UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
"Scienze Chimiche"

CICLO XXXIII

Peptide Nucleic Acid (PNAs) and Modified PNAs-based strategies for Advanced Diagnostic and Therapeutic applications

Coordinatore:
Chiar.ma Prof.ssa Alessia Bacchi

Tutore:
Chiar.mo Prof. Roberto Corradini
Chiar.mo Prof. Wolfgang Knoll

Approved on June 1st 2021
Prof. Alessia Bacchi

Dottorando: Martina Neri

Anni 2017/2018 – 2019/2020

Proprio nello sforzo enorme e coraggioso di vincere la fatica riusciamo a provare, almeno per un istante, la sensazione autentica di vivere. Raggiungiamo la consapevolezza che la qualità del vivere non si trova in valori misurabili in voti, numeri e gradi, ma è insita nell'azione stessa, vi scorre dentro.

[Haruki Muramaki, "L'arte di correre"]

Table of contents

Chapter 1

1 Introduction

1.1 Nucleic Acids: a general overview

All the hereditary information of living organisms is carried by the prime genetic molecule known as DNA (DeoxyriboNucleic Acid). The three-dimensional structure of this essential biological molecule was first proposed by the American biologist James D. Watson and the English physicist Francis H. C. Crick in 1953.[1] In their work the two scientists described the structure of DNA as two polynucleotide chains each coiled in a double helix shape, consisting of a phosphate group and a 2'-deoxyribose residue joined together by 3'-5' linkages. Four different nucleobases can be distinguished in DNA, specifically Adenine (A), Thymine (T), Guanine (G) and Cytosine (C) (**Figure 1.1**), which can be divided into two categories, pyrimidines (T,C) and purines (A,G). The bases are linked by a glycosidic bond to the sugar forming a nucleoside, then adding the phosphate residue by a phosphodiester linkage to a nucleoside generates a nucleotide. The two polynucleotide chains are held together in an anti-parallel orientation by hydrogen bonds occurring between pairs of nucleobases in their preferred tautomeric form, it follows that adenine (purine) can bind with thymine (pyrimidine), and guanine (purine) can bind with cytosine (pyrimidine). These non-covalent interactions, better known as Watson-Crick base pairs, highly contribute to the thermodynamic stability of the double helix structure together with a second important effect, the stacking interaction (π-π) between adjacent bases. The role of DNA double helical structure is crucial in life science not only for carrying genetic information but also in biological processes involved in DNA transcription and replication.

The RiboNucleic Acid (RNA) is a very similar molecule but differs from DNA in few respects. Firstly, RNA backbone contains the sugar ribose, which having one additional hydroxyl group on position 2' than deoxyribose makes the RNA more reactive, hence less stable, than DNA. Secondly, both DNA and RNA have four nucleobases each but thymine in DNA is replaced by Uracil (U) in RNA. Lastly, RNA is generally found as single strand helix composed of a shorter sequence of nucleotides, except for RNA present in some viruses. The reason why RNA is mostly existing as single chain is that it is regularly synthesized, used and recycled. Indeed, unlike DNA whose integrity is preserved thanks to a smaller dimension of the double helix grooves, RNA is more accessible for enzyme

attack because of the larger grooves of the single strand conformation (**Figure 1.1**).

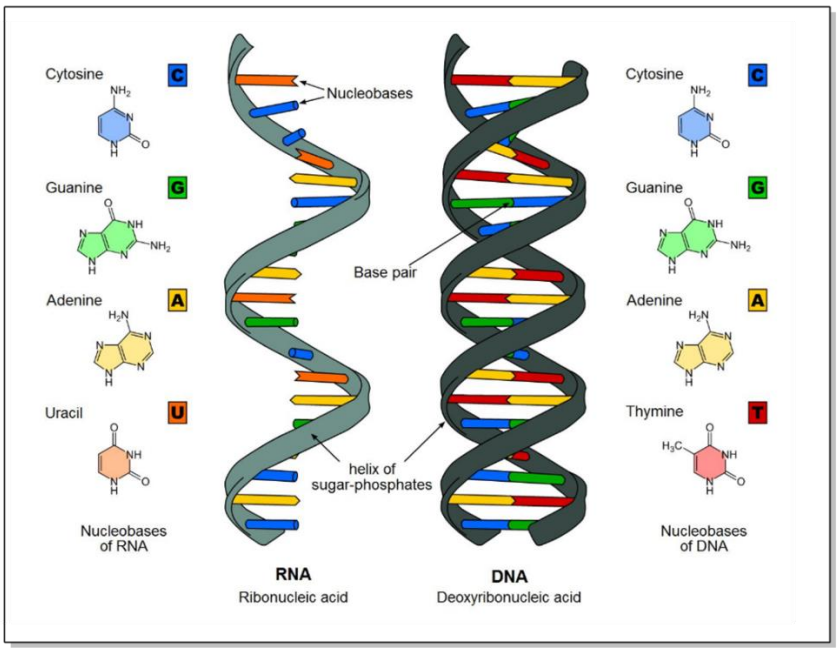


Figure 1.1. Structure of RNA and DNA.

The hybridization process, in which two complementary single-stranded nucleic acids (DNA or RNA) anneal each other to form a double-strand structure, is a fundamental part in many molecular biology techniques such as polymerase chain reaction (PCR) or Southern and Northern blots, as well as in biosensing field and for therapeutic purposes. In this respect, a first example of the potential of nucleic acid hybridization event was reported by Paterson *et al.* in 1977, who have developed a method for the identification of genes in DNA molecules based upon the inhibition of the translation phenomenon due to the binding occurring between the messenger RNA (mRNA) and its complementary DNA (cDNA) in eukaryotic cell-free system.[2] One year later, Zamecnik and Stephenson have demonstrated that an oligonucleotide, complementary to 13 nucleotide sequence of Rous sarcoma virus, can inhibit viral replication in cell cultures.[3] Since then, the progresses in antisense technology have significantly accelerated even leading to the commercialization of Vitravene (Fomivirsen) and Kynamro (Mipomersen), the first two antisense oligonucleotides (ASOs)

drugs approved by the FDA.[4,5] Antisense agents are defined as short (16-21 mer) single-stranded oligonucleotides complementary to specific mRNA able to interfere with the translation of the latter into protein. To perform their function two different mechanisms are proposed and universally accepted. A first mechanism involves the formation of a duplex complex through Watson-Crick base pairing between a target mRNA and an ASO, which can activate the enzyme RNase-H inducing the mRNA degradation. A second pathway includes ASO that are not able to induce RNase cleavage but can block the translation by steric hindrance of the ribosome in the cytoplasm. A further useful therapeutic strategy to control gene expression is the antigene therapy. In this case, the single-stranded oligonucleotide enters the nucleus of the cell, binds the DNA forming a double-strand DNA- antigene triplex complex through Hoogsteen hydrogen bonds, thereby leading to the inhibition of transcription of the target gene. In the last 40 years, the growing interest in oligonucleotide analogs as potential therapeutic agents has led to the development of emerging drugs targeting RNA with even more optimized pharmacokinetic properties. Oligonucleotide therapeutics include aptamers,[6] small interfering RNAs (siRNAs),[7] antisense oligonucleotides and ribozymes[8] that can be classified by their different mechanism of action.

1.2 Antisense Oligonucleotide

The key to be a good antisense candidate is to possess high specificity and affinity for the target sequences, furthermore they must show a sufficient nuclease resistance. Natural DNA or RNA oligonucleotides are quickly degraded by endogenous nucleases, consequently over the past years many chemically modified oligonucleotides have been designed and synthesized in order to extend *in vivo* half-life, to improve their target affinity, or to reduce cellular toxicity and other potential side effects. Modifications in the phosphodiester backbone, in the sugar moiety and in the nucleobase have been investigated in order to make oligonucleotides suitable for clinical purposes, moreover, they can be grouped into three categories according to the type of introduced modifications.

1.2.1 First-generation ASOs

The most representative DNA analogs belonging to the first generation of antisense oligonucleotides are certainly Phosphorothioate (PS) and Methylphosphonate (MP) nucleotides, in which a non-bridging oxygen atom of the phosphodiester backbone is replaced by sulfur and a methyl group, respectively (**Figure 1.2**). The interest in phosphorothioate biochemistry has started with the synthesis of some nucleoside 5'- phosphorothioates, which resulted extremely resistant to the action of alkaline phosphatase compared to nucleoside 5'-phosphates.[9] In light of these findings, the first PS oligonucleotides were synthesized by Eckstein *et al.* in order to study their cellular resistance and the ability to induce interferon production both *in vitro* and *in vivo*. This class of compounds was first tested as antisense agent by Matsukura and coworkers, who studied the inhibition of replication of human immunodeficiency virus (HIV) in human cells by phosphorothioates of different length.[10] These analogs have been developed with the aim to overcome the stability issues due to the degradation of the natural DNA/RNA backbone by cellular nuclease. Actually, in the past decades, it has been demonstrated that half-life in human serum of PS oligonucleotides is prolonged of few hours compared to an unmodified oligonucleotide.[11,12] So far, their attractive pharmacokinetic properties and the ability to recruit RNase-H make oligo-phosphorothioates the most widely used compounds in antisense therapy (e.g. Vitravene). Nevertheless, several drawbacks have been identified during exploration of their biological properties. The polyanionic nature of these compounds, if on the one hand make it possible the binding to plasma proteins, that seems to be the reason of a longer half-life in serum, on the other it also may cause potentially toxic side effects due to these non-specific interactions.[13] Oligo-phosphorothioates have also shown a lower binding affinity to RNA, approximately 0-4°C melting temperature in less regarding unmodified counterparts,[14] that is, however, partially compensated by an increase in hybridization specificity.[15] Furthermore, in the PS linkage the phosphor atom has four different substituents and thus is a chiral center. Since normally non enantiospecific methods are used for the formation of the phosphorothioate bond, PS drugs are generally obtained as mixture of 2^N stereoisomers, where N is the number of PS groups. This stereochemical issue is frequently overlooked, although it has been demonstrated that phosphor stereochemistry can affect DNA binding and interaction with proteins.

Methylphosphonate oligonucleotides have been the first chemically modified DNA analogs to be synthesized. Unfortunately, differently from oligo-PSs, their application in antisense technology has been prevented by their intrinsic features. In fact, the absence of charge on the backbone decreases their solubility and their cellular uptake, that do not occur by diffusion through membranes,[16] and, in addition, they are not capable to activate RNase H enzyme. Although the methylphosphonates have shown to inhibit gene expression, high concentrations (between 50 and 100 μM) are required to serve this purpose.[17] For all these reasons, their use in gene therapy was rather limited and other modifications have been investigated to try to improve the effectiveness and the safety of oligonucleotides mimic.

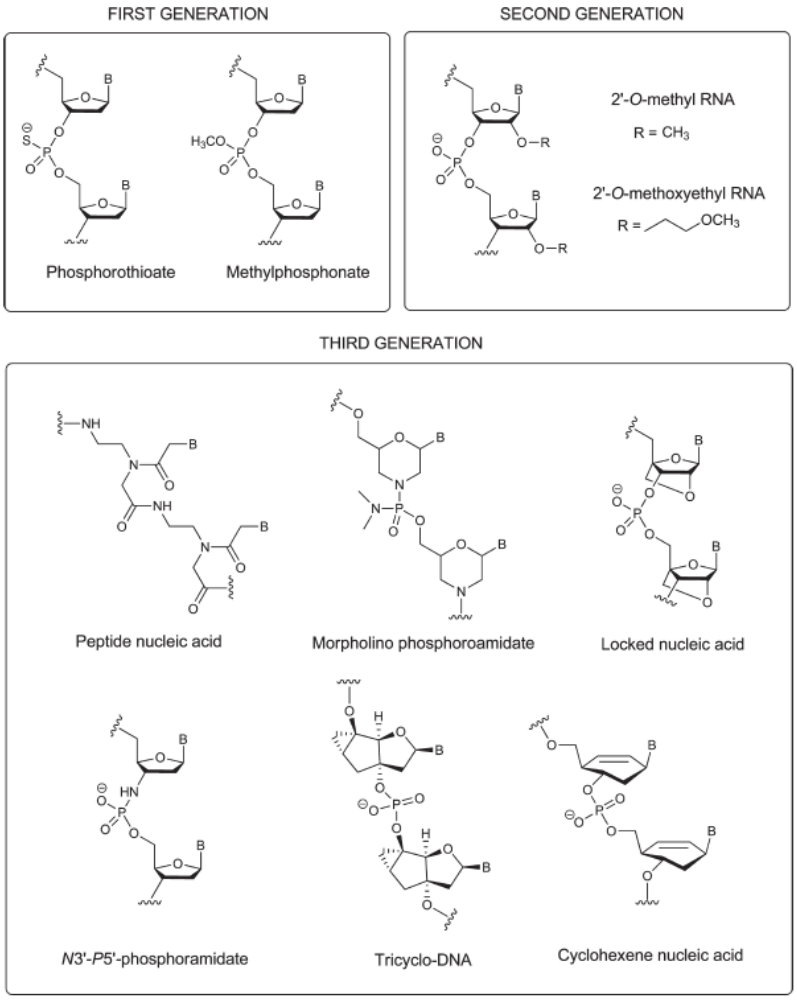


Figure 1.2. Different generation of oligonucleotide mimics (Antisense Oligonucleotides; ASOs).

1.2.2 Second-generation ASOs

The development of a second generation of antisense oligonucleotides, that include electronegative substituents at the 2'-position of the ribose, have in part solved the problems associated with the methylphosphonate and phosphorothioate analogs. The most relevant members of this class are 2'-O-methyl and 2'-O-methoxyethyl RNA (**Figure 1.2**). These type of ASOs have shown a reduced toxicity and an enhanced affinity towards their complementary RNAs, which appears to be due to an RNA-like C3'-endo conformation of the oligonucleotide conferred by alkyl substituents, present uniquely in A-type duplexes.[18] However, 2'-O-alkyl RNA are not able to recruit RNase-H, precisely because of this type of conformation. The RNase-H independent mechanism of these analogs was shown in 1997, when 2'-O-(2-methoxy)-ethyl antisense oligonucleotides were selected to inhibit the expression of human intercellular adhesion molecule 1 (ICAM-1) by blocking the translation event.[19] Although the antisense approach can proceed through different mechanisms, cleavage of target RNA by RNase-H is necessary in order to enhance antisense efficacy. This drawback has been addressed through the synthesis of the so-called "chimeric" oligonucleotides, that consist in a central RNase-compatible DNA gap, sufficient for activation of RNase-H, with 2'-O-alkyl modified nucleotides placed at each end, able to prevent nucleolytic degradation. In these "gapmers" the phosphodiester linkage can be also replaced by phosphorothioate, as in Mipomersen, the second antisense drug approved by the FDA.[20] Mipomersen (trade name Kynamro) is an ASO targeting apolipoprotein B, the main component of low-density lipoprotein (LDL) and very low-density-lipoprotein (VLDL), used to treat patient with homozygous familial hypercholesterolemia. The drug is administered by weekly subcutaneous injection and several side effects have been observed, especially with regard to the liver. Indeed, it can cause fatty liver disease and an abnormal increase of the transaminase level combined with flu-like symptoms and injection site reactions. In recent years, many second-generation antisense oligonucleotides, including the gapmers type chimera, are being tested in several clinical trials, in order to improve therapeutic properties.

1.2.3 Third- generation ASOs

A further improvement in the antisense technology has arisen with the development of the third-generation antisense oligonucleotides, which include analogs with modifications at the ribose sugar unit, at the phosphodiester linkage or a replacement of the latter with a completely different chemical moiety. Most notable among these are the morpholino-phosphoroamidates,[21] the locked nucleic acids (LNAs)[22] and the peptide nucleic acids (PNAs) (**Figure 1.2**).

Morpholino oligos, conceived by James Summerton in 1989,[21] consist in a DNA mimic with a phosphoroamidate linkage to replace the phosphate bonds and a morpholino moiety instead of ribose sugar. This class of ASOs has been devised with the aim to get around cost issues related to DNA analogs, in fact their synthesis involves less expensive ribonucleosides, which are converted to the desired morpholino structure via relatively simple synthetic steps (**Figure 1.3**).

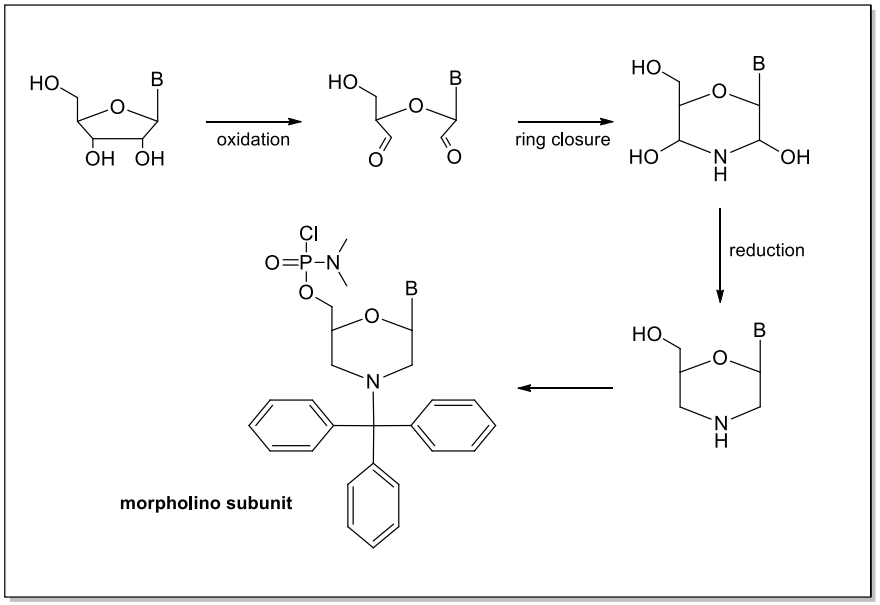


Figure 1.3. Synthesis of morpholino subunit.

Despite their nonionic phosphoroamidate backbone, morpholino oligonucleotides are very soluble in water because they show a strong base stacking, furthermore, because of the character of their backbone, unwanted interactions with nucleic acid-binding proteins are minimized. These analogs are extremely resistant to nuclease action and they exhibit a higher affinity

toward their complementary target RNA compared to the corresponding DNA/RNA duplexes, although lower than RNA binding obtained with other classes of modified DNA mimics. Since morpholino oligoes do not activate RNase-H, they fulfill the antisense function inhibiting translation of mRNA by targeting the 5' translational start region of the selected mRNA.[21] Morpholinos have played a fundamental part in the field of developmental biology, since sets of morpholino oligonucleotides were used in 'reverse genetics' to establish gene role, especially in morphogenesis.

Locked nucleic acids, first described in 1998, represent promising candidates of chemically modified antisense oligonucleotides.[23,24] LNAs contain a methylene bridge connecting the 2'-oxigen with the 4'-carbon of the ribose sugar and thanks to this conformational restricted structure they have displayed an exceptional hybridization affinity toward their complementary target DNA and RNA. NMR spectroscopic studies have shown that there is an increase in the A-type (RNA-like) character in LNA-RNA duplexes increasing the number of LNA nucleotides in the strand, as a result, the capacity to recruit RNase-H enzyme is inhibited. In order to restore an RNase-H activity, chimeric LNA-DNA gapmers incorporating 6-7 DNA monomers have been studied.[25]

The remarkable properties of these compounds have led to an extensive use either as antisense agents in gene therapy or for diagnostic purposes. Indeed, LNA oligonucleotides, in addition to their exceptional thermal stability, have exhibited excellent mismatch discrimination suitable for example, in genotyping studies related to Single-Nucleotide Polymorphism (SNP), which is associated to a wide number of genetic diseases. Methods as gel electrophoresis or enzymatic digestion are traditionally involved in scanning SNPs; in LNA-based assays, probes complementary to the mutated or the wild-type DNA are employed and the DNA sample (usually after amplification with PCR) is hybridized to the specific LNA probe. An example is represented by the work of Orum and coworkers that in 1999 have developed an ELISA-like assay for the screening of the factor V Leiden mutation by exploiting LNA hybridization properties.[26]

Among the large number of modified ASOs, it is also worth mentioning the N3'-P5'- phosphoroamidates (NPs), the cyclohexene nucleic acids (CeNA) and the tricyclo-DNAs (tcDNA) (**Figure 1.2**). In NPs oligoes 3'-amino group replace the 3'- hydroxyl group of the ribose ring and despite they do not activate RNase-H cleavage of target mRNA, their antisense effect by steric blocking of translation has been demonstrated both in cell culture[27] and *in vivo*[28]. Lastly, CeNA analogs

as well as tcDNAs have shown an enhanced binding affinity, to be resistant to nucleolytic degradation and consequently to be useful tools in the antisense technology.[29,30]

1.3 Aptamers, siRNAs and Ribozymes

Many other RNA-based therapeutics have been investigated, besides antisense oligonucleotides, and are currently under clinical trials. Aptamers are single-stranded DNA or RNA oligonucleotides that thanks to their unique three dimensional structure can bind target proteins as well as cells, metal ions, nanomaterials, viruses and small molecules.[31, 32] They are also called "chemical antibodies" since can be used in therapy similarly to protein monoclonal antibodies. However, aptamers present many advantages compared to these latter and some of them are reported below:

- Aptamers are produced by chemical synthesis in a scalable process.
- Chemical modifications can be easily introduced with the aim to improve the performance in terms of stability and functionality.
- Aptamers are generally non-immunogenic.
- Aptamers can readily enter into biological compartments thanks to their small size.

SELEX (Systematic Evolution of Ligand by Exponential Enrichment first introduced in 1990)[33] is the process by which aptamers are usually selected among a myriad of other sequences. The first step of this method requires the synthesis of oligonucleotides sequences, a library of about 10^{16} ssDNA that are repeatedly incubated with immobilized targets in the second step of the process. Once the oligonucleotide-target binding occurs, the unbound sequences are removed while the bound oligonucleotides are first eluted to facilitate the loss of the binding conformation and then amplified by polymerase chain reaction (PCR). The case of RNA SELEX is slightly different since the ssRNA library is obtained *by in vitro* transcription of a double stranded DNA template using T7 RNA polymerase. So far, SELEX technology allowed to select a large number of aptamers targeting leukemia cancer cells, breast cancer cells[34] and lung cancer cells,[35] and many other aptamers are under clinical trials for different diseases.[36,37] Despite their attractive properties, only one aptamer-

based drug has been approved by the US FDA in 2004, for the treatment of age-related macular degeneration.[38]

RNA interference (RNAi) is a biological defense mechanism in which RNA molecules can inhibit the expression of a gene, first discovered in 1998 in the nematode *Caenorhabditis elegans*.[39] In 2001, Tuschl and coworkers have observed RNAi gene silencing in mammalian cells for the first time.[40] The triggers of this phenomenon are known as small interference RNA (siRNA), long double-stranded RNA consisting in a guide strand and a passenger strand with 2 nucleotides overhang at the 3' side. The endonuclease enzyme called Dicer, belonging to the RNA III family, identify endogenous dsRNA and cleave it into these 21-23 long siRNAs that are subsequently incorporated into the RNA-Induced Silencing Complex (RISC) which promote the degradation of the target mRNA (**Figure 1.4**).

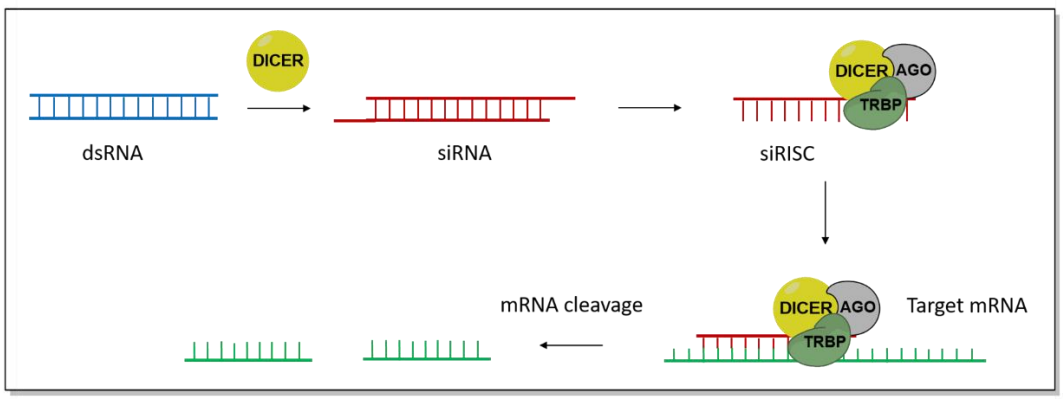


Figure 1.4. RNA interference mechanism.

Studies revealed that unmodified and naked siRNAs can be rapidly degraded into the bloodstream,[41] can induce off-target effects or activate unwanted immune stimulatory response.[42] As in the case of antisense oligonucleotides, various chemical modification geometries have been developed and tested in order to reduce potential side effects and increase the treatment potency,[43] for instance, the combination of sugar and backbone modifications, such as 2'-F, 2'-OMe and phosphorothioate, is well tolerated and provide an additional exonuclease resistance. Nevertheless, modifications that make the double strand too stable can interfere with gene silencing activity, avoiding the passenger strand removal and inhibiting the RISC loading of the guide strand.[44] Delivery of siRNA to the target system remains the major challenge for clinical applications since the drug needs to reach the circulatory system, resist to

enzymatic degradation, avoid renal filtration, cross cancer cells membrane and then escape the endosome to the cytoplasm. Chemically modified siRNA has partially addressed the issue regarding the stability against nucleases, however, the other questions need to be solved. Two different pathways are suitable for siRNA administration in the body. The local siRNA delivery involves direct delivery of siRNA to the target tissue, it has several advantages such as reduced undesired effects and high bio-availability, and to date, many siRNA drug candidates targeting skin diseases, ocular diseases, lung infections and others, via local delivery are currently in different trial phases. However, some other diseases such as viral infections, cancer and hypercholesterolemia require a systemic delivery approach, consequently many delivery strategies based on encapsulated or conjugated siRNA have been investigated. After two decades, the United States Food and Drug Administration approved ONPATTRO® (patirisan, ALN-TTR02) for the treatment of the polyneuropathy caused by hATTR amyloidosis and GIVLAARI ™ (girovisan, ALN-AS1) for the treatment of adults with acute hepatic porphyria (AHP) for commercial application.

Ribozyme is a term coined in the early 1980s to describe RNA molecules possessing an enzymatic activity. The catalytic role of RNA was discovered by Altman and coworkers in the enzyme called RNase P, in fact they observed that the RNA portion of the enzyme was able to convert a precursor tRNA into the mature tRNA in the absence of protein cofactors.[45] After that, many ribozymes capable of catalyze self-splicing or cleave reactions have been isolated in viruses, lower eukaryotes and bacteria and their mechanisms of action have been extensively described.[46,47] The therapeutic potential of ribozymes has led the researchers to develop a new class of human antisense agents, natural and artificial, due to the fact that ribozymes can, theoretically, selectively cleave any target RNA.

A summary schematic illustration of different oligonucleotide based therapeutic mechanisms is shown in the figure below (**Figure 1.5**).

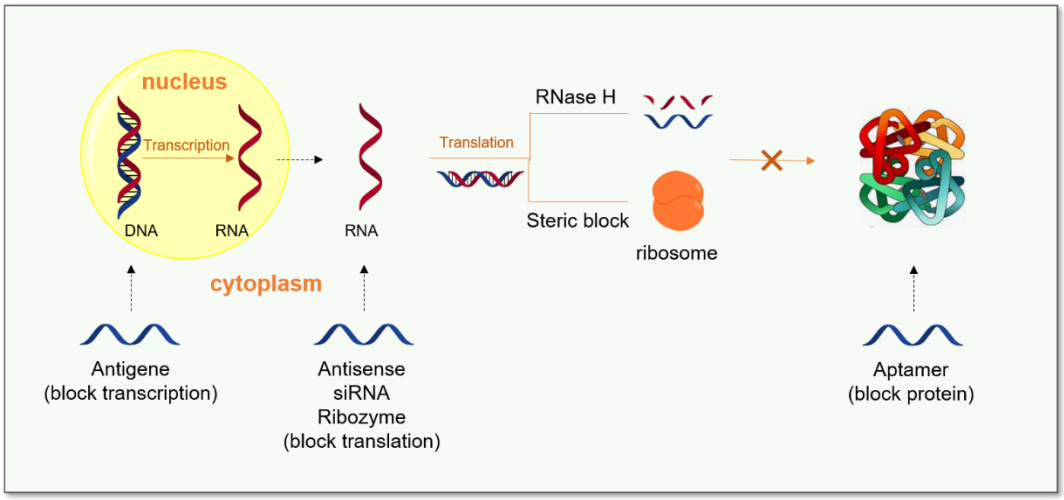


Figure 1.5. Different approaches in oligonucleotide-based therapy.

1.4 Peptide Nucleic Acids

Peptide nucleic acids are a class of DNA analogs introduced by Peter Nielsen and coworkers in 1991.[48] As stated previously, they belong to the third generation of antisense oligonucleotides, although their applications are not solely related to the field of antisense technology. In peptide nucleic acids, the natural sugar-phosphate backbone has been replaced with a pseudo-peptide backbone consisting of *N*-(2-aminoethyl)glycine units linked to the standard pyrimidine (C,T) and purine (A,G) nucleobases by a carbonyl methylene bridge (**Figure 1.6**). PNAs were designed initially by Nielsen and coworkers as ligands, whose purpose was to recognize specifically a double-stranded DNA forming a T-AT triple helix via Hoogsteen hydrogen interactions, but the promising results obtained in terms of affinity and specificity have prompted the researchers to explore extensively the DNA mimicking properties of PNA.

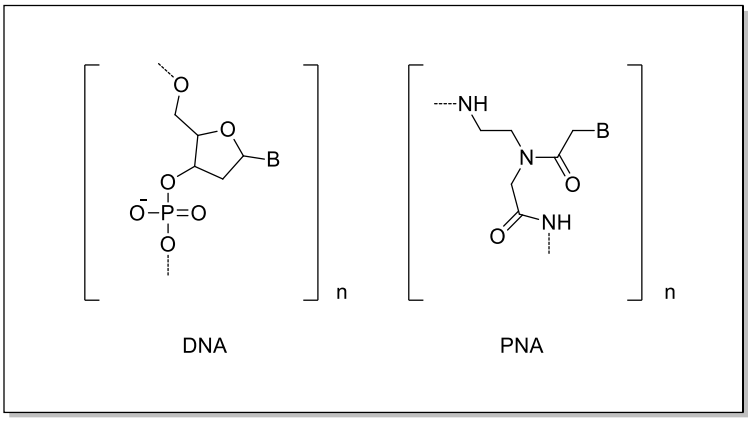


Figure 1.6. Schematic structures of DNA and PNA.

It has been found that the charge-neutral PNA backbone, thanks to the lack of electrostatic repulsion upon hybridization with another oligonucleotide strand, made the duplexes between PNA and DNA or RNA more stable than the corresponding DNA/DNA or DNA/RNA duplexes, with a melting temperature about 1.0-1.5 C° higher per base pair. In general, the thermal stability of duplexes follows the order: PNA/PNA > PNA/RNA > PNA/DNA, where the duplex between PNA and RNA is more stable compared to the duplex with DNA because of the A-form-like conformation adopted by the helical structure. A PNA strand can bind complementary DNA (RNA) strands in two different orientations, the antiparallel configuration with the amino terminal of PNA facing the 3'-end of the oligonucleotide and the parallel configuration that has

been shown to have the lower stability (typically, a ΔT_m of 1-2 C° per base pair). It is worth mentioning that different PNA binding modes have been observed during the last decades. As mentioned above, homo-pyrimidine PNA oligomers were designed originally with the aim to hybridize target homo-purine DNA duplexes by major-groove triplex formation, but they obtained, instead, a highly stable PNA-triplex inducing a strand displacement process, in which one PNA strand binds the complementary homo-purine DNA strand via Watson-Crick base pairing, while the homo-pyridine DNA strand is displaced as a single-stranded P-loop by a second PNA strand that bind this new PNA/DNA duplex via Hoogsteen hydrogen bonds (**Figure 1.7**). The main difference between the triplex and triplex-invasion binding modes lies in the on-rate kinetic, that is fast in the first case and thus not negatively affected by ionic strength, while the triplex-invasion binding is under kinetic control and can be adversely affected, for example, by the presence of cations (K^+, Mg^{2+}, spermine, etc.). All the dsDNA-PNA complexes so far discovered are shown in the figure below.

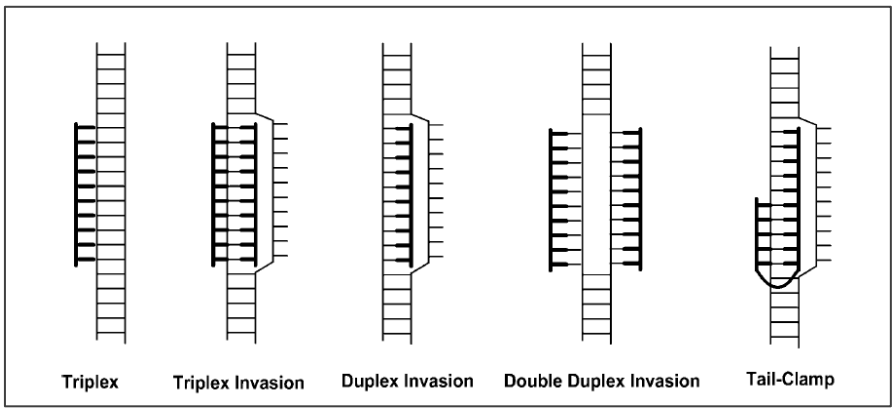


Figure 1.7. Four different types of PNA–dsDNA complexes. DNA is schematically drawn as a ladder, and the PNA oligomers are in bold. Reprinted with permission from P. E. Nielsen Copyright © 2010 Verlag Helvetica Chimica Acta AG, Zürich.

Peptide nucleic acids have proved to be resistant to nuclease and protease enzymes, which, instead, are the cause of rapid degradation of natural phosphate DNA/RNA backbone. Unfortunately, the noncharged nature of PNA backbone provides a poor water solubility to this class of oligonucleotides compared to DNA, and a propensity toward self-aggregation. A way to improve the solubility of PNA consists in the incorporation of positively charged lysine residues, both at the C-terminal or in lieu of glycine, which has proven to have also a favourable effect on DNA hybridization properties of PNA oligomers.[49]

Among the various classes of nucleic acid mimics, peptide nucleic acids are one of the most successful, mainly thanks to the biological properties they have exhibited during the past thirty years but also because of the easy synthetic pathway that can be used to assemble different PNA monomers.[50] PNAs can be synthesized by solid-phase manual or automated synthesis, a method that involves a three-reaction cycle which is deprotection, coupling and capping (**Figure 1.8**). The two main strategies involved in solid-phase synthesis are the Boc (*tert*-butyloxycarbonyl) and Fmoc (9-fluorenylmethoxycarbonyl) strategy, which represent temporary protecting groups (PG) of the primary amino function. Heterocyclic nitrogen atoms of nucleobases are protected with orthogonal semi-permanent protecting group, in order to avoid side reactions. Through this simple protocol, it has been possible introduce modifications both on the backbone and nucleobases, with the intention to enhance PNA applications in antigene/antisense therapy as well as in the field of genetic diagnostic.

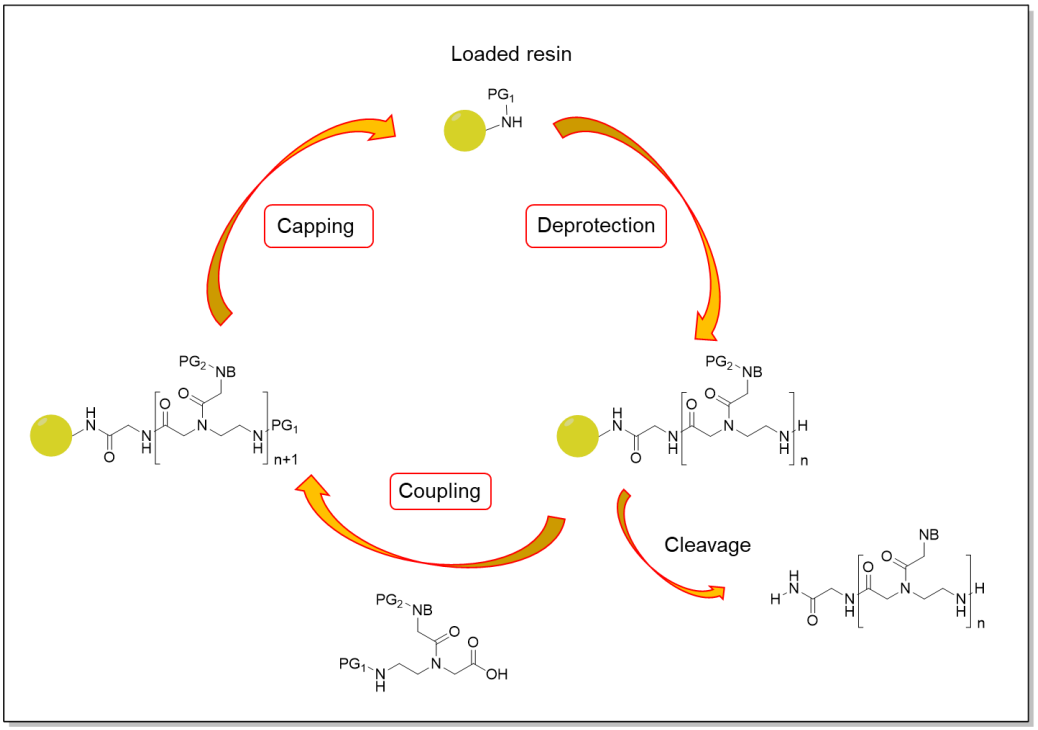


Figure 1.8. Schematic representation of solid-phase synthesis of PNA.

1.4.1 PNA targeting micro-RNA

The extraordinary hybridization properties exhibited by PNAs have led to the employment of these analogues in many fields. One showing increasing interest in the last decade is targeted modulation of microRNA-regulated gene expression.

MicroRNAs (miRNAs or miRs) are small noncoding RNAs, 19-24 nucleotides in length, first discovered by Lee and colleagues in *C. Elegans* in 1993.[51] In the cited work, and in other following works,[52–54] it was found that *let-7* and *lin-4* miRNAs can downregulate gene expression by binding the target sites in 3' untranslated region (3' UTR) of messenger RNA. Subsequently, a wide number of miRNAs have been identified in viruses, animals and plants, and it has become to be ever more clear their importance in controlling cellular processes.

The miRNA biogenesis starts with the transcription of the gene into an hairpin primary miRNAs (pri-miRNAs), generally mediated by RNA polymerase II (Pol II), although in some cases miRNAs can be transcribed by RNA polymerase III.[55] A microprocessor complex, consisting of RNase III enzyme Drosha and the RNA-binding protein DGCR8, cleaves pri-miRs into a shorter nucleotide (60-70 nt) called precursor miRNA (pre-miRNA) that is subsequently exported to the cytosol by Ran/GTP/Exportin 5 complex (**Figure 1.9**). For the production of pre-miRs, an alternative mechanism, in which miR-introns (miRtrons) are involved, has been observed. In this case, the Drosha cleavage is bypassed and the miRtrons are processed from unspliced intronic regions leading directly to precursor miRNAs.[56] In the cytosol, the both type of hairpin-shaped pre-miRs are processed by Dicer to produce a mature miRNA duplex, consisting of a guide strand and a passenger strand, the latter typically degraded once the mature miR is incorporated into the RISC complex. At this point, the post-transcriptional regulation of target mRNA can follow two potential mechanisms, according to the stability of association between the target messenger RNA and the guide miRNA strand. Indeed, target mRNA can be degraded in case of a near-perfect complementarity to the miRNA, otherwise gene silencing will occur through a translational repression.[57,58]

Over the past years, it has turned out that the dysregulation of microRNA expression is involved in different forms of cancer and in the onset of various other diseases. The first evidence of the miRNA implication in human cancer was reported by Prof. Croce group, which identified two miRNA genes, i.e., *miR-15a* and *miR-16-1*, at chromosome 13q14 region in B-cell chronic lymphocytic

leukemias (CLL), a region that is generally deleted in this type of disease. They have observed that both genes are deleted or downregulated in the ≈ 68% of clinical CLL cases.[59] In human cancers microRNAs play a dual role considering that they both can be overexpressed or underexpressed and so act as an oncogene miRNA[60,61] (oncomiR) or a tumor suppressor[62,63], respectively, by suppressing their mRNA counterparts. Nevertheless, the two roles are not mutually exclusive, as it happens in the case of miR-125b, which has been found to act as either tumor suppressor in various solid tumors and as an oncomiR in hematological malignancies.[64] This should not be considered unexpected since it is known that the 3'UTR of a mRNA can interact with multiple miRNA and tens to hundreds different messenger RNAs can be targeted by a single microRNA molecule. Given this, it is not surprising that the interest in controlling the expression of miRNAs for therapeutic purposes has recently been increasing exponentially.

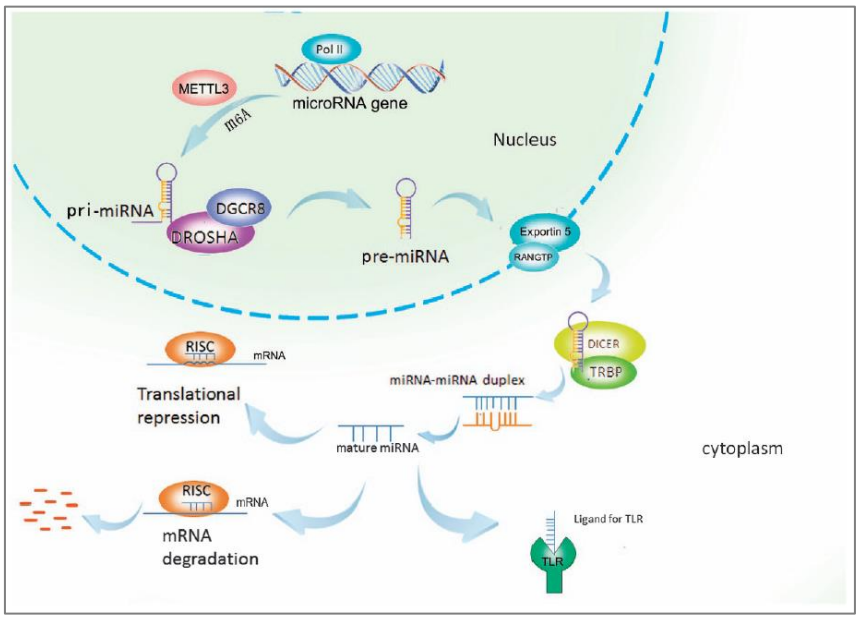


Figure 1.9. MicroRNAs biogenesis. Peng, Y., Croce, C. The role of MicroRNAs in human cancer. Sig Transduct Target Ther ***1,*** *15004 (2016) Copyright © 2016 West China Hospital, Sichuan University.*

In the past decade, many studies have confirmed the efficacy of peptide nucleic acids as modulator of the biological activity of miRNAs. The first example of this approach is represented by a study conducted in Dr. Gait group regarding a liver-specific microRNA (miR-122) involved in lipid metabolism, but especially in the Hepatitis C virus (HCV) replication. They evaluated the efficiency of miR-

122 inhibition by a peptide nucleic acid delivered via electroporation and a LNA/2'-O-methyl oligonucleotide mixmer delivered through lipofection in human and rat liver cells.[65] In the same work and in a following study,[66] it was also shown that conjugation of anti-miR-122 PNA with the cell penetrating peptide (CPP) R_6-Penetratin or use of lysine-derivatized PNA is sufficient to obtain a miR activity inhibition when these ONs are incubated in continuous cell culture for several days (gymnotic delivery) in the absence of transfection agents.[67]

miR-155 is one of the first oncogenic miRNA identified and one of the most studied. Overexpression of this microRNA is correlated to several B-cell malignancies, such as Burkitt's lymphoma (BL), Hodgkin's lymphoma (HL), diffuse large B-cell lymphoma (DLBCL) mucosa-associated lymphoid tissue lymphoma (MALT) and chronic lymphatic leukemia (CLL),[68–70] and therefore it has become a potential and promising biomarker in personalized medicine.[70] In 2010, Gait *et al.*[71] have proven the efficacy of K-PNA-K_3 oligonucleotide in blocking miR-155 function in primary murine B cells, both in culture and, for the first time *in vivo*, in the presence of lipopolysaccharide (LPS) that is has been demonstrated to stimulate miR-155 expression in cultured B cells.[72,73] A couple of years later Saltzman and Slack have generated an *in vivo* delivery system where a charge-neutral anti-miR155 PNA is encapsulated in poly(lactic-*co*-glycolic-acid)(PLGA) nanoparticles decorated with penetratin (ANTP-NPs).[74] They have shown that the treatment of pre-B cells in a mouse model of lymphoma with ANTP nanoparticles leads to a withdrawal of miR-155 bioavailability with a consequent rapid regression of tumor growth. Another delivery platform has been introduced by the same research group in 2015, which included the antimir-155 PNA linked by a disulfide bond with a pH-induced transmembrane structure peptide (pHLIP), in order to target acidic tumor microenvironment.[75] The peptide is known for the capacity to translocate impermeable molecules directly into the cytosol, thus skipping endosomal entrapment. Furthermore, it assumes a conformation suitable for transfection only at relatively low pH , the latter considered a hallmark of solid tumors.[76] Through this unique approach charge-neutral antimiR-155 PNA has been efficiently delivered into the tumor cells leading to a delaying of both tumor growth and the development of lymphadenopathy, in addition to the suppression of metastatic spread of neoplastic cells. It is worth mentioning that no evidence of systemic toxicity or renal damage has been observed in mice treated in the reported *in vivo* studies. A therapeutic protocol that involves a

combined treatment with two antimiR PNAs has also been investigated. miR155 and miR-221 are two oncomiRs highly expressed in gliomas and implicated in caspase-3 mRNA regulation. The co-administration of antimiR-155 and antimiR-221 PNAs was confirmed to induce apoptosis in temozolomide (TMZ)-resistant T98G glioma cell line and, in addition, to sensitize target tumor cells to TMZ, when the latter was used in combination with the two antimiR PNAs.[77] An analogue result was reported when the R8-antimiR-15b-5p PNA is administered with Sulforaphane for the treatment of colon cancer HT-29 cell lines.[78] Sulforaphane, a compound derived from cruciferous vegetables, is known for its chemopreventive activity.[79–81] The synergistic effect between this chemoprotective agent and the antimiR-15b PNA has led to an activation of the proapoptotic pathway in a percentage way higher than what has been observed for the singular administration.

1.5 PNA delivery methods

One of the main drawbacks for use peptide nucleic acids as antimiR agent for therapeutic purposes remains their poor cellular uptake. Many strategies have been investigated in order to overcome this limitation and many progresses have been made in this respect.[82]

1.5.1 Direct delivery

Microinjection is a direct approach to deliver foreign nucleic acids into target cells extensively used in a large number of applications. This technique involves the use of a glass microcapillary pipette, a microinjector and a micromanipulator, a device used to control the motion of the micropipette.
The first study regarding the intracellular direct delivery of PNA has been carried out by Hanvey *et al.* in 1992, in order to demonstrate an antisense/antigene effect of these oligonucleotide analogues. [83] A 15-mer and a 20-mer PNA were microinjected into fibroblast cells expressing SV40 large T antigen (T-Ag) and it was observed the suppression of the T-Ag protein expression. Although a few examples of microinjection technique applied to PNAs are reported in literature,[84,85] researchers have been discouraged by this laborious procedure only applicable to a small-scale set-up, so as to choose different PNA delivery methods.

Electroporation represents an alternative and feasible method in which an electrical field is applied to cells in order to open pores and increase the permeability of cell membrane. For instance, it has been used to deliver PNAs in cell culture system to inhibit the replication of the human immunodeficiency virus and to induce endogenous gamma-globin gene expression through the binding with 5' flanking region of the target gene.[86] Permeabilization of eukaryotic cells was also obtained by using streptolysin-O. In this study, a dimeric PNA molecule was delivered into cells with the aim to form a PNA/DNA/PNA triplex, via both Watson-Crick and Hoogsteen base pairing, that could cause mutagenesis within transgenic mouse fibroblasts.[87]
Corey *et al.* have presented an alternative strategy to transfect unmodified PNAs into living cells in culture through a PNA/DNA-lipid complex.[88,89] In fact, the use of cationic lipids is an established method to facilitate internalization of anionic nucleic acids. In this case, DNA represents a carrier for peptide nucleic acid cargoes that are delivered and subsequently released, in order to exhibit their exceptional antisense properties.

1.5.2 Conjugated-PNA based delivery

A promising approach widely explored for cellular delivery of PNA involves the conjugation to cell-penetrating peptides (CPP)[90–95] also known as protein translocation domains (PTD) or Trojan peptides (TP). CPPs can be generally classified into two categories: *i*) cationic CPPs, usually rich in positive-charge amino acids such as lysine or arginine. *ii*) amphipathic peptides, whose structure allows a direct translocation across the plasma membrane. Penetratin is a 16 amino-acid long peptide largely used for the delivery of peptide nucleic acids, obtained from *Drosophila antennapedia* homeoprotein. A first successful *in vivo* study regarding a PNA-penetratin conjugate has been reported by Pooga *et al.* in 1997, where the suppression of galanin receptor gene expression has been established and compared with the inhibition obtained with a different cell-penetrating peptide, Transportan, a 27 amino-acid long CPP containing galanin and mastoparan.[96] Although a large number of peptides has been employed to deliver PNA into live cells (**Table 1.1**), the major limitation of these platforms remains the endosomal trapping of molecular cargoes, hence more and more frequently a co-administration of endosomolytic agents, such as calcium ions or chloroquine, is required.

The conjugation of PNA to lipophilic moiety has also been investigated. In 1999 Nielsen and co-workers have described an improved cellular uptake of a 15mer-PNA conjugated with adamantyl acetic acid in human cells in culture, which turned out to be strongly dependent upon cell-type and PNA-sequence.[97] Muratovska *et al.* have exploited a lipophilic phosphonium cation to deliver a 11-mer PNA to mitochondria within 'myoclonic epilepsy and ragged red fibers' (MERRF) cells with the aim to inhibit MERRF mitochondrial DNA (mtDNA) replication. Indeed, triphenylphosphonium cations are taken up by mitochondria through the inner membrane, driven by the plasma membrane potential.[98,99]
Three different studies have been reported regarding PNA conjugated with cell-specific receptor ligands[100–102], but the results obtained in this respect do not represent a real step forward in cell-specific PNA delivery field.

Table 1.1. Peptides used for delivery of PNA. Adapted from Gupta et al.

Name of the peptide	Sequence of the peptide	Ref
Penetratin	RQIKIWFQNRRMKWKK	[93,94]
Transportan	GWTLNSAGYLLGKINLAALAKKIL	[92]
Retro-inverso penetratin	(D)-KKWKMRRNQFWIKIQR-(D)	[91]
Phenyl leucine	FLFLFLF	[82]
Nuclear localization peptide	PKKKRKV	[91]
Tat peptide	GRKKRRQRRRPPQ	[90]
Polylysine	KKKK	[82]
pH Low insertion peptide	AAEQNPIYWARYADWLFTTPLLLL DLALLVDADEGTCG	[75]

1.5.3 Nanoparticle based delivery

Polylactic Co-Glycolic acids (PLGA) represent a biodegradable copolymer already approved by both FDA and European Medicines Agency (EMA) and extensively used in drug delivery applications.[103] PLGA is constituted by different ratio of lactic- and glycolic acid esters that can be hydrolyzed under physiological conditions, moreover, its structure can be further adjusted by co-formulation with other polymers in order to create blended nanoparticles with tailored properties. PNA was successfully delivered into the cells by both regular and blended PLGA nanoparticles allowing to exploit its potential in

different fields, such as gene editing and anti-miRNA therapy.[74,104–108] For example, Glazer and coworkers reported a study regarding PNA targeting miR-210,[109] that is overexpressed in several cancer cells in response to hypoxia and associated to solid tumors.[110–112] In this work, the activity of an unmodified oligonucleotide was compared with that of a C (5)-modified PNA (MPγPNA), bearing polyethylene glycol chains (miniPEG) on its backbone. MPγPNAs generally display a higher binding affinity for complementary DNAs or RNAs as a result of the insertion of chiral centers with appropriate stereochemistry, which pre-organize the pseudo-peptide backbone, enhancing their features as antisense agents.[113,114] In fact, the treatment with PLGA NPs loaded with the anti-miR MPγPNA led to a significant delay in tumor growth in mice engrafted with human HeLa cells.

Mesoporous silica nanoparticles (MSNPs) represent a promising platform for various biomedical applications, such as bioimaging, biocatalysis and drug delivery. Their success is due to the good biocompatibility and degradability, in addition to the high loading capacity, the tunable pore structure, and the possibility to functionalize both the inner pore system and the outer particle surface. One of the most common type of MSNP (i.e. MCM-41) was initially synthesized at Mobil Corporation laboratories by a liquid crystal template mechanism with the aim to use them as molecular sieves because of their large internal surface area.[115] Six years later, well-ordered hexagonal mesoporous silica structures (SBA-15) with a tunable pore size between 50-300 Å were prepared at University of California, Santa Barbara.[116] SBA-15 was synthesized from tetraethoxysilane (TEOS), tetramethoxysilane (TMOS) and tetrapropoxysilane (TPOS) using amphiphilic poly (alkylene oxide) triblock copolymers as organic structure-directing agents in aqueous acidic conditions (pH ≈1). Through this method, the pore size and the thickness of the silica wall can be modulated by varying the copolymer composition or block sizes, the heating temperature and the time of SBA-15 in the reaction solution.

Since in 2001 the use of MSNPs as nanocarriers was first proposed to transport ibuprofen,[117] an anti-inflammatory drug, many efforts have been made to create multifunctional delivery platforms capable of encapsulating and carrying therapeutic cargo molecules, such as peptides, proteins and nucleic acids. Moreover, the properties of MPSNPs can be widely modulated by attaching molecular functionalities, including gatekeepers,[118] targeting ligands,[119] polymers,[120] and tracking markers,[121] both to external and/or internal surfaces with the aim to exploit their whole potential for nanomedicine development.

In 2014, fluorescent mesoporous silica nanoparticles were used to deliver an 18-mer antisense PNA within cancer cells for silencing of B-cell lymphoma 2 (Bcl-2) protein expression. In the mentioned work, Ma and coworkers demonstrated that a redox-triggered intracellular release of Cy5 labeled PNA is an effective strategy in order to achieve a controlled drug release and the subsequent protein silencing. To this end, a sequence-specific Cy5-PNA (Bcl-2) was linked to fluorescein isothiocyanate labeled MSNP (MSNP-FITC) previously functionalized with a disulfide bond (**Figure 1.10**). The release was monitored by confocal laser scanning microscopy and it was observed that in the presence of a natural reducing agent (i.e. Glutathione) inside HeLa cells, the disulfide bond is cleaved leading to the Cy5-PNA release into the cytoplasm. It is also worth mentioning that in the absence of disulfide bond, the release of Cy5-PNA was not observed, and in addition, the use of free Cy5-PNA at high concentration showed a low cellular uptake, confirming the efficacy of the PNA delivery through MSNPs. Lastly, the Bcl-2 protein expression in HeLa cell line was semiquantified by the Western blot assay, which showed a decreasing in the band intensity related to the protein expression upon increased concentration of MSNP-Cy5-PNA(Bcl-2).[122]

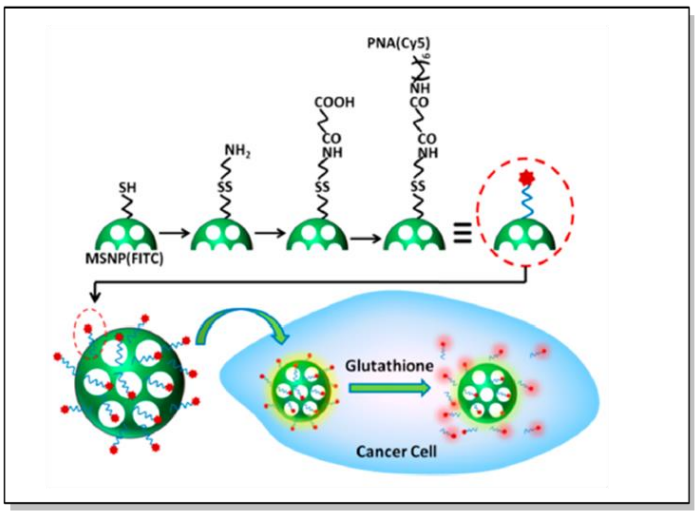


Figure 1.10. Illustration of the PNA Conjugation with FITC Labeled MSNP and Redox-Triggered Intracellular Release of Cy5 Labeled PNA. Reprinted with permission from ref. 122: Ma et al. © 2014 American Chemical Society.

A further interesting example of PNA delivery based on MSNPs was reported in a subsequent work by Bertucci *et al.*, which described the synthesis of a multifunctional nanosystem used to treat drug-resistant glioma cells. Temozolomide (TMZ) was loaded by impregnation into fluorescent MSNPs (Cy5 fluorophore was embedded in the silica matrix during the synthesis) that were subsequently decorated with a polyarginine antimiR-221 PNA by electrostatic binding on the particle surface. The presence of cationic PNA in addition to improving the cellular uptake of the functional nanosystem, can also serve as gatekeeper to control the release of temozolomide. The effect on cell viability was evaluated on TMZ-resistant T98G glioma cell line for TMZ-MSNPs, PNA-MSNPs and PNA-TMZ-MSNPs. The comparative analysis showed that the synergistic effect of PNA-TMZ-MSNPs led to a much higher increase of induced apoptosis of cancer cells compared to the values observed when the cells are treated with PNA-MSNPs and TMZ-MSNPs separately, namely 70.9% total of apoptotic cells using 0.5 mg mL^{-1} of PNA-TMZ-MSNPs.[123]

On the other hand, the proof of concept of PNA delivery through porous nanoparticles was reported for the first time by the same research group, in a study where a Zeolite-L-based multifunctional nanocarrier for the delivery of both organic molecules and PNA was outlined.[124] Zeolite-L, a porous crystalline aluminosilicate, was filled with a model fluorescent guest molecule (i.e. 4′,6-diamidino-2-phenylindole, DAPI) and then covalently functionalized with a PNA probe. In order to increase cellular uptake of the nanosystem, the crystalline structure was coated with the biodegradable cation polymer Poly-L-lysine (PLL). A rapid internalization into living cells of the hybrid nanoparticles (after 1h of incubation) was observed, while DAPI was not detected in the nucleus after 4h, indicating that the PLL coating inhibits the diffusion of the payload from the zeolite channels. Furthermore, DNA/RNA binding activity of PNA linked to the zeolites was tested by incubation of three different DNA strands (full match, single mismatch and 3-mismatch) and through the following UV/vis evaluation of the final DNA concentration in the solution, after the particles' dissolution. Thereby, the sequence selectivity of the PNA probe was confirmed also when attached to the zeolites.

Over the years, porous silicon nanomaterials (pSi) have shown to be suitable tools for biomedical applications as much as MSNPs. The current scientific interest in porous Si (**Figure 1.11, left panel**) is mainly due to the studies of Prof. Canham, who demonstrated its efficient visible photoluminescence and biocompatibility in a simulated body fluid in the mid 1990s.[125,126]

Electrochemical anodization of single crystalline Si wafers in hydrofluoric acid solution is the most common procedure to fabricate pSi materials, whose morphological properties, such as pore size, porosity and thickness, can be adjusted by varying the current density, HF concentration, etching time and wafer type and its crystalline orientation. The surface of freshly etched pSi contains hydrophobic Si hydrides (SiH, SiH_2, SiH_3) that are subject to oxidation over time in ambient conditions, thereby affecting the structural and optical properties as well as the degradation kinetic in aqueous media. Hence, the modification of surface with various chemical functionalities plays a major role in both stabilizing the pSi layer and tuning the dissolution rate of the nanomaterial. The ease with which pSi pore geometry can be tune during the fabrication process makes porous Si materials attractive and useful tools for drug delivery applications, thus a plethora of papers have been published in this respect starting from year 2000.

A first example of a novel PNA-pSi platform for applications in gene therapy and biosensing was reported by Duvall *et al.* in 2014.[127] Inspired by a previous work in which in situ DNA synthesis within pSi film showed a remarkable loading efficiency relative to the loading obtained with a presynthesized oligonucleotide, the researchers decided to replicate the study using a PNA designed to inhibit miR-122, a liver-specific miRNA largely expressed in hepatocytes and involved in cholesterol metabolism and liver function. By measuring the optical thickness of the loaded pSi film, they found that the loading capacity achieved using in situ synthesis was 8-fold higher than the one obtained with the conjugation of presynthesized PNA, but was comparable in the case of a physical adsorption within the pores. Next, the release profiles of physically adsorbed and in situ synthesized PNA were compared, proving that the covalent connection provided by in situ synthesis leads to more sustained release profile. In the light of these findings, the anti-miR activity, cellular uptake and cytotoxicity of PNA-pSi were evaluated in Huh7 (human hepatic carcinoma) cells, confirming the potential of this platform to enhance both the intracellular delivery and the bioactivity of therapeutics. Additionally, with the view of biosensing applications, the sequence selectivity of in situ synthesized 16-mer PNA on pSi biosensor towards a full complementary DNA target was established.

In a following work, with the aim to facilitate the endosomal escape of PNA-pSiNPs delivery platform, the same authors developed an endosomolytic polymer poly[(ethylene glycol)- *block* -(dimethylamino)ethyl methacrylate-*co* -

butyl methacrylate)] (PEG-DB) able to disrupt membranes in acidic environments.[128] For the scope, anti-miR-122 was loaded within pSiNPs by physical adsorption and, subsequently, the particles were coated electrostatically with the polymer PEG-DB (**Figure 1.11, right panel**).

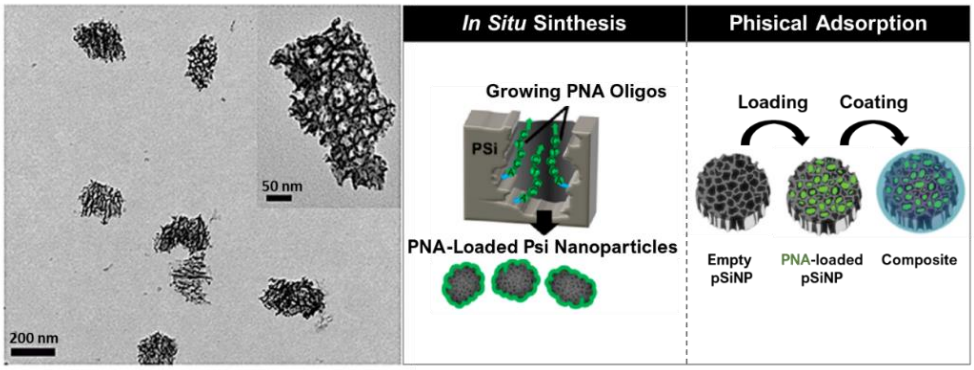


*Figure 1.11. Left panel: transmission electron microscope (TEM) image of unmodified pSiNPs; the inset shows a closer view of a single nanoparticle. Right panel: in situ synthesized PNA (**left**); physically adsorbed PNA (**right**). Reprinted with permission from Volpi et al. © 2020 by the authors. Licensee MDPI, Basel, Switzerland.*

From preliminary spectroscopic evaluations, the nanocomposite system showed a higher colloidal stability in physiological conditions compared to the uncoated pSiNPs, thus, cellular uptake and anti-miRNA activity were evaluated using Huh7 human liver cancer cells. The studies revealed that despite PNA uptake was lower in the case of coated nanoparticles, the latter improved the cytosolic delivery, miRNA inhibitory activity (10-fold higher), the blood circulation half-life of anti-miR122 PNA and, as a result, its bioavailability. Additionally, subsequent *in vivo* studies performed on female C57BL/6J mice have not only confirmed the results previously obtained (46% reduction in miR-122 levels relative to the empty carrier control) but have also shown that the treatment with nanocomposite particles did not cause liver or kidney toxicity.

Lastly, Duvall and coworkers have recently reported an interesting study regarding the effect of PEG-DB diblock copolymer composition on serum stability, endosome escape and bioavailability for the nanocomposite system mentioned above.[129] The endosomolytic polymer is formed by the PEG block that contributes to colloidal stability of the nanoparticles, and by the two monomers D ((dimethylamino)ethyl methacrylate) and B (butyl methacrylate) which are implied in the pH-responsive function. Since it has been assumed that

the efficiency of endosome escape is higher in early endosomal vesicles (pH=6.8),[130] various PEG-DB compositions (specifically increasing % of hydrophobic B content from 20 up to 70%) have been screened at different pH values in order to evaluate whether a robust membrane disruption at pH 6.8 correlates with the intracellular PNA bioactivity. In this study, they first observed that the stability of the polymer coating increases with hydrophobicity (% B), which also drives lipid bilayer membrane insertion. Next, it was found that the best balance between serum stability at physiological pH (7.4), pH-activated membrane disruption and a high bioactivity was provided by 40% B pSiNP composite. Indeed, although 30%B pSiNP composite showed a robust endosome disruption at the desired pH and the strongest bioavailability of the series, it also exhibited a premature release of the polymer at pH 7.4 that could cause hemotoxicity.

1.5.4 Modified PNAs based delivery

It has been proved that chemical modifications on PNA backbone which confer positive charge features can facilitate cellular uptake.[114,131] An example is represented by the positively charged group guanidinium that can be inserted on C-2 or C-5 position on the PNA backbone (**Figure 1.12**).

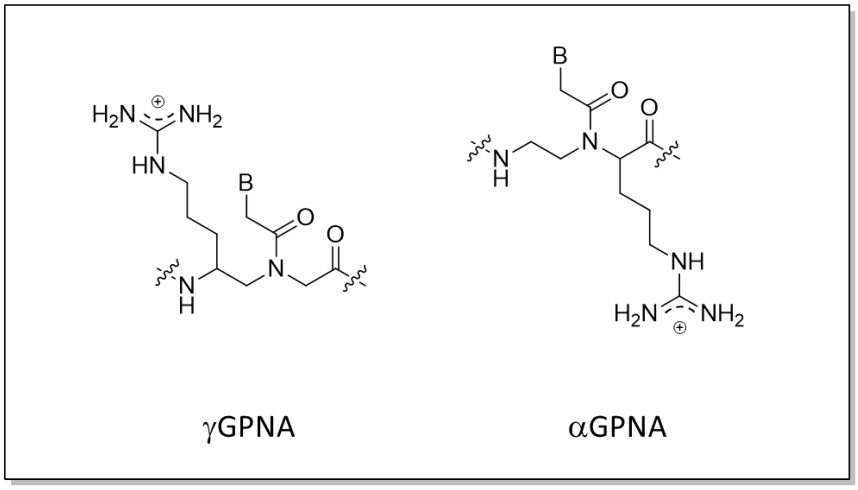


Figure 1.12. Chemical structure of γ GPNA and α GPNA.

In 2005, Danith Ly and coworkers have shown that GPNA can be readily taken up by mammalian cells,[131] one year later the researchers have designed a cell-permeable αGPNA in order to bind to 5′-Untranslated region of human E-

cadherin gene. Although biased by lack of control in enantiomeric purity of αPNA (see Chapter 3 for discussion) and therefore by the presence of a high number of diastereomers, the results reported in that study indicated that GPNA was able to induce an effective and sequence-specific antisense effect by blocking protein translation. Moreover, the cytotoxicity of GPNA oligomers was found to be lower than the corresponding PNA-peptide conjugates; likely the amphipathic nature of the latter may be responsible for the cytotoxic effect.

1.6 PNAs as diagnostic tools

The great chemical and biological stability as well as the high affinity of peptide nucleic acids towards complementary DNA/RNA has promoted the development of new diagnostic applications of PNAs in the biosensing field. A representative example of the value of PNA as diagnostic tool is provided by the fluorescence *in situ* hybridization (FISH) using peptide nucleic acid probes for rapid diagnosis of bloodstream infections.[132] Since PNA-FISH technique was first employed to determine the length of telomeres,[133] several further studies have been described in this respect, leading to development and commercialization of PNA-FISH kits for the identification of bacteria, yeast and other micro-organisms. The outstanding properties of PNAs promoted their use as bioreceptors driving to a dramatic increase of microarray, optical and electrochemical biosensors.

1.6.1 PNA-based electrochemical biosensor

One of the first example of PNA used as bioreceptor was reported by Wang *et al.* in 1996, where a 15-mer PNA was adsorbed onto a carbon paste electrode and by employment of $Co(phen)_3^{3+}$ as redox indicator they were able to detect a specific mutation in the p53 gene. Generally, the monitoring of the hybridization event between the PNA probe and a target DNA allows an electrochemically quantification of DNA and a discrimination of mismatches.
Voltammetry methods were widely used as electrochemical transductions in combination with different designed PNAs for a fast screening of sequence-specific DNA.[134–141] In particular, a large number of examples are reported in literature regarding the use of square wave voltammetry to detect electrochemical changes upon PNA/DNA hybridization. Aoki and Umezawa

reported a sensitive biosensor, i.e. femtomolar (~10^{-15} M) level, in which a gold electrode was modified with self-assembled monolayers of a 13-mer PNA probe and 8-amino-1-octanethiol for the detection of complementary oligonucleotides.[141] Square wave voltammetry technique was employed by Fu and colleagues with the aim to detect at the same time miRNA-21 and miRNA-155 from cancer cells, by using PNA-modified gold electrode combined with the target-catalyzed hairpin assembly (CHA) strategy. Target miRNAs were able to selectively trigger the CHA between two hairpins with one ferrocene (Fc) or methylene blue (MB) labelled that were later specifically captured by the PNA-21 or PNA-155 probes generating an enhanced electrochemical signals thanks to the proximity of the labels to the surface (**Figure 1.13**).[140]

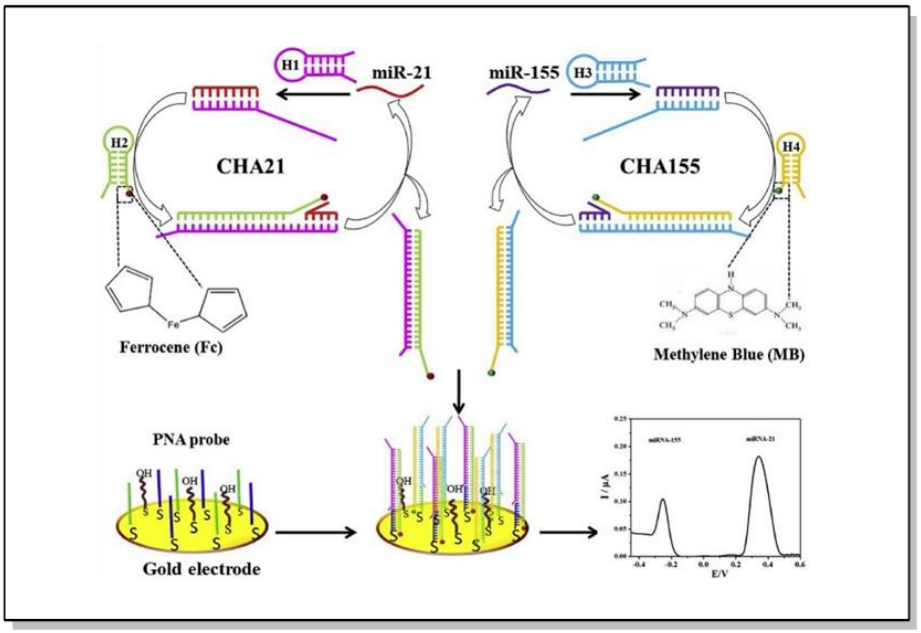


Figure 1.13. Schematic illustration of electrochemical and simultaneous detection of miRNA21 and miRNA155. Reprinted with permission from ref. [140]: Fu et al. © 2019 Elsevier B.V. All rights reserved.

