. *Eur. J. Org. Chem.* **2010**, 13, *2441-2444*.

[21] Tedeschi T., Sforza S., Dossena A., Corradini R., Marchelli R. *Chirality* **2005**, 17, S196-S204.

[22] Giannakakis A., Sandaltzopoulos R., Greshock J., Liang S., Huang J., Hasegawa K., Li C., O'Brien-Jenkins A., Katsaros D., Weber B. L., Simon C., Coukos G., Zhang L. *Cancer Biol. Ther.* **2008**, 7, 255-264.

[23] Bianchi N., Zuccato C., Lampronti I., Borgatti M., Gambari R. *BMB Rep.* **2009**, 42, 493-499.

[24] Faccini A., Tortori A., Tedeschi T., Sforza S., Tonelli R., Pession A., Corradini R., Marchelli R. *Chirality* **2008**, 20, 494-500.

[25] Dragulescu-Andrasi A., Rapireddy S., Frezza B. M., Gayathri C., Gil R. R., Ly D. H. *J. Am. Chem. Soc.* **2006**, 128, 10258-10267.

[26] Hyrup B., Nielsen P. E. *Bioorg. Med. Chem.* **1996**, 4, 5-23.

[27] Soliva R., Sherer E., Luque F. J., Laughton C. A., Orozco M. *J. Am. Chem. Soc.* **2000**, 122, 5997-6008.

[28] Demidov V. V., Potaman V. N., Frankkamenetskii M. D., Egholm M., Buchard O., Sonnichsen S. H., Nielsen P. E. *Biochemical Pharmacology* **1994**, 48 (6), 1310-1313.

[29] a) Zhou P., Dragulescu-Andrasi A., Bhattacharya B., O'Keefe H., Vatta P., Hyldig-Nielsen J. J., Ly D. H. *Bioorg. Med. Chem. Lett.* **2006**, 16, 4931-4935; b) Dragulescu-Andrasi A., Rapireddy S., He G., Bhattacharya B., Hyldig-Nielsen J. J. Zon G., Ly D. H. *J. Am. Chem. Soc.* **2006**, 128, 16104-16112.

[30] Huang X., Ding L, Bennewith K. L., Tong R. T., Welford S. M., Ang K. K., Story M., Le Q. T., Giaccia A. J. *Mol Cell.* **2009**, 35(6), 856-867.

[31] Gee H. E., Camps C., Buffa F. M., Patiar S., Winter S. C., Betts G., Homer J., Corbridge R., Cox G., West C. M., Ragoussis J., Harris A. L. *Cancer* **2010**, 116(9), 2148-2158.

[32] Greither T., Würl P., Grochola L., Bond G., Bache M., Kappler M., Lautenschläger C., Holzhausen H. J., Wach S., Eckert A. W., Taubert H. *Int J Cancer.* **2011**, doi: 10.1002/ijc.26109.

[33] Radojicic J., Zaravinos A., Vrekoussis T., Kafousi M., Spandidos D. A., Stathopoulos E. N. *Cell Cycle* **2011**, 10(3), 507-517.

[34] Corradini R., Manicardi A., Marchelli R., Sforza S., Tedeschi T., Fabbri E., Borgatti M., Bianchi N., Gambari R.. *Targets in Gene Therapy* – InTech – Open Access Publisher **2011**, ISBN 978-953-307-540-2, available from: http://www.intechopen.com/articles/show/title/gene-modulation-by-peptide-nucleic-acids-pnas-targeting-micrornas-mirs-

[35] Gambari R., Fabbri E., Borgatti M., Lampronti I., Finotti A., Brognara E., Bianchi N., Manicardi A., Marchelli R., Corradini R. *Biochem Pharmacol* **2011**, 82, 1416-1429.

[36] Fabbri E., Brognara E., Borgatti M., Lambprotti I., Finotti A., Bianchi N., Sforza S., Tedeschi T., Manicardi A., Marchelli R., Corradini R., Gambari R. *Ephigenomics* **2011**, 3 (6), 733-745.

[37] Fabbri E., Manicardi A., Tedeschi T., Sforza S., Bianchi N., Brognara E., Finotti A., Breveglieri G., Borgatti M., Corradini R., Marchelli R., Gambari R. *ChemMedChem* **2011**, 6 (12), 2192-2202.

Chapter 4 - Oligonucleotides mediated templated reaction

4.1 - Introduction

In biological systems all the processes that involve oligonucleotides (DNA or RNA) take advantage of their recognition abilities to promote reactions, in which the strand acts as a template to organize substrates in close proximity, increasing the effective molarity and promoting their reactions even at low concentration[1]. These properties can be exploited in nucleotide sensing technologies or in selective drug delivery systems. From the chemical point of view, there are a lot of different approaches that can be followed using oligonucleotide strands as template, most important of which are those related to chemical ligation, generation of fluorescence signals or release of functional molecules (Figure 4.1)[2].

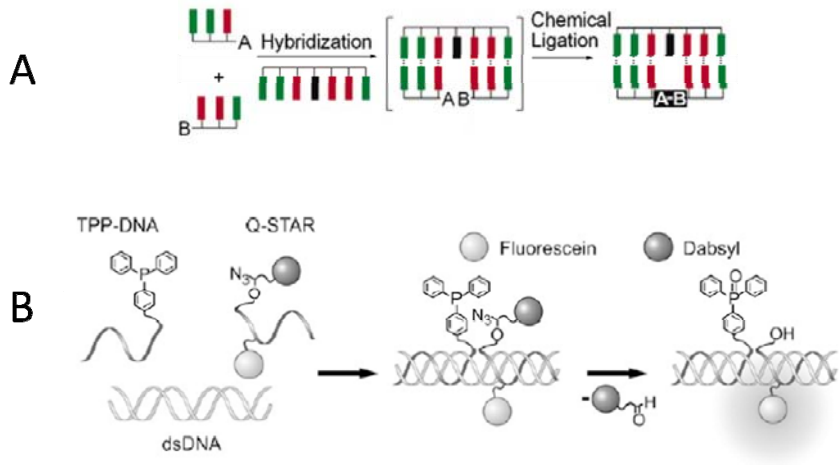


Figure 4.1: schematic representation of oligonucleotide templated reaction. A) chemical ligation ; B) generation of fluorescence signal combined with release of a functional molecule. (Ref. 8)

DNA-templated self-ligations are bond-forming reactions that require no added reagents to proceed. In such cases, the oligonucleotides being joined are modified during DNA synthesis to contain reactive ends that react when placed in close proximity, due to complexation with a complementary strand, and bring to the formation of a covalent bond between two different strands[3,4]. For the

generation of fluorescence signals upon hybridization two different strategies are reported: in the former case the reaction between the two reactive ends promotes the formation of a fluorescent molecule from a pro-fluorescent one[5,6], in the latter case the reaction between the two moieties leads to a cascade of reactions that end drifting apart a fluorophore from a quencher[7,8]. The last mechanism described is the same whereon the release of functional molecules is based[9,10].

An intriguing aspect that can be faced with the last two approaches, by opportunely tuning the recognition abilities, is the possibility to carry the reaction under catalytic conditions, which is often avoided in the case of ligation approaches, due to increased stability of the duplex formed between the templating strand and the product formed, as compared to the starting probes. The possibility to introduce a catalytic behavior in this kind of reaction leads to an accumulation of "signal" (fluorescence signal or concentration of released molecules), increasing the sensitivity of this techniques and opening the possibility of detecting low concentrated targets, such as oligonucleotides in living cells.

The discrimination ability between nucleotides is, therefore, the driving force of the templated reaction. Thus the sensitivity and selectivity issues are related to the complexation ability of the probes for the different strands. Based on this fact, the use of PNA instead of artificial DNA or other DNA analogs leads to a possible optimization of the reaction selectivity and sensitivity, related to their greater performance in oligonucleotide recognition. Furthermore, on account of the higher flexibility and stability of PNA, their synthesis and modifications, compared to other systems, the possibility to better tune the process arises, expanding the reaction pool and moving from *in vitro* to *in vivo* assays.

Moving from vials to biological systems, the enormous variety of molecules and systems that are present inside the "reaction flask" that can induce unwanted reaction must be taken in account. For this reason from the hypothetically large number of reactions that can be carried out under oligonucleotide templated reactions, only a few own the required bioorthogonality to be used in such complex systems. Examples of this kind of reactions are photoinduced reactions[11] or Staudinger reactions[12].

4.2 - DNA templated release of functional molecules

The ability to unleash the function of a small molecule in response to a signal is of broad interest not only for its practical application in probing biological systems and as "smart therapeutics", but

also from a more fundamental perspective in system chemistry. Strategies leveraged on the activity of particular enzymes such as proteases have been extensively studied for the conversion of prodrugs[13,14]. Alternatively, systems relying on external stimuli such as light (generally known as caged molecules) are also extensively used to unmask the function of a given molecule with spatial and temporal control[15]. Systems responding to a nucleic acid queues (DNA or RNA) have been reported using both molecular beacons (hairpin architectures) and nucleic acid templated reactions[16].

In a pioneering study, Taylor and Ma reported the first system[17] which relied on a DNA template to trigger the hydrolysis of an aryl esters catalyzed by imidazole[18]. A major focus on DNA/RNA templated reactions has been on nucleic acid sensing. While nucleophilic ligation reactions and Staudinger reactions have been used to release a quencher or a fluorophore[19,20], these approaches have not been used to release protein ligands or therapeutics. In parallel, molecular beacons bearing a photolabile group on one end and a quencher on the other have also been used to control the photo-release of prototypical small molecules in response to a given nucleic acid template in the presence of UV light[21,22]. Despite these notable advances, there remains a genuine need for nucleic acid templated reactions, which are broadly applied to unleash bioactive small molecules in a bio-orthogonal fashion.

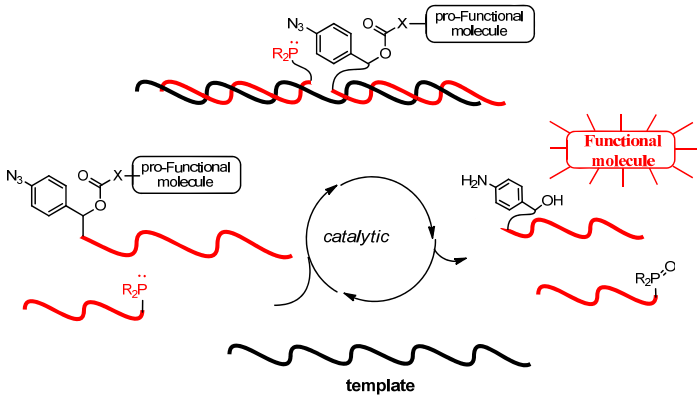


Figure 4.2: general mechanism of DNA-templated azide-triggered release of functional molecules.

In this chapter a novel and versatile design for the release of different functional molecules based on an azide-reduction triggered immolative linker is shown (Figure 4.2). Coupling the release of the functional molecule to an azide reduction was deemed most attractive by virtue of the bioorthogonality of the azide group and the well documented compatibility of the Staudinger reaction with cellular chemistry[23]. Indeed, it has been shown that nucleic acid templated reactions based on a Staudinger reaction are effective in live human cells and bacteria[20,24]. The immolative

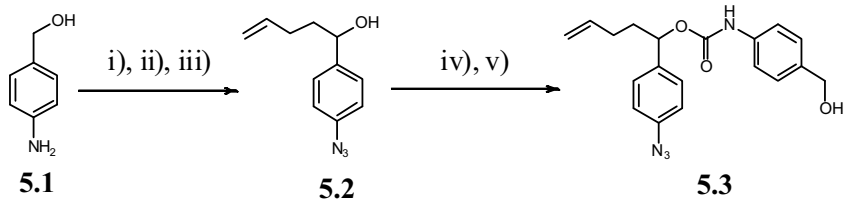
properties of the para-aminobenzyl moiety are well known[25] and the linker allows a broad variety of molecules to be coupled to the benzylic position via a carbonate or carbamate, which can in turn mask the function of the appended molecule, while providing a biologically stable connection.

4.2.1 - Design and synthesis of the probes

As a proof of principle, we selected a fluorophore (rhodamine) and a bioactive molecules (estradiol, a transcription factor activator and - see scheme 1 for structures). It is known that fluorescence of the rhodamine is quenched by conversion of both anilines to amides, ureas, carbamates or combination thereof[26,27]. In the other side, the phenolic moiety of the estradiol is required for protein binding and its function can be masked by derivatization of this phenol[28,29].

As shown in Figure 4.2, the immolative linker was envisioned to connect the different pro-functional molecules to a PNA strand. Given a second strand with a phosphine and a complementary template, the hybridization should increase the effective concentration of the reagents and promote the reduction of the azide, thus triggering the cleavage of the immolative linker and unleashing the functional molecule.

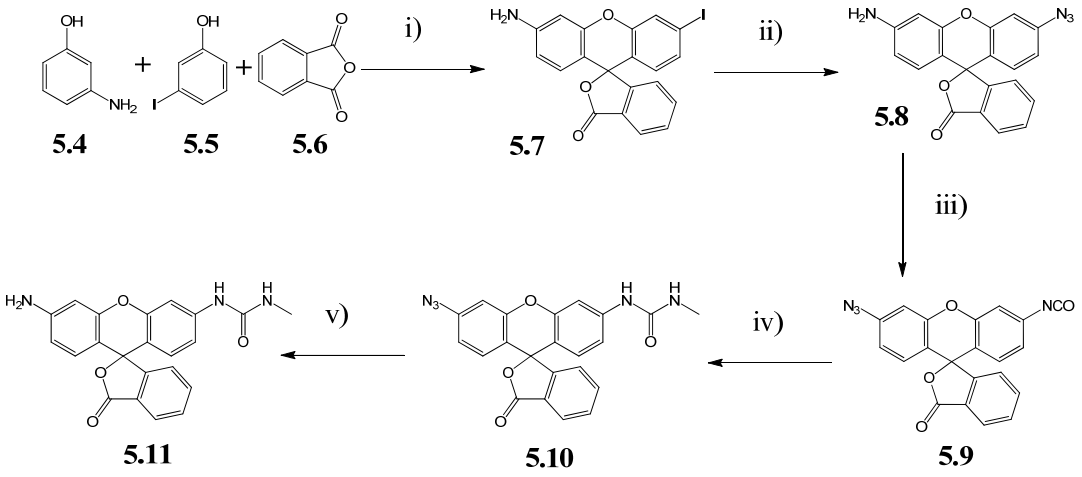
For this study PNAs bearing an arginine side chain at the C-2 position were chosen in order to confer cellular permeability to the system and possibly expand its application to the cellular compartment. For the connection between the PNA strand and the immolative linker the moiety was conjugated via an off-resin procedure, due to the impossibility of a number of interesting bioactive molecules to survive to PNA cleavage conditions. The connection between these two parts was realized via a chemoselective oxime formation between a PNA bearing a hydroxyl amine and an aldehyde on the immolative linker.



5.1 **5.2** **5.3**

Scheme 4.1: synthesis of azide-based immolative linker. i) $NaNO_2$, H_2SO_4, NaN_3 in water/THF, 90%; ii) PDC in DCM, 88%; iii) but-3-en-1-yl magnesium bromide in THF, 75%; iv) methyl 4-aminobenzoate, triphosgene, DIPEA in THF, 43%; DIBAL in DCM/toluene, 77%.

The synthesis of the linker started from 4-aminobenzyl alcohol (Scheme 4.1) which was converted to 4-azidobenzaldehyde according to a literature procedure[30] and treated with homoallyl magnesium bromide to obtain the secondary alcohol **5.2** in good yield. Efforts to conjugate different functional molecules directly to this secondary alcohol afforded poor yields, presumably due to steric congestion. This led us to add an immolative linker *via* a two step sequence (carbamate formation with methyl 4-isocyanatobenzoate followed by DIBAL reduction of the benzoate) to obtain the primary alcohol **5.3**.

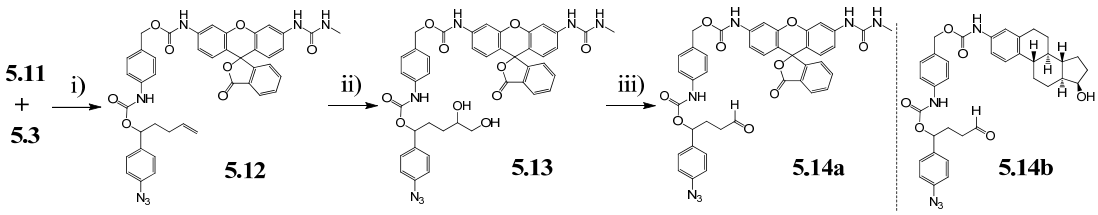


Scheme 4.2: synthesis of rhodamine moiety. i) methanesulfonic acid, 24%; ii) NaN_3, Sodium ascorbate, CuI, N,N'-dimethylethylendiamine in DMSO/water, 65%; iii) triphosgene, DIPEA in THF; iv) methylamine hydrochloride, DIPEA in THF, 76%; v) trimethylphosphine in THF/water.

For the rhodamine part, the synthesis (Scheme 4.2) starts from a large scale reaction with three inexpensive compounds such as phthalic anhydride , m-aminophenol and m-iodophenol, which condensate under strong acid condition to obtain **5.7** in moderate yield. The iodo group was then converted to an azide function in the presence of copper (I) in a good yield. The amine function was then converted to an ureidic group through a two step process involving the preliminary formation of isocianate **5.9** that is readily converted into **5.10** after the addition of the corresponding amine. The final amine **5.11** was obtained with a classical Staudinger reaction.

Compound **5.3** could now be coupled to the isocyanate of rhodamine methyl urea **5.11** (prepared *in situ* with triphosgene) followed by conversion of the terminal alkene to an aldehyde (OsO_4; $NaIO_4$) to obtain the rhodamine adduct **5.14a**. The overall process (Scheme 4.3) present a poor yield, due to the presence of a side product deriving from the formation of a carbamate between the rhodamine derivative and a propyl alcohol, that reduce the yield of the first step and induce the necessity of purification of the product after the conversion of the alkene to diol. To prepare the estradiol adduct

5.14b, the alcohol **5.3** was converted to an activated carbonate with para-nitrophenyl chloroformate and treated with estradiol followed by conversion of the terminal alkene to an aldehyde under similar conditions used for **5.14a**.



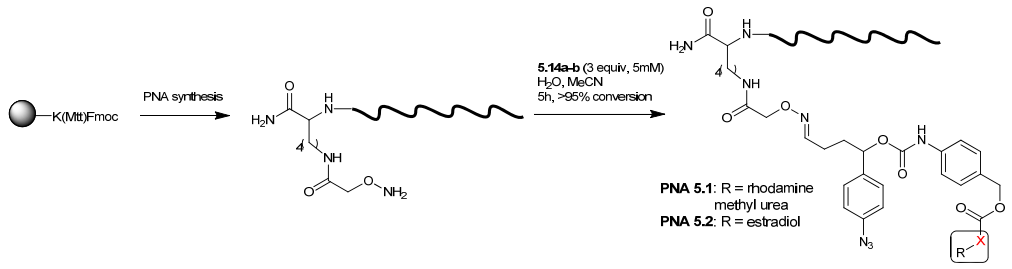
Scheme 4.3: conjugation of the immolative spacer to the rhodamine derivative. i) triphosgene, DIPEA in THF; ii) MMO, OsO_4 in acetone/water, 11%; iii) $NaIO_4$ in MeOH/water, 88%.

as the template to be interrogated with the PNA probes, the sequence was chosen corresponding to miR 21 (5'-TAG CTT ATC AGA CTG ATG TTG A-3'). From literature is know that 6-8mer PNA probes were suitable for DNA template reactions and that 2-4 nucleotide distance between the PNA probes was optimal[24,31]. With these criterions, we selected a 7mer sequence for the phosphine conjugate and 8mer for the immolative linker-conjugates using a 2D-Arg-PNA residue at every other position (Table 4.1).

Table 4.1: PNA sequences used for this study. Underlined letters are for arginine modified monomers.

Name	PNA sequence
AOOA-PNA	H-R-C<u>A</u>T<u>C</u>A<u>G</u>T<u>C</u>-K(AOOA)-NH_2
PNA-dmTCEP	dmTCEP-T<u>A</u>A<u>G</u>C<u>T</u>A-R-NH_2

We next investigate the coupling of the linker-conjugates **5.14a** and **5.14b** to PNA. PNAs were synthesized by standard Fmoc-based solid phase chemistry with Boc protected nucleobases[32] using a Fmoc-Lys(Mtt)-OH as a first residue. Prior to the cleavage of the last Fmoc, the Mtt group was cleaved (HFIP 50% in dichloroethane) and Boc-aminooxyacetic acid was introduced. Final Fmoc deprotection and cleavage from the resin (TFA) afforded the PNA for the chemoselective coupling with the desired aldehyde. Preliminary experiments for coupling the PNA to the adducts **5.14** in buffered conditions at pH 7 gave no results even in the presence of aniline as a catalyst[33]. However, efficient coupling was achieved when the PNAs were used directly as their TFA salts (obtained directly after cleavage or following an HPLC purification) without any buffer to obtain complete conversion into conjugates **PNA 5.1** and **PNA 5.2,** as verified by MALDI analysis of the reaction.



Scheme 4.4: synthesis of PNA probes bearing the two different immolative functions.

4.2.2 - Templated release

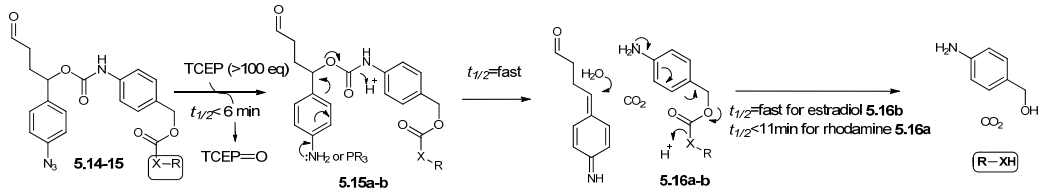


Figure 4.3: kinetic evaluation of functional molecule release upon azide reduction.

Before trying the template reaction in the presence of DNA, we assessed the kinetics of decomposition of the immolative linker-conjugates **PNA 5.1-5.2** upon reduction of the azide in physiologically relevant conditions. For this purpose, each product was dissolved in MeCN and diluted to a final concentration of 0.5 mM in PBS buffer at pH 7.4. The solutions were treated with a large excess of phosphine (tris(2-carboxyethyl)phosphine - TCEP, 50 equivalents) and the progress of the reaction was followed by LC-MS. Using a large excess of phosphine, the half life for the reduction of the azide was measured to be less than 6 minutes, based on the consumption of **5.14**. The intermediate **5.15** could not be observed suggesting that its immolation (or that of the aza-ylide intermediate) into compounds **5.16** is fast. The rate of reaction for the second self-immolative linker did vary depending on the nature of the leaving group. For estradiol, intermediate **5.16b** could not be observed suggesting that the reaction was fast. In the case of rhodamine (**5.16a**) the half life of the reaction was measured to be less than 11 minutes. These differences in reaction rate are consistent with the increased pKa of the leaving group. While there are notable differences in the rate of immolation, both reactions proceeded with kinetics that are compatible with the time frame of biological experiments.

To test the DNA templated reaction with PNA probes we first focused on rhodamine release with

conjugates **PNA 5.1**, as the reaction can be followed by continuous fluorescence monitoring. As shown in Figure 4.4A, in the absence of a template, the background reaction between the PNA is negligible. Using a large excess of TCEP, the fluorescence increases reaching a maximum within 30 min. The sigmoidal shape of the fluorescence increase is consistent with a two step cleavage of the immolative linker (Figure 4.3). In the presence of the perfect match DNA (PM), a fast reaction is observed reaching 30% conversion within 30 min, whereas a random DNA sequences (RD, 5'-CTG TGC GTG TGA CAG CGG CTG A-3') or a sequence bearing two mismatches sites (MM, 5'-TAC TTT ATC AGA GTG ATG TTG A-3') had significantly slower rate of reaction (<5% conversion after 30 min).

