Università degli Studi di Parma

Dipartimento di Chimica Organica ed Industriale

Dottorato di Ricerca in Scienze Chimiche

Ciclo XXII

Triennio 2007-2009

Molecular Engineering of PNA Using Modified Uracil Derivatives and Porphyrins

Tutor:

Prof. *Rosangela Marchelli*

Prof. *Roberto Corradini*

Dottorando:

Alessandro Accetta

2010

Abstract:

PNAs (peptide nucleic acids) are among the most powerful tools in molecular biology and biotechnology, also for their versatility to be changed by chemical modifications. In the first part of thesis we describe the design and synthesis of uracil dimers, accurately designed to reinforce binding ability of uracil. Symmetric dimers were found to be biologically active, showing antiproliferative effect and erytroid differentiation inducing activities. Dissymmetric dimers were incorporated into PNA oligomers and tested toward complementary and mismatched DNA. In both case we found a cooperativity between the two uracil moieties In the second part we explored the use of porphyrins as spectroscopic sensor for structure assessment. Preliminary result showed a strong influence of porphyrins on PNA duplex conformation, giving rise to unexplored PNA conformations.

Keyword: PNAs (peptide nucleic acids), Uracil dimers, Modified nucleobases, Adenine recognition, Erytroid differentiation, Antiproliferative effect, Porphyrins, parallel PNA:PNA.

Riassunto:

PNAs (acidi peptido nucleici) sono tra i più potenti strumenti utilizzati in biologia molecolare e in applicazioni biotecnologiche grazie anche alla possibilità di modificarne chimicamente la struttura. Nella prima parte della tesi è stata realizzata la sintesi di dimeri di uracile, progettati per aumentare le capacità di riconoscimento dell'adenina. Dimeri simmetrici hanno mostrato delle proprietà biologiche interessanti quali attività antiproliferativa e proprietà di differenziamento cellulare tipo eritroide. Alcuni dimeri dissimetrici sono stati incorporati come basi modificate in oligomeri di PNA di cui è stata valutata affinità e selettività nei confronti di DNA complementare fullmatch and mismatch. Sia per quanto riguarda i dimeri simmetrici che per le basi modificate è stata osservata una cooperatività tra le due unità di uracile. Nella seconda parte sono state utilizzate delle tetrapyrydyl porpfirine come gruppi reporter dell'elicità supramolecolare. Risultati preliminari hanno mostrato una forte influenza delle porfirine sulla conformazione del PNA:PNA parallelo dando origine a conformazioni non ancora conosciute.

Parole Chiavi: PNAs (acidi peptido nucleici), Dimeri di Uracile, Basi modificate, Riconoscimento di adenina, Differenziamento eritroide, Effetto antiproliferativo, Porfirine, PNA:PNA parallelo.

Contents

CHAPTER 2

New Uracil Dimers Showing Erytroid Differentiation Inducing Activities

CHAPTER 3

Design and Synthesis of "Tailor Made" Modified Uracil derivatives for Selective Adenine Recognition

CHAPTER 4

Synthesis and Properties of PNA Oligomers Based on Engineered Uracil Derivatives

CHAPTER 5

Porphyrins – PNAs Conjugates: Tools to Study and to Control Helicity in PNA Containing Duplexes

CHAPTER 1

Modified Nucleobases and Peptide Nucleic Acids

1.1 Preface

During the past centuries a great part of the work of organic chemists have been devoted to the isolation and purification of natural products from biological systems and determining their structures. As their knowledge grew, chemists focussed on even more large biomolecules such as proteins, nucleic acids and their supramolecular assembly in more complex structures. Indeed the important breakthrough that allowed the elucidation of the double helical structure of DNA occurred when it was realized that there are two separate strands of nucleotides connected through hydrogen bonds[1]. The knowledge of the DNA structure opened the possibility to understand, at the molecular level, not only the mechanism of storage and expression of the genetic information, but also the bases of the interactions of DNA with RNA, proteins, enzymes, small molecules and providing elements for designing new DNA binding molecules.