Hejazi *et al.* exploited differential pulse voltammetry (DPV) of methylene blue signal in order to detect a single nucleotide polymorphism (SNP) in the UGT1A9 gene by immobilizing 13-mer ssPNA onto a gold electrode. The biosensor showed a LOD of 22 nM.[142] An interesting example of immobilization-free biosensor was reported by Luo and Hsing.[143] In this strategy, the simultaneous detection of sequence-specific DNA and SNP was achieved by using a neutrally charged PNA probe labeled with Fc and MB (electroactive indicator) and a

negatively charged indium tin oxide electrode ITO. Once the negative duplex DNA/PNA was formed, the electrostatic repulsion between the latter and the negative surface of ITO electrode hampered the approaching of the electroactive indicators to the electrode leading to a suppression of the electrochemical signal.

1.6.2 PNA-based optical biosensor

Optical biosensors represent a beneficial alternative to other analytical techniques due to the label-free, real-time and direct determination of several chemical and biological targets.

A fluorescence based PNA biosensor was first reported by Wang and coworkers in 1999, which provided for the use of PNA as light-up probe. The PNA oligomer was labelled with a cyanine dye thiazole orange that upon binding to complementary DNA gave rise to a fluorescence increase of ~50-fold compared to a free probe.[144] This technique was mainly used for quantification of real time PCR and for the discrimination of single point mutations.[145,146] In addition to PNA-based fluorescence probes, more elaborated transductors have been developed over the years.

Optical fiber-based biosensors have arisen as promising alternatives for label-free biosensing thanks to their compactness, light weight, immunity to electromagnetic interference and the compatibility to optoelectronic devices.[147,148] Candiani *et al.* described a novel DNA sensing approach in which a peptide nucleic acid (PNA)-functionalized microstructured optical fiber (MOF) Bragg grating is used for targeting a DNA sequence involved in the cystic fibrosis disease bearing the single point mutation W1282X. In order to obtain signal amplifications, the authors exploited a sandwich-like system where oligonucleotide-functionalized gold nanoparticles (ON-AuNPs) were infiltrated after the permeation of the fiber capillaries with a solution of the mutated DNA. Several experiments were performed on the same fiber using identical experimental conditions showing a clear shift of the reflected high order mode of a Bragg grating for a 100 nM DNA solution and at the same time the opportunity of the reuse of the sensor.[149] This method was then extended to the direct detection of genomic DNA in certified samples containing different percentages of a genetically modified soybean (Roundup Ready), without prior PCR amplication.[147]

Surface Plasmon Resonance (SPR) is a representative label-free detection technique for monitoring biomolecular interactions in real-time. Introduced in the early 1990s, this optical technique was widely used to study DNA hybridization with PNA probes properly linked to a thin metal layer (typically gold or silver). Indeed, the change in the refractive index associated to the alterations of mass at the metal surface in response to biomolecular interactions, allows to control the DNA/PNA hybridization process. Sawata and colleagues reported a BIAcore™ 2000 biosensor based on SPR to detect a dsDNA encoding verotoxin 2 of *Escherichia coli* amplified by polymerase chain reaction (PCR) by using a N-term biotinylated PNA. The functionalization at the N-terminus was required for the immobilization of the PNA on sensor chip bearing streptavidin. The target DNA was detected with a LOD of 7.5 pmol over a range of 40–160 nM with a good reproducibility.[150] The peculiar property of PNA to detect SNPs was also exploited by SPR-based biosensors in order to identify several point mutations involved in different diseases. For instance, Sato *et al.* used a 15-mer PNA probe for recognizing a G-G mismatch in codon 2 of the K-*ras* gene for early detection of malignancies, and a high capability to detect mismatch in 11-mer DNA was reported.[151] Corradini and coworkers exploited the great ability of PNA for SNP recognition with the aim to identify W1282X point mutation of the cystic fibrosis gene. In this case, the PNA probe used to detect the target DNA consisted of a strand containing three chiral D-lysine monomers (the so-called 'Chiral Box'), which induce a preferential handedness of DNA/PNA duplexes. By using the SPR-based BIA technology, it was found that the selectivity observed for the chiral PNA was much higher than that with the control achiral PNA and moreover, the chiral probe was also able to discriminate between homozygous and heterozygous individuals.[152] In 2000, Surface-plasmon field-enhanced fluorescence spectroscopy (SPFS) was described as a novel technique for sensor applications.[153] One year later, this combination of surface plasmon and fluorescence spectroscopy was used by Knoll group to investigate the kinetics of DNA/DNA and DNA/PNA interactions.[154] The binding behavior of DNA/PNA was also studied by using a combination of electrochemical and SPFS techniques demonstrating the feasibility of this approach for the detection of DNA hybridization.[155,156] Interestingly, the high surface sensitivity provided by SPFS can be exploited in this respect together with the differences in the properties of DNA and PNA. Chu *et al.* described a method for the detection of DNA in which a positively charged polymer was deposited on the sensor surface and the different electrostatic

features of a fluorescently labeled PNA and target DNA allowed the detection of the latter with a LOD of 200 pM. Through this approach, labelling and/or immobilization of the DNA was not required.[157]
The use of sandwich type assays using gold nanoparticle allowed to improve sensitivity and develop ultrasensitive methods; this led to the Nanoparticle-Enhanced Surface Plasmon Resonance Imaging (NESPRI) version of this technique, developed by Spoto and collaborators,[158] which later on was found suitable to direct genomic DNA detection,[159] and ultimately, thanks to the effort of the ULTRAPLACAD EU project, led to the development of a suitable method for the detection of KRAS and NRAS mutations in plasma, in the so-called "liquid biopsy" approach.[160]

1.6.3 Nanostructured PNA biosensors

In the recent years, the use of nanomaterials in biosensor technology represented a remarkable strategy exploited to improve the analytical performances in both optical and electrochemical devices. Graphene is one of the most employed nanomaterial for biosensor design, thanks to its outstanding properties which include high thermal and electrical conductivity, intrinsic mobility, great flexibility and mechanical strength. Moreover, graphene can be functionalized to obtain, for instance, graphene oxide (GO), an oxidized form whose main advantage is to be dispersible in water. Graphene oxide was largely exploited as biosensing platform by using PNA probes for the detection of DNA, RNA and miRNAs.[161–164] In 2013, Guo and coworkers exploited the ability of graphene oxide as quencher to assemble a biosensor with a fluorescent dye-labelled PNA probe for detecting ssDNA with a high sensitivity and specificity. The labelled PNA probe was physically adsorbed on the GO sheet and after the PNA-DNA hybridization the fluorescence was recovered allowing the detection of the target DNA with a LOD of 8.0×10^{-10} mol L^{-1} and the discrimination of single-mismatched and non-complementary sequences.[165] A similar PNA-GO sensing platform was exploited by Lee and collaborators to detect short dsDNA, bypassing the denaturation step thanks to the invasion performed by the PNA probe towards target dsDNA that leads to a gradual strand displacement and therefore to the formation of the PNA/DNA duplex.[164]
The simultaneous detection of three different miRNAs in a living cell was achieved by Ryoo and colleagues who assembled a nanosized graphene oxide (NGO) biosensor using peptide nucleic acids as probes (**Figure 1.14**). The

presented sensor allowed to monitor multiple miRNAs with high sequence specificity and with the detection limit as low as ~1 pM.[163]

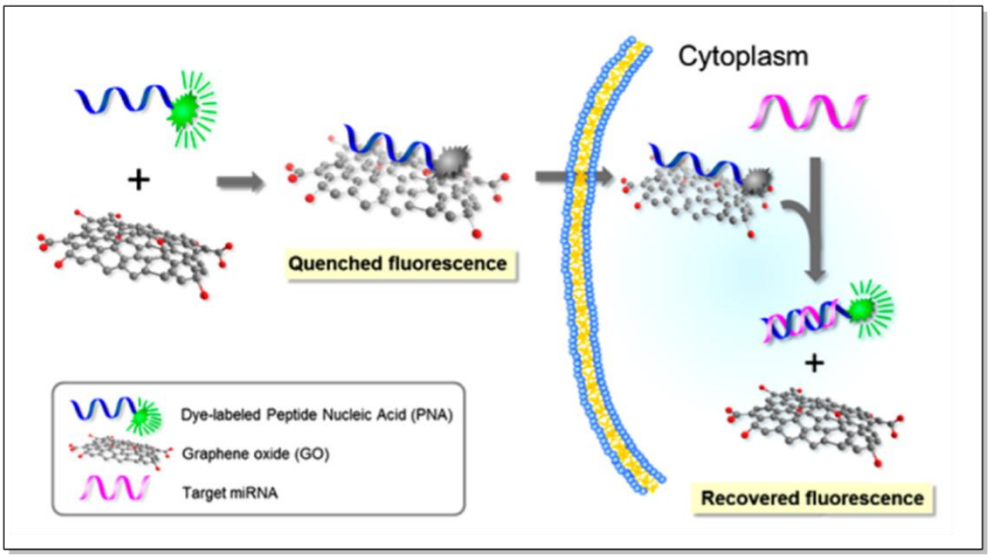


Figure 1. 14. Schematic of the PNA-NGO based miRNA biosensor. The fluorescence signal gets recovered when the fluorescent dye-labeled probes initially adsorbed onto the surface of NGO detach from NGO and hybridize with a target miRNA. Adapted with permission from ref. [155]: Ryoo et al. © 2013, American Chemical Society.

Tungsten disulfide (WS_2) nanosheet, a two-dimensional nanomaterial with graphite-like structure, was used as alternative to graphene oxide since the preparation of the former does not involve oxide treatment that can affect the performance of the fluorescence quenching.[166] In 2014, WS_2 nanosheet was exploited for detection of target ssDNA by assembling a platform that employs a FAM-labelled PNA at the C-terminal. The dye-labelled PNA probe was adsorbed on the surface of the WS_2 resulting in quenching of the fluorescence, which was then restored in the presence of the target DNA. The limit of detection exhibited by this DNA sensor was 500 pM.[167]

In a similar manner, carbon nitride nanosheets (CNNS) and a nano metal-organic framework were used in combination with fluorescent PNA probes with the aim to monitor the expression level of microRNAs in living cells.[168,169]

Yang *et al.* assembled a highly sensitive FRET-based DNA sensor by exploiting PNA probes and red color quantum dots (QDs). The system consisted in QDs conjugated with a sandwiched hybrid composed of a capture probe, two different type of reporter probes functionalized with Cy5 dyes at different

positions and target DNA. The developed sensor allowed the detection of target DNAs with a pico-molar detection limit.[170]
Gold nanorod-based platform was developed by Tadimety and colleagues for the sequence-specific detection of circulating tumor DNA (ctDNA) point mutations, namely G12V mutation in the KRAS gene. PNA probes were conjugated to the gold nanorods and by localized surface plasmon resonance (LSPR) measures over a range of concentrations of synthetic ctDNA they were able to detect ctDNA with a limit of detection of 2ng/mL and without the need for amplification or fluorescence labeling.[171]
An example of electrochemical biosensor exploiting nanomaterials features was recently reported by Fortunati *et al.,* who designed an amperometric genosensor for the determination of non-amplified DNA in genetically modified soy. In this platform a first PNA capture probe is covalently conjugated on the surface of single walled carbon nanotubes-screen printed electrodes (SWCNT-SPEs) and recognizes a portion of the target DNA. Next, the hybridization of a different portion of the DNA with a biotin-labeled PNA signalling probe allows to measure current signal by an Alkaline Phosphatase-streptavidin conjugate (ALP-Strp). The enzymatic substrate hydroquinone diphosphate (HQDP) was used to carry out the electrochemical detection. This platform provided a LOD of 64 pM and a high selectivity towards single-base mismatched and scrambled sequences.[172]

1.7 Theranostic applications

Theranostics represent the next generation of technologies to be used in the clinic. They provide a means to deliver a therapy while simultaneous reporting on the *in vivo* environment. These technologies would serve to reduce clinical costs, as only one platform would be required for treatment and detection instead of two distinct platforms. PNA is especially attractive in these applications, because of its proven antisense/antigene properties and the ease with which its backbone can be functionalized with labels.
A representative example was reported by Hwang *et al*. in 2017 that developed hyaluronic acid (HA) – graphene oxide (GO) sheets loaded with a dye-labeled PNA (HGP21).[173] Hyaluronic acid, a CD44 receptor-specific hydrophilic polysaccharide, was grafted onto GO surface to address the aggregation behavior of the latter in ionic solutions, such as PBS or cell culture medium. This platform was able to be internalized into MDA-MB-231 breast cancer cells by

CD44 receptor-mediated endocytosis and to target the oncogenic miR21, both visualizing the miRNA and inhibiting its tumorigenicity at the same time. Specifically, they functionalized the antisense miR-21 PNA with the Cy3, a cyanine fluorescing in the green-yellow region, and then the PNA probe was adsorbed onto GO surface leading to fluorescence quenching. Once the HGP21 nanoplatform entered the cytoplasm of MDA-MB-231 breast cancer cells, the Cy3-antiPNA21 reacted with the endogenous miR-21, resulting in restored fluorescence. Since no significant fluorescence was observed in the HeLa cells, where the miRNA-21 expression is lower in comparison with MDA-MB-231 cells, the process results highly sequence specific and sensitive. In fact, the expression of CD44 is comparable for both HeLa and MDA-MB-231 cells, indicating that the higher fluorescence in MDA-MB-231 cells is due to the sequence specificity of anti-PNA-21 toward endogenous miR-21. Lastly, a significant miR-21 knockdown was observed in the MDA-MB-231 cells treated with HGP21, nearly equivalent to that obtained using Lipofectamine® as transfecting agent.

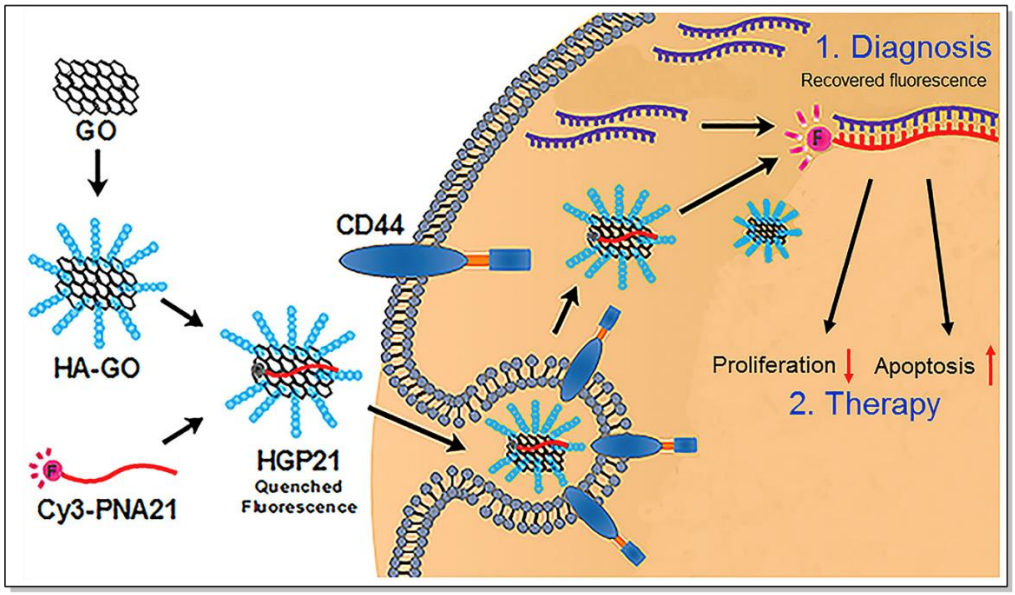


Figure 1.15. Schematic illustration of HGP synthesis and its theranostic effect by targeting oncogenic miR-21. From ref. [165] © 2017 Elsevier Ltd. All rights reserved.

From the above, it appears the broad applicability of PNA technology. While PNAs are nearly 30 years old, the interest towards the development of novel applications remained still high. The outstanding hybridization properties, the chemical versatility along with the exceptional biological stability of PNA

represent a key point for bringing important innovations in both sensing and therapeutic fields. In fact, considerable steps forward have been made to render this family of DNA mimic suitable for practical applications. The deep investigation of intrinsic features of PNA over time has provided the rules for triggering of both the affinity and selectivity for relevant cognate nucleic acids, by affording tailored chemical modifications. Moreover, the combination of PNA with a large number of nanomaterials has led to the development of several delivery platforms for leveraging at full the biological activity of these promising synthetic oligonucleotides.

1.8 References

1. a) Crick, F. & Watson, J. © 1953 Nature Publishing Group. *Nature* **171,** 737–738 (1953). b) Watson, J. & Crick, F. (2010). 1953. A structure for deoxyribose nucleic acid. In L. Garwin & T. Lincoln (Ed.), A Century of Nature (pp. 82-84). Chicago: University of Chicago Press.

2. Paterson, B. M., Roberts, B. E. & Kuff, E. L. Structural gene identification and mapping by DNA.mRNA hybrid-arrested cell-free translation. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 4370–4374 (1977).

3. Zamecnik, P. C. & Stephenson, M. L. Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc. Natl. Acad. Sci. U. S. A.* **75,** 280–284 (1978).

4. C. Marwick. First antisense drug will treat CMV retinitis. *J. Am. Med. Assoc.* **280,** 871 (1998).

5. Wong, E. & Goldberg, T. Mipomersen (Kynamro): A novel antisense oligonucleotide inhibitor for the management of homozygous familial hypercholesterolemia. *Pharm. Ther.* **39,** 119–122 (2014).

6. Keefe, A. D., Pai, S. & Ellington, A. Aptamers as therapeutics. *Nat. Rev. Drug Discov.* **9,** 537–550 (2010).

7. Burnett, J. C. & Rossi, J. J. RNA-based therapeutics: Current progress and future prospects. *Chem. Biol.* **19,** 60–71 (2012).

8. Christoffersen, R. E. & Marr, J. J. Ribozymes as Human Therapeutic Agents. *J. Med. Chem.* **38,** 2023–2037 (1995).

9. Eckstein, F. Nucleoside phosphorothioates. *J. Am. Chem. Soc.* **2303,** 4718–4723 (1970).

10. Matsukura, M., Shinozukat, K., Zont, G., Mitsuya, H., Reitz, M., Cohent, J. S. & Broder, S. Phosphorothioate analogs of oligodeoxynucleotides : Inhibitors of replication and cytopathic effects of human immunodeficiency virus. *Proc. Natl Acad. Sci. USA* **84,** 7706–7710 (1987).

11. Campbell, J. M., Bacon, T. A. & Wickstrom, E. Oligodeoxynucleoside phosphorothioate stability in subcellular extracts , culture media , sera and cerebrospinal fluid. *J. Biochem. Biophys. Methods* **20,** 259–267 (1990).

12. Phillips, M.I. Zhang, Y. C. Basic principles of using antisense oligonucleotides in vivo. *Methods Enzymol.* **313,** 46–56 (2000).

13. Levin, A. A. A review of issues in the pharmacokinetics and toxicology of phosphorothioate antisense oligonucleotides. *Biochim. Biophys. Acta* **1489,** 69–84 (1999).

14. Hoke, G. D., Draper, K., Freier, S. M., Gonzalez, C., Driver, V. B., Zounes, M. C., Ecker, D. J., Pharmaceuticals, I. & Avenue, F. Effects of phosphorothioate capping on antisense oligonucleotide stability , hybridization and antiviral efficacy versus herpes simplex virus infection. *Nucleic Acid Res* **19,** 5743–5748 (1991).

15. S.T. Crooke. Progress in Antisense Technology : The End of the Beginning. *METHODS Enzymol.* **313,** 3–45 (1999).

16. Shoji, Y., Akhtar, S., Periasamy, A., Herman, B. & Juliano, R. L. Mechanism of cellular uptake of modified oligodeoxynucleotides containing methylphosphonate linkages. *Nucleic Acids Res* **19,** 5543–5550 (1991).

17. Crooke, S. T. Therapeutic applications of oligonucleotides. *Annu. Rev. Pharmacol. Toxicol* **32,** 329–76 (1992).

18. Guschlbauer, W. & Jankowskit, K. Nucleoside conformation is determined by the electronegativity of the sugar substituent. *Nucleic Acids Res.* **8,** 1421–1433 (1980).

19. Baker, B. F., Lot, S. S., Condon, T. P., Cheng-flournoy, S., Lesnik, E. A., Sasmor, H. M. & Bennett, C. F. 2' ~O - (2-Methoxy) ethyl-modified Anti-intercellular Adhesion Molecule 1 (ICAM-1) Oligonucleotides Selectively Increase the ICAM-1 mRNA Level and Inhibit Formation of the ICAM-1 Translation Initiation Complex in Human Umbilical Vein Endothelial Cell.

J. Biol. Chem. **272,** 11994–12000 (1997).

20. Mcgowan, M. P., Tardif, J., Ceska, R., Burgess, L. J., Soran, H., Gouni-berthold, I., Wagener, G. & Chasan-taber, S. Randomized , Placebo-Controlled Trial of Mipomersen in Patients with Severe Hypercholesterolemia Receiving Maximally Tolerated Lipid-Lowering Therapy. *PLoS One* **7,** 1–10 (2012).

21. P. Stirchak, E., Summerton, J. & D. Weller, D. Morpholino antisense oligomers : the case for an RNase H-independent structural type. *Biochim. Biophys. Acta* **1489,** 141–158 (1999).

22. Vester, B. & Wengel, J. LNA (Locked Nucleic Acid): High-affinity targeting of complementary RNA and DNA. *Biochemistry* **43,** 13233–13241 (2004).

23. Wengel, J., Koshkin, A., Singh, S. K., Nielsen, P., Meldgaard, M., Rajwanshi, V. K., Kumar, R., Skouv, J., Nielsen, C. B., Jacobsen, J. P., Jacobsen, N. & Olsen, C. E. LNA (Locked Nucleic Acids) Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Nucleosides and Nucleotides* **18,** 1365–1370 (1999).

24. Obika, S., Nanbu, D., Hari, Y., Andoh, J. I., Morio, K. I., Doi, T. & Imanishi, T. Stability and structural features of the duplexes containing nucleoside analogues with a fixed N-type conformation, 2'-O,4'-C-methyleneribonucleosides. *Tetrahedron Lett.* **39,** 5401–5404 (1998).

25. Kurreck, J., Wyszko, E., Gillen, C. & Erdmann, V. A. Design of antisense oligonucleotides stabilized by locked nucleic acids. *Nucleic Acids Res.* **30,** 1911–1918 (2002).

26. Ørum, H., Jakobsen, M. H., Koch, T., Vuust, J. & Borre, M. B. Detection of the factor V Leiden mutation by direct allele-specific hybridization of PCR amplicons to photoimmobilized locked nucleic acids. *Clin. Chem.* **45,** 1898–1905 (1999).

27. Faria, M., Spiller, D. G., Dubertret, C., Nelson, J. S., White, M. R. H., Scherman, D., Hélène, C. & Giovannangeli, C. Phosphormidate oligonucleotides as potent antisense molecules in cells and in vivo. *Nat. Biotechnol.* **19,** 40–44 (2001).

28. Skorski, T., Perrotti, D., Nieborowska-Skorska, M., Gryaznov, S. & Calabretta, B. Antileukemia effect of c-myc N3′→P5′ phosphoramidate antisense oligonucleotides in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 3966–

3971 (1997).

29. Wang, J., Verbeure, B., Luyten, I., Lescrinier, E., Froeyen, M., Hendrix, C., Rosemeyer, H., Seela, F., Van Aerschot, A. & Herdewijn, P. Cyclohexene nucleic acids (CeNA): Serum stable oligonucleotides that activate RNase H and increase duplex stability with complementary RNA. *J. Am. Chem. Soc.* **122,** 8595–8602 (2000).

30. Renneberg, D., Bouliong, E., Reber, U., Schuümperli, D. & Leumann, C. J. Antisense properties of tricyclo-dna. *Nucleic Acids Res.* **30,** 2751–2757 (2002).

31. Shangguan, D., Meng, L., Cao, Z. C., Xiao, Z., Fang, X., Li, Y., Cardona, D., Witek, R. P., Liu, C. & Tan, W. Identification of liver cancer-specific aptamers using whole live cells. *Anal. Chem.* **80,** 721–728 (2008).

32. Famulok, M., Hartig, S. & Mayer, G. Functional Aptamers and Aptazymes in Biotechnology , Diagnostics , and Therapy. *Chem. Rev.* **107,** 3715–3743 (2007).

33. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346,** 818–822 (1990).

34. Kang, H. S., Huh, Y. M., Kim, S. & Lee, D. Isolation of RNA Aptamers Targeting HER -2-overexpressing Breast Cancer Cells Using Cell-SELEX. *Bull. Korean Chem. Soc.* **30,** 1827–1831 (2009).

35. Chen, H. W., Medley, C. D., Sefah, K., Shangguan, D., Tang, Z., Meng, L., Smith, J. E. & Tan, W. Molecular Recognition of Small-Cell Lung Cancer Cells Using Aptamers. *ChemMedChem* **3,** 991–1001 (2008).

36. Keefe, A. D. & Schaub, R. G. Aptamers as candidate therapeutics for cardiovascular indications. *Curr. Opin. Pharmacol.* **8,** 147–152 (2008).

37. Lee, J. F., Stovall, G. M. & Ellington, A. D. Aptamer therapeutics advance. *Curr. Opin. Chem. Biol.* **10,** 282–289 (2006).

38. Zhou, B. & Wang, B. Pegaptanib for the treatment of age-related macular degeneration. *Exp. Eye Res.* **83,** 615–619 (2006).

39. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391,** 806–811 (1998).

40. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. & Tuschl, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411,** 1–5 (2001).

41. Layzer, J. M., Mccaffrey, A. P., Tanner, A. K., Huang, Z. A. N., Kay, M. A. & Sullenger, B. A. In vivo activity of nuclease-resistant siRNAs. *RNA* **10,** 766–771 (2004).

42. Gil, W., Albuquerque, R. J. C., Kleinman, M. E., Tarallo, V., Greco, A., Nozaki, M., Green, M. G., Baffi, J. Z., Ambati, B. K., Falco, M. De, Alexander, J. S., Brunetti, A., Falco, S. De & Ambati, J. Small interfering RNA-induced TLR3 activation inhibits blood and lymphatic vessel growth. *PNAS* **106,** 1–6 (2009).

43. Khvorova, A. & Watts, J. K. The chemical evolution of oligonucleotide therapies of clinical utility. *Nat. Biotechnol.* **35,** 238–248 (2017).

44. Matranga, C., Tomari, Y., Shin, C., Bartel, D. P. & Zamore, P. D. Passenger-Strand Cleavage Facilitates Assembly of siRNA into Ago2-Containing RNAi Enzyme Complexes. *Cell* **123,** 607–620 (2005).

45. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35,** 849–857 (1983).

46. James, H. A. & Gibson, I. The therapeutic potential of ribozymes. *Blood* **91,** 371–382 (1998).

47. Doudna, J. A. & Cech, T. R. Site-specific RNA self-cleavage The chemical repertoire of natural ribozymes. *Nature* **418,** 222–228 (2002).

48. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science (80-.).* **254,** 1497–1500 (1991).

49. Haaima, G., Lohse, A., Buchardt, O. & Nielsen, P. E. Peptide nucleic acids (PNAs) containing thymine monomers derived from chiral amino acids: Hybridization and solubility properties of D-lysine PNA. *Angew. Chemie (International Ed. English)* **35,** 1939–1942 (1996).

50. Nielsen, P. E. PNA technology. *Appl. Biochem. Biotechnol. - Part B Mol. Biotechnol.* **26,** 233–248 (2004).

51. Feinbaum, R., Ambros, V. & Lee, R. The C. elegans Heterochronic Gene lin-

4 Encodes Small RNAs with Antisense Complementarity to lin-14. *Cell* **116,** 843–854 (2004).

52. Slack, F. J., Basson, M., Liu, Z., Ambros, V., Horvitz, H. R. & Ruvkun, G. The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol. Cell* **5,** 659–669 (2000).

53. Reinhart, B. J., Slack, F. J., Basson, M., Pasquienelll, A. E., Bettlnger, J. C., Rougvle, A. E., Horvitz, H. R. & Ruvkun, G. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403,** 901–906 (2000).

54. Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75,** 855–862 (1993).

55. Borchert, G. M., Lanier, W. & Davidson, B. L. RNA polymerase III transcribes human microRNAs. *Nat. Struct. Mol. Biol.* **13,** 1097–1101 (2006).

56. Ruby, J. G., Jan, C. H. & Bartel, D. P. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448,** 83–86 (2007).

57. Fabian, M. R., Sonenberg, N. & Filipowicz, W. Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* **79,** 351–379 (2010).

58. Gregory, R. I., Chendrimada, T. P., Cooch, N. & Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123,** 631–640 (2005).

59. Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F. & Croce, C. M. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 15524–15529 (2002).

60. Björk, J. K., Sandqvist, A., Elsing, A. N., Kotaja, N. & Sistonen, L. miR-18, a member of Oncomir-1, targets heat shock transcription factor 2 in spermatogenesis. *Development* **137,** 3177–3184 (2010).

61. Jia, M., Shi, Y., Li, Z., Lu, X. & Wang, J. MicroRNA-146b-5p as an oncomiR promotes papillary thyroid carcinoma development by targeting CCDC6.

Cancer Lett. **443,** 145–156 (2019).

62. Zhang, L., Liao, Y. & Tang, L. MicroRNA-34 family: A potential tumor suppressor and therapeutic candidate in cancer. *J. Exp. Clin. Cancer Res.* **38,** 1–13 (2019).

63. Hong, B. S., Ryu, H. S., Kim, N., Kim, J., Lee, E., Moon, H., Kim, K. H., Jin, M. S., Kwon, N. H., Kim, S., Kim, D., Chung, D. H., Jeong, K., Kim, K., Kim, K. Y., Lee, H. B., Han, W., Yun, J., Kim, J. Il, Noh, D. Y. & Moon, H. G. Tumor suppressor miRNA-204-5p regulates growth, metastasis, and immune microenvironment remodeling in breast cancer. *Cancer Res.* **79,** 1520–1534 (2019).

64. Svoronos, A. A., Engelman, D. M. & Slack, F. J. OncomiR or tumor suppressor? The duplicity of MicroRNAs in cancer. *Cancer Res.* **76,** 3666–3670 (2016).

65. Fabani, M. M. & Gait, M. J. miR-122 targeting with LNA/2′-O-methyl oligonucleotide mixmers, peptide nucleic acids (PNA), and PNA-peptide conjugates. *Rna* **14,** 336–346 (2008).

66. Torres, A. G., Threlfall, R. N. & Gait, M. J. Potent and sustained cellular inhibition of miR-122 by lysine-derivatized peptide nucleic acids (PNA) and phosphorothioate locked nucleic acid (LNA)/2'-O-methyl (OMe) mixmer anti-miRs in the absence of transfection agents. *Artif. DNA PNA XNA* **2,** 71–78 (2011).

67. Stein, C. A., Hansen, J. B., Lai, J., Wu, S. J., Voskresenskiy, A., Høg, A., Worm, J., Hedtjärn, M., Souleimanian, N., Miller, P., Soifer, H. S., Castanotto, D., Benimetskaya, L., ørum, H. & Koch, T. Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents. *Nucleic Acids Res.* **38,** 1–8 (2009).

68. Eis, P. S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M. F., Lund, E. & Dahlberg, J. E. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 3627–3632 (2005).

69. Kluiver, J., Poppema, S., de Jong, D., Blokzijl, T., Harms, G., Jacobs, S., Kroesen, B. J. & van den Berg, A. BIC and miR-155 are highly expressed in Hodgkin, primary mediastinal and diffuse large B cell lymphomas. *J. Pathol.* **207,** 243–249 (2005).

70. Due, H., Svendsen, P., Bødker, J. S., Schmitz, A., Bøgsted, M., Johnsen, H. E., El-Galaly, T. C., Roug, A. S. & Dybkær, K. MiR-155 as a Biomarker in B-Cell

Malignancies. *Biomed Res. Int.* **2016,** (2016).

71. 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. Efficient inhibition of miR-155 function in vivo by peptide nucleic acids. *Nucleic Acids Res.* **38,** 4466–4475 (2010).

72. Vigorito, E., Perks, K. L., Abreu-Goodger, C., Bunting, S., Xiang, Z., Kohlhaas, S., Das, P. P., Miska, E. A., Rodriguez, A., Bradley, A., Smith, K. G. C., Rada, C., Enright, A. J., Toellner, K. M., MacLennan, I. C. M. & Turner, M. microRNA-155 Regulates the Generation of Immunoglobulin Class-Switched Plasma Cells. *Immunity* **27,** 847–859 (2007).

73. Thai, T., Calado, D., Casola, S., Ansel, K., Xiao, C., Xue, Y., Murphy, A., Frendewey, D., Valenzuela, D., Kutok, J., Schmidt-Supprian, M., Rajewsky, N., Yancopoulos, G., Rao, A. & Rajewsky, K. Regulation of the Germinal Center. *Science (80-.).* **316,** 604–609 (2007).

74. Babar, I. A., Cheng, C. J., Booth, C. J., Liang, X., Weidhaas, J. B., Saltzman, W. M. & Slack, F. J. Nanoparticle-based therapy in an in vivo microRNA-155 (miR-155)-dependent mouse model of lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 1695–1704 (2012).

75. Cheng, C. J., Bahal, R., Babar, I. A., Pincus, Z., Barrera, F., Liu, C., Svoronos, A., Braddock, D. T., Glazer, P. M., Engelman, D. M., Saltzman, W. M. & Slack, F. J. MicroRNA silencing for cancer therapy targeted to the tumour microenvironment. *Nature* **518,** 107–110 (2015).

76. Vaupel, P., Kallinowski, F. & Okunieff, P. Blood Flow, Oxygen and Nutrient Supply, and Metabolic Microenvironment of Human Tumors: A Review. *Cancer Res.* **49,** 6449–6465 (1989).

77. Milani, R., Brognara, E., Fabbri, E., Manicardi, A., Corradini, R., Finotti, A., Gasparello, J., Borgatti, M., Cosenza, L. C., Lampronti, I., Dechecchi, M. C., Cabrini, G. & Gambari, R. Targeting miR-155-5p and miR-221-3p by peptide nucleic acids induces caspase-3 activation and apoptosis in temozolomide-resistant T98G glioma cells. *Int. J. Oncol.* **55,** 59–68 (2019).

78. Gasparello, J., Papi, C., Gambari, R., Finotti, A., Gambari, L., Rozzi, A., Manicardi, A. & Corradini, R. High Levels of Apoptosis Are Induced in the Human Colon Cancer HT-29 Cell Line by Co-Administration of Sulforaphane and a Peptide Nucleic Acid Targeting miR-15b-5p. *Nucleic Acid Ther.* **30,** 164–174 (2020).

79. Bessler, H. & Djaldetti, M. Broccoli and human health: immunomodulatory effect of sulforaphane in a model of colon cancer. *Int. J. Food Sci. Nutr.* **69,** 946–953 (2018).

80. Tafakh, M. S., Saidijam, M., Ranjbarnejad, T., Malih, S., Mirzamohammadi, S. & Najafi, R. Sulforaphane, a Chemopreventive Compound, Inhibits Cyclooxygenase-2 and Microsomal Prostaglandin e Synthase-1 Expression in Human HT-29 Colon Cancer Cells. *Cells Tissues Organs* **206,** 46–53 (2019).

81. Langner, E., Lemieszek, M. K. & Rzeski, W. Lycopene, sulforaphane, quercetin, and curcumin applied together show improved antiproliferative potential in colon cancer cells in vitro. *J. Food Biochem.* **43,** 1–12 (2019).

82. Koppelhus, U. & Nielsen, P. E. Cellular delivery of peptide nucleic acid (PNA). *Adv. Drug Deliv. Rev.* **55,** 267–280 (2003).

83. Hanvey, J. C., Peffer, N. J., Bisi, J. E., Thomson, S. A., Cadilla, R., Josey, J. A., Ricca, D. J., Hassman, C. F., Bonham, M. A., Au, K. G., Carter, S. G., Bruckenstein, D. A., Boyd, A. L., Noble, S. A. & Babiss, L. E. Antisense and Antigene Properties of Peptide Nucleic Acids. *Science (80-.).* **258,** 1481–1485 (1992).

84. Mercurio, S., Cauteruccio, S., Manenti, R., Candiani, S., Scarì, G., Licandro, E. & Pennati, R. Mir-7 knockdown by peptide nucleic acids in the ascidian ciona intestinalis. *Int. J. Mol. Sci.* **20,** 5127 (2019).

85. Mercurio, S., Cauteruccio, S., Manenti, R., Candiani, S., Scarì, G., Licandro, E. & Pennati, R. Exploring mir-9 involvement in ciona intestinalis neural development using peptide nucleic acids. *Int. J. Mol. Sci.* **21,** (2020).

86. Wang, G., Xu, X., Pace, B., Dean, D. A., Glazer, P. M., Chan, P., Goodman, S. R. & Shokolenko, I. Peptide nucleic acid (PNA) binding-mediated induction of human γ -globin gene expression. *Nucleic Acids Res.* **27,** 2806–2813 (1999).

87. Faruqi, F. A., Egholm, M. & Glazer, P. M. Peptide nucleic acid-targeted mutagenesis of a chromosomal gene in mouse cells. *Proc. Natl Acad. Sci. USA* **95,** 1398–1403 (1998).

88. Doyle, D. F., Braasch, D. A., Simmons, C. G., Janowski, B. A. & Corey, D. R. Inhibition of Gene Expression Inside Cells by Peptide Nucleic Acids : Effect of mRNA Target Sequence , Mismatched Bases , and PNA Length.

Biochemistry **40,** 53–64 (2001).

89. Janowski, B. A., Kaihatsu, K., Huffman, K. E., Schwartz, J. C., Ram, R., Hardy, D., Mendelson, C. R. & Corey, D. R. Inhibiting transcription of chromosomal DNA with antigene peptide nucleic acids. *Nat. Chem. Biol.* **1,** 210–215 (2005).

90. Dragulescu-andrasi, A., Rapireddy, S. & He, G. Cell-Permeable Peptide Nucleic Acid Designed to Bind to the 5 ′ -Untranslated Region of E-cadherin Transcript Induces Potent and Sequence-Specific Antisense Effects. *J. Am. Chem. Soc.* **128,** 16104–16112 (2006).

91. Turner, J. J., Ivanova, G. D., Verbeure, B., Williams, D., Arzumanov, A., Abes, S., Lebleu, B. & Gait, M. J. Cell-penetrating peptide conjugates of peptide nucleicacids (PNA) as inhibitors of HIV-1 Tat-dependenttrans-activation in cells. 6837–6849 (2005).

92. Cogoi, S., Codognotto, A., Rapozzi, V. & Xodo, L. E. ANTIGENE PROPERTY OF PNA CONJUGATED TO THE NUCLEAR LOCALIZATION SIGNAL PEPTIDE. *Nucleosides , Nucleotides Nucleic Acids* **24,** 971–974 (2005).

93. Chaubey, B., Tripathi, S., Ganguly, S., Harris, D., Casale, R. A. & Pandey, V. N. A PNA-transportan conjugate targeted to the TAR region of the HIV-1 genome exhibits both antiviral and virucidal properties. *Virology* **331,** 418–428 (2005).

94. Turner, J. J., Ivanova, G. D., Owen, D., Williams, D., Arzumanov, A., Clair, P., Gait, M. J. & Lebleu, B. Efficient splicing correction by PNA conjugation to an R 6 -Penetratin delivery peptide. *Nucleic Acids Res.* **35,** 4495–4502 (2007).

95. Chaubey, B., Tripathi, S. & Pandey, V. N. Single Acute-Dose and Repeat-Doses Toxicity of anti-HIV-1 PNATAR–Penetratin Conjugate after Intraperitoneal Administration to Mice. *Oligonucleotides* **18,** 9–20 (2008).

96. Pooga, M., Soomets, U., Hallbrink, M., Valkna, A., Saar, K., Rezaei, K., Kahl, U., Hao, J., Xu, X., Wiesenfeld-hallin, Z., Hokfelt, T. & Bartfai, T. Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. *Nat. Biotechnol.* **16,** 857–861 (1998).

97. Ljungstrøm, T., Knudsen, H. & Nielsen, P. E. Cellular Uptake of Adamantyl Conjugated Peptide Nucleic Acids. *Bioconjug. Chem.* **10,** 965–972 (1999).

98. Liberman, E. A., Topaly, V. P., Tsofina, L. M., Jasaiti, A. A. & Skulachev, V. P. Mechanism of Coupling of Oxidative Phosphorylation and the Membrane Potential of Mitochondria. *Nature* **222,** 1076–1078 (1969).

99. Murphy, M. P. Selective targeting of bioactive compounds to mitochondria. *Trends Biotechnol.* **15,** 326–330 (1997).

100. Boffa, L. C., Scarfi', S., Mariani, M. R., Damonte, G., Allfrey, V. G., Benatti, U. & Morris, P. L. Dihydrotestosterone as a selective cellular/nuclear localization vector for anti-gene peptide nucleic acid in prostatic carcinoma cells. *Cancer Res.* **60,** 2258–2262 (2000).

101. Zhang, X., Simmons, C. G. & Corey, D. R. Liver cell specific targeting of peptide nucleic acid oligomers. *Bioorganic Med. Chem. Lett.* **11,** 1269–1272 (2001).

102. Basu, S. & Wickstrom, E. Synthesis and characterization of a peptide nucleic acid conjugated to a D-peptide analog of insulin-like growth factor 1 for increased cellular uptake. *Bioconjug. Chem.* **8,** 481–488 (1997).

103. Danhier, F., Ansorena, E., Silva, J. M., Coco, R., Le Breton, A. & Préat, V. PLGA-based nanoparticles: An overview of biomedical applications. *J. Control. Release* **161,** 505–522 (2012).

104. Cheng, C. J. & Saltzman, W. M. Polymer nanoparticle-mediated delivery of MicroRNA inhibition and alternative splicing. *Mol. Pharm.* **9,** 1481–1488 (2012).

105. Bahal, R., McNeer, N. A., Ly, D. H., Saltzman, W. M. & Glazer, P. M. Nanoparticle for delivery of antisense γPNA oligomers targeting CCR5. *Artif. DNA PNA XNA* **4,** 49–57 (2013).

106. McNeer, N. A., Chin, J. Y., Schleifman, E. B., Fields, R. J., Glazer, P. M. & Saltzman, W. M. Nanoparticles deliver triplex-forming PNAs for site-specific genomic recombination in CD34+ human hematopoietic progenitors. *Mol. Ther.* **19,** 172–180 (2011).

107. McNeer, N. A., Schleifman, E. B., Cuthbert, A., Brehm, M., Jackson, A., Cheng, C., Anandalingam, K., Kumar, P., Shultz, L. D., Greiner, D. L., Mark Saltzman, W. & Glazer, P. M. Systemic delivery of triplex-forming PNA and donor DNA by nanoparticles mediates site-specific genome editing of human hematopoietic cells in vivo. *Gene Ther.* **20,** 658–669 (2013).

108. Ricciardi, A. S., Bahal, R., Farrelly, J. S., Quijano, E., Bianchi, A. H., Luks, V.

L., Putman, R., López-Giráldez, F., Coşkun, S., Song, E., Liu, Y., Hsieh, W. C., Ly, D. H., Stitelman, D. H., Glazer, P. M. & Saltzman, W. M. In utero nanoparticle delivery for site-specific genome editing. *Nat. Commun.* **9,** 1–11 (2018).

109. Gupta, A., Quijano, E., Liu, Y., Bahal, R., Scanlon, S. E., Song, E., Hsieh, W. C., Braddock, D. E., Ly, D. H., Saltzman, W. M. & Glazer, P. M. Anti-tumor Activity of miniPEG-γ-Modified PNAs to Inhibit MicroRNA-210 for Cancer Therapy. *Mol. Ther. - Nucleic Acids* **9,** 111–119 (2017).

110. Grosso, S., Doyen, J., Parks, S. K., Bertero, T., Paye, A., Cardinaud, B., Gounon, P., Lacas-Gervais, S., Noël, A., Pouysségur, J., Barbry, P., Mazure, N. M. & Mari, B. MiR-210 promotes a hypoxic phenotype and increases radioresistance in human lung cancer cell lines. *Cell Death Dis.* **4,** 1–13 (2013).

111. Ho, A. S., Huang, X., Cao, H., Christman-Skieller, C., Bennewith, K., Le, Q. T. & Koong, A. C. Circulating miR-210 as a novel hypoxia marker in pancreatic cancer. *Transl. Oncol.* **3,** 109–113 (2010).

112. Qin, Q., Wei, F. & Li, B. Multiple functions of hypoxia-regulated miR-210 in cancer. *J. Exp. Clin. Cancer Res.* **33,** 1–10 (2014).

113. Sahu, B.; Sacui, I.; Rapireddy, S.; Zanotti, K. J.; Bahal, R.; Armitage, A. A.; Ly, D. H. Synthesis and characterization of conformationally peptide nucleic acids with superior hybridization properties and water solubility . *J. Org. chem.* **76,** 5614–5627 (2011).

114. Dragulescu-Andrasi, A., Rapireddy, S., Frezza, B. M., Gayathri, C., Gil, R. R. & Ly, D. H. A simple γ-backbone modification preorganizes peptide nucleic acid into a helical structure. *J. Am. Chem. Soc.* **128,** 10258–10267 (2006).

115. Beck, J. S., Vartuli, J. C., Roth, W. J., Leonowicz, M. E., Kresge, C. T., Schmitt, K. D., Chu, C. T. W., Olson, D. H., Sheppard, E. W., McCullen, S. B., Higgins, J. B. & Schlenker, J. L. A New Family of Mesoporous Molecular Sieves Prepared with Liquid Crystal Templates. *J. Am. Chem. Soc.* **114,** 10834–10843 (1992).

116. Zhao, D., Feng, J., Huo, Q., Melosh, N., Fredrickson, G. H., Chmelka, B. F. & Stucky, G. D. Triblock copolymer syntheses of mesoporous silica with periodic 50 to 300 angstrom pores. *Science (80-.).* **279,** 548–552 (1998).

117. Vallet-Regi, M., Ramila, A., del Real, R. P. & Perez-Pariente, J. A New

Property of MCM-41 : Drug Delivery System. *Chem. Mater.* **13,** 308–311 (2001).

118. Climent, E., Martínez-mµæez, R., Sancenón, F., Marcos, M. D., Soto, J., Maquieira, A. & Amorós, P. Controlled Delivery Using Oligonucleotide-Capped Mesoporous Silica. *Angew. Chemie - Int. Ed.* **49,** 7281–7283 (2010).

119. Mackowiak, S. A., Schmidt, A., Weiss, V., Argyo, C., Schirnding, C. Von, Bein, T. & Bra, C. Targeted Drug Delivery in Cancer Cells with Red-Light Photoactivated Mesoporous Silica Nanoparticles. *Nano Lett.* **13,** 2576–2583 (2013).

120. Cauda, V., Argyo, C. & Bein, T. Impact of different PEGylation patterns on the long-term bio-stability of colloidal mesoporous silica nanoparticles. *J. Mater. Chem.* **20,** 8693–8699 (2010).

121. Sauer, A. M., Schlossbauer, A., Ruthardt, N., Cauda, V. & Bein, T. Role of Endosomal Escape for Disulfide-Based Drug Delivery from Colloidal Mesoporous Silica Evaluated by Live-Cell Imaging. *Nano Lett.* **10,** 3684–3691 (2010).

122. Ma, X., Devi, G., Qu, Q., Toh, D. F. K., Chen, G. & Zhao, Y. Intracellular delivery of antisense peptide nucleic acid by fluorescent mesoporous silica nanoparticles. *Bioconjug. Chem.* **25,** 1412–1420 (2014).

123. Bertucci, A., Prasetyanto, E. A., Septiadi, D. & Manicardi, A. Combined Delivery of Temozolomide and Anti-miR221 PNA Using Mesoporous Silica Nanoparticles Induces Apoptosis in Resistant Glioma Cells. *Small* **11,** 5687–5695 (2015).

124. Bertucci, A., Lülf, H., Septiadi, D., Manicardi, A., Corradini, R. & De Cola, L. Intracellular Delivery of Peptide Nucleic Acid and Organic Molecules Using Zeolite-L Nanocrystals. *Adv. Healthc. Mater.* **3,** 1812–1817 (2014).

125. Canham, L. T. Silicon quantum wire array fabricaiton by electrochemical. *Appl. Phys. Lett.* **57,** 1046–1048 (1990).

126. Canham, L. T. Bioactive Silicon Structure Fabrication Through Nanoetching Techniques. *Adv. Mater.* **7,** 1033–1037 (1995).

127. Beavers, K. R., Mares, J. W., Swartz, C. M., Zhao, Y., Weiss, S. M. & Duvall, C. L. In situ synthesis of peptide nucleic acids in porous silicon for drug delivery and biosensing. *Bioconjug. Chem.* **25,** 1192–1197 (2014).

128. Beavers, K. R., Werfel, T. A., Shen, T., Kavanaugh, T. E., Kilchrist, K. V., Mares, J. W., Fain, J. S., Wiese, C. B., Vickers, K. C., Weiss, S. M. & Duvall, C. L. Porous Silicon and Polymer Nanocomposites for Delivery of Peptide Nucleic Acids as Anti-MicroRNA Therapies. *Adv. Mater.* **28,** 7984–7992 (2016).

129. Kelly, I. B., Fletcher, R. B., McBride, J. R., Weiss, S. M. & Duvall, C. L. Tuning Composition of Polymer and Porous Silicon Composite Nanoparticles for Early Endosome Escape of Anti-microRNA Peptide Nucleic Acids. *ACS Appl. Mater. Interfaces* **12,** 39602–39611 (2020).

130. Evans, B. C., Fletcher, R. B., Kilchrist, K. V., Dailing, E. A., Mukalel, A. J., Colazo, J. M., Oliver, M., Cheung-Flynn, J., Brophy, C. M., Tierney, J. W., Isenberg, J. S., Hankenson, K. D., Ghimire, K., Lander, C., Gersbach, C. A. & Duvall, C. L. An anionic, endosome-escaping polymer to potentiate intracellular delivery of cationic peptides, biomacromolecules, and nanoparticles. *Nat. Commun.* **10,** 5012 (2019).

131. Dragulescu-Andrasi, A., Zhou, P., He, G. & Ly, D. H. Cell-permeable GPNA with appropriate backbone stereochemistry and spacing binds sequence-specifically to RNA. *Chem. Commun.* 244–246 (2005). doi:10.1039/B412522C

132. Stender, H. PNA FISH: An intelligent stain for rapid diagnosis of infectious diseases. *Expert Rev. Mol. Diagn.* **3,** 649–655 (2003).

133. Lansdorp, P. M., Verwoerd, N. P., Van De Rijke, F. M., Dragowska, V., Little, M. T., Dirks, R. W., Raap, A. K. & Tanke, H. J. Heterogeneity in telomere length of human chromosomes. *Hum. Mol. Genet.* **5,** 685–691 (1996).

134. Reisberg, S., Dang, L. A., Nguyen, Q. A., Piro, B., Noel, V., Nielsen, P. E., Le, L. A. & Pham, M. C. Label-free DNA electrochemical sensor based on a PNA-functionalized conductive polymer. *Talanta* **76,** 206–210 (2008).

135. Jolly, P., Batistuti, M. R., Miodek, A., Zhurauski, P., Mulato, M., Lindsay, M. A. & Estrela, P. Highly sensitive dual mode electrochemical platform for microRNA detection. *Sci. Rep.* **6,** 36719 (2016).

136. Gaiji, H., Jolly, P., Ustuner, S., Goggins, S., Abderrabba, M., Frost, C. G. & Estrela, P. A Peptide Nucleic Acid (PNA)-DNA Ferrocenyl Intercalator for Electrochemical Sensing. *Electroanalysis* **29,** 917–922 (2017).

137. Mat Zaid, M. H., Abdullah, J., Yusof, N. A., Sulaiman, Y., Wasoh, H., Md Noh, M. F. & Issa, R. PNA biosensor based on reduced graphene oxide/water

soluble quantum dots for the detection of Mycobacterium tuberculosis. *Sensors Actuators B Chem.* **241,** 1024–1034 (2017).

138. Pournaghi-Azar, M. H., Ahour, F. & Hejazi, M. S. Direct detection and discrimination of double-stranded oligonucleotide corresponding to hepatitis C virus genotype 3a using an electrochemical DNA biosensor based on peptide nucleic acid and double-stranded DNA hybridization. *Anal. Bioanal. Chem.* **397,** 3581–3587 (2010).

139. Kongpeth, J., Jampasa, S., Chaumpluk, P., Chailapakul, O. & Vilaivan, T. Immobilization-free electrochemical DNA detection with anthraquinone-labeled pyrrolidinyl peptide nucleic acid probe. *Talanta* **146,** 318–325 (2016).

140. Fu, P., Xing, S., Xu, M., Zhao, Y. & Zhao, C. Peptide nucleic acid-based electrochemical biosensor for simultaneous detection of multiple microRNAs from cancer cells with catalytic hairpin assembly amplification. *Sensors Actuators B Chem.* **305,** 127545 (2020).

141. Aoki, H. & Umezawa, Y. Trace analysis of an oligonucleotide with a specific sequence using PNA-based ion-channel sensors. *Analyst* **128,** 681–685 (2003).

142. Arami, S., Alipour, E., Pournaghi-Azar, M. H. & Hejazi, M. S. Voltammetric detection of uridin diphosphate glucuronosyl transferase 1A9 (UGT1A9) gene corresponding oligonucleotide covering promoter region from −268 to −280 including (A/T) polymorphism at position −275 and optimization of the detection factors. *J. Iran. Chem. Soc.* **10,** 399–406 (2013).

143. Luo, X. & Hsing, I.-M. Immobilization-free multiplex electrochemical DNA and SNP detection. *Biosens. Bioelectron.* **25,** 803–808 (2009).

144. Svanvik, N., Westman, G., Wang, D. & Kubista, M. Light-Up Probes: Thiazole Orange-Conjugated Peptide Nucleic Acid for Detection of Target Nucleic Acid in Homogeneous Solution. *Anal. Biochem.* **281,** 26–35 (2000).

145. Isacsson, J., Cao, H., Ohlsson, L., Nordgren, S., Svanvik, N., Westman, G., Kubista, M., Sjöback, R. & Sehlstedt, U. Rapid and specific detection of PCR products using light-up probes. *Mol. Cell. Probes* **14,** 321–328 (2000).

146. Wolffs, P., Knutsson, R., Sjöback, R. & Rådström, P. PNA-Based Light-Up Probes for Real-Time Detection of Sequence-Specific PCR Products.

Biotechniques **31,** 766–771 (2001).

147. Bertucci, A., Manicardi, A., Candiani, A., Giannetti, S., Cucinotta, A., Spoto, G., Konstantaki, M., Pissadakis, S., Selleri, S. & Corradini, R. Detection of unamplified genomic DNA by a PNA-based microstructured optical fiber (MOF) Bragg-grating optofluidic system. *Biosens. Bioelectron.* **63,** 248–254 (2015).

148. Candiani, A., Sozzi, M., Cucinotta, A., Selleri, S., Veneziano, R., Corradini, R., Marchelli, R., Childs, P. & Pissadakis, S. Optical Fiber Ring Cavity Sensor for Label-Free DNA Detection. *IEEE J. Sel. Top. Quantum Electron.* **18,** 1176–1183 (2012).

149. Candiani, A., Bertucci, A., Giannetti, S., Konstantaki, M., Manicardi, A., Pissadakis, S., Cucinotta, A., Corradini, R. & Selleri, S. Label-free DNA biosensor based on a peptide nucleic acid-functionalized microstructured optical fiber-Bragg grating. *J. Biomed. Opt.* **18,** 057004 (2013).

150. Sawata, S., Kai, E., Ikebukuro, K., Iida, T., Honda, T. & Karube, I. Application of peptide nucleic acid to the direct detection of deoxyribonucleic acid amplified by polymerase chain reaction. *Biosens. Bioelectron.* **14,** 397–404 (1999).

151. Sato, Y., Fujimoto, K. & Kawaguchi, H. Detection of a K-ras point mutation employing peptide nucleic acid at the surface of a SPR biosensor. *Colloids Surfaces B Biointerfaces* **27,** 23–31 (2003).

152. Corradini, R., Feriotto, G., Sforza, S., Marchelli, R. & Gambari, R. Enhanced recognition of cystic fibrosis W1282X DNA point mutation by chiral peptide nucleic acid probes by a surface plasmon resonance biosensor. *J. Mol. Recognit.* **17,** 76–84 (2004).

153. Liebermann, T. & Knoll, W. Surface-plasmon field-enhanced fluorescence spectroscopy. *Colloids Surfaces A Physicochem. Eng. Asp.* **171,** 115–130 (2000).

154. Kambhampati, D., Nielsen, P. E. & Knoll, W. Investigating the kinetics of DNA-DNA and PNA-DNA interactions using surface plasmon resonance-enhanced fluorescence spectroscopy. *Biosens. Bioelectron.* **16,** 1109–1118 (2001).

155. Liu, J., Tiefenauer, L., Tian, S., Nielsen, P. E. & Knoll, W. PNA–DNA Hybridization Study Using Labeled Streptavidin by Voltammetry and

Surface Plasmon Fluorescence Spectroscopy. *Anal. Chem.* **78,** 470–476 (2006).

156. Liu, J., Tian, S., Tiefenauer, L., Nielsen, P. E. & Knoll, W. Simultaneously Amplified Electrochemical and Surface Plasmon Optical Detection of DNA Hybridization Based on Ferrocene–Streptavidin Conjugates. *Anal. Chem.* **77,** 2756–2761 (2005).

157. Chu, L.-Q., Förch, R. & Knoll, W. Surface-Plasmon-Enhanced Fluorescence Spectroscopy for DNA Detection Using Fluorescently Labeled PNA as "DNA Indicator". *Angew. Chemie Int. Ed.* **46,** 4944–4947 (2007).

158. D'Agata, R., Corradini, R., Grasso, G., Marchelli, R. & Giuseppe, S. Ultrasensitive Detection of DNA by PNA and Nanoparticle-Enhanced Surface Plasmon Resonance Imaging. *ChemBioChem* **9,** 2067–2070 (2008).

159. D'Agata, R., Corradini, R., Ferretti, C., Zanoli, L., Gatti, M., Marchelli, R. & Spoto, G. Ultrasensitive detection of non-amplified genomic DNA by nanoparticle-enhanced surface plasmon resonance imaging. *Biosens. Bioelectron.* **25,** 2095–2100 (2010).

160. D'Agata, R., Bellassai, N., Allegretti, M., Rozzi, A., Korom, S., Manicardi, A., Melucci, E., Pescarmona, E., Corradini, R., Giacomini, P. & Spoto, G. Direct plasmonic detection of circulating RAS mutated DNA in colorectal cancer patients. *Biosens. Bioelectron.* **170,** 112648 (2020).

161. Tian, M., Qiao, M., Shen, C., Meng, F., Frank, L. A., Krasitskaya, V. V, Wang, T., Zhang, X., Song, R., Li, Y., Liu, J., Xu, S. & Wang, J. Highly-sensitive graphene field effect transistor biosensor using PNA and DNA probes for RNA detection. *Appl. Surf. Sci.* **527,** 146839 (2020).

162. Sabale, P. M., George, J. T. & Srivatsan, S. G. A base-modified PNA–graphene oxide platform as a turn-on fluorescence sensor for the detection of human telomeric repeats. *Nanoscale* **6,** 10460–10469 (2014).

163. Ryoo, S.-R., Lee, J., Yeo, J., Na, H.-K., Kim, Y.-K., Jang, H., Lee, J. H., Han, S. W., Lee, Y., Kim, V. N. & Min, D.-H. Quantitative and Multiplexed MicroRNA Sensing in Living Cells Based on Peptide Nucleic Acid and Nano Graphene Oxide (PANGO). *ACS Nano* **7,** 5882–5891 (2013).

164. Lee, J., Park, I.-S., Jung, E., Lee, Y. & Min, D.-H. Direct, sequence-specific detection of dsDNA based on peptide nucleic acid and graphene oxide

without requiring denaturation. *Biosens. Bioelectron.* **62,** 140–144 (2014).

165. Guo, S., Du, D., Tang, L., Ning, Y., Yao, Q. & Zhang, G.-J. PNA-assembled graphene oxide for sensitive and selective detection of DNA. *Analyst* **138,** 3216–3220 (2013).

166. Pawbake, A. S., Waykar, R. G., Late, D. J. & Jadkar, S. R. Highly Transparent Wafer-Scale Synthesis of Crystalline WS2 Nanoparticle Thin Film for Photodetector and Humidity-Sensing Applications. *ACS Appl. Mater. Interfaces* **8,** 3359–3365 (2016).

167. Wang, S., Zhang, Y., Ning, Y. & Zhang, G.-J. A WS2 nanosheet-based platform for fluorescent DNA detection via PNA–DNA hybridization. *Analyst* **140,** 434–439 (2015).

168. Liao, X., Wang, Q. & Ju, H. A peptide (1) Liao, X.; Wang, Q.; Ju, H. A Peptide Nucleic Acid-Functionalized Carbon Nitride Nanosheet as a Probe for in Situ Monitoring of Intracellular MicroRNA. Analyst 2015, 140 (12), 4245–4252. https://doi.org/10.1039/C5AN00128E.nucleic acid-functio. *Analyst* **140,** 4245–4252 (2015).

169. Wu, Y., Han, J., Xue, P., Xu, R. & Kang, Y. Nano metal–organic framework (NMOF)-based strategies for multiplexed microRNA detection in solution and living cancer cells. *Nanoscale* **7,** 1753–1759 (2015).

170. Yang, L.-H., Ahn, D. J. & Koo, E. Ultrasensitive FRET-based DNA sensor using PNA/DNA hybridization. *Mater. Sci. Eng. C* **69,** 625–630 (2016).

171. Tadimety, A., Zhang, Y., Kready, K. M., Palinski, T. J., Tsongalis, G. J. & Zhang, J. X. J. Design of peptide nucleic acid probes on plasmonic gold nanorods for detection of circulating tumor DNA point mutations. *Biosens. Bioelectron.* **130,** 236–244 (2019).

172. Fortunati, S., Rozzi, A., Curti, F., Giannetto, M., Corradini, R. & Careri, M. Novel amperometric genosensor based on peptide nucleic acid (PNA) probes immobilized on carbon nanotubes-screen printed electrodes for the determination of trace levels of non-amplified DNA in genetically modified (GM) soy. *Biosens. Bioelectron.* **129,** 7–14 (2019).

173. Hwang, D. W., Kim, H. Y., Li, F., Park, J. Y., Kim, D., Park, J. H., Han, H. S., Byun, J. W., Lee, Y. S., Jeong, J. M., Char, K. & Lee, D. S. In vivo visualization of endogenous miR-21 using hyaluronic acid-coated graphene oxide for targeted cancer therapy. *Biomaterials* **121,** 144–154 (2017).

Chapter 2

2 Loading of PNA into Porous Silicon Nanoparticles

2.1 Introduction

As reported in Chapter 1, a great interest has arisen in the development of novel oligonucleotide-based functional hybrid nanomaterials for programmable drug delivery and diagnostic applications over the past years. These materials possess unique properties derived from the synergistic combination of the physicochemical properties of both organic and inorganic components, additionally there is the possibility to systematically tailor the properties of the hybrid nanomaterial to design systems with enhanced and tailored performances.

In 2017, seven research groups based in six different countries (Italy, Spain, France, US, Canada and Cuba) have joined an international collaborative project (MSCA-RISE-778133-Nano-OligoMed) funded by the European Commission through the Marie Skłodowska-Curie actions, whose aim is the generation of efficient and safe hybrid structured oligonucleotide vehicles by conjugation with nanostructure materials, as silica, silicon and quatsomes for biomedical applications.

Porous silicon (pSi) is a versatile material largely and successfully applied in drug delivery and diagnostic. In the mid 1950s, this material was accidentally discovered at Bell Laboratories during electrochemical experiments on silicon wafers, nevertheless it did not generate much enthusiasm. Indeed, it was only 30 years later that the interest in the properties of porous silicon exploded. As already mentioned in Chapter 1, the most common procedure to fabricate porous silicon is by a top-down approach in which a single crystalline silicon wafer is electrochemically etched in an aqueous hydrofluoric acid solution. In the two-electrode cell, where the electrochemical reaction occurs, silicon represents the working electrode (anode) at which surface an oxidation reaction takes place, the counter-electrode (cathode) is instead typically platinum and it is possible observe the release of hydrogen gas caused by the water electrolysis reaction (**Figure 2.1a**). The current-potential curve for electrochemical etch of silicon in an HF electrolyte shows three different regions defining the oxidation process of silicon (**Figure 2.1b**). Electropolishing regime, controlled by the 4-electron half reaction defined by Equation 2.1,

occurs when a high current density is applied at the more positive electrode potential and no porous silicon formation is observed.

$$Si + 6\ F^- + 4h^+ \rightarrow SiF_6^{2-} \quad (2.1)$$

Porous silicon formation occurs instead at lower applied potential and the two-electron reaction shown in Equation 2.2 is the predominant process.

$$Si + 6\ F^- + 2\ H^+ + 2\ h^+ \rightarrow SiF_6^{2-} + H_2 \quad (2.2)$$

Valence band holes (depicted as the lower case h^+ in the Equations above) represent the oxidizing equivalents that initiate the process, which is generally ruled by the rate of delivery of valence band holes to the surface (current density) and by the diffusion of the F^- species. Indeed, valence band holes make Si surface atoms susceptible to nucleophilic attack, so when the HF concentration in solution is too low and hence current density exceeds the rate of HF diffusion, water molecules attack oxidized silicon atoms in place of F^- generating an insulating oxide that terminates the propagation of the pore (electropolishing). It follows that porous silicon formation occurs when sufficiently high HF concentration is present in the electrolyte.

The electrochemical corrosion of silicon allows to obtain a wide range of pore diameters by variation of the process parameters, such as applied voltage, electrolyte composition, dopant type, etc. When the etching conditions (current density, electrolyte, and etch duration) are held constant, the morphology of pSi is affected by the types of silicon: the p-type is typically boron-doped and the n-type is phosphorous-doped silicon. Specifically, p-type silicon tends to generate meso- to micropores, highly doped p+, p++ or n+ wafers yield mesopores, whereas n-type material gives macropores.

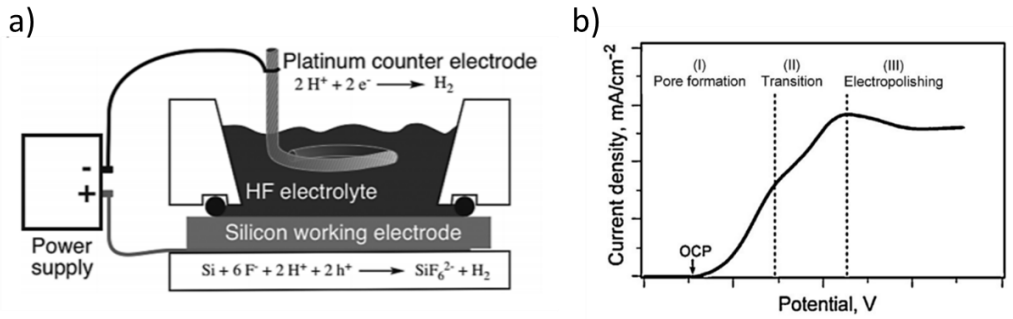


Figure 2.1. a) Schematic of a two-electrode electrochemical cell used to fabricate porous silicon; b) General current density versus applied potential curve. Reprinted with permission from M. J. Sailor Copyright © 2011, John Wiley and Sons (a); Reprinted with permission from F. A. Harraz Copyright © 2014 Elsevier B.V. All rights reserved (b).

The nanometer-scale structure of porous silicon electrochemically generated is strictly correlated to photoluminescence properties; these features have led to a massive production of silicon-based optoelectronic switches, lasers and displays in the early 1990s. Indeed, the porous silicon material emits a bright red-orange photoluminescence related to a quantum confinement effect that arises when the pores evolve into a nanometer-scale silicon filament. The photoluminescence of pSi is dependent on fabrication conditions, especially etching current density (**Figure 2.2**), and it can be originated not only by the quantum confinement effect but also by surface species or defects. A lift-off process of the etched silicon layer produces a dispersible material which can be then converted into nanoparticles under ultrasound treatment.

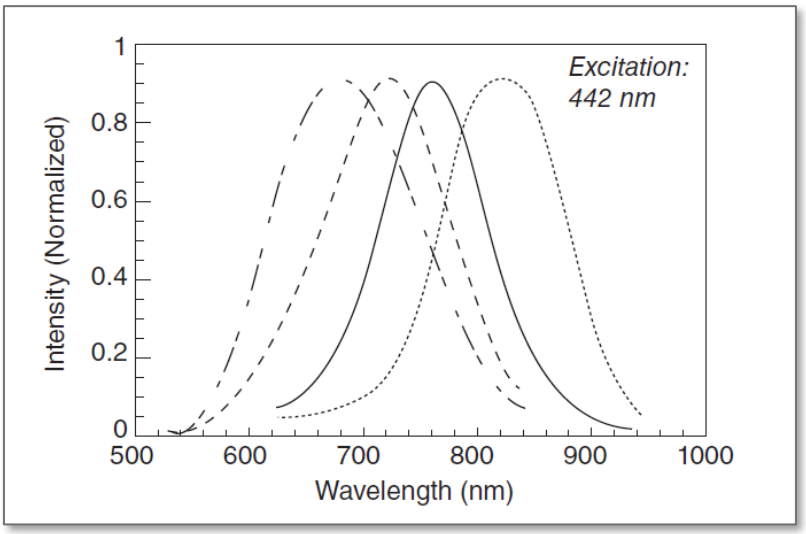


Figure 2. 2. Photoluminescence spectra of four n-type porous silicon samples, prepared by etching at different current density. Reprinted with permission from M. J. Sailor Copyright © 2011, John Wiley and Sons.

The intrinsic photoluminescence, together with its biocompatibility and biodegradability, have made porous silicon nanoparticles attractive tools for drug delivery and imaging, in contrast, for example, with quantum dots, gold nanoparticles or carbon nanotubes that have shown to be potent tools for *in vivo* imaging but that are not suitable for a clinical translation because of their potential toxic side effects. On the other hand, the biodegradation product of pSi is the orthosilicic acid $Si(OH_4)$ that is generally adsorbed by humans and, if administered, can be excreted from the body in the urine.