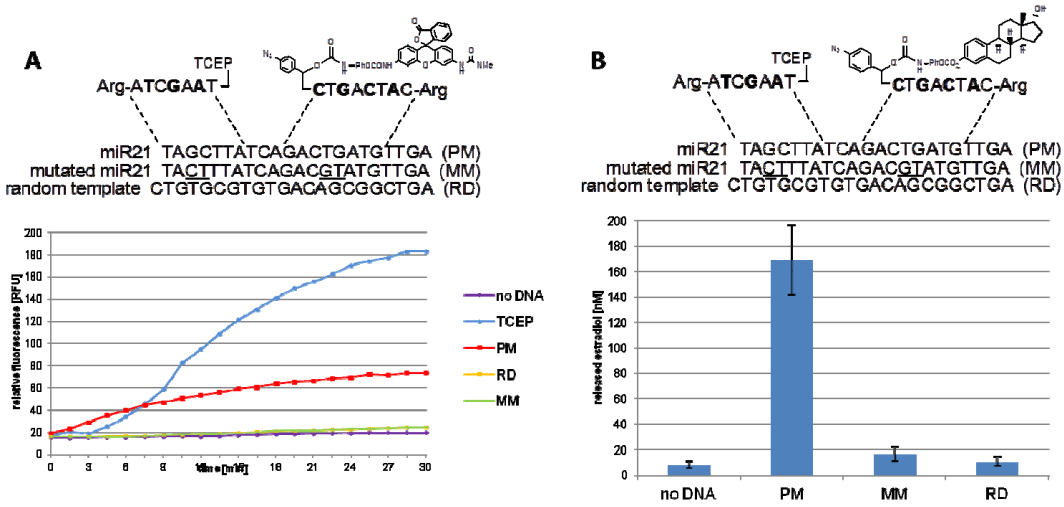


Figure 4.4: templated release of functional molecules. Conditions: 200nM immolative-PNA probe, 800nM dmTCEP-PNA probe, 100nM DNA, in PBS, pH 7.4 with 0.1% BGG at 37 °C. Bold letters denote modified monomer residues. PM: perfect match; MM: mismatched; RD: random. A) Methylureidorhodamine release; B) estradiol release.

Having established the viability of the present azide-triggered immolative linker in a templated reaction, we then turned our attention to the release of a bioactive small molecule. While estradiol have not a fluorescence efficiency which enables real time monitoring of the reaction, it is possible to quantify estradiol accurately by HPLC using a fluorescence detector at concentration down to low nM[34]. To this end, we quantified the amount of estradiol released after 30 minutes in the presence of the perfect match DNA template, a mismatched template, a random DNA template and in the absence of template. After 30 min, the reactions were quenched by addition of H_2O_2 (to a final concentration of 10%) to oxidize the phosphine probe; the estradiol was extracted and quantified by HPLC. As shown in Figure 4.4B, while the mismatched template afforded a concentration of estradiol below 20 nM, the perfect matched template afforded 180 nM concentration of estradiol, which is in excess relatively to the concentration of template (100 nM),

suggesting a turnover by the template. This represents an 85% yield with respect to the concentration of **PNA 5.2** within 30 min of reaction. The higher conversion obtained after 30 min with estradiol compared to rhodamine is consistent with the faster kinetics of linker decomposition.

4.2.3 - Conclusions

This is the first study in which a templated Staudinger reaction is used to reduce an azide reduction-triggered immolative spacer, based on p-aminobenzyl alcohol, to unleash a functional molecule (here demonstrated for a modified rhodamine as well as for estradiol). The viability of this DNA-templated reaction in combination with the mild conditions used for chemoselectively linking the immolative moiety and the possibility to easily modify the PNA strand, open the application of this concept to biomedical applications. It is in fact theoretically possible to link a broad variety of biologically active molecules, deliver them into cell masking their activity and release them only in presence of a specific target, suggesting the application of this system as a smart drug.

4.3 - miR templated detection in cells

The detection of nucleic sequences directly in cells can play an important role in different fields, from molecular biology to clinical diagnosis. For example, the possibility to detect specific sequences in live cells allows to specifically select cells that undergo to effective gene correction from the unrepaired ones. Direct detection of oligonucleotide sequences was already achieved with fluorescent in situ hybridization (FISH)[35], molecular beacons[36] or nucleic acid templated reactions[37,38]. Advantages of the last techniques compared to the others are the possibility of faster detection, due to the fact that hybridization or washing step before the reading are not required, the reduction of aspecific signals, due to the minor length of the probes, that avoids complexation of non-matching targets and the possibility to have signal amplification, related to the fact that the target sequence plays a catalytic role in the process.

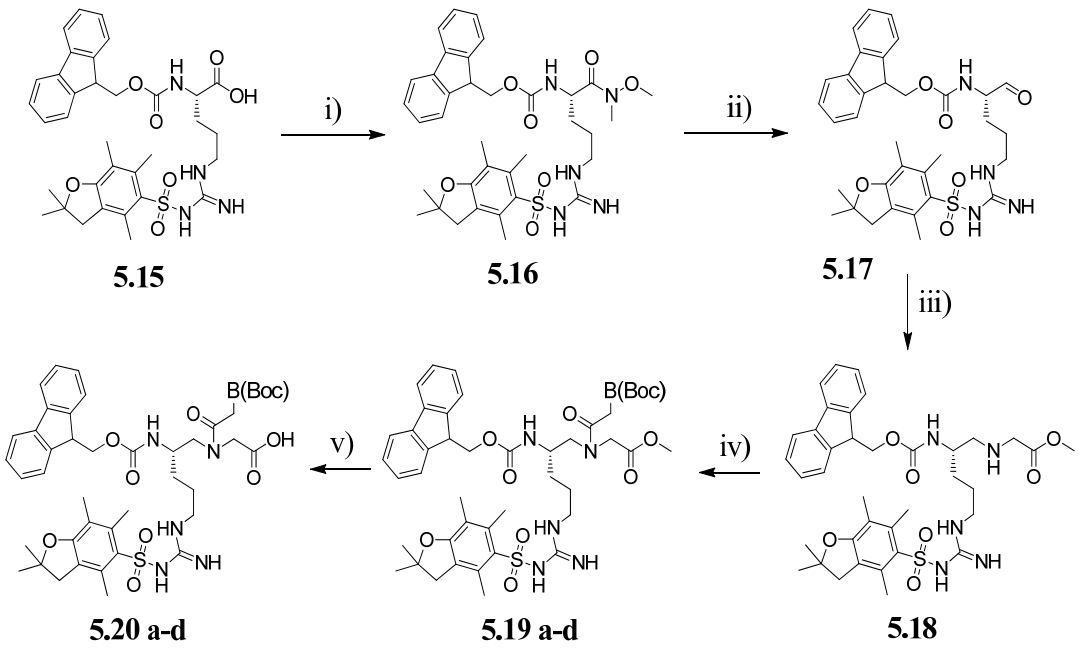
Here we propose the synthesis of a library of compounds targeting two different sequences (miR210 and miR21) and the synthesis of a set of 5L-Arginine modified PNA monomers for the use in the standard Fmoc-based solid phase synthesis.

4.3.1 - Design and synthesis

As previously described in Chapter 3, the introduction of arginine side chains embedded in the PNA strand strongly encourages the cellular internalization of PNA probes. Based on these results we designed a set of PNAs bearing arginine side chains placed with different motifs along the strand to evaluate their ability to enter inside the cell and to undergo to a templated reaction using a miR strand as a template.

The synthesis of all PNAs for this study was achieved by automated Fmoc-based solid phase synthesis. The Fmoc-2D-Arg monomers were previously synthesized for another study[24], whereas the 5L-modified were synthesized as showed for other C5-modified monomers (Chapter 2 and 3).

The synthesis of the Fmoc-based monomers (Scheme 4.5) started from the commercially available Fmoc-Arg(Pbf)-OH, that was reduced to the corresponding aldehyde **5.17** through preliminary conversion to the Weinreb amide **5.16**, to avoid the excessive reduction to the undesired alcohol.



Scheme 4.5: synthesis of the Fmoc-5L-Arg- PNA monomers 5.20 a-d: i) N,O-dimethylhydroxylamine hydrochloride, DIPEA, HBTU in DMF, 99%; ii) $LiAlH_4$ in THF, 49%; iii) glycine methyl ester hydrochloride, $NaBH_3CN$, AcOH in MeOH, 45%; iv) DIC, DhBtOH, CMA(Boc) [a]/ CMC(Boc) [b]/ CMG(Boc) [c]/CMT [d] in DMF, 81%[a]/ 79%[b]/ 79%[c]/ 85%[d]; v) $Ba(OH)_2 \cdot 8H_2O$ in THF/water 1:1, 90%[a]/ 97%[b]/ 95%[c]/ 85%[d].

The aldehyde was then converted into the backbone **5.18** by reductive amination, and on this secondary amine the four different bases were linked to obtain the fully protected monomers **5.19a-**

d, that were then converted to the corresponding acids by chemoselective hydrolysis in the presence of $Ba(OH)_2$ to obtain the desired PNA monomers **5.20a-d**. The four different monomers were synthesized with fairly good yields, except for the reaction involving the guanine as a nucleobase due to the poor solubility of compound **5.19c**. Differently from commercially available nucleobases that use Bhoc as protective group for the exocyclic amine, we chose to use Boc protected bases that were already demonstrated[32] to give better results in term of product purity and consequently of product yield due to less purification work needed.

A rhodamine derivative with two different azide functions was used, instead of a previously reported mono azide rhodamine[24a], as precursor of the fluorophore for its greater variation in fluorescence intensity upon reduction, in order to obtain better sensitivity as well as a higher signal to noise ratio.

Table 4.2: PNAs synthesized in this study. Single underlined letters correspond to phospine PNAs matching region, double underlined letters correspond to rhodamine PNAs. Bold letters are for modified monomers.

PNA	Sequence	PNA	Sequence	Info
miR 210	5'-CUG UGC GUG UGA CAG CGG CUG A-3'			
5.3	H-R-T**C**A**G**C**C**G**C**T**G**T**C**AC-K(Rh)-NH_2	5.8	TCEP-C**G**C**A**C**A**G-R-NH_2	2D-Arg / Mod 1
5.4	H-R-TC**A**GC**C**GC**T**GT**C**AC-K(Rh)-NH_2	5.9	TCEP-C**G**CA**C**AG-R-NH_2	2D-Arg / Mod 2
5.5	H-R-T**C**A**G**C**C**G**C**T**G**T**C**AC-K(Rh)-NH_2	5.10	TCEP-C**G**C**A**C**A**G-R-NH_2	5L-Arg / Mod 1
5.6	H-R-TC**A**GC**C**GC**T**GT**C**AC-K(Rh)-NH_2	5.11	TCEP-C**G**CA**C**AG-R-NH_2	5L-Arg / Mod 2
5.7	H-TCAGCCGCTGTCAC-K(Rh)KKKK-NH_2	5.12	TCEP-CGCACAG-KKKKK-NH_2	poly-K
miR 21	5'-UAG CUU AUC AGA CUG AUG UUG A-3'			
5.13	H-R-T**C**AA**C**A**T**C**A**G**T**CTG-K(Rh)-NH_2	5.18	TCEP-T**A**A**G**C**T**A-R-NH_2	2D-Arg / Mod 1
5.14	H-R-TC**A**AC**A**TC**A**GT**C**TG-K(Rh)-NH_2	5.19	TCEP-T**A**AG**C**TA-R-NH_2	2D-Arg / Mod 2
5.15	H-R-T**C**AA**C**A**T**C**A**G**T**CTG-K(Rh)-NH_2	5.20	TCEP-T**A**A**G**C**T**A-R-NH_2	5L-Arg / Mod 1
5.16	H-R-TC**A**AC**A**TC**A**GT**C**TG-K(Rh)-NH_2	5.21	TCEP-T**A**AG**C**TA-R-NH_2	5L-Arg / Mod 2
5.17	H-TCAACATCAGTCTG-K(Rh)KKKK-NH_2	5.22	TCEP-TAAGCTA-KKKKK-NH_2	poly-K

Thus, in order to obtain a strong increment in fluorescence intensity the two PNAs involved in the template reaction were designed in order to present different stabilities upon hybridization with the target sequence. In particular, the probes bearing the azide function must form a more stable complex than those presenting the reducing units, to allow the possibility to have double reduction on the templating strand before melting. In this context probes with different length were designed: a 7mer for reductive probes and 14mer for pro-fluorescent probes. Moreover, in order to identify the correct modification that can allow fast cellular uptake and good catalytic behavior three different models were synthesized: one with a poly-lysine at the C-term and two bearing both 2D- or 5L-modified monomers, in which the modifications are introduced every two or three units. All

sequences synthesized are show in The aldehyde was then converted into the backbone **5.18** by reductive amination, and on this secondary amine the four different bases were linked to obtain the fully protected monomers **5.19a-d**, that were then converted to the corresponding acids by chemoselective hydrolysis in the presence of $Ba(OH)_2$ to obtain the desired PNA monomers **5.20a-d**. The four different monomers were synthesized with fairly good yields, except for the reaction involving the guanine as a nucleobase due to the poor solubility of compound **5.19c**. Differently from commercially available nucleobases that use Bhoc as protective group for the exocyclic amine, we chose to use Boc protected bases that were already demonstrated[32] to give better results in term of product purity and consequently of product yield due to less purification work needed.

A rhodamine derivative with two different azide functions was used, instead of a previously reported mono azide rhodamine[24a], as precursor of the fluorophore for its greater variation in fluorescence intensity upon reduction, in order to obtain better sensitivity as well as a higher signal to noise ratio.

Table 4.2.

All PNA syntheses were carried out using automated Fmoc-based solid phase chemistry, using preformed modified monomers together with standard monomers, with the corresponding partial racemization of the C2 derived monomers. In all cases, the rhodamine derivative was introduced after the removal of the Mtt protective group (HFIP/DCE 1:1) using HCTU/DIPEA/lutidine as activating mixture. For all PNAs with modified backbone the monomers were introduced before polymer elongation, while for those bearing the poly-lysine chain the pro-fluorescent group was introduced after PNA synthesis, to avoid the preparation of a monomer with the modified lysine. All PNAs were subsequently purified by LPLC using an automated system.

4.3.2 - Preliminary cellular results

Before the evaluation of the ability of these PNA systems to undergo to template reaction in cells their ability to effectively permeate inside the cellular compartment were first evaluated. At this purpose, formaldehyde fixed MCF-7 cells were incubated with a 100nM solution of PNA for 3.5 hours, and after a washing step, the azide functions were reduced by addition of a 1mM solution of tmTCEP.

As shown in Figure 4.5 **PNA 5.16** exhibits better cellular uptake as compared to the other systems,

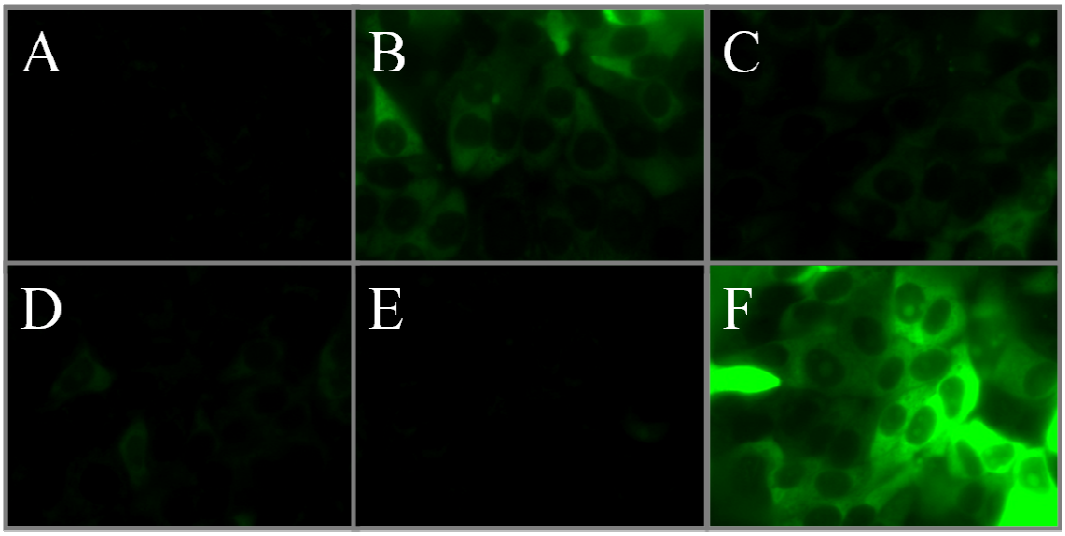


Figure 4.5: cellular uptake of miR21 targeting PNAs on MCF-7 cells: A) no PNA; B) PNA 5.13; C) PNA 5.14; D) PNA 5.15; E) PNA 5.16; F) PNA 5.17. Experiment were conducted on fixed cells, incubated with 100nM PNA for 3.5 hours, then reduced with 1mM solution of tmTCEP

whereas **PNA 5.13** shows the best cellular permeation among all the backbone-modified PNAs.These preliminary results are in contrast with those obtained in Chapter 3. This can be due to different cell lines and number of arginine residues, since it was proposed that the optimal number of guanidinium units for a good cellular permeation should be at least seven[39]. Comparison between the backbone-modified PNAs suggests that the cellular uptake of 2D-derivatives is higher than that of 5L, and that model 1 is superior to model 2, thus confirming the necessity to have a correct spacing between charged residues in order to achieve cellular internalization.

To exclude eventual differences due to different cell lines, the cellular uptake of PNAs targeting miR210 were tested on K562 cells. Experiments are still in progress.

4.3.3 - Conclusions

Based on the results obtained from previous studies we have synthesized a set of PNA probes, bearing both 2D- or 5L-arginine side chain, that are hypothetically able to permeate inside cells and selectively detect the presence of a miR sequence.

To this aim a new set of monomers for Fmoc-based solid phase synthesis, bearing 5L-Arginine backbone modifications was synthesized, and five different model probes were obtained for the targeting of two different miR sequences.

Preliminary results in cellular permeation showed a better internalization for the PNA bearing a poly-lysine chain, and showed interesting differences in internalization of backbone modified PNAs as a function of substituent position. Further investigation about cellular uptake and template reactions in cell are in progress.

4.4 - Experimental Section

General. Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, Acros and used without further purification. All reactions were carried out under a nitrogen atmosphere with dry solvents under anhydrous conditions, unless otherwise noted. Anhydrous solvents were obtained by passing them through a commercially available alumina column (Innovative technology, MA). DNA were purchased from Eurogentec as desalted 100 μM solutions in deionized water. Reactions were monitored by TLC carried out on 0.25 mm E. Merck silica-gel plates (60F-254) by using UV light as visualizing agent and ninhydrin solution and heat as developing agents. E. Merck silica gel (60, particle size 0.040–0.063 mm) was used for flash-column chromatography. PNA were prepared with standard Fmoc-based automatic peptide chemistry using MultiPep RS (Intavis) synthesizer. NMR spectra were recorded on Bruker Avance 400 or Bruker AC 300 instruments and calibrated by using residual undeuterated solvent as an internal reference. The following abbreviations were used to explain the multiplicities: s=singlet, d=doublet, t=triplet, q=quartet, m=multiplet, and br=broad. IR spectra were measured using a FT-IR Thermo Nicolet 5700, in transmission mode using KBr or NaCl or using a FT-IR Nicolet Magna 200 spectrophotometer, in reflective mode (ATR) directly from solid or liquid samples. LCMS were recorded by using an Agilent 1100 HPLC with a Surveyor MSQ spectrometer equipped with a Supelco C18 (2.1x50 mm, 5 μm particles) column (method A, linear elution gradient from 95% H_2O 0.01% TFA to 100% MeCN 0.01% TFA in 8 minutes at a flow rate of 0,7 ml/min) or by using a Thermo Scientific Accela HPLC with a Surveyor MSQ Plus spectrometer equipped with a Thermo C18 (2.1x50 mm, 1.9 μm particles) Hypersil gold column (method B, linear elution gradient for 95% H_2O 0.01% TFA to 90% MeCN 0.01% TFA in 3.6 minutes at a flow rate of 1.0 ml/min; method C, linear elution gradient for 95% H_2O 0.01% TFA to 90% MeCN 0.01% TFA in 2.2 minutes at a flow rate of 1.0 ml/min). If mass spectrum was not obtained in LCMS conditions compounds were analysed by direct infusion into LCQ Fleet three-dimensional ion trap mass spectrometer (Thermo Scientific) or an ACQ-SQD ESI-Q mass spectrometer (Waters). The MALDI spectra were measured using Bruker Daltonics Autoflex TOF/TOF spectrometer. PNA oligomers were purified by reverse-phase chromatography using a Biotage Isolera ONE equipped with a Biotage SNAP Cartridge KP-C18-HS 12g (linear gradient from 100% H_2O 0.01% TFA to 30% MeCN 0.01% TFA in 22.5 minutes with a flow rate of 10 ml/min) or using an Agilent 1100 series HPLC equipped with DAD and with a Agilent ZORBAX Eclipse XDB-C18 (4.6x250 mm, 5μm) column (linear gradient from 100% H_2O 0.1% TFA to 50% MeCN 0.1% TFA in 30 minutes with a flow rate of 1 ml/min).

4-azidobenzyl alcohol. Solution of 4-aminobenzyl alcohol (**5.1**, 2.15g, 17.5mmol) in 60mL of water and 25mL of THF was cooled to 0°C and 4.8ml of concentrated sulfuric acid was added slowly. Then solution of sodium nitrite (1.45g, 21.0mmol) in 10ml of water was added drop-wise. After next 1h at 0°C, solution of sodium azide (1.82g, 28.9mmol) in 5ml of water was added drop-wise and the reaction mixture was stirred overnight while warming to room temperature. The reaction was quenched with brine and extracted with DCM (3x100ml). Combined organic layers were dried over Na_2SO_4 and concentrated under reduced pressure. Crude product was purified by flash chromatography (hexanes:EtOAc 4:1) to give the desired product as yellow oil (2.33g, 90%); **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 7.33 (2H, d, J= 8.4Hz), 7.00 (2H,

d, J= 8.4Hz), 4.64 (2H, s); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 139.6, 137.8, 128.8, 119.4, 64.9; **LCMS (ESI):** t_r = 4.05 min (Method A) *m/z* calcd for $C_7H_7N_3O$ [M]: 149.05891, found: 122.1 $[M+H-N2]^+$; **IR ν(cm^{-1}):** 3539.5, 2113.1, 1610.6, 1508.4, 1417.7, 1291.4, 1204.6, 1181.4, 1129.4, 1042.6, 1013.6, 822.7.

4-azidobenzyl aldehyde. To solution of alcohol 4-azidobenzyl alcohol (2.33g, 15.6mmol) in DCM (200ml) was added pyridinium dichromate (7.05g, 18.7mmol) and the reaction was stirred for 4h at room temperature. After the reaction mixture was filtered through 130 g of silica and silica was washed with additional 1.5l of DCM. Evaporation of solvent under reduced pressure and drying under high vacuum yielded the desired product as pale yellow oil (2.03g, 88%); **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 9.94 (1H, s), 7.88 (2H, d, J= 8.4Hz), 7.16 (2H, d, J= 8.4Hz); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 190.5, 146.2, 133.2, 131.5, 119.4; **LCMS (ESI):** t_r = 3.31 min (Method A) *m/z* calcd for $C_7H_4N_3O$ [M]: 147.04326, found: 147.9 $[M+H]^+$; **IR ν(cm-1):** 2116.0, 1693.6, 1598.1, 1502.6, 1422.6, 1391.7, 1285.6, 1212.3, 1667.0, 1126.5, 826.5, 787.0.