Progresses in organic synthesis have also provided very important tools in the manipulation of biological systems. For example, organic chemists provided easy and efficient phosphoramidite (or H-phosphonate) solid phase synthesis[2] for producing artificial DNA molecules, and this was one of the key elements in the outburst of modern biotechnology and molecular biology. Synthetic oligonucleotides (ODN) were used not only to substitute natural DNA sequences, but also found to alter specifically gene expression, as shown in seminal studies by Stephenson and Zamecnik[3], who used an *antisense* DNA oligonucleotide (i.e. an artificial DNA tract complementary to a specific RNA) to prevent translation of the Rous Sarcoma virus RNA. Later on the possibility to directly hit a gene on double stranded DNA (dsDNA) opened the road to the so-called *antigene approach*. Antisense and antigene strategies helped molecular biologists to discover new gene functions and to lay the base for future gene therapy approaches. However, the application of unmodified oligonucleotides in

several *in vivo* applications was limited by their poor stability and bioavailability. The negatively charged phosphate backbone is recognized as largely responsible for the unsuitability of DNA for gene targeting, limiting the cellular uptake and favouring degradation by naturally occurring nucleases[4]. These drawbacks stimulated the quest for better modified probes[5] obtained from sugar phosphate backbone modifications, or even replacing the deoxyribose ring with alternative structures, leading to a variety of modified nucleic acids and analogues; however, only few of the proposed variations turned out to have improved properties when compared with unmodified ODNs, though in many cases they allowed to better understand DNA properties and function. The backbone of ODNs and of their modifications turned out to be not simply a rigid scaffold to hold the nucleobases in place, but also an important recognition element that is central to gene expression and function.

Thus "*DNA organic chemistry*" stimulated, during the years, the biology discoveries and *vice versa* biology demanded to organic chemistry and supramolecular chemistry big efforts to create complex systems that mimic the natural ones, in what is called "*biomimetic chemistry*"[6].

More recently, the broader approach of *"synthetic biology"* , i.e. the synthesis of unnatural synthetic organic molecules able to perform functions in living systems has become increasingly important:[7] "*Synthetic biology has a broader scope, however, in that it attempts to recreate in unnatural chemical systems the emergent properties of living systems, including inheritance, genetics and evolution" (Benner S.A., 2003)*[8].

The chemical approach to synthetic biology is to redesign basic molecules of life (nucleic acids, amino acids and carbohydrates) and insert them back into living systems, in order to explore the functions and the mechanisms of natural systems[9]. Redesign of nucleic acids is an important pillar of the synthetic biology approach.

In this chapter we describe an overview of the most significant and updated literature on nucleic acid modifications, focusing on base modifications as well as peptide nucleic acids (PNAs), in particular on the chemical modifications of PNAs, enabling rational "engineering" of their properties.

1.2 Modified Oligonucleotides

The need of oligonucleotide modifications came from the unsuitability of DNA for gene targeting due to the instability of the phosphate linkage in biological fluids. Among the first modified oligonucleotide phosphorothiates and methylphosphonates, in which one sulphur atom or methyl groups replacing one oxygen of the phosphate internucleotidic linkage (*Figure 1.1a-c*), were particularly important; these derivatives were able to recognize complementary DNA with improved resistance to nucleases. Moreover, the replacement of one oxygen of the phosphate group with a sulfur or a methyl group led to two possible diastereomers. The diastereomerically pure products were used to investigate the role of electrostatic effects in nuclease-DNA interactions[10]. Subsequently, a wide number of phosphate-ribose surrogates have been proposed[11]; few examples are depicted in *Figure1.1d–h*.