2.1.1 Cystic Fibrosis

Cystic Fibrosis (CF) is an autosomal recessive genetic disease deriving from misfunctioning of cystic fibrosis transmembrane conductance regulator (CFTR) protein, due to genetic defects in its coding DNA sequence. Among the roughly 1000 cystic fibrosis-causing mutations, the most common results being ΔF508, corresponding to a loss of phenylalanine (F) on chromosome 7. Due to amino acid deletions or substitution, the defective protein is not able to be folded and processed in a proper way, and is hence not properly translocated to the outer layer of cellular membrane of epithelial cells, thus leading to a defective transport of chloride ions (**Figure 2.3**).

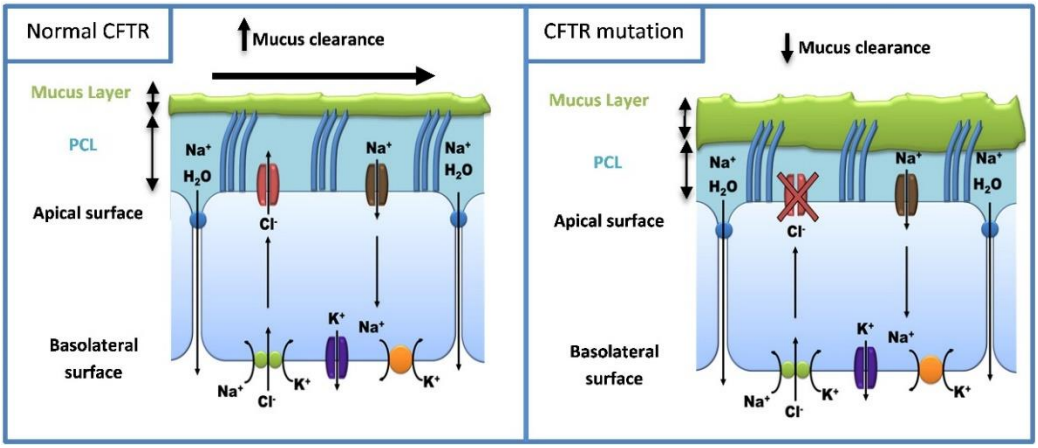


Figure 2.3. Differences in the mechanism of a functional and a mutated cystic fibrosis transmembrane conductance regulator (CFTR) protein. Reprinted with permission from Kumar et al. Copyright © 2014 European Federation of Internal Medicine. Published by Elsevier B.V. All rights reserved.

This leads to dramatic changes in the fluid regulation of epithelial cells in various organs, most evidently in the lungs,[1] with problems for respiration and increased infection risk. Recently, correctors of the CFTR function have been approved by the FDA (e.g. VX-809, known under the commercial name of Lumacaftor)[2] but these are very expensive and suitable alternatives would be very interesting both for studies on CFTR properties and as potential candidates for treatments.

In this respect, several microRNAs have been described as downregulators of CFTR expression,[3] hence a strategy that involves the use of anti-miR agents for targeting microRNAs could represent an effective choice. Numerous studies have already demonstrated the efficacy of peptide nucleic acids in modulating the level of microRNAs involved in the CFTR gene regulation.[4,5] Particularly interesting was the use of PNAs as protector of microRNA targets as alternative to the largely used anti-miRNA strategy. Indeed, the latter could cause unknown side effects by interfering with other targets of the miRNA addressed. This possibility was evaluated by the Prof. Oliviero group that, firstly, confirmed the formation of a stable heteroduplex between a negatively charged PNA and its complementary sequence situated in the 3'UTR region of CFTR mRNA by spectroscopic techniques and molecular dynamic simulations, and finally, tested the efficacy of this strategy on A549 human lung carcinoma cells.[6] A similar approach was reported by Sultan *et al.*, in which a specific PNA sequence was not designed to directly bind the complementary microRNA, namely miR-145-5p, but to mask its binding site within the 3'UTR region of the target

mRNAs (**Figure 2.4**).[7] Through this strategy, the researchers observed an accumulation of CFTR mRNA that led to a CFTR protein increase of approximately 7-fold for the cells treated with miR145-masking PNA.

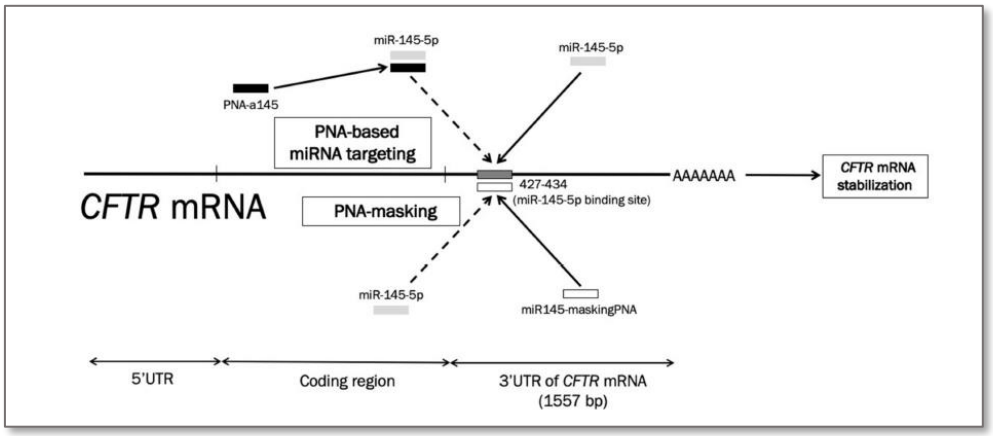


Figure 2.4. Comparison of the peptide nucleic acid (PNA)-based miRNA-targeting (upper part of the panel) and the PNA-masking (lower part of the panel) strategies to inhibit miR-145-5p biological functions. Dark grey box: the miR-145-5p binding site; light grey boxes: miR-145-5p; white box: miR145-maskingPNA; black boxes: the anti-miR-145-5p PNA-a145. Dotted arrows: inhibition/interference. Reprinted with permission from ref. [7] Sultan et al. © 2020 Creative Commons Attribution 4.0 International License.

One of the most important challenge in PNA technology is the development of suitable strategies for delivering these synthetic oligonucleotides into the cells. Indeed, poor cellular internalization of PNAs, mainly due to their neutral-charged backbone, represents an impediment to their use in the biomedical field as both therapeutic agents and diagnostic tools.

Conjugation of PNA with polyarginine pendants is a largely investigated approach to this aim. For example, an octaarginine-antimiR PNA has been delivered to Calu-3 cells, where was able to enhance CFTR gene expression by targeting its complementary miR-145-5p.[4] However, these carriers are not optimal since toxic effect could arise when used at high concentration, which might be the case for *in vivo* tests.

It is worth noting that the simple overexpression of CFTR protein could not be beneficial in all cases. In fact, in individuals expressing mutated and malfunctioning protein overexpression turns to be detrimental. Thus, it is interesting to induce biological processes that might lead to correction of the CFTR function.

A strategy evaluated to address CF disease with PNA-based technology includes the encapsulation of PNAs into nanoparticles systems. In a work of McNeer and

co-workers, an appropriate tail-clamp PNA and a donor DNA have been loaded into polymer nanoparticles (PLGA) in order to correct ΔF508 mutation both *in vitro* and *in vivo*.[8] The use of blended particles, i.e. PLGA/PBAE (poly-(beta-amino ester) NPs decorated with the MPG peptide, restored a normal chloride efflux in 25% of human bronchial epithelial (HBE) cells, furthermore, the evaluation of the *in vivo* correction of ΔF508 in murine model showed an enhanced gene modification in the lungs with low off-target effects.

Another approach, more recently proposed, is that of inducing the overproduction of 'corrector' scaffolding proteins, that are primarily localized at the apical surfaces of epithelial cells and which are able to assist the correct protein folding of defective CFTR, thus restoring its chloride transport efficiency.[9] This approach was shown to be effective when using anti-miR 335 PNA modified with polyarginine (R8-PNA-a335-5p), which was shown to enhance the sodium-hydrogen exchanger regulatory factor 1 (NHERF1) scaffolding protein level in bronchial epithelial Calu-3 cells. The same PNA was shown to be able to restore halogenide transport in defective cells bearing the F508del mutation, with effects similar to those obtained with the commercial drug VX 809 Lumacaftor (**Figure 2.5**).

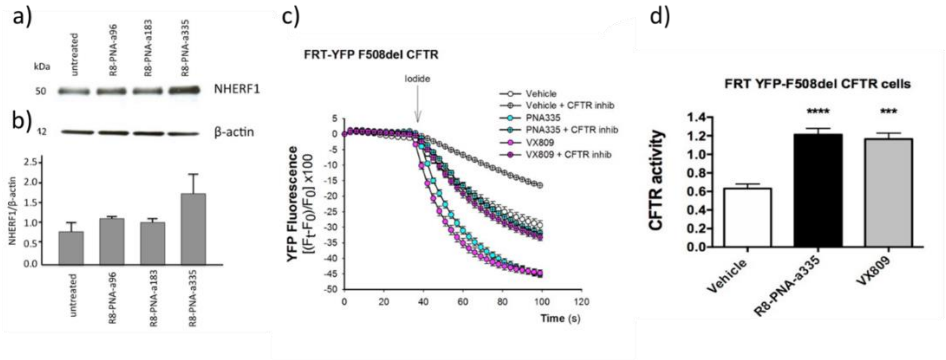


Figure 2.5. (a)(b) Changes in NHERF1 gene expression analyzed by Western blotting in samples from Calu-3 cells cultured for 72 h in the absence or in the presence of R8-PNA-a335-5p, or of PNAs directed to different targets (miR-96 and miR-183) as indicated; (c) (d) Effects of R8-PNA-a335 on CFTR halide ion transport in FRT-YFP F508del CFTR cells. Cells were cultured untreated, or were treated with R8-PNA-a335 (4 μM) or VX 809 (5 μM) for 48 h. CFTR function was assessed by single-cell fluorescence imaging in the presence of the protein kinase A activator forskolin (20 μM) and genistein (50 μM). Representative traces are shown in panel (c). The summary of three independent experiments is shown in panel (d). Results are presented as transformed data to obtain the signal variation (Fx) relative to the time of addition of iodide, according to the equation: Fx ([Ft Fo]/Fo) 100, where Ft and Fo are the fluorescence values at the time t and at the time of addition of the iodide, respectively. Adapted with permission from ref. [9] © 2021 by the authors. Licensee MDPI, Basel, Switzerland.

2.2 *Aim of the study*

In the last years, an emerged and promising strategy for the delivery of PNA turned out to be the encapsulation of the oligonucleotide in nanoparticle carriers. In this study, a calcium-silicate trapping protocol was employed for fast loading of a small library of differently charged anti-miR PNAs in pSiNPs (**Figure 2.6**). The release properties of the different calcium-silicate capped pSiNPs-PNA formulations were investigated in aqueous buffer. Eventually, as proof-of-concept of the potential of these formulations in drug delivery applications, we performed preliminary experiments to verify cellular internalization of PNA-loaded pSiNPs and their ability to inhibit the expression of cognate microRNA targets.

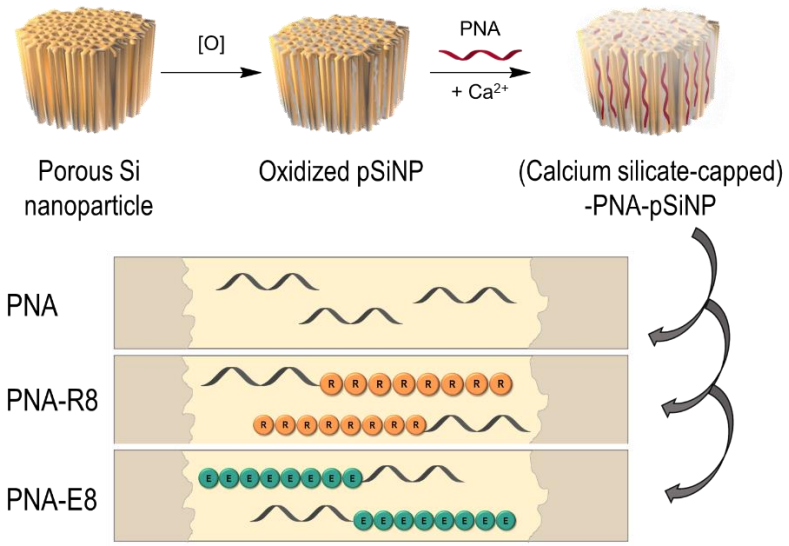


Figure 2.6. Schematic representation of the calcium silicate trapping protocol used to produce PNA-loaded pSiNPs.

As model system, PNAs targeting miRs involved in the regulation of the expression of proteins involved in the cystic fibrosis cellular regulation (miR 145, miR 101 and miR-335) were chosen to evaluate the translational potential of this technology.

2.3 Results and discussion

2.3.1 Synthesis of anti-miR PNAs

Previous studies carried out on specific microRNAs involved in the CF disease, in collaboration with Prof. Gambari at University of Ferrara, showed that using PNA targeting microRNAs is a potential strategy to enhance CFTR production and activity. Based on these findings, three different antimiR PNA sequences were synthesized, namely antimiR-335, antimiR-145 and antimiR-101. In order to test whether differently charged PNA could induce a different effect on CFTR expression, each antimiR PNA was conjugated at the N-terminus to a polyarginine pendant, a glutamic acid chain and neutral rhodamine. A list of PNAs produced is shown in **Table 2.1**.

Table 2. 1. List of the anti-miR PNA sequences synthesized in this study.

Sequence name	Sequence
AntimiR-335 Rho	Rho-TTT CGT TAT TGC TCT TGA-Gly-NH_2
AntimiR-335 E_8	H-E_8-TTT CGT TAT TGC TCT TGA-Gly-NH_2
AntimiR-335 R_8	H-R_8-TTT CGT TAT TGC TCT TGA-Gly-NH_2
AntimiR-145 Rho	Rho-AGG GAT TCC TGG GAA AAC-Gly-NH_2
AntimiR-145 E_8	H-E_8-AGG GAT TCC TGG GAA AAC-Gly-NH_2
AntimiR-145 R_8	H-R_8-AGG GAT TCC TGG GAA AAC-Gly-NH_2
AntimiR-101 Rho	Rho-AGT TAT CAC AGT ACT GTA-Gly-NH_2
AntimiR-101 E_8	H-E_8-AGT TAT CAC AGT ACT GTA-Gly-NH_2
AntimiR-101 R_8	H-R_8-AGT TAT CAC AGT ACT GTA -Gly-NH_2

The mentioned PNAs were synthesized following the standard Fmoc-based solid-phase synthesis (SPS) protocol and after their cleavage from the resin, they were purified by RP-HPLC and characterized by UPLC/ESI-MS and UV-Vis techniques. Samples of these antimiR-PNAs were used during secondment at University of California San Diego (UCSD) in Prof. Michael J. Sailor's laboratory, where I spent 5 months of the doctorate program as visiting scholar in the frame of the NanoOligomed Project, in order to test their loading into porous silicon nanoparticles.

2.3.2 *Preparation of porous silicon nanoparticles (pSiNPs)*

Porous silicon nanoparticles were produced by a well-established protocol developed in the research group of Prof. Michael Sailor at UCSD. The protocol involves the electrochemical etching of a single-crystal silicon wafer in an electrolyte HF acid solution by applying an appropriate current density. A highly doped p++ silicon wafer was chosen to our scope since it typically yields pores in the range between 2 and 50 nm (mesopores) suitable for the incorporation of biomolecules. As mentioned above, the valence band holes represent the oxidizing equivalents at the beginning of the process, so the choice of the type-dopant in the silicon wafer is crucial also because it defines the availability of these holes. Once the porous layer is formed, it can be removed from the Si substrate by applying a current pulse for several seconds and by replacing the electrolyte solution with one at a lower HF concentration (electropolishing or 'lift off'). A vigorous rinse or tweezers can be used to remove the free-standing porous Si film that is subsequently subjected to an ultrasonication treatment for 18 hours in order to obtain the desired nanoparticles (**Figure 2.7**).

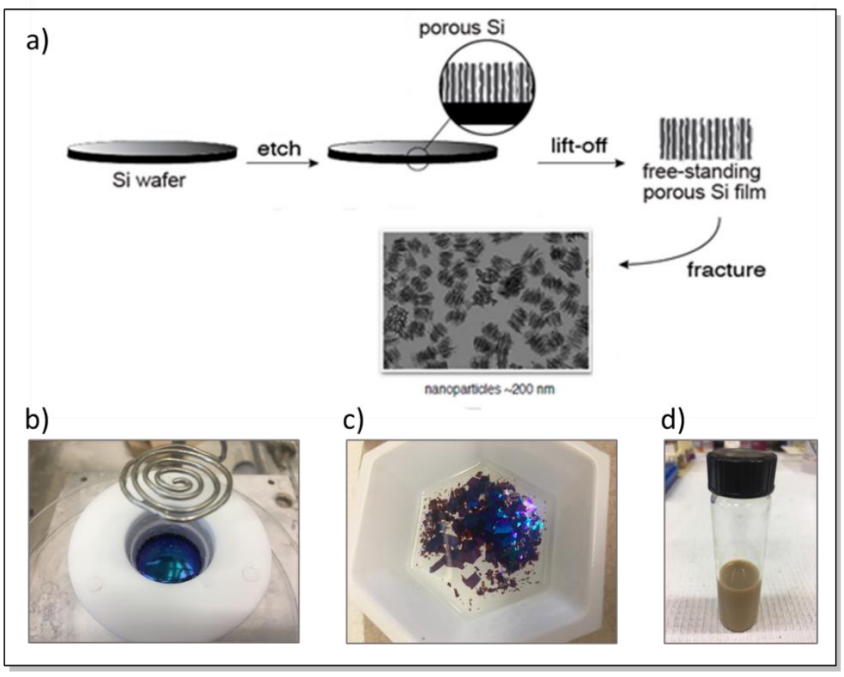


Figure 2.7. Electrochemical etching of pSiNPs. a) Schematic representation of the process; b) closer view of Teflon cell after applying the etching current; c) free-standing porous silicon film; d) pSiNPs after overnight sonication.

The surface of freshly etched porous silicon is quite susceptible to air oxidation, as already reported by K.H. Beckmann in 1965,[10] leading to a continuous change in optical and structural properties over time. In order to stabilize PSi, a thin oxide layer is grown on the surface by a mild oxidation by air in an aqueous sodium tetraborate solution (0.8 mM) at room temperature.
The hydrodynamic diameter of the nanoparticles was measured to be 215 ± 8 nm by dynamic light scattering (DLS) and was confirmed by TEM measurements. Spectroscopic liquid infiltration method (SLIM)[11] was used to determine the porous layer porosity, that showed to be 58% ± 4. Furthermore, data collected by cryogenic nitrogen adsorption isotherm allowed to determine the average pore size (14.5 nm) as depicted in **Figure 2.8**.

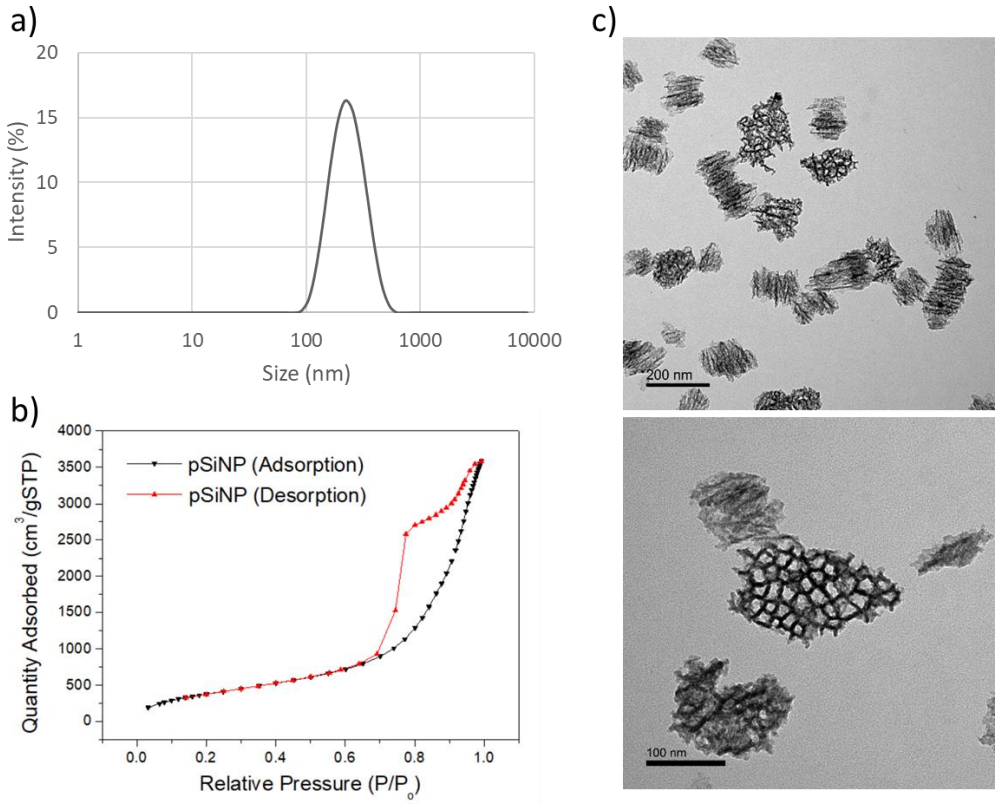


Figure 2.8. Characterization of pSiNPs. a) Intensity weighted size distribution from dynamic light scattering of freshly etched pSiNPs; b) Cryogenic nitrogen adsorption–desorption isotherm of the empty, unmodified pSiNPs; c) Transmission electron microscope (TEM) image of freshly etched pSiNPs (scale bar = 200 nm) (top); close-up of a single particle (scale bar = 100 nm) (bottom).

It worth mentioning that the oxides on the thin shallow layer of porous Si can be also exploited to functionalize the surface with different chemical species using traditional silane chemistry.

Three types of nanoparticles were prepared in order to investigate the loading achievable by electrostatic adsorption with the differently charged antimiR PNAs previously synthesized (see **Table 2.1**). The first type of NPs were prepared following the novel 'ring-opening click' reaction proposed by Prof. Sailor and coworkers that allows to obtain positively-charged nanoparticles.[12] The protocol was developed as alternative to the most common hydrolytic condensation of organotrialkoxy-silanes on the surface, since this type of functionalization can necessitate high temperatures or long reaction times, the onset of byproducts can interfere with the process and the coverage can be hampered by cross-linking reactions. Instead, this protocol showed to proceed to completion within 1-2 hr with no significant side products and above all, the coverage occurs preserving the open pore structure of the silicon nanomaterial. The reaction involves the modification of hydroxyls on the surface with a diaza-silane, namely 2,2-dimethoxy-1,6-diaza-2-silacyclooctane, which leads to the generation of primary amine moieties (**Figure 2.9**).

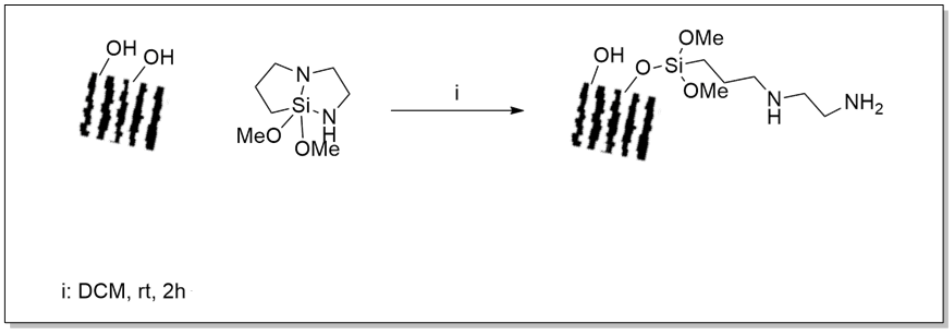


Figure 2.9. Schematic of the synthesis of amino-modified pSiNPs.

The reaction was confirmed by infrared spectroscopy and, moreover, an increase in zeta potential, between +10 mV and +15 mV, was observed for this type of nanoparticles, indicating a net positive surface charge.

Amine-modified pSiNPs can be exploited to prepare porous silicon nanoparticles with a terminal carboxylic acid functionality when the reaction mixture contains succinic anhydride (**Figure 2.10**). As in the first case, data collected using Fourier transform infrared (FTIR) technique pointed out the generation of the expected amide group (1400-1490 cm^{-1} stretching C-N,

bending N-H, 1600 cm^{-1} stretching C=O) and the carboxylic acid moiety (1750 cm^{-1} stretching C=O. See **Figure 2.12**). The nanoparticles exhibited a net negative surface charge, with a zeta potential equal to -27 mV.

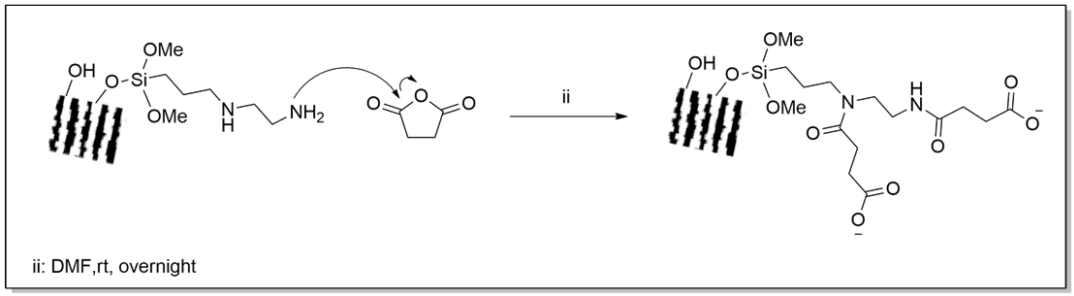


Figure 2.10. Schematic of the synthesis of carboxylate-modified pSiNPs.

In a previous work, pSiNPs modified with a phosphonate group yielded the highest loading efficiency of a peptide payload compared to other modifications.[13] Following the protocol reported in the cited work, phosphonate-modified pSiNPs were prepared in order to be tested in the planned loading experiments (**Figure 2.11**). FTIR spectrum of phosphonate NPs revealed a band at 2900 cm^{-1} corresponding to the stretching of aliphatic C-H and the typical stretching band at 1300 cm^{-1} relative to the P=O bond. Zeta potential measurements of this type of nanoparticles showed a negative surface potential equal to -16 mV.

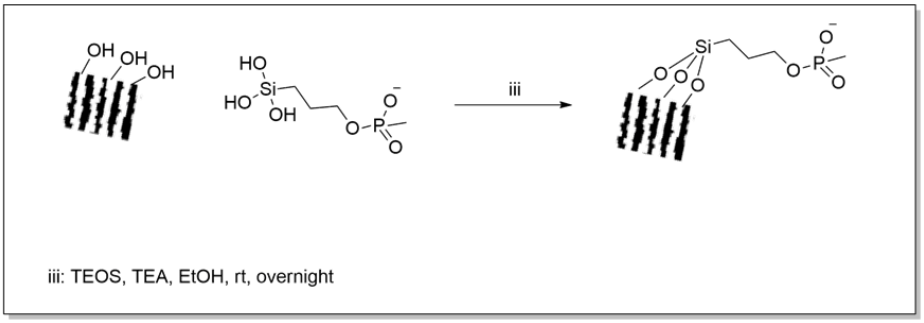


Figure 2.11. Schematic of the synthesis of phosphonate-modified pSiNPs.

The figure below (**Figure 2.12**) summarizes the FTIR and zeta potential data collected for the all three type of modified porous silicon nanoparticles.

a)

	FTIR analysis
pSiNP	1900, 800 cm^{-1}
Amino-pSiNP	3300, 2900 cm^{-1}
Phosp-pSiNP	2900, 1650, 1300 cm^{-1}
Carboxy-pSiNP	2950, 1750, 1600, 1490 cm^{-1}

b)

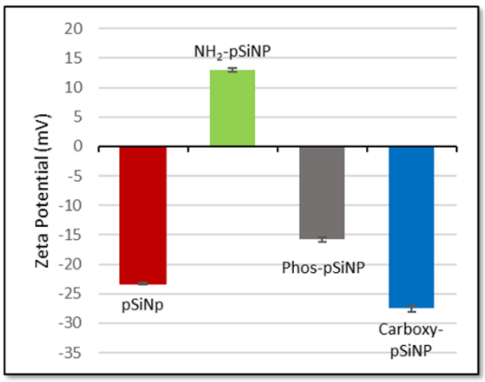


Figure 2.12. Characterization of modified pSiNPs. a) Diagnostic bands observed in Fourier transform infrared (FTIR) spectroscopy of amine, carboxylate, and phosphonate modified pSiNPs; b) ζ-potential measurements of oxidized pSiNPs and modified pSiNPs.

2.3.3 Loading of antimiR-PNAs into pSiNPs

According to the literature, only a few examples of PNAs loaded into pSiNPs are reported. In a first work, PNA was synthesized *in situ* from a porous silicon substrate,[14] whereas in a following one, a non-covalent physical adsorption approach was described,[15] in which PNA-loaded nanoparticles were coated with the copolymer PEGDB developed to facilitate the endosomal escape.
With the aim to improve the loading of PNA into pSiNPs for antimiR therapy in the context of CF disease, different loading experiments were planned.
Carboxylate-modified NPs were chosen to test the electrostatic adsorption of the positively charged antimiR-335 R_8. The nanoparticles were incubated in phosphate buffered saline (PBS, pH 7.4) with PNA for 3 hr at room temperature and then, the supernatant was recovered by centrifugation in order to assess the loading efficiency by UV-Vis measurements thereof. To this aim, a calibration curve was constructed (**Figure 2.13**), and concentration of the PNA in the supernatant was calculated using this curve. The loading value achieved by this protocol, calculated as the mass of the PNA divided by the total mass of the PNA + Si nanoparticles, was estimated to be between 0-1.9% w/w, with an encapsulation efficiency (PNA nmol encapsulated/ PNA nmol incubated) up to

7.7%. These results showed that simple adsorption, even if charge-assisted, is not sufficient to achieve the high loading values needed for therapeutic applications. For this reason a different loading approach was attempted.

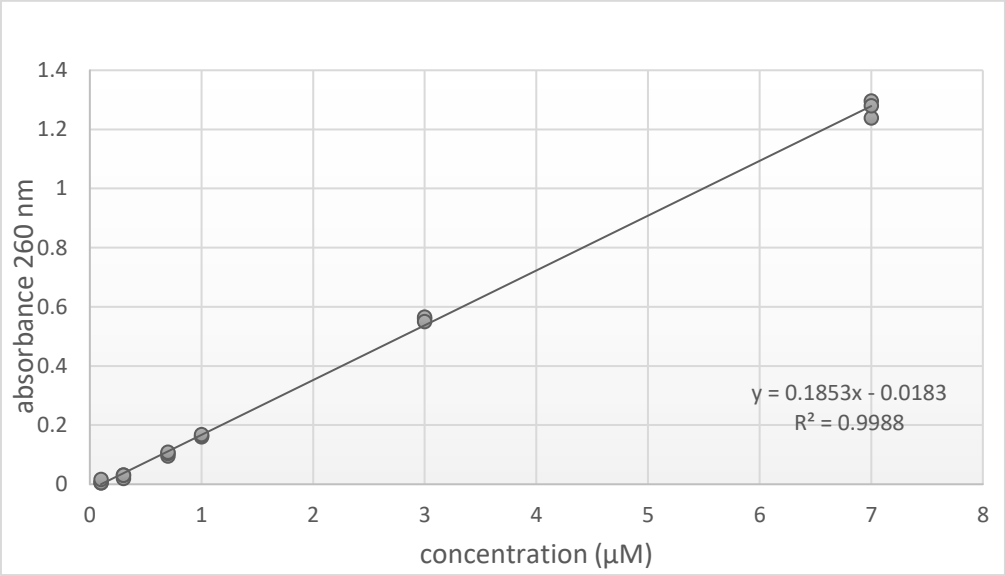


Figure 2.13. Calibration curve of PNA antimiR-335 R8.

In the group of Prof. Sailor, a novel protocol for loading oligonucleotides into pSiNPs was developed,[16] which includes the embedding of high concentrations of siRNA simultaneously to a precipitation of an insoluble shell of calcium silicate Ca_2SiO_4 (**Figure 2.14**). It was demonstrated that the formation of the shell occurs at the nanoparticles surface as a result of partial dissolution of the oxidized porous silicon matrix ($Si(OH)_4$) in the presence of a solution containing high concentrations of calcium (II) ion, and furthermore, that this calcium silicate shell is able to slow the degradation of the nanoparticles and to protect the payload from an early release. The use of calcium coating is also interesting for systemic administration, since the calcium silicate layer is stable under high calcium concentrations (such as in serum or extracellular matrices, where it is in the millimolar range), and can be more rapidly dissolved in the intracellular medium in which the calcium concentration is much lower (in the range of 100nM, though is subject to variations due to intracellular processes and stimuli).

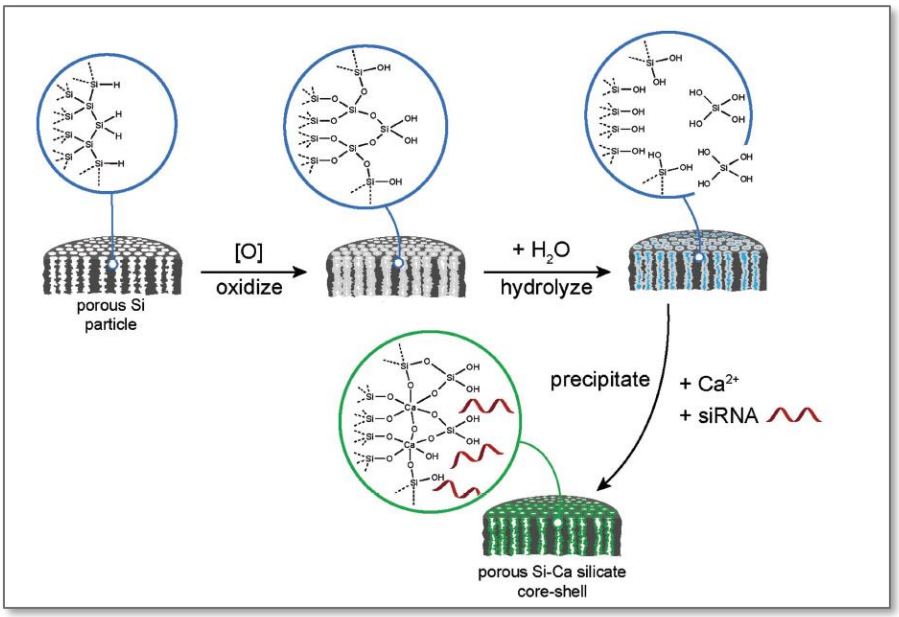


Figure 2.14. Schematic illustration of the mechanistic steps involved in preparation of siRNA-loaded, calcium silicate-coated porous silicon nanoparticles (Ca-pSiNP-siRNA). Reprinted with permission from ref. [12] © 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

The first loading test was performed using the antimiR-101 E_8, which with its glutamate tail could have mimic the behavior of the negatively charged siRNA. PNA was sealed into the porous silicon nanostructure by stirring the nanoparticles in an aqueous solution containing the PNA and a 4M concentration of $CaCl_2$. After 45' the supernatant was recovered by centrifugation and the isolated pSiNPs were washed three times. The UV analysis of the supernatant was performed using a new calibration curve registered in the same medium as the final supernatant solution. The results obtained showed that this procedure allowed to achieve an encapsulation efficiency of ≈99% and a loading capacity of 17.7% ±0.2 (n=3), corresponding to 29.7 nmol/mg (nanomoles of PNA divided by milligrams of pSiNPs). In the light of these findings, the same approach was used to test the loading of the positively charged antimiR-335 R_8 and the neutral antimiR-101 Rho. In the case of antimiR-335 R_8, the obtained results were almost comparable with the ones achieved with the negatively charged PNA, in fact the amount of loaded PNA in nanoparticles was found to be 28.6 nmol with a similar loading value (17.6% ± 1.8, n=3). Differently from what observed in the first two cases, Ca-pSiNPs displayed a lower loading value for the neutral-Rho PNA (8.7% ± 2.4, n=3) with a total amount of loaded PNA equals to 16.4 nmol/mg. This behaviour could be

ascribed to the magnitude of electrostatic interactions between the PNA and the nanosystem. In fact, in this case the lack of charges on the tested PNA could cause a decreasing in the loading rate, while for charged PNAs the electrostatic forces seem to be strong enough to promote the incorporation process.

Calcium silicate-based trapping of PNA led to a slight increase in the average hydrodynamic diameter of the nanoparticles (up to ~ 270 nm) and a shift in the ζ-potential to more positive values. We registered negative values of ζ-potential for all the three classes of PNA (-15.8, -9.7, -7.4 mV for neutral, anionic and cationic PNA, respectively), indicating that the trapping procedure results in an efficient encapsulation of the different PNA payloads in the nanoparticle pores, and the final surface charge of the loaded nanoparticles is only determined by the presence of the formed calcium silicate coating.

Since the use of this novel protocol has led to satisfactory results for all the three types of antimiR PNAs, the set of the remaining PNAs was loaded into pSiNPs following the calcium-silicate trapping approach. In the **Table 2.2** below are shown the results for each type of antimiR PNA, with results consistent with those obtained for the anti-miR 335 sequence, suggesting that the loading efficiency is not determined by the PNA sequence, but rather on the polarity of the PNA as determined by the nature of the conjugate tail.

Table 2.2. Loading of different antimiR PNAs payloads.

AntimiR-PNA payload	Loading (nmol/mg)	Loading value (%)
AntimiR-335 R8	29 ± 3	17.6
AntimiR-335 E8	30 ± 1	17.7
AntimiR-335 Rho	16.4 ± 0.8	8.7
AntimiR-101 R8	27 ± 2	16.7
AntimiR-101 E8	29.7 ± 0.3	17.7
AntimiR-101 Rho	16 ± 4	8.7
AntimiR-145 R8	27 ± 2	16.7
AntimiR-145 E8	29.2 ± 0.7	17.7
AntimiR-145 Rho	16.1 ± 4	8.8

2.3.4 *PNA release profiles*

In view of a biological validation of the antimiR PNAs for therapeutic applications in the CF disease, we systematically studied the release kinetics of pSiNP systems loaded with differently charged PNAs. The release of the three PNAs from Ca-pSiNPs was monitored by incubation in phosphate buffered saline (PBS) at 37°, in order to mimic physiological conditions in the absence or with low calcium ion concentrations. The release process was monitored by measuring UV absorbance intensity (λ=260 nm) of released PNA into the supernatant after centrifugation of pSiNPs over time; thus it was possible to obtain a cumulative release profile for the three antimiR PNAs with negative, positive or neutral tail (**Figure 2.15 a-d**).

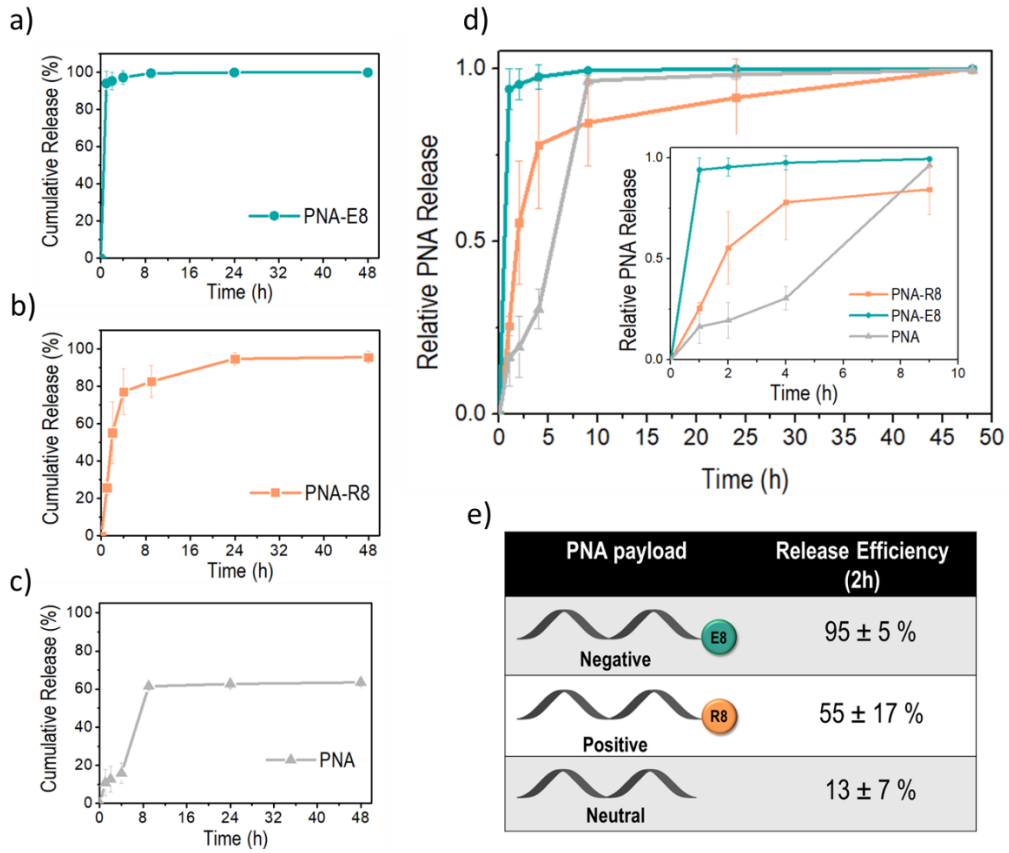


Figure 2.15. Release kinetics of pSiNPs loaded with differently charged PNA payloads. Cumulative release % of PNA-E8 (a), PNA-R8 (b) and PNA (c) over 48 h. The fractional quantity of PNA released from pSiNPs shows a faster release kinetic for anionic PNA-E8 followed by cationic PNA-R8 and neutral PNA. Relative PNA release was normalized to the same scale, where 1.0 equals to the total amount of PNA released in 48 h (d). Three distinct release rates are expressed as release efficiency (defined as mass of released PNA divided by mass of loaded PNA) at the 2-hour timepoint (e).

Three distinct release profiles were obtained based on the net charge of the PNA payload. The release profile of the anionic PNA-E8 obtained in this study matches with the degradation profile of Ca-pSiNPs loaded with locked nucleic acid (LNA)[17] and siRNA,[16] showing a burst release of ~95% of the total PNA loaded within the nanoparticles in the first 2 hours. Nearly 100% efficiency release (defined as mass of released PNA divided by mass of loaded PNA) was observed in this case (**Figure 2.14 a**). PSiNPs loaded with cationic PNA-R8 showed a slower release kinetics compared the previous pSiNP-PNA-E8-conjugates. In this case, nearly 80% of the loaded PNA was released within the first 8 hours, and 95% over 24 hours (**Figure 2.14 b**). Nearly quantitative release of the PNA-R8 payload was observed within 24 hours. The release kinetics of pSiNPs loaded with neutral PNA showed another different, slower temporal profile compared to PNA-R8 and PNA-E8. No major burst release was observed in this case, with an average of 55% of the total PNA loaded from Si nanoparticles released in the first 8 hours (**Figure 2.14 c**). We assume that the absence of electrostatic interactions between the nanoparticles and the oligonucleotide may play a crucial role in its release kinetic. Contrary to the first two cases, the neutral PNA payload was released with an efficiency of ~ 64%. We hypothesize this may be caused by the lower solubility of neutral PNA, when compared to the negatively or positively charged one, leading to the formation of insoluble aggregates that are removed by centrifugation from the supernatant. The release rate of the three formulations has been expressed as the release efficiency after 2 hours and showed that pSiNPs release the anionic payload (PNA-E8) ~ 2-fold and ~ 7-fold faster than cationic PNA-R8 and neutral PNA, respectively. (**Figure 2.15 e**). The difference in the release rate of the three formulations suggests that PNA delivery tuned in time may be achieved by playing on the net charge of the specific PNA payload used. This could result in a great advantage of this type of combination, since, unlike other oligonucleotide analogs such as LNA or 2-O-alkyl-RNA, the PNA net charge (and hence the release rate) can easily be regulated by conjugation with appropriate peptide sequences.

2.3.5 Cellular uptake and Silencing of microRNAs

To verify the effective and potential internalization of the PNA-pSiNPs formulations and their ability to recognize and silence cognate microRNA targets, preliminary *in vitro* experiments were carried out in human bronchial

epithelial cells by Prof. Roberto Gambari group at University of Ferrara. Typically, the lack of charge in the canonical PNA backbone results in very poor cellular internalization properties. Since positively charged PNAs can be internalized also without additional carriers, whereas very poor uptake is normally observed for neutral PNAs, IB3-1 cells were incubated with neutral rhodamine-labeled PNA-loaded pSiNPs. Cellular internalization was evaluated in preliminary experiments by means of fluorescence-activated cell sorting (FACS) analysis, which showed effective internalization of the pSiNP-PNA conjugates in all the three cases based on the fluorescence signal due to rhodamine B (**Figure 2.16 a)**.

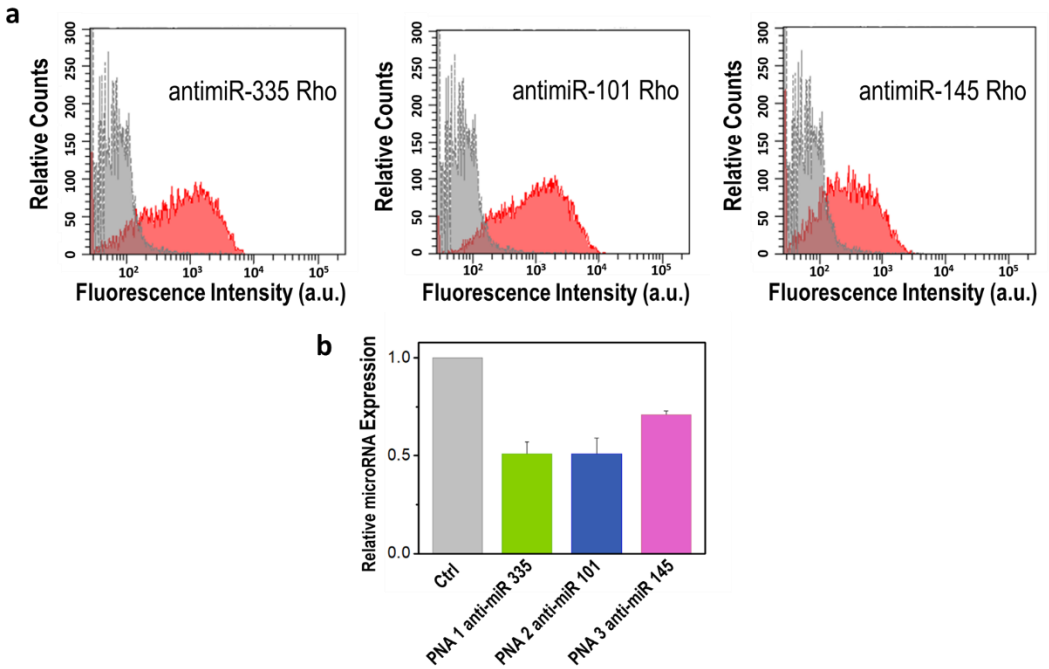


Figure 2.16. a) Flow cytometry graphs of untreated IB3-1 cells (grey) and IB3-1 cells treated with doses of pSiNPs corresponding to 4 μM PNA (red). The fluorescent signal derives from the rhodamine B tag of the PNA payloads. b) Relative microRNA expression evaluated by reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) for miR-335, miR-101 and miR-145 in IB3-1 cells treated with pSiNPs carrying PNA 1 anti-miR-335, PNA 2 anti-miR-101 and PNA 3 anti-miR-145, respectively.

The ability of each PNA sequence released in the cytosol to bind and silence its cognate target microRNA was then measured by quantitative reverse transcription polymerase chain reaction (RT-qPCR). Incubation of IB3-1 cells with each of the three different pSiNP-PNA conjugates, respectively, led to a reduction of the bioavailability of the cognate microRNA target of ~ 50 % in the

case of miR-335 and miR-101, and of ~ 30 % in the case of miR-145, relatively to untreated control samples (**Figure 2.16 b**).

2.3.6 Combined loading of two negatively charged PNAs

To expand the application window of our method, we also demonstrated that combined loading and release of two PNAs with different sequences is readily obtainable. As a proof-of-principle demonstration, we set out to co-load and release two PNAs bearing different sequences. A preliminary experiment was performed using the two negative PNAs PNA1-E8 and PNA2-E8, which were simultaneously co-loaded in the particles following the same trapping protocol. The loading value (calculated as the mass of the oligonucleotides PNA1-E8 + PNA2-E8 divided by the total mass of the PNAs + Si nanoparticles) was comparable to that found for a single-PNA loading (~ 17%). Release of PNA payloads was carried out in PBS, pH 7.4 at 37° over 24 hours. The presence of both species in the solution was confirmed by UPLC-ESI-MS analysis of the supernatant, which demonstrated that already after the first 2 hours both PNAs had started to be released (**Figure 2.17**).

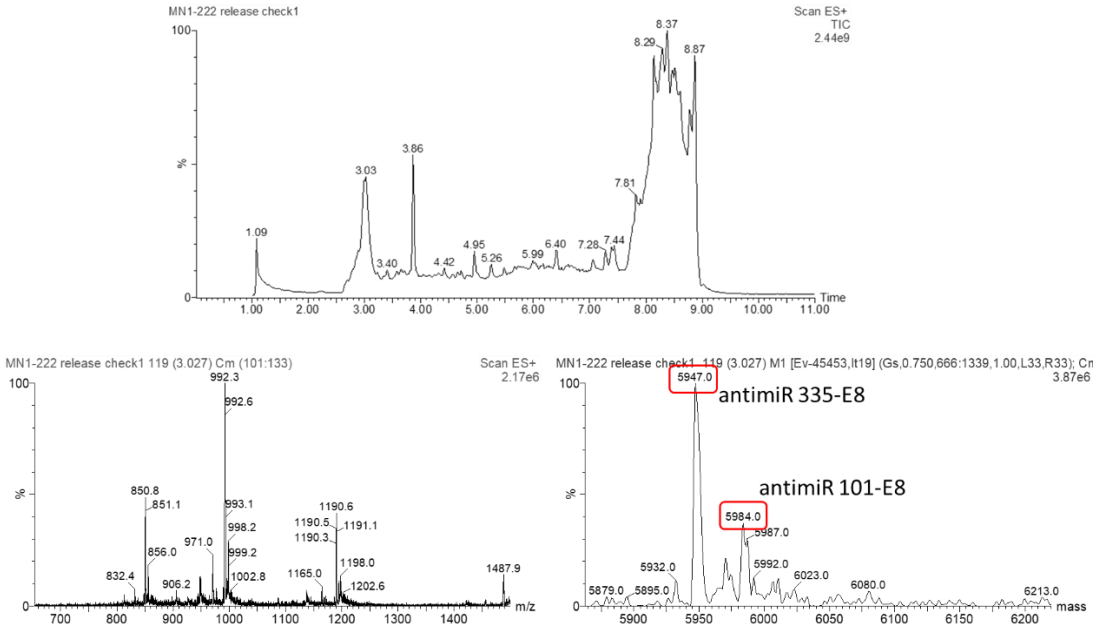


Figure 2.17. UPLC-MS chromatogram of supernatant (top) and ESI-MS spectrum for the peak at 3.03 min with the corresponding mathematic deconvolution of the multicharged signals that shows the mass of the two loaded PNAs: for antimiR 335-E8 corresponding to 5947.0 and antimiR 101-E8 corresponding to 5984.0 (bottom).

2.4 Conclusions

In conclusion, we have demonstrated that PNA can be efficiently loaded into porous silicon nanoparticles following a simple, rapid approach based on calcium silicate trapping. The versatile chemistry of PNA allows for synthesizing oligonucleotides with custom net charge and this can be used to optimize the efficiency of the trapping process. We have shown that calcium silicate trapping is compatible with both a negative, positive and a neutral PNA payload, and that charged PNAs can be encapsulated with loading values of ~ 30 nmol /mg pSi and with an efficiency of > 90%. The charge of the PNA payload then determines the release kinetics of the nanoparticle system, with the release rate increasing from a neutral PNA payload (slowest release) to a negative PNA payload (fastest release. PSiNPs loaded with neutral PNA payloads successfully delivered their cargo in a model cell line relevant to cystic fibrosis and enabled silencing of three specific target microRNAs. These results all demonstrate a strategy for formulating PNA into pSiNPs for therapeutic applications in precision medicine. Especially, the finding that the rate of PNA release can be finely tuned by the type of PNA conjugate can be of great interest for the development of treatments in which the release process is either burst-type or sustained. A combination of these might be also possible by co-delivery of different PNAs.

2.5 Experimental section

2.5.1 General

The reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe, Link Tech., Spectrum Chemicals and used without further purification. DMF was dried over 4 Å molecular sieves and purged with nitrogen to avoid the presence of dimethylamine. Diaza-silane and (hydroxysilyl)phosphonate were supplied by Gelest, Inc. and were used without further purification. Silicon wafers (≈1 mΩ cm resistivity, 100 mm diameter) were purchased from Virginia Semiconductor, Inc.

2.5.2 *Instrumental*

UPLC-ESI-MS was carried out by using a Waters Acquity Ultra Performance LC with Waters Acquity SQ Detector and with ESI interface, and equipped with a Waters Acquity UPLC BEH 300 (50x2.1 mm, 1.7 μm, C18) (UPLC, 0.90 minutes in H_2O 0.2% FA, then linear gradient to 50% MeCN 0.2% FA in 5.70 minutes at a flow rate of 0.25 mL/min. PNA oligomers were purified with RP-HPLC using a Phenomenex Jupiter C18 (5 μm, 300 Å, 250x10 mm) (HPLC1, linear gradient from H_2O 0.1% TFA to 50% MeCN 0.1 % TFA in 30 minutes at a flow rate of 4.0 ml/min). Etching waveforms were generated in a computer program written in Labview (National Instruments, Inc.), and the electric current was driven by a Keithley 2651A Sourcemeter power supply interfaced to the LabView program. A 50T ultrasonic bath (VWR International) was used for ultrasonic fracture of the pSi films. Particles were collected by centrifugation (Eppendorf Centrifuge Model 5424R). Attenuated total reflectance Fourier transform infrared (ATR-FTIR) spectra were recorded using a Thermo Scientific Nicolet 6700 FTIR instrument fitted with a Smart iTR diamond ATR fixture. Dynamic light scattering (DLS) and zeta-potential of nanoparticles were determined using a Malvern Instruments Zetasizer Nano ZS90. Transmission electron microscope (TEM) images were obtained with a JEOL-1200 EX II 120 kV instrument. Nitrogen adsorption-desorption isotherms of the pSi microparticles were recorded at 77 K using a Micro-meritics ASAP 2020 instrument. Prior to the nitrogen adsorption experiment, the pSi microparticles were degassed under vacuum overnight. The surface area of the particles was determined using the BET (Brunauer-Emmett-Teller) method.

2.5.3 *Fmoc-based Solid-Phase Synthesis*

All the PNAs were synthesized at the 5 μmol scale using standard manual Fmoc-based chemistry with HBTU/DIPEA coupling on a Rink amide resin loaded with Fmoc-Gly-OH as first monomer (0.2mmol/g). Commercially available Fmoc-PNA(Bhoc)-OH monomers were used for the manual synthesis. Fmoc protocol was composed of the following steps: a) deprotection 2x8min ; b) DCM Wash; c) dry DMF Wash; d) Kaiser test (1 min, should be positive); e) coupling 1x30 min (activation 2 min, activation solution: 5 eq of monomer and activator, 10 eq of DIPEA in dry DMF); f) DMF wash; Kaiser test (1min, should be negative); g) capping 2x1min; h) DMF wash; i) DIPEA wash 2x2 min; l) DMF wash; m) DCM

wash. The following solutions were used for the solid-phase synthesis of PNA: deprotection: 20% piperidine in DMF; capping: acetic anhydride/DIPEA/dry DMF 5/6/89: DIPEA wash: 5% DIPEA in DMF; cleavage: TFA/*m*-cresol 9/1. Cleavage of the PNA from the resin was performed by treatment with a TFA/m-cresol (9:1 v:v) solution (2 x 1h); the resin was eliminated by filtration and the filtrate was treated to precipitate PNA by adding at least 10 volumes of diethyl ether. The purity and identity of the PNAs were evaluated by UPLC-ESI/MS, using columns and gradient described elsewhere.[18] The PNAs were purified by RP-HPLC using the following conditions: gradient 5 minutes in H2O + 0.1% TFA, then linear gradient from H2O + 0.1% TFA to 50% MeCN + 0.1 % TFA in 30 minutes at a flow rate of 4.0 ml/min. The purity of purified batches was again checked by UPLC-MS. Quantification was performed by dissolving the PNA obtained in 1 mL of water and measuring the absorbance at 260 nm by UV-Vis spectroscopy, using additive base contributions for the calculation of molar absorptivity (ε_{260}: T = 8600, C = 6600, A = 13700, G = 11700 $M^{-1}cm^{-1}$, Rho B = 23173 $M^{-1}cm^{-1}$).

PNA AntimiR-145 R_8. R8-AGG GAT TCC TGG GAA AAC-Gly-NH_2, ε (260 nm): 198000 $M^{-1}cm^{-1}$, 4.8 % yield. UPLC-MS (ESI): r.t. = 2.50 min, m/z: [M] Calcd 6275.4; Found 1256.4 $[M+5H]^{5+}$, 1046.9 $[M+6H]^{6+}$, 897.7 $[M+7H]^{7+}$, 785.5 $[M+8H]^{8+}$, 698.4 $[M+9H]^{9+}$, 628.7 $[M+10H]^{10+}$

PNA AntimiR-145 E_8. E8- AGG GAT TCC TGG GAA AAC-Gly-NH_2, ε (260 nm): 198000 $M^{-1}cm^{-1}$, 1 % yield. UPLC-MS (ESI): r.t. = 2.86 min, m/z: [M] Calcd 6058.3; Found 1213.1 $[M+5H]^{5+}$, 1010.9 $[M+6H]^{6+}$, 866.7 $[M+7H]^{7+}$, 758.5 $[M+8H]^{8+}$

PNA AntimiR-145 Rho. Rho B-AGG GAT TCC TGG GAA AAC-Gly-NH_2, ε (260 nm): 221173 $M^{-1}cm^{-1}$, 1.2 % yield.UPLC-MS (ESI): r.t. = 5.09 min, m/z: [M] Calcd 5451.4; Found 1363.8 $[M+4H]^{4+}$, 1091.1 $[M+5H]^{5+}$, 909.5 $[M+6H]^{6+}$, 779.8 $[M+7H]^{7+}$, 682.4 $[M+8H]^{8+}$

PNA AntimiR-335 R_8. R8- TTT CGT TAT TCT TGA-Gly-NH_2, ε (260 nm): 168300 $M^{-1}cm^{-1}$, 3.1 % yield. UPLC-MS (ESI): r.t. = 2.64 min, m/z: [M] Calcd 6164.3; Found 1234.0 $[M+5H]^{5+}$, 1028.5 $[M+6H]^{6+}$, 881.7 $[M+7H]^{7+}$, 771.5 $[M+8H]^{8+}$, 686.0 $[M+9H]^{9+}$, 617.4 $[M+10H]^{10+}$, 561.4 $[M+11H]^{11+}$

PNA AntimiR-335 E_8. E8- TTT CGT TAT TCT TGA-Gly-NH_2, ε (260 nm): 168300 $M^{-1}cm^{-1}$, 11 % yield. UPLC-MS (ESI): r.t. = 3.0 min, m/z: [M] Calcd 5947.6; Found 1190.9 $[M+5H]^{5+}$, 992.6 $[M+6H]^{6+}$, 850.8 $[M+7H]^{7+}$, 744.7 $[M+8H]^{8+}$

PNA AntimiR-335 Rho. Rho B-TTT CGT TAT TGC TCT TGA- Gly-NH_2, ε (260 nm): 191473 $M^{-1}cm^{-1}$, 14.6 % yield. UPLC-MS (ESI): r.t. = 5.48 min, m/z: [M] Calcd 5340.3; Found 1335.9 $[M+4H]^{4+}$, 1069.0 $[M+5H]^{5+}$, 890.9 $[M+6H]^{46+}$, 763.9 $[M+7H]^{7+}$, 668.5 $[M+8H]^{8+}$, 594.4 $[M+9H]^{9+}$

PNA AntimiR-101 R_8. R8- AGT TAT CAC AGT ACT GTA-Gly-NH_2, ε (260 nm): 188700 $M^{-1}cm^{-1}$, 13.6 % yield. UPLC-MS (ESI): r.t. = 2.69 min, m/z: [M] Calcd 6200.3; Found 1241.3 $[M+5H]^{5+}$, 1035.3 $[M+6H]^{6+}$, 887.4 $[M+7H]^{7+}$, 776.8 $[M+8H]^{8+}$, 690.8 $[M+9H]^{9+}$, 621.7 $[M+10H]^{10+}$

PNA AntimiR-101 E_8. E8- AGT TAT CAC AGT ACT GTA-Gly-NH_2, ε (260 nm): 188700 $M^{-1}cm^{-1}$, 10.1 % yield. UPLC-MS (ESI): r.t. = 2.91 min, m/z: [M] Calcd 5983.7; Found 1497.3 $[M+4H]^{4+}$, 1198.0 $[M+5H]^{5+}$, 998.4 $[M+6H]^{6+}$, 856.0 $[M+7H]^{7+}$, 749.1 $[M+8H]^{8+}$, 666.0 $[M+9H]^{9+}$

PNA AntimiR-101 Rho. Rho B-AGT TAT CAC AGT ACT GTA-Gly-NH_2, ε (260 nm): 211873 $M^{-1}cm^{-1}$, 5.8 % yield. UPLC-MS (ESI): r.t. = 5.04, 5.72, 6.57 min, m/z: [M] Calcd 5376.4; Found 1345.1 $[M+4H]^{4+}$, 1076.0 $[M+5H]^{5+}$, 897.1 $[M+6H]^{6+}$, 769.4 $[M+7H]^{7+}$, 673.6 $[M+8H]^{8+}$, 598.0 $[M+9H]^{9+}$

2.5.4 Porous Silicon Nanoparticle Fabrication

PSi films were prepared by electrochemical perforation etching of a single crystal, (100)-oriented p-type silicon wafer in an electrolyte consisting of 3:1 (v:v) of 48% aqueous HF: absolute ethanol. The preparation followed a published "perforated etch" procedure.[19] Etching was carried out in a Teflon etch cell using a platinum coil counter electrode. Prior to preparation of the porous silicon layers, the wafer surface was cleaned using a sacrificial etch consisting of electrochemical anodization (60 sec, 46 mA/cm^2) in an electrolyte consisting of 3:1 (v:v) 48% aqueous HF : absolute ethanol, followed by ethanol rinse, then dissolution of the porous film with aqueous KOH (1 M). The wafer was rinsed with water, then ethanol. The etching waveform consisted of a square wave in which a lower value of current density of 46 mA/cm^2 was applied for 1.82 s, followed by an upper value of current density of 365 mA/cm^2 applied for 0.363 s. The layered porous nanostructure was removed from the

crystalline silicon substrate by application of current pulse of 3.7 mA/cm^2 for 250 s in an electrolyte consinsting of 1:29 (v:v) of 48% aqueous HF: absolute ethanol. The freestanding Psi film was then fractured by ultrasonication in deionized water (DI H2O, 7 mL) overnight and the resulting pSiNPs were dispersed in an aqueous solution of sodium tetraborate 0.8 mM. The resulting porous silicon nanoparticles were collected using centrifugation (15,000 rpm, 10 min) and washed 3 times with ethanol. The hydroxylated surface functionality (Si-OH, Si-O-Si) was confirmed by ATR-FTIR.

2.5.5 Surface Functionalization of pSiNPs

Amine modification. The pSiNPs were twice suspended in dichloromethane (DCM) and isolated by centrifugation (15,000 rpm, 10 min) in order to remove ethanol. After the washing steps, the pSiNPs (~0.5 mg) were suspended in DCM (200 μL), an aliquot of 2,2-dimethoxy-1,6-diaza-2-silacyclooctane (50 μL) was added, and the mixture was agitated for 2 hr at room temperature. The resulting particles were then washed twice by centrifugation (15,000 rpm, 10 min) from DCM and resuspended in absolute ethanol.
ATR-FTIR: 3300, 2900 cm^{-1}

Carboxylate modification. The amine-modified pSiNPs (~0.7 mg) were isolated by centrifugation (15,000 rpm, 10 min) and suspended in dimethylformamide (DMF, 1 mL). An aliquot of succinic anhydride (20 mg) was added and the mixture was agitated overnight at room temperature. The resulting particles were then washed 2 times by centrifugation (15,000 rpm, 10 min) from DMF and resuspended in absolute ethanol.
ATR-FTIR: 2950, 1750, 1600, 1490 cm^{-1}

Phosphonate modification. To a suspension of pSiNPs (~0.7 mg) in ethanol was added 1.25 μL of tetraethyl orthosilicate and a catalytic amount of TEA at room temperature for 1 h. Subsequently, three aliquots of 3-(trihydroxysilyl)propyl methylphosphonate (2.5 μL x 3) was added and the mixture was agitated overnight. The resulting particles were then washed 4 times by centrifugation (15,000 rpm, 10 min) and resuspended in absolute ethanol.
ATR-FTIR: 2900, 1650, 1300 cm^{-1}

2.5.6 AntimiR-335 R_8 Loading in Carboxy-modified pSiNPs

AntimiR-335 R_8 PNA was loaded into carboxy-pSiNPs by incubating 33% (w:w) PNA:carboxyl pSiNP for 2 h at room temperature in PBS at a final PNA concentration of 1 mg mL^{-1}. PNA–pSiNPs were purified three times by centrifugation and resuspension in deionized water and then absolute ethanol. Percent PNA loading was quantified by measuring absorbance of the PNA recovered in the supernatant after the first centrifugation, using a calibration curve performed with PNA samples in the range 0.1-10 μM in the same medium.

2.5.7 Preparation of PNA-Ca-pSiNPs

A stock solution 4M in calcium chloride ($CaCl_2$) was prepared by adding 2.25 g of solid $CaCl_2$ to 5 mL of RNAse-free water. The solution was centrifuged to remove any precipitates and stored at 4 °C before use.
A solution of anti-miR PNA 150 μM in water (50 μL) was mixed with 0.25 mg pSiNP dispersed in 200 μL of ethanol. To this mixture 250 μL of a $CaCl_2$ solution were added and the reaction was carried out for 45' under mixing. The nanoparticles were washed three times by centrifugation, first using deionized (DI) water (200 μL), then 70% ethanol, and finally absolute ethanol, and the first supernatant was recovered. The analysis of the supernatant was performed using a calibration curve registered in the same medium as the final supernatant solution (200 μL ethanol, 250 μL $CaCl_2$, 50 μL water). The loading was calculated by the difference between the starting quantity of PNA in the loading solution and that of the supernatant after loading.

2.5.8 PNA Release Measurements

PNA-loaded Ca-pSiNPs (0.25 mg, n = 3) were dispersed in 1 mL of PBS, pH 7.4 at 37° with mild shaking. The supernatant containing released PNA was collected after centrifugation at the indicated time points and analyzed by optical absorbance spectroscopy (λ = 260 nm).

For completeness, the procedures used by Prof. Gambari group for cellular studies are briefly summarized hereinafter. Full details are provided in a paper in preparation (Neri *et al.* Tuning the loading and release properties of microRNA-silencing porous silicon nanoparticles by using chemically diverse peptide nucleic acid payloads**,** *ACS Biomater. Sci. Eng.*)

2.5.9 FACS Analysis

Uptake of rhodamine labelled PNAs was evaluated, after 24 hours from transfection, using FACS Canto II (BD, Becton Dickinson, Franklin Lakes, New Jersey, USA), using PE (Phycoerythrin) channel. Cells were detached with trypsin and collected by centrifugation at 1,200 rpm for 10 minutes at room temperature, washed with DPBS 1X, re-suspended in 200 µL of DPBS 1X and analyzed by FACS analysis. For each sample, 10,000 events were acquired, and data analysis was performed using BD FACSDiva Software (BD, Becton Dickinson, Franklin Lakes, New Jersey, USA).

2.5.10 In Vitro RT-qPCR

To verify miRNAs content within cells, RNA was extracted and reverse transcribed using TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA) and miRNA specific stem-loop primers (hsa-miR-101-3p, ID: 002253; hsa-miR-145-5p, ID: 002278; hsa-miR-335-5p, ID: 000546). Reverse transcription quantitative polymerase-chain reaction (RT-qPCR) was performed according to the manufacturer's protocol. All RT reactions, including RT-minus controls and no-template controls, were run in duplicate using the CFX96 Touch Real-Time PCR Detection System (BioRad, Hercules, CA, USA) using TaqMan Universal PCR Master Mix, no AmpErase UNG 2x (Applied Biosystems, Foster City, CA, USA). The relative expression was calculated using the comparative cycle threshold method (ΔΔCT) and using simultaneously U6 snRNA (hsa U6 snRNA, ID:001973) and hsa-let-7c (hsa-let-7c, ID:000379) as endogenous controls.