1-(4-azidophenyl)pent-4-en-1-ol (5.2). Magnesium turnings (0.48g, 19.8mmol) were suspended in THF (22ml) in a flask equipped with vacuum condenser and 4-bromo-1-butene (1.6ml, 16.5mmol) was added drop-wise. The mixture was stirred for 12 h at room temperature. The Grignard reagent was next added dropwise (20ml/h) to the solution of aldehyde 4-azidobenzyl aldehyde (2.03g, 13.8mmol) in THF (50ml) at -78°C. When addition was finished, the reaction mixture was stirred next 2h at -78°C, after quenched with saturated ammonium chloride solution and extracted with DCM (3x100ml). Combined organic layers were washed with saturated ammonium chloride and brine, dried over anhydrous sodium sulfate and concentrated in vacuum. Purification of residue by flash chromatography (hexanes:EtOAc 10:1) gave **5.2** as yellow solid (2.1g, 75%); **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 7.31 (2H, d, J= 8.6Hz), 6.99 (2H, d, J= 8.1Hz), 5.86÷5.79 (1H, m), 5.06÷4.98 (1H, m), 4.67÷4.62 (2H, m), 2.47 (1H, s), 2.13÷2.08 (2H, m), 1.89÷1.59 (2H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 141.4, 139.0, 137.9, 127.3, 118.9, 115.0, 73.2, 38.0, 29.9; **LCMS (ESI):** t_r = 3.66 min (Method A) *m/z* calcd for $C_{11}H_{14}N_3O$ [M]: 203.10586, found: 176.1 $[M+H-N2]^+$, 204.1 $[M+H]^+$; **IR ν(cm-1):** 3730.5, 2937.7, 2121.8, 1741.8, 1699.3, 1645.3, 1605.8, 1507.4, 1290.4, 1128.4, 1067.6, 1010.7, 918.2, 838.1.

Methyl 4-((((1-(4-azidophenyl)pent-4-en-1-yl)oxy)carbonyl)amino)benzoate. 4-aminobenzoic acid methyl ester (93mg, 0.61mmol) and triphosgene (93mg, 0.31mmol) were dissolved in THF (5ml) and DIPEA (310µl, 1.84mmol) was added drop-wise. The reaction mixture was stirred 40 minutes at room temperature. Then solution of **5.2** (500mg, 2.46mmol) in THF (2ml) was added and the reaction was stirred 18h at 60°C. After the reaction mixture was cooled to ambient temperature, quenched with saturated ammonium chloride solution and extracted with DCM (3x10ml). Combined organic layers were washed with saturated ammonium chloride, dried over sodium sulfate and concentrated under reduced pressure. Flash chromatography (hexanes:EtOAc 5:1) yielded the desired product as yellow oil (101.2mg, 43%); **^{1}H NMR (acetone-d^6, 400MHz) δ(ppm):** 7.96 (2H, d, J= 8.6Hz), 7.46 (2H, d, J= 9.1Hz), 7.32 (2H, d, J= 8.6Hz), 6.98 (2H, d, J= 8.0Hz), 5.83÷5.76 (1H, m), 5.72÷5.70 (1H, m), 5.04÷4.99 (2H, m), 3.91 (3H, s), 2.09÷2.03 (3H, m), 1.89÷1.84 (1H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 166.7, 152.4, 142.2, 139.8, 137.0, 136.9, 130.8, 128.0, 124.6, 119.1, 117.5, 115.5, 76.2, 51.9, 35.2, 29.6; **LCMS:** t_r = 3.55 min (Method A); **HRMS (MALDI-TOF):** could not be measured due to molecule fragmentation; **IR ν(cm-1):** 3625.3, 2122.7, 1723.5, 1708.0, 1652.1, 1646.3, 1549.9, 1542.1, 1515.1, 1364.7, 1289.5, 1210.4.

Immolative spacer (5.3). Solution of methyl 4-((((1-(4-azidophenyl)pent-4-en-1-yl)oxy)carbonyl)amino) benzoate (101.2mg, 0.266mmol) in DCM was cooled to -78°C and a 1.0M solution of DIBAL in toluene (910µl, 0.91mmol) was added drop-wise over 1h. After next 1h at -78°C supplementary portion of DIBAL (266µl, 0.266mmol) was added drop-wise over 20 minutes to complete the reaction. The reaction mixture was stirred for 40 minutes more, quenched with ice-cold 20% NaOH solution (10ml) when still at -78°C and after allowed to warmed up to ambient temperature. Then layers where separated, aqueous phase was

extracted with DCM (4x10ml), combined organic layers were dried over anhydrous sodium sulfate and concentrated under reduced pressure. Residue was redissolved in DCM and evaporated with silica. Flash chromatography (hexanes:EtOAc from 8:1 to 3:1) yielded **5.3** as pale yellow oil (69.2mg, 77% yield); **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 7.33 (4H, d, J= 8.0Hz), 7.23 (2H, d, J= 8.6Hz), 7.09 (1H, br s), 6.98 (2H, d, J= 8.6Hz), 5.84÷5.74 (1H, m), 5.71÷5.68 (1H, m), 5.03÷4.97 (2H, m), 4.58 (2H, s), 3.85 (1H, s), 2.11÷1.99 (3H, m), 1.88÷1.83 (1H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 152.8, 139.6, 137.3, 137.2, 137.1, 136.0, 128.0, 127.7 (x2), 119.0, 115.3, 75.8, 64.6, 35.3, 29.6; **LCMS:** t_r = 4.19 min (Method A); **HRMS (MALDI-TOF):** could not be measured due to molecule fragmentation **IR ν(cm-1):** 3730.5, 2123.7, 1731.2, 1652.1, 1605.8, 1533.5, 1364.7, 1291.4, 1221.0, 1046.2, 920.1, 835.2.

Amino-iodo-rhodamine (5.7). 3-Aminophenol (2.00g, 18.32mmol), 3-iodophenol (4.03g, 18.32mmol) and phtalic anhydride (2.71g, 18.32mmol) were dissolved in 16ml methanosulfonic acid at room temperature. The reaction mixture was stirred for 40 h at 140°C. After cooling to r.t. the reaction mixture was poured onto 400g of ice, the resulting solution was neutralized with 10% NaOH and extracted with DCM. The combined organic fractions were washed with 10% NaOH, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. Flash chromatography (hexane:AcOEt 5:1, then 3:1) yielded **5.7** as orange solid (1.95g, 24%); **^{1}H NMR (DMSO-d^{6}, 400MHz) δ(ppm):** 8.00 (1H, d, J= 7.6Hz), 7.80 (1H, t, J= 7.2Hz), 7.78 (1H, d, J= 1.6Hz), 7.48 (1H, t, J= 7.6Hz), 7.43 (1H, dd, J= 8.0, 1.6Hz), 7.31 (1H, d, J= 7.6Hz), 6.51 (1H, d, J= 8.4Hz), 6.46 (1H, d, J= 2.0Hz), 6.40÷6.36 (2H, m), 5.73 (2H, br s); **^{13}C NMR (DMSO-d^{6}, 100MHz) δ(ppm):** 168.6, 152.2, 151.6, 151.5, 151.3, 135.7, 132.5, 130.2, 129.6, 128.5, 126.0, 125.4, 124.7, 124.0, 119.1, 111.6, 104.7, 99.1, 96.2, 82.7; **LCMS (ESI, MeOH):** t_r = 4.29 min (Method A) *m/z* calcd for $C_{20}H_{12}INO_3$ [M]: 440.98169, found: 442.39 $[M+H]^+$; **HRMS (MALDI-TOF):** *m/z* found: 441.9979 $[M+H]^+$, 463.9757 $[M+Na]^+$; **IR ν(cm^{-1}):** 3457.3, 3371.6, 2915.2, 2850.8, 1761.5, 1633.0, 1594.1, 1448.5, 1403.5, 1285.2, 1255.8, 1228.7, 1107.7, 1083.0.

Amino-azido-rhodamine (5.8). 5.7 (140mg, 0.317mmol), sodium azide (42mg, 0.634mmol) and sodium ascorbate (3.2mg, 0.016mmol) were dissolved in 2ml/0,4ml DMSO/H_2O mixture at room temperature and degassed under vacuum. Copper iodide (11.5mg, 0.064mmol) and N, N'- dimethylethylenediamine (5.5 µL, 0.048mmol) were added and reaction mixture was stirred until complete conversion of the starting material as judged by LCMS analysis (8h). The reaction was dilute with brine and extracted with AcOEt. The combined organic fractions were washed with brine, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. Flash chromatography (from hexane:AcOEt 5:1 to 1:1) afforded **5.8** as red solid (74.0mg, 65%); **^{1}H NMR (MeOD-d^{4}, 400MHz) δ(ppm):** 8.07 (1H, d, J= 7.9Hz), 7.84 (1H, t, J= 7.4Hz), 7.77 (1H, t, J= 7.4Hz), 7.28 (1H, d, J= 7.6Hz), 7.05 (1H, br s), 6.85 (2H, br s), 6.63 (1H, br s), 6.53÷6.48 (2H, m); **^{13}C NMR (acetone-d^{6}) δ(ppm):** 169.4, 153.8, 153.3, 153.1, 152.1, 143.1, 136.0, 130.7, 130.6, 129.6, 127.8, 125.4, 124.9, 117.5, 115.5, 112.6, 107.8, 107.4, 100.8, 83.5; **LCMS (ESI):** t_r= 3.95min (Method A) *m/z* calcd for $C_{20}H_{12}N_4O_3$ [M]: 356.09094, found: 357.49 $[M+H]^+$, 329.37 $[M+H-N2]^+$; **HRMS (MALDI-TOF):** *m/z* found: 379.1012 $[M+Na]^+$; **IR ν(cm^{-1}):** 3465.8, 3363.2, 2973.9, 2866.9, 2115.8, 1763.0, 1634.5, 1610.7, 1513.3, 1497.6,1452.9, 1427.7, 1284.6, 1255.1, 1225.3, 1106.9, 1066.1.

Azido-isocyanate-rhodamine (5.9). 5.8 (74mg, 0.208mmol) and triphosgene (31.3mg, 0.104mmol) were dissolved in 5ml THF at room temperature and DIPEA (105µl, 0.624mmol) was added dropwise. The reaction mixture was stirred for 1h after which, the starting material was cleanly converted to the isocyanate (conversion was monitored by quenching an aliquot of the reaction with EtOH and analyzing the product by LCMS). The solution of **5.9** in THF was used immediately for subsequent reaction without further purification. **LCMS (ESI):** for an aliquot quenched with ethanol t_r = 4.97 min (Method A) *m/z* calcd for $C_{23}H_{16}N_4O_5$ [M+EtOH]: 428.11207, found: 429.56 $[M+EtOH+H]^+$.

Azido-methylureido-rhodamine (5.10). To the **5.9** solution (0.047mmol in 1.5ml THF), methylamine hydrochloride (9.5mg, 0.141mmol) and DIPEA (23.3 µL, 0.141mmol) were added and resulting mixture was

stirred for 17h. After reaction quenching with brine the aqueous phase was extracted with DCM (3 times) and the combined organic fractions were dried over Na_2SO_4 and concentrated under reduced pressure. Flash chromatography (from hexane:AcOEt 2:1 to AcOEt) yielded **5.10** as orange solid (14.9mg, 76%);**^{1}H NMR (MeOD-d^4, 400MHz) δ(ppm):** 8.08 (1H, d, J= 7.6Hz), 7.81 (1H, t, J= 7.6Hz), 7.77 (1H, t, J= 7.2Hz), 7.66 (1H, s), 7.25 (1H, d, J= 7.6Hz), 7.06 (1H, s), 7.01 (1H, d, J= 8,4Hz), 6.86 (2H, s), 6.60 (1H, d, J= 8.0Hz), 2.84 (3H, s); **^{13}C NMR (MeOD-d^4, 400MHz) δ(ppm):** 171.2, 158.3, 154.3, 153.6, 152.8, 144.2, 143.9, 136.7, 131.3, 130.7, 129.2, 127.6, 125.9, 125.1, 117.0, 116.0, 115.8, 112.9, 108.2, 106.7, 84.2, 26.8; **LCMS (ESI):** t_r = 3.87 min (Method A) *m/z* calcd for $C_{22}H_{15}N_5O_4$ [M]: 413.11240, found: 414.56 $[M+H]^+$, 386.46 $[M+H-N_2]^+$; **HRMS (MALDI-TOF):** *m/z* found: 414.1179 $[M+H]^+$; **IR ν(cm^{-1}):** 2119.8, 1764.9, 1741.8, 1701.3, 1614.5, 1538.3, 1501.6, 1411.0, 1284.6, 1220.0, 1105.3, 891.1.

Amino-methylureido-rhodamine (5.11). 5.10 (14.9mg, 0.036mmol) dissolved in 3ml THF and a 1.0M trimethylphospine in THF (72µl, 0.072mmol) was added. After 1h of stirring at room temperature, water (0.5 ml) was added and the reaction mixture was stirred next 30 minutes. Then brine and DCM were added, transferred in a separatory funnel, aqueous phase was extracted with DCM (3x10ml), the combined organic layers were dried over Na_2SO_4 and concentrated under reduced pressure to yield **5.11** as red solid (20.0mg crude) which was used for next step without further purification; **^{1}H NMR (acetone-d^6) δ(ppm):** 7.94 (1H, d, J= 7.2Hz), 7.79 (1H, s), 7.77 (1H, t, J= 8.4Hz), 7.70 (1H, t, J= 7.6Hz), 7.25 (1H, d, J= 7.6Hz), 6.92 (1H, dd, J 1= 8.4Hz, 2Hz), 6.60÷6.55 (2H, m), 6.52÷6.40 (2H, m), 5.90 (1H, br s), 5.15 (1H, br s), 2.73 (3H, s); **^{13}C NMR (acetone-d^6) δ(ppm):** 171.7, 156.4, 154.1, 152.9, 151.8, 143.9, 135.9, 131.9, 130.4, 129.5, 128.8, 128.1, 125.1, 124.8, 117.9, 114.3, 112.4, 112.0, 105.5, 100.8, 85.2, 26.8; **LCMS (ESI):** t_r = 2.90 min (Method A) *m/z* calcd for $C_{22}H_{17}N_3O_4$ [M]: 387.12191, found: 388.40 $[M+H]^+$; **HRMS (MALDI-TOF):** *m/z* found: 388.1308 $[M+H]^+$; **IR ν(cm^{-1}):** 2928.0, 2855.7, 1692.6, 1601.9, 1552.8, 1500.7, 1458.2, 1411.9, 1336.7, 1289.5, 1177.6, 1136.1.

Immolative-Rhodamine-alkene (5.12). 5.11 (11.5mg, 29.3 µmol) and triphosgene (4.4mg, 14.6 µmol) were dissolved in 1ml THF and DIPEA (15 µL, 87.9mmol) was added drop-wise. The reaction mixture was stirred 30 minutes at r.t., then solution of **5.3** (10.5mg, 3.12mmol, 1.05 equiv.) in THF (1 ml) was added and the reaction was stirred 20h at 60°C. The reaction mixture was quenched with saturated ammonium chloride solution, after cooling back to r.t., and extracted with DCM (3x10ml). Combined organic layers were washed with saturated ammonium chloride, dried over sodium sulfate and concentrated under reduced pressure. The resulting residue was purified by flash chromatography (from hexane:AcOEt 3:1 to AcOEt) which gave 5mg of mixture of **5.12** with a side product (propyl carbamate). That mixture was used without further purification for next step; **LCMS (ESI):** t_r=2.13 min (Method B) *m/z* calcd for $C_{42}H_{35}N_7O_8$ [M]: 765.25471, found: 766.04 $[M+H]^+$, t_r= 1.75 min *m/z* found: 474.06 [side product+H]$^+$; **HRMS (MALDI-TOF):** *m/z* found: 788.2470 $[M+Na]^+$; **IR ν(cm^{-1}):** 2957.9, 2928.0, 2853.8, 2118.9, 1692.6, 1602.9, 1546.0, 1508.4, 1412.9, 1287.5, 1231.6, 1110.7.

Immolative-Rhodamine-diol (5.13). The mixture containing **5.12** (5mg, max. 6.5 µmol) was dissolved in 1ml acetone. N-methyl morpholine oxide monohydrate (3mg, 19.6 µmol), osmium tetroxide (4% solution in water, 2.1 µL, 0.33 µmol) were solubilized in 0.5ml H_2O and added to the solution containing the alkene. The resulting mixture was stirred at r.t. for 44h. The reaction mixture was then cooled to 0°C, quenched with 2ml saturated $Na_2S_2O_5$and stirred for next 30 minutes. After the reaction mixture was extracted with DCM (4x5ml), the combined organic fractions washed with brine, dried over Na_2SO_4 and concentrated under reduced pressure. Flash chromatography (from hexane:AcOEt 1:1 to 9:1 AcOEt:MeOH) yielded **5.13** as orange oil (2.5mg, 11% for two steps); **^{1}H NMR (MeOD-d^4) δ(ppm):** 8.08 (1H, d, J= 7.2Hz), 7.84 (1H, t, J= 7.2Hz), 7.78 (1H, t, J= 7.2Hz), 7.66 (2H, d, J= 2.0Hz), 7.51÷7.48 (4H, m), 7.41 (2H, d, J= 8.4), 7.28 (1H, d, J= 7.6Hz); 7.14÷7.12 (3H, m), 7.02 (1H, dd, J= 8.4Hz, 2.0Hz); 6.73-6.69 (2H, m), 5.80÷5.74 (1H, m), 5.20 (2H, s), 3.68÷3.65 (2H, m), 3.51÷3.49 (1H, m), 2.85 (3H, s), 1.78÷1.62 (2H, m), 1.56÷1.49 (2H, m); **^{13}C NMR (acetone-d^6) δ(ppm):** 169.3, 154.2, 154.0, 153.8, 152.9, 152.7, 152.6, 142.4, 140.2, 140.1, 139.5,

136.1, 131.5, 130.7, 130.5, 130.0, 129.4, 129.0, 128.9, 127.6, 125.4, 124.9, 119.8, 115.0, 114.8, 114.3, 106.3, 105.6, 83.2, 74.8, 72.3, 70.0, 67.0, 27.8, 26.7, 26.3; **LCMS (ESI):** t_r = 3.89 min (Method A) *m/z* calcd for $C_{42}H_{37}N_7O_{10}$ [M]: 799.26019, found: 800.80 $[M+H]^+$; **HRMS (MALDI-TOF):** *m/z* found: 800.2743 $[M+H]^+$; **IR ν(cm^{-1}):** 3350.5, 2968.6, 2928.0, 2858.6, 2117.9, 1694.5, 1602.9, 1546.0, 1412.9, 1287.5, 1233.5, 1049.3.

Immolative-Rhodamine-aldehyde (5.14a). To a solution of **5.13** (13.2mg, 16.5 µmol) in 3ml MeOH was added a solution of $NaIO_4$ (70.6mg, 330.0 µmol) in 1ml water. After 30 minutes the organic solvent was evaporated under reduced pressure and the resulting solution was diluted with 5ml brine and 5ml water, then extracted with DCM (5x10 ml) dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to yield **5.14** (11.2mg, 88%) as a pink/red solid; **^{1}H NMR (acetone-d^6) δ(ppm):** 9.72 (1H, d), 9.02 (1H, s), 8.90 (1H, s), 8.25 (1H, s), 7.78 (1H, d, J= 7.6Hz), 7.81 (1H, d, J= 2.0Hz), 7.78 (1H, dd, J= 7.6Hz, 0.8Hz), 7.74÷7.72 (2H, m), 7.54 (2H, d, J= 8.4Hz), 7.45 (2H, d, J= 8.4Hz), 7.36 (2H, d, J= 8.4Hz), 7.29 (1H, d, J= 7.6Hz), 7.19 (1H, dd, J= 8.8Hz, 2.0Hz), 7.09 (2H, d, J= 8.4Hz), 7.00 (1H, dd, J= 8.4Hz, 2.0Hz), 6.95 (1H, s), 6.74 (2H, d, J= 8.8Hz), 6.65 (2H, d, J= 8.4Hz), 5.78÷5.75 (1H, m), 5.12 (2H, s), 2.74 (3H, s), 2.55 (2H, t, J= 6.8Hz), 2.32÷2.25 (1H, m), 2.19-2.13 (1H, m); **^{13}C NMR (acetone-d^6) δ(ppm):** 201.9, 169.5, 154.2, 154.0, 153.6, 152.7, 152.6, 151.6, 144.0, 142.4, 140.4, 138.7, 136.1, 131.7, 130.7, 130.0, 129.7, 129.3, 129.0, 128.8, 127.6, 125.4, 124.9, 119.9, 115.0, 114.9, 114.3, 112.6, 106.3, 105.7, 83.2, 75.6, 67.0, 40.4, 29.7, 26.7; **LCMS (ESI):** t_r = 4.14 min (Method A) *m/z* calcd for $C_{41}H_{33}N_7O_9$ [M]: 767.23398, found: 768.97 [M+H]+; **HRMS (MALDI-TOF):** *m/z* found: 768.2390 $[M+H]^+$; **IR ν(cm^{-1}):** 2971.4, 2930.0, 2119.8, 1729.2, 1609.7, 1546.0, 1530.6, 1411.9, 1288.5, 1230.6, 1049.3, 876.7.

Immolative-Estadiol-Aldehyde (5.14b). To a solution of the corresponding diol (20.0mg, 31.1µmol) in MeOH (3ml) was added a solution of $NaIO_4$ (133.1mg, 0.622mmol) in water (1ml). After 1 hour the organic solvent was evaporated under reduced pressure and the resulting solution was diluted with 5ml brine and 5ml water, then extracted with DCM (3x10ml) dried over anhydrous Na_2SO_4 and concentrated under reduced pressure to yield **5.14b** (12.5 mg, 62%) as a pale yellow oil; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 9.80 (1H, s), 7.41÷7.38 (6H, m), 7.30 (1H, d, J= 8.4Hz), 7.05 (2H, d, J= 8.4Hz), 6.94 (1H, dd, J= 9.6Hz, 2.4Hz), 6.90 (1H, d, J= 2.0Hz), 6.71 (1H, bs), 5.22 (2H, s), 3.85 (1H, d, J= 6.0Hz), 2.89÷2.87 (m, 2H), 2.60÷2.56 (1H, m), 2.40÷2.35 (1H, m), 2.32÷2.25 (2H, m), 2.23÷2.18 (2H, m), 1.95÷1.85 (2H, m), 1.76 (1H, dd, J= 12.8Hz, 4.2 Hz), 1.68÷1.60 (4H, m), 1.58÷1.49 (2H, m), 1.47÷1.45 (1H, m), 0.71 (3H, s); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 199.4, 153.9, 148.8, 138.3, 137.0, 136.9, 129.7, 127.9, 126.4, 120.9, 119.1, 118.5, 117.9, 80.1, 77.2, 69.8, 47.7, 45.4, 43.7, 38.6, 32.4, 31.4, 29.6, 29.2, 27,8, 26.0, 24.2, 16.9; **LCMS (ESI):** t_r= 2.22 min, 2.31 min (Method B) – 2 diasteroisomers; *m/z* calcd for $C_{37}H_{41}N_4O_7$ [M]: 652.28970, found: 635.40 [M+H+-H2O]; **HRMS (MALDI-TOF)** could not be measured due to molecule fragmentation; **IR ν(cm-1):** 3515.4, 3023.5, 2121.8, 1758.2, 1747.6, 1730.2, 1602.9, 1590.4, 1443.8, 1374.3, 1228.7, 1008.8.

Dimethyltricarboxyethylphosphine (dmTCEP). Tricarboxyethylphosphine hydrochloride (200mg, 0.21mmol) was stirred with 200mg of sulfonic acid resin (Amberlyst) in 3 ml of methanol at room temperature for 40 minutes. The resin was then removed by filtration and the filtrates, containing a mixture of mono-, di-and tri-methylester, were fractionated by flash chromatography (from DCM to DCM/MeOH 85:15) to give **dmTCEP** as colorless oil (67mg, 34%); **LCMS (ESI):** t_r = 0.71 min (Method B) *m/z* calcd for $C_{11}H_{19}O_6P$ [M]: 278.0919, found: 279.2 $[M+H]^+$.