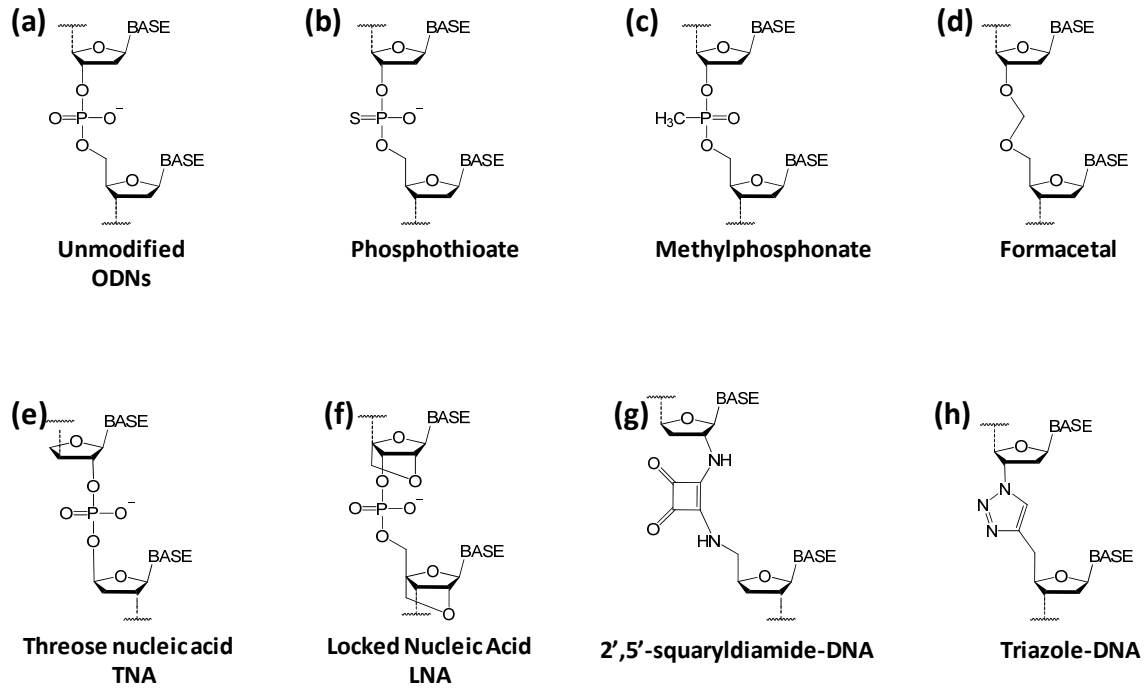


Figure 1.1. Backbone modifications: (a) Unmodified ODNs; (b) Phosphothioates; (c) Methylphosphonates; (d) Formacetals DNA[12]; (e) Threose nucleic acids[13]; (f) Locked nucleic acids[14]; (g) 2'-5'-Squaryldiamides[15]; (h) Triazole DNA[16].

Among the analogues depicted above (*Figure 1.1*), the LNA is one of the best in terms of improved recognition properties and increased binding affinity toward complementary DNA, making this modification very useful for gene targeting. Very promising results were reported for the triazole DNA analogues (h); a polythymine triazole analogue showed higher affinity towards polyadenosine oligonucleotides compared to the unmodified DNA. For synthetic biology, ideally modified oligonucleotides should mimic DNA or RNA functions as much possible[17], including the recognition by polymerase enzymes. To this scope TNA showed very interesting properties. A 200-mer TNA was synthesized by *DNA dependent Therminator polymerase* using a DNA template and TNA triphosphates[18] with acceptable fidelity; using an ad hoc TNA dependent polymerase, it was possible to transcript back the information from TNA to DNA.

Peptide nucleic acids (widely discussed below) are not recognized by polymerases, but are considered promising tool in synthetic biology[17] for template reactions. Recently, a thioester peptide was shown to incorporate appropriate thioester-carboxymethyl nucleobase derivatives, the overall process being driven by a DNA template (*Figure 1.2*)[19].

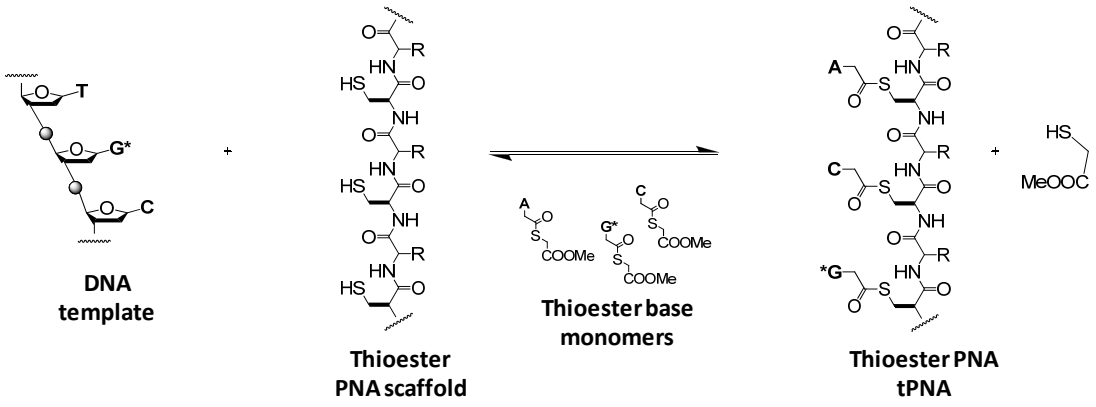


Figure 1.2. Dynamic self-assembly of *t*PNA through a DNA template.