2.6 *Appendix*

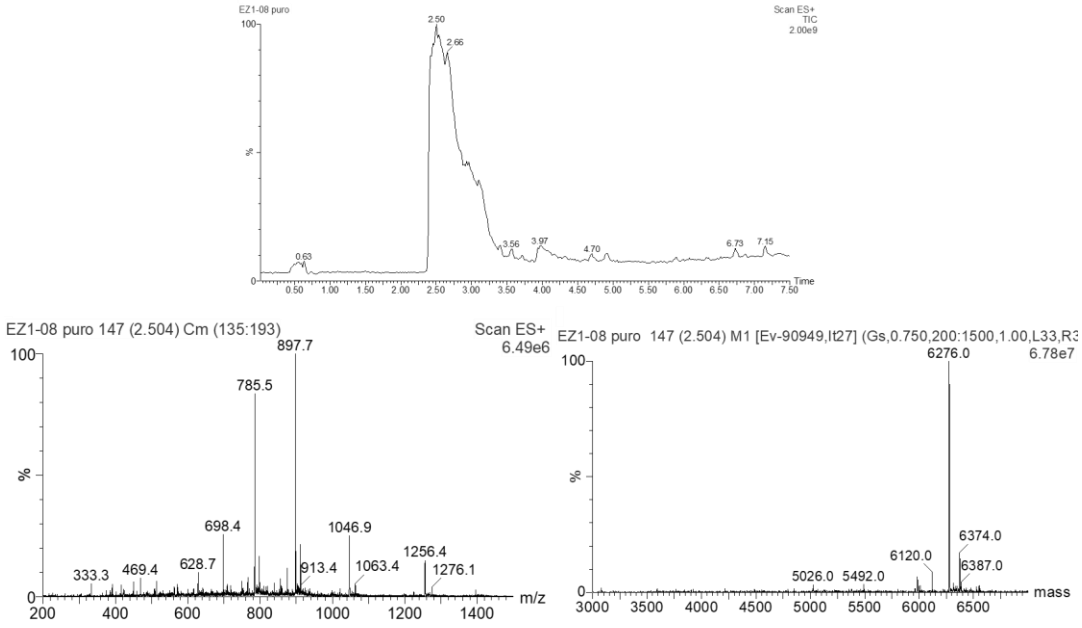


Figure 2.18. UPLC-MS chromatogram of pure anti-miR 145 PNA-R8 (top) and ESI-MS spectrum for the peak at 2.50 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

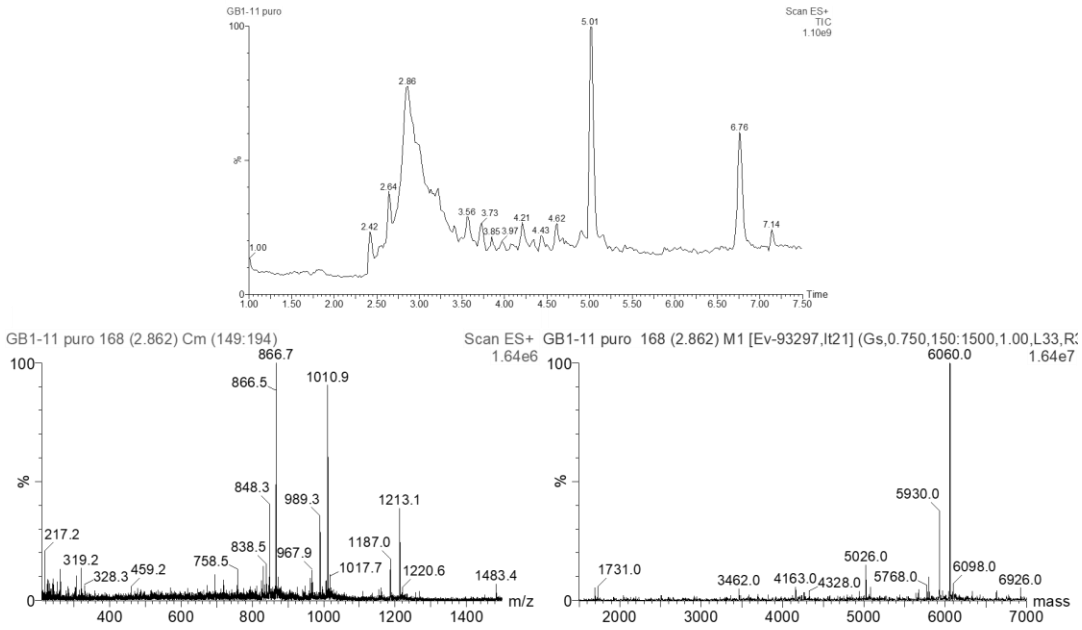


Figure 2. 19. UPLC-MS chromatogram of pure anti-miR 145 PNA-E8 (top) and ESI-MS spectrum for the peak at 2.86 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

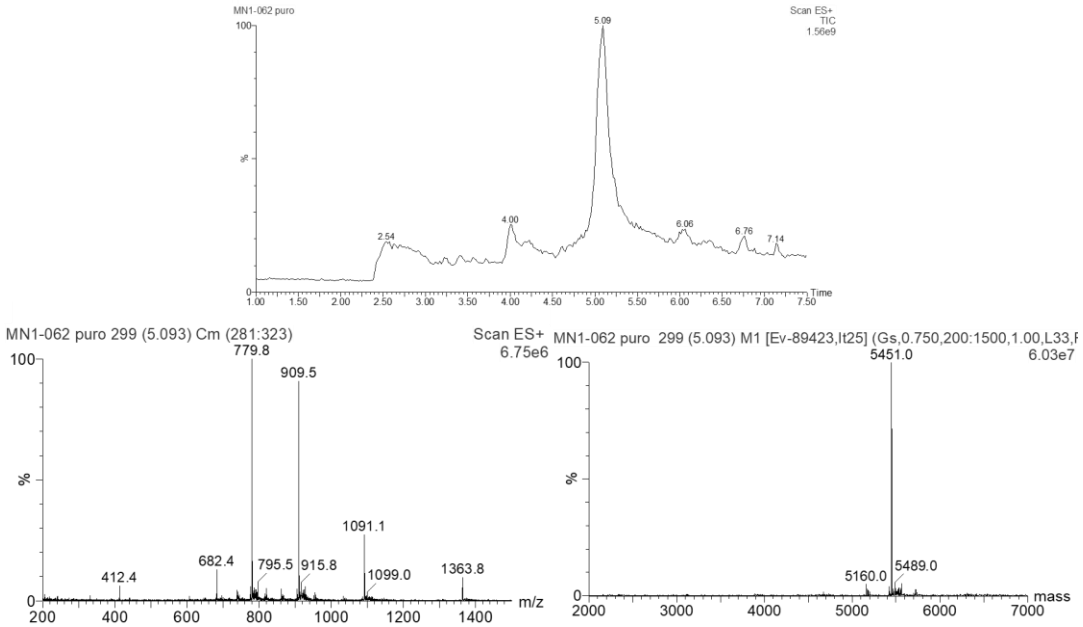


Figure 2. 20. UPLC-MS chromatogram of pure anti-miR 145 PNA Rho (top) and ESI-MS spectrum for the peak at 5.09 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

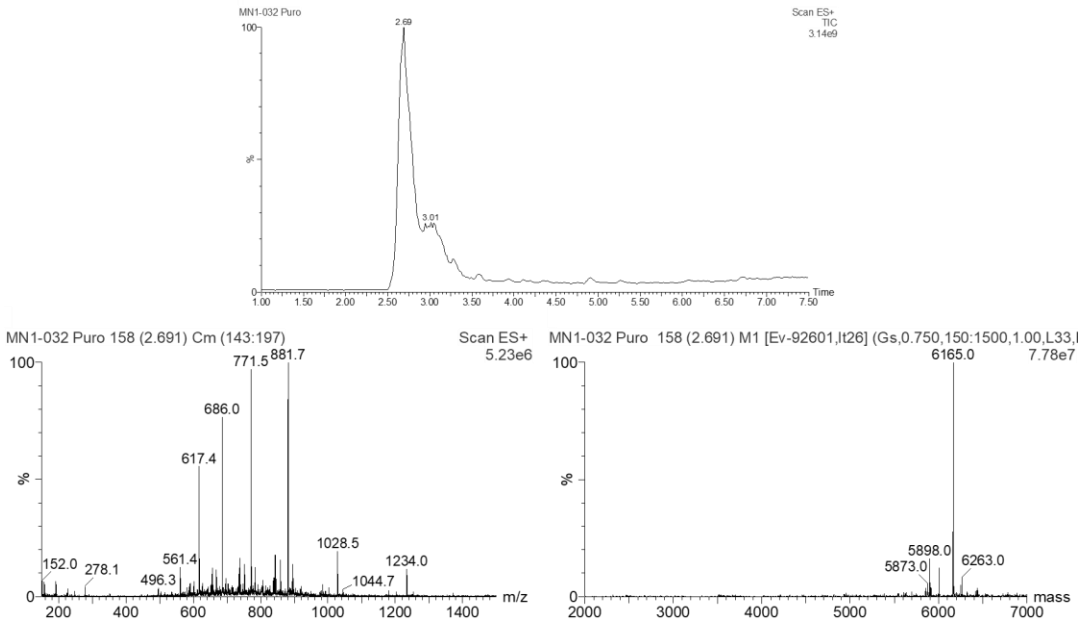


Figure 2. 21. UPLC-MS chromatogram of pure anti-miR 335 PNA-R8 (top) and ESI-MS spectrum for the peak at 2.69 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

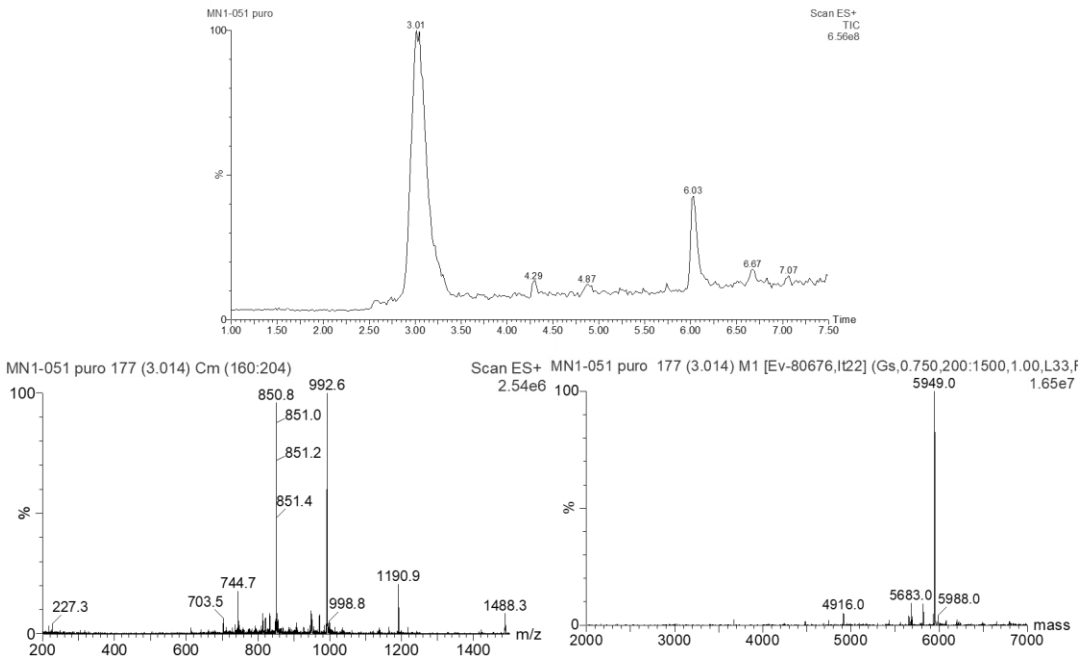


Figure 2.22. UPLC-MS chromatogram of pure anti-miR 335 PNA-E8 (top) and ESI-MS spectrum for the peak at 3.01 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

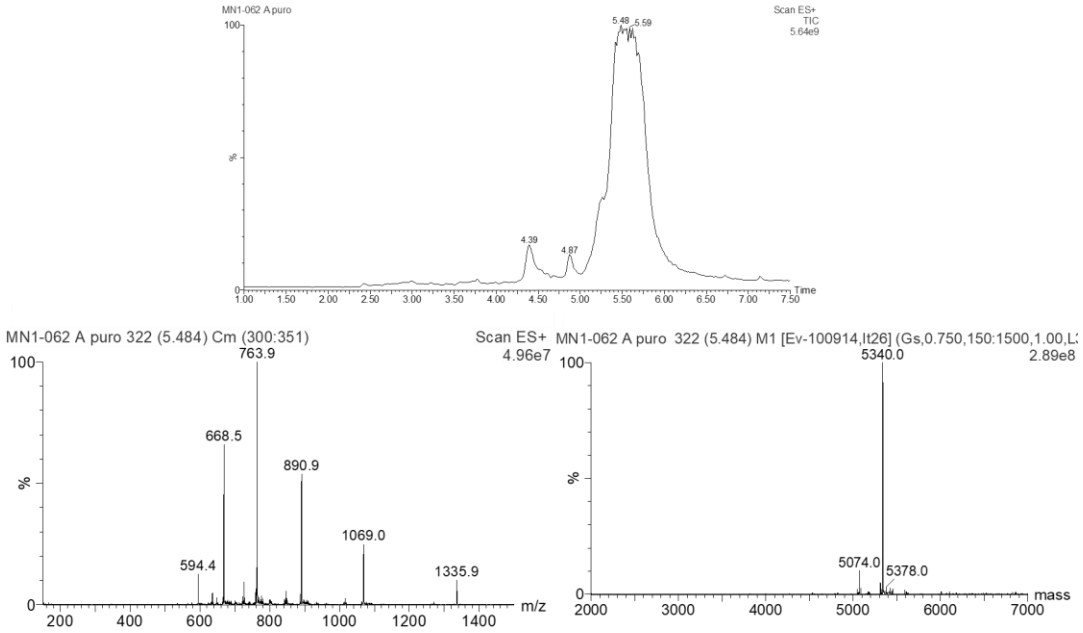


Figure 2.23. UPLC-MS chromatogram of pure anti-miR 335 PNA Rho (top) and ESI-MS spectrum for the peak at 5.48 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

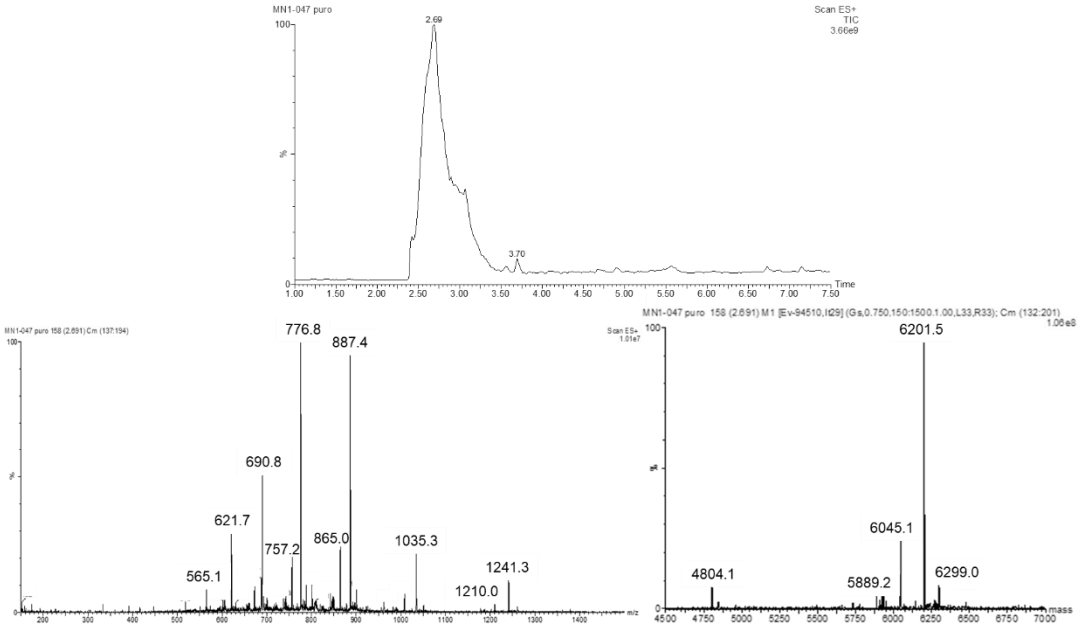


Figure 2.24. UPLC-MS chromatogram of pure anti-miR 101 PNA-R8 (top) and ESI-MS spectrum for the peak at 2.69 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

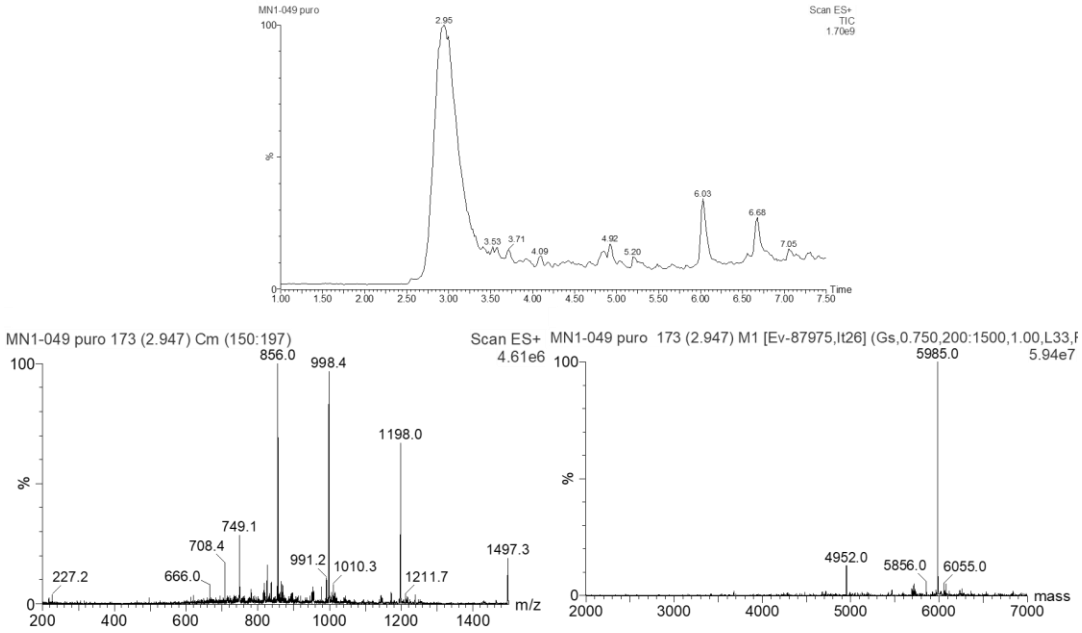


Figure 2.25. UPLC-MS chromatogram of pure anti-miR 101 PNA-E8 (top) and ESI-MS spectrum for the peak at 2.95 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

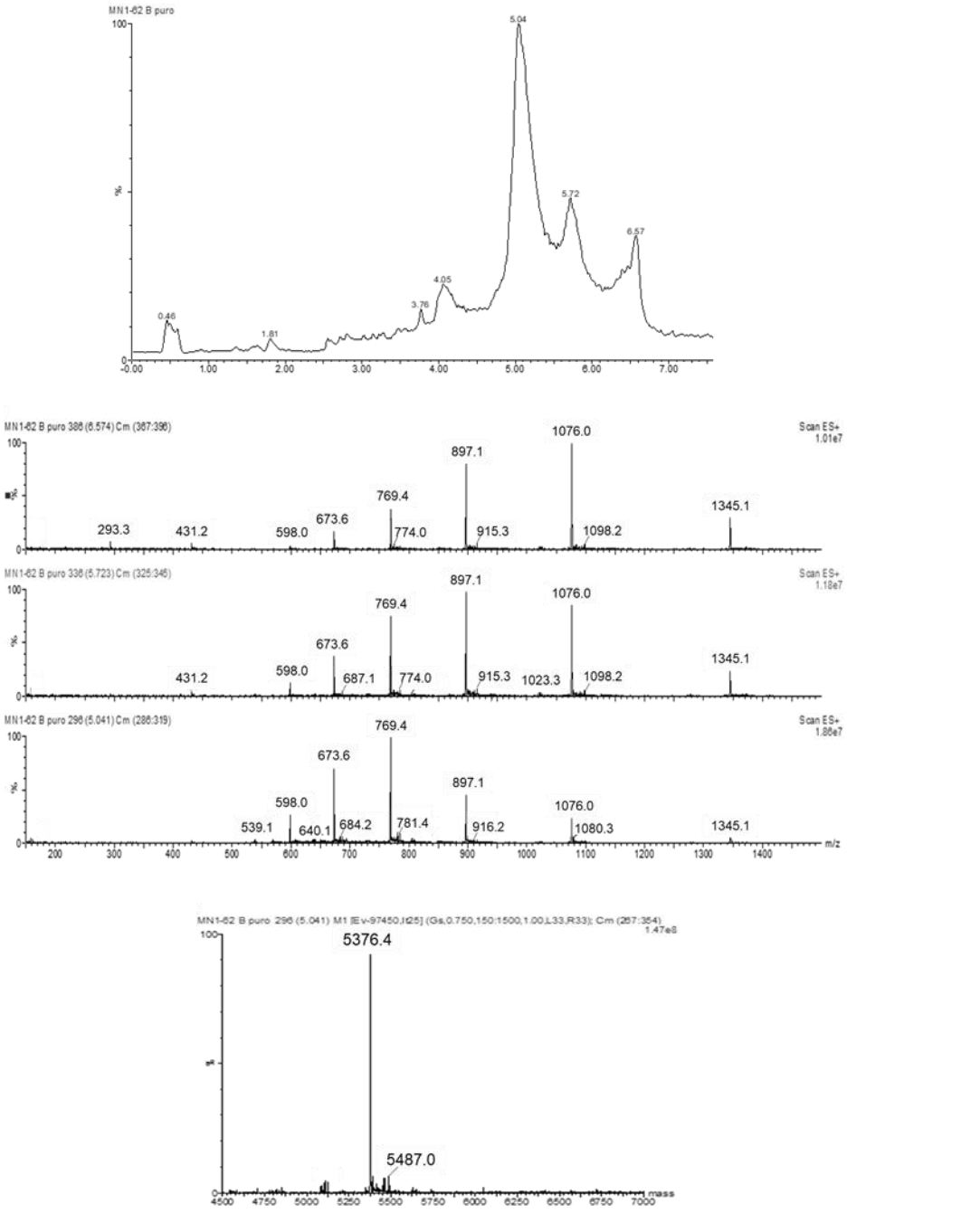


Figure 2.26. UPLC-MS chromatogram of pure anti-miR 101 PNA Rho (top) and ESI-MS spectrum (central) for the peak at 5.04 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).

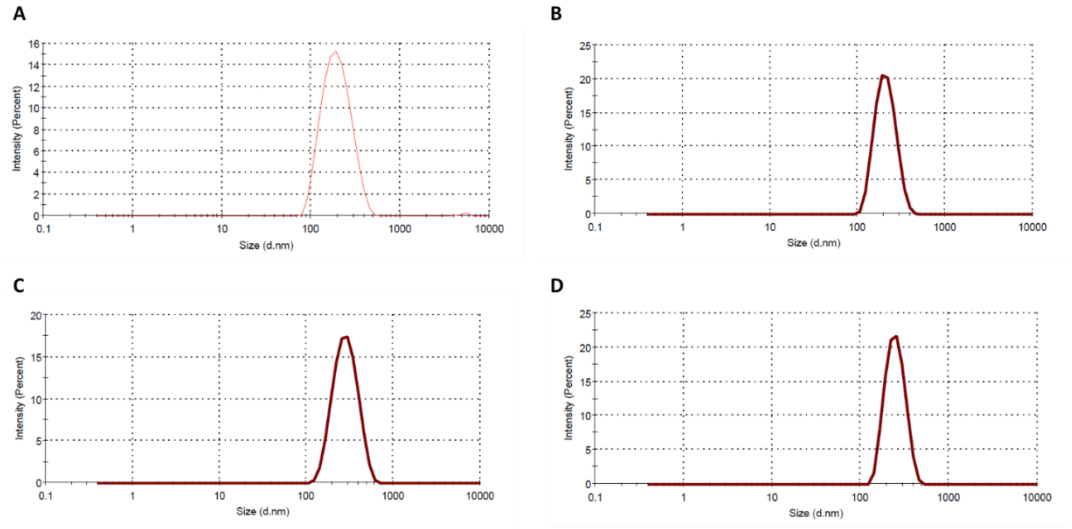


Figure 2.27. Intensity weighted size distribution from dynamic light scattering of freshly etched pSiNPs (A), and after loading with negatively charged anti-miR 145 PNA-E8(B), positively charged anti-miR 145 PNA-R8 (C) and neutral anti-miR 101PNA Rho (D).

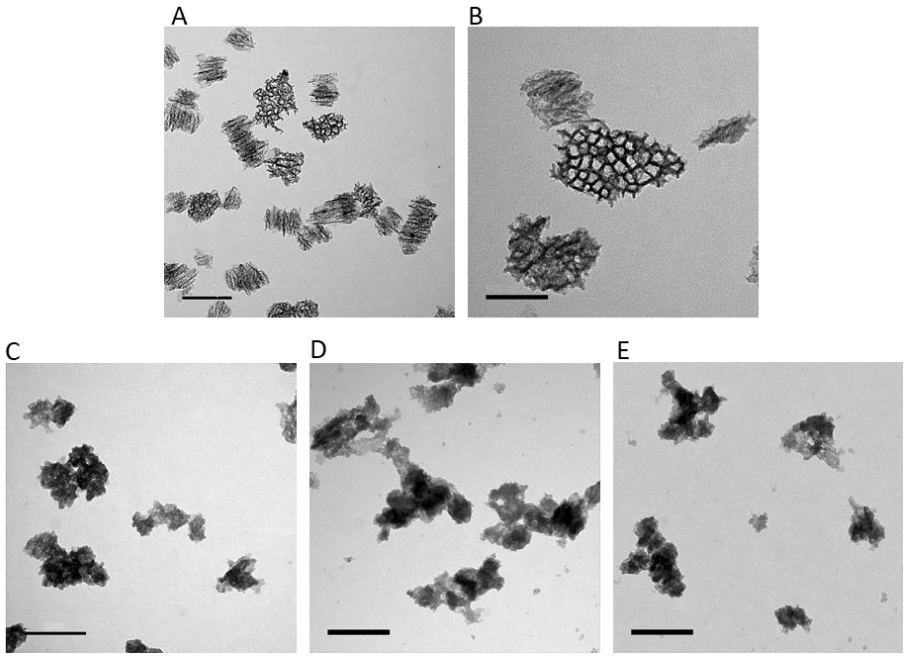


Figure 2.28. Transmission electron microscope images of freshly etched pSiNPs (A) and (B), Ca-pSiNPs-anti-miR 145 PNA-E8 (C), Ca-pSiNPs-anti-miR 145 PNA-R8 (D) and Ca-pSiNPs-anti-miR 101 PNA Rho (E). Scale bar = 100 nm for (B); Scale bar = 200 nm for (A), (B), (C) and (E).

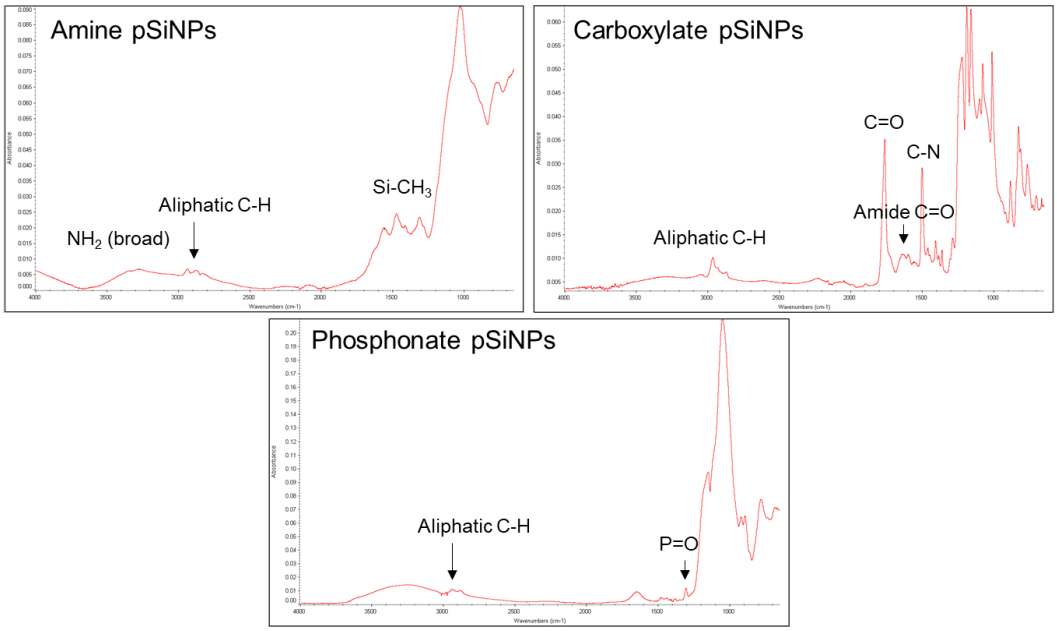


Figure 2.29. Fourier transform infrared (FTIR) spectroscopy of amine, carboxylate, and phosphonate pSiNPs.

2.7 *References*

1. Steven M. Rowe, M.D., Stacey Miller, B.S., and Eric J. Sorscher, M. . Cystic fibrosis Cystic fibrosis. *Cent. Dis. Control* **196,** 1–6 (2000).

2. Wainwright, C. E., Elborn, J. S., Ramsey, B. W., Marigowda, G., Huang, X., Cipolli, M., Colombo, C., Davies, J. C., De Boeck, K., Flume, P. A., Konstan, M. W., McColley, S. A., McCoy, K., McKone, E. F., Munck, A., Ratjen, F., Rowe, S. M., Waltz, D. & Boyle, M. P. Lumacaftor-ivacaftor in patients with cystic fibrosis homozygous for phe508del CFTR. *N. Engl. J. Med.* **373,** 220–231 (2015).

3. Mckiernan, P. J. & Greene, C. M. MicroRNA Dysregulation in Cystic Fibrosis. *Mediators Inflamm.* **2015,** 1–7 (2015).

4. Fabbri, E., Tamanini, A., Jakova, T., Gasparello, J., Manicardi, A., Corradini, R., Sabbioni, G., Finotti, A., Borgatti, M., Lampronti, I., Munari, S., Dechecchi, M. C., Cabrini, G. & Gambari, R. A peptide nucleic acid against microrna miR-145-5p enhances the expression of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) in Calu-3 Cells. *Molecules* **23,** (2018).

5. Fabbri, E., Tamanini, A., Jakova, T., Gasparello, J., Manicardi, A., Corradini,

R., Finotti, A., Borgatti, M., Lampronti, I., Munari, S., Dechecchi, M. C., Cabrini, G. & Gambari, R. Treatment of human airway epithelial Calu-3 cells with a peptide-nucleic acid (PNA) targeting the microRNA miR-101-3p is associated with increased expression of the cystic fibrosis Transmembrane Conductance Regulator (CFTR) gene. *Eur. J. Med. Chem.* **209,** 112876 (2020).

6. Zarrilli, F., Amato, F., Morgillo, C. M., Pinto, B., Santarpia, G., Borbone, N., D'Errico, S., Catalanotti, B., Piccialli, G., Castaldo, G. & Oliviero, G. Peptide nucleic acids as miRNA target protectors for the treatment of cystic fibrosis. *Molecules* **22,** 1–16 (2017).

7. Sultan, S., Rozzi, A., Gasparello, J., Manicardi, A., Corradini, R., Papi, C., Finotti, A., Lampronti, I., Reali, E., Cabrini, G., Gambari, R. & Borgatti, M. A Peptide Nucleic Acid (PNA) Masking the miR-145-5p Binding Site of the 3'UTR of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) mRNA Enhances CFTR Expression in Calu-3 Cells. *Molecules* **25,** 1–12 (2020).

8. McNeer, N. A., Anandalingam, K., Fields, R. J., Caputo, C., Kopic, S., Gupta, A., Quijano, E., Polikoff, L., Kong, Y., Bahal, R., Geibel, J. P., Glazer, P. M., Mark Saltzman, W. & Egan, M. E. Nanoparticles that deliver triplex-forming peptide nucleic acid molecules correct F508del CFTR in airway epithelium. *Nat. Commun.* **6,** 1–11 (2015).

9. Tamanini, A., Fabbri, E., Jakova, T., Gasparello, J., Manicardi, A., Corradini, R., Finotti, A., Borgatti, M., Lampronti, I., Munari, S., Dechecchi, M. C., Cabrini, G. & Gambari, R. A Peptide-Nucleic Acid Targeting miR-335-5p Enhances Expression of Cystic Fibrosis Transmembrane Conductance Regulator (<i>CFTR</i>) Gene with the Possible Involvement of the CFTR Scaffolding Protein NHERF1. *Biomedicines* **9,** (2021).

10. Beckmann, K. H. Investigation of the chemical properties of stain films on silicon by means of infrared spectroscopy. *Surf. Sci.* **3,** 314–332 (1965).

11. Segal, E., Perelman, L. A., Cunin, F., Di Renzo, F., Devoisselle, J.-M., Li, Y. Y. & Sailor, M. J. Confinement of Thermoresponsive Hydrogels in Nanostructured Porous Silicon Dioxide Templates. *Adv. Funct. Mater.* **17,** 1153–1162 (2007).

12. Kim, D., Zuidema, J. M., Kang, J., Pan, Y., Wu, L., Warther, D., Arkles, B. & Sailor, M. J. Facile Surface Modification of Hydroxylated Silicon Nanostructures Using Heterocyclic Silanes. *J. Am. Chem. Soc.* **138,** 15106–15109 (2016).

13. Kwon, E. J., Skalak, M., Bertucci, A., Braun, G., Ricci, F., Ruoslahti, E., Sailor,

M. J. & Bhatia, S. N. Porous Silicon Nanoparticle Delivery of Tandem Peptide Anti-Infectives for the Treatment of Pseudomonas aeruginosa Lung Infections. *Adv. Mater.* **29,** 1–9 (2017).

14. Beavers, K. R., Mares, J. W., Swartz, C. M., Zhao, Y., Weiss, S. M. & Duvall, C. L. In situ synthesis of peptide nucleic acids in porous silicon for drug delivery and biosensing. *Bioconjug. Chem.* **25,** 1192–1197 (2014).

15. Beavers, K. R., Werfel, T. A., Shen, T., Kavanaugh, T. E., Kilchrist, K. V., Mares, J. W., Fain, J. S., Wiese, C. B., Vickers, K. C., Weiss, S. M. & Duvall, C. L. Porous Silicon and Polymer Nanocomposites for Delivery of Peptide Nucleic Acids as Anti-MicroRNA Therapies. *Adv. Mater.* **28,** 7984–7992 (2016).

16. Kang, J., Joo, J., Kwon, E. J., Skalak, M., Hussain, S., She, Z. G., Ruoslahti, E., Bhatia, S. N. & Sailor, M. J. Self-Sealing Porous Silicon-Calcium Silicate Core–Shell Nanoparticles for Targeted siRNA Delivery to the Injured Brain. *Adv. Mater.* **28,** 7962–7969 (2016).

17. Bertucci, A., Kim, K. H., Kang, J., Zuidema, J. M., Lee, S. H., Kwon, E. J., Kim, D., Howell, S. B., Ricci, F., Ruoslahti, E., Jang, H. J. & Sailor, M. J. Tumor-Targeting, MicroRNA-Silencing Porous Silicon Nanoparticles for Ovarian Cancer Therapy. *ACS Appl. Mater. Interfaces* **11,** 23926–23937 (2019).

18. Mejia-Ariza, R., Rosselli, J., Breukers, C., Manicardi, A., Terstappen, L. W. M. M., Corradini, R. & Huskens, J. DNA Detection by Flow Cytometry using PNA-Modified Metal–Organic Framework Particles. *Chem. – A Eur. J.* **23,** 4180–4186 (2017).

19. Qin, Z., Joo, J., Gu, L. & Sailor, M. J. Size Control of Porous Silicon Nanoparticles by Electrochemical Perforation Etching. *Part. Part. Syst. Charact.* **31,** 252–256 (2014).

Chapter 3

3 Modified 'Chiral Box' PNA probes for WT and mutated KRAS

3.1 Introduction

Peptide nucleic acids have shown to be suitable tools in several biomedical applications, specifically in diagnostics they have been exploited for detecting DNA and RNA sequences thanks to their excellent hybridization properties with a particular focus on discrimination of single-point mutations.[1,2]

Kirsten rat sarcoma (KRAS) gene encodes for the protein called K-*ras* (G-protein) that is a part of the RAS/MAPK pathway. This protein acts as a molecular switch and cyclically passes from an activated form to an inactivated one by binding guanosine triphosphate (GTP). The Epidermal Growth Factor Receptor (EGFR) belongs to the ErbB family of surface cellular receptors, which in normal cells are able to activate different signal transduction pathway including the mitogen RAS-activated protein kinase. These pathways can regulate the differentiation, proliferation, angiogenesis and cell apoptosis, so in the case of some point mutations occur in the KRAS gene, these can determine an anomaly of the KRAS proteins, which remain always in the active form, leading to an abnormal cell growth. KRAS mutations occur in 35-45% of colorectal cancers.[3] In the remaining 60%, the tumor does not present KRAS mutations and is therefore defined as wild-type KRAS. In colorectal carcinomas, further 5,700 different point mutations of the KRAS gene have been identified and a high percentage of these mutations occur in codons 12 or 13. It is worth mentioning that KRAS mutations are known to bypass upstream pharmacological inhibitors of the signaling pathway RAS/MAPK, such as monoclonal anti-EGFR antibodies, making it imperative for fast detection of these biomarkers in order to obtain fundamental prognostic information regarding the efficacy of specific drugs. In the past years, peptide nucleic acids have already been demonstrated to be suitable in the detection of KRAS mutations.[4–7] An interesting example was reported by D'Agata and coworkers, who developed an ultrasensitive nanoparticle-enhanced plasmonic method for detecting mutated or wild-type RAS sequences by exploiting a liquid biopsy approach and taking advantage of the use of PNA probes. This non-invasive approach allowed to

discriminate between KRAS mutated ctDNAs (circulating tumor DNA) over wild-type cfDNA (cell-free DNA) with a sensitivity of 100% and specificity of 83.33% without the need of amplification of tumor DNA from plasma.
Although these results are highly relevant, the method would greatly benefit from the further increase in sequence-selectivity and in the ability of probes to discriminate even single-point mutations. In fact, in early detection of cancer it is likely that only a small fraction of mutated DNA is present together with a large excess of wild-type DNA, and this situation need the maximum sequence selectivity possible to provide reliable results.

3.2 Aim of the study

As already mentioned in Chapter 1, modified PNAs bearing positively charged side chains on their backbone can exhibit improved performances both in their cellular uptake and in their hybridization ability with cognate DNA/RNA. Specifically, the conjugation of C-2 position of the PNA backbone with D-amino acid side chains have been shown to enhance the selectivity for single-mismatched sequences,[8,9] a crucial feature for the diagnosis of genetic diseases or malignancies. The effect is increased when stretches of three consecutive C2- D-modified chiral monomers, the so-called 'Chiral Box' (CB), are placed in the middle of the strand (**Figure 3.1**). Moreover, the stereogenic centres have proved to induce a strong preference for a right-handed helix of the PNA/DNA duplex provided by the high chiral restraint in the middle of the PNA strand with a complete direction control (only antiparallel complex is formed).

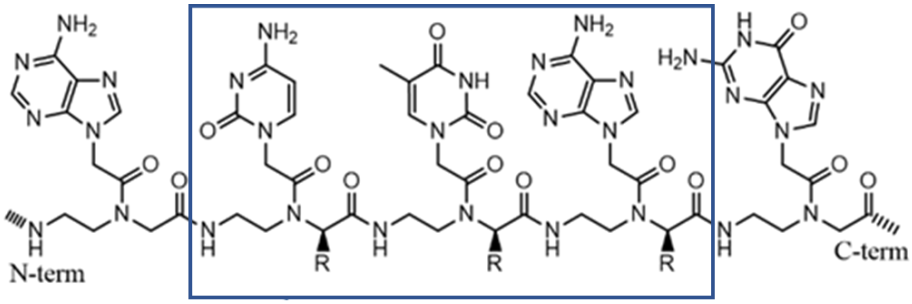


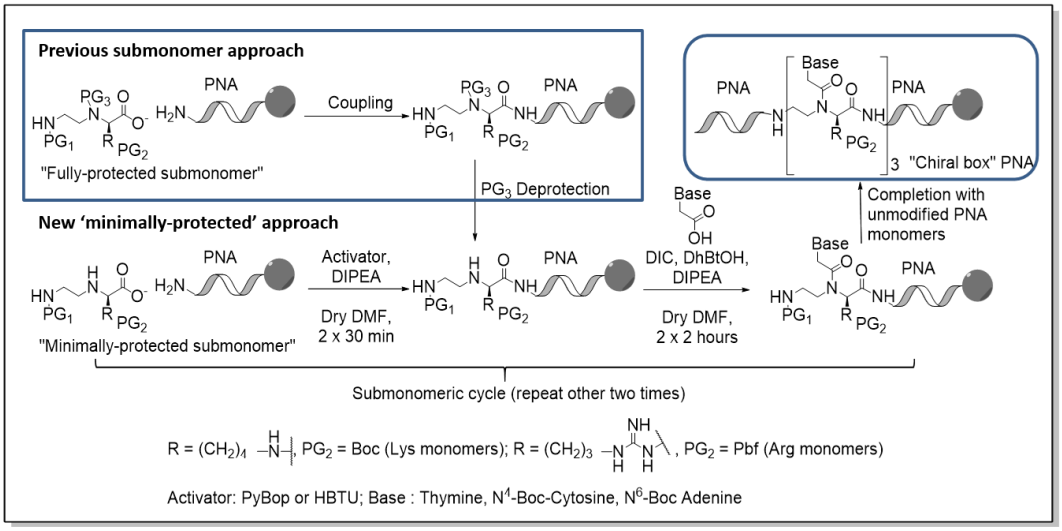
Figure 3.1. Schematic of 'Chiral Box' PNA.

On the other hand, the synthesis of C-2 modified PNAs remains one of the major drawbacks for their use since epimerization reactions can occur during the process leading to a mixture of different stereoisomers.[10] In this work a new synthetic approach for the preparation of 2D-Chiral Box PNA is presented and used to synthesize two PNA chiral probes of KRAS exon 2 G12D mutation, which is extremely relevant for monitoring the efficacy of antibody-based therapies in colorectal cancer.[11] Preliminary DNA/PNA hybridization studies were also performed on sensor surface in order to establish the potential of these probes in discriminating mismatched sequences. This work have been obtained as part of that described in a paper published in *Organic Letters*,[12] and was designed as a part of the ULTRAPLACAD EU project for the development of new instrumentation and protocols for early colorectal cancer diagnosis, in which both the University of Parma and the Austrian Institute of Technology (AIT) were involved as partners.

3.3 Results and discussion

3.3.1 The 'minimally-protected' approach

A "submonomeric approach", which provides for a sequential incorporation of the backbone (the submonomer) and of the nucleobase on a growing PNA chain directly on the solid support, can be exploited to minimize the epimerization process.[13,14] Nevertheless, this protocol remains challenging from a synthetic point of view since requires the introduction of three orthogonal protecting groups on the submonomer (**Scheme 3.1, top**), thus creating limitations on the choice of reaction conditions that can be used with the additional risk of unwanted deprotection. Furthermore, the combination of certain protecting groups is not compatible with most of commercial automatic synthesizers since the latter are designed for Fmoc/Bhoc chemistry. In the host laboratory, a shorter strategy for the production of optically-pure 'Chiral Box' PNAs was recently proposed in a previous PhD Thesis (Rozzi. A., PhD Thesis, 2018) (**Scheme 3.1, center**). This is based on the use of minimally-protected submonomers that feature protecting groups only at the N-6 position and on the side chain linked to C-2 position.[12]



Scheme 3.1. Comparison of Solid Phase Synthesis of "Chiral Box" PNAs (Top, Right Panel) with Fully and Minimally Protected Submonomers (Top Left and Bottom Routes, Respectively). Adapted with permission from ref. [13] Copyright © 2021 The Authors. Published by American Chemical Society.

Indeed, the synthesis of PNAs by the submonomeric approach showed that the coupling of carboxymethylnucleobase is much slower than that involving the primary amine at the N-term with the incoming submonomer, implying that also an eventual addition of the PNA backbone on itself should be hindered because of the bulkiness of both the nucleophile and the electrophile. For this reason, our group reasoned that building blocks lacking protecting groups on their N-3 position could be suitable to perform the former reaction. Accordingly, also the protection of the C-term carboxylate could be unnecessary during the synthesis of the 'minimally-protected' submonomers, which could be directly performed in a single reductive amination step.

In a previous work (Rozzi. A., PhD Thesis, 2018) different "Chiral Box" PNAs presenting a fully complementary sequence for the G12D point mutation of the *KRAS* gene were produced according to the newly developed synthetic approach (**Figure 3.2, left**), which have been completed with other sequences synthesized in the present work . An ATC stretch of nucleobases, attached on three consecutive modified backbones, formed the "Chiral Box" portion. The central nucleobase T was expected to face the single point mutation of the target DNA. The minimally-protected building blocks depicted in **Figure 3.2 (right)** were attached on the N-term of a growing oligomer and then, the appropriated (Boc-protected)

carboxymethylnucleobase[26] (i.e. A, C, and T) was directly coupled on the modified backbone. Most importantly, the developed method was shown to preserve the enantiomeric purity of the PNA monomer, as evaluated by chiral GC analysis performed with a Chirasil-Val column on the piperazine-2-one derivatives obtained by the PNAs after hydrolysis.

List of PNA sequences:

14-mer achiral	H-CTA CGC CAT CAG CT-Gly-NH_2	**PNA1**
14-mer CB-D-Lys	H-CTA CGC CAT CAG CT-Gly-NH_2	**PNA2**
14-mer CB-D-Arg	H-OO-CTA CGC CAT CAG CT-Gly-NH_2	**PNA3**
11-mer achiral	H-O-CGC CAT CAG CT-Gly-NH_2	**PNA4**
11-mer CB-D-Lys	H-O-CGC CAT CAG CT-Gly-NH_2	**PNA5**

DNA sequences:

KRAS G12D mutation	5'-AG CTG ATG GCG TAG-3'	**DNA1-FM**
	5'-AG CTG ATG GCG-3'	**DNA2-FM**
KRAS wild type	5'-AG CTG GTG GCG TAG-3'	**DNA1-MM**
	5'-AG CTG GTG GCG-3'	**DNA2-MM**

*O = 2-(2-aminoethoxy)ethoxyacetic acid

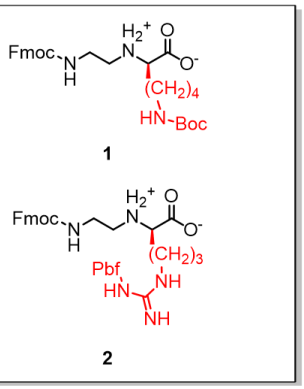


*Figure 3.2. "Chiral Box" (CB) PNA, unmodified PNA and DNA sequences used in the previous work (ref. [13]). C(2)-modified monomers are highlighted in red (Left panel); Minimally protected 2D-Lys **1** and 2D-Arg **2** submonomers (Right panel).*

The recognition properties of these PNAs were tested by evaluation of the thermal stabilities of the complexes formed with complementary DNA strands presenting full-matched (FM, G12D-mutated) and mismatched (MM, wild type) sequences (**Figure 3.2, left**). The use of the 14-mer **PNA 2** and **PNA 3** in place of the corresponding unmodified **PNA 1** showed a higher performance of the C(2)-modified oligomers for the discrimination of single point mutations, resulting in increasing of the difference in melting temperature between the two types of duplexes (ΔT_m) by 1.6° C and 0.9° C, respectively. It is worth noting that the best discrimination ability was obtained for the Lys-based 11-mer **PNA 5** with a remarkable ΔT_m value of 19.2°C between the full-matched and the mismatched complexes, 3.5°C higher than that afforded by its achiral analog **PNA 4.** It is reasonable to assume that a shorter sequence of the probes increases the influence of the 'Chiral Box' stretch and thus the selectivity. Additionally, by comparing the melting temperature of the adducts that 'Chiral Box' PNAs and their unmodified version form with complementary DNAs (i.e. T_m = 59.9°C *vs* 65.9°C for **PNA 4** and **5**, respectively) we can suggest that the sum of destabilizing steric effects generated by the three adjacent modified-monomers leads to less stable duplexes. This apparent disadvantage is

balanced by the higher selectivity of the PNA/DNA interaction, resulting in a higher sensitivity for a single mismatch in cognate DNA strands.

Table 3.1. Comparison of melting temperatures for unmodified and 'Chiral box' PNAs with full-matched and mismatched target DNA. ΔTm = Tm (Full Match)-Tm (Mismatch). Adapted with permission from ref [13] Copyright © 2021 The Authors. Published by American Chemical Society.

Hybrid type	DNA	Tm [°C]	ΔTm [°C]
PNA 1	DNA1-FM	76.7±0.9	14.1
PNA 1	DNA1-MM	62.6±0.6	
PNA 2	DNA1-FM	72.4±0.8	15.7
PNA 2	DNA1-MM	56.7±0.8	
PNA 3	DNA1-FM	73.0±2.0	15.0
PNA 3	DNA1-MM	58.0±0.9	
PNA 4	DNA2-FM	65.9±0.5	15.9
PNA 4	DNA2-MM	50.0±0.5	
PNA 5	DNA2-FM	59.9±0.2	19.2
PNA 5	DNA2-MM	40.7±0.5	

Given the increased performances of 'Chiral Box' PNAs in terms of mismatch recognition for complementary DNA strands, the same sequences were selected to test their discrimination ability on surface in view of a potential development of optical sensing devices.

3.3.2 Synthesis of biotinylated PNA probes

Surface Plasmon Resonance (SPR) biosensors are the most commonly employed optical biosensors for the detection and characterization of biological targets. This technique involves the use of a catcher molecule that is immobilized to a sensor surface whereas the target analyte is injected for a real-time monitoring of the binding event, thanks to small local changes in refractive index occurring at the sensor surface upon biomolecular interactions.

As catcher probes, 'Chiral box' (CB) PNAs containing D-Lys (**PNA 6, Figure 3.3a**) and D-Arg (**PNA 7, Figure 3.3b**) side chains on C-2 position were synthesized according to the recently developed 'minimally-protected' protocol.

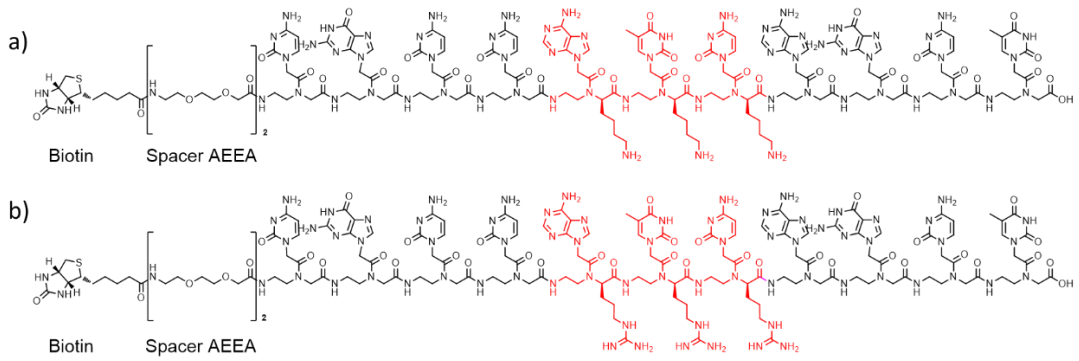


Figure 3. 3. 'Chiral Box' PNA sequences synthesized in this work. a) 11-mer CB D-Lys PNA; b) 11-mer CB D-Arg PNA. 'Chiral Box' stretch is highlighted in red.

We decided to focus on 11-mer-long sequences as the best results in terms of discrimination were previously obtained for these sequences. The synthesis of the two CB PNAs started with the introduction of the first four monomers, namely AGCT, by following standard Fmoc/Bhoc protocol described in Chapter 1. Next, minimally protected building blocks **1** or **2** were attached on the N-term of the growing oligomer. The minimally-protected D-submonomers (**1** and **2**) were already available in our laboratory and they were obtained following two distinct reductive amination methods.[12] As shown in **Scheme 3.1** (**center**), the introduction of modified submonomers requires an activation step right before the coupling. In this case, HBTU was employed as activating agent since similar overall efficiency was found compared to the use of PyBop, a further commercial activator. Compound **1** or **2** (5 equivalents) were previously dissolved in DMF showing a slightly different solubility. In fact, D-Lys backbone **1** needed a strong sonication step at 40°C for several minutes, in contrast to D-Arg compound **2**. After the dissolution of the submonomers, HBTU (5 equivalents) was added and the coupling of the primary amine at the N-term with the incoming minimally protected submonomers was readily obtained. The possibility of double backbone attachment was previously verified for the first coupling of both Lys-based **1** and Arg-based **2** submonomers on **PNA 2** and **PNA 3**, respectively.[12] As UPLC-MS analysis did not show any traces of the aforementioned double backbone coupling, the present synthesis (**PNA 6** and **PNA 7**) were checked after each nucleobase addition by UPLC-MS. Following backbone introduction, coupling with the appropriate (Boc-protected) carboxymethylnucleobase (in this case A, C and T) was performed using a combination of strong coupling agents (DIC/DhBtOH). As expected,

this step confirmed to be slower than the addition of the backbone (**1** or **2**) and longer reaction times (2x 2 hours) were required to finalize the coupling. In particular, the synthesis of the D-Lys CB PNA worked better than the D-Arg CB PNA (**Figures 3.4 d** and **3.5 d**). In fact, in the latter case the N3 secondary amine showed a very low reactivity and some unreacted backbone was present even after several treatments with different CMNBs. It is worth mentioning that in both present cases the starting PNA (pre-CB sequence) was not recovered differently from the synthesis of CB sequences used in previous studies.

Analysis of UPLC-MS spectra after the completion of the 'Chiral Box' part of both PNAs showed only small traces of acetylated derivatives both on terminal amine and on the secondary N3-amino group of the PNA backbone. In **Figures 3.4** and **3.5**, the results of PNA composition after the completion of the 'Chiral Box' part is reported. From a comparison of the present data (**Figures 3.4** and **3.5**) and those reported in the previous work[12] appears a slight increase in the synthesis of the chiral stretch as unreacted starting PNA was not found and the presence of acetylated derivatives was not significant.

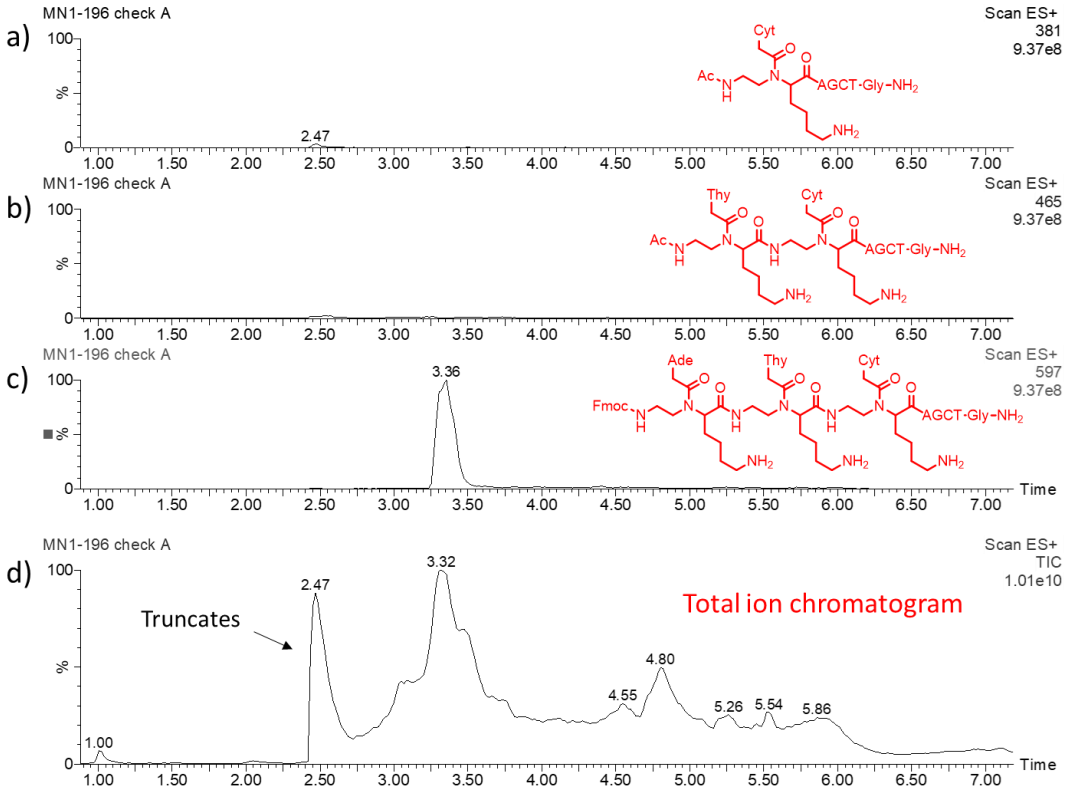


*Figure 3.4. UPLC-MS analysis of **PNA 6** after the completion of its 'Chiral box' stretch. The traces generated by the $[M+4H]^{4+}$ ions of derivatives containing a) one, b) two or c) three C(2)-modified monomers were extracted and normalized to the same intensity scale for evaluating the efficiency of chain elongation (Ac = acetyl group, Ade = Adenine, Thy = Thymine, Cyt = cytosine); d) Total ion current (TIC) chromatogram of a Fmoc-protected crude sample.*

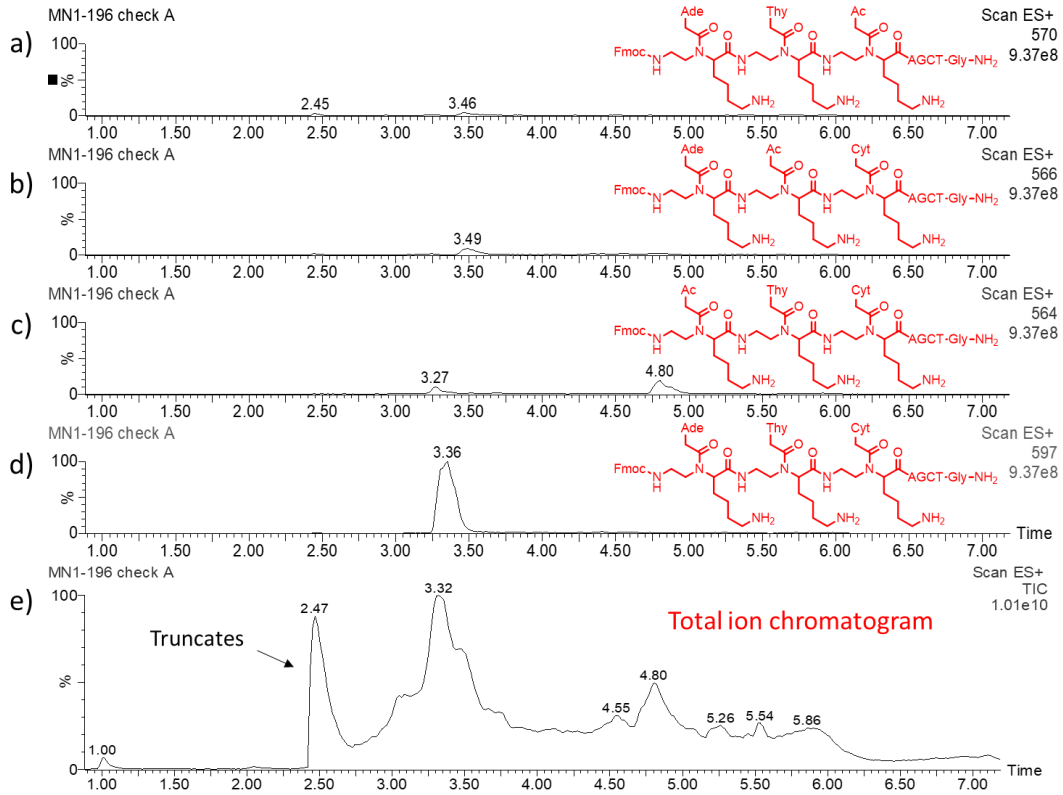


*Figure 3.5. UPLC-MS analysis of **PNA 6** after the completion of its 'Chiral box' stretch. The traces generated by the $[M+4H]^{4+}$ ions of derivatives acetylated on a) - c) one of their N(3) positions or d) none were extracted and normalized to the same intensity scale for evaluating the efficiency of the nucleobase insertion (Ac = acetyl group, Ade = Adenine, Thy = Thymine, Cyt = cytosine). e) Total ion current (TIC) chromatogram of a Fmoc-protected crude sample.*

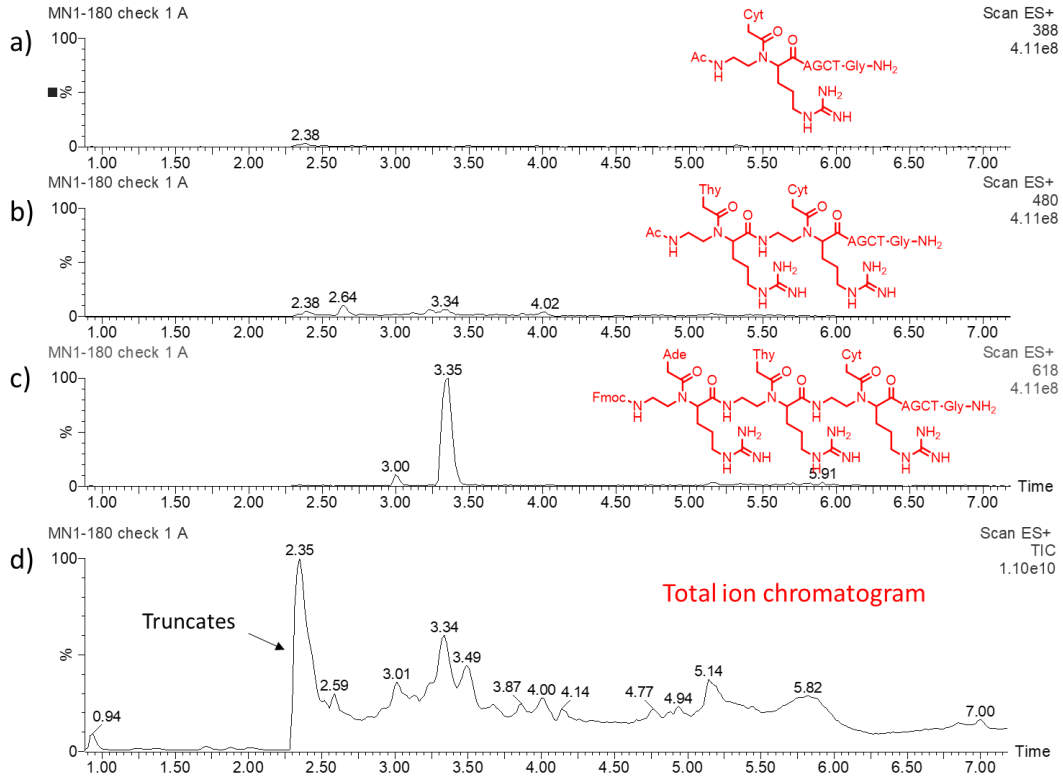


*Figure 3.6. UPLC-MS analysis of **PNA 7** after the completion of its 'Chiral box' stretch. The traces generated by the $[M+4H]^{4+}$ ions of derivatives containing a) one, b) two or c) three C(2)-modified monomers were extracted and normalized to the same intensity scale for evaluating the efficiency of chain elongation (Ac = acetyl group, Ade = Adenine, Thy = Thymine, Cyt = cytosine); d) Total ion current (TIC) chromatogram of a Fmoc-protected crude sample.*

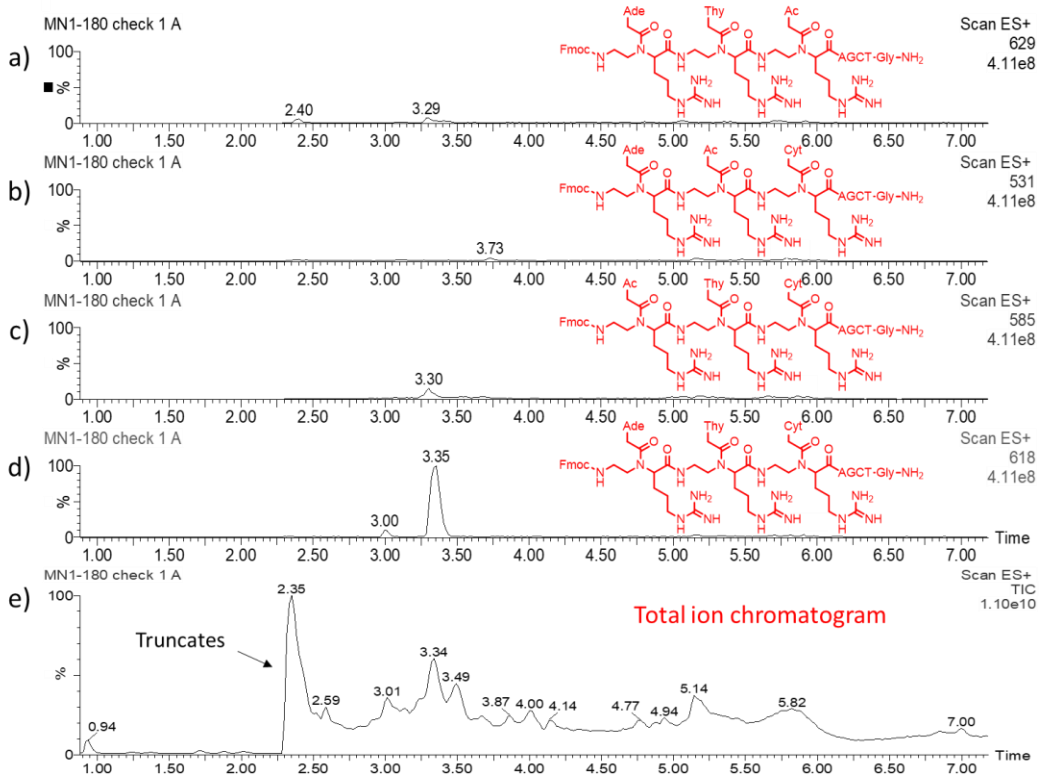


*Figure 3.7. UPLC-MS analysis of **PNA 7** after the completion of its 'Chiral box' stretch. The traces generated by the $[M+4H]^{4+}$ ions of derivatives acetylated on a) - c) one of their N(3) positions or d) none were extracted and normalized to the same intensity scale for evaluating the efficiency of the nucleobase insertion (Ac = acetyl group, Ade = Adenine, Thy = Thymine, Cyt = cytosine). e) Total ion current (TIC) chromatogram of a Fmoc-protected crude sample.*

The two CB PNA sequences were then completed following the standard Fmoc/Bhoc procedure and a biotin was coupled at the end of each strand to allow the immobilization of the chiral PNA probes on the sensor surface.
To investigate the discrimination ability towards mismatched sequences of the CB PNAs, unmodified achiral PNA was synthetized as control (**PNA 8**). The synthesis was performed by an automatic peptide synthesizer and the biotin was then attached manually according to standard Fmoc-based solid-phase synthesis (SPS) protocol.
PNA 6, **PNA 7** and **PNA 8** were then cleaved from the resin, purified by RP-HPLC and characterized by UPLC/ESI-MS (**Figure 3.8- 3.10**).

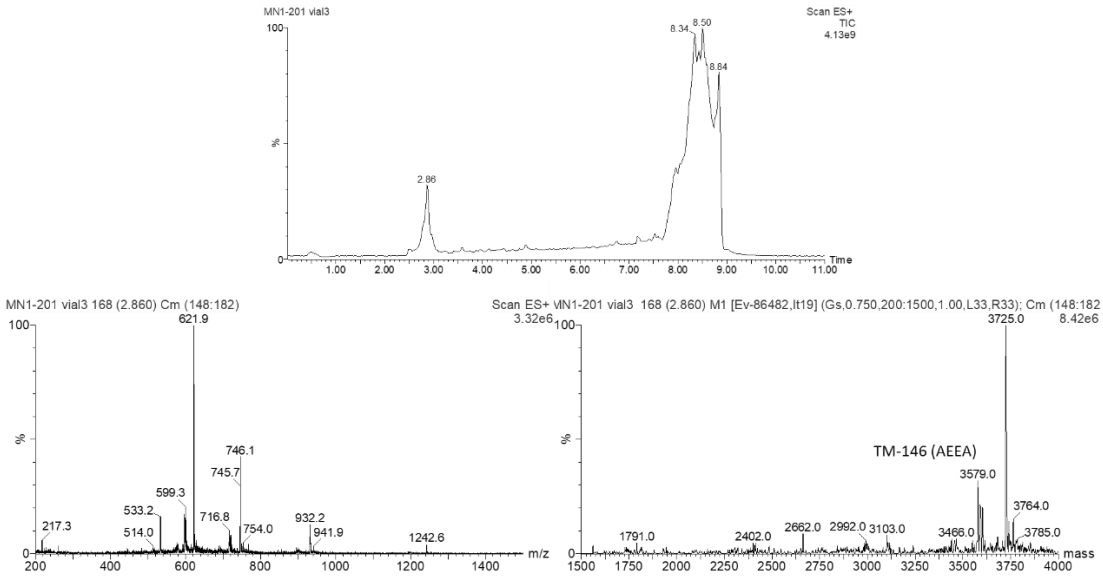


*Figure 3.8. UPLC-MS chromatogram of pure **PNA6** (top) and ESI-MS spectrum for the peak at 2.86 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).*

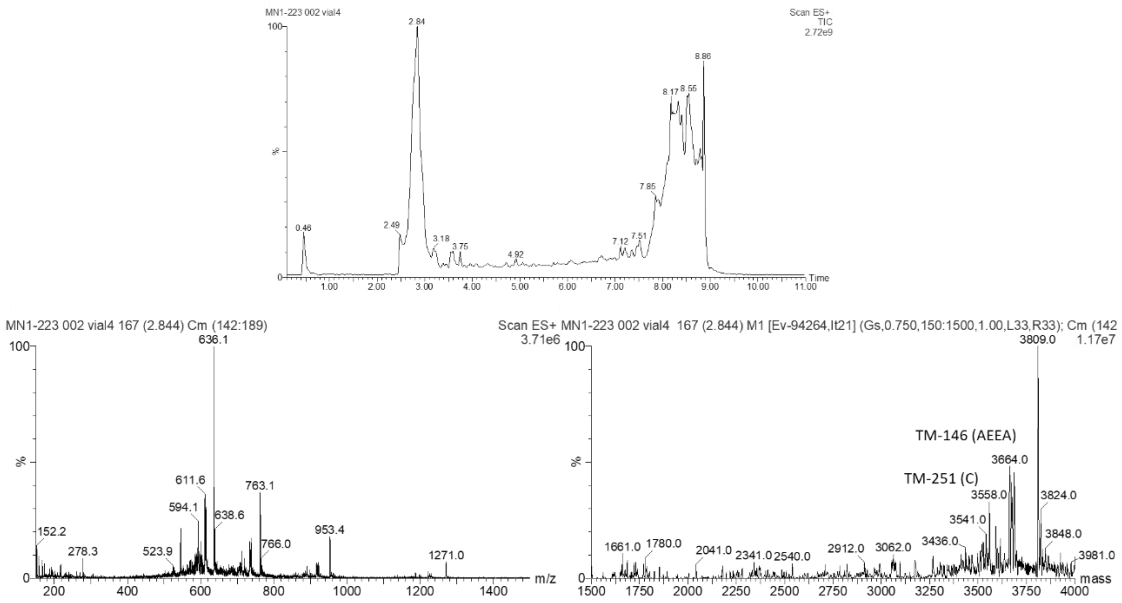


*Figure 3.9. UPLC-MS chromatogram of pure **PNA7** (top) and ESI-MS spectrum for the peak at 2.84 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).*

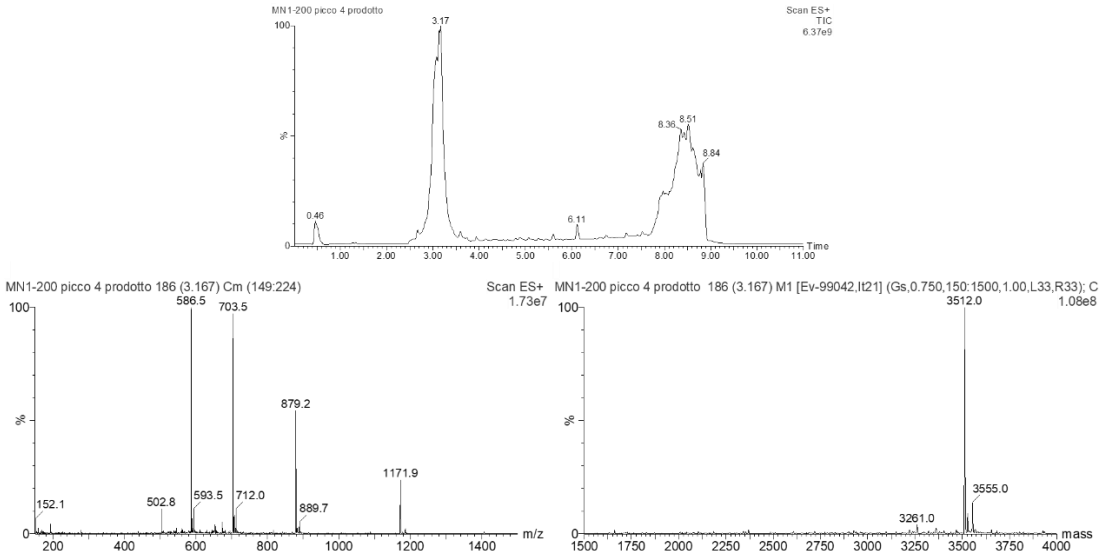


*Figure 3.10. UPLC-MS chromatogram of pure **PNA8** (top) and ESI-MS spectrum for the peak at 3.17 min with the corresponding mathematic deconvolution of the multicharged signals (bottom).*

3.3.3 DNA/PNA hybridization studies

SPR spectroscopy is a versatile optical technique largely used for monitoring interfacial binding reactions, generally in Kretschmann configuration. The interaction of the light with a noble metal surface (usually gold or silver) can efficiently generate an evanescent light-wave, when the energy carried by the photons equals that of the collective oscillations of valence electrons of the metal, i.e. surface plasmons, and these two waves are phase-matched along the surface. The incident light beam is send towards the plane of interface between the metal and a dielectric medium at an angle θ. Above the critical angle, a total internal reflection (TIR) event occurs and an evanescent field is formed at the upper interface (**Figure 3.8 a**). When the thin metallic film is present at the surface, a decrease of the intensity of the reflected light beam occurs at the angle where it is phase matched the surface plasmon waves travelling at the outer metal surface. This resonant angle varies with the refractive index of the medium contacted with the metal surface, for example due to capture of molecules (**Figure 3.8 b**). By recording the reflected intensity at a fixed angle of incidence as a function of time is possible to register dynamic measurements useful to establish physical

parameters of the capture process. It becomes clear that SPR can be exploited to detect specific biomolecules by using a proper sensing architecture.