N-methoxy-N-methyl-N$^{\alpha}$-Fmoc-N$^{\omega}$-Pbf-L-argininamide (5.16). In a round bottom flask N$^{\alpha}$-Fmoc-N$^{\omega}$-Pbf-L-arginine (5.63g, 8.68mmol) was dissolved in 25ml DMF, cooled down to 0°C with an ice bath then HBTU (2.29g, 8.68mmol) and DIPEA (3.0ml, 17.4mmol) were added. The reaction was left to react for 30 minutes. Finally N-methoxy-N-methyl-hydroxylamine (0.85g, 8.68mmol) was added and the reaction left to stir for 10 minute at 0°, then left to react overnight at r.t.. The DMF was removed under reduced pressure and the resulting oil was taken up with 400ml AcOEt, transferred in a separatory funnel and washed with saturated

$KHSO_4$ (2x400ml), saturated $NaHCO_3$ (2x400ml) and brine (400ml). The organic layer was dried over Na_2SO_4 and removed under reduced pressure to afford **5.16** as a white foamy solid (5.99g, 99%); **TLC (AcOEt) Rf:** 0.29, **MP (°C):** 106.2÷108.3°C; **^{1}H NMR ($CDCl_3$, 300MHz) δ(ppm):** 7.76 (2H, d, J= 7.5Hz), 7.63÷7.53 (2H, m), 7.39 (2H, t, J= 7.6Hz), 7.33÷7.24 (2H, m), 6.21 (2H, br s), 5.94 (1H, d, J= 8.5Hz), 4.71 (1H, br s), 4.43÷4.29 (2H, m), 4.17 (2H, t, J= 6.9Hz), 3.73 (3H, s), 3.44÷3.22 (2H, m), 3.19 (3H, s), 2.92 (2H, s), 2.60 (3H, s), 2.53 (3H, s), 2.09 (3H, s), 1.81÷1.52 (4H, m), 1.44 (6H, s); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 172.3, 158.7, 156.7, 156.1, 143.7, 141.2, 138.4, 133.0, 132.3, 127.8, 127.1, 125.1, 124.6, 120.0, 117.5, 86.3, 67.1, 61.7, 50.4, 47.1, 43.2, 40.9, 32.1, 30.6, 28.6, 24.8, 19.3, 17.9, 12.5; **MS (ESI, MeOH):** *m/z* calcd for $C_{36}H_{45}N_5O_7S$ [M]: 691.30397, found: 692.2$[M+H]^+$, 714.2 $[M+Na]^+$, 730.2 $[M+K]^+$, 726.0 $[M+Cl]^-$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 692.31076 $[C_{36}H_{46}N_5O_7S]^+$.

N$^{\alpha}$-Fmoc-N$^{\omega}$-Pbf-L-arginal (5.17). In a 3 neck round bottom flask **5.16** (6.07g, 8.77mmol) was solubilized in 100ml THF, cooled down to 0°C with an ice bath and a 1M solution of $LiAlH_4$ in THF (10.5ml, 10.5mmol) was added. The reaction mixture was left to react for 5 minutes at 0°C and further 40 minutes at r.t.. The reaction was then quenched with 60ml saturated $KHSO_4$ and the organic fraction evaporated under reduced pressure. To the resulting water solution 400ml AcOEt were added, the mixture transferred into a separatory funnel and the organic layer washed with saturated $KHSO_4$ (2x400ml), saturated $NaHCO_3$ (2x400ml) and brine (400ml). The organic layer was then over Na_2SO_4 and removed under reduced pressure to afford yielded **5.17** as a white foamy solid (5.48g, 49%); **TLC (AcOEt) Rf:** 0.56, **MP (°C):** 95.5÷97.7°C; **^{1}H NMR ($CDCl_3$, 300MHz) δ(ppm):** 9.52 (1H, s), 7.77 (2H, d, J= 7.5Hz), 7.58 (2H, t, J= 6.9Hz), 7.40 (2H, t, J= 7.4Hz), 7.30 (2H, t, J= 7.4Hz), 6.59 (1H, s), 5.65 (1H, s), 5.28 (1H, d, J= 9.2Hz), 4.58 (1H, br s), 4.38 (2H, d, J= 7.0Hz), 4.21 (1H, t, J=6.9Hz), 3.67 (2H, m), 2.96 (2H, s), 2.57 (3H, s), 2.51 (3H, s), 2.11 (3H, s), 1.89÷1.50 (4H, m), 1.47 (6H, s); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 158.0, 155.9, 155.3, 144.4, 141.2, 138.0, 134.3, 132.0, 128.1, 127.5, 125.8, 120.6, 116.9, 86.8, 66.1, 51.7, 47.1, 42.9, 39.4, 28.8, 24.3, 19.4, 18.1, 12.7; **MS (ESI, MeOH):** *m/z* calcd for $C_{34}H_{40}N_4O_6S$ [M]: 632.26686, found: 633.1$[M+H]^+$, 655.1 $[M+Na]^+$, 671.1 $[M+K]^+$, 667.1 $[M+Cl]^-$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 633.27405 $[C_{34}H_{41}N_4O_6S]^+$.

N$^{\alpha}$-Fmoc-Ψ-[L-arginin-(N$^{\omega}$-Pbf)]-glycine methylester (5.18). In a round bottom flask glycine methylester hydrochloride (1.73mg, 13.80mmol) and **5.17** (8.73g, 13.80mmol) were solubilized in 150ml MeOH and left to stir 30 minutes together with tryethylamine (1.92ml, 13.8). The temperature was lowered to 0°C with an ice bath and $NaBH_3CN$ (1.04g, 16.56mmol) and acetic acid (947µl, 16.55mmol) were added. The reaction mixture was left to warm to r.t. and to react for 6h. The solvent was then removed under reduced pressure, the resulting oil was taken up with 300ml AcOEt, transferred in a separatory funnel and washed with saturated $NaHCO_3$ (2x200ml) and brine (200ml). The organic layer was then dried over Na_2SO_4 and removed under reduced pressure. Flash chromatography (from AcOEt to AcOEt/MeOH 9:1) yielded **5.18** as a white foamy solid (4.39g, 45%); **TLC (AcOEt) Rf:** 0.28, **MP (°C):** 69.5-71.4°C; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 7.71 (2H, d, J= 7.5Hz), 7.56 (2H, d, J= 6.5Hz), 7.34 (2H, t, J= 7.4Hz), 7.22 (2H, t, J= 7.4Hz), 6.30 (3H, s), 5.63 (1H, d, J= 8.2Hz), 4.33 (2H, d, J= 6.7Hz), 4.13 (1H, t, J= 6.7Hz), 3.64 (4H, br s), 3.39÷3.26 (2H, m), 3.18 (2H, br s), 2.88 (2H, s), 2.58 (5H, m), 2.48 (3H, s), 2.05 (3H, s), 1.86 (1H, br s), 1.43 (10H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 173.2, 158.6, 156.9, 156.3, 143.8, 141.15, 138.2, 132.8, 132.1, 127.6, 127.0, 125.1, 124.6, 119.8, 117.4, 86.3, 66.5, 53.1, 51.8, 50.4, 47.1, 43.1, 40.8, 30.1, 28.5, 25.5, 19.2, 17.9, 12.4; **MS (ESI, MeOH):** *m/z* calcd for $C_{37}H_{47}N_5O_7S$ [M]: 705.31962, found: 706.3 $[M+H]^+$, 728.2 $[M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 706.3258 $[C_{37}H_{48}N_5O_7S]^+$.

General procedure for the synthesis of Fmoc-PNA(5L-Arg$_{(Pbf)}$)-Base(Boc)-OMe. EDC·HCl (2eq), Carboxymethylnucleobase (2eq) and DhBtOH (2eq) were dissolved in 10ml DMF at 0°C, DIPEA (2eq) was added and the solution was left to stir for 10 minutes at 0°C, for other 20 minutes at r.t. and then **5.18** (1eq) dissolved in 3ml DMF was added and left to react for 6h. The DMF was evaporated under reduced pressure and the resulting oil was taken up with 200ml AcOEt and washed with saturate NH_4Cl (2x100ml), saturate

$NaHCO_3$ (2x100ml) and brine (100ml). The organic layer was dried over Na_2SO_4 and the organic solvent evaporated to afford an oil that was purified by flash chromatography (from AcOEt/MeOH 95:5 to AcOEt/MeOH 9:1) to afford the desired product foamy solid. **Fmoc-PNA(5L-Arg$_{(Pbf)}$)-A$_{(Boc)}$-OMe (5.19a):** 1.12g, 81%; **TLC (AcOEt/MeOH 9:1) Rf:** 0.33; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 10.34 (1H, s), 8.56 (1H, s), 8.53 (1H, s), 7.63 (2H, d, J= 6.4Hz), 7.47 (2H, d, J= 6.8Hz), 7.27 (2H, br m), 7.12 (2H, br m), 6.40 (3H, br s), 5.32÷5.05 (1H, m), 5.01÷4.76 (1H, m), 4.33÷4.21 (3H, m), 4.10÷4.00 (1H, m), 3.95÷3.35 (7H, m), 3.25÷2.95 (2H, m), 2.81 (2H, s), 2.53 (3H, s), 2.44 (3H, s), 1.99 (3H, s), 1.65÷1.15 (19H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 169.3, 169.2, 167.4, 166.9, 158.6, 156.4, 156.2, 152.3, 151.1, 149.9, 149.3, 143.9, 143.4, 140.9, 137.9, 132.4, 131.8, 127.4, 126.8, 124.8, 124.5, 120.8, 119.7, 117.3, 86.2, 81.9, 66.5, 52.6, 52.0, 49.9, 48.5, 46.9, 42.9, 40.4, 29.2, 27.9, 25.9, 19.1, 17.7, 12.2; **LCMS (ESI):** t_r = 1.12 min (Method C) *m/z* calcd for $C_{49}H_{60}N_{10}O_{10}S$ [M]: 980.42146, found: 981.3 $[M+H]^+$. **Fmoc-PNA(5L-Arg$_{(Pbf)}$)-C$_{(Boc)}$-OMe (5.19b):** 1.10g, 79%; **TLC (AcOEt/MeOH 9:1) Rf:** 0.46; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 8.46 (1H, br s), 7.68 (2H, d, J= 6.4Hz), 7.57÷7.42 (2H, m), 7.31 (2H, br s), 7.18 (3H, br s), 6.75÷6.32 (3H, br m), 4.99 (1H, br s), 4.81÷4.39 (2H, m), 4.35÷4.00 (4H, m), 3.68 (2H, s), 3.59 (3H, s), 3.35÷2.87 (2H, m), 2.87 (2H, s), 2.58 (3H, s), 2.50 (3H, s), 2.04 (3H, s), 1.75÷1.25 (19H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 169.3, 167.5, 163.2, 158.3, 156.4, 156.3, 156.2, 156.0, 151.0, 149.8, 143.5, 140.9, 137.9, 133.0, 131.8, 127.4, 126.8, 124.9, 124.3, 119.6, 117.1, 66.5, 52.4, 51.9, 50.3, 49.7, 48.5, 46.9, 42.9, 40.1, 28.3, 27.6, 25.5, 25.0, 19.1, 17.7, 13.9, 12.2; **LCMS (ESI):** t_r = 1.16 min (Method C) *m/z* calcd for $C_{48}H_{62}N_8O_{11}S$ [M]: 958.42588, found: 959.3 $[M+H]^+$. **Fmoc-PNA(5L-Arg$_{(Pbf)}$)-G$_{(Boc)}$-OMe (5.19c):** 1.14g, 79%; **TLC (AcOEt/MeOH 95:5) Rf:** 0.29; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 7.79 (1H, s), 7.75÷7.65 (2H, m), 7.59÷7.45 (2H, m), 7.35÷7.07 (4H, m), 5.09 (2H, br s), 4.42÷4.02 (4H, m), 3.79 (2H, s), 3.66 (3H, s), 3.47÷3.05 (4H, m), 2.90 (2H, s), 2.60 (3H, s), 2.47 (3H, s), 2.02 (3H, s), 1.63÷1.25 (19H, m); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 169.9, 169.7, 168.1, 167.8, 158.4, 157.4, 155.6, 153.9, 150.1, 147.8, 143.8, 143.5, 141.2, 140.7, 137.9, 133.0, 132.0, 127.4, 126.8, 124.9, 124.6, 119.5, 118.4, 117.0, 86.3, 83.1, 66.2, 52.4, 51.9, 51.3, 50.3, 49.3, 48.9, 43.8, 42.5, 40.1, 29.2, 27.3, 26.9, 18.2, 17.0, 11.1; **LCMS (ESI):** t_r = 1.11 min (Method C) *m/z* calcd for $C_{49}H_{60}N_{10}O_{11}S$ [M]: 996.41037, found: 997.3 $[M+H]^+$. **Fmoc-PNA(5L-Arg$_{(Pbf)}$)-T-OMe (5.19d):** 1.07g, 85%; **TLC (AcOEt/MeOH 9:1) Rf:** 0.38; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 12.30 (1H, br s), 7.72 (2H, d, J= 7.5Hz), 7.64÷7.60 (2H, m), 7.36 (2H, t, J= 7.4), 7.22 (2H, t, J= 7.4Hz), 7.08 (1H, s), 6.51 (1H, br s), 6.12 (1H, br s), 5.69 (1H, br s), 4.40÷4.10 (6H, m), 4.05÷3.75 (2H, m), 3.69 (3H, s), 3.65÷3.51 (2H, br m), 3.35÷3.30 (2H, m), 2.91 (2H, s), 2.60 (3H, s), 2.53 (3H, s), 2.06 (3H, s), 2.05÷1.81 (2H, m), 1.79 (3H, s), 1.45 (8H, br s); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 168.8, 167.4, 158.6, 156.8, 156.0, 151.5, 143.9, 143.0, 141.2, 138.0, 133.4, 131.9, 127.6, 127.0, 125.3, 125.2, 124.6, 119.8, 117.5, 110.0, 86.4, 66.7, 52.3, 51.8, 50.1, 49.6, 48.1, 47.2, 43.1, 41.3, 30.0, 28.5, 25.3, 19.3, 18.0, 12.4; **LCMS (ESI):** t_r = 1.16 min (Method C) *m/z* calcd for $C_{44}H_{53}N_7O_{10}S$ [M]: 871.35746, found: 872.1 $[M+H]^+$.

General procedure for the synthesis of Fmoc-PNA(5L-Arg$_{(PBf)}$)-Base(Boc)-OH. To a solution of the corresponding ester (1eq) in 60ml THF cooled at 0°C with an ice bath was added a solution of $Ba(OH)_2 \cdot 8H_2O$ (1.5eq) in 60ml H_2O. After 10 minutes the reaction was quenched with 3.5ml HCl 1M, the organic fraction removed and the pH adjusted to 3.5. The solution was then cooled for 2h at 4°C and the desired product was collected over Büchner. **Fmoc-PNA(5L-Arg$_{(Pbf)}$)-A$_{(Boc)}$-OH (5.20a):** 1.03g, 90%; **TLC (AcOEt/MeOH 9:1) Rf:** 0.00; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 10.04 (1H, d, J= 9.8Hz), 8.48 (1H, d, J= 10.4Hz), 8.22 (1H, d, J= 9.6Hz), 7.88 (2H, d, J= 7.2Hz), 7.67 (2H, t, J= 7.5Hz), 7.47÷7.17 (4H, m), 7.10÷6.49 (3H, br m), 5.14 (1H, d, J= 16.8Hz), 5.07 (1H, d, J= 16.8Hz), 4.43÷4.16 (4H, m), 4.15÷3.87 (2H, m), 3.58÷3.37 (2H, m), 3.08÷3.00 (2H, m), 2.92 (2H, s), 2.41 (3H, s), 1.97 (3H, s), 1.55÷1.20 (19H, m); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 171.1, 170.3, 167.2, 166.6, 157.3, 156.0, 155.8, 152.2, 151.3, 151.4, 149.5, 144.8, 143.7, 140.6, 137.1, 134.1, 131.3, 127.5, 126.9, 125.1, 124.2, 122.8, 120.0, 116.1, 86.2, 79.9, 65.2, 51.3, 49.3, 47.9, 46.7, 43.9, 42.3, 29.0, 28.2, 27.8, 25.3, 18.9, 17.5, 12.2; **LCMS (ESI):** t_r = 1.06 min (Method C) *m/z* calcd for $C_{48}H_{58}N_{10}O_{10}S$ [M]: 966.40581, found: 967.3 $[M+H]^+$. **Fmoc-PNA(5L-Arg$_{(Pbf)}$)-C$_{(Boc)}$-OH (5.20b):** 1.05g, 97%; **TLC (AcOEt/MeOH 9:1) Rf:** 0.00; **^{1}H NMR (DMSO-d^6, 400MHz)**

δ(ppm): 10.31 (1H, br s), 7.88 (3H, m), 7.68 (2H, d, J= 7.1Hz), 7.40 (2H, t, J= 7.4Hz), 7.31 (3H, m), 7.23 (d, J= 7.7Hz), 7.06÷6.48 (4H, br m), 4.62 (1H, d, J= 15.5Hz), 4.55 (1H, d, J= 15.5Hz), 4.41÷4.14 (4H, m), 4.01÷3.71 (2H, m), 3.04 (2H, br s), 2.93 (3H, s), 2.42 (3H, s), 1.99 (3H, s), 1.53÷1.23 (19H, m); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 171.0, 170.4, 167.7, 167.1, 163.1, 157.3, 156.0, 151.9, 150.6, 143.7, 140.6, 137.1, 134.2, 131.3, 127.5, 126.9, 125.1, 124.2, 120.0, 116.1, 65.1, 51.0, 49.9, 49.3, 48.9, 48.1, 46.7, 42.3, 28.9, 28.2, 27.7, 25.7, 18.8, 17.5, 12.2; **LCMS (ESI):** t_r = 1.11 min (Method C) *m/z* calcd for $C_{47}H_{60}N_8O_{11}S$ [M]: 944.41037, found: 945.2 $[M+H]^+$. **Fmoc-PNA(5L-Arg$_{(Pbf)}$)-G$_{(Boc)}$-OH (5.20c):** 0.92g, 95%; **TLC (AcOEt/MeOH 9:1) Rf:** 0.00; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 11.35 (1H, br s), 7.86 (2H, d, J= 5.8Hz), 7.74 (1H, s), 7.66 (2H, d, J= 5.8), 7.45÷7.22 (4H, m), 6.73 (2H, br s), 4.84 (2H, br s), 4.45÷4.15 (3H, m), 4.02÷3.72 (3H, m), 3.06 (2H, br s), 2.92 (2H, s), 2.41 (3H, s), 1.98 (3H, s), 1.76÷1.15 (19H, m); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 172.9, 172.6, 171.8, 167.2, 157.3, 156.2, 155.8, 155.0, 149.6, 143.8, 140.6, 140.2, 137.1, 134.2, 131.3, 128.8, 127.5, 127.2, 126.9, 125.1, 121.3, 119.9, 118.7, 116.1, 86.1, 83.1, 65.1, 46.7, 43.8, 42.4, 28.2, 27.6, 25.0, 18.8, 17.5, 12.2; **LCMS (ESI):** t_r = 4.03 min (Method A) *m/z* calcd for $C_{48}H_{58}N_{10}O_{11}S$ [M]: 982.40072, found: 983.6 $[M+H]^+$. **Fmoc-PNA(5L-Arg$_{(Pbf)}$)-T-OH (5.20d):** 0.89g, 85%; **TLC (AcOEt/MeOH 9:1) Rf:** 0.00; **^{1}H NMR (DMSO-d^6, 400MHz, major rotamer) δ(ppm):** 12.25 (1H, s), 7.88 (2H, d, J= 7.2Hz), 7.68 (2H, d, J= 7.4), 7.40 (2H, t, J= 7.3Hz), 7.35÷7.11 (3H, m), 7.07 (1H, s), 6.63 (2H, br s), 4.45 (1H, br s), 4.41÷4.16 (3H, m), 4.03÷3.65 (6H, m), 3.04 (2H, br s), 2.93 (3H, s), 2.42 (3H, s), 1.98 (3H, s), 1.70 (3H, s), 1.52÷1.23 (10H, m); **^{13}C NMR (DMSO-d^6, 100MHz, major rotamer) δ(ppm):** 171.5, 167.7, 164.3, 157.3, 156.1, 150.9, 143.7, 142.5, 142.0, 140.6, 137.1, 134.2, 131.3, 128.3, 127.5, 126.9, 125.11, 124.2, 120.0, 116.1, 107.9, 86.1, 65.1, 51.9, 51.2, 49.2, 48.8, 47.7, 46.7, 42.3, 28.9, 28.6, 28.2, 25.0, 18.8, 17.5, 12.2; **LCMS (ESI):** t_r = 1.00 min (Method C) *m/z* calcd for $C_{43}H_{51}N_7O_{10}S$ [M]: 857.34181, found: 858.1 $[M+H]^+$.

Kinetics of small molecule release upon TCEP treatment. Aldehydes **5.14a-b** dissolved in acetonitrile were added to PBS and release of functional molecule after addition of TCEP (final concentration of aldehyde: 500μM, TCEP 25mM, 100mM PBS, pH=7.4) was followed by LCMS (5μL injections).

Time of injection (min)	Integral **5.14a**	Integral **5.16a**	Integral **5.14b**	Integral **5.16b**
0	330548	0	118985	0
1	200995	49823	52386	49312
6	85496	302269	7703	81115
11	37040	516143	0	85377
16	18766	652717	-	-
23	0	688557	-	-
31	0	783900	-	-

Synthesis and characterization of PNAs. AOOA-PNA synthesis was performed by standard Fmoc chemistry as previously described on NovaPEG (NovaBiochem) Rink amide resin loaded with Fmoc-Lys(Mtt)-OH (0,2 mmol/g). Mtt from side chain of lysine was deprotected before removal of terminal Fmoc from PNA strand and additional coupling to introduce aminooxyacetyl moiety on liberated amino group was done without capping at the end. Then terminal Fmoc was removed and product was cleaved from resin. **PNA-dmTCEP** synthesis was prepared by standard Fmoc chemistry, deprotected at the N-terminus, was treated with **dmTCEP** (6.0 equiv) in DMF previously activated for one minute with DIC (24 equiv) and HOBt (12 equiv) for 30 min. Then resin was washed first with 5% AcOH in DCM then with DCM and DMF. **N_3RhN_3-PNAs** synthesis was performed by standard Fmoc chemistry as previously described on NovaPEG (NovaBiochem) Rink amide resin loaded with Fmoc-Lys(Mtt)-OH (0,2 mmol/g), Mtt from side chain of lysine was deprotected and coupled with 5,6-carboxy-di-azidorhodamine (5.0 equiv) using

HCTU/DIPEA/lutidine as activating mixture, for PNAs bearing the poly-lysine chain the introduction of the pro-fluorescent group was achieved after PNA synthesis by selective Mtt deprotection and introduction of the modification before resin cleavage. PNA were purified on C18 using Biotage ISOLERA flash chromatography system and characterized by MALDI. **AOOA-PNA:** calcd 2897.53 [M]: *m/z* found 2898.57 $[M+H]^+$; **PNA-dmTCEP:** calcd 2630.33 [M]: *m/z* found 2632.03 $[M+H]^+$; **PNA 5.3:** calcd 5032.23 [M]: *m/z* found 4977.03 $[M+H-2N_2]^+$; **PNA 5.4:** calcd 4834.07 [M]: *m/z* found 4779.16 $[M+H-2N_2]^+$; **PNA 5.5:** calcd 5032.23 [M]: *m/z* found 4977.09 $[M+H-2N_2]^+$; **PNA 5.6:** calcd 4834.07 [M]: *m/z* found 4778.99 $[M+H-2N_2]^+$; **PNA 5.7:** calcd 4794.03 [M]: *m/z* found 4739.00 $[M+H-2N_2]^+$; **PNA 5.8:** calcd 2616.20 [M]: *m/z* found 2617.53 $[M+H]^+$; **PNA 5.9:** calcd 2517.12 [M]: *m/z* found 2518.36 $[M+H]^+$; **PNA 5.10:** calcd 2616.20 [M]: *m/z* found 2617.26 $[M+H]^+$; **PNA 5.11:** calcd 2517.12 [M]: *m/z* found 2518.38 $[M+H]^+$; **PNA 5.12:** calcd 2803.33 [M]: *m/z* found 2804.77 $[M+H]^+$; **PNA 5.13:** calcd 5055.24 [M]: *m/z* found 5000.26 $[M+H-2N_2]^+$; **PNA 5.14 :** calcd 4857.08 [M]: *m/z* found 4802.26 $[M+H-2N_2]^+$; **PNA 5.15:** calcd 5055.24 [M]: *m/z* found 5000.52 $[M+H-2N_2]^+$; **PNA 5.16 :** calcd 4857.08 [M]: *m/z* found 4802.32 $[M+H-2N_2]^+$; **PNA 5.17:** calcd 4817.04 [M]: *m/z* found 4762.30 $[M+H-2N_2]^+$; **PNA 5.18:** calcd 2630.20 [M]: *m/z* found 2632.03 $[M+H]^+$; **PNA 5.19:** calcd 2531.12 [M]: *m/z* found 2533.26 $[M+H]^+$; **PNA 5.20:** calcd 2630.20 [M]: *m/z* found 2631.65 $[M+H]^+$; **PNA 5.21:** calcd 2531.12 [M]: *m/z* found 2531.59 $[M+H]^+$; **PNA 5.22:** calcd 2817.33 [M]: *m/z* found 2818.78 $[M+H]^+$.

General Procedure for the aldehyde coupling. 50 µL of a 5 mM solution of **AOOA-PNA** (250 nmol, 1.0 equiv) in H_2O/MeCN 1:1 were diluted with 125 µL of MeCN, then 75 µL of a 10 mM solution of aldehyde **5.14** (750 nmol, 3.0 equiv) in MeCN was added. After 5 hours the reaction was stopped adding 400 µL of H_2O and lyophilizing. The product was then isolated by HPLC purification and characterized by MALDI. **PNA 5.1:** calcd 3646.76 [M]: *m/z* found 3647.93 $[M+H]^+$; **PNA 5.2:** 79%, calcd 3531.82 [M]: *m/z* found 3532.48 $[M+H]^+$.

Templated reactions for rhodamine release. Black 96-well plates (500 µl/well) were used to perform the templated reactions. Stock PNA and DNA solutions at 10µM and 1µM were prepared in deionized water and diluted to their final concentrations with buffer containing 10mM PBS, 154mM NaCl, 25mM $MgCl_2$, 0.1mg/ml Bovine γ-Globuline (BGG) at pH=7,4. First the wells were filled with 240µl of the **PNA-dmTCEP** solution containing appropriate DNA. After 240µl of **PNA 5.2** solutions was added using multichannel pipette and the plate was placed in a fluorometer (SpectraMax GeminiXS, Molecular Devices) pre-heated to 37°C and the fluorescence readout was recorded immediately. All experiments were performed in duplicates and each individual experiment included a positive control (**PNA-dmTCEP** solution was replace with 240µl of 20mM TCEP solution in buffer), two negative control (no DNA template and random DNA template), and a background florescence measurement (reaction without **PNA-dmTCEP**). The fluorescence level corresponding to 100% conversion was measured by treating all reactions with large excess of TCEP. The fluorescence measurements of the templated reactions was performed using 495nm (excitation) and 525nm (emission) at 37°C. The curves shown represent the average of the duplicate measurements.

Templated reactions for estradiol release. Stock solutions and buffer were prepared identically as for fluorescence experiments. 250µl of the **PNA-dmTCEP** solution containing appropriate DNA template was added to eppendorf. After 250µl of **PNA 5.2** solution was added using multichannel pipette and the rack with tubes was shaken for 30 minutes at 37°C. Then all reaction mixtures were quenched by addition of 100µl of 35% H_2O_2 solution. After 10 minutes at room temperatures resulting solutions were lyophilized. Released estradiol was extracted from dry residue with 1:1 DCM:MeCN. Tubes were sonicated 10 minutes after addition of 1ml of mixture. Then samples were spun down, supernatant was transferred to fresh tubes and evaporated to dryness in speedvac. Residue was then re-dissolved in 50µL of MeCN and 10µL portions of samples were injected to HPLC (Thermo Scientific Accela HPLC with a Thermo C18 (50x2.1mm, 1.9µm particles) hypersil gold column, linear elution gradient for 95% H_2O 0.01% TFA to 90% MeCN 0.01% TFA

in 3.6 minutes at a flow rate of 1.0 ml/min) equipped with a fluorescence detector (Jasco X-LCTM 312 OFP). All experiments were performed in duplicates or triplicates and each samples was injected twice. The fluorescence measurements were performed using 290nm (excitation) and 310nm (emission). Each individual experiment included a positive control (reaction mixture with random DNA template spiked with known amount of estradiol) and two negative controls (no DNA template and random DNA template), and a background florescence measurement (reaction without **PNA-dmTCEP**). The conversion of measured peak areas to estradiol concentrations was done using a calibration curve. Calibration curve was recorded following identical preparation, incubation, quenching and extraction procedures but **PNA 5.2** was replaced by the same amount of non-labelled PNA oligomer and known amount of estradiol was added before starting incubation.

Preliminary cellular uptake evaluation on fixed cells. Cells were seeded on 35 mm glass bottom dishes 12h before the experiment. Then they were washed with PBS and fixed with 3.7% formaldehyde (20'). After washing with PBS (3x5') they were treated with 100 nM PNA-azidorhodamine probe in PBS (or PBS alone for control) for 3.5h at 37°C. Then incubation solution was removed, cells were washed with PBS and treated with 1mM tmTCEP) for 2h at 37°C. After final washing with PBS cells were imaged with 60x oil immersion objective using a Nikon Eclipse Ti inverted fluorescent microscope (filter set: excitation 460-500 nm, emission 510-560 nm, 1s exposure time).

4.5 - References

[1] a) Orgel L. E. *Nature* **1992**, *358*, 203-209; b) von Kiedrowski G. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 932-935; c) Sievers D., von Kiedrowski G. *Nature* **1994**, *369*, 221-224; d) Orgel L. E. *Acc. Chem. Res.* **1995**, *28*, 109-118; e) Joyce G. F. *Cold Spring Harbor Symposia on Quantitative Biology*; Cold Spring Harbor Press: New York, 1987; Vol. 52, pp 41-51; f) Wintner E. A., Conn M. M., Rebek J. *Acc. Chem. Res.* **1994**, *27*, 198-203; g) Anderson S., Anderson H. L., Sanders J. K. M. *Acc. Chem. Res.* **1993**, *26*, 469-475; h) Kanavarioti A. *J. Org. Chem.* **1998**, *63*, 6830-6838; i) Kurz M., Gobel K.; Hartel C., Gobel M. W. *Angew. Chem., Int. Ed. Engl.* **1997**, 36 (8), 842-845.

[2] Silverman A. P., Kool E. T. *Chem. Rev.* **2006**, 106, 3775-3789.

[3] Gryaznov S. M., Letsinger R. L. *Nucleic Acids Res.* **1993**, *21*, 1403-1408.

[4] Dose C., Seitz O. *Org. Biomol. Chem.* **2004**, 2, 59-65.

[5] Cai J., Li, X. Yue X., Taylor J. S. *J. Am. Chem. Soc.* **2004**, 126, 16324-16325.

[6] Abe H., Wang J., Furukawa K., Oki K., Uda M., Tsuneda S., Ito Y. *Bioconjugate Chem.* **2008**, 19, 1219-1226.

[7] Abe H., Kool E. T. *J. Am. Chem. Soc.* **2004**, 126, 13980-13986.

[8] Li H., Franzini R. M., Bruner C., Kool E. T. *ChemBioChem* **2010**, 11, 2132-2137.

[9] Ma Z., Taylor J.-S. *Proc. Nat. Acad. Sci. USA* **2000**, 97, 11159-11163.

[10] Brunner J., Mokhir A., Kraemer R. *J. Am. Chem. Soc.* **2003**, 125, 12410-12411.

[11] Fujimoto K., Matsuda S., Hayashi M., Saito I. *Tetrahedron Lett.* **2000**, 41, 7897-7900.

[12] Lemieux G. A., De Graffenried C. L., Bertozzi C. R. *J. Am. Chem. Soc.* **2003**, 125, 4708-4709.

[13] Rooseboom M, Commandeur J. N., Vermeulen N. P. *Pharmacol. Rev.* **2004**, 56, 53-102.

[14] Kratz F., Muller I. A., Ryppa C., Warnecke A. *ChemMedChem* **2008**, 3, 20-53.

[15] Lee H. M., Larson D. R., Lawrence D. S. *ACS Chem. Biol.* **2009**, 4, 409-427.

[16] a) Silverman A. P., Kool E. T. *Chem Rev* **2006**, *106*, 3775-3789; b) Jacobsen M. F., Clo E., Mokhir A., Gothelf K. V. *ChemMedChem* **2007**, 2, 793-799; c) S. Tyagi, *Nat. Methods* **2009**, 6, 331-338.

[17] Ma Z., Taylor J. S. *Proc. Nat. Acad. Sci. USA* **2000**, *97*, 11159-11163.

[18] a) Ma Z., Taylor J. S. *Bioorg. Med. Chem.* **2001**, *9*, 2501-2510; b) Ma Z., Taylor J. S. *Bioconjug. Chem.* **2003**, *14*, 679-683; c) Cai J., Li X., Taylor J. S. *Org Lett* **2005**, *7*, 751-754.

[19] Sando S., Kool E. T. *J. Am. Chem. Soc.* **2002**, *124*, 2096-2097.

[20] Franzini R. M., Kool E. T. *J. Am. Chem. Soc.* **2009**, *131*, 16021-16023.

[21] Okamoto A., Tanabe K., Inasaki T., Saito I. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 2502-2504.
[22] Tanabe K., Nakata H., Mukai S., Nishimoto S. *Org. Biomol. Chem.* **2005**, *3*, 3893-3897.
[23] Sletten E. M., Bertozzi C. R. *Angew. Chem. Int. Ed. Engl.* **2009**, 48, 6974-6998.
[24] a) Pianowski Z., Gorska K., Oswald L., Merten C. A., Winssinger N. *J. Am. Chem. Soc.* **2009**, 131, 6492-6497; b) Furukawa K., Abe H., Hibino K., Sako Y., Tsuneda S., Ito Y. *Bioconjug. Chem.* **2009**, 20, 1026-1036.
[25] Carl P. L., Chakravarty P. K., Katzenellenbogen J. A. *J. Med. Chem.* **1981**, 24, 479-480.
[26] Lavis L. D., Chao T. Y., Raines R. T. *ACS Chem. Biol.* **2006**, 1, 252-260.
[27] Leytus S. P., Melhado L. L., Mangel W. F. *Biochem. J.* **1983**, 209, 299-307.
[28] Cruz F. G., Koh J. T., Link K. H. *J. Am. Chem. Soc.* **2000**, 122, 8777-8778.
[29] Li H., Hah J. M., Lawrence D. S. *J. Am. Chem. Soc.* **2008**, 130, 10474-10475.
[30] Spletstoser J. T., Flaherty P. T., Himes R. H., Georg G. I. *J. Med. Chem.* **2004**, 47, 6459-6465.
[31] Pianowski Z., Winssinger N. *Chem. Commun.* **2007**, 3820-3822.
[32] Pothukanuri S., Pianowski Z., Winssinger N. *Eur. J. Org. Chem.* **2008**, 3141-3148.
[33] Dirksen A., Dirksen S., Hackeng T. M., Dawson P. E. *J. Am. Chem. Soc.* **2006**, *128*, 15602-15603.
[34] Yoon Y., Westerhoff P., Snyder S. A., Esparza M. *Water Res.* **2003**, *37*, 3530-3537.
[35] Tanke H. J., Dirks R. W., Raap T. *Curr. Opin. Biotechnol.* **2005**, 16, 49-54.
[36] Tyagi S. *Nat. Methods* **2009**, 6, 331-338.
[37] Silverman A. P., Kool E. T. *Chem. Rev.* **2006**, 106, 3775-3789.
[38] Furukawa K., Abe H., Hibino K., Sako Y., Tsuneda S., Ito Y. *Bioconjugate Chem.* **2009**, 20, 1026-1036.
[39] Rothbard J. B., Kreider E., VanDeusen C. L., Wright L., Wylie B. L., Wander P. A. *J. Med. Chem.* **2002**, 45, 3612-3618.

Chapter 5 - Artificial RNase PNA

5.1 - Introduction

Ribonucleases (RNase) are a class of nucleases that catalyze the hydrolysis of RNA strands. Different classes of RNase are present in all organisms and play a key role in the maturation of mRNA or other non-coding RNA and in the regulation of its content inside the cell; furthermore, they are the first defense against RNA viruses[1]. Natural RNases show a very different activity as a function of their role, for example non specific RNases (RNase A, RNase V1, RNase L, oligoribonuclease, …) are involved in the regulation of the RNA level within the cell by degrading their targets without any specificity, whereas specific RNases (RNase III, RNase P, RNase T, …) play an important role in the maturation of RNA sequences degrading them at specific positions.

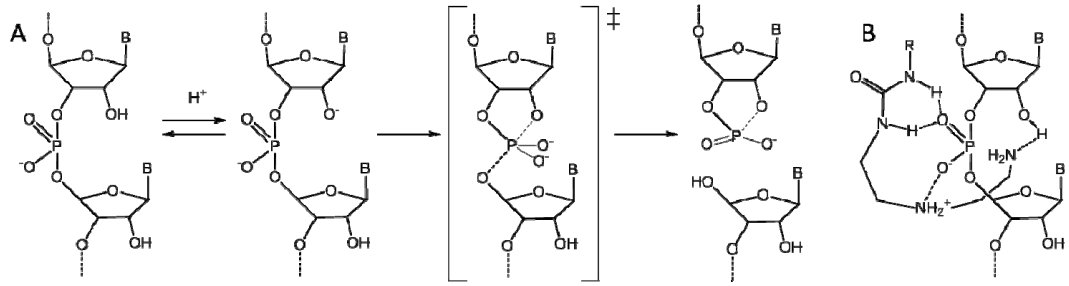


Figure 5.1: A) mechanism of basic hydrolysis of the RNA linkage; B) placement of the DETA fragment around the phosphate unit.

The formal mechanisms by which RNA cleavage occurs can be categorized into two general classes based upon the nature of the products obtained. In most cases the reactions proceed as transesterifications with the attacking nucleophile being either a 2'-hydroxyl or a terminal 3'-hydroxyl. In some enzymes the 2'-hydroxyl adjacent to the scissile phosphodiester is employed as an internal nucleophile, but it can be possible also to find enzymes that use external groups (from co-factors, water or other part of the RNA sequence that are placed outside of the catalytic pocket)[2]. In Figure 5.1A the mechanism of cleavage by transesterification using the 2'-hydroxyl group is shown: the initial deprotonation of the alcoholic function (pKa ~ 13) leads to the activation of the nucleophile that attacks the phosphate group forming a bipyramidal intermediate and inducing the break of the axial (trans) bond, forming a pentacyclic phosphodiester that can be successively hydrolyzed. The catalytic activity of the enzyme is executed by three different actions: increase of

the acidity of the hydroxyl group, reduction of the electron density on phosphorous and attainment of the correct geometry for the formation of the right intermediate bipyramidal complex.

Based on this mechanism, different systems able to play the role of the catalytic pocket of the enzyme were proposed. These artificial RNases are artificial molecules with the ability of cleaving the phosphodiester bond within RNA nucleotides. During the last years they have received a great interest for the possibility to use them as RNA restriction enzymes or for gene silencing[3]. More in detail, the reasons for their interest in these completely artificial enzymes, compared to natural enzyme or semi-artificial ribozimes, are their greater stability to other enzymes, the possibility of selectively cleaving a chosen position without the need of recognizing specific sequences and the possibility of efficiently cleave RNA with complex tertiary structures thanks to their reduced steric hindrance.

The general model for artificial RNases is composed by two main domains. The former part is the catalytic site, an organic or inorganic group able to catalyze the cleavage of the phosphodiester bond. The latter is formed by an oligonucleotide sequence (or analog) that is able to specifically recognize the target sequence and allows the placement of the catalytic moiety in close proximity of the desired cleavage site.

Looking at the literature, a lot of different approaches have been proposed for the synthesis of these systems based on differences in the nature of the recognizing sequence, the catalytic site or the strategy of action. Concerning the recognition part, a lot of work is being done using the discriminating ability of DNA sequences[4,5], but also work based on the use of PNA can be found[3,6].

The role of the catalytic site is the stabilization of the transition state leading to the intermediate in the process of hydrolysis, the reduction of the electron density of the phosphate to enhance its electrophilicity, and/or the activation of the attacking nucleophile. This can be performed by a broad variety of complexes bearing different ions, such as lanthanide[7], copper[8], zinc[9]; however, one of the major drawbacks of these systems is the possible leak of the metal once introduced in the cellular system. To overcome this problem ion-independent catalytic sites were proposed either bearing diethylentriamino (DETA, its coordination around the phosphate group is show in Figure 5.1) moieties[5,6b] or histidine residues[10,11] or other organocatalytic moieties were proposed (Figure 5.2)[12].

All artificial RNases described above work placing the active species in close proximity of the phosphodiester bond to be cut, but there are other kinds of systems in which the recognizing moiety introduces a residue (called pinpoint activator) that induces a bulge formation in the double stranded structure able to activate a specific bond to be cleaved by simply addition of a lanthanide ion

(Lu(III), called molecular scissor) to the solution, or by addition of a poor catalyst such as Zn(II), Mn(II) or Mg(II)[13,14].

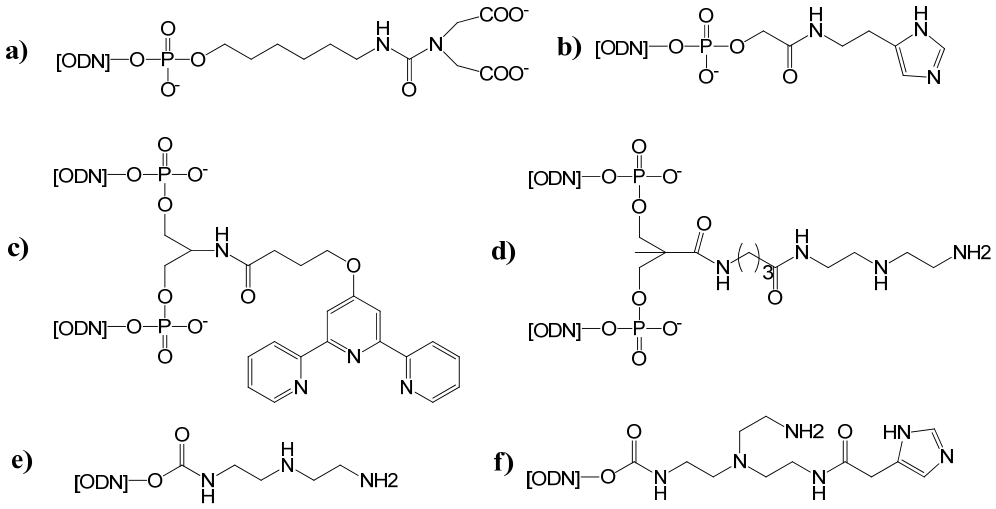


Figure 5.2: artificial RNases proposed; a) terminal chelating ligand for lanthanide ions, b) terminal ligand for Zn^{2+}, c) central ligand for Cu^{2+}, d) central DETA derivative, e) terminal DETA derivative, f) terminal amino-imidazole derivative (From Ref. 12).

Unluckily all the systems proposed so far did not have catalytic properties, due to the strong interaction between the recognizing moiety and the target RNA strand, that is kept after the cleavage of the target bond. For this reason, all these systems can be used as efficient tool in bench practices, but the introduction of their use in gene silencing they did not induce an effective gain from the normal antisense therapy.

5.2 - Results and discussion

In this chapter the synthesis of a series of building blocks for the construction of PNA-based artificial RNases with catalytic properties is reported.

As a catalytic site we chose to use a DETA derivative that was already reported to be effective in RNA[5,6b] cleavage by an acid-base mechanism of activation of the 2'-hydroxyl group and hydrogen bonding at the phosphate unit (Figure 5.1B). To induce a catalytic behavior in the systems we decided to not simply conjugate the catalytic function to an extremity of the sequence, as reported by van Boom's and Messere's groups, but to modify a PNA monomer in order to place the cleavage system in the middle of the recognition probe and introduce multiple units along the strand, and to

possibly induce a catalytic behavior (Figure 5.3). To achieve this target, we decided to explore two different approaches: i) modification of the PNA backbone to place the catalytic group close to the phosphate unit; ii) modification at the nucleobase level, to reach the phosphate through the major groove of the PNA:DNA double strand.

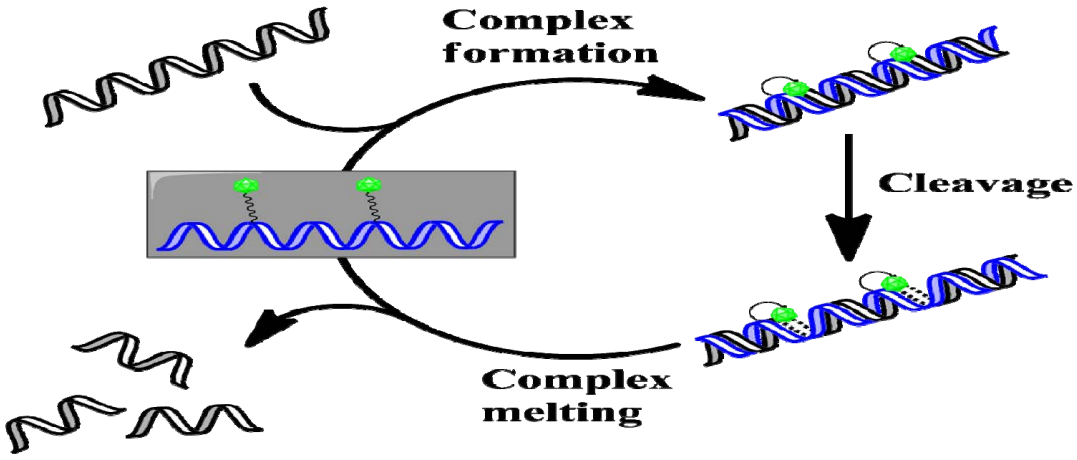
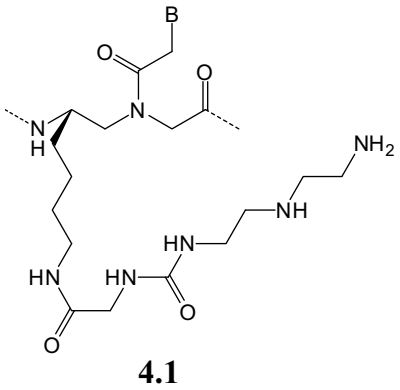


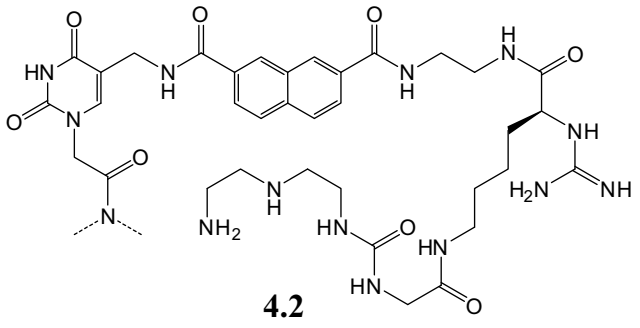
Figure 5.3: proposed mechanism of action of PNAs.