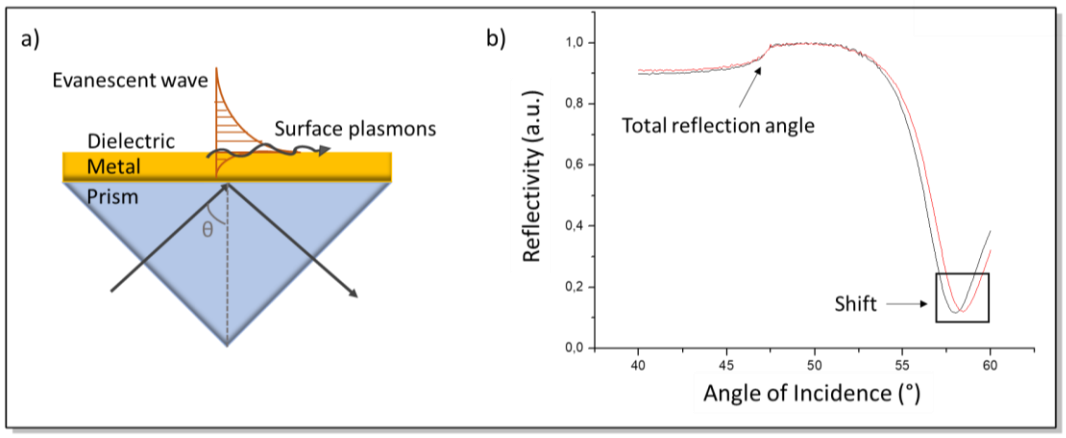


Figure 3.11. a) Total internal reflection (TIR) construct for surface plasmon excitation in Kretschmann geometry; b) Reflectivity curves (angular scans) of surface plasmon spectroscopy (SPR) before (black) and after (red) binding of analyte on the sensor surface.

Despite the great potential of the SPR technique for monitoring interfacial binding reactions, for short oligonucleotides (10-30 nt), change of refractive index is not significant to be analyzed in detail. The combination of surface plasmon and fluorescence spectroscopy (SPFS) can be used to increase the sensitivity and allow the characterization of molecular interactions at the interface between a catcher probe (i.e. in our case PNA) and target analyte labelled with a fluorophore.[15] In SPFS technique, the surface plasmon evanescent waves are used to excite the chromophore-labelled analyte close to the metal surface and to monitor the fluorescent intensity during the hybridization process.

In the work done during the secondments at the Austrian Institute of Technology (AIT), under the supervision of Prof. Wolfgang Knoll and Dr. Jakub Dostalek, the biotinylated D-Lys 'Chiral Box' PNA (**PNA 6**) and its achiral homologous (**PNA 8**) were used as catcher probes in a SPFS sandwich-type assay, a standard assay already established for these type of studies in the host laboratory (**Figure 3.12**). This was used in previous works for studying the ability of PNA probes to detect oligonucleotides and PCR products, and for assessing the thermodynamic and kinetic aspects of the sensing processes.[16–18] It is worth noting that the selected architecture ensures a proper distance between the labeled target DNA and the gold

surface that allows to minimize Förster quenching of the emitted fluorescence.

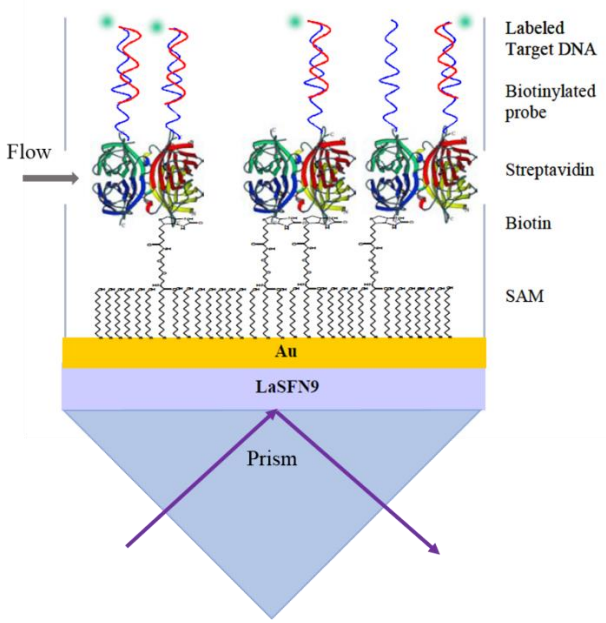


Figure 3.12. Schematic of sandwich-type architecture used in surface plasmon-enhanced fluorescence spectroscopy (SPFS) binding studies. Adapted with permission from Park et al. Copyright © 2006, American Vacuum Society.

A 50 nm thick layer of gold was deposited onto glass slides by thermal evaporation, preceded by 2 nm of chromium film useful to improve the adhesion of the gold film to the glass substrate. The gold surface was incubated in ethanolic solution with a thiol carrying biotin headgroup (HS-C11-EG3-Biotin) and a thiol with ethylene glycol head group (HS-C11-EG3-OH) mixed at a molar ratio 1:4 (**Figure 3.10**) in order to form the self-assembled monolayer (SAM). The protocol is optimized in to minimize nonspecific adsorption and to obtain a suitable dilution of biotin sites on the surface that allows a proper streptavidin layer formation. In fact, a too high density of biotinylated thiol was observed to reduce the accessibility of the binding sites for the protein.[16]

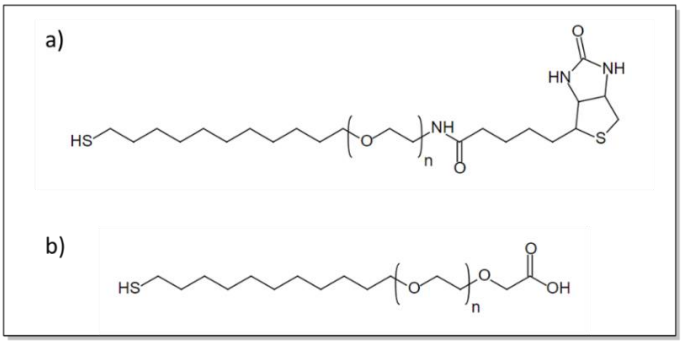


Figure 3.13. Molecular structure of a) biotinylated thiol (n=3) and b) spacer thiol (n=3).

Since binding events are detected as SPR changes at the sensor surface, the latter is clamped against a flow cell that allows a controlled injection of the samples. Using this set-up (see **Figures 3.16** and **3.17** for instrumental components), a streptavidin solution (1.86 μM) is injected into the flow cell at a flow rate of 30μL/s to allow the immobilization of biotinylated PNA probes at the binding matrix. Subsequently, a solution of PNA (**PNA 6** or **PNA 8**, 1μM) was rinsed through the flow cell and coupled to the streptavidin layer thanks to the extraordinary affinity between the latter and the biotin moiety of PNA ($K_A \sim 10^{14}$ M^{-1}). SPR measurements confirmed the immobilization of both streptavidin layer and PNA probes (**Figure 3.14**).
A fully-complementary Cy5-labeled DNA (Cy5-5'-AGC TGA TGG CG-3') and its single mismatched analog (Cy5-5'-AGC TG**G** TGG CG-3') were used as targets to assess the affinity and sequence selectivity of the chiral probe **PNA 6** compared to an achiral control **PNA 8**. By injecting the chromophore-labeled DNA target (10 nM, 100 nM) before the PNA immobilization, it was possible to check unspecific binding with the SAM-streptavidin layer. As depicted in **Figure 3.14 a**, it was not observed any unexpected adsorption of DNA. Indeed, after the injection of DNA target the fluorescence signal increased up to 5.8 x 10^3 cps (in the case of 100 nM) and then decreased immediately to the baseline level (1.7 x 10^3) due to the replacement of the target solution with fresh buffer solution. This signal increase is observed as the SPR evanescent wave causes the excitation of the fluorescent chromophore in the bulk.

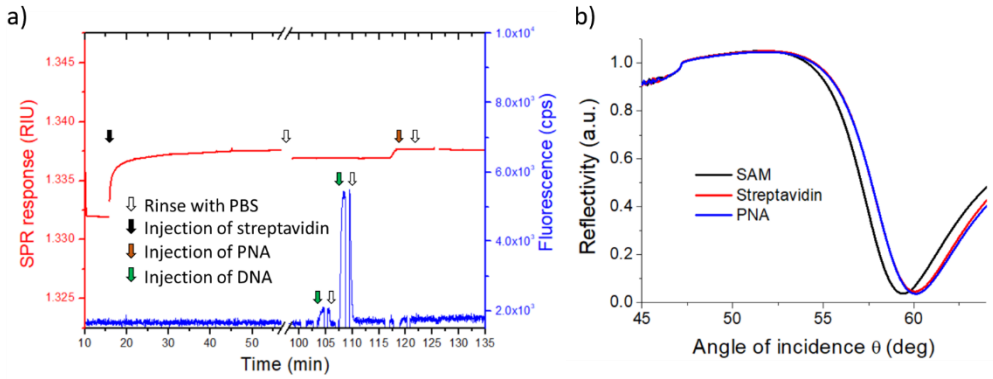
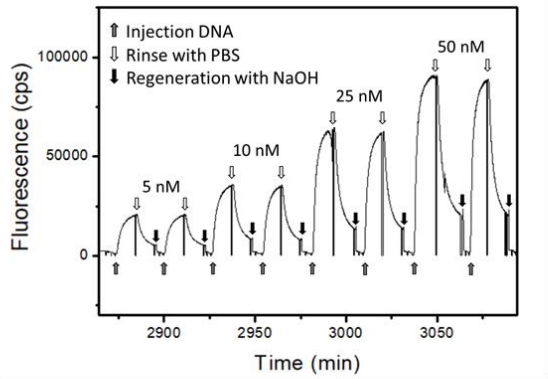


Figure 3.14. a) SPFS measurements for the immobilization process of both streptavidin and PNA, and to check the unspecific binding of fluorescently labelled DNA target to the sensor surface; b) Angular scan curves of the reflectivity R for the self-assembled monolayers (SAM) and after the incubation of streptavidin and PNA.

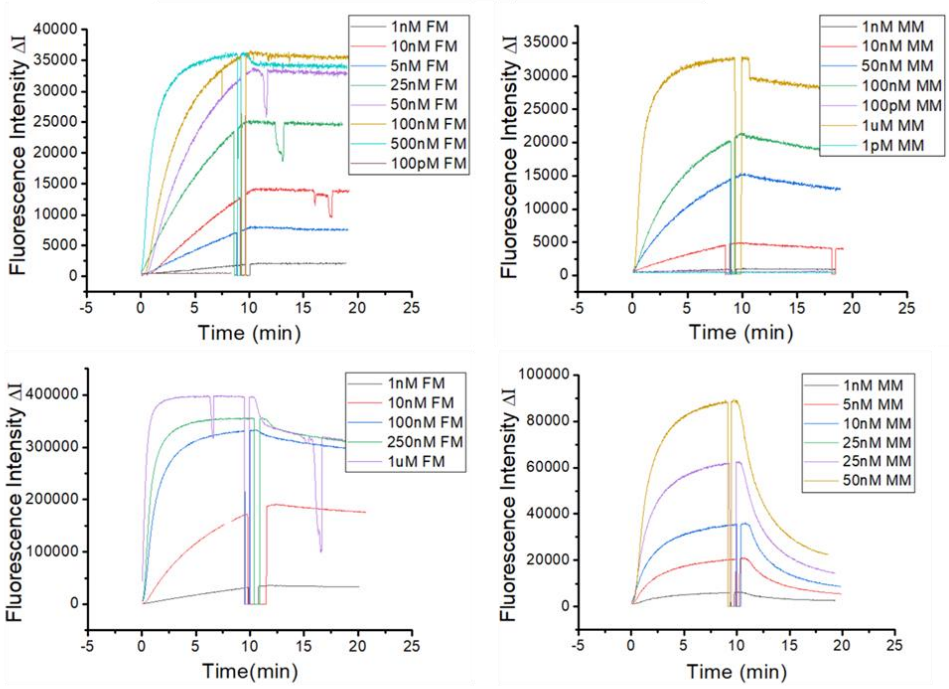
The hybridization between PNA probe and its target DNA was monitored at a fixed incident angle of 56.6°, that is generally chosen as a position in the linear slope of the measured scan curve (e.g. 30% reflectivity) recorded with time. The association of Cy-5 labeled DNA (FM or MM) target with the PNA (chiral or unmodified) functionalized sensor surface was observed at different concentration of the target DNA (from 1 pM up to 1μM). Each DNA target solution was injected into the flow cell for 10 minutes and it was afterwards replaced with fresh phosphate buffer solution to monitor the dissociation (10 minutes) from the surface (**Figure 3.15 b**). To allow multiple analysis cycles at the same sensor surface, the latter was regenerated by treatment with 10 mM NaOH. In this way, the remaining bound target DNA was readily removed (**Figure 3.15 a**). In SPFS experiments, the fluorescence intensity carries kinetic information with respect to the hybridization event between two species and can be generally analyzed in terms of kinetic constants for association rate (k_a), dissociation rate (k_d) and the equilibrium affinity constant (K_D). The evaluation of association (k_a) and dissociation (k_d) rate constants was performed by fitting to the binding data using the Langmuir model and global fitting method implemented in the tool Scrubber (from BioLogic Software Pty Ltd, Australia). The K_D (calculated as k_d divided by k_a) values obtained in this preliminary experiment suggest that the performances of the CB probe in terms of mismatch discrimination for target DNA result to be increased compared to unmodified probes (**Figure 3.15 c**), confirming the results previously obtained in solution. In fact, the K_a values indicate that the duplex

formation is slightly faster when the unmodified PNA **8** is compared to the chiral PNA **6** using a full complementary DNA target (7.80×10^4 M^{-1} s^{-1} vs 1.30×10^5 M^{-1} s^{-1} respectively), although a net difference was not observed when a FM and MM DNA sequences are considered (see **Figure 3.15 c**). As a result of the destabilizing effects generated by the three adjacent modified monomers, the dissociation process for the chiral PNA showed to be faster than its achiral homologous (1.23×10^{-4} s^{-1} vs 2.95×10^{-5} s^{-1} for PNA **6** and PNA **8** respectively). However, unlike the association process, in this case both PNAs displayed a different behavior when a mismatched DNA is used resulting in K_d constants of 3.67×10^{-3} s^{-1} for CB PNA and 1.49×10^{-4} s^{-1} for the unmodified one. From the evaluation of the affinity constants K_D the increased performance of the 'Chiral Box' PNA for the discrimination of single point mutations resulted more evident. In fact, a comparison between the ratio in the equilibrium affinity constant K_D(FM)/K_D(MM) for the four hybridization events has showed that the discrimination ability of the Lys-based PNA **6** was observed to be 2.3-fold higher than that afforded by the achiral PNA **8**, confirming the superior sensitivity for single-mismatched DNA sequences of these chiral PNA probes.

a)



b)



c)

PNA Probe	Target DNA	K_a (M^{-1} s^{-1})	K_d (s^{-1})	K_D (10^{-9} M)
PNA **6**	FM	1.30×10^{5}	1.23×10^{-4}	9.46×10^{-1}
PNA **6**	MM	1.67×10^{5}	3.67×10^{-3}	2.20×10
PNA **8**	FM	7.80×10^{4}	2.95×10^{-5}	3.78×10^{-1}
PNA **8**	MM	3.91×10^{4}	1.49×10^{-4}	3.81

*Figure 3.15. Hybridization kinetics for 'Chiral Box' (CB) PNA **6** and mismatch DNA. The sensor surface was regenerated with 10 mM NaOH for the next experiment (a); Kinetic curves recorded by SPFS for unmodified PNA **8** (top) and for CB PNA **6** (bottom) at different concentration of full-match (FM) and mismatch (MM) DNA (b); Comparison between K_a, K_d and K_D values for unmodified and CB PNAs with FM and MM target DNA (c).*

3.4 Conclusion

The novel submonomeric strategy developed in our laboratory and based on the use of minimally-protected building blocks allowed to synthesize 2D-'Chiral Box' PNAs in a simpler and faster way respect to the traditional submonomeric approach. The study of the processes occurring during 'minimally protected submonomeric synthesis were further investigated in this thesis. By evaluation of the change in melting temperature between the different duplexes, these PNAs showed increased performances in terms of mismatch discrimination for cognate DNA sequences compared to unmodified PNAs. However, a detailed description of the kinetics and thermodynamic behavior of the probes once immobilized on a sensor surface could be obtained by performing SPFS experiments. The preliminary test performed with D-Lys based CB PNA has demonstrated that also in this case the chiral probe was able to better discriminate a mismatched DNA sequence compared to the unmodified PNA probe, and that this effect is mainly due to faster dissociation of the mismatched DNA from the CB-PNA sensing surface. Thus, this approach is likely to produce better results in ultrasensitive sensing of the G12D KRAS point mutation in the so-called 'liquid biopsy' format. The same approach could be extended to other mutated sequences for which sub-optimal discrimination from the wild type is observed when using unmodified PNA probes.

3.5 Experimental section

3.5.1 General

The Fmoc-PNA(Bhoc)-OH monomers for PNA synthesis were purchased from LGC link, Inc., while the DNA oligomers for hybridization studies were purchase from Biomers.net and used without further purification. Thymine-1-acetic acid (CMT), Biotin and all the other commercially available reagents and solvents were bought from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe and used without further purification. Adenine (Boc)-1-acetic acid (CMA) and Cytosine (Boc)-1-acetic acid (CMC) were already available in the laboratory. DMF was dried over 4 Å molecular sieves and purged with nitrogen to avoid the presence of dimethylamine. The biotinylated thiol (HS-

C11-EG3-Biotin) and a spacer thiol (HS-C11-EG3-OCH2-COOH) were purchased from Prochimia Surfaces and used without further purification.

3.5.2 Instrumental

Reversed-phase HPLC purifications were performed on an Agilent 1200 Series LC system with a semi-preparative Jupiter column (C18, 5 mm, 300 Å, 10 x 250 mm) by Phenomenex. UPLC-ESI-MS was carried out by using a Waters Acquity Ultra Performance LC with Waters Acquity SQ Detector and with ESI interface, and equipped with a Waters Acquity UPLC BEH 300 column (50x2.1 mm, 1.7 µm, C18) (Gradient for PNAs: 0.90 minutes in H_2O + 0.2% FA, then linear gradient to 50% MeCN + 0.2% FA in 5.70 minutes at a flow rate of 0.25 mL/min). UV-Vis measurements were performed by a Thermo Scientific Evolution 260 BIO spectrophotometer. SPFS measurements were performed with a home-made optical system described below (**Figure 3.17**).

3.5.3 Synthesis of PNA probes

All the PNAs were synthetized on Rink Amide-ChemMatrix resin loaded with Fmoc-Gly-OH (0.15 mmol/g) in 5 µmol scale, by using the following solutions: Deprotection: 20% piperidine in DMF; Capping: acetic anhydride/DIPEA/dry DMF 5/6/89;DIPEA wash: 5% DIPEA in DMF; Cleavage: TFA/*m*-cresol 9/1.

3.5.3.1 Chiral Box PNAs

'Chiral Box' PNAs (**PNA 6 and 7**) were synthetized manually following standard Fmoc/Bhoc synthesis for the attachment of the first unmodified monomers. The synthesis of the chiral stretch was instead performed following a submonomeric approach: a) Fmoc deprotection 2 x 8 min; b) DCM wash; c) DMF wash; d) Kaiser test (1 min, positive);[19] e) coupling of the minimally protected submonomer (**1** or **2**) 2 x 30min (activation for 2 min; activation solution: 5 eq of submonomer, 4.9 eq of HBTU as activator, 10 eq of DIPEA in dry DMF); f) DMF wash; g) coupling of the carboxymethylnucleobase (CMA, CMT or CMC) 2 x 2 h (activation for 15min; Activation solution: 7 eq of carboxymethylnucleobase, 7 eq of DIC, 7 eq of

DhBtOH and 1 eq of DIPEA in dry DMF); h) DMF wash; i) DCM wash. These steps were repeated for all the three C2-modified monomers and then the PNA sequences were completed by using standard Fmoc/Bhoc synthetic strategies.
Aliquots of the resin were separated during intermediate steps of the 'chiral box' part in order to monitor the progression of the synthesis. These underwent the cleavage procedure, followed by precipitation of the PNA and UPLC-MS analysis.
A final treatment of the dried resins with the cleavage solution (2 x 1 h) resulted in the release of the oligomers and in the removal of the protecting groups from the nucleobases and the side-chains on C(2)-position.

The crude samples were precipitated with 50 volumes of Et_2O at -20°C and purified by reversed-phase HPLC. The purity of the resulting chiral PNAs was checked by UPLC-MS analysis, while their concentration was extracted from the UV absorbance of diluted solutions at λ =260 nm, assuming an additive contribution of all nucleobases to the total ε (260 nm) of the oligomers. The molar extinction coefficients were calculated by considering the following values for each nucleobase: T = 8600 $M^{-1}cm^{-1}$, C = 6600 $M^{-1}cm^{-1}$, A = 13700 $M^{-1}cm^{-1}$, G = 11700 $M^{-1}cm^{-1}$.

3.5.3.2 Unmodified PNA

The unmodified **PNA 8** was obtained by automatic synthesis on a Biotage Syro I system according to the standard manufacturer protocols, consisting in the following steps: a) Deprotection 2 x 8 min; b) DCM wash; c) DMF Wash; d) Coupling 2 x 1 h (3 eq of monomer, 3eq of HBTU, 6 eq of DIPEA in dry DMF); e) DMF Wash; f) Capping 2 x 2 min; g) DMF Wash. A periodic control of the synthesis progress was performed by UPLC-MS analysis.
Unmodified PNA was removed from the resin, purified, characterized and quantified according to the general procedures reported above for the 'Chiral Box' PNAs.

PNA 6: Biotin-OO-CGC $CA_{D\text{-}Lys}T_{D\text{-}Lys}C_{D\text{-}Lys}$AG CT-Gly-$NH_2$, ε (260 nm): 101000 cm^{-1} M^{-1}, 5.8 % yield. [M] Calcd 3725.5 Da.
UPLC-MS (ESI): r.t. = 2.86 min, m/z found (calculated); 1242.6 (1243.0) $[M + 3H]^{3+}$, 932.2 (932.5)$[M + 4H]^{4+}$, 746.1 (746.2) $[M + 5H]^{5+}$, 621.9 (622.0)$[M + 6H]^{6+}$, 533.1 (533.3)$[M + 7H]^{7+}$.

PNA 7: Biotin-OO-CGC $CA_{D\text{-}Arg}T_{D\text{-}Arg}C_{D\text{-}Arg}AG$ CT-Gly-NH_2, ε (260 nm): 101000 cm^{-1} M^{-1}, 3.2 % yield. [M] Calcd 3810.0 Da.
UPLC-MS (ESI): r.t. = 2.86 min, m/z: found (calculated): 1271.0 (1271.0)$[M + 3H]^{3+}$, 953.8 (953.5) $[M + 4H]^{4+}$, 762.8 (763.0)$[M + 5H]^{5+}$, 636.1 (636.0) $[M + 6H]^{6+}$, 545.4 (545.3) $[M + 7H]^{7+}$.
PNA 8: Biotin-OO-CGC CAT CAG CT-Gly-NH_2, ε (260 nm): 101000 cm^{-1} M^{-1}, 4.1 % yield. [M] Calcd 3512.5 Da.
UPLC-MS (ESI): r.t. = 3.17 min, m/z: found (calculated) 1172.3 (1171.9)$[M + 3H]^{3+}$, 879.2 (879.1)$[M + 4H]^{4+}$, 704.1 (703.5)$[M + 5H]^{5+}$, 587.9 (586.4)$[M + 6H]^{6+}$, 503.1 (502.8)$[M + 7H]^{7+}$.

3.5.4 Preparation of sensor matrix on gold substrate

The glass substrates were carefully cleaned the by following procedure:

- sonication for 10 min each in 2% alkaline detergent solution (Helmanex, Helma, Germany; sonification apparatus Super RK510, Sonorex, Germany),
- sonication for 15 min in MilliQ water (18 MΩ / cm),
- 15 min sonification in ethanol,
- cleaned samples were blown dry in a stream of nitrogen.

Gold (99.9999%, Balzers) was deposited onto clean LaSFN9 slides by thermal
evaporation at a deposition rate of 0.1 nm/s under UHV conditions (p = 10^{-6} -10^{-7} mbar) in an evaporation apparatus (Edwards). To improve the adhesion of the gold film to the glass substrate a chromium film of approximately 2 nm was evaporated if necessary.
The gold surface was incubated overnight in a mixed 1:4 thiol solution of a biotinylated thiol (HS-C11-EG3-Biotin) and a spacer thiol (HS-C11-EG3-OCH2-COOH) (**Figure 3.13**) in absolute ethanol for obtaining the self-assembled monolayers (SAM). The streptavidin solution (1.86 μM) was injected into the flow cell to allow for binding to the self-assembled thiol layer at a flow rate of 30 μL/sec. Subsequently, biotinylated PNAs (1 μM) were immobilized on the streptavidin layer as catcher probes via the streptavidin/biotin binding.

3.5.5 SPFS set-up

An optical home-made system combining SPR and SPFS was used for investigation of the PNA/DNA interaction. The detection was performed in the Kretschmann configuration, which is based on a high refractive index prism, 50 nm thermally evaporated gold glass slide, optically matched to the prism through matching oil. A HeNe laser beam (5 mW, λ = 633 nm) passed through a chopper that was connected to a lock-in amplifier (EG&G) to allow for reduced noise and daylight independent measurements of the reflected intensity. The intensity and polarization of the incident light were adjusted by two polarizers (Glan-Thompson). A laser beam at wavelength of 633 nm was coupled to the prism and allowed to resonantly excite surface plasmons at the outer interface of thin gold film with an intensity that was controlled by the angle of incidence. A flow cell (**Figure 3.16**) with a volume of 4 µl was clamped against the gold sensor surface to flow liquid samples with a flow rate of 30 µl/min. The reflected beam intensity was measured with a lock-in amplifier to track changes in the SPR signal. The enhanced electromagnetic field intensity generated by surface plasmons produced the fluorescence signal measured on the SPFS modality. The emitted fluorescence light (at 670 nm) from the sensor surface was collected through the flow cell by a lens and detected by a photomultiplier (H6240-01, Hamamatsu, Japan) that was connected to a counter. The fluorescence light emitted at a wavelength of about 670 nm was spectrally separated from the excitation light at 633 nm by using a set of laser notch filter and fluorescence band pass filter (**Figure 3.17**).

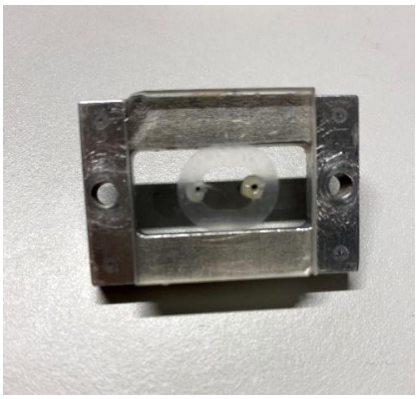


Figure 3.16. Flow cell for surface plasmon-enhanced fluorescence spectroscopy experiments.

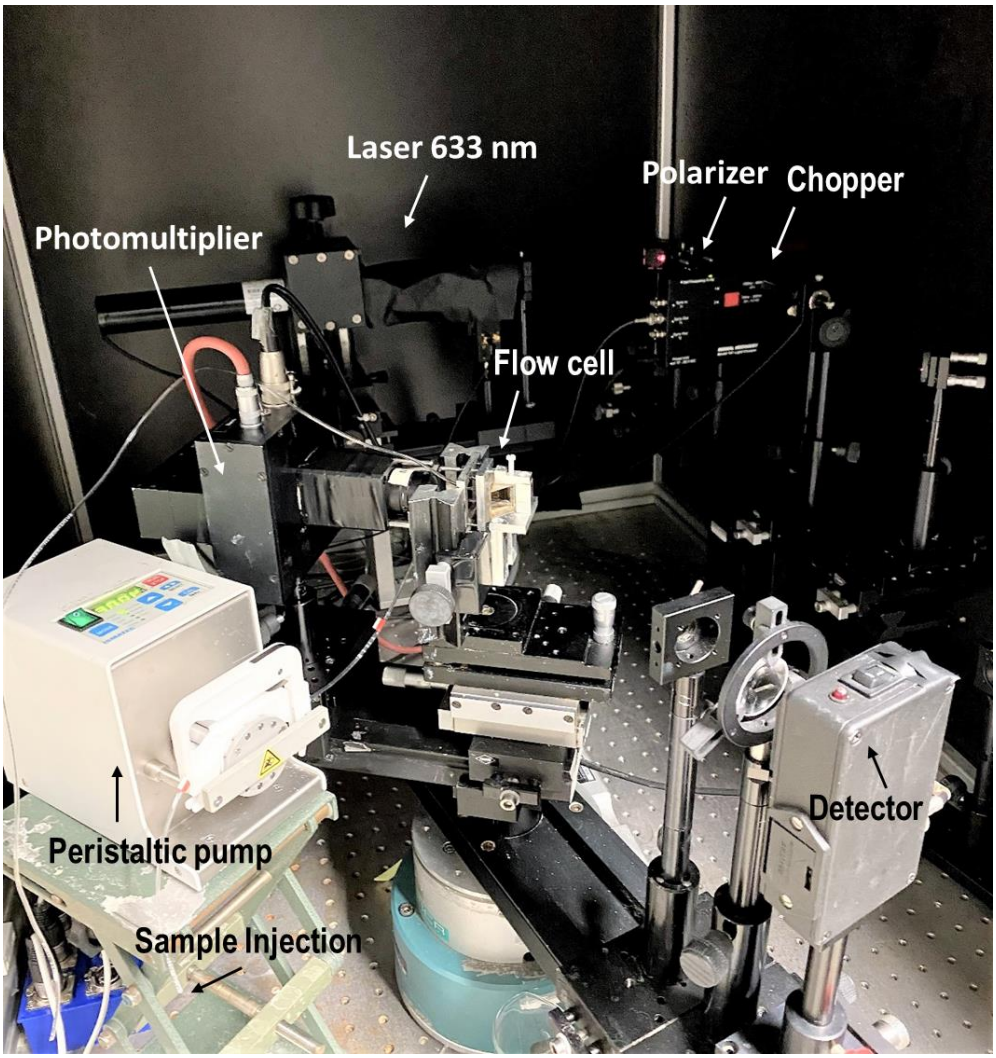


Figure 3.17. Surface plasmon-enhanced fluorescence spectroscopy set-up used for hybridization studies.

3.6 References

1. Dong, B., Nie, K., Shi, H., Chao, L., Ma, M., Gao, F., Liang, B., Chen, W., Long, M. & Liu, Z. Film-Spotting chiral miniPEG-γPNA array for BRCA1 gene mutation detection. *Biosens. Bioelectron.* **136,** 1–7 (2019).

2. Zhang, N. & Appella, D. H. Colorimetric detection of anthrax DNA with a peptide nucleic acid sandwich-hybridization assay. *J. Am. Chem. Soc.* **129,** 8424–8425 (2007).

3. Cunningham, D., Atkin, W., Lenz, H. J., Lynch, H. T., Minsky, B., Nordlinger, B. & Starling, N. Colorectal cancer. *Lancet* **375,** 1030–1047 (2010).

4. Itonaga, M., Matsuzaki, I., Warigaya, K., Tamura, T., Shimizu, Y., Fujimoto, M., Kojima, F., Ichinose, M. & Murata, S. ichi. Novel methodology for rapid detection of KRAS mutation using PNA-LNA mediated loop-mediated isothermal amplification. *PLoS One* **11,** 1–12 (2016).

5. Oh, J. E., Lim, H. S., An, C. H., Jeong, E. G., Han, J. Y., Lee, S. H. & Yoo, N. J. Detection of Low-Level KRAS Mutations Using PNA-Mediated Asymmetric PCR Clamping and Melting Curve Analysis with Unlabeled Probes. *J. Mol. Diagnostics* **12,** 418–424 (2010).

6. Tatsumi, K., Mitani, Y., Watanabe, J., Takakura, H., Hoshi, K., Kawai, Y., Kikuchi, T., Kogo, Y., Oguchi-Katayama, A., Tomaru, Y., Kanamori, H., Baba, M., Ishidao, T., Usui, K., Itoh, M., Cizdziel, P. E., Lezhava, A., Ueda, M., Ichikawa, Y., Endo, I., Togo, S., Shimada, H. & Hayashizaki, Y. Rapid Screening Assay for KRAS Mutations by the Modified Smart Amplification Process. *J. Mol. Diagnostics* **10,** 520–526 (2008).

7. Han, J.-Y., Choi, J.-J., Kim, J. Y., Han, Y. L. & Lee, G. K. PNA clamping-assisted fluorescence melting curve analysis for detecting EGFR and KRAS mutations in the circulating tumor DNA of patients with advanced non-small cell lung cancer. *BMC Cancer* **16,** 627 (2016).

8. Mitra, R. & Ganesh, K. N. Aminomethylene Peptide Nucleic Acid (am-PNA): Synthesis, Regio-/Stereospecific DNA Binding, And Differential Cell Uptake of (α/γ,R/S)am-PNA Analogues. *J. Org. Chem.* **77,** 5696–5704 (2012).

9. Sforza, S., Tedeschi, T., Corradini, R. & Marchelli, R. Induction of helical handedness and DNA binding properties of peptide nucleic acids (PNAs) with two stereogenic centres. *European J. Org. Chem.* 5879–5885 (2007). doi:10.1002/ejoc.200700644

10. Corradini, R., Sforza, S., Dossena, A., Palla, G., Rocchi, R., Filira, F., Nastri, F. & Marchelli, R. Epimerization of peptide nucleic acids analogs during solid-phase synthesis: optimization of the coupling conditions for increasing the optical purity. *J. Chem. Soc., Perkin Trans. 1* 2690–2696 (2001). doi:10.1039/B104146K

11. Mullard, A. Cracking KRAS. *Nat. Rev. Drug Discov.* **18,** 887–891 (2019).

12. Volpi, S., Rozzi, A., Rivi, N., Neri, M., Knoll, W. & Corradini, R. Submonomeric Strategy with Minimal Protection for the Synthesis of C(2)-Modified Peptide Nucleic Acids. *Org. Lett.* **23,** 902–907 (2021).

13. Sforza, S., Tedeschi, T., Corradini, R., Ciavardelli, D., Dossena, A. & Marchelli, R. Fast, solid-phase synthesis of chiral peptide nucleic acids with a high optical purity by a submonomeric strategy. *European J. Org. Chem.* 1056–1063 (2003). doi:10.1002/ejoc.200390148

14. Tedeschi, T., Sforza, S., Maffei, F., Corradini, R. & Marchelli, R. A Fmoc-based submonomeric strategy for the solid phase synthesis of optically pure chiral PNAs. *Tetrahedron Lett.* **49,** 4958–4961 (2008).

15. Liebermann, T. & Knoll, W. Surface-plasmon field-enhanced fluorescence spectroscopy. *Colloids Surfaces A Physicochem. Eng. Asp.* **171,** 115–130 (2000).

16. Park, H., Germini, A., Sforza, S., Corradini, R., Marchelli, R. & Knoll, W. Kinetic and affinity analyses of hybridization reactions between peptide nucleic acid probes and DNA targets using surface plasmon field-enhanced fluorescence spectroscopy. *Biointerphases* **1,** 113–122 (2006).

17. Yao, D., Yu, F., Kim, J., Scholz, J., Nielsen, P. E., Sinner, E.-K. & Knoll, W. Surface plasmon field-enhanced fluorescence spectroscopy in PCR product analysis by peptide nucleic acid probes. *Nucleic Acids Res.* **32,** e177–e177 (2004).

18. Knoll, W., Park, H., Sinner, E.-K., Yao, D. & Yu, F. Supramolecular interfacial architectures for optical biosensing with surface plasmons. *Surf. Sci.* **570,** 30–42 (2004).

19. Kaiser, E., Colescott, R. L., Bossinger, C. D. & Cook, P. I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **34,** 595–598 (1970).

Chapter 4

4 Synthesis of new dimeric modified nucleobase

4.1 Introduction

The synthesis of artificial analogs of the four natural nucleobases has been widely investigated in the past decades, leading to a large number of applications in different research fields. For example, the modification of canonical nucleobases, or their replacement with synthetic ones, can be used as an accurate strategy for tuning the interactions of oligonucleotides with the corresponding molecular targets or for allowing their post-synthetic conjugation, labelling or imaging with appropriate (bio)molecules.[1–7] Also, the use of non-canonical nucleobase has been exploited in synthetic biology to afford nucleic acids with extended functions by the expansion of the genetic code.[8–10] In medicinal chemistry, several nucleobase derivatives have been tested as candidates in the treatment of different diseases and some of them are currently used as antiviral, antibacterial or antitumor agents.[11–14] Recently, an increasing number of applications is also emerging from the field of nanotechnology, where the self-assembly of nucleobases or their analogs is effectively used for the creation of different architectures, including nanomaterials or supramolecular polymers.[15–17]

The binding of the major groove of DNA- or RNA duplexes is related to several of these topics, being reported as a strategy for the modulation of gene expression, the detection and/or isolation of important biological targets, and the functionalization of DNA nanostructures.[18–20] One of the most relevant mode of interaction with the major groove consists in the formation of triple helical structures by an additional oligonucleotide able to install Hoogsteen base pairing with its double-stranded counterpart. Naturally occurring triplexes normally include polypyrimidine oligomers bound in parallel orientation to a polypurine stretch on the target double helix leading to the formation of C^+-GC or T-AT triplets (**Figure 4.1**), despite also purine-rich strands interacting in antiparallel orientation can give rise to less stable A-AT or G-GC complexes.[13,21]

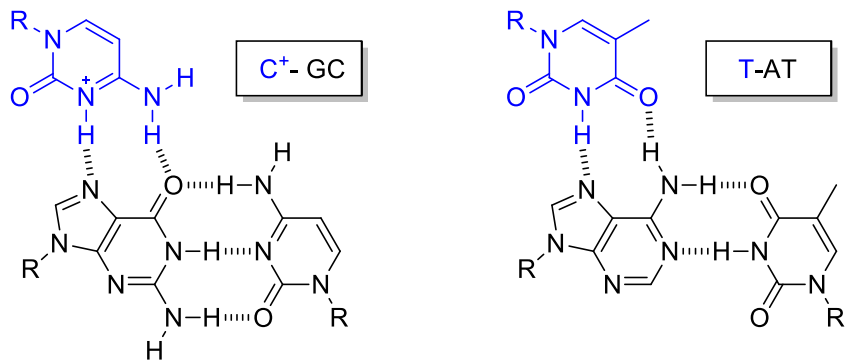


Figure 4.1. Structures of C⁺-GC and T-AT triplexes exploiting both Watson-Crick and Hoogsteen base-pairing.

These systems suffer of several limitations, such as the low binding affinity and the slow kinetics of association at physiological pH, and the significant loss of stability when purine-pyrimidine inversions are present in the double-stranded substrate. For solving these issues, many efforts have been dedicated to the development of the so-called Triplex Forming Oligonucleotides (TFO), that is DNA (or RNA) analogs presenting modifications at a nucleobase and backbone level able to promote the formation of stable triple helical adducts with a given nucleic acid. Of course the role of the nucleobase is crucial in this view and several strategies to obtain artificial mimics have been proposed, leading in some cases to the creation of oligonucleotides able to bind mixed sequences with high affinity.[7,13,14,21–23] However, the production of effective TFO remains often a challenging process which requires long and specific synthetic procedures depending on the features of the target duplex.

A simpler approach to target the major groove of a double stranded nucleic acid could consist in involving the nucleobases of one of the two strands in a triplex-like motif with their cognate one, thus avoiding the necessity of assembly two different strands for binding. This purpose can be achieved by extending the hydrogen bond pairing of the given nucleobase(s) to the Hoogsteen domain of its counterpart, eventually providing additional stacking and hydrophobic interactions by widening its aromatic core. The combination of these principles has led in some case to a remarkable increase of both affinity and selectivity of oligonucleotides including the resulting artificial nucleobase for their complementary sequences. For example, the replacement of cytosine units with its analogs [24–27] in nucleic acid duplexes (**Figure 4.2 a**) has been reported to enhance the stability of the system with excellent sequence selectivity. For both nucleobases the improved binding performances arose from an additional

hydrogen bond afforded by an amino or guanidino pendant protruding to the Hoogsteen domain of target guanine and from the stacking ability of the phenoxazine or pyrrole rings fused with the original pyrimidine skeleton, respectively. The most significant results were obtained with compound **b** **Figure 4.2**, otherwise known as G-clamp, with an enhancement of melting temperature in the resulting duplex by up to 20°C when included in appropriate oligonucleotide mimics.[25,26]

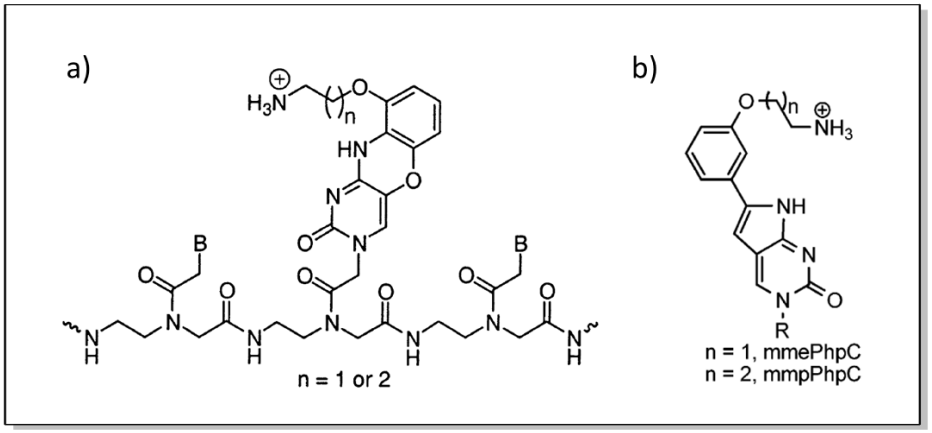


Figure 4.2. Chemical structures of cytosine analogs. a) PNA monomer containing the tricyclic cytosine analogues phenoxazine, 9-(2-aminoethoxy)phenoxazine (G-clamp, n=1), and 9-(3-aminopropoxy)phenoxazine (propyl-G-clamp, n=2). Reprinted with permission from ref. 25: Rajeev et al. © 2002, American Chemical Society. b) meta-substituted [mono-m-(aminoethoxy)phenyl]pyrrolocytosine (mmePhpC) and [mono-m-(aminopropoxy)phenyl]pyrrolocytosine (mmpPhpC). Reprinted with permission from ref. 27: Wojciechowski & Hudson © 2009, American Chemical Society.

A similar strategy could be exploited for targeting adenine, but interestingly the synthesis of thymine (or uracil) analogs with increased affinity for a biological target has been mainly exploited for the creation of TFOs rather than for mimicking the triplex binding mode by modification of double stranded systems.[7,13,14,21–23] Since the methyl group of thymine points to the major groove of the resulting nucleic acid duplex when it is involved in Watson-Crick base pairs, it appears as a convenient point for performing modifications aimed to extend its hydrogen bonding pattern. Moreover, the functionalization of the C-5 position of thymine has been explored for a wide range of applications and protocols for the conjugation with different types of (bio)molecules are known in literature.[1,2,6,7]

An extended work was carried out by Prof. Julius Rebek and his group regarding the study and the development of model receptors for adenine derivatives. In

fact, inspired by the naturally occurring base-pairing process and the stacking interactions existing between adjacent base pairs, a new model system was introduced in which these two forces can act simultaneously from perpendicular directions for binding adenine derivatives (**Figure 4.3**).[28–31]

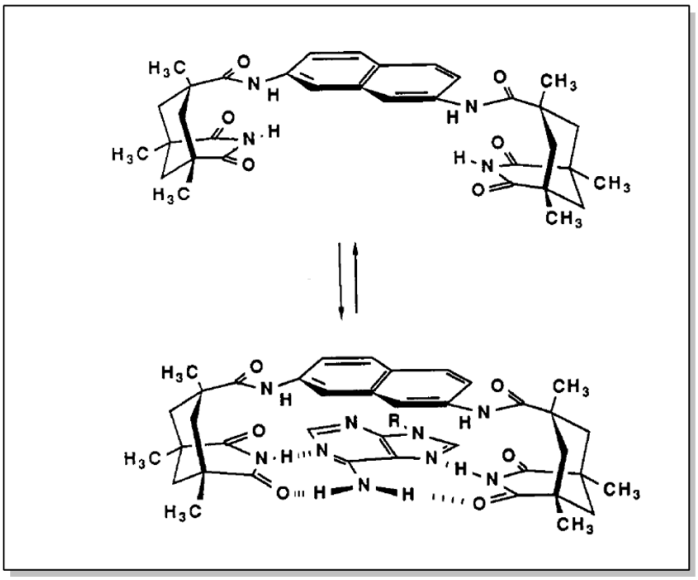


Figure 4.3. New model receptor proposed by Rebek et al. binding adenine derivatives. Reprinted with permission from Ref. 31 © 1987, American Chemical Society.

Several NMR titrations studies and selective extraction experiments confirmed the great potential of these so-called "Rebek imide-type" receptors for selective recognition of adenine derivatives.[32] However, these systems would be hardly inserted into a nucleotide analog, and therefore are limited to monomeric adenine derivative binding.

Starting from existing PNAs triplex structures has led our group to develop a model which could be suited for incorporation into a PNA strand. The analysis of the structure of a PNA-DNA-PNA complex present in the Protein Data Bank (1 PNN) indicates that the distance between the methyl groups of the thymine units composing their T-AT triplets is nearly conserved and stands around 9.5 Å. A naphthalene-based spacer (**Figure 4.4**) was suggested to be a good candidate for mimicking this binding motif, since it can in principle arrange to two pyrimidinic rings at a similar distance and avoid the collapse of the system by formation of intramolecular hydrogen bonds thanks to its rigidity.

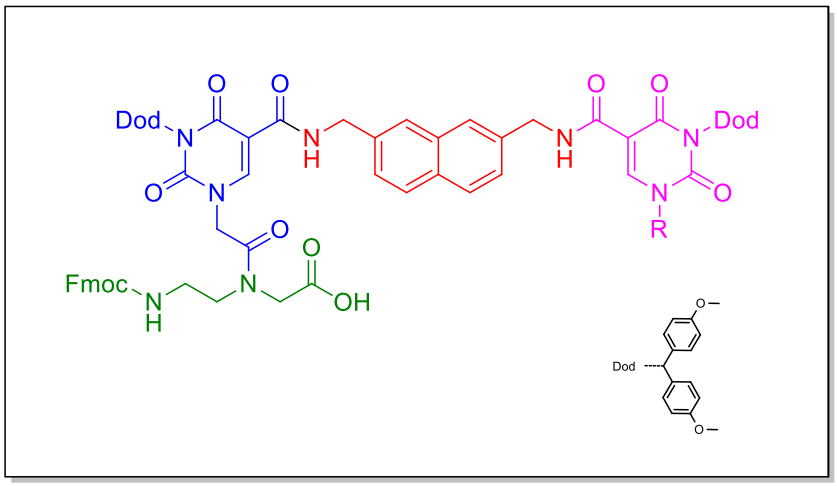


Figure 4.4. Structure of PNA monomer where two uracil moieties are linked by 2,7-bis(aminomethyl)-naphathalene spacer.

Accordingly, our research group designed a series of dimeric nucleobases consisting in two thymine units connected on the C-5 with different spacers by amide bonds,[33] in the frame of a project focused on the creation of PNA with increased affinity for biologically relevant targets. Homopyrimidine PNAs are known to perform strand invasion on double stranded DNAs with complementary sequences by forming very stable PNA-DNA-PNA triplexes[34–36] and the tested dimeric compounds were suggested to mimic the formation of the T-AT triplets found in these complexes. The dimeric compounds were found to induce erythroid differentiation and cytotoxicity on K562 leukemic cells when a 2,7-bis(aminomethyl)-naphathalene unit was used as spacer (**Figure 4.4**), with an increase of activity respect to a monofunctional control compound which indicates cooperation between the two pyrimidine units.[33]

The system was supposed to target proper regions of the genome of the tested cell-line by binding adenine moieties with both Hoogsteen and Watson-Crick base pairing, but once inserted into a PNA oligomer lead only to a moderate increase of affinity for a complementary DNA sequence.[37]

On the other hand, more recent studies on molecular modelling pointed out that the presence of a residual flexibility on its amidomethyl arms could disadvantage the correct orientation of the two uracil units once in contact with a target adenine residue, and for this reason two alternative structures bearing more rigid α, β-unsaturated or triazole pendants were proposed (Verona, M.D., PhD Thesis). The simulation performed on the model system using tested molecular dynamics (MD) methods[38] indicated also that alternative structures

containing the triazole ring would be destabilizing, causing a disruption of the duplex. Finally, a promising model was offered by the system bearing a planar α, β-unsaturated carbonyl unit depicted in **Figure 4.5**, for which a stable triplex-like structure was observed using a 200 ns MD simulation.

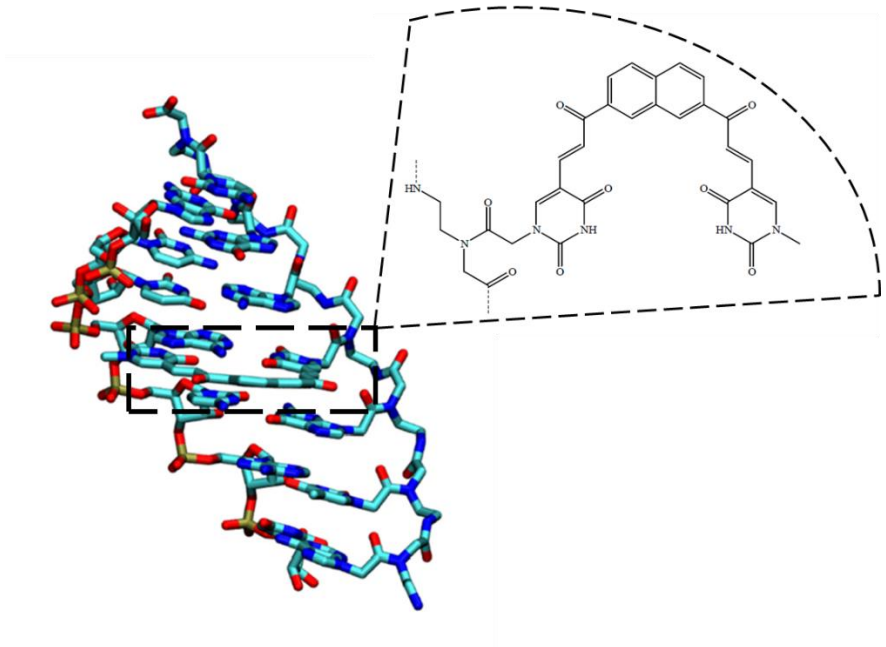


Figure 4.5. Structure of the system containing the modified dimer selected for targeting adedine. [Verona, M.D., PhD Thesis]

This model, while preserving a certain mobility, showed to possess a rather stable duplex in which the second uracil was able to constantly interact with the target adenine, therefore it was selected as a promising candidate to our aim. However, in this work only modelling of the nucleobase was performed, but synthetic attempts to obtain the model systems were not at that time successful. The synthetic challenge was successively undertaken by myself in collaboration with the post-doctoral fellow Dr. Saša Korom, in the frame of the ULTRAPLACAD project aiming at obtaining very specific probes for detection of adenine in point mutations (such as that of G12D KRAS, see chapter 3). The present chapter reports the first encouraging data obtained in the synthesis and properties of this model.

4.2 Aim of the study

In this work a synthetic strategy for the production of the new dimeric thymine analog designed in the previous PhD thesis work is reported. The main strategy is based on Wittig- Horner reaction to ensure the presence of a *trans*-double bond, and thus allowing the two nucleobases to be connected with a perfectly planar system to improve their binding ability. A symmetric dimer (**11**) was produced as a model to test the ability to form both Watson-Crick and Hoogsteen hydrogen bonds, and was found to be able to perform synergistic recognition adenine derivatives, with induced chemical shift changes in the NMR spectrum in the order of 6 ppm. In order to be able to link this modified nucleobase on a PNA monomer, the synthesis of a precursor of an asymmetric dimer (**27**) with one alkylated uracil on one side and one bearing a carboxymethyl linker for linking to the PNA backbone was also carried out.

4.3 Results and discussion

4.3.1 Retrosynthetic analysis

First, we planned to obtain a symmetric version of the dimeric model reported in **Figure 4.6**, in order to test the overall strategy, then the desymmetrization of the system to obtain a suitable PNA monomer was planned. A crucial point was that, according to the model, the alkene formed should be obtained with *E*-geometry (*trans*). Thus, a simple Wittig reaction or aldol condensation would require careful control of reaction conditions to obtain mixtures of *Z*- and *E*-stereoisomers with preferential *E*-isomer content, but long and tiresome separation procedures would have been necessary to obtain pure *E*-derivatives. Thus, the Wittig–Horner variant was chosen, since this allows to obtain exclusively *E*-geometry with high yields. For this a bis-α-ketophosphonate reacting with the appropriate uracil-derived aldehyde was needed. The correct geometry for the aromatic part of the linker required a 2,7-substitute naphthalene derivative with bromoacetyl substituents that could then be transformed into phosphonate by Arbuzov reaction. This could be obtained by bromination of the corresponding bis-acetylated naphthalene derivative. The low cost 2,7-dihydroxynaphthalene was therefore suggested as good starting material, taking onto account the easy transformation of OH residues into

triflate leaving groups. The two-carbon residue needed for the acetyl part could be introduced using acetylene and Sonogashira coupling, followed by regiospecific hydration.

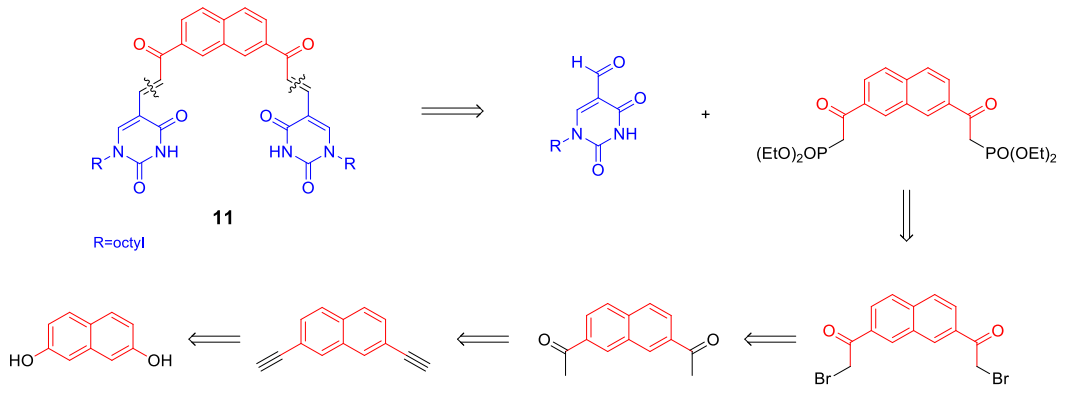


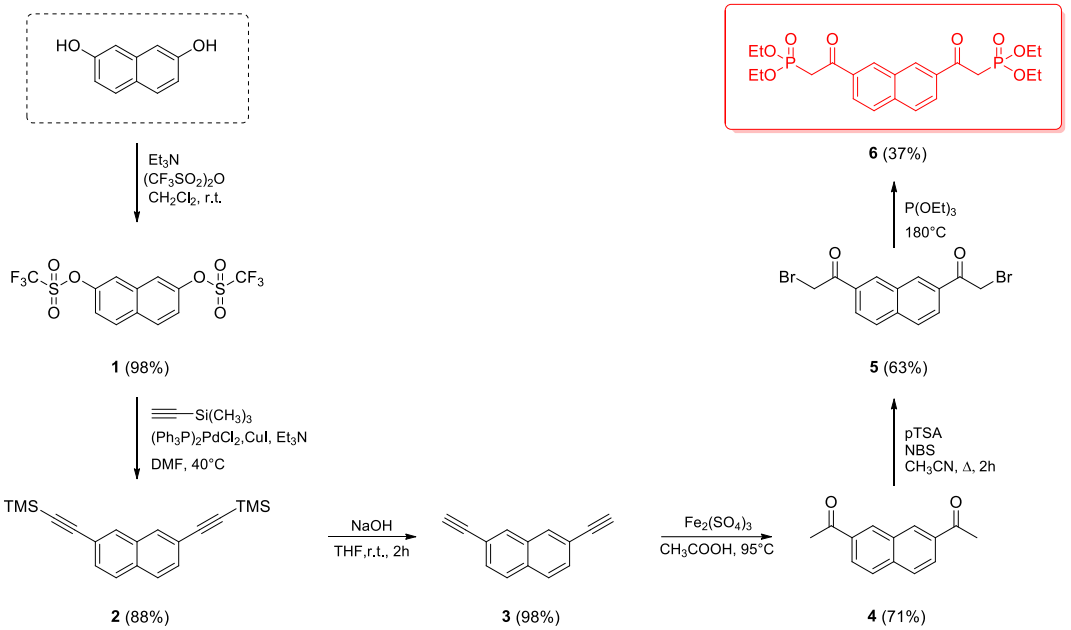
Figure 4.6. Structure of the final symmetric model system (11) and retrosyntesis.

4.3.2 Synthesis of the symmetric model system

The symmetrical dimeric model system (**Figure 4.6, compound 11**) was constructed following the synthetic strategy mainly based on Wittig-Horner reaction described in the previous paragraph.

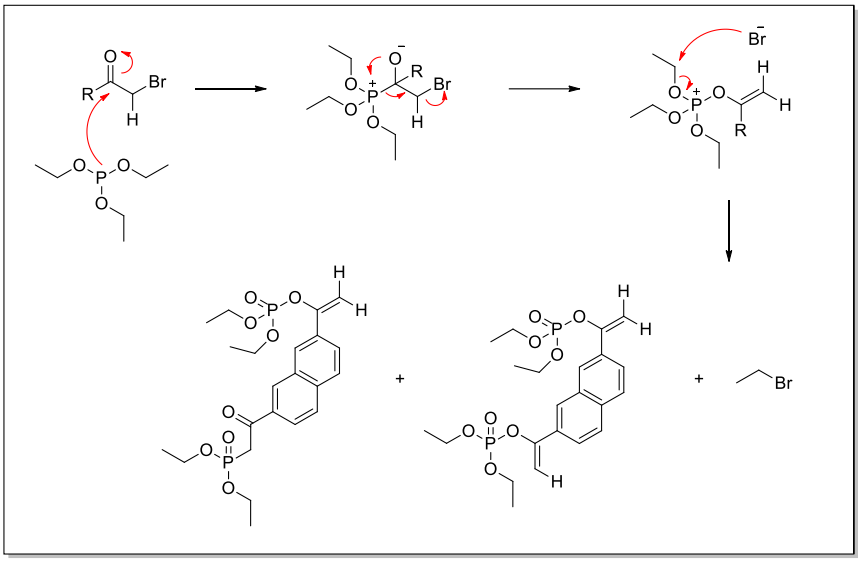
The central naphthalene linker (in red) was synthesized starting from 2,7-dihydroxynaphthalene, a non-expensive commercial compound, that was converted into a triflate derivative using $(CF_3SO_2)_2O$ (**Scheme 4.1**). Product **2** was obtained by a Sonogashira coupling between compound **1** and trimethylsilylacetylene in the presence of $(Ph_3P)_2PdCl_2$ and CuI. A following cleavage of silyl groups under basic conditions led to product **3**. Hydration of the aromatic terminal alkynes was performed in acetic acid, a Brønsted acid source, using a catalytic amount of iron (III) sulfate that generated regioselectively the inner enol, which spontaneously rearranged to the corresponding methyl ketone **4** (by keto-enol rearrangement) with good yields[39]. Finally, product **5** was obtained by a selective alpha-halogenation of **4** using N-bromosuccinimide (NBS). As mentioned above, the transformation needed for obtaining the *E* carbon-carbon double bond using the Wittig-Horner-type reaction was the introduction of an alkyl phosphonate group in the alpha-position to the carbonyl groups of the linker. Hence, the crucial step of

this synthetic pathway was represented by the conversion of **5** into naphthalene 2,7-bis-phosphonate **6** through the so-called Arbuzov reaction.



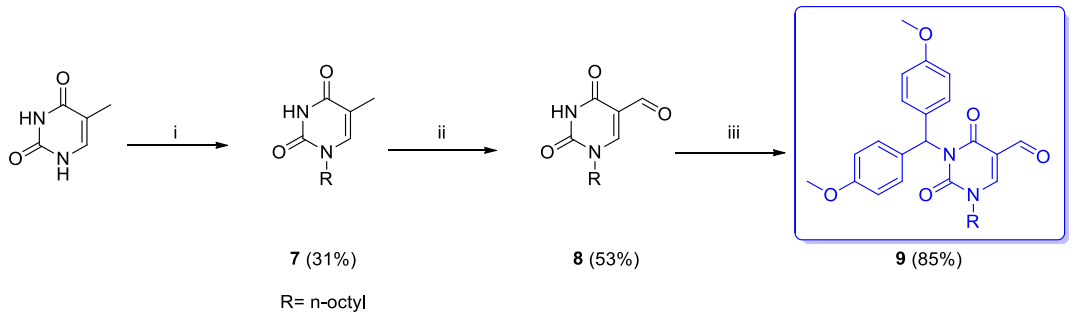
Scheme 4.1. Synthetic pathway for the synthesis of tetraethyl(naphthalene-2,7-diylbis(2-oxoethane-2,1-diyl)) bis(phosphonate) (in red).

This procedure yielded compound **6** with yields up to 37%, probably due to the Perkow side-reaction that usually occurs with α-halogen ketones. In fact, triethylphosphite used in Arbuzov conditions, can also perform a nucleophilic addition to the carbonyl carbon of the aryl halide leading to the formation of a vinyl diethyphosphate (**Scheme 4.3**). In order to optimize the synthetic route to the phosphonate compound, other strategies (e.g. phosphonate synthesis using ester groups reacted with diethyl methylphosphonate as reported in the literature) can therefore be explored.[40,41]



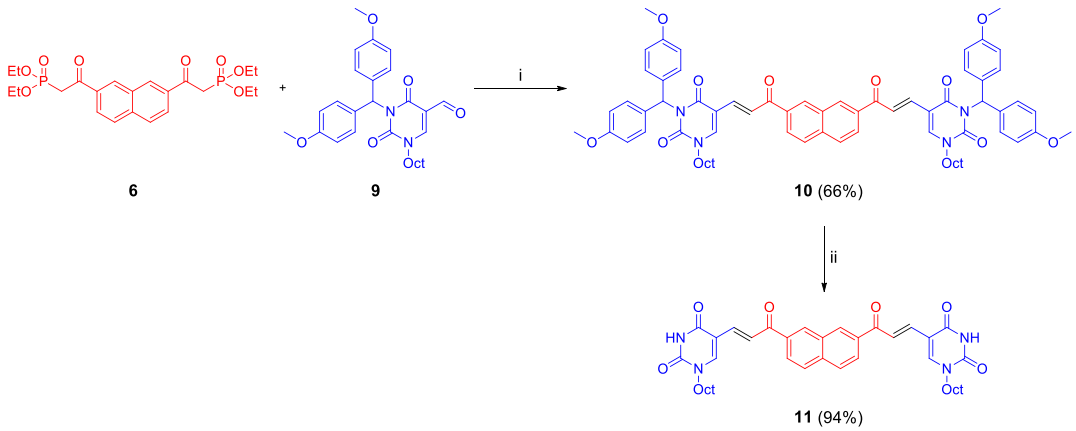
Scheme 4.2. Mechanism of the Perkow reaction.

The next step for obtaining the symmetric model system involved the synthesis of the uracil derivative functionalized on carbon-5 with a formyl group. Product **7** was obtained by a selective alkylation of thymine using 1-bromooctane, although about the same percentage of dialkylated product was recovered. Methyl group of compound **7** was converted into an aldehyde by $K_2S_2O_8$ leading to **8**,[42] and then Dod protective group was installed in N^3-position (**Scheme 4.4**).



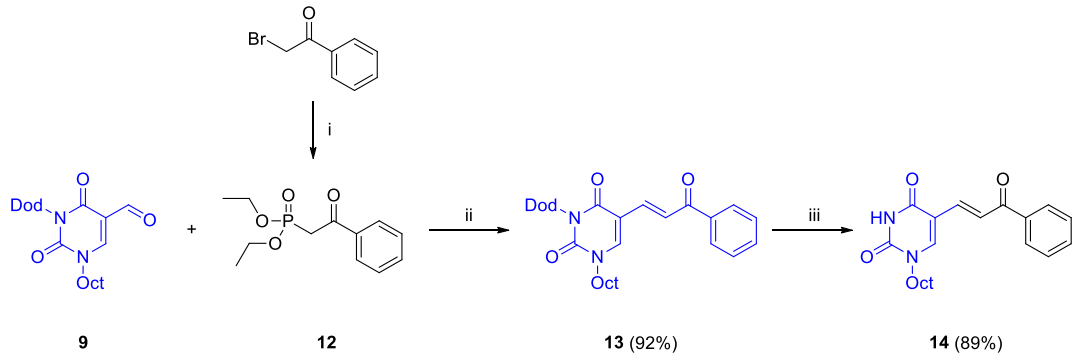
Scheme 4.3. Synthesis of 1-octyl-3-Dod-5-formyl uracil. (i) 1-Bromooctane, NaH, dry DMF, 5 h, 80°C; (ii) $K_2S_2O_8$, 2,6- lutidine, $CuSO_4$, MeCN, 1 h, 80°C; (iii) Dod-Cl, K_2CO_3, dry DMF.

Synthon **9** was then used in a Wittig-Horner reaction with the bis-phosphonate **6** in 2:1 ratio to obtain **10** that was then treated with TFA in DCM to cleave Dod protection (**Scheme 4.5**). The formation of the desired *E* geometry was confirmed by the J coupling constant between alkene hydrogens found for compound **11**, i.e. 15.2 Hz (typical range 12-18 Hz).



Scheme 4.4. Synthesis of the symmetrical dimeric model system. (i) 1 M NaOH, THF, 2 h, r.t.; (ii) TFA, 2 h, r.t.

Subsequently, a monomeric uracil derivative was synthesized as a control system, using an equivalent strategy starting from compound **9** and bromoacetophenone in order to perform comparative binding studies with adenine derivatives using NMR spectroscopy (**Scheme 4.6**).



*Scheme 4.5. Synthesis of model monomer. Compound **12** was obtained by Arbuzov reaction starting from bromoacetophenone. i) $P(EtO)_3$, 180°C, overnight (60%); ii) 1 M NaOH, THF, 2 h, r.t.; iii) TFA, 45 min, r.t.*

4.3.3 NMR binding studies with 9-ethyladenine

Both model molecules (**11** and **14**) were used for *in vitro* studies of binding interactions with 9-ethyl adenine. Presence of octyl group in N^1 position of model dimer **11** and model monomer **14** and ethyl group in N^9 position in adenine ensured an adequate solubility in non-polar solvent such as $CDCl_3$, in which the hydrogen bonding ability can easily be evidenced. Titration of the

dimer model (1.84 μM) with 9-ethyl adenine (**Figure 4.7**) showed significant shift of NH group in uracil moiety involved in hydrogen bonding with the adenine derivative, shifting signal from 8.32 to 14.37 ppm upon addition of 1 equivalent of adenine. Smaller extent of chemical shift changes was seen for vinylic protons as well vinylic protons from uracil moiety and singlet signal from naphthalene bridge suggesting geometrical changes, compared to free dimer model molecule, compatible with a rearrangement of the conformation to give rise to a cleft-like shape. Moreover, the dimeric system showed a stronger binding and gave rise to a larger chemical shift change than the corresponding monomer (**Figure 4.67**, Appendix). Additional NOE experiments were performed in order to verify proposed geometry of the formed 1:1 dimer model/adenine complex. Selective 1D NOE experiments (data not shown) confirmed the predicted cleft-like geometry of dimer model, with cross-signals in agreement with the formation of both Watson-Crick and Hoogsteen type of interactions. These findings confirmed that dimension and geometry of this bridging unit is appropriate for adenine binding.

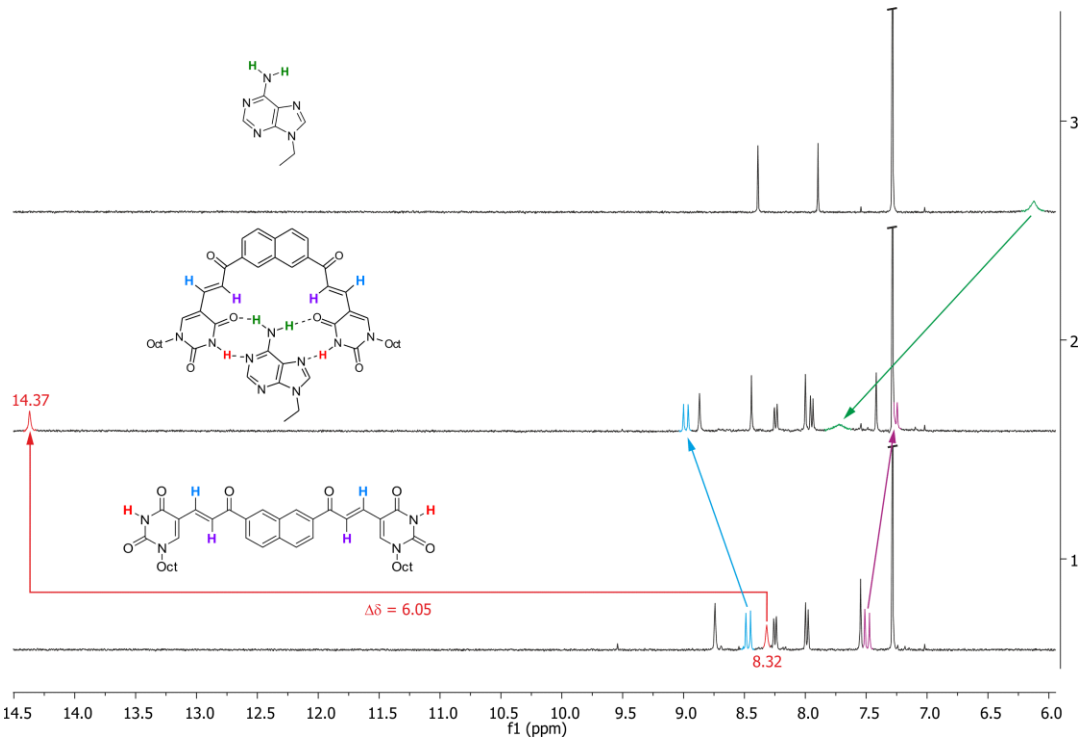
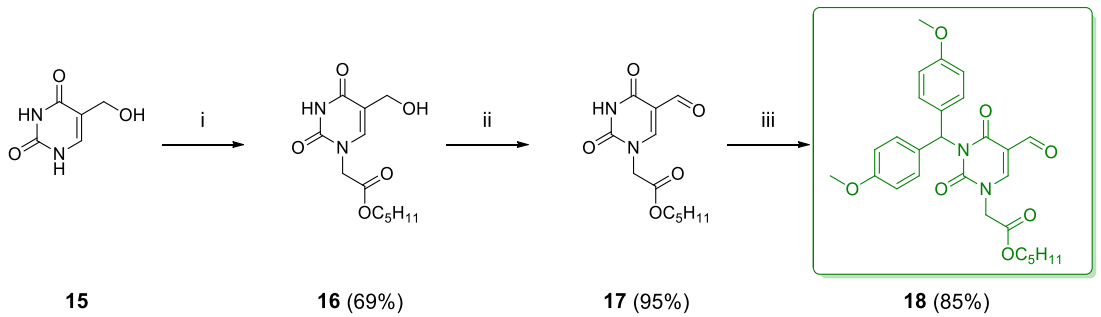


Figure 4.7. ^1H NMR spectra of 9-ethyladenine (1.84 μM), dimeric base (1.84 μM) and 1:1 mixture of 9-ethyladenine and dimeric base in $CDCl_3$

4.3.4 Synthesis of the asymmetric dimer

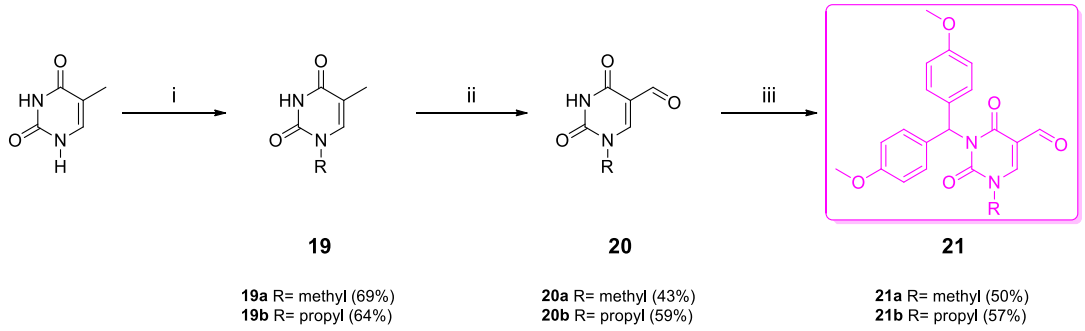
Based on success of making the dimeric model molecule **11** by exploiting Wittig-Horner reaction to produce main structural framework, the same strategy was used to obtain an asymmetric derivatization of the naphthalene bridge in order to allow only one uracil to be coupled to the PNA backbone.
The central *bis*-phosphonate compound **6** was obtained following the protocol described in **Scheme 4.1**. Next, the synthesis of two different uracil synthons were planned. For the preparation of the first modified uracil derivative compound **15** was used as starting material, that was previously obtained following a literature protocol involving an hydroxymethylation of uracil with paraformaldehyde.[43] The regioselective alkylation of the N^1 position for the introduction of the carboxymethylene linker was achieved by pentyl bromoacetate to obtain **16**, which was converted into the aldehydic derivative **17** by oxidation with Dess-Martin periodinane (DMP), an oxidant reagent widely used for the mild conditions required, the short reaction times, high yields and the tolerance of sensitive functional group. Finally, compound **17** was protected in position N^3 with Dod chloride to obtain product **18** (**Scheme 4.7**).



Scheme 4.6. Synthesis of pentyl (3-Dod-5-formyl uracil-1-yl) acetate. (i) pentyl bromoacetate, K_2CO_3, dry DMF, r.t. (ii) DMP, DCM, r.t. 3h. (iii) Dod-Cl, K_2CO_3, dry DMF.

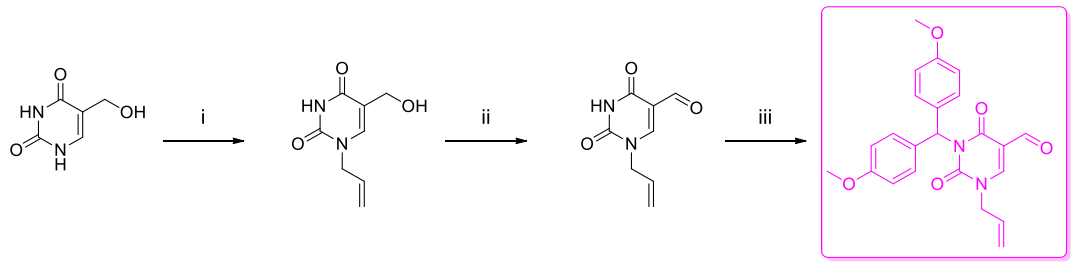
For the synthesis of the second uracil derivative, three different functionalizations of the N^1 position were achieved. First, a methyl group was introduced because of its small size, which allows to minimize the steric hindrance that could affect the base-pairing once obtained the PNA monomer. Then, a propyl group was proposed in order to improve the solubility of the uracil derivative with respect to the methyl-functionalized one. Finally, with the aim to introduce further modifications, an allyl function was exploited.

The alkylation of thymine with iodomethane or 1-iodopropane in the presence of hexamethyldisilazane (HMDS) and trimethylsilyl chloride (TMS-Cl) allowed to obtain, respectively, compounds **19a** and **19b** (**Scheme 4.8**). These derivatives were oxidized by following the protocol described in the Scheme 2.4 for the oxidation of 1-octyl thymine, yielding products **20a** and **20b**, which after the functionalization on N-3 position with the Dod protective group, gave **21a** and **21b** derivatives.



Scheme 4.7. Synthesis of 3-Dod-1-alkyl-5-formyl uracil derivatives. (i) Methyl iodide or propyl iodide, HMDS, TMS-Cl, Δ, 6h; (ii) $K_2S_2O_8$, 2,6- lutidine, $CuSO_4$, MeCN, 1 h, 80°C; (iii) Dod-Cl, K_2CO_3, dry DMF.

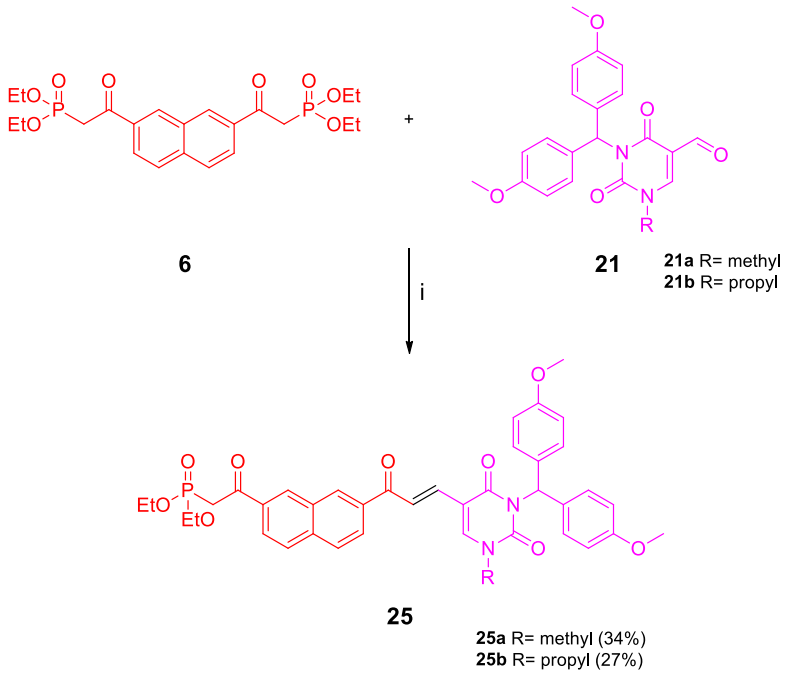
The allyl group was introduced following a slightly different synthetic pathway. In fact, in this case similarly to the synthesis of the uracil derivative **18**, the procedure started from 5-hydroxymethyl uracil **15** (**Scheme 4.9**). Product **22** was obtained by alkylation of N-1 position with allyl bromide in the presence of hexamethyldisilazane (HMDS) and trimethylsilyl chloride (TMS-Cl). Next, oxidation of **22** with DMP in dichloromethane yielded the product **23** that was readily protected at N^3 position with the Dod group, giving the uracil derivative **24**.



Scheme 4.8. Synthesis of 3-Dod-1-allyl-5-formyl uracil. (i) Allyl bromide, HMDS, TMS-Cl, Δ, 6h; (ii) DMP, DCM, r.t. 3h. (iii) Dod-Cl, K_2CO_3, dry DMF.

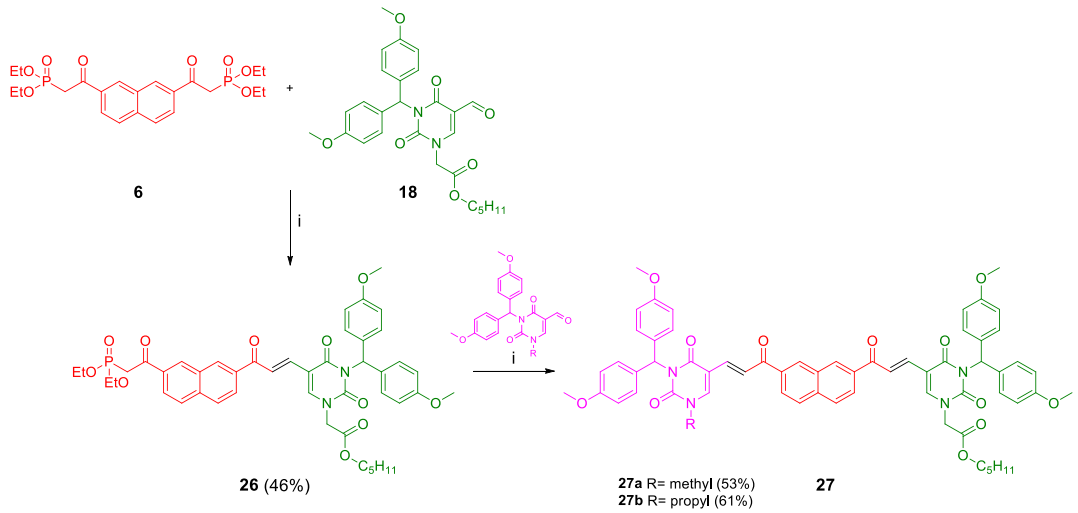
Once obtained the various uracil synthons, a Wittig-Horner reaction was performed in order to couple them for the preparation of the modified PNA monomer. The desymmetrization of compound **6** was found to be relatively easy, since mono-derivatives could be obtained simply by careful choice of stoichiometry of reagents.

The *bis*-phosphonate **6** was first coupled to the methyl uracil derivative **21a** in the presence of K_2CO_3 and then the same Wittig-Horner coupling was performed using the propyl uracil derivative **21b**, yielding the products **25a** and **25b**, respectively (**Scheme 4.10**). In both cases a certain amount of the symmetric product was recovered, ensuing from the coupling of both phosphonate moieties of **6** with the respective aldehyde. Specifically, 10% in the case of the methyl derivative and 17% for the propyl one. Using NMR it was possible to confirm the presence of desired products thanks to the appearance of the coupling between the two *trans* alkene hydrogens (Figure **4.54**, **4.56** and **4.57**, **4.59**, Appendix).



*Scheme 4.9. Synthesis of the monosubstituted intermediate **25** by Wittig-Horner reaction. (i) K_2CO_3, THF, r.t., overnight.*

The same synthetic strategy was used to connect the *bis*-phosphonate **6** to the uracil derivative bearing the carboxymethylene linker on the N^1 position (**18**). The Wittig-Horner coupling yielded in this case the monosubstituted product **26** with higher yield (49%) compared to the previous couplings, with 29% of the symmetric disubstituted product recovered. A following coupling, using the same reaction conditions, was performed on the monosubstituted intermediate **26** with the two different aldehydes **21a** and **21b** in order to obtain the precursors of the asymmetric dimer **27** (**Scheme 4.11**).
In this way it will be possible to obtain precursors with different reactive groups on N^1 position of one uracil and finally, after hydrolysis of the ester moiety of the second uracil derivative, to couple these asymmetric synthons to the PNA backbone.



*Scheme 4.10. Synthesis of the asymmetric precursor **27**. (i) K_2CO_3, THF, r.t. overnight.*

4.4 Conclusions

This part of the work represents a step forward towards the synthesis of rationally designed nucleobases for adenine recognition in PNA. Following previous theoretical predictions, the developed synthetic strategy has been proved to give the possibility to obtain the designed modified nucleobases. Despite the poor solubility of thymine in organic solvents it has been possible to obtain several uracil derivatives with different alkyl groups at N-1 position

and protected with Dod group at N-3 position suitable for preliminary studies and for the synthesis of asymmetric nucleobase derivatives.
It has been demonstrated that the model dimeric nucleobase **11** confirmed the predictions previously obtained by Molecular Dynamics simulations, and therefore to possess the appropriate geometry for the selective recognition of adenine derivatives, at least in simple 1:1 complexes formed in organic solvents. Synthesis of important precursors of modified PNA monomers has been achieved by exploiting a Wittig-Horner reaction-based strategy. Indeed, the use of the latter allowed to obtain dimeric modified nucleobase with the intended α, β-unsaturated geometry.
The aforementioned precursors will be used in the host laboratory with the aim to obtain the final modified PNA monomers. It should be emphasized that the complexation of the dimer model in apolar solvent maximizes the effect of hydrogen bonding and, therefore, the subsequent binding ability in water can be affected by other important effect, primarily solubility, solvent competition for hydrogen bonding groups, and stacking interactions, which can only be verified once the modified nucleobase will be inserted into a PNA strand. However, the tight binding observed for the dimer model in chloroform is a sign that the molecular design perfectly fits with the 'A-clamp' model that guided this research.

4.5 Experimental section

4.5.1 General

All solvents and chemical reagents were purchased from Sigma-Aldrich, VWR International, Alfa Aesar, TCI Europe, Novabiochem, Zentek, Strem in the highest purity available. DMF was dried over 4 Å molecular sieves and purged with nitrogen to avoid the presence of dimethylamine. THF was dried by distillation over sodium and stored over 3 Å molecular sieves. TLCs were run on Merck 5554 silica 60 aluminium sheets. Column chromatography was performed as flash chromatography on Merck 9385 silica 60 (0.040- 0.063 mm).

4.5.2 Instrumental

NMR spectra were recorded on a Bruker Advance 300 and 400 spectrometers operating at room temperature. Chemical shifts are reported in parts per

million (ppm) relative to the residual solvent peak. Multiplicities are reported as singlet (s), doublet (d), doublet of doublets (dd), triplet (t) or multiplet (m). Coupling constants were expressed in hertz (Hz). UPLC-ESI-Q data were collected on a Waters Acquity UPLC system equipped with a Waters Acquity UPLC BEH C18 column (2.1x50 mm, 1.7 μm) at 35°C. A flow rate of 0.25 mL/min was used with the following solvent systems: (A): 0.2% FA in H_2O and (B): 0.2% FA in MeCN (FA = formic acid). The column was flushed for 0.9 min with solvent A, then a gradient from 0 to 50% B in 5.7 min was used.

4.5.3 Experimental protocols

Naphthalene-2,7-diyl bis(trifluoromethanesulfonate) (1). Naphthalene-2,7-diol (1g, 6.24 mmol) was suspended in methylene chloride (40 mL), followed by addition of triethylamine (2 mL, 14.35 mmol, 2.3 eq) and slow drop-wise addition of trifluoromethanesulfonic anhydride (2.16 mL, 13.1 mmol, 2.1 eq). The solution was left to stir overnight at room temperature, first changing colour into light-red and then into dark-red. The solvent was evacuated, residue dissolved in ethyl acetate, washed with brine and dried over anhydrous sodium sulphate. Product was purified by column chromatography on silica with 10% ethyl acetate in hexane yielding 2.592 g (98%) of white solid product. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 8.03 (d, *J* = 9.1 Hz, 2H, Ar*H-meta*), 7.83 (d, *J* = 2.3 Hz, 2H, Ar*H-orto'*), 7.50 (dd, *J* = 9.1, 2.3 Hz, 2H, Ar*H-orto*). Full characterization is reported in literature.[44]

2,7-bis((trimethylsilyl)ethynyl) naphthalene (2). Compound **1** (2 g, 5.1 mmol), palladium (II) bis-triphenylphosphine chloride (104 mg, 148 μmol) and copper(I) iodide (44 mg, 230 μmol) were put in a vessel and this was evacuated and purged with nitrogen. Dry DMF (20 mL) was added and solution stirred till dissolved, followed by addition of ethynyltrimethylsilane (1.7 mL, 2.95 mmol) and dry triethylamine (3 mL, 5.30 mmol). Solution was stirred at 40°C overnight. The solution was transferred into extraction funnel containing $KHSO_4$ solution and crude product extracted with ethyl acetate. The product was purified by column chromatography on silica using hexane as eluent. Purification yielded 1.437 g (88%) of product. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 7.93 (s, 2H, Ar*H-orto'*), 7.74 (d, *J* = 8.5 Hz, 2H, Ar*H-meta*), 7.52 (dd, *J* = 8.5, 1.0 Hz, 2H, Ar*H-orto*), 0.31 (s, 18H, C*H*$_3$Si). Full characterization is reported in literature.[44]

2,7-diethynylnaphthalene (3). Compound **2** (734 mg, 2.29 mmol) was dissolved in THF (22 mL) followed by addition of 1 M solution of KOH (22 mL). The solution was stirred for 2 h at rt. The product was extracted with EtOAc, dried over Na_2SO_4 and solvent removed, yielding pure product (98%, 395 mg). ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 7.99 (s, 2H, Ar*H-orto'*), 7.79 (d, *J* = 8.5 Hz, 2H, Ar*H-meta*), 7.56 (dd, *J* = 8.5, 1.2 Hz, 2H, Ar*H-orto*), 3.19 (s, 2H, *H*C≡C). Full characterization is reported in literature.[44]

1,1'-(naphthalene-2,7-diyl) diethanone (4). Compound **3** (600 mg, 3.4 mmol) and iron (III) sulfate pentahydrate (180 mg, 450 μmol) were dissolved in a glacial acetic acid (15 mL) and solution heated at 95°C over weekend. Then, the solution was transferred in separatory funnel containing saturated solution of $NaHCO_3$ and EtOAc and carefully shaken with frequent opening to allow release of CO_2. The organic phase was separated, dried over Na_2SO_4, the solvent was removed and the residue was purified by column chromatography on silica (3:7 EtOAc/Hex) yielding 512 mg (71%) of pure product. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 8.59 (s, 2H, Ar*H-orto'*), 8.16 (dd, *J* = 8.7, 1.6 Hz, 2H, Ar*H-orto*), 7.95 (t, *J* = 8.7 Hz, 2H, Ar*H-meta*), 2.73 (s, 6H, C*H*$_3$C=O). Full characterization is reported in literature.[45]

1,1'-(naphthalene-2,7-diyl) bis(2-bromoethanone) (5). To a solution containing carbonyl compound **4** (120 mg, 0.565 mmol) and p-toluenesulfonic acid monohydrate (323 mg, 1.696 mmol) in acetonitrile (40 mL) was slowly added NBS (201 mg, 1.131 mmol). The reaction mixture was stirred for 2 h in reflux. Reaction mixture was cooled down to rt and the solvent was evacuated. Residue was dissolved in methylene chloride (10 mL), washed with water (2 x 5 mL) and dried over anhydrous sodium sulfate. After solvent evacuation, compound was purified by column chromatography on silica with methylene chloride as an eluent, yielding pure product (131 mg, 63 %). ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.68 (s, 2H, Ar*H-orto'*), 8.20 (d, *J* = 8.6 Hz, 2H, Ar*H-orto*), 8.02 (d, *J* = 8.6 Hz, 2H, Ar*H-meta*), 4.60 (s, 4H, C*H*$_2$Br). ^{13}C NMR (101 MHz, $CDCl_3$) *δ* (ppm): 190.84, 138.16, 132.36, 131.65, 128.87, 127.31, 30.41. OrbiTrap-MS (-): calculated for $C_{14}H_{10}Br_2O_2$ - H^+ 368.9120; found 366.8968, 368.8947, 370.8926 (isotopic distribution of the two bromide atoms).

Tetraethyl(naphthalene-2,7-diylbis(2-oxoethane-2,1-diyl))bis(phosphonate) (6). In a small reaction vial compound **5** (200 mg, 0.54 mmol) and triethyl phosphite (3 mL, 17.5 mmol, 32.4 eq) were heated overnight at 180°C. Excess

of reagent was removed by vacuum and the crude material was purified by column chromatography on silica (1:1 => 3:2 acetonitrile / ethyl acetate) yielding 97 mg (37%) of product. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.73 (s, 2H, Ar*H-orto'*), 8.22 (d, *J* = 8.1 Hz, 2H, Ar*H-orto*), 7.97 (d, *J* = 8.6 Hz, 2H, Ar*H-meta*), 4.21-4.14 (m, 8H, OCH_2CH_3), 3.78 (d, *J*= 22.8 Hz, $POCH_2$), 1.30 (t, *J* = 7.1 Hz, 12H, OCH_2CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) *δ* (ppm): 191.43 (d, *J* = 6.5 Hz), 138.00, 134.81 (d, *J* = 1.8 Hz), 132.99, 131.62, 128.39, 127.15, 62.82 (d, *J* = 6.6 Hz), 39.41 (d, *J* = 130.1 Hz), 16.32 (d, *J* = 6.4 Hz). ^{31}P NMR (162 MHz, $CDCl_3$) δ (ppm): 19.58 (s). OrbiTrap-MS (+): calculated for $C_{22}H_{30}O_8P_2$ + H^+ 485.1488; found 485.1486, calculated for $C_{22}H_{30}O_8P_2$ + Na^+ 507.1308; found 507.1302.

1-octyl-thymine (7). Thymine (1 g, 7.92 mmol) and sodium hydride (347 mg, 8.69 mmol) were dissolved in dry DMF (12 mL) and stirred at 80°C for 1 h under nitrogen atmosphere. Then, 1-bromooctane (1.35 mL, 7.9 mmol) was added to the solution and left to stir at 80°C for 4 hr. The reaction mixture was poured into methylene chloride/water and the organic layer was extracted three times with water to remove the DMF. After drying with sodium sulfate and evaporation of the solvent the residue was purified by column chromatography on silica with hexane/ethyl acetate (4:1) as the eluent. Product **7** was isolated in 31% yield (585 mg). ^{1}H NMR (300 MHz, acetone-d^6) δ (ppm): 9.85 (s, 1H, N*H*), 7.42 (s, 1H, C*H* vinylic), 3.72 (t, *J* = 6.5 Hz, 2H, NCH_2), 1.82 (s, 3H, CH_3 thymine), 1.71-1.66 (m, 2H, NCH_2CH_2 chain) 1.35-1.27 (m, 10H, CH_2 chain), 0.88 (d, *J* = 7.0 Hz, 3H, CH_3 chain). Full characterization is reported in literature.[33]

1-octyl-5-formyl uracil (8). 1-octyl thymine **7** (200 mg, 0.839 mmol) and 2,6-lutidine (340 µL, 2.94 mmol, 3.5 eq) in acetonitrile (3.5 mL) were added to a water (3.5 mL) solution of potassium persulfate (454 mg, 1.678 mmol, 2 eq) and copper(II) sulfate (53.6 mg, 0.336 mmol, 0.4 eq). Solution was stirred at 80°C for 1 h while and reaction progress tracked by TLC (4:1 hexanes / ethyl acetate). Half of the solution volume was evacuated and product extracted with ethyl acetate, dried over sodium sulfate and purified by column chromatography (4:1 => 2:3 hexanes / ethyl acetate) yielding 112 mg (53%) of product. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 10.03 (s, 1H, C*H*O), 8.84 (s, 1H, N*H*), 8.10 (s, 1H, C*H* (6) uracil), 3.89 – 3.82 (m, 2H, NCH_2), 1.81 – 1.71 (m, 2H, NCH_2CH_2 chain), 1.38 – 1.26 (m, 10H, CH_2 chain), 0.90 (t, *J* = 6.8 Hz, 3H, CH_3 chain). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 186.05, 161.87, 149.60, 149.18, 110.96, 50.18, 31.67, 29.12, 29.03, 29.01, 26.34, 22.58, 14.04. OrbiTrap-MS (+): calculated for $C_{13}H_{20}N_2O_3$ +

H^+ 253.1552; found 253.1546, calculated for $C_{13}H_{20}N_2O_3$ + Na^+ 275.1372; found 275.1367.

1-octyl-3-Dod-5-formyl uracil (9). 1-octyl-5-formyl uracil (109 mg, 0.459 mmol) and potassium carbonate (76 mg, 0.551 mmol, 1.2 eq) were dissolved in dry DMF (2 mL) and to it added Dod-Cl (133 mg, 0.505 mmol, 1.1 eq) dissolved in dry DMF (2 mL) was added. The reaction was carried out overnight at rt. Then, DMF was evacuated and the residue purified by column chromatography (3:2 hexane / ethyl acetate) yielding 187 mg (85%) of viscous transparent liquid. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 10.05 (s, 1H, C*H*O), 8.06 (s, 1H, C*H*), 7.34 (d, *J* = 8.6 Hz, 4H, C*H* Ar-Dod), 7.32 (s, 1H, Ar-C*H*-Ar Dod), 6.89 (d, *J* = 8.8 Hz, 4H, C*H* Ar-Dod), 3.82 (s, 6H, OCH_3 Dod), 3.81 – 3.79 (m, 2H, NCH_2), 1.74-1.68 (m, 2H, NCH_2CH_2 chain), 1.29 (d, *J* = 8.2 Hz, 10H, CH_2 chain), 0.90 (t, *J* = 6.9 Hz, 3H, CH_3 chain). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 187.06, 162.09, 158.98, 147.08, 129.95, 129.89, 127.74, 113.85, 113.65, 110.42, 55.23, 50.92, 31.64, 29.14, 29.04, 29.00, 26.34, 22.56, 14.01. OrbiTrap-MS (+): calculated for calculated for $C_{56}H_{68}N_4O_{10}$ + Na^+ 979.4828; found 979.4838.

Dod-protected model dimer (10). Tetraethyl (naphthalene-2,7-diylbis(2-oxoethane-2,1-diyl))bis(phosphonate) **6** (50 mg, 0.103 mmol) and 1 M NaOH (1.032 mL, 1.032 mmol, 10 eq) were stirred in a reaction tube and then compound **9** dissolved in THF (1 mL) was added. . The reaction was conducted for 2 h at room temperature during which time the color of the solution changed from pale yellow, over intense yellow to yellow-orange. The reaction was cooled down with ice bath and quenched by addition of water (50 mL). The crude product was extracted with ethyl acetate, the organic phase dried over anhydrous sodium sulfate and the solvent was evacuated. The crude material was purified by column chromatography using 1:4 hexane/ methylene chloride followed by 3:2 hexane/ethyl acetate yielding 85 mg (66%) of pure product. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.75 (s, 2H, C*H* orto' napthalene), 8.46 (d, *J* = 15.2 Hz, 2H, C*H* trans), 8.25 (dd, *J* = 8.6, 1.2 Hz, 2H, C*H* meta napthalene), 7.97 (d, *J* = 8.6 Hz, 2H, C*H* orto napthalene), 7.58 (s, 2H, C*H* (6) uracil), 7.58 – 7.52 (m, 2H, C*H* trans), 7.45 (s, 2H, Ar-C*H*-Ar Dod), 7.36 (d, *J* = 8.7 Hz, 8H, C*H* Ar-Dod), 6.89 (d, *J* = 8.8 Hz, 8H, C*H* Ar-Dod), 3.83 – 3.81 (m, 4H, NCH_2), 3.80 (s, 12H, OCH_3 Dod), 1.72 (t, *J* = 6.4 Hz, 4H, NCH_2CH_2), 1.31 – 1.28 (m, 20H, CH_2 octyl chain), 0.92 (t, *J* = 6.8 Hz, 6H, CH_3 octyl chain). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 190.00, 161.50, 158.84, 146.40, 137.50, 137.29, 136.23, 131.96, 131.57, 130.83, 130.36, 129.91, 128.35, 127.04, 122.17, 113.60, 109.40, 55.21, 53.81, 50.39,

31.69, 29.70, 29.28, 29.21, 29.11, 26.42, 22.60, 14.07. OrbiTrap-MS (+): calculated for $C_{70}H_{76}N_4O_{10}$ + Na^+ 1155.5454; found 1155.5464.

Dimer model (11). To the Dod-protected dimer **10** (260.9 mg, 0.230 mmol) in a small vial, TFA (3 mL) was added and reaction mixture stirred for 2 h at room temperature. Volatiles were removed by vacuum and crude material triturated by acetonitrile separating pale yellow solid product in yield of 147.1 mg (94%) of pure product. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.74 (s, 2H, C*H* orto' napthalene), 8.49 (d, *J* = 15.2 Hz, 2H, C*H* trans), 8.42 (s, 2H, N*H*), 8.25 (d, *J* = 8.6 Hz, 2H, C*H* meta napthalene), 7.99 (d, *J* = 8.6 Hz, 2H, C*H* orto napthalene), 7.55 (s, 2H, C*H* (6) uracil), 7.51 (d, *J* = 15.2 Hz, 2H, C*H* trans), 3.87 (t, *J* = 7.3 Hz, 4H, NC*H*$_2$), 1.81-1.74 (m, 4H, NCH$_2$C*H*$_2$), 1.33 (dd, *J* = 23.5, 12.3 Hz, 20H, C*H*$_2$ octyl chain), 0.92 (t, *J* = 6.8 Hz, 6H, C*H*$_3$ octyl chain). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 189.80, 161.16, 149.17, 147.73, 136.10, 136.06, 131.75, 128.42, 127.00, 122.89, 110.27, 49.59, 31.71, 29.18, 29.09, 26.43, 22.60, 14.06 (2 carbons overlapped, as inferred from HSQC). OrbiTrap-MS (+): calculated for $C_{40}H_{48}N_4O_6^+$ 680.3574; found 680.4797 - calculated for $C_{40}H_{48}N_4O_6$ + Na^+ 703.3472; found 703.3559.

Diethyl (2-oxo-2-phenylethyl) phosphonate (12). 2-bromo-1-phenylethan-1-one (120 mg, 0.603 mmol) was placed in a small reaction vial together with triethyl phosphite (3 mL, 17.50 mmol), the vial was sealed and the reaction conducted at 180°C overnight. Excess of reagent was removed by vacuum and the residue purified by column chromatography (3:1 ethyl acetate/hexane => 100% ethyl acetate) yielding 94 mg (60%) of pale yellow oily product. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.05 (d, *J* = 7.2 Hz, 2H, C*H* orto), 7.63 (t, *J* = 7.4 Hz, 1H, C*H* para), 7.52 (t, *J* = 7.7 Hz, 2H, C*H* meta), 4.19 (m, 4H, OC*H*$_2$CH$_3$), 3.68 (d, 2H, *J* = 22.8, P-C*H*$_2$), 1.32 (t, *J* = 7.1 Hz, 6H, OCH$_2$C*H*$_3$). ^{31}P NMR (162 MHz, $CDCl_3$) δ (ppm): 19.87. ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 192.02 (d, *J* = 6.7 Hz), 136.57 (d, *J* = 2.0 Hz), 133.68, 129.06, 128.62, 62.70 (d, *J* = 6.5 Hz), 39.16 (d, *J* = 130.5 Hz), 16.28 (d, *J* = 6.4 Hz). OrbiTrap-MS (+): calculated for $C_{12}H_{17}O_4P$ + H^+ 257.0943; found 257.0935; calculated for $C_{12}H_{17}O_4P$ + Na^+ 279.0762; found 279.0754.

Dod-protected model monomer (13). Excess of diethyl (2-oxo-2-phenylethyl)phosphonate (62.2 mg, 0.243 mmol, 1.32 eq) was dissolved in 1 M NaOH (2.023 mL, 2.023 mmol, 11 eq) and to it added THF (~1.5 mL) solution of uracil derivative **9** (88 mg, 0.184 mmol). The reaction was conducted at room

temperature for 1 h. The solution was transferred into a separatory funnel containing brine and ethyl acetate and the organic phase was separated and then dried over anhydrous sodium sulfate, the solvent was removed and the crude material was purified by column chromatography on silica using 2:1 hexanes / ethyl acetate yielding 98.6 mg (92%) of pure product. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.32 (d, *J* = 15.3 Hz, 1H, C*H* trans), 8.08 (d, *J* = 7.2 Hz, 2H, C*H* orto), 7.58 (t, *J* = 7.3 Hz, 1H, C*H* para), 7.52 (s, 1H, C*H* (6) uracil), 7.49 (t, *J* = 7.5 Hz, 2H, C*H* meta), 7.47 (d, *J* = 15.3 Hz, 1H, C*H* trans), 7.41 (s, 1H, ArC*HAr* Dod), 7.36 (d, *J* = 8.7 Hz, 4H, ArC*H* Dod), 6.90 (d, *J* = 8.8 Hz, 4H, ArC*H* Dod), 3.82 (s, 6H, OC*H*$_3$ Dod), 3.81 – 3.77 (m, 2H, NC*H*$_2$), 1.71 (d, *J* = 6.4 Hz, 2H, NCH2C*H*$_2$), 1.31 – 1.23 (m, 10H, C*H*$_2$ chain), 0.92 (t, *J* = 6.9 Hz, 3H, C*H*$_3$ chain). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 190.63, 161.43, 158.82, 149.66, 146.19, 138.09, 136.77, 132.81, 130.39, 129.89, 128.65, 128.56, 122.56, 113.58, 109.45, 58.73, 55.24, 50.37, 31.69, 30.94, 29.19, 29.11, 26.42, 22.60, 14.07. OrbiTrap-MS (+): calculated for $C_{36}H_{40}N_2O_5$ + Na^+ 603.2829; found 603.2832.

Model monomer (14). To the Dod protected model monomer **13** (59 mg, 0.102 mmol) TFA (2 mL) was added and solution stirred at room temperature for 45 min. The reaction was quenched with methanol and the reagents were removed under vacuum. The residue was purified by column chromatography (5:1 methylene chloride / ethyl acetate) yielding 32 mg (89%) of product. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.60 (s, 1H, N*H*), 8.35 (d, *J* = 15.3 Hz, 1H, C*H* trans), 8.09 (d, *J* = 7.3 Hz, 2H, C*H orto*), 7.60 (t, *J* = 7.3 Hz, 1H, C*H para*), 7.51 (t, *J* = 5.8 Hz, 2H, C*H meta*), 7.51 (s, 1H, C*H* (6) uracil), 7.44 (d, *J* = 15.3 Hz, 1H, C*H* trans), 3.85 (t, *J* = 7.4 Hz, 2H, NC*H*$_2$), 1.78 – 1.74 (m, 2H, NCH$_2$C*H*$_2$), 1.36 (dd, *J* = 17.9, 7.9 Hz, 10H, C*H*$_2$ chain), 0.92 (t, *J* = 6.8 Hz, 3H, C*H*$_3$ chain). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 190.45, 161.27, 149.29, 147.65, 137.96, 135.70, 132.94, 128.65, 128.62, 123.18, 110.31, 49.54, 31.71, 29.15, 29.10, 29.08, 26.41, 22.59, 14.06 - OrbiTrap-MS (+): calculated for $C_{21}H_{26}N_2O_3$ + H^+ 355.2022; found 355.2013; calculated for $C_{21}H_{26}N_2O_3$ + Na^+ 377.1841; found 377.1832.

Pentyl (5-hydroxymethyluracil-1-yl) acetate (16). Compound **15** (150 mg, 1.05 mmol) and K_2CO_3 (145 mg, 1.05 mmol) were suspended in dry DMF (2 mL). Pentyl bromoacetate (232 mg, 1.1 mmol) was added to the mixture that was stirred overnight at room temperature. K_2CO_3 was removed by filtration and the DMF was removed under vacuum, thus yielding a white solid. Product **14** was obtained by filtration and drying (196 mg, 69%). ^{1}H NMR (400 MHz, DMSO) δ (ppm): 11.38 (s, 1H, N*H*), 7.55 (s, 1H, C*H* (6) uracil), 5.07 (t, *J* = 5.2 Hz, 1H, O*H*),

4.55 (s, 2H, NC*H*$_2$CO), 4.16 (d, *J* = 4.1 Hz, 2H, C*H*$_2$OH), 4.12 (t, *J* = 6.6 Hz, 2H, OC*H*$_2$ pentyl chain), 1.61-1.56 (m, 2H, OCH$_2$C*H*$_2$ pentyl chain), 1.31-1.28 (m, 4H, C*H*$_2$C*H*$_2$CH$_3$ pentyl chain), 0.89 (t, *J* = 6.9 Hz, 3H, C*H*$_3$ pentyl chain). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 168.76, 163.50, 151.14, 142.35,113.98, 65.44, 56.28, 49.32, 28.17, 27.87, 22.17, 14.29. OrbiTrap-MS (+): calculated for $C_{12}H_{18}N_2O_5$ + H^+ 271.1286; found 271.1291.

Pentyl (5-formyl uracil-1-yl) acetate (17). To a solution of **16** (200 mg, 0.74 mmol) in DCM (10 mL) was slowly added DMP (627 mg, 1.48 mmol) under nitrogen atmosphere at 0°C for 30 min. The mixture was stirred at room temperature for further 2h and then was taken up with Et_2O (20 mL) was added and the reaction was quenched with aq. $Na_2S_2O_3$ (10 mL) .The reaction mixture was extracted with ethyl acetate and the organic layer was dried over anhydrous sodium sulfate, the solvent was removed and the crude material was purified by column chromatography on silica using 2:8 hexanes / ethyl acetate. Product **17** was isolated in 95% yield (188 mg) as white solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 10.04 (s, 1H, C*H*O), 8.02 (s, 1H, C*H* (6) uracil), 4.58 (s, 2H, NC*H*$_2$), 4.26 (t, *J*= 6.7 Hz, 2H, OC*H*$_2$ pentyl chain), 1.72-1.70 (m, 2H, OCH$_2$C*H*$_2$ pentyl chain), 1.38-1.35 (m, 4H, C*H*$_2$C*H*$_2$CH$_3$ pentyl chain), 0.96 (t, *J* = 6.7 Hz, 3H, C*H*$_3$ pentyl chain). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 186.82, 168.05, 162.70, 152.25, 150.50, 110.79, 65.76, 49.93, 28.12, 27.86, 22.16, 14.30. OrbiTrap-MS (+): calculated for $C_{12}H_{16}N_2O_5$ + H^+ 269.11320; found 269.11322; calculated for $C_{12}H_{16}N_2O_5$ + Na^+ 291.0951; found 291.0950.

Pentyl (3-Dod-5-formyl uracil-1-yl) acetate (18). Compound **17** (190 mg, 0.708 mmol) and potassium carbonate (196 mg, 1.42 mmol, 2 eq) were dissolved in dry DMF (2 mL) and to it added Dod-Cl (186 mg, 0.708 mmol, 1.1 eq) dissolved in dry DMF (2 mL) was added. The reaction was carried overnight at r.t. Then, DMF was evacuated and the residue purified by column chromatography (7:3 hexane / ethyl acetate) yielding 298 mg (85%) of viscous transparent liquid. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 10.02 (s, 1H, C*H*O), 8.11 (s, 1H, C*H* (6) uracil), 7.33-7.30 (m, 5H, Ar-*ortho* Dod, ArC*H*Ar Dod), 6.88 (m, 4H, Ar-*meta* Dod), 4.54 (s, 2H, NC*H*$_2$), 4.19 (t, *J*= 6.8 Hz, 2H, OC*H*$_2$ pentyl chain), 3.80 (s, 6H, OC*H*$_3$ Dod), 1.66-1.61 (m, 2H, OCH$_2$C*H*$_2$ pentyl chain), 1.34-1.27 (m, 4H, C*H*$_2$C*H*$_2$CH$_3$ pentyl chain), 0.94 (t, *J* = 6.9 Hz, 3H, C*H*$_3$ pentyl chain). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm): 186.89, 166.56, 161.92, 158.98, 149.98, 147.37, 129.90, 129.61, 113.64, 110.97, 66.71, 58.88, 55.25, 50.74, 28.08, 27.83, 22.23, 13.92. OrbiTrap-MS (+):

calculated for $C_{12}H_{16}N_2O_5$ + H^+ 269.11320; found 269.11322; calculated for $C_{27}H_{30}N_2O_7$ + Na^+ 517.1945; found 517.1943.

1-methyl-thymine (19a). Thymine (1g, 7.93 mmol), hexamethyldisilazane (10.4 mL, 49.9 mmol, 6.3 eq) and trimethylsilyl chloride (0.6 mL, 4.76 mmol, 0.6 eq) were added in a round-bottom flask and allowed to stir for 4 hours at 120°C under nitrogen atmosphere. After the reaction, the solution was cooled to 40°C and methyl iodide (7.17 mL, 115 mmol, 16.4 eq) was added to it and the mixture was let to stir overnight. Then, volatiles were removed under vacuum and to the reaction mixture ice water (10 mL) and acetic acid (10 mL) were added. After 20 minutes, the solution was concentrated under reduce pressure thus yielding a white solid. Product **19a** was collected by filtration and drying (69%, 767 mg). ^{1}H NMR (300 MHz, DMSO) δ (ppm): 11.21 (s, 1H, N*H*), 7.50 (s, 1H, C*H* (6) thymine), 3.20 (s, 3H, NC*H*$_3$), 1.74 (s, 3H, C*H*$_3$ thymine). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 165.17, 151.70, 142.86, 108.66, 35.41, 12.28. OrbiTrap-MS (+): calculated for $C_6H_8N_2O_2$ + H^+ 141.0658; found 141.0658.

1-methyl-5-formyluracil (20a). N^1-methyl thymine **19a** (200 mg, 1.4 mmol) and 2,6-lutidine (570 µL, 4.9 mmol, 3.5 eq) in acetonitrile (3.5 mL) were added to a water (3.5 mL) solution of potassium persulfate (772 mg, 2.85 mmol, 2 eq) and copper(II) sulfate (91 mg, 0.57 mmol, 0.4 eq). The solution was stirred at 80°C for 1 h and the reaction progress tracked by TLC (12:1 chloroform / methanol). The solvent was evacuated under reduced pressure and the product was crystallized from acetonitrile yielding 101 mg (43%) of product. ^{1}H NMR (300 MHz, DMSO) δ (ppm): 11.74 (s, 1H, N*H*), 9.77 (s, 1H, C*H*O), 8.49 (s, 1H, C*H* (6) uracil), 3.37 (s, 3H, NC*H*$_3$). ^{13}C NMR (75 MHz, DMSO) δ (ppm): 186.71, 162.89, 153.33, 150.94, 110.30, 36.78. OrbiTrap-MS (+): calculated for $C_6H_6N_2O_3$ + Na^+ 177.0270; found 177.0269.

3-Dod-1-methyl-5-formyluracil (21a). Compound **20a** (195 mg, 1.16 mmol) and potassium carbonate (320 mg, 2.32 mmol, 2 eq) were dissolved in dry DMF (2 mL) and to it Dod-Cl (305 mg, 1.16 mmol, 1.1 eq) dissolved in dry DMF (2 mL) was added. The reaction was carried overnight at r.t., then DMF was evacuated and the residue purified by column chromatography (8:2 toluene / ethyl acetate) yielding 188 mg (50%) of viscous transparent liquid. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 10.04 (s, 1H, C*H*O), 8.08 (s, 1H, C*H* (6) uracil), 7.36-7.33 (m, 5H, Ar$_{ortho}$, ArC*H*Ar Dod), 6.90-6.87 (m, 4H, Ar$_{meta}$ Dod), 3.82 (s, 6H, OC*H*$_3$ Dod), 3.48 (s, 3H, NC*H*$_3$). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm): 186.99, 162.15,

158.98, 150.37, 147.89, 129.96, 129.77, 113.64, 110.42, 58.85, 55.25, 38.17. OrbiTrap-MS (+): calculated for $C_{21}H_{20}N_2O_5$ + Na^+ 403.1264; found 403.1264.

1-propyl-thymine (19b). Thymine (500 mg, 3.96 mmol), hexamethyldisilazane (5.2 mL, 24.9 mmol, 6.3 eq) and trimethylsilyl chloride (301 μL, 2.37 mmol, 0.6 eq) were added in a round-bottom flask and allowed to stir for 4 hours at 120°C under nitrogen atmosphere. After the reaction solution was cooled to 90°C, propyl iodide (6.3 mL, 64.9 mmol, 16.4 eq) was added to it and the mixture was let to stir overnight. Then, volatiles were removed by vacuum and to the reaction mixture were added ice water (5 mL) and acetic acid (5 mL). After 20 minutes, the solution was concentrated under reduce pressure thus yielding a white solid. Product **19b** was collected by filtration and drying (64%, 426 mg). ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 8.59 (s, 1H, N*H*), 6.98 (s, 1H, C*H* (6) uracil), 3.68-3.63 (m, 2H, NC*H*$_2$), 1.92 (s, 3H, C*H*$_3$ thymine), 1.72-1.66 (m, 2H, $NCH_2CH_2CH_3$), 0.98-0.93 (m, 3H, C*H*$_3$ propyl). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm): 164.10, 150.78, 140.46, 110.47, 50.08, 22.35, 12.33, 10.90. OrbiTrap-MS (+): calculated for $C_8H_{12}N_2O_2$ + H^+169.0972; found 169.0972.

1-propyl-5-formyluracil (20b). N^1-propyl thymine **19b** (503 mg, 2.97 mmol) and 2,6-lutidine (1.21 mL, 10.4 mmol, 3.5 eq) in acetonitrile (3.5 mL) were added to a water (3.5 mL) solution of potassium persulfate (1.6 g, 5.94 mmol, 2 eq) and copper(II) sulfate (190 mg, 1.19 mmol, 0.4 eq). The solution was stirred at 80°C for 1 h and the reaction progress tracked by TLC (8:2 ethyl acetate/ toluene). The solvent was evacuated under reduced pressure and the residue purified by column chromatography (9:1 ethyl acetate/ toluene) yielding 319 mg (59%) of product. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 10.01 (s, 1H, C*H*O), 8.99 (s, 1H, N*H*), 8.09 (s, 1H, C*H* (6) uracil), 3.84-3.79 (m, 2H, NC*H*$_2$) 1.82-1.75 (m, 2H, $NCH_2CH_2CH_3$), 1.01 (t, *J* = 7.4 Hz, 3H, C*H*$_3$ propyl). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 186.18, 162.42, 149.90, 149.34, 110.84, 51.63, 22.41, 10.81. OrbiTrap-MS (+): calculated for $C_8H_{10}N_2O_3$ + Na^+ 205.0583; found 205.0584.

3-Dod-1-propyl-5-formyluracil (21b). Compound **21a** (230 mg, 1.26 mmol) and potassium carbonate (349 mg, 2.52 mmol, 2 eq) were dissolved in dry DMF (2 mL) and to it added Dod-Cl (331 mg, 1.26 mmol, 1.1 eq) dissolved in dry DMF (2 mL). The reaction was carried overnight at r.t. Then, DMF was evacuated and the residue purified by column chromatography (8:2 toluene / ethyl acetate) yielding 293 mg (57%) of viscous transparent liquid. ^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 10.02 (s, 1H, C*H*O), 8.04 (s, 1H, C*H* (6) uracil), 7.32 (d, *J*= 8.1 Hz, 4H,

Ar_{ortho} Dod), 7.29 (s, 1H, ArC*H*Ar Dod), 6.88 (d, *J*= 7.9 Hz, 4H, Ar_{meta} Dod), 3.80 (s, 6H, OC*H*$_3$ Dod), 3.79 (t, *J*= 7.5 Hz, 2H, NC*H*$_2$), 1.76-1.68 (m, 2H, $NCH_2CH_2CH_3$), 0.95 (t, *J*= 7.3 Hz, 3H, C*H*$_3$ propyl). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 187.11, 162.11, 158.95, 149.95, 147.21, 129.91, 129.89, 113.64, 110.35, 58.77, 55.26, 52.46, 22.46, 10.81. OrbiTrap-MS (+): calculated for $C_{23}H_{24}N_2O_5$ + Na^+ 431.1577; found 431.1580.

1-allyl-5-hydroxymethyluracil (22). Compound **15** (400 mg, 2.8 mmol), hexamethyldisilazane (3.7 mL, 17.7 mmol, 6.3 eq) and trimethylsilyl chloride (200 μL, 1.68 mmol, 0.6 eq) were added in a round-bottom flask and allowed to stir for 4 hours at 120°C under nitrogen atmosphere. After the reaction solution was cooled to 66°C, allyl bromide (3.9 mL, 45.9 mmol, 16.4 eq) was added to it and the mixture was let to stir overnight. Then, volatiles were removed under vacuum and to the reaction mixture ice water (5 mL) and acetic acid (5 mL), were added. After 20 minutes, the solution was treated twice with water (5 mL) and evaporated each time. Product **22** was crystallized from ethanol and collected by filtration as white solid (70%, 360 mg). ^{1}H NMR (300 MHz, DMSO) δ (ppm): 11.30 (s, 1H, N*H*), 7.45 (s, 1H, C*H* (6) uracil), 5.88 (ddt *J* = 15.8, 10.6, 5.4Hz, 1H, C*H* allyl), 5.24-5.11 (m, 2H, C*H*$_2$ allyl), 4.98 (t, *J* = 5.3 Hz, 1H, O*H*), 4.39-4.26 (m, 2H, NC*H*$_2$), 4.14 (dd, *J* = 5.3, 1.1 Hz, 2H, C*H*$_2$OH). Full characterization is reported in literature.[46]

1-allyl-5-formyluracil (23). To a solution of **22** (111 mg, 0.61 mmol) in DCM (5 mL) was slowly added DMP (517 mg, 1.22 mmol) under nitrogen atmosphere at 0°C for 30 min. The mixture was stirred at room temperature for further 2h and then was taken up with Et_2O (10 mL) and quenched with aq. $Na_2S_2O_3$. The reaction mixture was extracted with ethyl acetate and the organic layer was dried over anhydrous sodium sulfate, the solvent was removed and the product was isolated by crystallization from ethanol in 61% yield (67 mg) as white solid. ^{1}H NMR (400 MHz, DMSO) δ (ppm): 11.78 (s, 1H, N*H*), 9.78 (s, 1H, C*H*O), 8.41 (s, 1H, C*H* (6) uracil), 5.97-5.86 (m, 1H, C*H* allyl), 5.26-5.20 (m, 2H, C*H*$_2$ allyl), 4.46 (d, *J* = 5.5 Hz, 2H, N*H*$_2$).^{13}C NMR (101 MHz, DMSO) δ (ppm): 187.14, 162.64, 153.17, 150.38, 132.25, 119.09, 111.09, 50.91. OrbiTrap-MS (+): calculated for $C_8H_8N_2O_3$ + H^+ 181.0608; found 181.0609.

3-Dod-1-allyl-5-formyluracil (24). Compound **23** (67 mg, 0.37 mmol) and potassium carbonate (103 mg, 0.74 mmol, 2 eq) were dissolved in dry DMF (1 mL) and to it added Dod-Cl (97 mg, 0.37 mmol, 1.1 eq) dissolved in dry DMF (1

mL). The reaction was carried overnight at r.t. Then, DMF was evacuated and the residue purified by column chromatography (8:2 toluene / ethyl acetate) yielding 130 mg (86%) of viscous transparent liquid. ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 10.02 (s, 1H, C*H*O), 8.05 (s, 1H, C*H* (6) uracil), 7.33 (d, *J*= 8.2 Hz, 5H, , Ar_{ortho} Dod, ArC*H*Ar Dod), 6.88 (d, *J*= 7.9 Hz, 4H, , Ar_{meta} Dod), 5.91-5.82 (m, 1H, C*H* allyl), 5.39-5.28 (m, 2H, CH_2 allyl), 4.42 (d, *J* = 5.7 Hz, 2H, NH_2), 3.80 (s, 6H, OCH_3 Dod). ^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm): 187.04, 161.99, 158.97, 150.00, 146.65, 130.47, 129.93, 129.82, 121.12, 113.65, 110.69, 58.98, 55.26, 52.18. OrbiTrap-MS (+): calculated for $C_{23}H_{22}N_2O_5$ + Na^+ 429.1421; found 429.1423.

General procedure for monosubstituted intermediates: Bis-phosphonate **6** (1 eq.) was stirred with aldehyde **21** or **18** (1.2 eq.) in a solution of aq. 6 M K_2CO_3 (206 mg, 3.67 mmol, 20 eq.) in THF (3 mL). The reaction was carried overnight under vigorous stirring. Then, the organic phase was extracted by ethyl acetate, dried over anhydrous sodium sulfate and purified by column chromatography on silica using ethyl acetate as eluent.

Monosubstituted intermediate (25a). Yellow solid (25 mg, 0.035 mmol, 34% yield).
^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.70 (s, 2H, Naphth *H (1,8)*), 8.44 (d, *J* = 15.4 Hz, 1H, vinylic C*H trans'*), 8.24 (d, *J* = 8.7 Hz, 1H, Naphth *H (3 or 6)*), 8.17 (d, *J* = 8.9 Hz, 1H, Naphth *H (3 or 6))*, 7.94 (d, *J* = 6.6 Hz, 2H, Naphth *H (4,5)*), 7.58 (s, 1H, C*H* (6) uracil), 7.52 (d, *J* = 15.2 Hz, 1H, vinylic C*H trans*), 7.44 (s, 1H, ArC*H*Ar Dod), 7.38 (d, *J* = 8.50, 4H, Ar_{ortho} Dod), 6.90 (d, *J* = 8.30, 4H, Ar_{meta} Dod) 4.20-4.10 (m, 4H, OCH_2CH_3), 3.80 (s, 6H, OCH_3 Dod), 3.78-3.75 (m, 2H, O=CCH_2P=O), 3.45 (s, 3H, NCH_3), 1.28 (t, *J* = 7.1 Hz, 6H, OCH_2CH_3). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 191.57 (d, *J* = 6.6 Hz), 189.79, 165.34, 159.03, 147.12, 137.76, 137.18, 136.41, 134.57, 132.92, 132.88, 132.33, 131.75, 131.73, 131.66, 131.10, 130.32, 130.30, 130.21, 129.98, 128.41, 128.34, 127.44, 126.60, 122.16, 113.63, 102.23, 62.81 (d, *J* = 6.7 Hz), 55.26, 39.32, 37.54, 16.34 (d, *J* = 6.4 Hz). ^{1}H-^{1}H COSY NMR (**Figure 4.56,** Appendix). OrbiTrap-MS (+): calculated for $C_{39}H_{38}N_2O_8P$ + Na^+ 688.1951; found 688.1952

Monosubstituted intermediate (25b). Yellow solid (18 mg, 0.024 mmol, 27% yield).
^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.73 (s, 2H, Naphth *H (1,8)*), 8.44 (d, *J* = 15.2 Hz, 1H, vinylic C*H trans'*), 8.27 (d, *J* = 8.6 Hz, 1H, Naphth *H (3 or 6)*), 8.20 (d, *J* = 8.6 Hz, 1H, Naphth *H (3 or 6)*), 7.97 (d, *J* = 8.5 Hz, 2H, Naphth *H (4,5)*), 7.60 (s,

1H, C*H* (6) uracil), 7.55 (d, *J* = 15.1 Hz, 1H, vinylic C*H trans*), 7.44 (s, 1H, ArC*H*Ar Dod), 7.38 (d, *J* = 8.4 Hz, 4H, Ar_{ortho} Dod), 6.91 (d, *J* = 8.3 Hz, 4H, Ar_{meta} Dod), 4.22-4.15 (m, 4H, OC$H_2$$CH_3$), 3.82 (s, 6H, OC$H_3$ Dod), 3.81 – 3.78 (m, 4H, O=CCH_2P=O, NCH_2), 1.79-1.73 (m, 2H, $NCH_2CH_2CH_3$), 1.33 (t, *J* = 7.1 Hz, 3H, OCH_2CH_3), 0.97 (t, *J* = 7.3 Hz, 3H, CH_3 propyl). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 191.59 (d, *J* = 6.6 Hz), 189.84, 158.90, 146.51, 137.77, 137.41, 136.41, 134.60 (d, *J* = 1.9 Hz), 132.91, 131.75, 131.66, 130.33, 129.89, 128.42, 128.37, 127.57, 126.59, 122.03, 113.62, 109.31, 69.53, 62.82 (d, *J* = 6.6 Hz), 55.28, 53.78, 51.97, 39.27 (d, *J* = 136.1 Hz), 29.28, 22.57, 16.35 (d, *J* = 6.3 Hz), 10.79. ^{1}H-^{1}H COSY NMR (**Figure 4.59**, Appendix). OrbiTrap-MS (+): calculated for $C_{37}H_{34}N_2O_8P$ + Na^+ 716.2264; found 716.2262

Monosubstituted intermediate (26). Yellow solid (119 mg, 0.144 mmol, 46% yield).
^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.67 (s, 1H, Naphth *H (1 or 8)*, 8.65 (s, 1H, Naphth *H (1 or 8)*), 8.37 (d, *J* = 15.1 Hz, 1H, vinylic C*H trans'*), 8.16 (d, *J* = 8.7 Hz, 1H, Naphth *H (3 or 6)*), 8.12 (d, *J* = 8.7 Hz, 2H, Naphth *H (3 or 6)*), 7.88 (d, *J* = 7.7 Hz, 2H, Naphth *H (4,5)*), 7.67 (s, 1H, C*H* (6) uracil), 7.46 (d, *J* = 15.1 Hz, 1H, vinylic C*H trans*), 7.38 (s, 1H, ArC*H*Ar Dod), 7.30 (d, *J* = 8.7 Hz, 4H, Ar_{ortho} Dod), 6.84 (d, *J* = 8.7 Hz, 4H, Ar_{meta} Dod), 4.52 (s, 2H, NCH_2), 4.15 – 4.09 (m, 6H, OC$H_2$$CH_3$, OCOC$H_2$), 3.79 – 3.72 (m, 8H, OCH_3 Dod, O=CCH_2P=O), 1.60 – 1.55 (m, 2H, $OCOCH_2CH_2$), 1.28-1.19 (m, 10H, OCH_2CH_3, $OCOCH_2CH_2CH_2CH_2$), 0.84 (t, *J* = 6.7 Hz, 3H, CH_3 chain). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): 191.57 (d, *J* = 6.6 Hz), 189.69, 166.93, 161.39, 158.87, 152.59, 146.38, 137.78, 136.93, 136.32, 134.63, 134.61, 132.90, 131.75 (d, *J* = 7.4 Hz), 130.05, 129.90, 128.42, 128.32, 127.54, 126.62, 122.66, 113.62, 66.55, 62.81 (d, *J* = 6.5 Hz), 55.25, 50.29, 39.30 (d, *J* = 130.4 Hz), 28.10, 27.86, 22.24, 16.34 (d, *J* = 6.8 Hz), 13.87. ^{1}H-^{1}H COSY NMR (**Figure 4.62**, Appendix). OrbiTrap-MS (+): calculated for $C_{43}H_{44}N_2O_{10}P$ + Na^+ 802.2632; found 802.2636

General procedure for asymmetric dimers: Monosubstitute intermediate **26** (1 eq.) was stirred with aldehyde **21 (a,b)** (1.2 eq.) in a solution of aq. 6 M K_2CO_3 (206 mg, 3.67 mmol, 20 eq.) in THF (3 mL). The reaction was carried overnight under vigorous stirring. Then, the organic phase was extracted by ethyl acetate, dried over anhydrous sodium sulfate and purified by column chromatography on silica using ethyl acetate as eluent.

Asymmetric dimer (27a). Yellow solid (67 mg, 0.064 mmol, 53% yield).
^{1}H NMR (400 MHz, $CDCl_3$) δ (ppm): 8.74 (s, 2H, Naphth *H (1,8)*), 8.49-8.42 (m, 2H, vinylic C*H trans'*), 8.26 (d, *J* = 8.8 Hz, 2H, Naphth *H (3,6)*), 7.99 (d, *J* = 8.6 Hz, 2H, Naphth *H (4,5*)), 7.60 (s, 1H, C*H* (6) uracil), 7.54 (s, 1H, C*H* (6) uracil), 7.56 (d, *J* = 15.2 Hz, 2H, vinylic C*H trans*), 7.45 (s, 2H, ArC*H*Ar Dod), 7.39-7.34 (m, 8H, Ar_{ortho} Dod), 6.89 (d, *J* = 8.0 Hz, 8H, , Ar_{meta} Dod), 4.53 (s, 2H, NCH_2), 4.19 (t, *J* = 7.0 Hz 2H, OC$H_2$$CH_3$), 3.80 (s, 12H, OC$H_3$ Dod), 3.48 (s, 3H, NCH_3) 1.65 – 1.62 (m, 2H, $OCOCH_2CH_2$), 1.34-1.28 (m, 4H, $OCH_2CH_2CH_2$), 0.91 (t, *J* = 6.6 Hz, 3H, CH_3 pentyl chain).^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm): δ 189.99, 166.96, 162.81, 158.88, 153.03, 149.91, 137.52, 136.23, 136.15, 131.94, 131.57, 130.21, 130.05, 129.99, 129.91, 128.40, 127.08, 122.84, 113.61, 66.49, 60.05, 55.22, 50.91, 37.70, 28.10, 27.86, 22.23, 13.86. ^{1}H-^{1}H COSY NMR (**Figure**). OrbiTrap-MS (+): calculated for $C_{62}H_{58}N_4O_{12}$ + Na^+ 1073.3949; found 1073.3953

Asymmetric dimer (27b). Yellow solid (~ 28 mg, ~ 0.026 mmol, ~61% yield, since material was not pure enough).
The crude product was characterized by NMR but showed the presence of unreacted aldehyde. However, it was possible to identify the peaks relative to the product as reported below, by analogy with the previous product **27a**.

^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm): 8.73 (s, 2H, Naphth *H (1,8)*), 8.48 (d, *J* = 15.2 Hz, 1H, C*H trans'*), 8.43 (d, *J* = 15.2 Hz, 1H, vinylic C*H trans'*), 8.25-8.21 (m, 2H, Naphth *H (3,6)*), 7.97 (d, *J* = 8.7 Hz, 2H, Naphth *H (4,5)*), 7.60 (s, 1H, C*H* (6) uracil), 7.58 (s, 1H, C*H* (6) uracil), 7.57 (d, *J* = 15.3 Hz, 2H, vinylic C*H trans*), 7.44 (s, 2H, ArC*H*Ar Dod), 7.38 – 7.32 (m, 10H, Ar_{ortho} Dod), 6.90-6.86 (m, 8H, Ar_{meta} Dod), 4.53 (s, 2H, NCH_2), 4.23 – 4.11 (m, 4H, OC$H_2$$CH_3$, NC$H_2$$CH_2CH_3$), 3.83 – 3.75 (m, 12H, OCH_3 Dod), 1.74 – 1.58 (m, 4H $OCOCH_2CH_2$ pentyl chain, $NCH_2CH_2CH_3$ propyl chain), 1.34-1.28 (m, 4H, $OCH_2CH_2CH_2$ pentyl chain), 0.98 – 0.88 (m, 6H, CH_3 pentyl chain, $NCH_2CH_2CH_3$ propyl chain).
This batch should be further purified to complete the characterization and proceed towards the synthesis of the PNA monomer.