5.2.1 - Design and synthesis of the building blocks

In chapter 2 we have described the higher stability of PNAs with an arginine side chain in the 5L-position, as compared to the 2D- modification, as a factor of proximity of the positively charge guanidinium group in close proximity to the negatively charged phosphate group. Based on this



4.1



4.2

Figure 5.4: models proposed in this study.

concept, we decided to synthesize a monomer based on 5L-lysine, and use the ε-amino group to anchor the catalytic function during solid phase synthesis (**4.1** inFigure 5.4). When the modification was introduced at the nucleobase level, we started from uracil and built up a complex system for the interaction with the complementary adenine base to correctly place the catalytic function in close proximity of the target phosphodiester bond (**4.2** in Figure 5.4).

Both systems were designed using a strategy which allows introduction of the catalytic moiety during the PNA synthesis on the solid phase, allowing the possibility to easily implement the number of modifications that can be introduced only by variation of a single catalytic unit, with no need to synthesize of an entire PNA monomer for each modification. Thus the synthons **4.3**, **4.4**, **4.5**, **4.6** and **2.25** (Figure 5.5) were obtained and combined to construct "engineered" complex PNA derivatives.

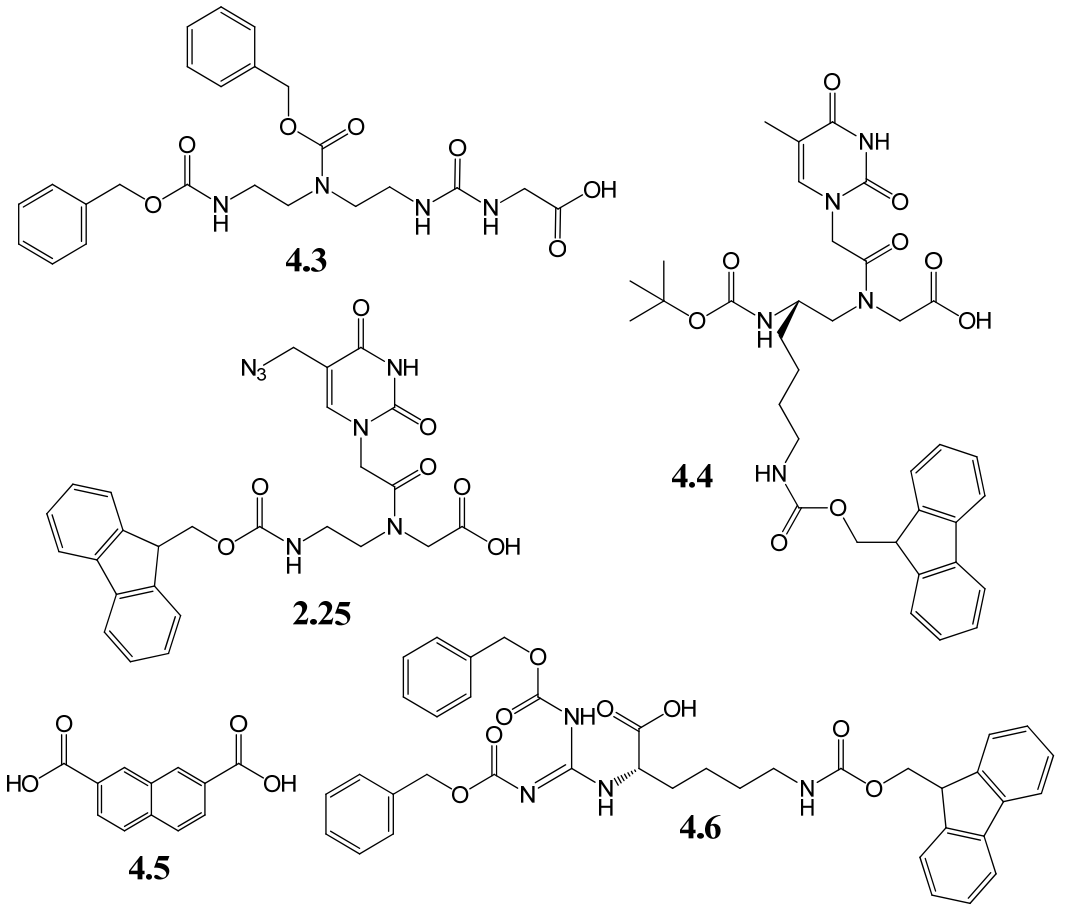
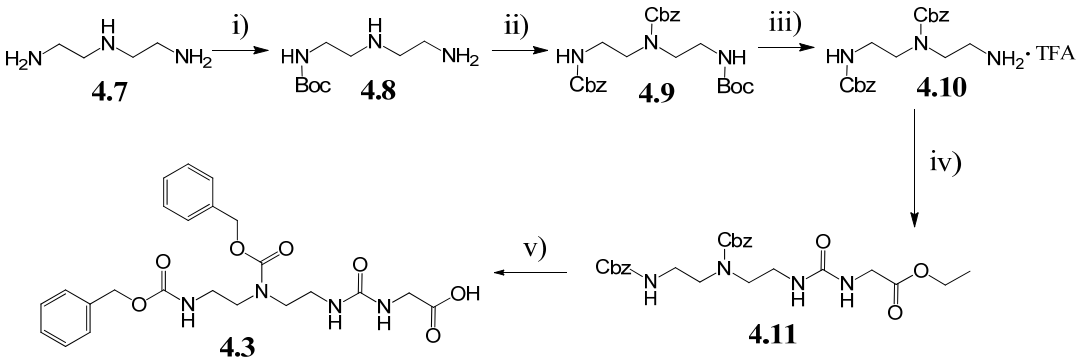


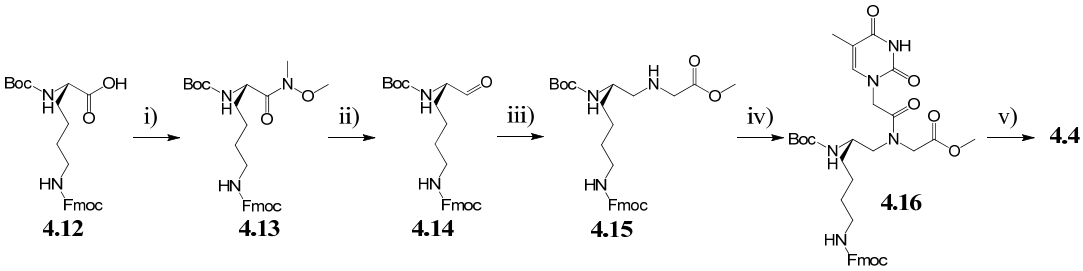
Figure 5.5: synthons used for the realization of the proposed models.

The synthesis of the building block containing the catalytic center (Scheme 5.1) starts with the N1-mono protection of the diethylenetriamine, following a similar procedure reported for the synthesis of mono-protected ethylenediamine[15], with good yield. After the first Boc-protection, the other two amino groups were protected with Cbz group and the more labile Boc-protective group was removed under acidic conditions to obtain the diprotected amine **4.10**. The trifluoroacetate salt was then converted to an urea derivative, coupling it with an isocyanate to obtain the protected synthon **4.11**. The final ester hydrolysis under basic conditions gave the desired building block with the fully protected catalytic center **4.3**. The relatively low yield obtained during the second and fourth step are related to the presence of traces of N1,N4-Boc protected product or N4-Boc isomer so that a supplementary purification step was needed.



Scheme 5.1: synthesis of the catalytic site monomer 4.3: i) Boc_2O in DCM, 90%; ii) Cbz-Cl, DIPEA in DCM, 65%; iii) TFA in DCM, quant.; iv) ethyl isocyanato-acetate, K_2CO_3 in DCM, 79%; v) NaOH in THF/H_2O 1:1, 92%.

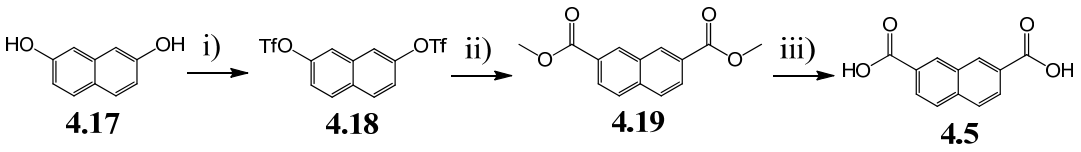
As previously described in chapter 2 and chapter 3 the synthesis of 5L-modified monomer (Scheme 5.2) started from the corresponding amino acid **4.12** that was converted into aldehyde **4.14** by reduction of the corresponding Weinreb amide **4.13** with $LiAlH_4$. The aldehyde was then linked by reductive amination to the backbone **4.15,** which was coupled to carboxymethylthimine to obtain



Scheme 5.2: synthesis of the Boc-5L-$Lys_{(Fmoc)}$-T-PNA monomer 4.4: i) N,O-dimethylhydroxylamine hydrochloride, HBTU, DIPEA in DMF, 99%; ii) $LiAlH_4$ in THF, 49%; iii) glycine methyl ester hydrochloride, $NaBH_3CN$, CH_3COOH in MeOH, 72%; iv) CMT, EDC·HCl, DHBtOH in DMF, 80%; v) $Ba(OH)_2$ in THF/H_2O = 1: 1, 80%.

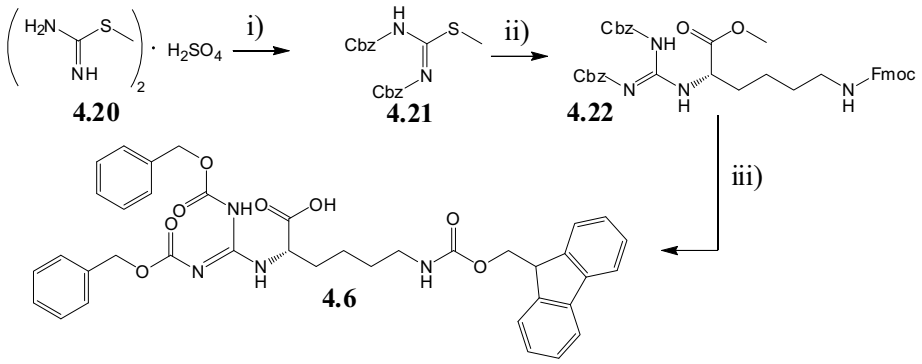
the protected monomer **4.16**. Hydrolysis with $Ba(OH)_2$ yielded the final chiral monomer **4.4**. Compared to previously described C5-modified monomers, the presence of the Fmoc protective group on the ε-amino group gave rise to by-products during the reduction step so that it was necessary to introduce a supplementary purification step.

For the insertion of the synthon **4.2** inside the PNA strand we followed an on-resin built protocol, using the previously described monomer **2.25** for the introduction of an orthogonally protected amine and the following functionalization of the corresponding amine. The protocol moreover avoids the necessity of a solution desymmetrization of the symmetric spacer **4.5** which could induce poor yields of the desired compound. The synthesis of the symmetric spacer **4.5** (Scheme 5.3) starts from the commercially available naphthalene-2,7-diol which is activated by conversion of the hydroxyl groups to a better living group such as trifluoromethanesulfonate, by reaction with trifluoromethanesulfonic anhydride[16] to obtain **4.18**. Carboxymethylation of this compound was achieved following a described procedure for similar compounds [17]. The low yield obtained in the synthesis of this intermediate is related to the side production of the monocarbonylated derivative (30%), which could however be useful for a solution desymmetrization protocol. The desired bridging unit **4.5** was obtained by basic hydrolysis of the symmetric ester.



Scheme 5.3: synthesis of the spacer 4.5: i) trifluoromethanesulfonic anhydride, Et_3N in DCM, 95%; ii) MeOH, $Pd(Ac)_2$, PPh_3, Et_3N, CO in DMF, 43%; iii) NaOH in H_2O/MeOH, 85%.

For the spacer with the guanidine group (inserted for the recognition of the N7 of the adenine) the synthesis (Scheme 5.4) started from the diprotected S-methylisothiourea **4.20** that was coupled with a lysine derivative bearing a protection on the ε amino group and on the carboxylate, prior to the final hydrolysis of the ester function to obtain the desired product **4.6**. The protection of the guanylating agent followed a previous described protocol[18] and the coupling of this product with the desired amino acid was favored in presence of a mercury salt that form a precipitate in the presence of the sulfide generated during the reaction. Final hydrolysis of the ester yielded the desired spacer **4.6** in low yield due to the side formation of the mono-Cbz protected target. The same behavior was also observed during the cleavage of the methyl ester function in analogous derivatives bearing the double protected guanidinium function in α-position to the ester to be hydrolyzed (data not shown).



Scheme 5.4: synthesis of the spacer 4.6: i) NaOH, Cbz-Cl in water/DCM, ii) H-Lys(Fmoc)-OMe·HCl, DIPEA, $HgCl_2$ in DMF, 87%; iii) $(BaOH)_2 \cdot 8H_2O$ in water/THF,

5.2.2 - Synthesis of PNAs

The built of the final targets **4.1** and **4.2** followed an on-resin modification protocol. In the former case the monomer **4.4** was inserted in the PNA strand during the Boc-based solid phase synthesis of the polymer and the Fmoc protecting group was removed under basic conditions (20% piperidine in DMF) to generate the amine function to be used as branching point for the introduction of **4.6 (**Figure 5.6**)**.

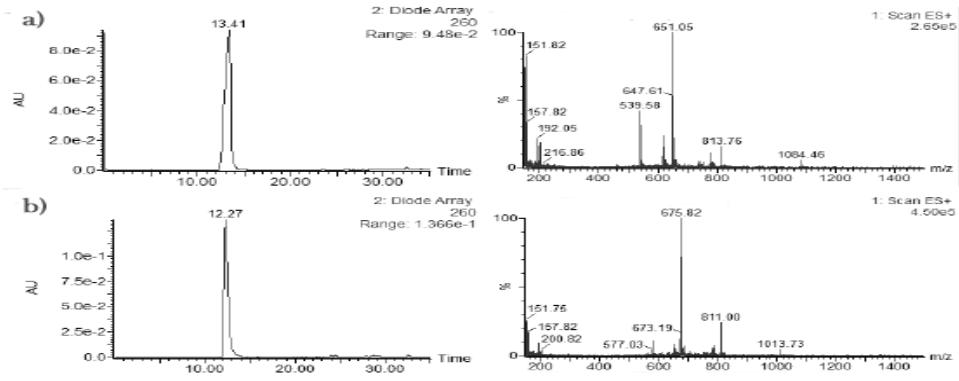
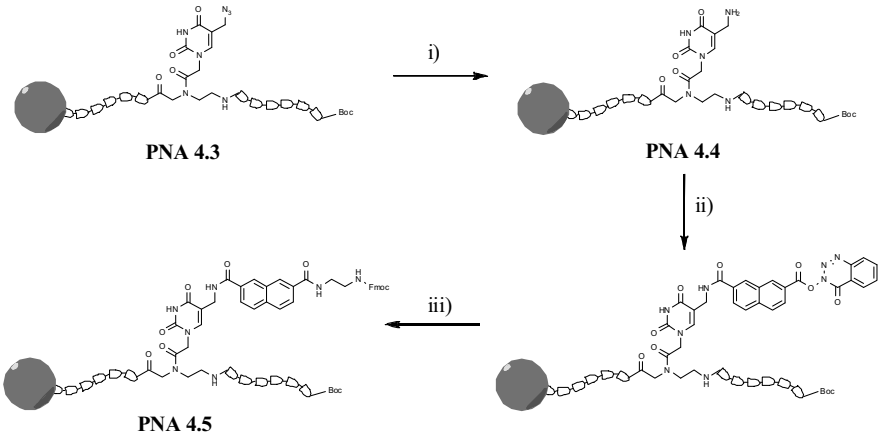


Figure 5.6: HPLC trace and mass spectrum of the chromatographic peak of PNAs bearing 4.2; a) PNA 4.1, b) PNA 4.2.

The full synthesis of the target **4.2** was not achieved due to the low quantities of the synthon **4.6** obtained; however, a solid-phase procedure for the orthogonal building of this systems was designed (Scheme 5.5). We first introduced monomer **2.25** in the elongation of the sequence in Boc-based solid phase synthesis. The azide function can be reduced in an orthogonal way, as described

in chapter 2, and this amine was used to induce the desymmetrization of the symmetric acid **4.5**. This acid was pre-activated with an excess of DIC and DhBtOH to ensure the fast conversion of the acidic function to the activated ester, and the solution was transferred in the resin vessel. On the solid phase only one of the two acidic functions can be readily coupled with the amine coming from the reduction of the azide function, while the second activated ester can be used to link the free amine function of the Fmoc-ethylenediamine, to obtain a final product than can be modified in the side direction using Fmoc-based solid phase synthesis or elongated in the strand direction using Boc-based approach.



Scheme 5.5: solid phase desymmetrization of 4.5: i) PMe_3 in THF/H_2O 1:1, 2x10'; ii) 4.5, DIC, DhBtOH in DMF, 2h (10' pre-activation); iii) Fmoc-ethylenediamine hydrobromide, DIPEA in DMF, 3h.

Thus both syntheses can be carried out during polymer elongation or at the end of the chain growth. Sequence of the PNA synthesized and their corresponding models are shown in Table 5.1.

Table 5.1: PNA sequences used for this study. Underlined letters are for modified monomers.

Name	PNA sequence	Catalytic model
PNA 4.1	H-CGCTGTCACAC-Gly-NH_2	4.1
PNA 4.2	H-GCCGCTGTCACAC-Gly-NH_2	4.1
PNA 4.3	Boc-GT(N_3)GACAC-Gly-$MBHA_{res}$	For 4.2 synthesis
PNA 4.4	Boc-GT(NH_2)GACAC-Gly-$MBHA_{res}$	For 4.2 synthesis
PNA 4.5	Boc-GT(Napt+EDA)GACAC-Gly-$MBHA_{res}$	For 4.2 synthesis

The synthesis of the final PNA bearing the full catalytic unit as in **4.2** is still under development.

5.2.3 - Evaluation of the melting properties

All PNAs synthesized are designed for placing the reactive center in the middle of the probe in order to induce a strong destabilization of the complex after cleavage of the phosphodiester bond of the target RNA. As shown in Figure 5.3, the idea is to have a strong interaction between the PNA and the full length RNA, to allow the correct recognition between the two different strands, but a very weak interaction between the two strands after cleavage, thus inducing the dissociation of the complex, and leaving the PNA strand free to recognize another oligonucleotide sequence.

To evaluate the possibility to implement this catalytic cycle, we evaluated the melting profile of the PNAs in the presence of different DNA oligonucleotides, simulating the different complexes to be formed (sequence show in Table 5.1).

Table 5.2: DNA sequences used of the estimation of the catalytic behavior of the PNA probes. The cleavage site was chosen in the bond flanking the modified monomer, and it is denoted in DNA 4.1 as a tilde. Double underlined bases are those complementary to both PNAs, whereas single underlined bases are those complementary only to PNA 4.2.

Name	DNA sequence	DNA type
DNA 4.1	5'-CTG TGC GTG TGA C~AG CGG CTG A-3'	Full length DNA
DNA 4.2	5'-AGC GGC TGA-3'	3'-side after cleavage
DNA 4.3	5'-CTG TGC GTG TGA C-3'	5'-side after cleavage

The cleavage site was chosen arbitrarily between all the possible neighboring bonds, in order to obtain the hypothetically more stable fragment that can be generated after hydrolysis (longer fragment formed). UV melting analysis were conducted in PBS buffer at 5μM strand concentration, for evaluating the stability of the three different complexes and a fourth condition in which **DNA 4.2** and **DNA 4.3** were contemporary present. Denaturation profiles are shown in Figure 5.7.

From the melting profile it is possible to see the effect of the stabilization induced by the presence of the polycationic chain on stability; in fact **PNA 4.1-2**:**DNA 4.1** complexes showed a higher melting temperature if compared to theoretical values (64°C and 73°C respectively)[19].

Looking back at **PNA 4.1**, the cleavage of the bond in the middle of the DNA sequence would induce a strong drop in the complex stability for both fragments, introducing the possibility to use this PNA in a catalytic fashion in biological compatible conditions (37°C). For **PNA 4.2** is possible to see a similar behaviour as for as the stability drop, but the higher stability induced by the second polycationic chain on the complex formed with one of the two fragment (**DNA 4.2**) will prevent its

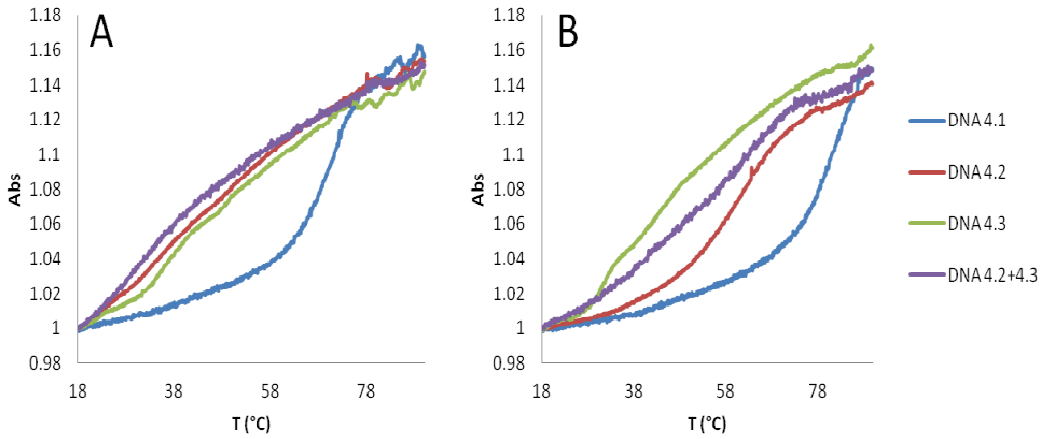


Figure 5.7: normalized UV melting curves for the complex formed between PNA 4.1 (A) and PNA 4.2 (B) in the presence of DNA 4.1 (blue lines), DNA 4.2 (red lines), DNA 4.3 (green lines) and an equimolar quantity of DNA 4.2 and DNA 4.3 (purple lines). All measurements were done in PBS buffer, pH 7; concentration of each strand was 5 μM.

recycle in the catalytic cycle at biological temperature; however the strong denaturating environment present inside the cells can help in the destabilization of the complex allowing to use also this derivative as a catalytic RNase model. Reasons of this increased stability toward **DNA 4.2** can be explained by a sort of clamping effect of the two polycationic DETA chains toward the DNA backbone, beside the increased stability due to the formation of two additional G-C pairing. The melting temperature relative to all complexes are summarized in Table 5.3

Table 5.3: melting temperatures of the PNA:DNA complexes.

DNA	PNA 4.1	PNA 4.2
DNA 4.1	71.6 °C	82.1 °C
DNA 4.2	33.8 °C	31.5 °C
DNA 4.3	35.8 °C	63.0 °C
DNA 4.2 + DNA 4.3	30.3 °C	61.5 °C

5.3 - Conclusions

On the basis of the results obtained in chapter 4, concerning the use of a nucleic acid strands as template for the reaction between two different strands, we moved forward through the use of RNA strands as a substrate of the reaction, combining the possibility to use a reactive center in connection with the ability of the PNA strand to selectively recognize the molecular target. In this contest we proposed the construction of two different systems bearing a DETA derivative, which

already showed RNase ability[6b], using a modification approach on solid phase that can be rapidly used for the generation of library of compounds without the need of solution synthesis of a huge number of different monomers.

We showed the synthesis of some of the building blocks needed for the construction of the proposed system, and a general method for a fast solid-phase introduction of the modification. Moreover was tested the effective possibility of two different PNA systems to undergo catalytic recycling after the cleavage of a phosphodiester bond, was tested using a model system based on DNA. A remarkable result of this explorative preliminary study was the finding that the introduction of a single polycationic catalytic center induced a moderately high increase in the PNA:DNA complex stability.

Future developments will be addressed with a new synthon for the synthesis of derivative **4.6** and with the test of the effective activity of the different proposed models as effective catalytic RNase mimics.