4.6 *Appendix*

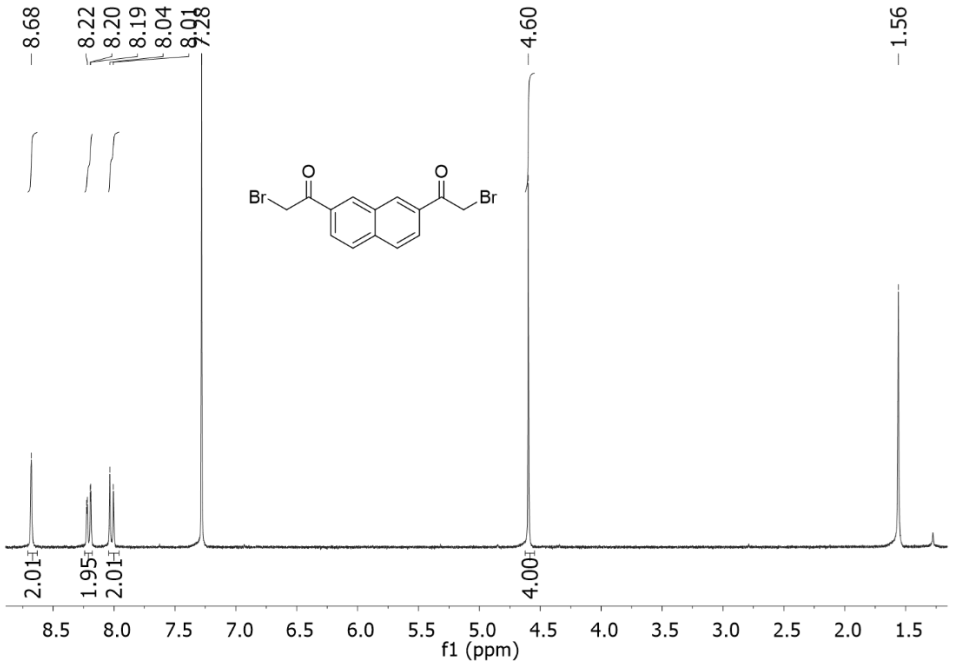


*Figure 4.8. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **5**.*

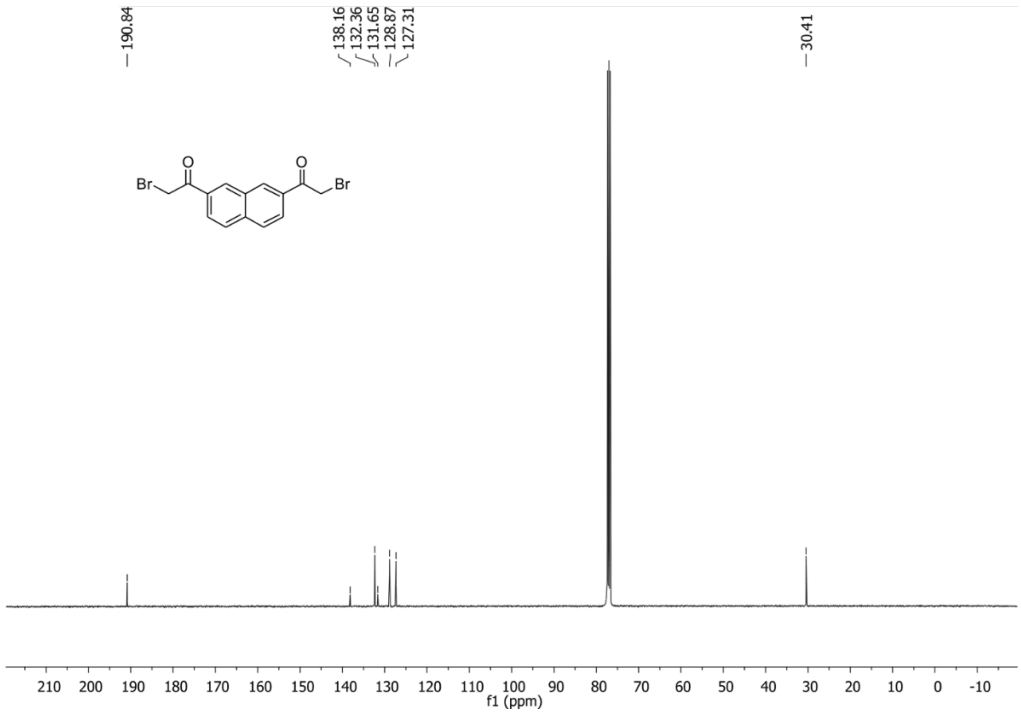


*Figure 4.9. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **5**.*

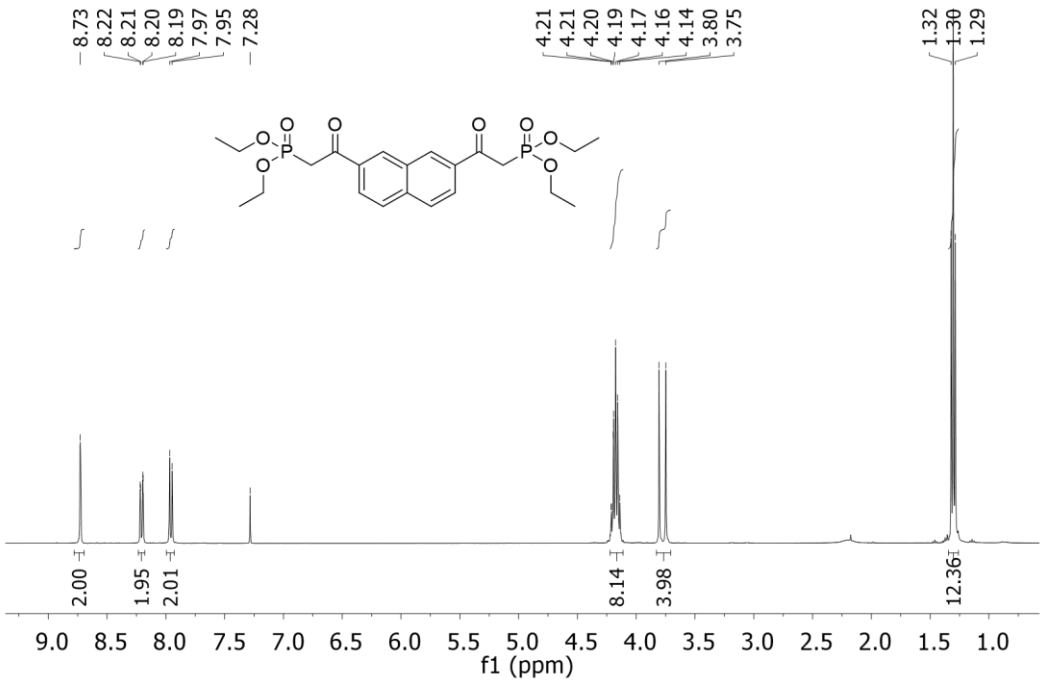


*Figure 4.10. [1]H NMR (400 MHz, $CDCl_3$) of compound **6.***

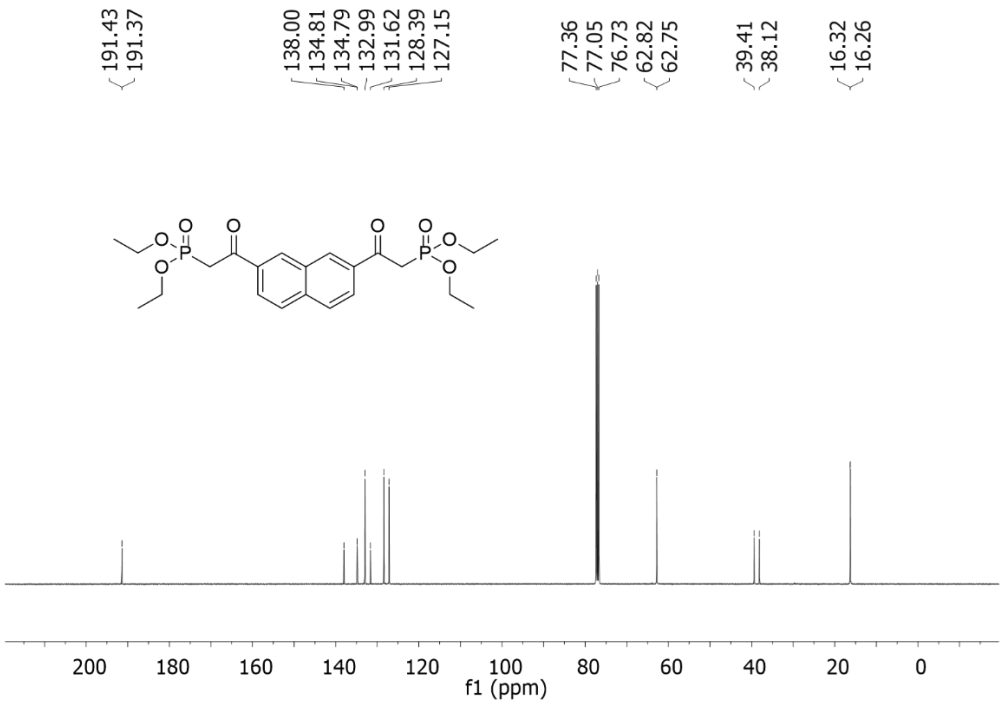


*Figure 4.11. [13]C NMR (101 MHz, $CDCl_3$) of compound **6.***

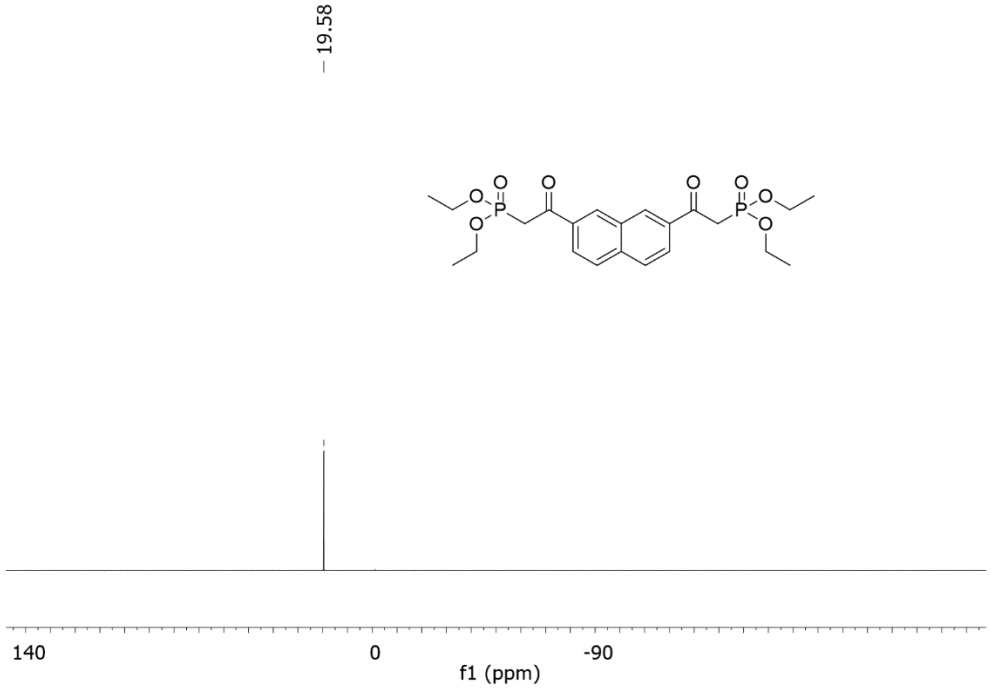


*Figure 4.12. ^{31}P NMR (162 MHz, $CDCl_3$) of compound **6**.*

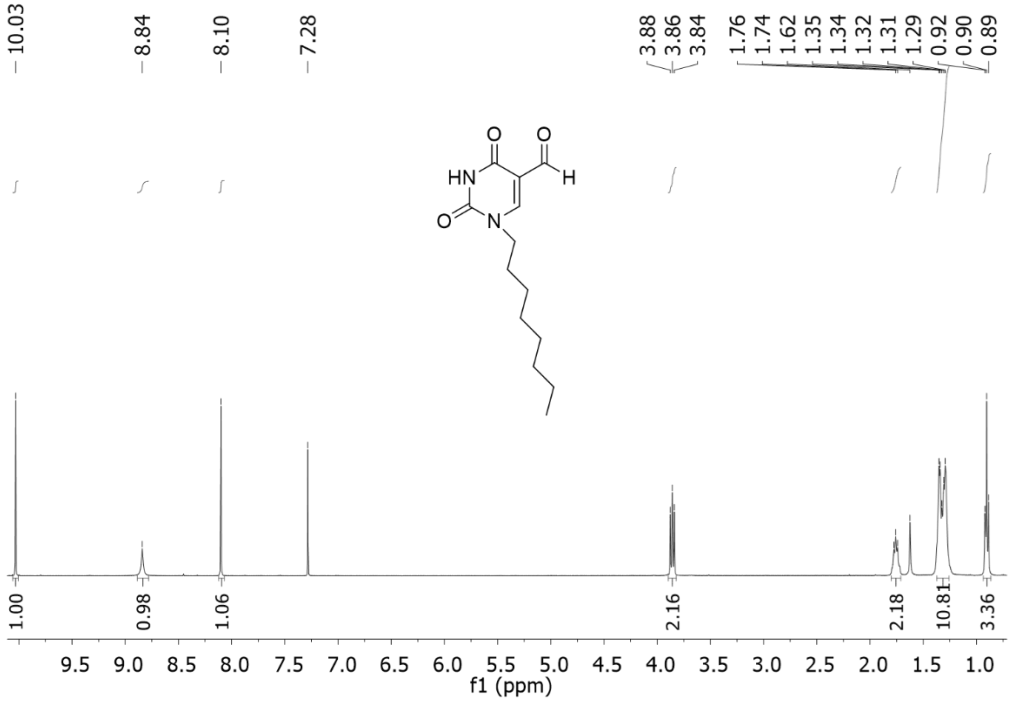


*Figure 4.13. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **8**.*

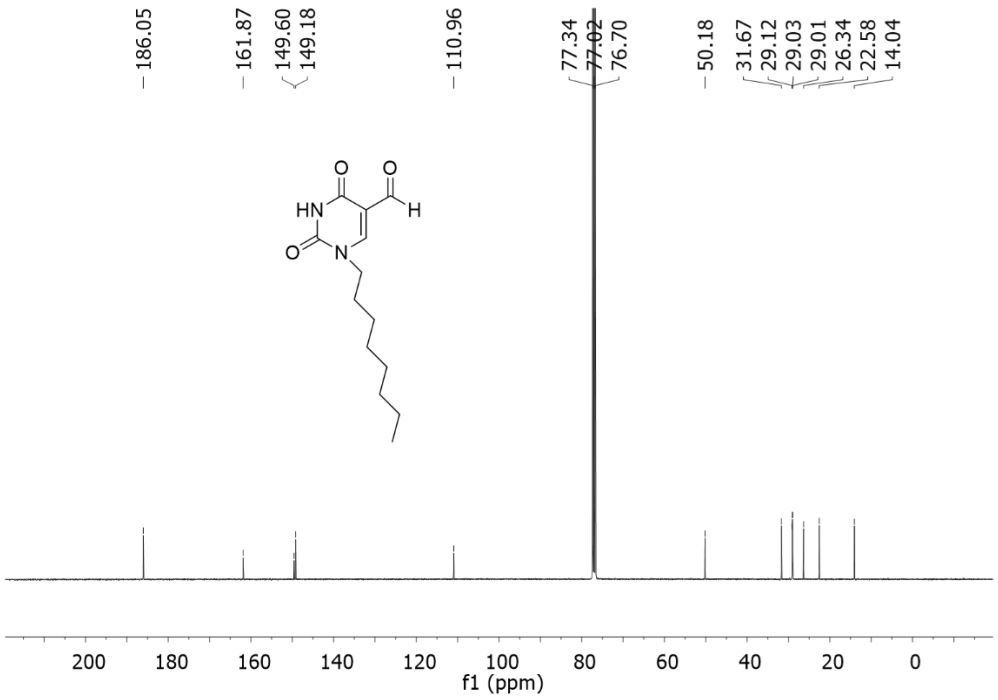


*Figure 4.14. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **8**.*

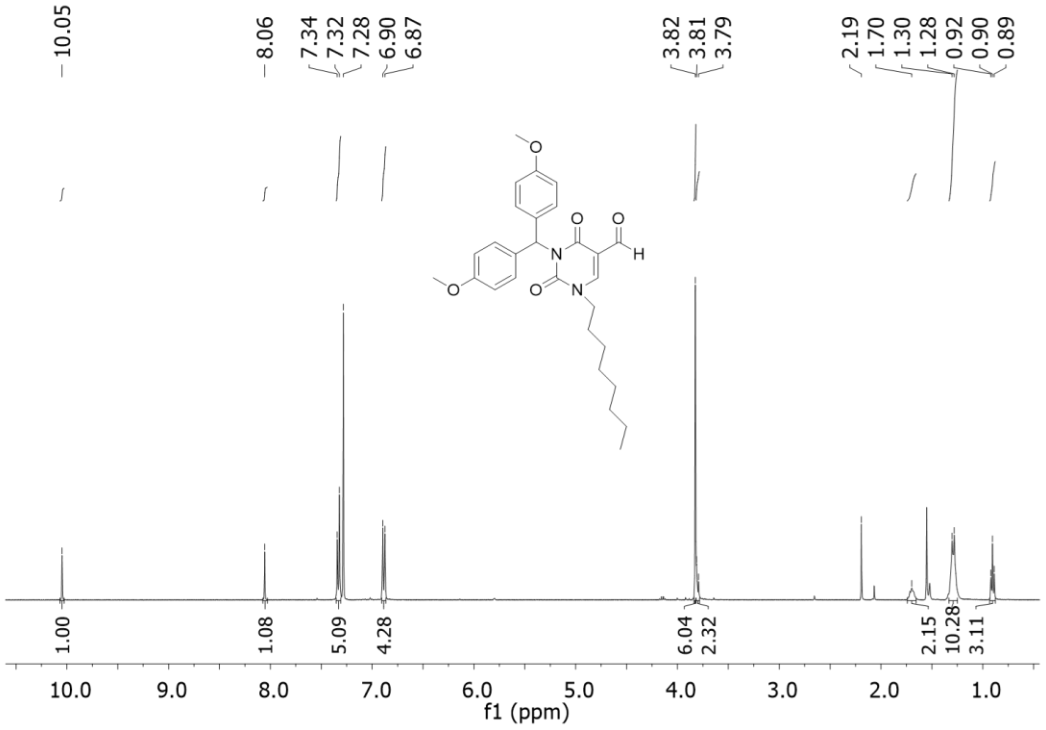


*Figure 4.15. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **9**.*

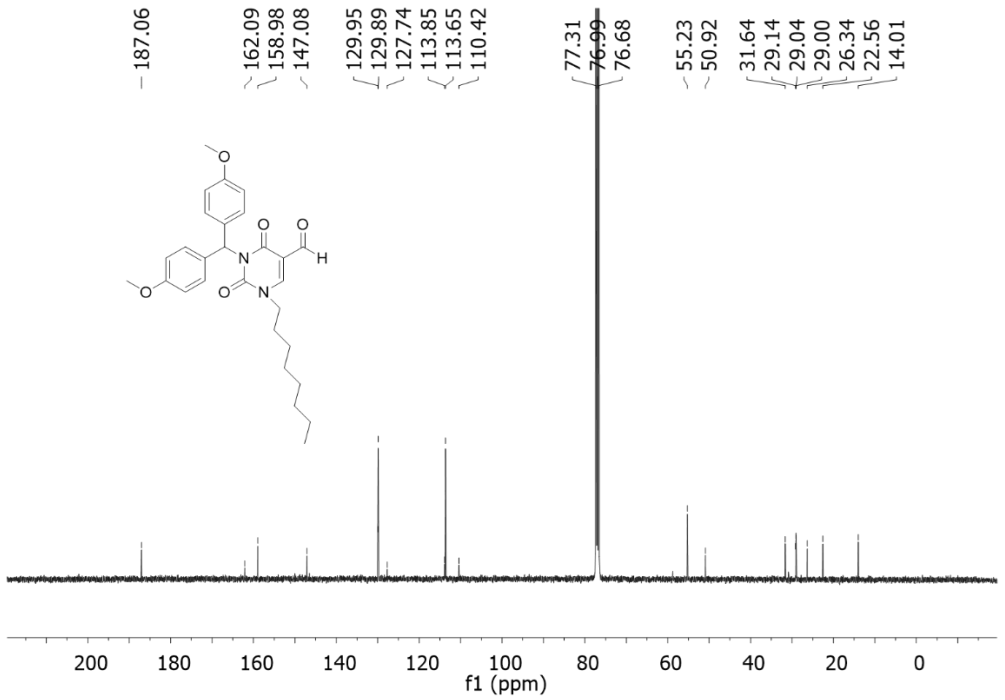


*Figure 4.16. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **9**.*

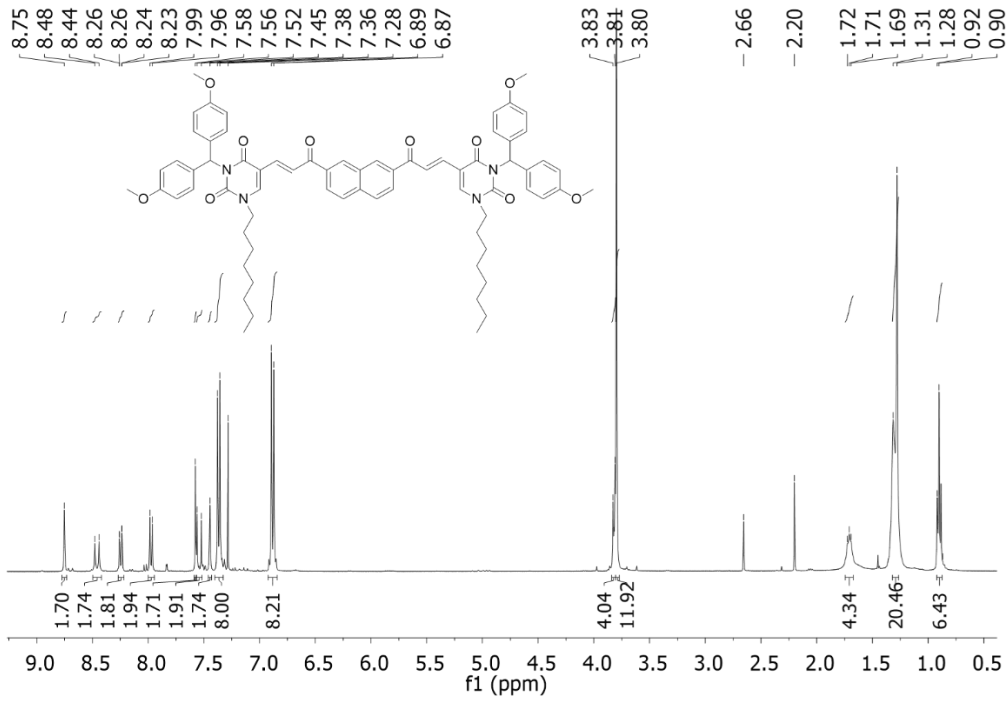


*Figure 4.17. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **10**.*

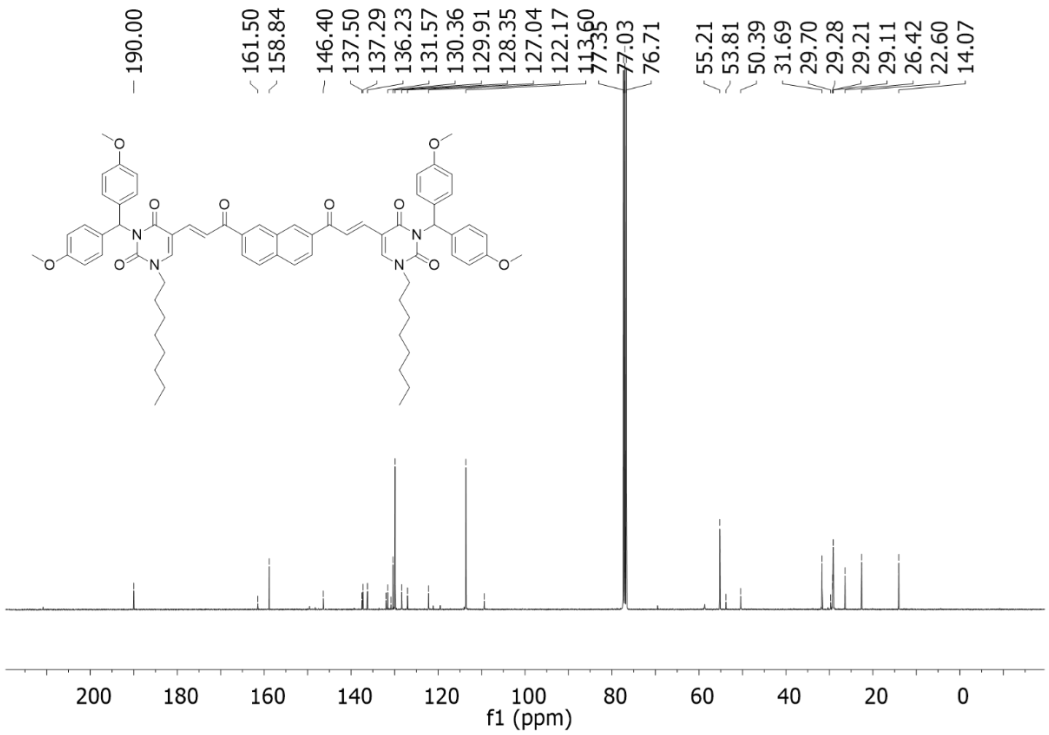


*Figure 4.18. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **10**.*

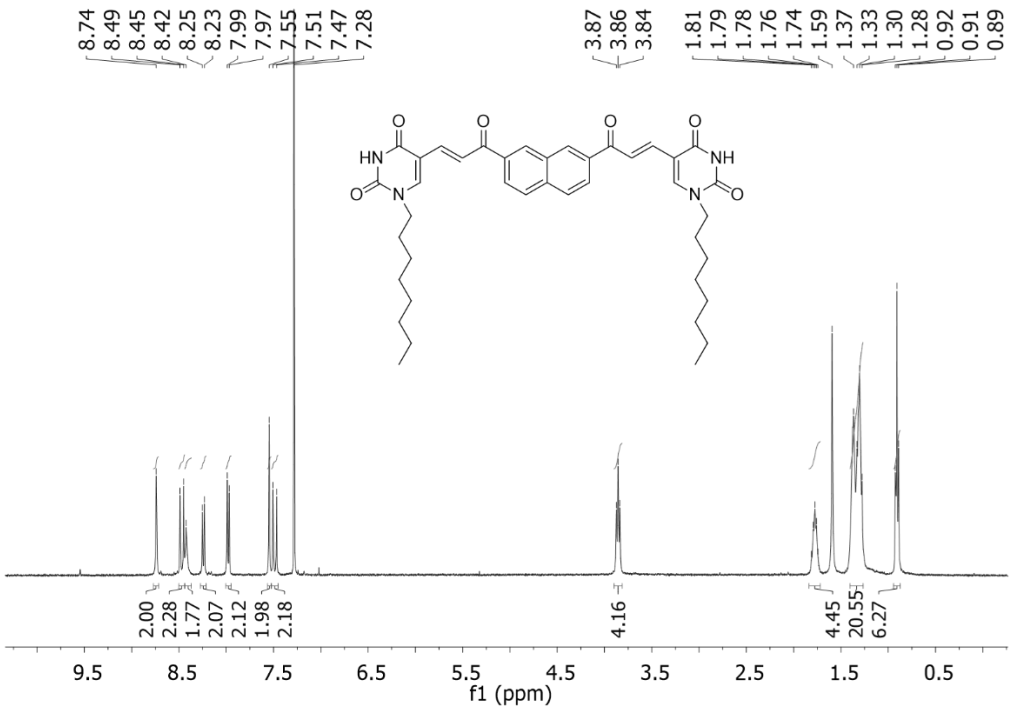


*Figure 4.19. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **11**.*

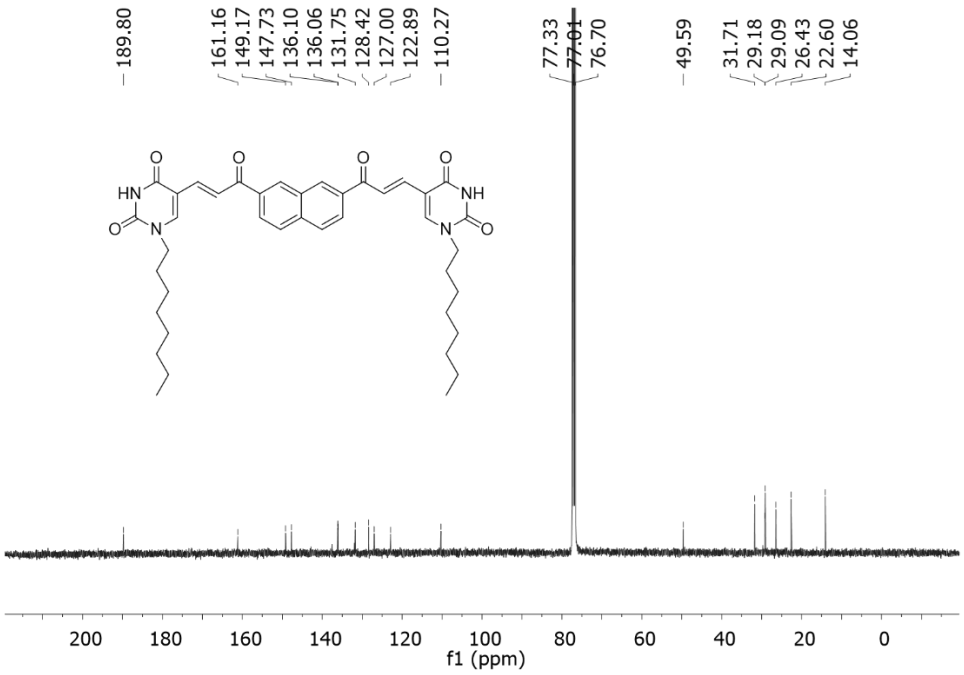


*Figure 4. 20. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **11.***

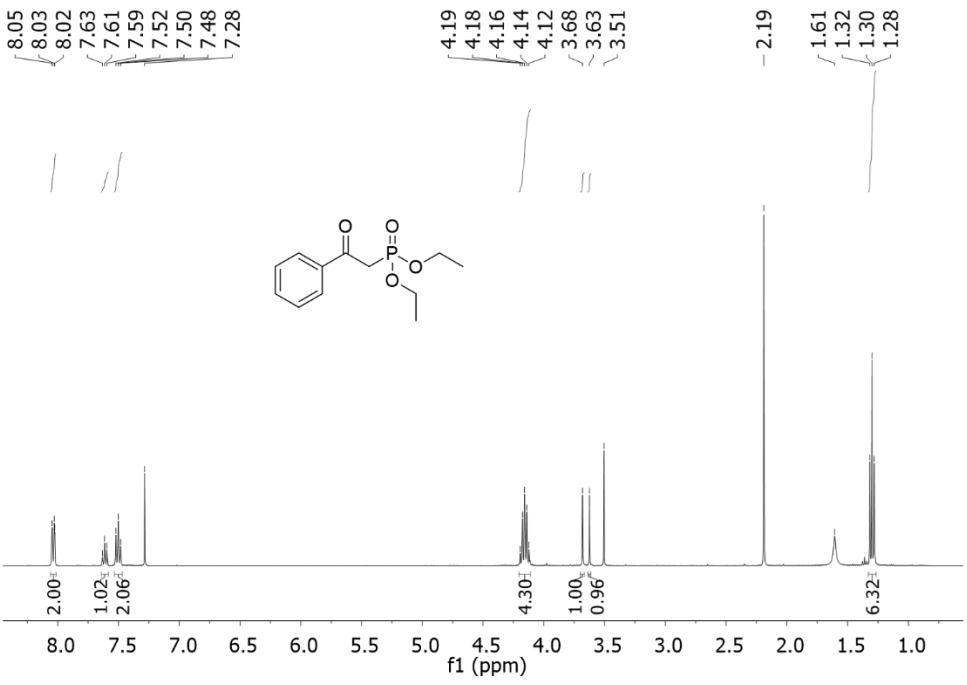


*Figure 4. 21. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **12**.*

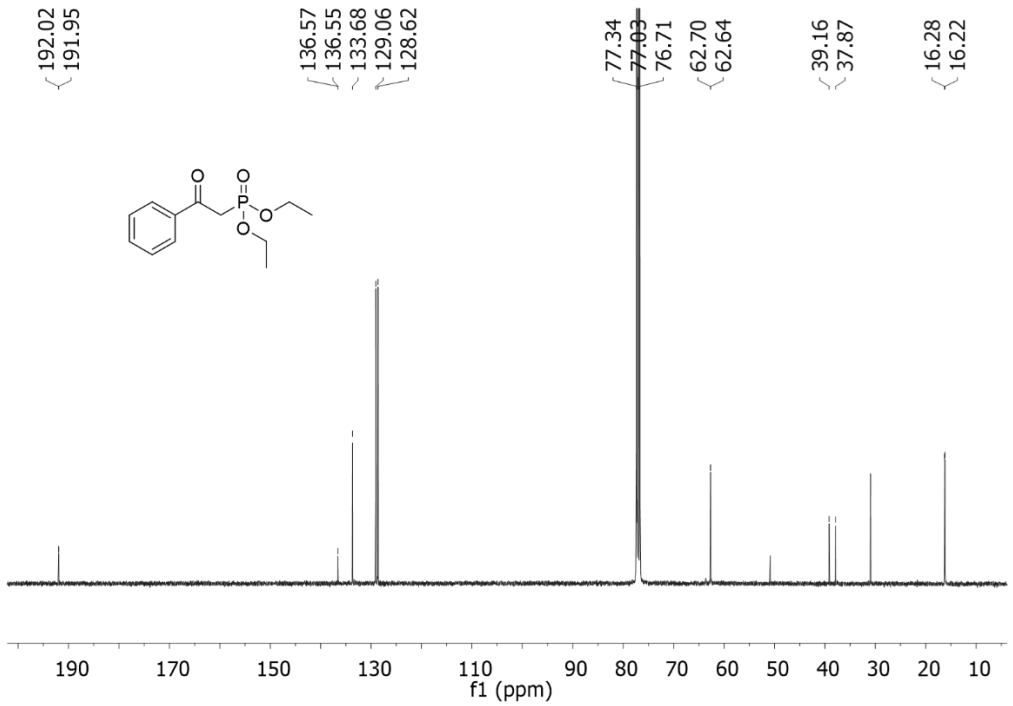


*Figure 4. 22. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **12**.*

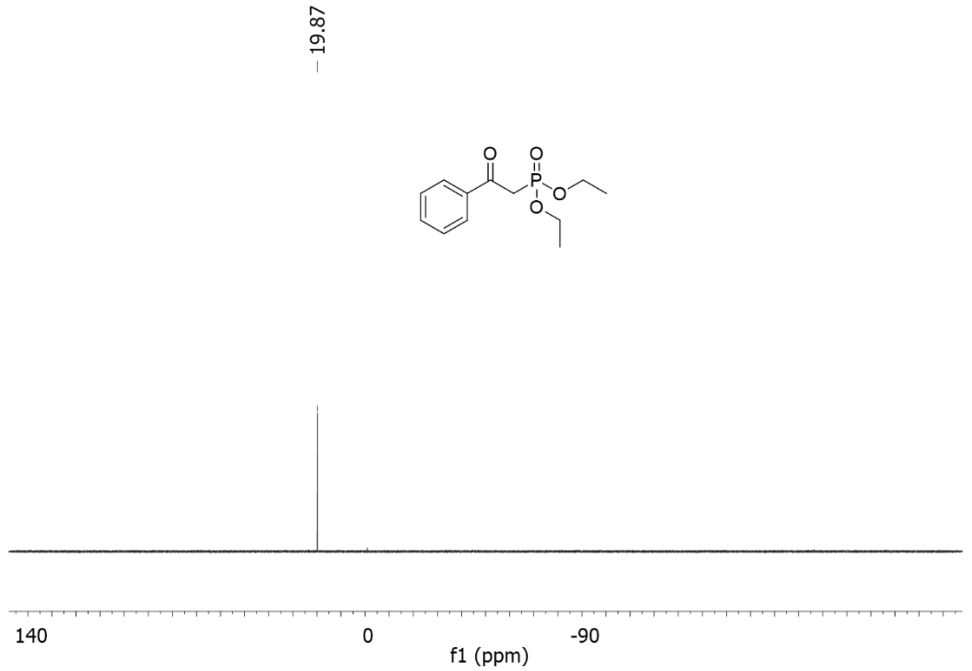


*Figure 4.23 . ^{31}P NMR (162 MHz, $CDCl_3$) of compound **12**.*

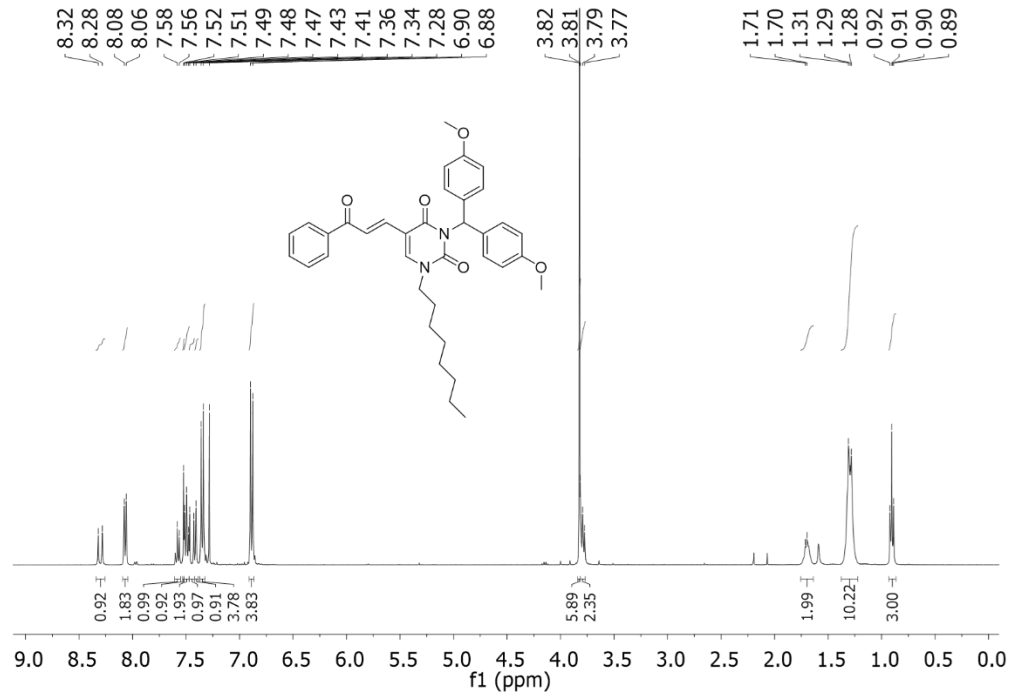


*Figure 4.24. [1]H NMR (400 MHz, $CDCl_3$) of compound **13**.*

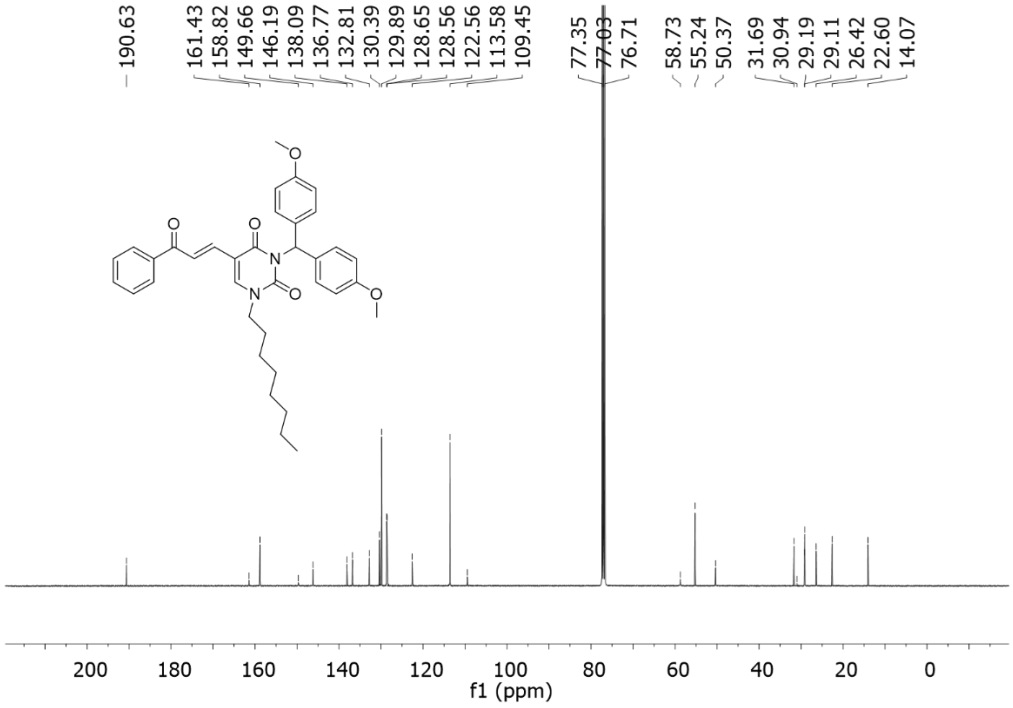


*Figure 4.25. [13]C NMR (101 MHz, $CDCl_3$) of compound **13**.*

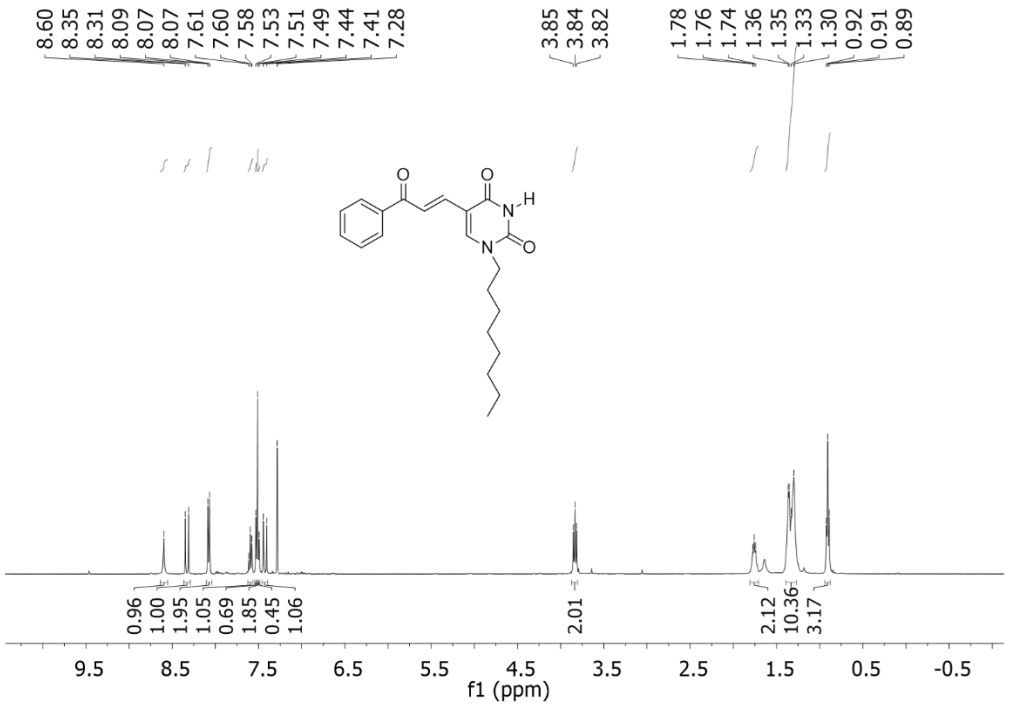


*Figure 4. 26. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **14**.*

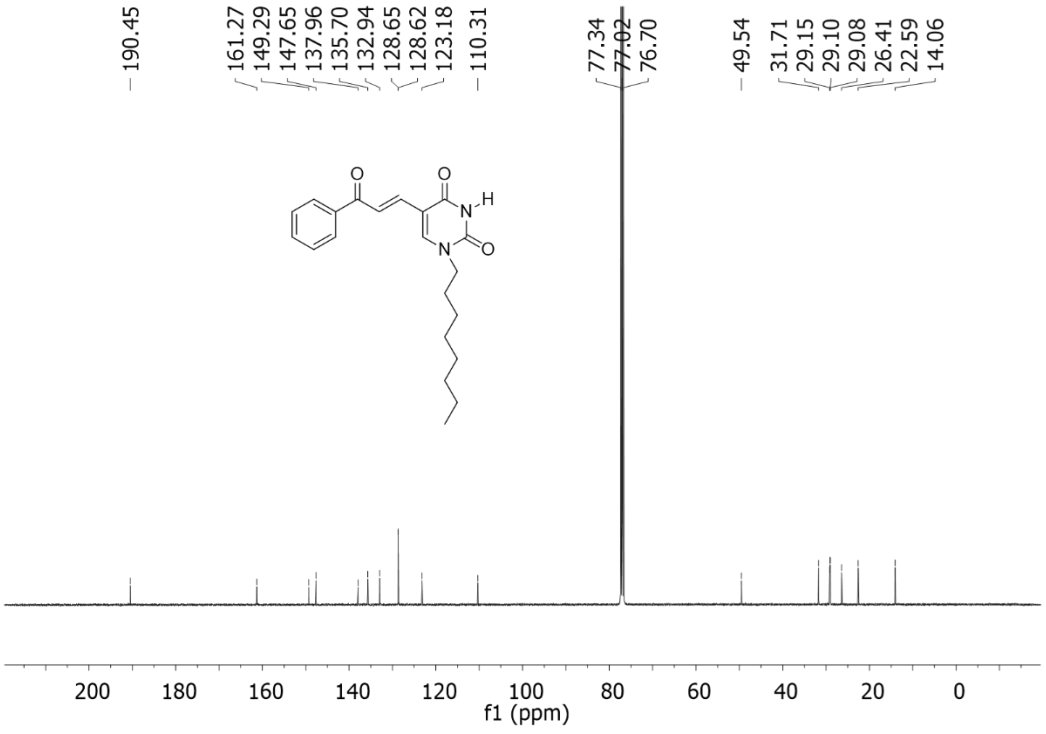


*Figure 4. 27. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **14**.*

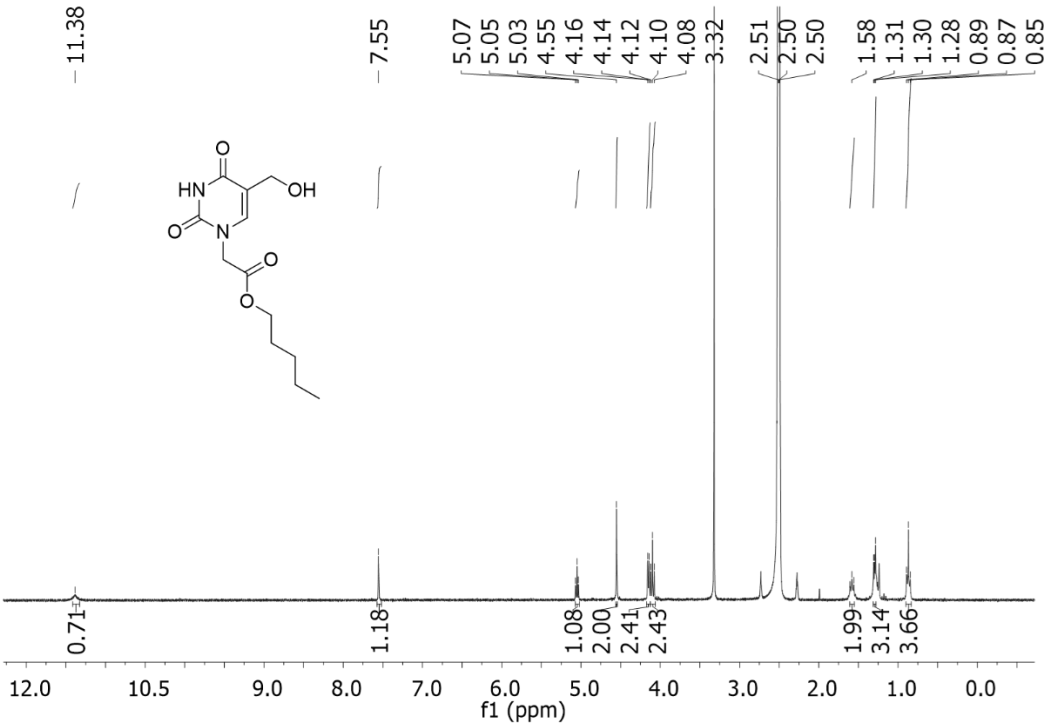


*Figure 4. 28. [1]H NMR (400 MHz, DMSO) of compound **16**.*

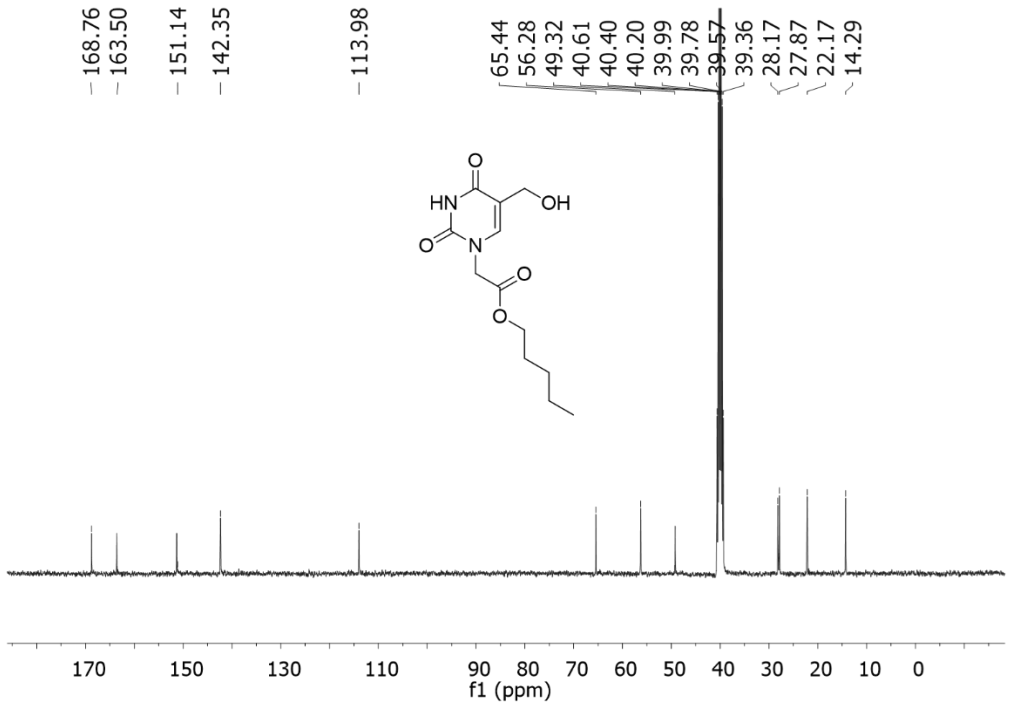


*Figure 4. 29. [13]C NMR (101 MHz, DMSO) of compound **16**.*

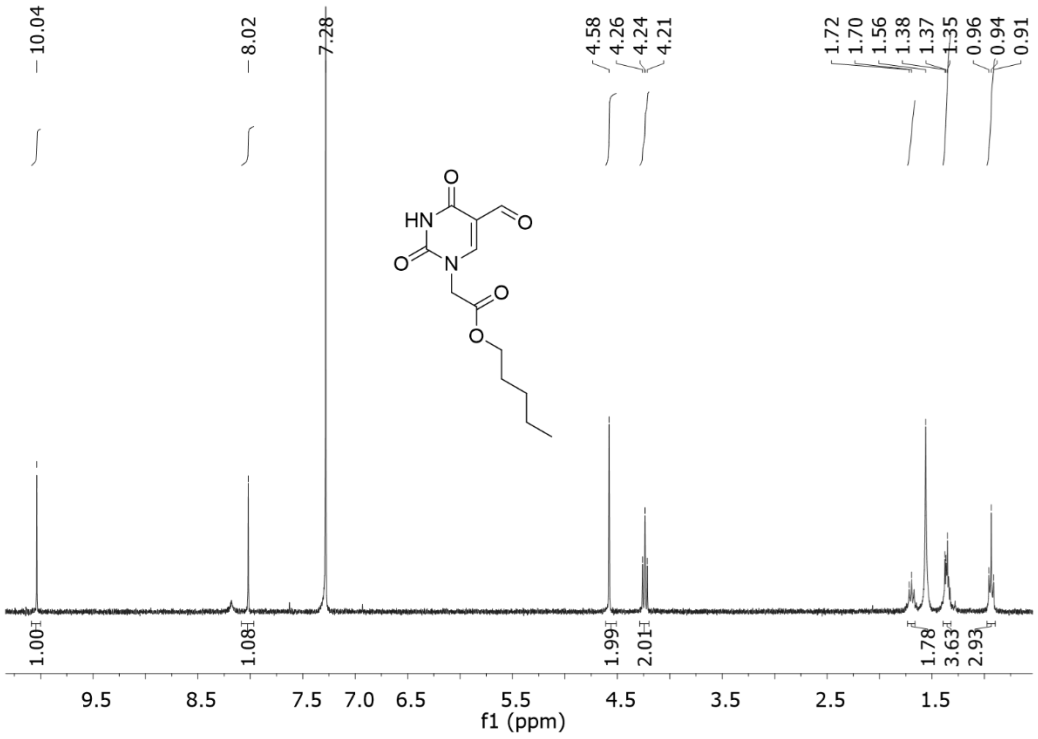


*Figure 4. 30. [1]H NMR (300 MHz, $CDCl_3$) of compound **17**.*

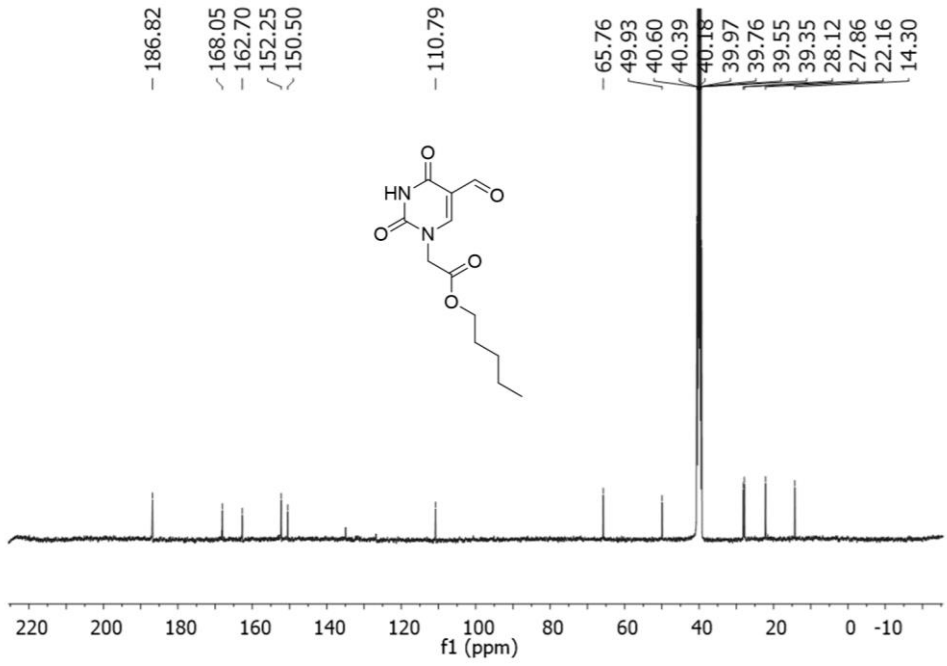


*Figure 4. 31. [13]C NMR (75 MHz, DMSO) of compound **17**.*

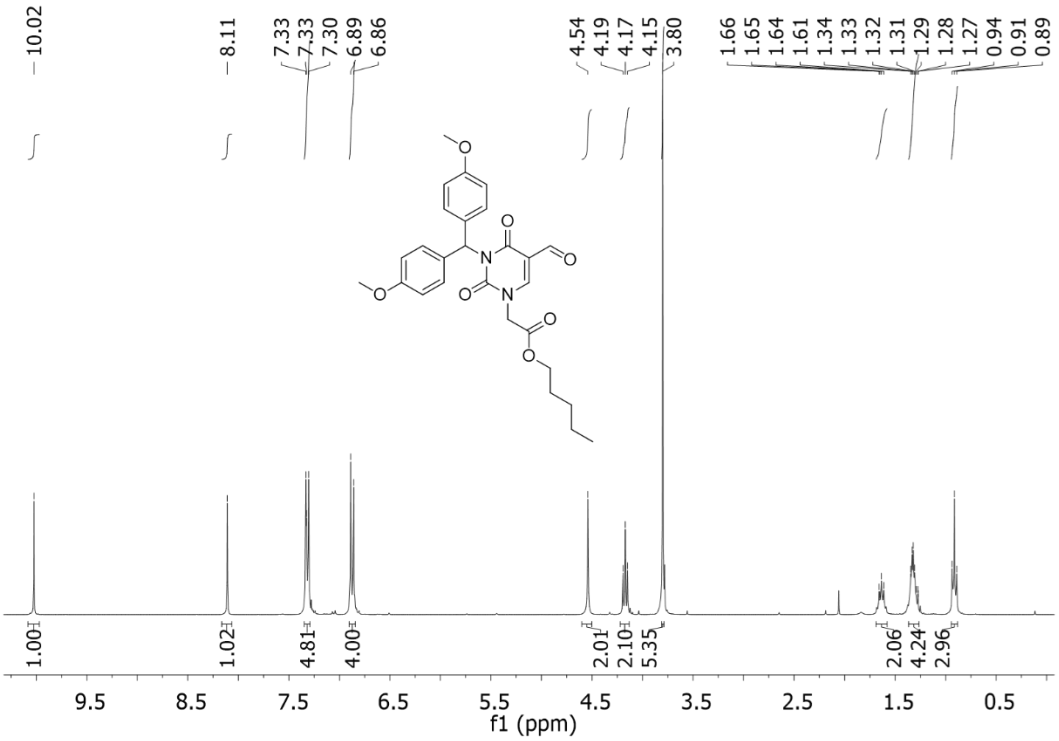


*Figure 4. 32. ^{1}H NMR (300 MHz, $CDCl_3$) of compound **18**.*

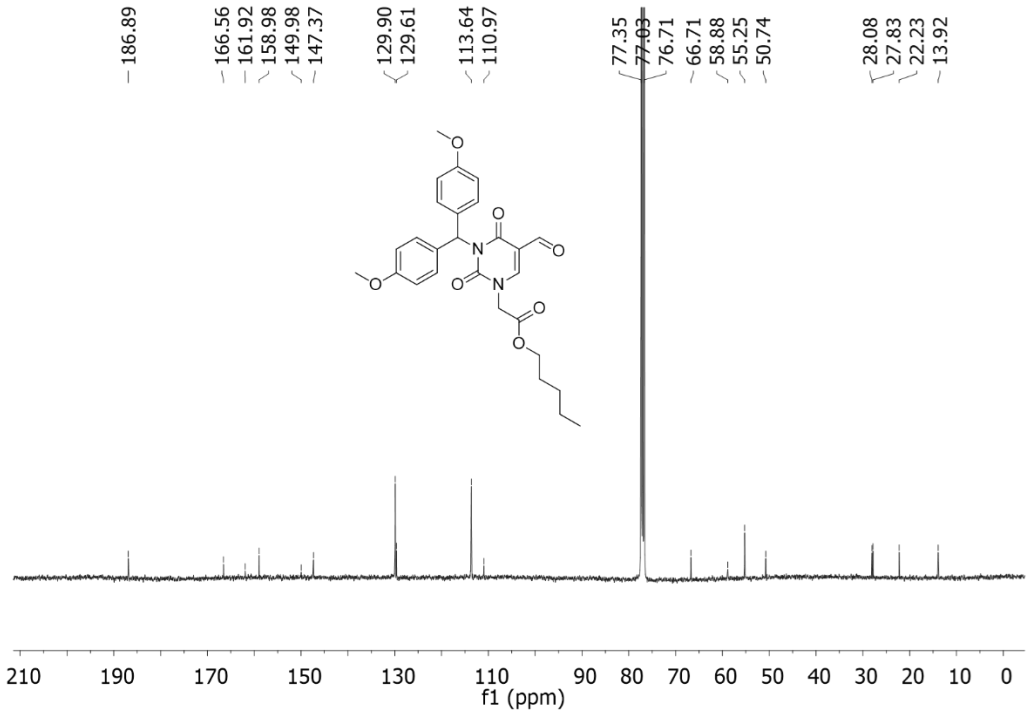


*Figure 4. 33. ^{13}C NMR (75 MHz, $CDCl_3$) of compound **18**.*

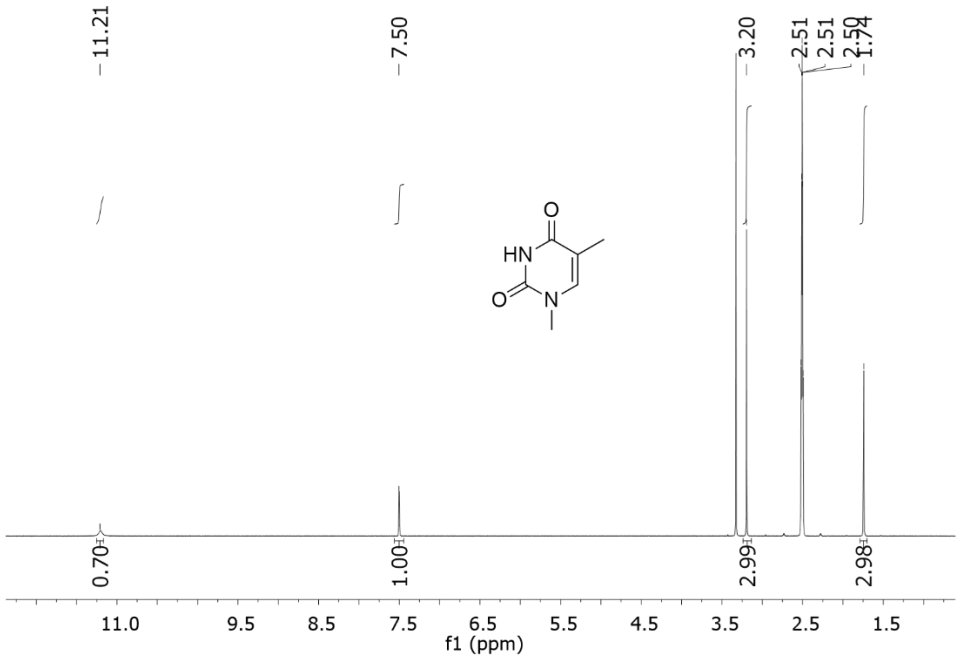


*Figure 4. 34. [1]H NMR (300 MHz, DMSO) of compound **19a**.*

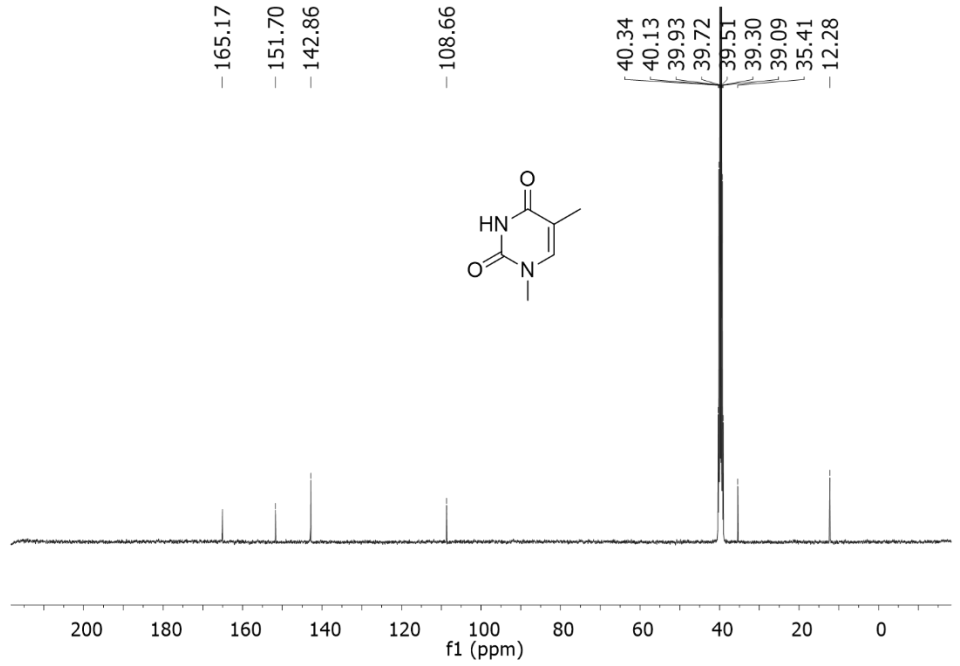


*Figure 4. 35. [13]C NMR (75 MHz, DMSO) of compound **19a**.*

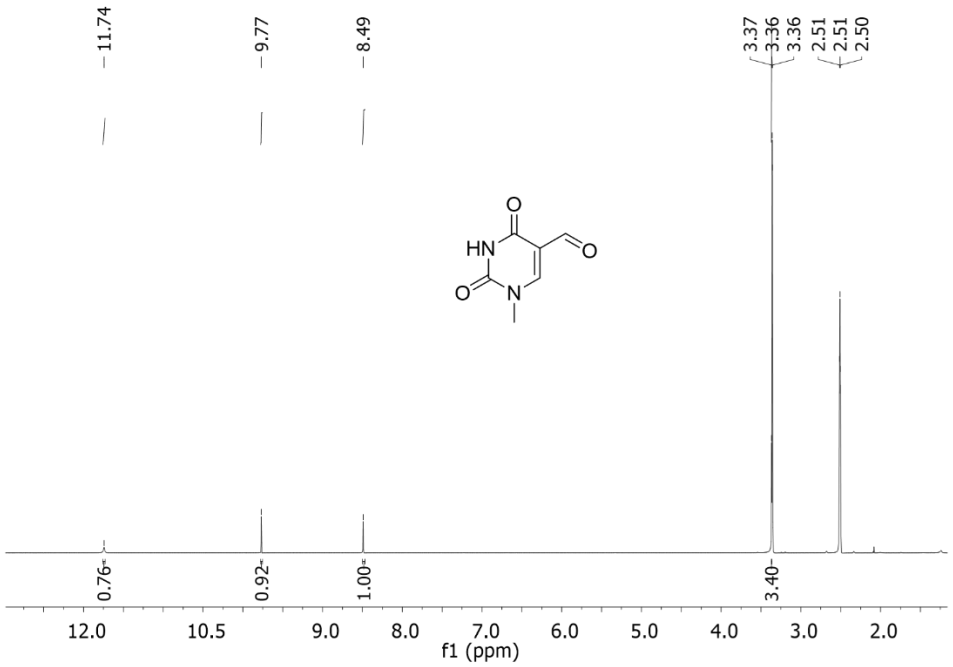


*Figure 4. 36. [1]H NMR (300 MHz, DMSO) of compound **20a**.*

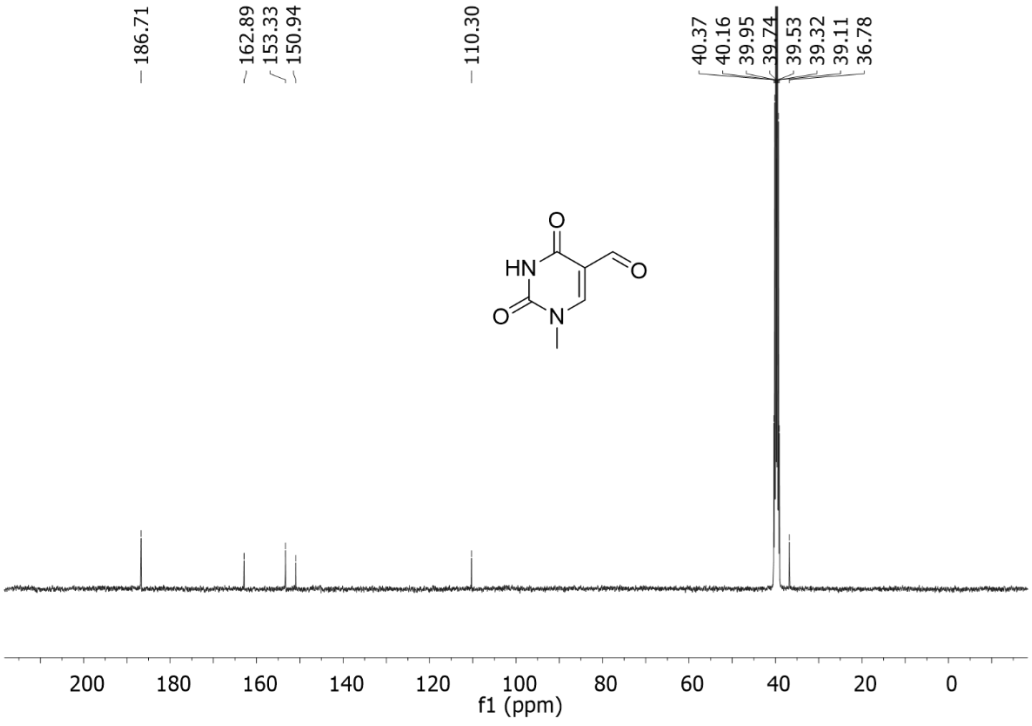


*Figure 4. 37. [13]C NMR (75 MHz, DMSO) of compound **20a**.*

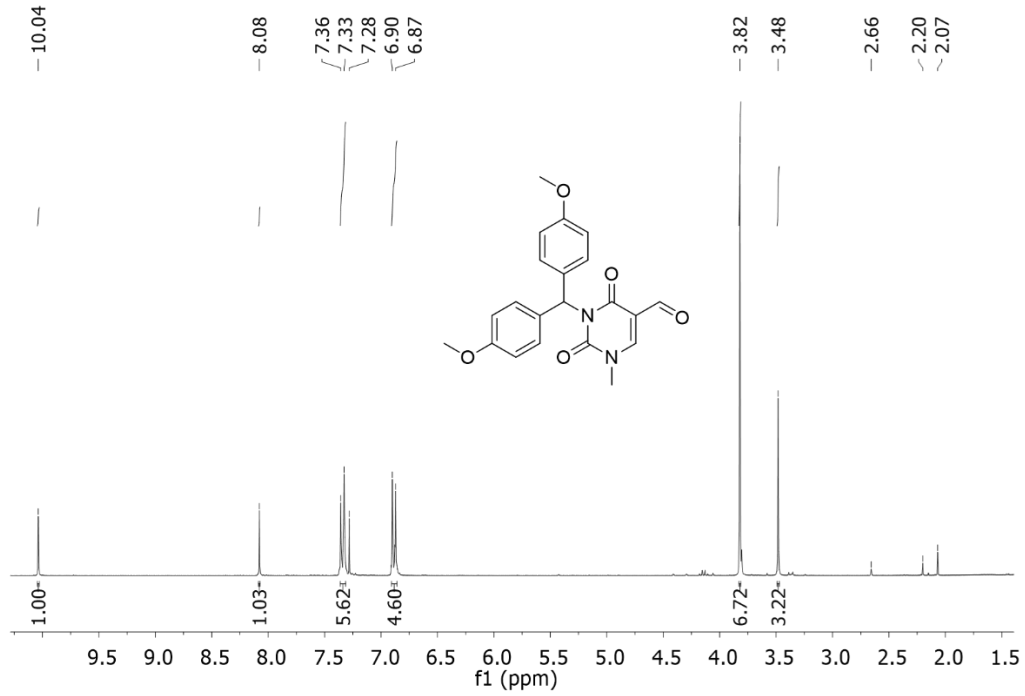


*Figure 4. 38. ^{1}H NMR (300 MHz, $CDCl_3$) of compound **21a**.*

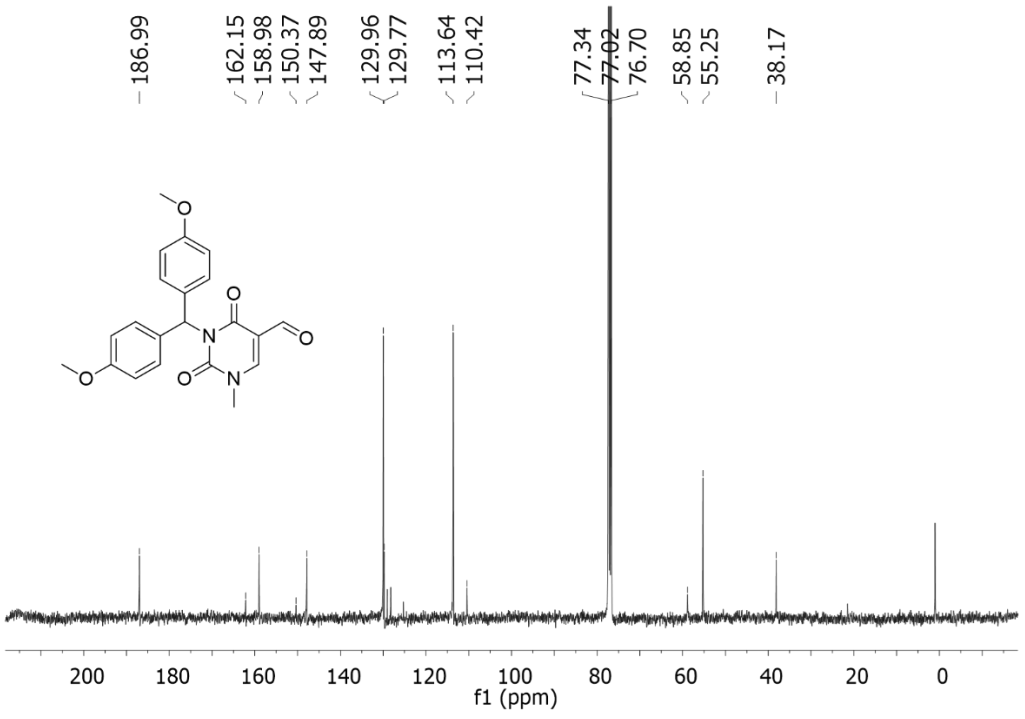


*Figure 4. 39. ^{13}C NMR (75 MHz, $CDCl_3$) of compound **21a**.*

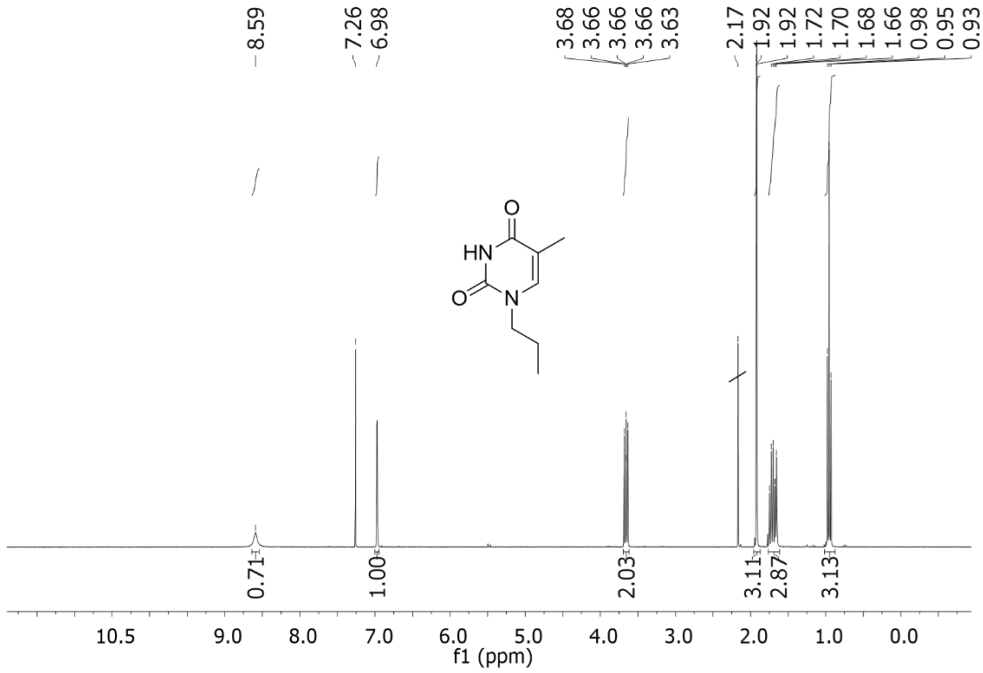


*Figure 4. 40. ^{1}H NMR (300 MHz, $CDCl_3$) of compound **19b**.*

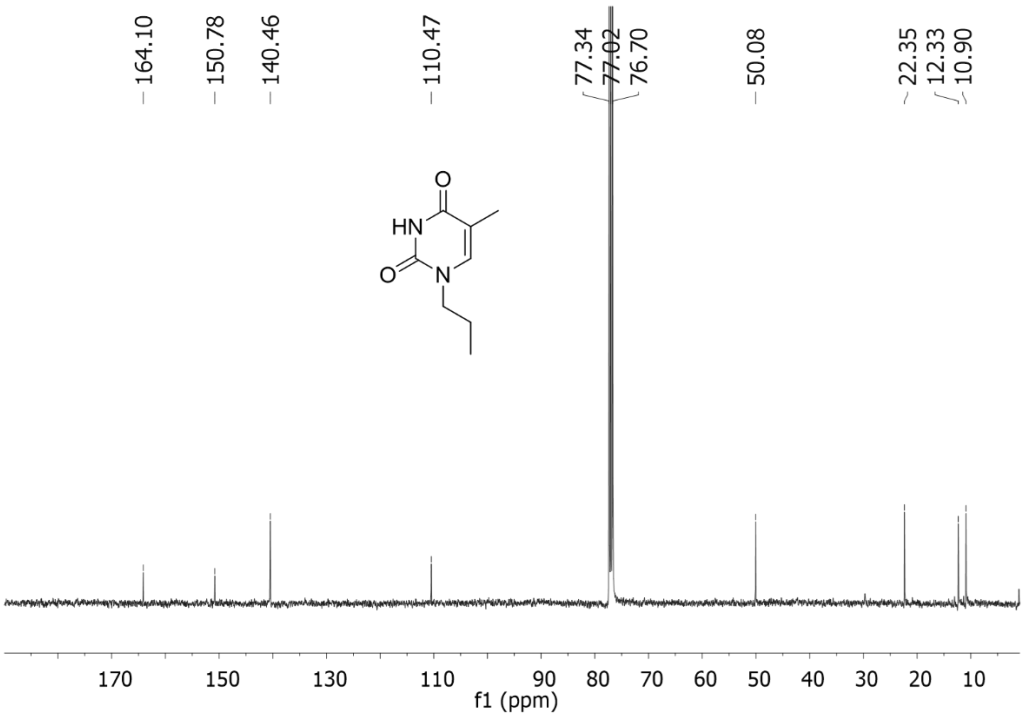


*Figure 4. 41. ^{13}C NMR (75 MHz, $CDCl_3$) of compound **19b**.*

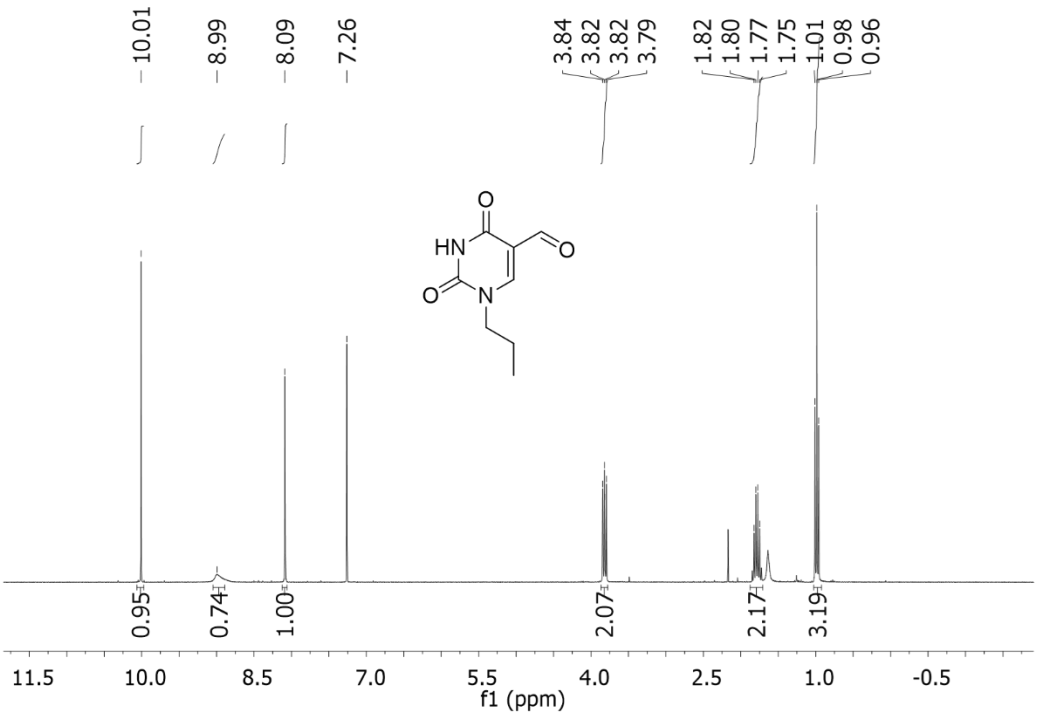


*Figure 4. 42. [1]H NMR (400 MHz, $CDCl_3$) of compound **20b**.*

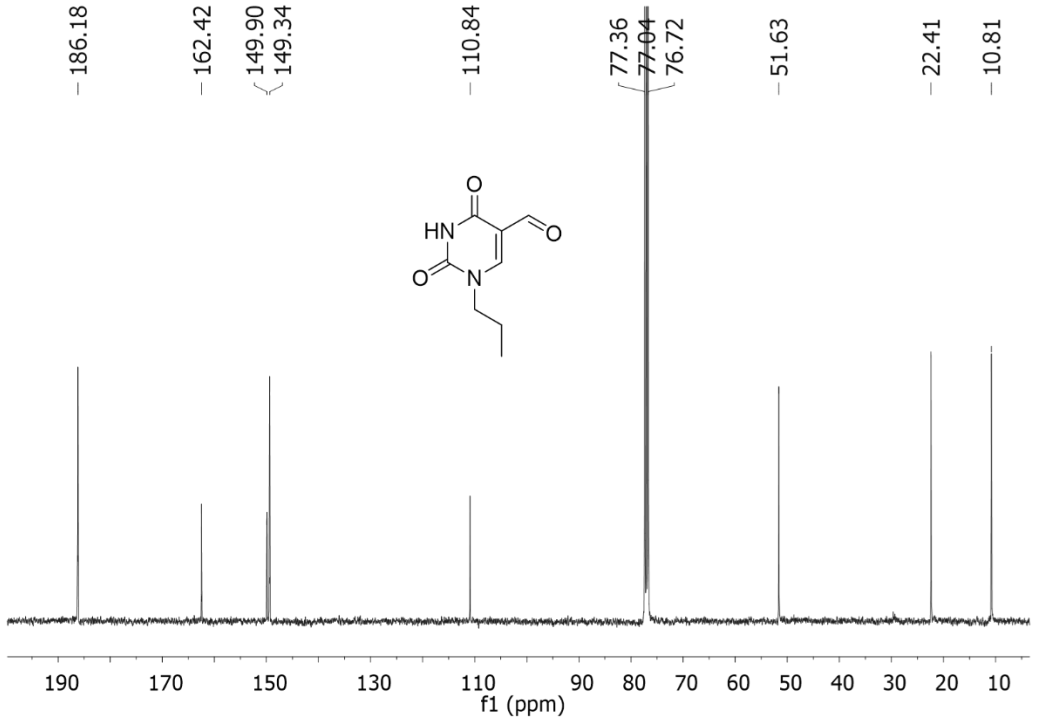


*Figure 4. 43. [13]C NMR (101 MHz, $CDCl_3$) of compound **20b**.*

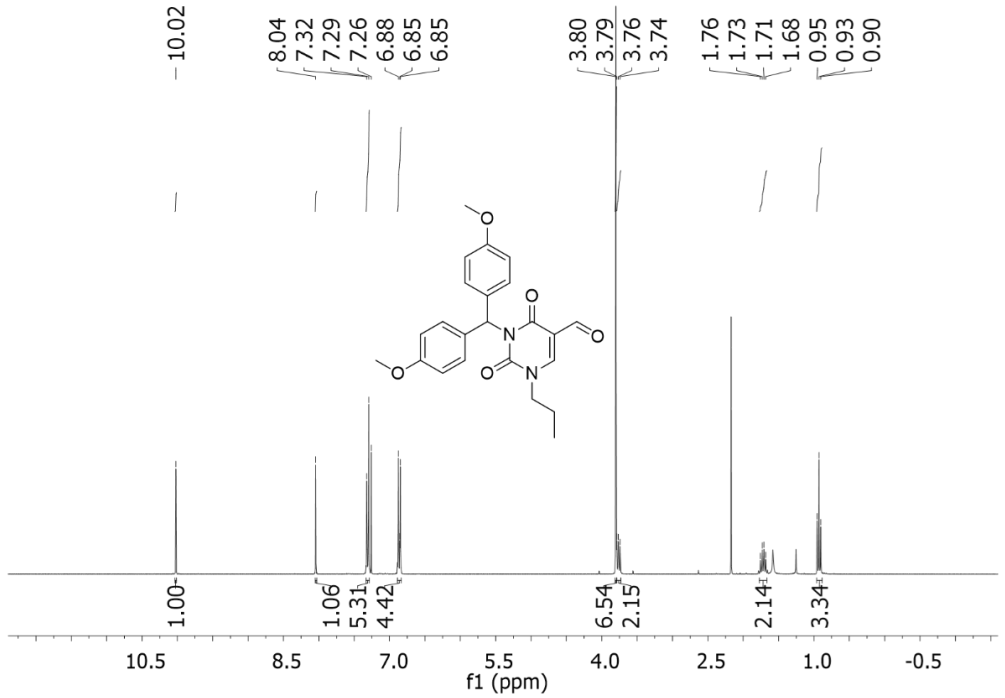


*Figure 4. 44. [1]H NMR (400 MHz, $CDCl_3$) of compound **21b**.*

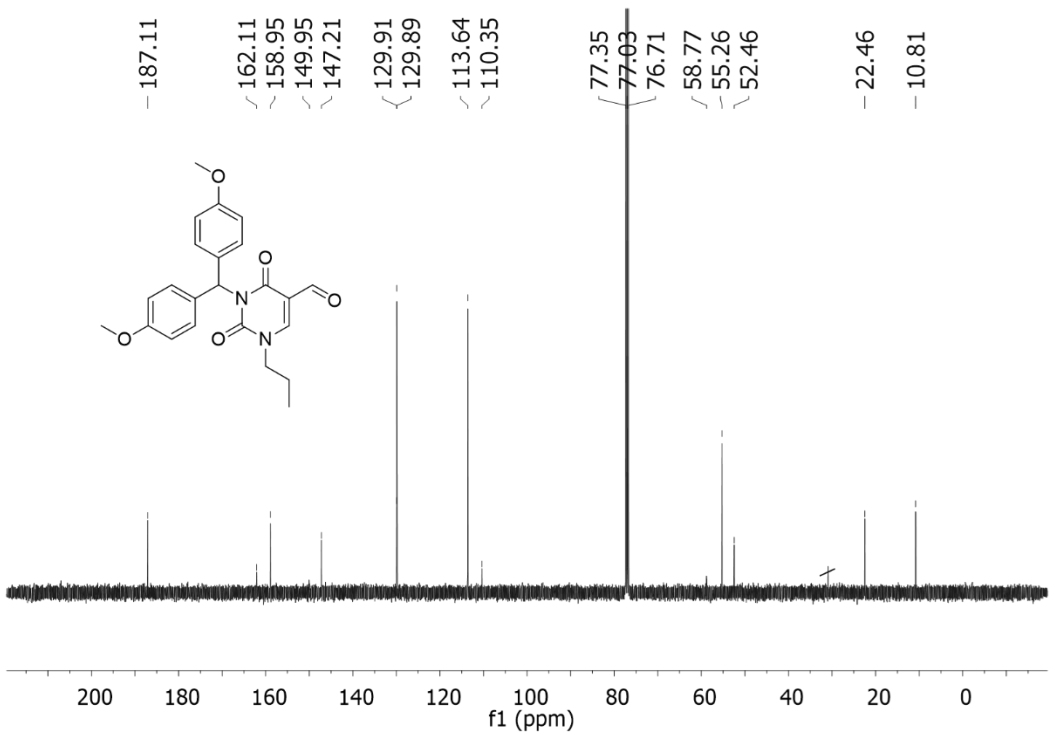


*Figure 4. 45. [13]C NMR (101 MHz, $CDCl_3$) of compound **21b**.*

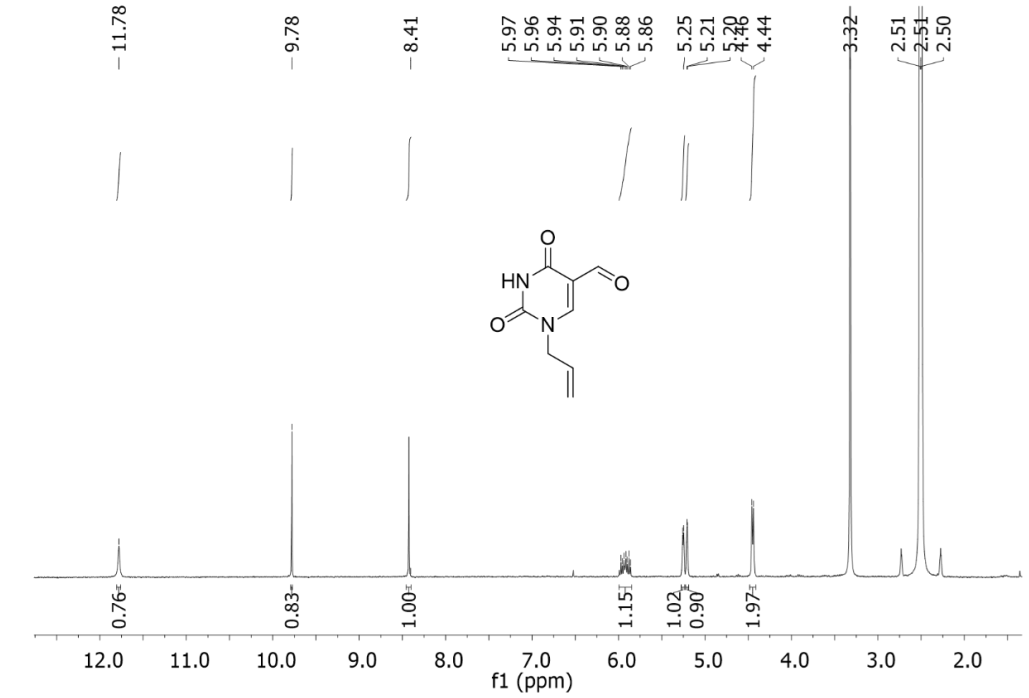


*Figure 4. 46. [1]H NMR (400 MHz, DMSO) of compound **23**.*

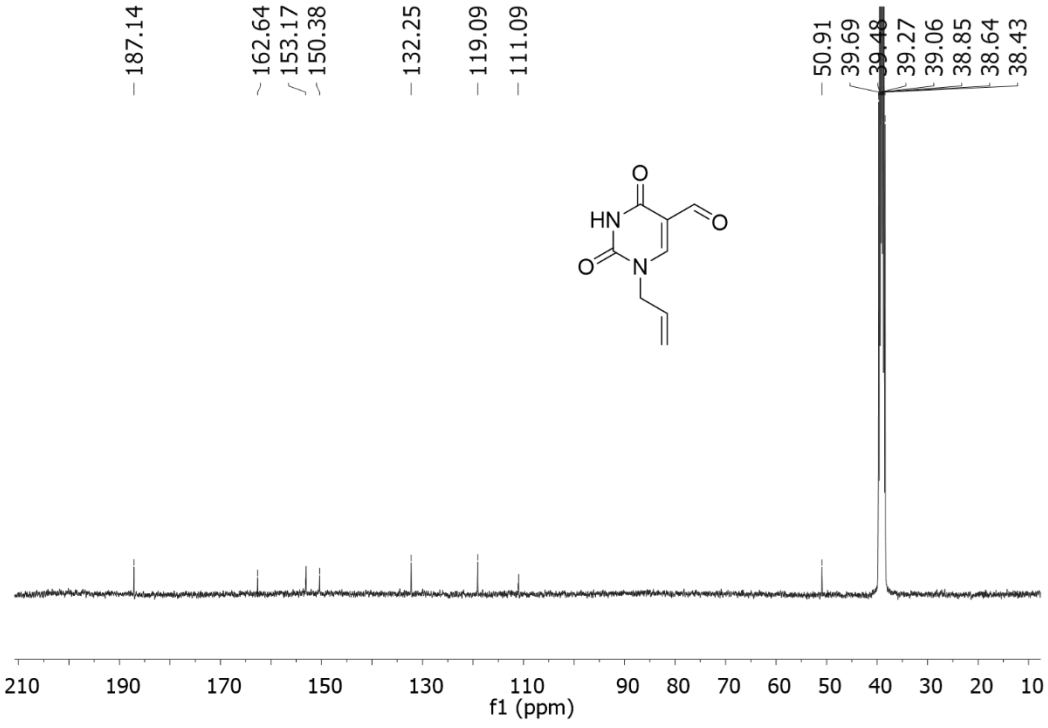


*Figure 4. 47. [13]C NMR (101 MHz, DMSO) of compound **23**.*

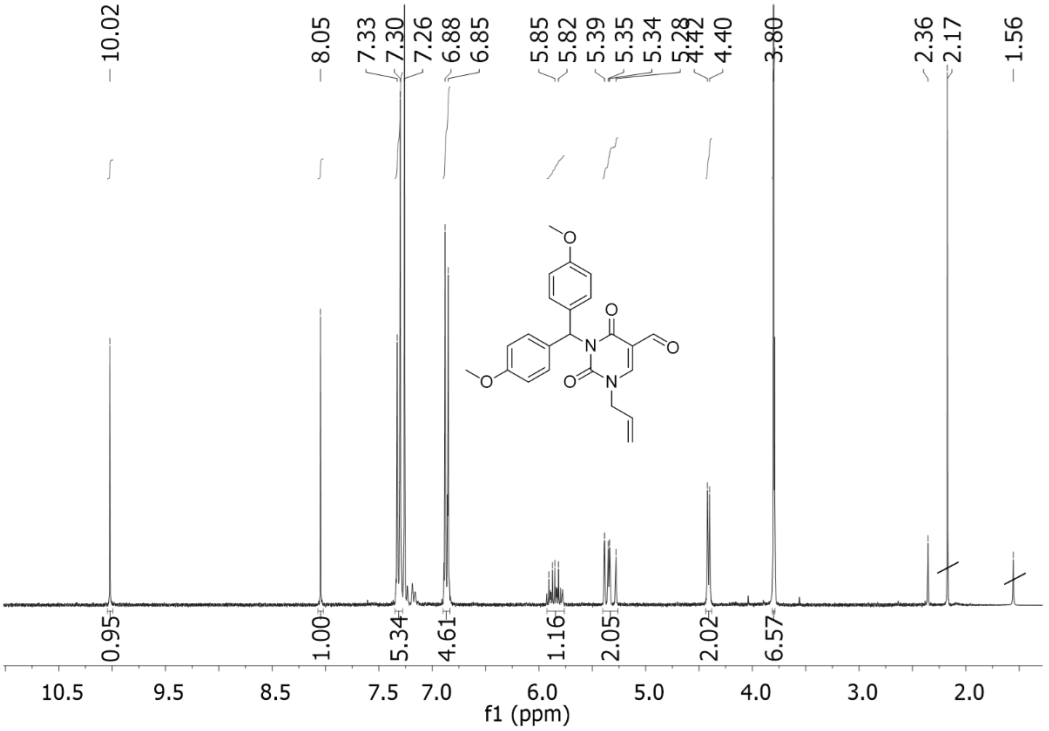


*Figure 4. 48 . ^{1}H NMR (300 MHz, $CDCl_3$) of compound **24**.*

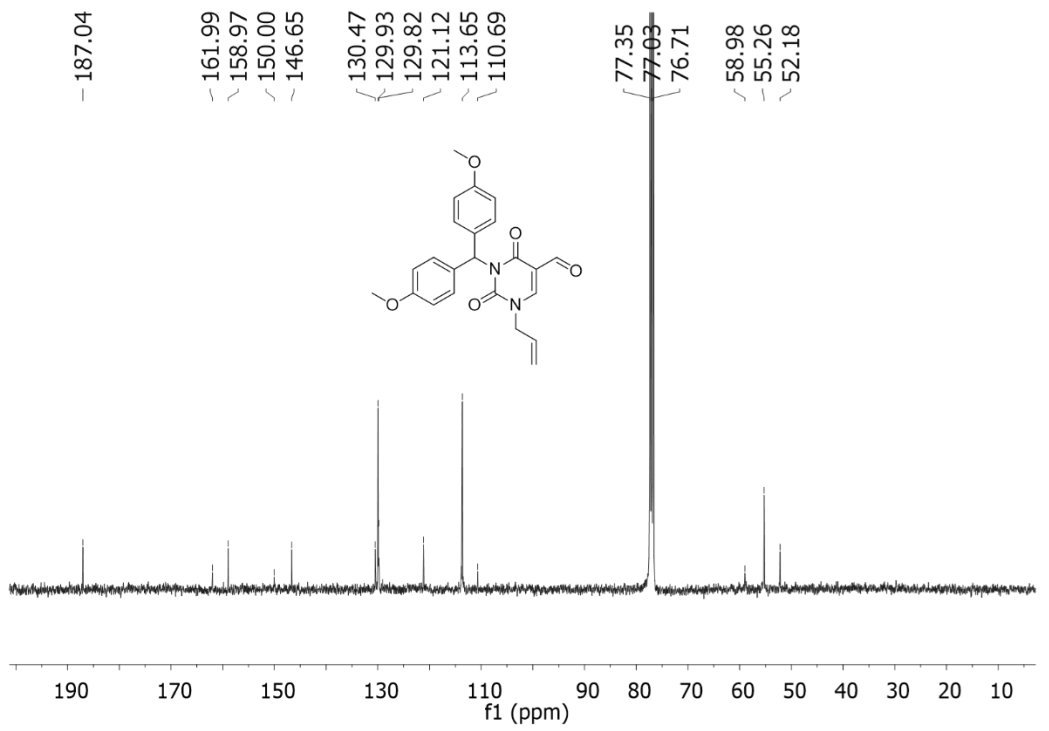


*Figure 4. 49. ^{13}C NMR (75 MHz, $CDCl_3$) of compound **24**.*

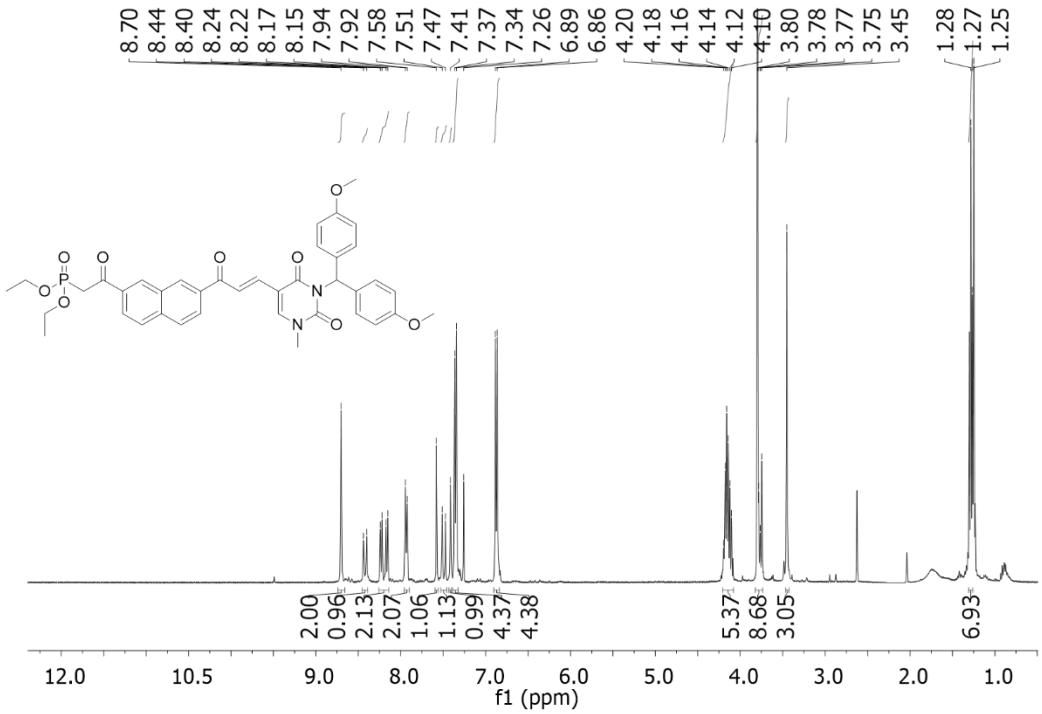


*Figure 4. 50. [1]H NMR (400 MHz, $CDCl_3$) of compound **25a**.*

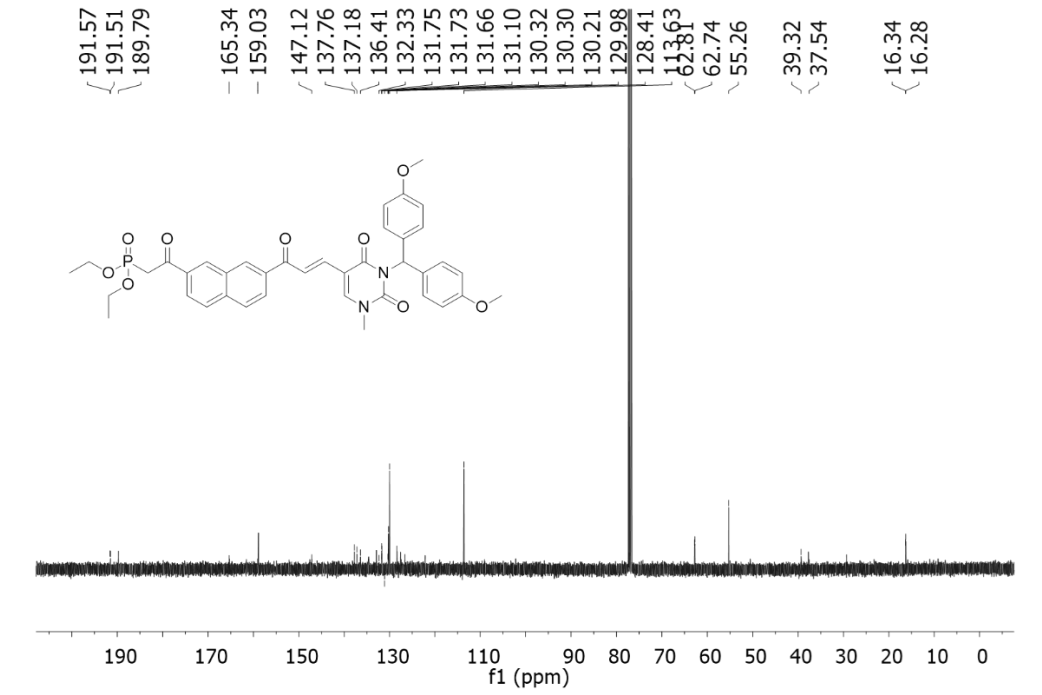


*Figure 4. 51. [13]C NMR (101 MHz, $CDCl_3$) of compound **25a**.*

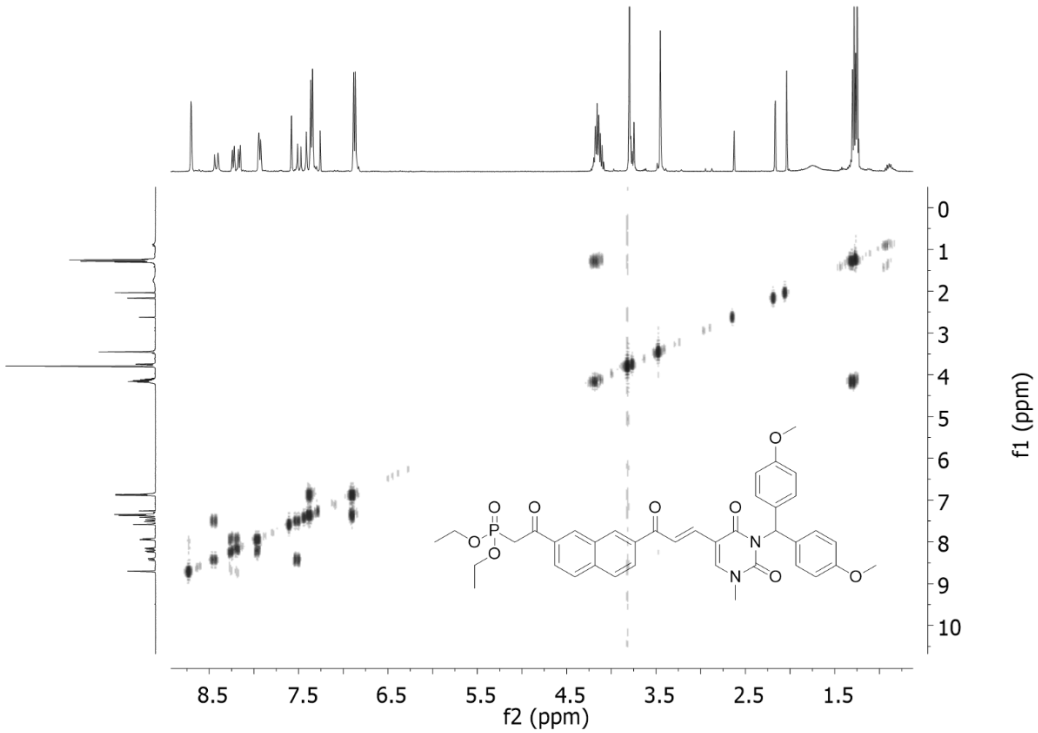


*Figure 4. 52. ^{1}H-^{1}H COSY NMR ($CDCl_3$) of compound **25a**.*

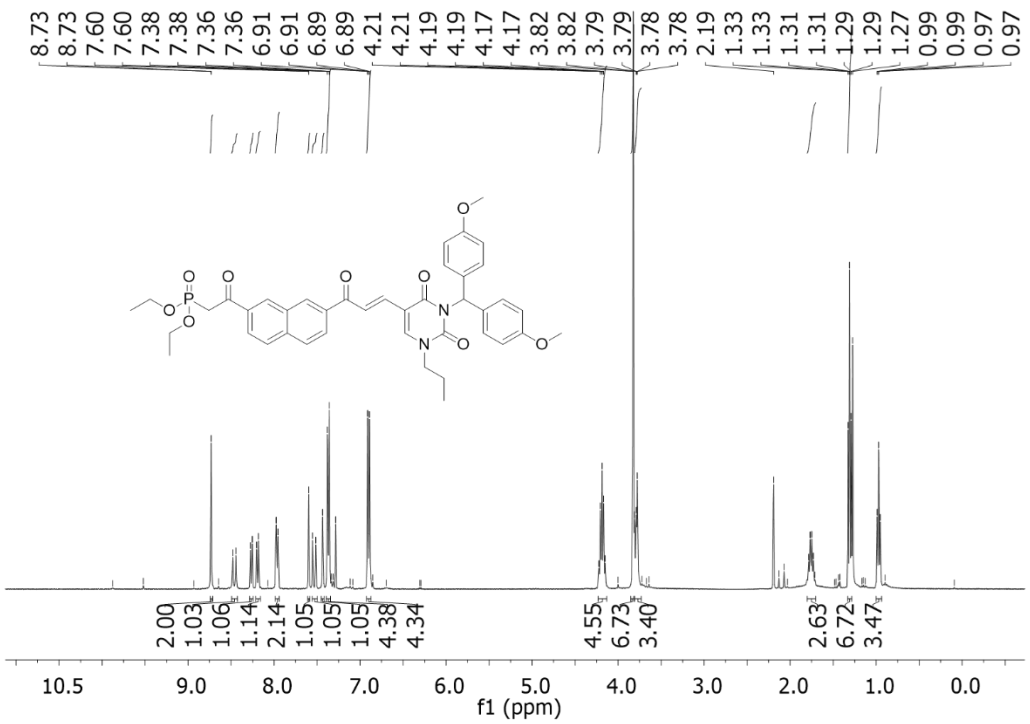


*Figure 4. 53. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **25b***

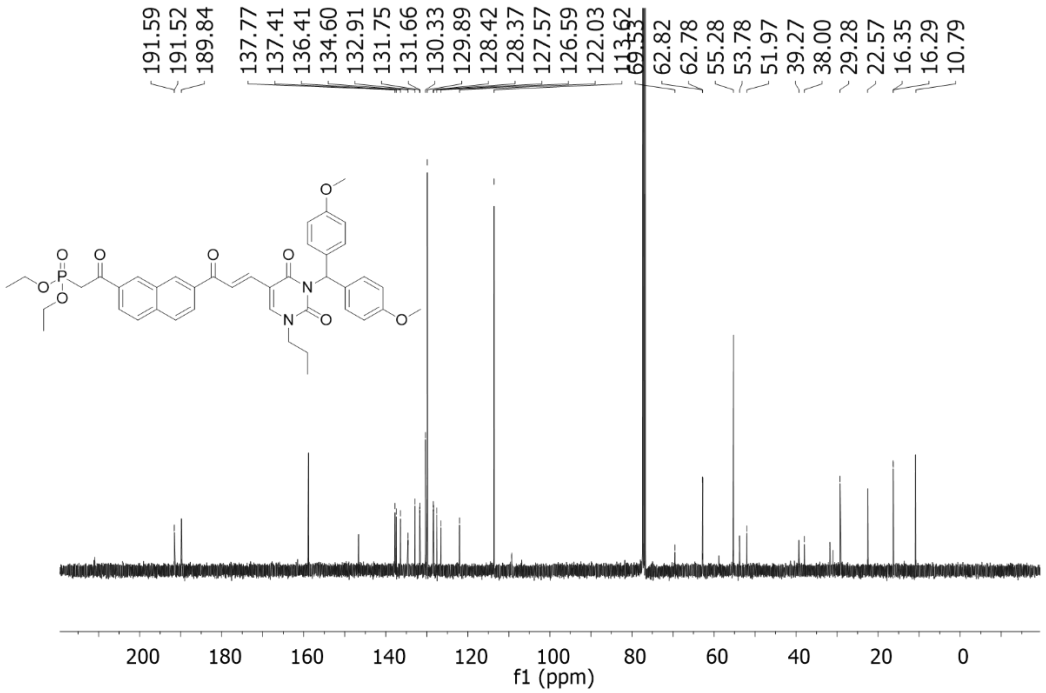


*Figure 4. 54. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **25b***

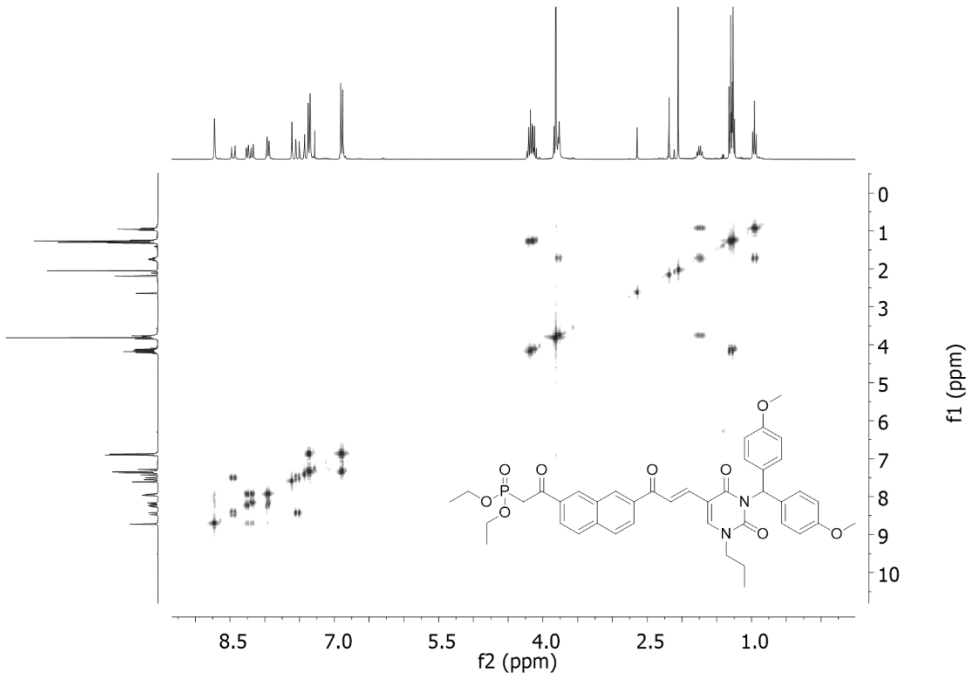


*Figure 4. 55. ^{1}H-^{1}H COSY NMR ($CDCl_3$) of compound **25b***

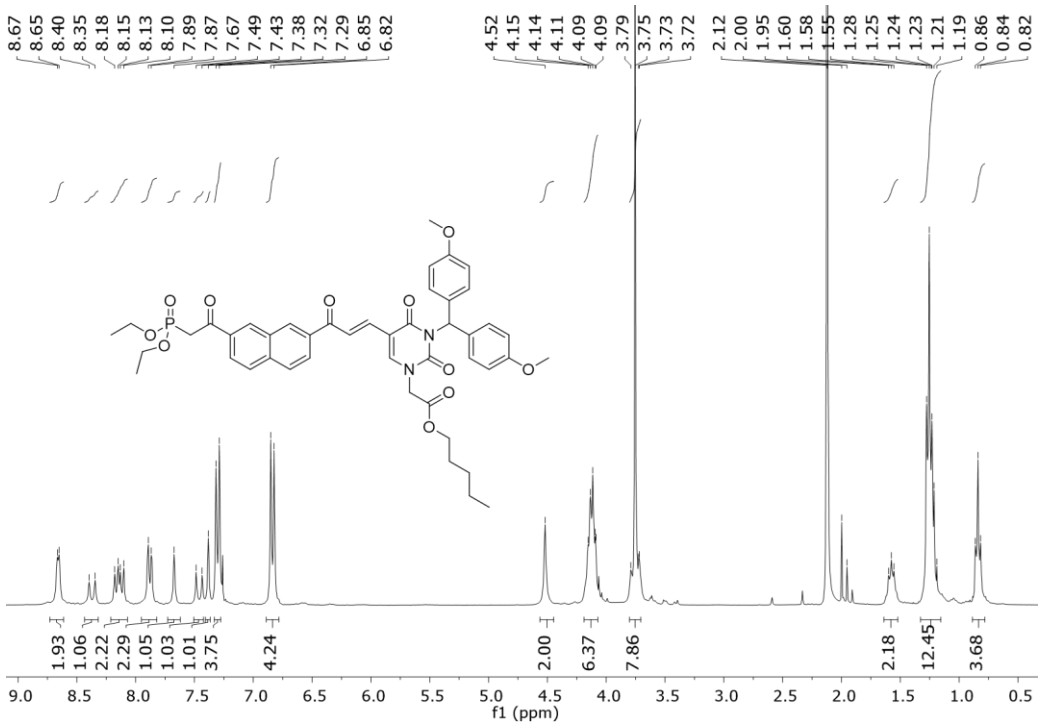


*Figure 4. 56. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **26***

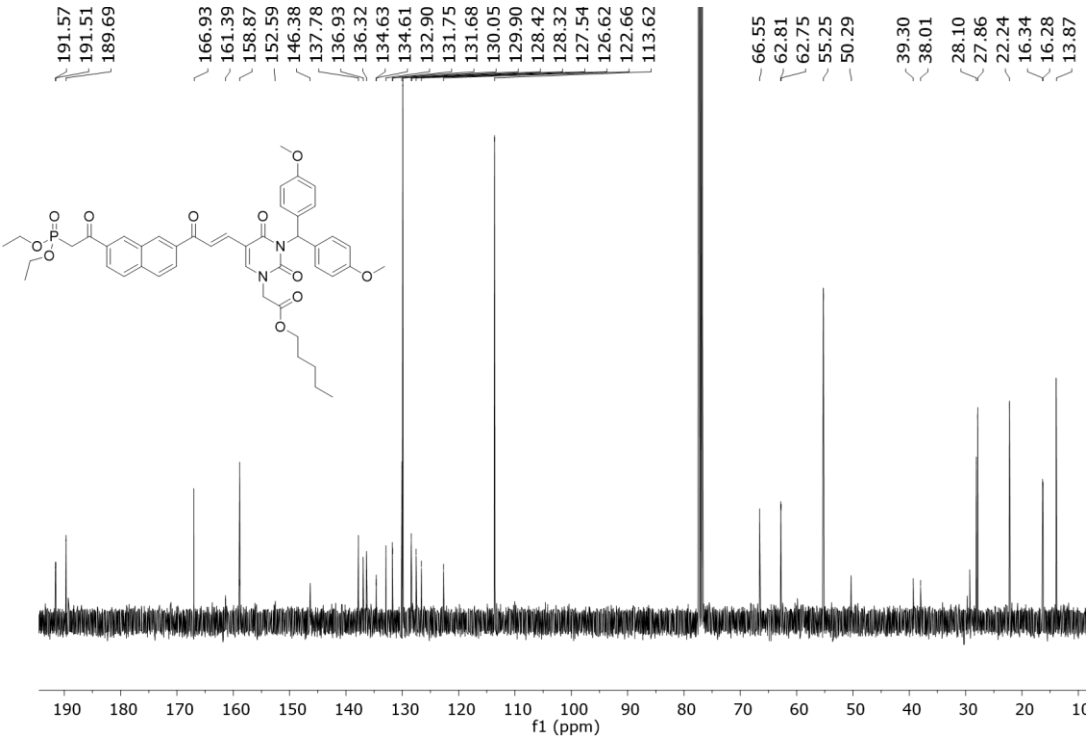


*Figure 4. 57. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **26***

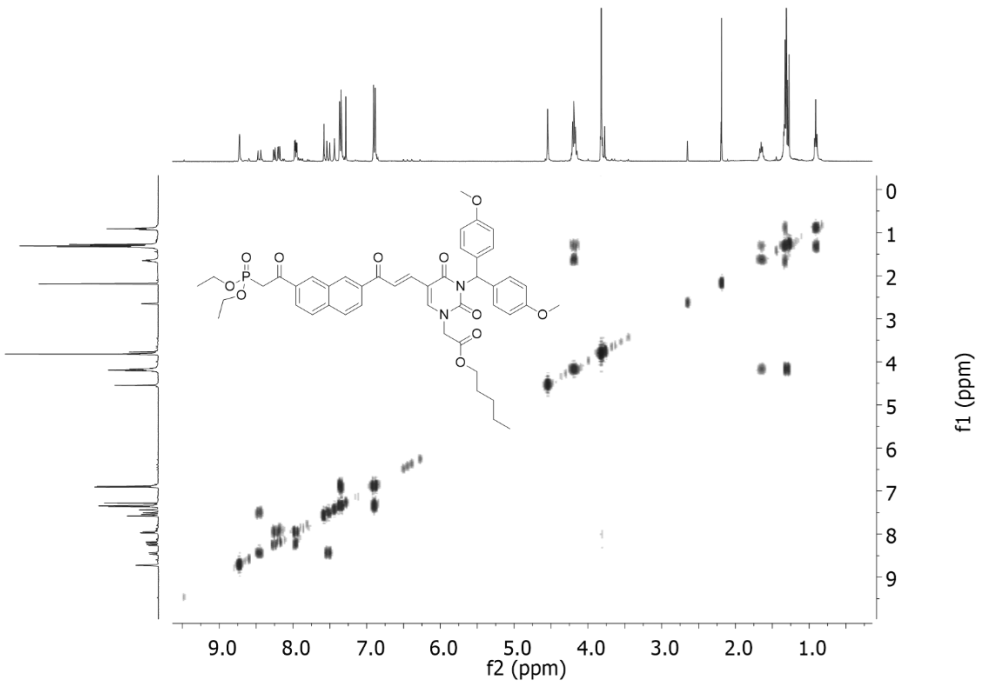


*Figure 4. 58. ^{1}H-^{1}H COSY NMR ($CDCl_3$) of compound **26***

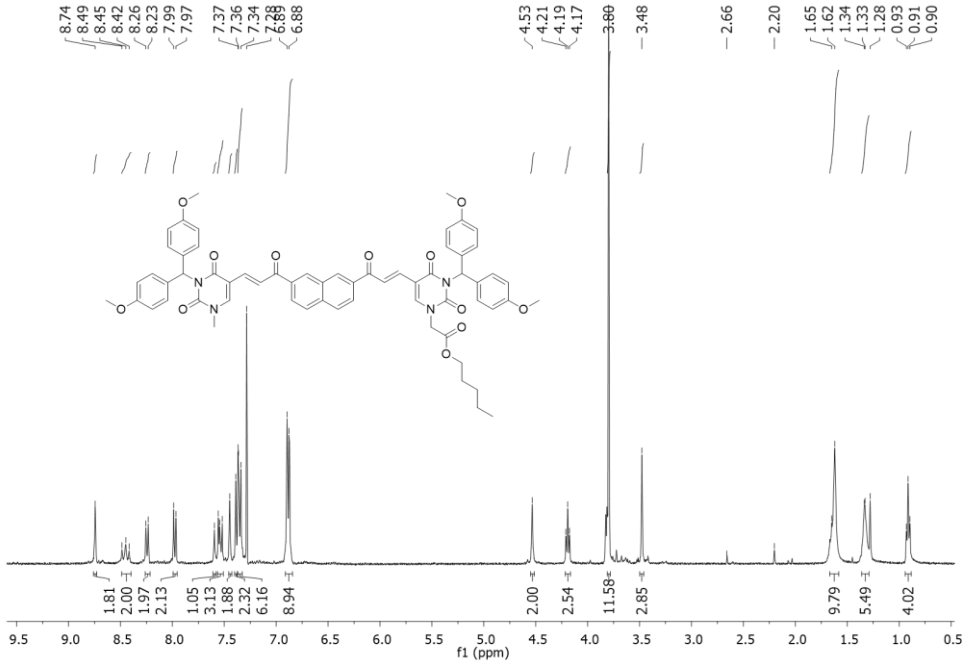


*Figure 4. 59. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **27a***

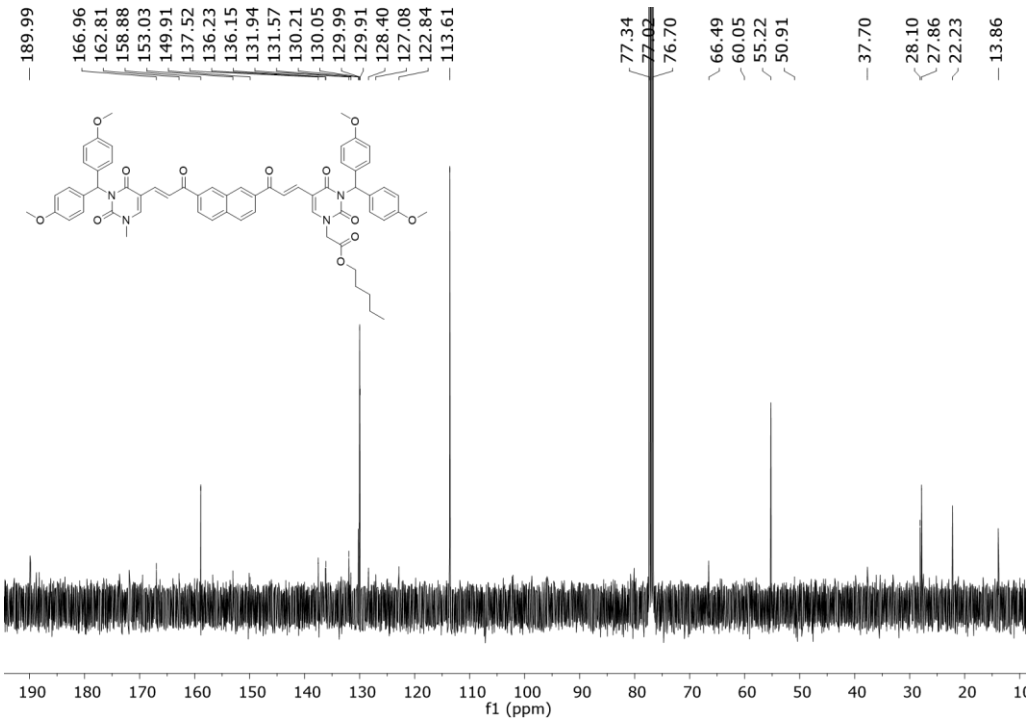


*Figure 4. 60. ^{13}C NMR (101 MHz, $CDCl_3$) of compound **27a***

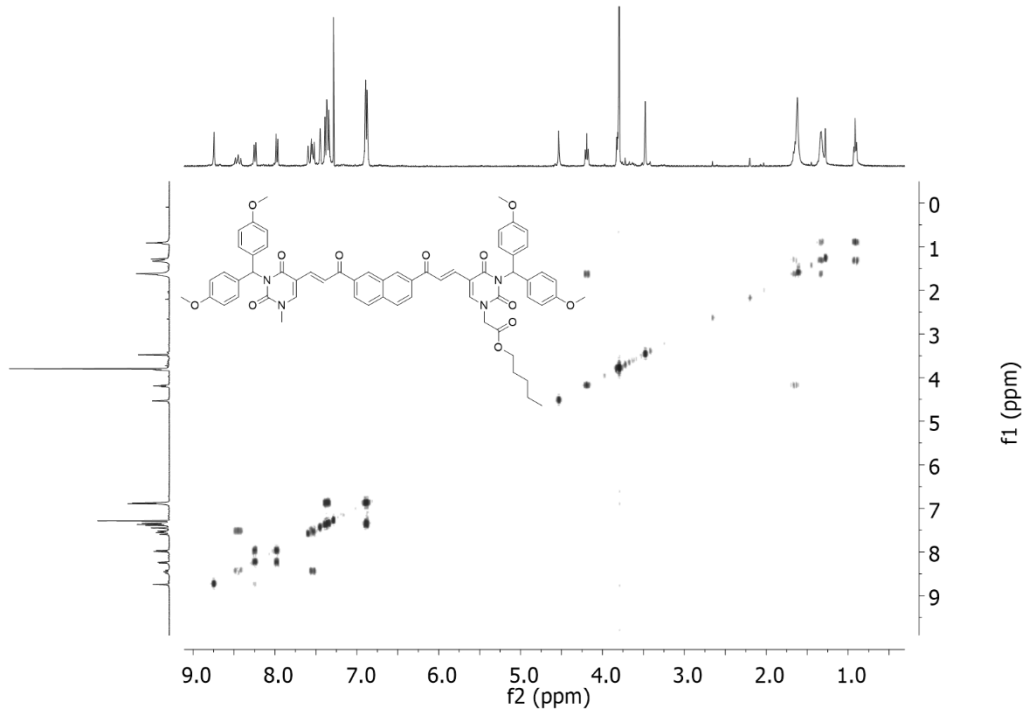


*Figure 4. 61. ^{1}H-^{1}H COSY NMR ($CDCl_3$) of compound **27a***

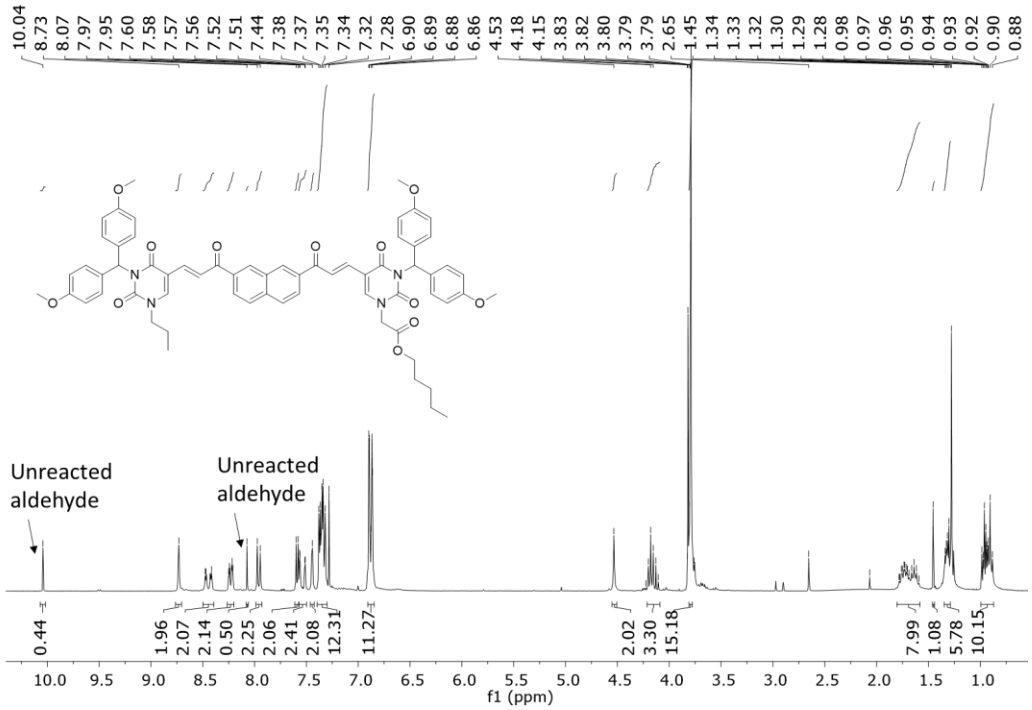


*Figure 4. 62. ^{1}H NMR (400 MHz, $CDCl_3$) of compound **27b***

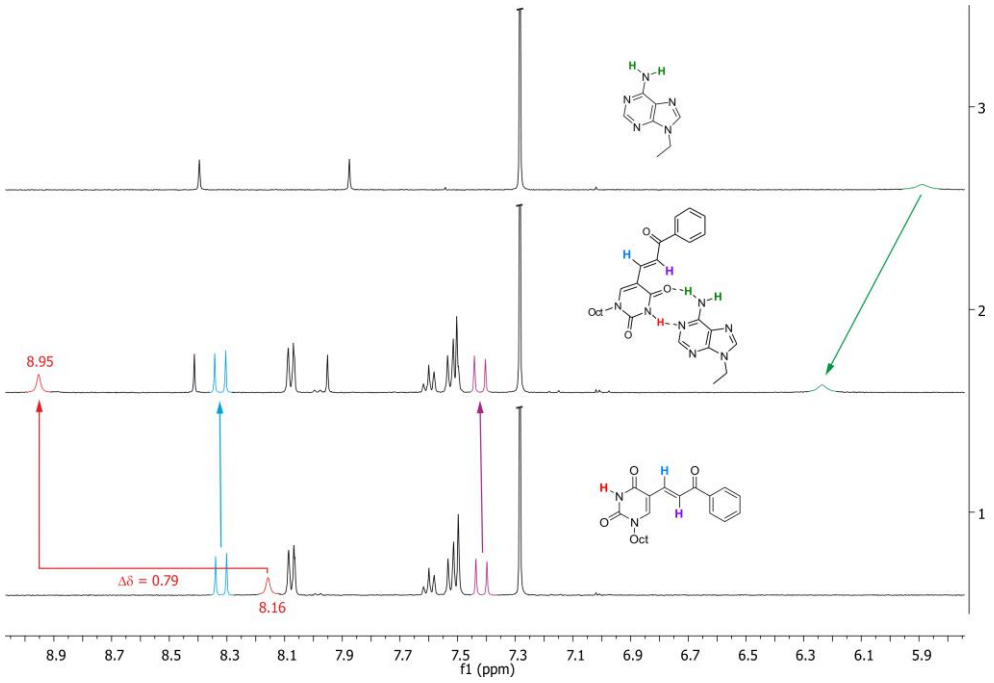


Figure 4.63. ^{1}H NMR spectra of 9-ethyladenine (1.84 μM), monomeric base (3.68 μM, two monomeric base units correspond to one dimeric base unit) and 1:1 mixture of 9-ethyladenine and monomeric base in $CDCl_3$.

4.7 References

1. Ivancová, I., Leone, D. L. & Hocek, M. Reactive modifications of DNA nucleobases for labelling, bioconjugations, and cross-linking. *Curr. Opin. Chem. Biol.* **52,** 136–144 (2019).

2. Hocek, M. Enzymatic Synthesis of Base-Functionalized Nucleic Acids for Sensing, Cross-linking, and Modulation of Protein-DNA Binding and Transcription. *Acc. Chem. Res.* **52,** 1730–1737 (2019).

3. Xu, W., Chan, K. M. & Kool, E. T. Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **9,** 1043–1055 (2017).

4. Spiridoula, M. in *Photoinduced Phenom. Nucleic Acids I. Top. Curr. Chem.* (eds. Barbatti, M., Borin, A. C. & Ullrich, S.) **355,** 209–243 (Springer International Publishing, 2014).

5. Hocek, M. & Fojta, M. Nucleobase modification as redox DNA labelling for electrochemical detection. *Chem. Soc. Rev.* **40,** 5802–5814 (2011).

6. Manicardi, A., Rozzi, A., Korom, S. & Corradini, R. Building on the peptide nucleic acid (PNA) scaffold : a biomolecular engineering approach. *Supramol. Chem.* **0278,** 1–12 (2017).

7. Wojciechowski, F. & E. Hudson, R. Nucleobase Modifications in Peptide Nucleic Acids. *Curr. Top. Med. Chem.* **7,** 667–679 (2007).

8. Hirao, I., Kimoto, M. & Yamashige, R. Natural versus artificial creation of base pairs in DNA: Origin of nucleobases from the perspectives of unnatural base pair studies. *Acc. Chem. Res.* **45,** 2055–2065 (2012).

9. Karalkar, N. B. & Benner, S. A. The challenge of synthetic biology. Synthetic Darwinism and the aperiodic crystal structure. *Curr. Opin. Chem. Biol.* **46,** 188–195 (2018).

10. Feldman, A. W. & Romesberg, F. E. Expansion of the Genetic Alphabet: A Chemist's Approach to Synthetic Biology. *Acc. Chem. Res.* **51,** 394–403 (2018).

11. Saftić, D., Ban, Ž., Matić, J., Tumirv, L.-M. & Piantanida, I. Conjugates of Classical DNA/RNA Binder with Nucleobase: Chemical, Biochemical and Biomedical Applications. *Curr. Med. Chem.* **26,** 5609–5624 (2018).

12. Silverman, R. B. & Holladay, M. W. in *Org. Chem. Drug Des. Drug Action (Third Ed.* (eds. Silverman, R. B. & Holladay, M. W.) 275–331 (Academic Press, 2014). doi:https://doi.org/10.1016/B978-0-12-382030-3.00006-4

13. Fox, K. R., Brown, T. & Rusling, D. A. in *DNA-targeting Mol. as Ther. Agents* 1–32 (The Royal Society of Chemistry, 2018). doi:10.1039/9781788012928-00001

14. Wojciechowska, M., Równicki, M., Mieczkowski, A., Miszkiewicz, J. & Trylska, J. Antibacterial Peptide Nucleic Acids—Facts and Perspectives. *Molecules* **25,** 559–565 (2020).

15. McHale, R. & OReilly, R. K. Nucleobase containing synthetic polymers: Advancing biomimicry via controlled synthesis and self-assembly. *Macromolecules* **45,** 7665–7675 (2012).

16. Peters, G. M. & Davis, J. T. Supramolecular gels made from nucleobase, nucleoside and nucleotide analogs. *Chem. Soc. Rev.* **45,** 3188–3206 (2016).

17. Pu, F., Ren, J. & Qu, X. Nucleobases, nucleosides, and nucleotides: Versatile biomolecules for generating functional nanomaterials. *Chem. Soc. Rev.* **47,** 1285–1306 (2018).

18. Seidman, M. M. & Glazer, P. M. The potential for gene repair via triple helix formation. *J. Clin. Invest.* **112,** 487–494 (2003).

19. Rusling, D. A. & Fox, K. R. Sequence-specific recognition of DNA nanostructures. *Methods* **67,** 123–133 (2014).

20. Potaman, V. N. Applications of triple-stranded nucleic acid structures to DNA purification, detection and analysis. *Expert Rev. Mol. Diagn.* **3,** 481–496 (2003).

21. Devi, G., Zhou, Y., Zhong, Z., Toh, D. F. K. & Chen, G. RNA triplexes: From structural principles to biological and biotech applications. *Wiley Interdiscip. Rev. RNA* **6,** 111–128 (2015).

22. Malnuit, V., Duca, M. & Benhida, R. Targeting DNA base pair mismatch with artificial nucleobases. Advances and perspectives in triple helix strategy. *Org. Biomol. Chem.* **9,** 326–336 (2011).

23. Brodyagin, N., Hnedzko, D., MacKay, J. A. & Rozners, E. in *Pept. Nucleic Acids Methods Protoc.* (ed. Nielsen, P. E.) 157–172 (Springer US, 2020). doi:10.1007/978-1-0716-0243-0_9

24. Lin, K. Y. & Matteucci, M. D. A cytosine analogue capable of clamp-like binding to a guanine in helical nucleic acids [8]. *J. Am. Chem. Soc.* **120,** 8531–8532 (1998).

25. Rajeev, K. G., Maier, M. A., Lesnik, E. A. & Manoharan, M. High-Affinity

Peptide Nucleic Acid Oligomers Containing Tricyclic Cytosine Analogues. *Org. Lett.* **4,** 4395–4398 (2002).

26. Ortega, J. A., Blas, J. R., Orozco, M., Grandas, A., Pedroso, E. & Robles, J. Binding affinities of oligonucleotides and PNAs containing phenoxazine and G-clamp cytosine analogues are unusually sequence-dependent. *Org. Lett.* **9,** 4503–4506 (2007).

27. Wojciechowski, F. & Hudson, R. H. E. Peptide nucleic acid containing a meta-substituted phenylpyrrolocytosine exhibits a fluorescence response and increased binding affinity toward RNA. *Org. Lett.* **11,** 4878–4881 (2009).

28. Rebek, J., Askew, B., Ballester, P., Buhr, C., Costero, A., Jones, S. & Williams, K. Molecular Recognition: Watson-Crick, Hoogsteen, and Bifurcated Hydrogen Bonding in a Model for Adenine Recognition. *J. Am. Chem. Soc.* **109,** 6866–6867 (1987).

29. Askew, B., Ballester, P., Buhr, C., Jeong, K. S., Jones, S., Parris, K., Williams, K. & Rebek, J. Molecular Recognition with Convergent Functional Groups. 6. Synthetic and Structural Studies with a Model Receptor for Nucleic Acid Components. *J. Am. Chem. Soc.* **111,** 1082–1090 (1989).

30. Jeong, K. S., Tjivikua, T., Muehldorf, A., Deslongchamps, G., Famulok, M. & Rebek, J. Convergent Functional Groups.10. Molecular Recognition of Neutral Substrates. *J. Am. Chem. Soc.* **113,** 201–209 (1991).

31. Rebek, J., Askew, B., Ballester, P., Buhr, C., Jones, S., Nemeth, D. & Williams, K. Molecular Recognition: Hydrogen Bonding and Stacking Interactions Stabilize a Model for Nucleic Acid Structure. *J. Am. Chem. Soc.* **109,** 5033–5035 (1987).

32. Faraoni, R., Blanzat, M., Kubicek, S., Braun, C., Schweizer, W. B., Gramlich, V. & Diederich, F. New Rebek imide-type receptors for adenine featuring acetylene-linked π-stacking platforms. *Org. Biomol. Chem.* **2,** 1962–1964 (2004).

33. Accetta, A., Corradini, R., Sforza, S., Tedeschi, T., Brognara, E., Borgatti, M., Gambari, R. & Marchelli, R. New uracil dimers showing erythroid differentiation inducing activities. *J. Med. Chem.* **52,** 87–94 (2009).

34. Betts, L., Josey, J. A., Veal, J. M. & Jordan, S. R. A Nucleic Acid Triple Helix Formed by a Peptide Nucleic Acid-DNA Complex. *Science (80-.).* **270,** 1838–1841 (1995).

35. Nielsen, P. E. Peptide nucleic acid. A molecule with two identities. *Acc. Chem. Res.* **32,** 624–630 (1999).

36. Griffith, M. C., Risen, L. M., Greig, M. J., Lesnik, E. A., Sprankle, K. G., Griffey, R. H., Kiely, J. S. & Freier, S. M. Single and Bis Peptide Nucleic Acids as Triplexing Agents: Binding and Stoichiometry. *J. Am. Chem. Soc.* **117,** 831–832 (1995).

37. Accetta, A. *Molecular Engineering of PNA Using Modified Uracil Derivatives and Porphyrins*. (Università degli Studi di Parma, PhD Thesis, 2010).

38. Verona, M. D., Verdolino, V., Palazzesi, F. & Corradini, R. Focus on PNA Flexibility and RNA Binding using Molecular Dynamics and Metadynamics. *Sci. Rep.* **7,** 1–11 (2017).

39. Bassetti, M., Ciceri, S., Lancia, F. & Pasquini, C. Hydration of aromatic terminal alkynes catalyzed by iron(III) sulfate hydrate under chlorine-free conditions. *Tetrahedron Lett.* **55,** 1608–1612 (2014).

40. Holt, D. A., Levy, M. A., Ladd, D. L., Oh, H. J., Erb, J. M., Heaslip, J. I., Brand, M. & Metcalf, B. W. Steroidal a Ring Aryl Carboxylic Acids: A New Class of Steroid 5α-Reductase Inhibitors. *J. Med. Chem.* **33,** 937–942 (1990).

41. Somu, R. V., Boshoff, H., Qiao, C., Bennett, E. M., Barry, C. E. & Aldrich, C. C. Rationally-designed nucleoside antibiotics that inhibit siderophore biosynthesis of Mycobacterium tuberculosis. *J. Med. Chem.* **49,** 31–34 (2006).

42. Itahara, T., Fujii, Y. & Tada, M. Oxidation of Thymines by Peroxosulfate Ions in Water. *J. Org. Chem.* **53,** 3421–3424 (1988).

43. Hudson, R. H. E., Liu, Y. & Wojciechowski, F. Hydrophilic modifications in peptide nucleic acid - Synthesis and properties of PNA possessing 5-hydroxymethyluracil and 5-hydroxymethylcytosine. *Can. J. Chem.* **85,** 302–312 (2007).

44. Lombardo, C. M., Welsh, S. J., Strauss, S. J., Dale, A. G., Todd, A. K., Nanjunda, R., Wilson, W. D. & Neidle, S. A novel series of G-quadruplex ligands with selectivity for HIF-expressing osteosarcoma and renal cancer cell lines. *Bioorg. Med. Chem. Lett.* **22,** 5984–5988 (2012).

45. Katz, T. J., Liu, L., Willmore, N. D., Fox, J. M., Rheingold, A. L., Shi, S., Nuckolls, C. & Rickman, B. H. An Efficient Synthesis of Functionalized Helicenes. *J. Am. Chem. Soc.* **119,** 10054–10063 (1997).

46. Eschenfelder, V. & Brossmer, R. Mono-, Di- und Triallyliertes 5-(Hydroxymethyl)uracil. *Justus Liebigs Ann. Chem.* **1977,** 1969–1975 (1977).

Acknowledgements

Here we are. The end of these challenging three years has finally (?) arrived.

First of all, I would like to thank the sponsor, Cariparma, and the supervisors of this PhD project.
Prof. Roberto Corradini for giving me the opportunity to join his lab, for his patient guidance in the world of PNA technology and for encouraging me to step outside my comfort zone and always prove myself. Prof. Wolfgang Knoll, that despite his busy time has been always kind and caring for me and my work.

Secondly, I would like to thank Prof. Michael Sailor for hosting me in his lab in the beautiful city of San Diego. The Summer School for Silicon Nanotechnology is a powerful opportunity both in terms of scientific and individual growth.

I would also like to thank the Research and Innovation Staff Exchange (RISE) program and all the people involved with it that allowed me to be a part of this wonderful experience. Indeed, I want to thank Prof. Alessandro Bertucci (Sandro). A familiar shape for a scared first-year PhD student so far from home. A mentor, friendly and yet demanding. I can't forget that look on your face after some of my 'cavolata'. Thank you to "the Italian gang", Prof. Alessandro Porchetta, Elisa, Mariannina (my partner in crime), Jacopo (the sleepy in the bathtube) and Chiara, that shared with me from lectures about porous silicon to pool parties at Porchetta's place. I can definitely say that it would not have been the same without you all.

I would like to thank Prof. Jakub Dostalek that supervised my work at Austrian Institute of Technology (AIT). Last year pandemic situation was unexpected and complicated for anyone but nevertheless you found time to discuss with me and help me with my time-limited experiments. For the same reasons, a big 'thank you' goes to the BST/CEST family. Thank you Chiara and Valeria for hosting me without hesitation (I owe you a dinner, hopefully my next time in Vienna) and again Anil, Jiao, Patrick, Simone, Bernadette, Paolo, Jakob, Agathe, Maddalena, Simone, Filippo (thanks for the desk), Sandra, Matthias, Katarina, Roger, Stefan, Ciril. I wish I could have spent more time with you all.

I would also like to thank the SCVSA family. Starting from former and current lab mates, thanks Andre for your outstanding patience, Saša, Silvia, Domi, Luca, Michelino, Red, Sabraina, I am glad to leave you the flume hood, Fede, the kindest person I ever met and last but not least Ste, joy and sorrow. Thank you for the stinkbugs you got rid of in my place and the ones you 'gifted' me but mostly for your dedicated support during these years in the lab and outside.
I want to thank the girls of "Il Lazzaretto" (pretty accurate name). Silvia, Fra, Angie, Claudia (I will never forget your pasticcio ferrarese), Cri, Giulia, Marghe (since day one) for the chat, the aperitivi, the laughs, the emotional support and the endless complaints you had to hear from me. The same for Nico and Simo for staying even during the armocromia chats (please don't forget your seasons). Thanks Spez and Ema, it seems a hundred years ago from our weekly dinners but they will always remain in my heart.
I know, it was not easy to welcome this weird 'terrona' in your lives (by the way Zagarolo is central Italy) but I am glad you did.

I am also thankful for all the people I met over my undergraduate years. Filli, Edo, Lollo, Pier, Fede, Georgette, Danish, Ruggio, Sara, Capu, Laura, because anytime I come back home it is like nothing changed.

I would like to thank a special person. JZ, you have been a friend, a partner, a co-author (hopefully), my biggest supporter throughout these years. Your love and your unshakeable faith in me have been essential to make it to the end of this pathway.

How to forget my beloved long-time friends. Giuli, Rosi, Bè, you are pieces of my heart. Although the scientific life split us apart the certainty of our friendship has never wavered. Fra, after all these years (29, if I am not wrong) you are still here to celebrate together this goal and the ones will be next.

Lastly, I am extremely thankful to my family. Tutto questo non sarebbe stato possibile senza il vostro incondizionato supporto. In tutti questi anni la vostra fiducia mai una volta ha vacillato e discretamente con amore mi siete stati accanto in ogni mia scelta. Mamma, Papà, Ila, Nonna, grazie per aver creduto in me anche quando nemmeno io ci credevo.

Vi abbraccio tutti con affetto.