5.4 - Experimental Section

General. Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, Link, ASM and used without further purification. All reactions were carried out under a nitrogen atmosphere with dry solvents under anhydrous conditions, unless otherwise noted. Anhydrous solvents were obtained by distillation or anhydrification with molecular sieves. Reactions were monitored by TLC carried out on 0.25mm E. Merck silica-gel plates (60F-254) by using UV light as visualizing agent and ninhydrin solution and heat as developing agents. E. Merck silica gel (60, particle size 0.040÷0.063 mm) was used for flash-column chromatography. NMR spectra were recorded on Bruker Avance 400, Bruker Avance 300 or Brucker AC 300 instruments and calibrated by using residual undeuterated solvent as an internal reference. The following abbreviations were used to explain the multiplicities: s=singlet, d=doublet, t=triplet, q=quartet, m=multiplet, and br=broad. IR spectra were measured using a FT-IR Thermo Nicolet 5700, in transmission mode using KBr or NaCl. HPLC-UV-MS were recorded by using a Waters Alliance 2695 HPLC with Micromass Quattro microAPI spectrometer, a Waters 996 PDA and equipped with a Phenomenex Jupiter column (250x4.6mm, 5µ, C18, 300Å) (method A, 5 minutes in H_2O 0.2% FA, then linear gradient to 50% MeCN 0.2% FA in 30 minutes at a flow rate of 1ml/min). PNA oligomers were purified with RP-HPLC using a XTerra Prep RP_{18} column (7.8x300mm, 10µm) (method B, linear gradient from H_2O 0.1% TFA to 50% MeCN 0.1 % TFA in 30 minutes at a flow rate of 4.0ml/min).

N-Boc-N'-(2-aminoethyl)ethane-1,2-diamine (4.8). To a vigorous stirred solution of N'-(2-aminoehtyl)ethane-1,2-diamine (7.69ml, 71.18mmol) in 15ml DCM a solution of Boc-anhydride (2.22g, 10.2mmol) in 50ml DCM was added dropwise over 10h. After 20h the solvent was removed under reduced pressure, the resulting oil was taken up with 100ml of saturated $NaHCO_3$ and the product was extracted with DCM (3x100ml). The combined organic fractions were dried over Na_2SO_4 and the solvent removed under reduced pressure to give **4.8** as a pale yellow oil (1.85g, 90%); **^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 6.71 (1H, br s), 3.32÷2.76 (11H, m), 1.41 (9H, s); **^{13}C NMR (DMSO-d^6, 75MHz) δ(ppm):** 156.0, 77.9, 52.7, 49.3, 49.0, 41.9, 28.6; **MS (ESI, MeOH):** *m/z* calcd for $C_9H_{21}N_3O_2$ [M]: 203.16338, found: 204.5 $[M+H]^+$, 226.5 $[M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 204.17065 $[C_9H_{22}N_3O_2]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3608.3 (w), 3488.7 (w), 32886 (w), 2932.2 (m), 2876.6 (m), 1694.9 (s), 1594.5 (m), 1392.4 (w), 1366.6 (s).

N-Boc-N'-Cbz-N'-(2-Cbz-aminoethyl)ethane-1,2-diamine (4.9). To a solution of **4.8** (1.33g, 6.55mmol) in 40ml DCM, Cbz-Cl (1.87ml, 13.10mmol) and DIPEA (2.28ml, 13.10mmol) were added at 0°C. The mixture was left to stir at 0°C for 30 minutes and for further 3h at r.t.. The organic solvent was then removed and the resulting oil was taken up with 100ml AcOEt, transferred to a separatory funnel and washed with saturated $KHSO_4$ (3x100ml), H_2O (100ml) and brine (100ml). The organic layer was dried over Na_2SO_4 and evaporated under reduced pressure. Flash chromatography (from AcOEt/hexane 3:7 to AcOEt/hexane 5:5) yielded **4.9** as an oil (2.02g, 65%); **TLC (AcOEt) Rf:** 0.53; **^{1}H NMR ($CDCl_3$, 300MHz) δ(ppm):** 7.61÷7.36 (10H, m), 5.43 (1H, br s), 5.09 (4H, br m), 4.97 (1H, br m), 3.35 (8H, br m), 1.26 (9H, s); **^{13}C NMR ($CDCl_3$, 75MHz) δ(ppm):** 156.8, 136.3, 128.5, 128.4, 128.1, 128.0, 127.9, 79.5, 67.5, 66.6, 53.1, 40.0, 39.4, 28.3; **MS (ESI, MeOH):** *m/z* calcd for $C_{25}H_{33}N_6O_3$ [M]: 471.23694, found: 472.4 $[M+H]^+$, 494.5 $[M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 472.24421 $[C_{25}H_{34}N_3O_6]^+$, 494.22670 $[C_{25}H_{33}N_3O_6Na]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3135.8 (m), 2918.3 (w), 2849.4 (w), 1695.4 (s), 1455.5 (w), 1418.2 (w).

N-Cbz-N-(2-Cbz-aminoethyl)ethane-1,2-diamine trifluoroacetate salt (4.10). A solution of **4.9** (2.01g, 4.26mmol) was solubilized in 30ml DCM and cooled to 0°C with an ice bath, then 20ml TFA were added. The solution was left to stir for 5 minutes at 0°C and for 1h at r.t.. The solvent was then co-evaporated under reduced pressure in presence for MeOH and $CHCl_3$ to gave **4.10** as oil in quantitative yield; **TLC (AcOEt/hexane 1:1) Rf:** 0.00; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 7.89 (1H, br m),7.42÷7.28 (10H, m), 5.32 (2H, s), 4.59 (3H, br m), 3.59÷3.21 (8H, br m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 161.9, 161.6, 157.5, 1362, 135.3, 128.6, 128.5, 128.3, 128.2, 116.2, 68.2, 66.9, 48.6, 46.9, 39.8, 39.5; **MS (ESI, MeOH):** *m/z* calcd for $C_{20}H_{25}N_3O_4$ [M]: 371.18451, found: 372.3 $[M+H]^+$, 394.4 $[M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 372.19178 $[C_{20}H_{26}N_3O_4]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3341.7 (w), 2923.7 (w), 1682.2 (s), 1479.9 (w), 1455.5 (w), 1426.1 (w).

Ethyl N-Cbz-N-(2-Cbz-aminoethyl)ethane-1,2-diamine-N'-ureidoacetate (4.11). To a solution of **4.10** (0.96g, 1.98mmol) in 50ml DCM K_2CO_3 (0.27g, 1.98mmol) and ethyl isocyanato-acetate (226µl, 1.98mmol) were added. The mixture was left to react overnight and subsequently quenched with 100ml H_2O, transferred

into a separatory funnel and the organic layer was then washed with saturated $KHSO_4$ (2x100ml) and brine (100ml). The organic layer was dried over Na_2SO_4 and evaporated under reduced pressure. Flash chromatography (from AcOEt/hexane 5:5 to AcOEt) yielded **4.11** as an oil (0.78g, 79%); **TLC (AcOEt) Rf:** 0.32; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 7.54÷7.36 (10H, m), 5.38÷5.26 (2H, br m), 5.11 (4H, s), 4.22÷4.11 (4H, m), 3.50÷3.33 (10H, br m), 1.28 (3H, t, J= 7.0Hz); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 171.2, 136.7, 128.5, 128.2, 128.0, 67.5, 66.9, 61.1, 50.1, 49.1, 42.1, 39.9, 37.2, 14.1; **MS (ESI, MeOH):** *m/z* calcd for $C_{25}H_{32}N_4O_7$ [M]: 500.22710, found: 501.4 $[M+H]^+$, 523.4 $[M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 501.23492 $[C_{25}H_{33}N_4O_7]^+$, 524.22469 $[C_{25}H_{33}N_4O_7]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3.387.4 (s), 1586.6 (s), 1121.7 (m).

N-Cbz-N-(2-Cbz-aminoethyl)ethane-1,2-diamine-N'-ureidoacetic acid (4.3). To a solution of **4.11** (748mg, 1.49mmol) in 10ml THF at 0°C a solution of NaOH (597mg, 14.9 mmol) in 10ml H_2O was added. The mixture was left to stir for 30 minutes at 0°C then allowed to warm to r.t. and react overnight. The organic fraction was the removed under reduced pressure and the pH lowered to 2.5. The solution containing a precipitate was cooled for 2h at 4°C. **4.3** was collected as a white solid through Buchner filtration (651mg, 92%); **TLC (AcOEt) Rf:** 0.00, **MP (°C):** decompose with melting at 153°C; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 12.42 (1H, br s), 7.47÷7.30 (11H, m), 6.27÷6.12 (2H, br m), 5.10÷4.90 (4H, m), 3.69 (2H, s), 3.15÷2.99 (8H, br m); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 172.9, 158.4, 156.7, 156.0, 137.6, 137.4, 128.1, 127.7, 66.6, 65.7, 49.1, 48.3, 47.8, 47.3, 40.6; **MS (ESI, MeOH):** *m/z* calcd for $C_{23}H_{28}N_4O_7$ [M]: 472.19580, found: 495.3 $[M+Na]^+$, 471.4 $[M+H]^-$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 471.18829 $[C_{23}H_{27}N_4O_7]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3321.7 (m), 3143.8 (w), 2931.8 (w), 1765.7 (s), 1682.1 (s), 1623.8 (s), 158.0 (s).

N-methoxy-N-methyl-N$^{\alpha}$-Boc-N$^{\omega}$-Fmoc-L-lysinamide (4.13). In a round bottom flask N$^{\alpha}$-Boc-N$^{\omega}$-L-lysine (5.27g, 11.2mmol) was dissolved in 50ml DMF, cooled down to 0°C with an ice bath then HBTU (4.79g, 12.6mmol) and DIPEA (6.0ml, 34.5mmol) were added. The reaction was left to react for 15 minutes at 0°C then allowed to warm to r.t. and react for other 15 minutes. Finally N-methoxy-N-methyl-hydroxylamine (2.24g, 23.0mmol) was added and the reaction left to stir for further 3.5h, then the DMF was evaporated under reduced pressure and the resulting oil was taken up with AcOEt (400ml), transferred in a separatory funnel and washed with saturated $KHSO_4$ (2x500ml), saturated $NaHCO_3$ (2x500ml) and brine (500ml). The organic layer was dried over Na_2SO_4 and removed under reduced pressure to afford **4.13** as a white foamy solid (5.69g, 99%); **TLC (AcOEt) Rf:** 0.57, **MP (°C):** 52.3÷56.0°C; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 7.78 (2H, d, J= 7.4Hz), 7.62 (2H, d, J= 7.2Hz), 7.41 (2H, t, J= 7.4Hz), 7.30 (2H, t, J= 6.8Hz), 5.30 (1H, d, J= 8Hz), 5.04 (1H, s), 4.69 (1H, s), 4.39 (2H, d, J= 6.8Hz), 4.22 (1H, t, J= 6.7), 3.77 (3H, s), 3.21 (5H, br s), 1.72 (1H, m), 1.54 (2H, m), 1.45 (11H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 173.1, 156.5, 155.7, 144.0, 141.3, 127.7, 127.0, 125.1, 119.9, 79.6, 66.5, 61.6, 50.1, 47.3, 40.8, 32.1, 29.3, 28.4, 22.5; **MS (ESI, MeOH):** *m/z* calcd for $C_{28}H_{37}N_3O_6$ [M]: 511.26824, found: 534.3 $[M+Na]^+$, 550.3 $[M+K]^+$, 1045.8 $[2M+Na]^+$, 1061.7 $[2M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 534.2569 $[C_{28}H_{37}N_3O_6Na]^+$; **FT-**

IR (KBr) ν(cm^{-1}): 3342.4 (m), 3007.0 (w), 2975.3 (m), 2938.7 (m), 2864.0 (m), 1707.4 (s), 1655.0 (s), 1523.2 (s), 1390.5 (m), 1366.2 (m), 1248.2 (s), 1169.0 (s), 759.8 (s), 741.6 (s).

N$^{\alpha}$-Boc-N$^{\omega}$-Fmoc-L-lysinal (4.14). In a 3 neck round bottom flask $LiAlH_4$ (0.164g, 4.10mmol) was solubilized in 4.1ml THF and cooled down to 0°C with an ice bath, **4.13** (1.75g, 3.42mmol) was added dropwise over 1h dissolved in 80ml THF. The reaction mixture was left reacting for further 40 minutes in which the solution turn orange. The reaction was then quenched with 65ml saturated $KHSO_4$ and the organic fraction evaporated under reduced pressure. To the resulting water solution 200ml of DCM were added, the mixture transferred into a separatory funnel and the organic layer washed with saturated $KHSO_4$ (2x200ml), saturated $NaHCO_3$ (2x200ml) and brine (200ml). The organic layer was then over Na_2SO_4 and removed under reduced pressure. Flash chromatography (from AcOEt/hexane 4:6 to AcOEt/hexane 5:5) yielded **4.14** as a white foamy solid (0.76g, 49%); **TLC (AcOEt) Rf:** 0.76; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 9.53 (1H, s), 7.73 (2H, d, J= 7.4Hz), 7.52 (2H, d, J= 7.4Hz), 7.30 (2H, t, J= 6.3Hz), 7.27 (2H, t, J= 7.2Hz), 5.42 (1H, d, J= 6.2Hz), 5.22 (1H, br s), 4.40 (2H, d, J= 5.4Hz), 4.15 (2H, m), 3.15 (2H, m), 1.61 (1H, m), 1.45 (14H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 199.9, 156.6, 155.7, 143.9, 141.3, 127.7, 127.1, 125.0, 119.9, 80.2, 66.6, 59.6, 47.3, 40.4, 29.6, 28.7, 28.3, 22.2; **MS (ESI, MeOH):** *m/z* calcd for $C_{26}H_{32}N_2O_5$ [M]: 452.23112, found: 507.3 $[M+Na]^+$, 523.3 $[M+K]^+$, 992.7 $[2M+Na]^+$, 1008.7 $[2M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 475.24554 $[C_{26}H_{32}N_2O_5Na]^+$.

N$^{\alpha}$-Boc-Ψ-[L-lysin-(N$^{\omega}$-Fmoc)]-glycine methylester (4.15). In a round bottom flask glycine methylester hydrochloride (177mg, 2.82mmol) and **4.14** (1.32g, 2.35mmol) were solubilized in 30ml MeOH. The temperature was lowered to 0°C with an ice bath and $NaBH_3CN$ (169mg, 2.82mmol) and acetic acid (158µl, 2.82mmol) were added. The reaction mixture was left to react for 40 minutes at 0°C and for further 4.5h at r.t.. The solvent was then removed under reduced pressure, the resulting oil was taken up with 200ml AcOEt, transferred in a separatory funnel and washed with saturated $NaHCO_3$ (3x100ml) and brine (100ml). The organic layer was then dried over Na_2SO_4 and removed under reduced pressure. Flash chromatography (from AcOEt/hexane 7:3 to AcOEt) yielded **4.15** as a white foamy solid (889mg, 72%); **TLC (AcOEt) Rf:** 0.28, **MP (°C):** decompose before melt at 110.5°C; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 7.75 (2H, d, J= 7.4Hz), 7.59 (2H, d, J= 7.2Hz), 7.39 (2H, t, J= 7.3Hz), 7.31 (2H, t, J= 7.4Hz), 5.13 (1H, s), 4.85 (1H, s), 4.38 (2H, d, J= 6.9Hz), 4.21 (1H, t, J= 6.6Hz), 3.70 (3H, s), 3.66 (1H, m), 3.44 (2H, d, J= 17.3Hz), 3.39 (1H, d, J= 17.5), 3.18 (2H, m), 2.63 (2H, m), 2.02 (1H, m), 1.44 (15H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 172.9, 155.9, 156.3, 144.0, 141.3, 127.7, 127.0, 125.0, 119.9, 79.2, 66.5, 52.9, 51.8, 50.7, 50.1, 47.3, 40.7, 32.8, 29.6, 28.4, 22.9; **MS (ESI, MeOH):** *m/z* calcd for $C_{29}H_{39}N_3O_6$ [M]: 525.28389, found: 526.4 $[M+H]^+$, 548.3 $[M+Na]^+$, 564.3 $[M+K]^+$, 1051.7 $[2M+H]^+$, 1073.8 $[2M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 534.2569 $[C_{29}H_{40}N_3O_6]^+$; **FT-IR (KBr) ν(cm^{-1}):** 3342.3 (s), 2979.6 (s), 2915.6 (m), 2876.2 (m), 1743.9 (s), 1686.3 (s), 1677.9 (s), 1530.4 (s), 1437.6 (m), 1368.0 (m), 1264.4 (s), 1248.3 (s), 1170.9 (s), 761.2 (s), 738.2 (s).

Boc-PNA(5L-Lys$_{(Fmoc)}$)-T-OMe (4.16). In a sealed test tube 1-thymineacetic acid (1.11g, 6.03mmol) was solubilized together with EDC·HCl (1.16g, 6.03mmol) and DhBtOH (0.98g, 6.03mmol) in 6ml DMF at 0°C, then DIPEA (1.05ml, 6.03mmol) was added and the mixture was left to stir for 30 minutes, and transferred in a round bottom flask containing **4.15** (1.59g, 3.02mmol) dissolved in 4ml DMF. The reaction was allowed to react for further 3h, then the reaction was then poured into AcOEt (700ml) and washed with saturated $KHSO_4$ (2x500ml), saturated $NaHCO_3$ (2x500ml) and brine (2x300ml). The organic layer was then dried over Na_2SO_4 and removed under reduced pressure. Flash chromatography (from AcOEt to AcOEt/MeOH 95:5) yielded **4.16** as a white foamy solid (1.65g, 80%); **TLC (AcOEt/MeOH 95:5) Rf:** 0.28, **MP (°C):** decompose without melting at 87.2°C; **^{1}H NMR (DMSO-d^6, 400MHz, major rotamer) δ(ppm):** 11.29 (1H, s), 7.88 (2H, d, J= 7.5Hz), 7.68 (2H, d, J= 7.4Hz), 7.42 (2H, t, J= 7.4Hz), 7.33 (2H, t, J= 7.4Hz), 7.28 (1H, m), 7.22(1H, s), 6.80 (1H, m), 4.60 (1H, d, J= 16.8), 4.53 (1H, d, J= 16.4Hz), 4.29 (2H, d, J= 6.8Hz), 4.21 (1H, t, J= 6.7Hz), 4.08 (1H, d, J= 17.2Hz), 4.01 (1H, d, J= 17.3Hz), 3.62 (3H, s), 3.52 (1H, m), 3.27 (2H, m), 2.99 (2H, m), 1.76 (3H, s), 1.38 (6H, m), 1.36 (9H, m); **^{13}C NMR (DMSO-d^6, 100MHz, major rotamer) δ(ppm):** 169.9, 167.9, 164.8, 156.5, 156.1, 151.4, 144.4, 142.2, 141.2, 128.1, 127.5, 125.6, 120.6, 108.6, 78.4, 65.6, 52.7, 52.2, 49.2, 48.4, 48.2, 47.3, 31.5, 29.7, 28.7, 23.4, 12.4; **MS (ESI, MeOH):** *m/z* calcd for $C_{36}H_{45}N_5O_9$ [M]: 691.32173, found: 714.4 $[M+Na]^+$, 730.4 $[M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 690.3113 $[C_{36}H_{44}N_5O_9]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3350.0 (m), 3063.6 (w), 3008.1 (w), 2935.9 (m), 2863.7 (m), 1744.4 (m), 1687.5 (s), 1521.8 (m), 1469.2 (m), 1451.3 (m), 1366.5 (m), 1248.7 (s), 1170.9 (m), 759.8 (m), 742.6 (m).

Boc-PNA(5L-Lys$_{(Fmoc)}$)-T-OH (4.4). In a round bottom flask **4.16** (1.71g, 2.47mmol) was solubilized in 90ml THF at 0°C then a solution of $Ba(OH)_2$·8 H_2O (0.86g, 2.73mmol) in 90ml H_2O was added and the mixture was left to stir for 25 minutes before quench the reaction with 1M HCl (5.5ml). The organic solvent was removed under reduced pressure, the pH adjusted to 2.5. **4.4** was collected over buckner as dirty white solid and dried under reduced pressure over P_2O_5 (1.34g, 80%); **TLC (AcOEt/MeOH 9:1) Rf:** 0.00, **MP (°C):** 206.3÷211.4°C; **^{1}H NMR (DMSO-d^6, 400MHz, major rotamer) δ(ppm):** 11.30 (1H, s), 7.88 (2H, d, J= 7.5Hz), 7.68 (2H, d, J= 7.4Hz), 7.41 (2H, t, J= 7.4Hz), 7.33 (2H, t, J= 7.4Hz), 7.28 (1H, m), 7.21(1H, s), 6.80 (1H, m), 4.72 (1H, d, J= 16.6), 4.62 (1H, d, J= 16.7Hz), 4.29 (2H, d, J= 6.5Hz), 4.21 (1H, t, J= 6.3Hz), 3.93 (1H, d, J= 17.1Hz), 3.89 (1H, d, J= 17.1Hz), 3.51 (1H, m), 3.22 (2H, m), 2.97 (2H, m), 1.75 (3H, d, J= 3.04Hz), 1.36 (13H, m), 1.26 (2H, m); **^{13}C NMR (DMSO-d^6, 100MHz, major rotamer) δ(ppm):** 169.9, 167.9, 164.8, 156.6, 156.1, 151.5, 144.4, 142.3, 141.2, 128.1, 127.5, 125.6, 120.6, 108.6, 78.4, 65.7, 55.4, 51.9, 48.1, 47.2, 31.6, 29.7, 28.7, 23.3, 12.4; **MS (ESI, MeOH):** *m/z* calcd for $C_{35}H_{43}N_5O_9$ [M]: 677.30608, found: 700.4 $[M+Na]^+$, 716.4 $[M+K]^+$, 1378.0 $[2M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 676.30176 $[C_{35}H_{42}N_5O_9]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3403.4 (m), 3345.5 (m), 3035.6 (w), 3064.8 (w), 2977.2 (m), 2861.4 (m), 1689.8 (s), 1530.7 (m), 1476.8 (m), 1451.3 (m), 1367.6 (s), 1252.0 (m), 1170.4 (m), 759.2 (m), 741.7 (m).

Naphtalene-2,7-diyl bis(trifluoromethanesulfonate) (4.18). 2,7-dihydroxynaphtalene (0.96g, 6.00mmol) was solubilized with 30ml DCM in a round bottom flask. Triethylamine (2.47ml, 17.72mmol) was added, the reaction mixture cooled to -20°C with a $CaCl_2$/ice bath and trifluomethanesulfonic anhydride (5,00g, 17.22mmol) was added dropwise over 10 minutes. After 2 hours the reaction was quenched with 30ml H_2O, transferred in a separatory funnel and extracted with DCM (2x25ml). The combined organic layers were dried over Na_2SO_4 and removed under reduced pressure. Flash chromatography (AcOEt/hexane 1:9) yielded **4.18** as a white solid (2.42g, 95%); **TLC (AcOEt/hexane 1:1) Rf:** 0.72, **MP (°C):** 60.2÷60.7°C; **^{1}H NMR ($CDCl_3$, 300MHz) δ(ppm):** 8.00 (2H, d, J= 8.8 Hz), 7.81 (2H, d, J= 2.4Hz), 7.48 (2H, dd, J= 8.8Hz, 2.4Hz); **^{13}C NMR ($CDCl_3$, 75MHz) δ(ppm):** 148.2, 133.5, 131.2, 103.7, 121.1, 119.4, 118.7 (q, J= 319Hz); **EI-MS:** *m/z* calcd for $C_{12}H_6O_6F_6S_2$ [M]: 423.95100, found: 291 (100%) $[M\text{-}SO_2CF_3]^+$, 424 (92%) $[M^\cdot]^+$, 130 (89%) $[M\text{-}2SO_2CF_3\text{-}CO]^+$, 199 (64%) $[M\text{-}SO_2CF_3\text{-}CO\text{-}SO_2]^+$, 102 (61%) $[M\text{-}2SO_2CF_3\text{-}2CO]^+$, 263 (45%) $[M\text{-}SO_2CF_3\text{-}CO]^+$, 69 (38%) $[CF_3^\cdot]^+$, 275 (19%) $[M\text{-}SO_2CF_3]^+$; **Elemental composition:** calcd %C 33.97, %H 1.43, %N 0.00, found %C 33.78, %H 1.23, %N 0.00; **FT-IR (KBr) ν(cm^{-1}):** 3095 (w), 1414 (s), 1255÷1180 (s), 1145÷1110 (s).

Dimethyl naphtalene-2,7-dicarboxylate (4.19). In a tree neck round bottom flask **4.18** (2.39g, 5.62mmol), triphenilphosphine (295mg, 1.12mmol), palladium acetate (126mg, 0.56mmol), triethylamine (3.13ml, 22.49mmol) and methanol (11.4ml, 281mmol) were solubilized in 15ml DMF, heated to 60°C and CO was bubbled in the reaction mixture for 6 hours. The reaction was then quenched with 150ml 0.1M $KHSO_4$ and extracted with AcOEt (3x50ml). The combined organic layers were dried over Na_2SO_4 and removed under reduced pressure. Flash chromatography (AcOEt/hexane 2:8) yielded **4.18** as a light brown solid (589mg, 43%); **TLC (AcOEt/hexane 2:8) Rf:** 0.23, **MP (°C):** 130.4÷131.9°C; **^{1}H NMR ($CDCl_3$, 300MHz) δ(ppm):** 8.70 (2H, d, J= 1.1Hz), 8.16 (2H, dd, J= 8.5Hz, 1.1Hz), 7.92 (2H, d, J= 8.5Hz), 3.99 (6H, s); **^{13}C NMR ($CDCl_3$, 75MHz) δ(ppm):** 166.7, 137.4, 132.2, 131.6, 128.2, 128.0, 127.5, 52.3; **EI-MS:** *m/z* calcd for $C_{14}H_{12}O_4$ [M]: 244.07256, found: 231 (100%) $[M\text{-}OCH_3]^+$, 244 (77%) $[M^\cdot]^+$, 185 (29%) $[M\text{-}COOCH_3]^+$, 126 (15%) $[M\text{-}2COOCH_3]^+$, 154 (13%) $[M\text{-}COOCH_3\text{-}OCH_3]^+$, 170 (7%) $[M\text{-}COOCH_3\text{-}CH_3]^+$; **Elemental composition:** calcd %C 68.85, %H 4.65, %N 0.00, found %C 68.21, %H 4.71, %N 0.00; **FT-IR (KBr) ν(cm^{-1}):** 3075 (w), 3031 (w), 2965 (w), 1724 (s).

Naphtalene-2,7-dicarboxylic acid (4.5). In a round bottom flask **4.19** (254mg, 1.04mmol), was dispersed in 20ml MeOH, then 10ml of a 1M solution of NaOH were added. The reaction was left to react for 24 hours with changes from yellow to orange solution. The organic solvent was then evaporated under reduced pressure, the resulting aqueous phase diluted with 40ml $KHSO_4$ and extracted with AcOEt (3x150ml). The combined organic layers were dried over Na_2SO_4 and removed under reduced pressure to afford **4.18** as a light brown solid (191mg, 85%); **TLC (DCM/MeOH 9:1 + 1% AcOH) Rf:** 0.38; **MP (°C):** decompose without melt at 290°C; **^{1}H NMR (DMSO-d^6, 300MHz) δ(ppm):** 11.20 (2H, s), 8.78 (2H, br s), 8.10 (4H, br s); **^{13}C NMR (DMSO-d^6, 75MHz) δ(ppm):** 167.0, 136.6, 131.7, 131.3, 128.8, 128.0, 127.2; **MS (ESI, MeOH):** *m/z* calcd for $C_{12}H_8O_4$ [M]: 216.04226, found: 215.1 $[M\text{-}H]^-$; **Elemental composition:** calcd %C

66.67, %H 3.73, %N 0.00, found %C 66.48, %H 3.90, %N 0.05; **FT-IR (KBr) ν(cm^{-1}):** 2989 (m), 2827 (m), 1685 (s).

1,3-dibeziloxycarbonil-2-methylisothiourea (4.21). In a round bottom flask S-methylisothiourea (1.00g, 3.61mmol) was dispersed in 25ml DCM and cooled to 0°C with an ice bath. Cbz-Cl (2.59ml, 18.2mmol) was then added to the solution followed by addition of a 1M solution of NaOH in water (25.2ml, 25.2mmol). The reaction mixture was left to react for 24 hours, then diluted with 20ml DCM and washed with water (3x90ml). The organic layer was dried over Na_2SO_4 and the solvent evaporated under reduced pressure. Flash chromatography (AcOEt/hexane 1:5) yielded **4.21** as a white solid (1.92g, 74%); **TLC (AcOEt/hexane 1:5) Rf:** 0.30; **MP (°C):** 52.5÷55.0°C; **^{1}H NMR (DMSO-d^6, 400MHz) δ(ppm):** 11.87 (1H, s), 7.38 (10H, m), 5.23 (2H, s), 5.21 (2H, s), 2.44 (3H, s); **^{13}C NMR (DMSO-d^6, 100MHz) δ(ppm):** 172.9, 161.1, 151.6, 135.8, 134.6, 128.8, 128.7, 128.6, 128.3, 68.4, 68.1, 14.6; **MS (ESI, MeOH):** *m/z* calcd for $C_{18}H_{18}N_2O_4S$ [M]: 358.09875, found: 359.2 $[M+H]^+$, 381.2 $[M+Na]^+$, 397.1 $[M+K]^+$, 739.3 $[2M+Na]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 359.1060 $[C_{18}H_{19}N_2O_4S]^-$; **FT-IR (KBr) ν(cm^{-1}):** 3032 (w), 2966 (w), 2925 (w), 1750 (s), 1657(m), 1575 (s), 1456 (m), 1445 (m), 1418 (m), 1380 (m), 1192 (s).

(di-Cbz-guanidino)-Lys(Fmoc)-OMe (4.22). In a round bottom flask H-Lys$_{(Fmoc)}$-OMe·HCl (399mg, 0.952mmol) and **4.21** (312mg, 0.871mmol) were dissolved in 10ml DMF, followed by the addition of DIPEA (383µl, 2.20mmol) and mercury chloride (238mg, 0.879mmol). The reaction mixture was left to react overnight, then diluted with 200ml AcOEt and washed with saturated $KHSO_4$ (2x200ml), $NaHCO_3$ (2x200ml) and brine (200ml). The organic layer was dried over Na_2SO_4 and the solvent removed under reduced pressure. Flash chromatography (AcOEt/hexane 1:2) yielded **4.22** as a white solid (524mg, 87%); **TLC (AcOEt/hexane 1:5) Rf:** 0.30; **^{1}H NMR ($CDCl_3$, 400MHz) δ(ppm):** 11.72 (1H, s), 8.84 (1H, d, J= 7.5Hz), 7.78 (2H, d, J= 7.7Hz), 7.61 (2H, d, J= 7.4Hz), 7.48÷7.24 (14H, m), 5.20 (2H, s), 5.13 (2H, s), 4.84 (2H, m), 4.42 (2H, d, J= 6.8Hz), 4.24 (1H, t, J= 6.1Hz), 3.77 (3H, s), 3.20 (2H, q, J= 6.0Hz), 1.96÷1.92 (1H, m), 1.85÷1.75 (1H, m), 1.65÷1.50 (2H, m), 1.43÷1.27 (4H, m); **^{13}C NMR ($CDCl_3$, 100MHz) δ(ppm):** 171.9, 163.5, 156.4, 155.6, 153.7, 144.0, 141.3, 136.7, 134.5, 128.8, 128.7, 128.5, 128.4, 128.1, 128.0, 127.7, 127.0, 125.0, 120.0, 68.3, 67.3, 66.5, 53.3, 52.5, 47.3, 40.7, 31.7, 29.5, 22.2; **MS (ESI, MeOH):** *m/z* calcd for $C_{39}H_{40}N_4O_8$ [M]: 678.26896, found: $[M+Na]^+$, $[M+K]^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 693.2911 $[C_{39}H_{41}N_4O_8]^+$.

(di-Cbz-guanidino)-Lys(Fmoc)-OH (4.6). In a round bottom flask **4.22** (484mg, 0.699mmol) was solubilized in 50ml THF, cooled to 0°C with an ice bath then a solution of $Ba(OH)_2·8H_2O$ (331mg, 1.048mmol) in 50ml water was added. The mixture was left to react for 2 hour at 0°C and for further 24 hours at r.t. before quench with a diluted solution of HCl. The organic solvent was then removed and the pH adjusted to 3.5. The precipitate was left to form at 4°C for 2 hours and collected through buckner filtration to gave a mixture of compound containing **4.6** and a side product missing one Cbz protection (319mg);**TLC (AcOEt/hexane 1:5) Rf:** 0.0; **MS (ESI, MeOH):** *m/z* calcd for $C_{39}H_{40}N_4O_8$ [M]: 678.26896, found: 679.2 $[M+H]^+$, 545.3 [biproduct+H]$^+$; **HRMS (LTQ-Orbitrap, MeOH)** *m/z* found: 693.2911 $[C_{39}H_{41}N_4O_8]^+$.

Desymmetrization protocol. During the synthesis of the **PNA 4.5**, the **PNA 4.3** was obtained with monomer **2.25** incorporated. The azide group was reduced following the protocol reported in the experimental part of chapter 2, to the amine that was formed (**PNA 4.4**). After reduction **4.5** (2eq), was preactivated with DIC (15eq) and DhBtOH (15eq) for 10 min in DMF; this solution (60µl/µmol of active site) was then added to the reactor vessel, which was left to shake for 2 hours, then a solution of Fmoc-ethylenediamine bromohydrate (10eq) and DIPEA (20eq) in DMF (20µl/µmol of active site) was added and the vessel left to shake further 3 hours.

Synthesis and characterization of PNAs. The synthesis of the PNAs was performed with standard manual Boc-based solid-phase synthesis with HBTU/DIPEA as coupling agent, using the preformed monomer **4.4** and **2.25** in addition to commercially available Boc-PNA monomers on a MBHA resin loaded with Fmoc-Gly-OH as first monomer (0.2mmol/g). The introduction of the catalytic center for **PNA 4.1** and **PNA 4.2** was performed deprotecting the Fmoc group with a 20% solution of piperidine in DMF and coupling **4.3** with HBTU/DIPEA protocol (5 equivalents). **PNA 4.1** and **PNA 4.2** purification were performed by RP-HPLC with UV detection at 260nm. The purity and identity of the purified PNAs were determined by HPLC-UV-MS, for **PNA 4.4** and **PNA 4.5** the characterization was performed using HPLC-UV-MS on the crude product of the cleavage from resins. **PNA 4.1:** 23%; t_r: 13.4min; ESI-MS: *m/z* calcd 3251.3875 [M]: 1084.5 $[MH_3]^{3+}$, 813.8 $[MH_4]^{4+}$, 651.0 $[MH_5]^{5+}$, 542.6 $[MH_6]^{6+}$; **PNA 4.2:** 25%; t_r: 12.3min; ESI-MS: *m/z* calcd 4050.7825 [M]: 1351.5 $[MH_3]^{3+}$, 1013.7 $[MH_4]^{4+}$, 811.0 $[MH_5]^{5+}$, 675.8 $[MH_6]^{6+}$, 579.2 $[MH_7]^{7+}$; **PNA 4.4:** t_r: 12.7min; ESI-MS: *m/z* calcd 1989.9617 [M]: 996.0 $[MH_2]^{2+}$, 664.0 $[MH_3]^{3+}$, 498.1 $[MH_4]^{4+}$;**PNA 4.5:** t_r: 23.6min; ESI-MS: *m/z* calcd 2452.1175 [M]: 1227.0 $[MH_2]^{2+}$, 817.9 $[MH_3]^{3+}$.

UV measurements. Stock solutions of PNA and DNA synthetic oligonucleotides (Thermo-Fisher Scientific, HPLC-grade) were prepared in bidistilled water, and the PNA concentration was calculated by UV absorbance with the following extinction coefficients (ε_{260} $[M^{-1}cm^{-1}]$) for the nucleobases: T 8600, C 6600, A 13700, and G 11700. For DNA the data provided by the producer were used. From these, solutions containing single stranded PNA and DNA, PNA–DNA duplexes were prepared. Measurement conditions: [PNA]=[DNA]= 5µM in PBS (100mM NaCl, 10mM $NaH_2PO_4 \cdot H_2O$, 0.1mM EDTA, pH 7.0). All the samples were first incubated at 90°C for 5 min, then slowly cooled to room temperature. Thermal denaturation profiles (A_{260} versus T) of the hybrids were measured with a UV/Vis Lambda Bio 20 spectrophotometer equipped with a Peltier temperature programmer PTP6 interfaced to a personal computer. For the temperature range 18÷90°C, A_{260} values were recorded at 0.1°C increments. A melting curve was recorded for each duplex. The melting temperature (Tm) was determined from the maximum of the first derivative of the melting curves.

5.5 - References

[1] D'Alessio G., Riordan J.F., *Ribonucleases: Structures and Functions*, **1997** , Academic Press.

[2] Kuimelis R. G., Mc Laughlin L. W. *Chem. Rev.* **1998**, 98, 1027-1044.

[3] Murtola M., Wenska M., Strömberg R. *J. Am. Chem. Soc.* **2010**, 132 (26), 8984-8990.

[4] a) Komiyama M., Yoshinari K. *J. Org. Chem.* **1997**, 62, 2155-2160; b) Holmes S. C., Gait M. J. *Eur. J. Org. Chem.* **2005**, 5171-5183.

[5] Kalek M., Benediktson P., Vester B., Wengel J. *Chem. Commun.* **2008**,762-764.

[6] a) Verheijen J. C., Deiman B. A. L. M., Yeheskiely E., van der Marel G. A., van Boom J. H. *Angew. Chem. Int. Ed.* **2000**, 39 (2), 369-372; b) Gaglione M., Milano G., Chambery A., Moggio L., Romanelli A., Messere A. *Mol. BioSyst.* **2011**, 7, 2490-2499.

[7] a) Matsumura K., Endo M., Komiyama M. *Chem. Commun.* **1994**, 2019-2020; b) Magda D., Miller R. A., Sessler J. L., Iverson B. L. *J. Am. Chem. Soc.* **1994**, 116, 7439-7440; c) Magda D., Crofts S., Lin A., Miles D., Wright M., Sessler J. L. *J. Am. Chem. Soc.* **1994**, 119, 2293-2294.

[8] a) Bashkin J. K., Frolova E. I., Sampath U. *J. Am. Chem. Soc.* **1994**, 116, 5981-5982; b) Inoue H., Furukawa T., Tamura T., Kamada A., Ohtsuka E. *Nucleosides, Nucleotides Nucleic Acids* **2001**, 20, 833-835; c) Sakamoto S., Tamura T., Furukawa T., Komatsu Y., Ohtsuka E., Kitamura M., Inoue H. *Nucleic Acids Res.* **2003**, 31, 1416-1425.

[9] a) Putnam W. C., Bashkin J. K. *Chem. Commun.* **2000**, 767-768; b) Whitney A., Gavory G., Balasubramanian S. *Chem. Commun.* **2003**, 36-37; c) Matsuda S., Ishikubo A., Kuzuya A., Yashiro M., Komiyama M. *Angew. Chem., Int. Ed.* **1998**, 37, 3284-3286.

[10] Beloglazova N. G., Fabani M. M., Zenkova M. A., Vlassov V. V. *FEBS Lett.* **2000**, 481, 277-280.

[11] Ushijima K., Takaku H. *Biochim. Biophys. Acta* **1998**, 1379, 217-223.

[12] Niittymaki T., Lonnberg H. *Org. Biomol. Chem.* **2006**, 4, 15-25.

[13] Kuzuya A., Mizoguchi R., Machida K., Komiyama M. *J. Am. Chem. Soc.* **2002**, 124, 6887-6894.

[14] Nakano S., Uotani Y., Uenishi K., Fujii M., Sugimoto N. *J. Am. Chem. Soc.* **2005**, 127, 518-519.

[15] Kneeland D. M., Ariga K., Lynch V. M., Huang C.-Y., Anslyn E. V. *J. Am. Chem. Soc.* **1993**, 115, 10042-10055.

[16] Brimble M. A., Lai M.Y.H. *Org. Biomol. Chem.* **2003**, 1, 2084-2095.

[17] Cacchi S., Ciattini P.G., Morera E., Ortar G. *Tetrahedron Lett.* **1986**, 27(33), 3931-3934.

[18] WO 2005/079791

[19] Giesen U. Kleider W., Berding C., Geiger A., Ørum H., Nielsen P. E. *Nucl. Acids Res.* **1998**, 26 (21), 5004-5006.

Acknowledgements

First of all I would like to thank Prof. Roberto Corradini and Prof. Rosangela Marchelli for the possibility they gave me to undertake the challenging way of a Ph D. I would like to thanks them especially for their support during the three years, from the approach to the projects to the stimuli they gave me to overcome all the problems encountered in this period.

The second great thanks is for Prof. Nicolas Winssinger for the chance he gave me to join his lab in ISIS (Institut de Science et d'Ingénierie Supramoléculaires) for seven month and for the introduction to a new way to carry reactions, by the use of a nucleic acid strand. The same for Katarzyna Gorska for her help during the approach to this new field and for some measurements.

I also thank to Prof. Arnaldo Dossena, Prof. Stefano Sforza and Dr. Tullia Tedeschi for their advices and suggestion in normal laboratory life, and Prof. Roberto Gambari and his group for all the cellular experiment. A special thanks also to Dr. Andrea Faccini for the HR-MS analysis and for his patience during my introduction to the ESI-MS world.

How to forget all the people that satellite around me in the labs during these years. First the “people from 50” from former Ph.D students to bachelor and master students what worked with (for?) me, then all the “cousins from 96” for the competition/help in the use of HPLC and UPLC, and finally the “people from the 4th” (aka “the NW’s”) for the help and support they gave me in the lab and during the time I spent in Strasbourg, and together with the other “people from ISIS” helping me to enjoying it.

Finally I would say a big thanks to my parents and my sister, to all my friends (in Italy or abroad), and to all the others for their endurance during the writing of the thesis.

Contributions

Publications

1. A. Manicardi, A. Calabretta, M. Bencivenni, T. Tedeschi, S. Sforza, R. Corradini, R. Marchelli. Affinity and selectivity of C2- and C5-substituted "Chiral-box" PNA in solution and on microarrays. *Chirality* **2010**, 22(1E), 161-172.

2. R. Corradini, S. Sforza, T. Tedeschi, F. Totsingan, A. Manicardi, R. Marchelli. Peptide nucleic acid with a structurally biased backbone. Updated review and emerging challenges. *Current topics in medicinal chemistry* **2011**, 11(12), 1535-1554.

3. K. Gorska, A. Manicardi, S. Barluenga, N. Winssinger. Azide-Reduction Triggered Immolative Linker and Application to DNA-Templated Release of Functional Molecules. *ChemCommun* **2011**, 47(15), 4364-4366.

4. T. Tedeschi, A. Calabretta, M. Bencivelli, A. Manicardi, G. Corrado, R. Corradini, R. Rao, S. Sforza, R. Marchelli. A PNA microarray for tomato genotyping. *Molecular BioSystems* **2011**, 7(6), 1902-1907.

5. R. Corradini, A. Manicardi, R. Marchelli, S. Sforza, T. Tedeschi, E. Fabbri, M. Borgatti, N. Bianchi, R. Gambari. Gene modulation by Peptide Nucleic Acids (PNAs) targeting microRNA (miR). *Targets in Gene Therapy – InTech – Open Access Publisher **2011**, ISBN 978-953-307-540-2, available from: http://www.intechopen.com/articles/show/title/gene-modulation-by-peptide-nucleic-acids-pnas-targeting-micrornas-mirs-*

6. R. Gambari, E. Fabbri, M. Borgatti, I. Lampronti, A. Finotti, E. Brognara, N. Bianchi, A. Manicardi, R. Marchelli, R. Corradini. Targeting microRNAs involved in human diseases: A novel approach for modification of gene expression and drug development. *Biochem Pharmacol* **2011**, 82, 1416-1429.

7. C. De Cola, A. Manicardi, R. Corradini, I. Izzo, F. De Riccardis. Carboxyalkyl Peptoid PNAs: Synthesis and Hybridization Properties. *Thetrahedron* **2011**, 68, 499-506.

8. E. Brognara, I. Lamprotti, G. Breveglieri, A. Accetta, R. Corradini, A. Manicardi, M. Borgatti, A. Canella, C. Multineddu, R. Marchelli, R. Gambari. C(5) Modified Uracil Derivatives Showing Antiproliferative and Erythroid Differentiation Inducing Activities on Human Chronic Myelogenous Leukemia K562 Cells. *European Journal of Pharmacology* **2011**, 672 (1-3), 30-37.

9. E. Fabbri, E. Brognara, M. Borgatti, I. Lamprotti, A. Finotti, N. Bianchi, S. Sforza, T. Tedeschi, A. Manicardi, R. Marchelli, R. Corradini, R. Gambari. miRNA therapeutics: delivery and biological activity of peptide nucleic acids targeting miRNAs. *Ephigenomics* **2011**, 3 (6), 733-745.

10. E. Fabbri, A. Manicardi, T. Tedeschi, S. Sforza, N. Bianchi, E. Brognara, A. Finotti, G. Breveglieri, M. Borgatti, R. Corradini, R. Marchelli, R. Gambari. Modulation of the Biological Activity of microRNA-210 with Peptide Nucleic Acids (PNAs). *ChemMedChem* **2011**, 6 (12), 2192-2202.

Communications

1. Molecular engineering of PNA using Tailor-Made Uracil Derivatives - R. Corradini, A. Accetta, A. Manicardi, S. Sforza, T. Tedeschi, R. Marchelli - Poster presentation – 3rd EuCheMS Chemistry Congress

2. Backbone-modified Peptide Nucleic Acids: Cellular Uptake and Anti-micro-RNA activity - R. Corradini, A. Manicardi, T. Tedeschi, S. Sforza, E. Fabbri, N. Bianchi, E. Brognara, R. Gambari, R. Marchelli - Poster presentation – XXXIII CoDiCO – Congress of Organic Chemistry Division of the Italian Chemical Society

3. miR-Targeting using templated Staudinger Reaction - A. Manicardi, K. Gorska, N. Winssinger, R. Corradini, R. Marchelli - Poster presentation – XXXIII CoDiCO – Congress of Organic Chemistry Division of the Italian Chemical Society

4. miR-Targeting using templated Staudinger Reaction - A. Manicardi, K. Gorska, N. Winssinger, R. Corradini, R. Marchelli - Poster presentation – IASOC 2010 – Ischia Advanced School of Organic Chemistry

5. Use of Azide Function in PNAs Chemistry - A. Manicardi - Poster presentation – X Chemistry day of Emilia Romagna

6. New highly efficient switching fluorescent probes for DNA detection - A. Manicardi, A. Ghidini, T. Tedeschi, A. Tonelli, S. Sforza, R. Corradini, R. Marchelli - Poster presentation – BioPhotonics 2011

7. New highly efficient switching fluorescent probes for DNA detection - A. Manicardi, A. Ghidini, T. Tedeschi, A. Tonelli, S. Sforza, R. Corradini, R. Marchelli - Poster presentation – XXXVI Attilio Corbella summer school

8. Oligonucleotide signaling using PNA based systems - A. Manicardi, K. Gorska, N. Winssinger, R. Gambari, R. Corradini, R. Marchelli - Oral presentation – XXXVI Attilio Corbella summer school

9. Development of chiral arginine-based PNA microarrays for the selective identification of tomato DNA - S. Sforza, T. Tedeschi, A. Calabretta, M. Bencivenni, A. Manicardi, R. Corradini, R. Marchelli - Poster presentation – XXIV National Congress of the Italian Chemical Society

10. PNA bearing multiple pyrene units for DNA detection - Manicardi A., Ghidini, A., Tedeschi T., Sforza, S., Corradini R., Marchelli R. - Poster presentation – X CNCS – National Congress of Supramolecular Chemistry