1.3 Why Replace DNA Bases[20] ?

The four natural DNA nucleobases, plus uracil in RNA, allow to store and transfer of genetic information of all forms of life. However, base replacement is a very powerful strategy to widen the function of natural oligonucleotides. In the literature many examples of base replacement have been reported; these can be classified in two broad classes: a) base modifications made in order to change function and properties of DNA, and b) using the phosphate-ribose scaffold to organize complementary functional molecules, using as a model the DNA supramolecular recognition properties. As examples of the first class we chose two works coming from synthetic biology, concerning two different approaches aimed at the expansion of the natural genetic code by base modifications (*Figure 1.3*).

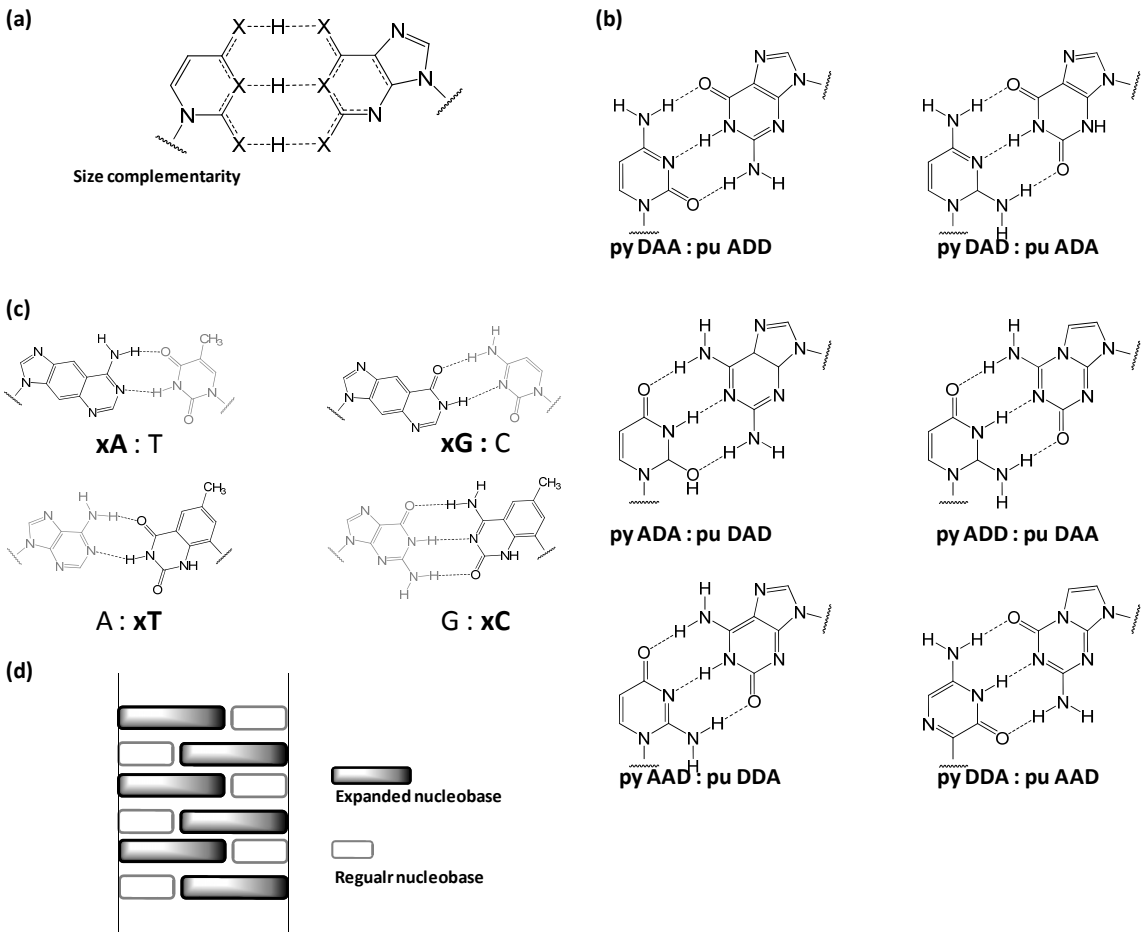


Figure 1.3. (a)Size complementarity and (b) modified purine and pyrimidine in Benner's Artificially Expanded Genetic Information Systems (AEGIS); (c)Kool's xDNA base pairing and their applications into eight bases (d) artificial genetic code. Py (pyrimidine), Pu(purine), D(H-bond donors) and A(H-bond acceptors).

Expansion of all possibility of hydrogen bonding donors (D)-acceptors (A) recognition patterns obeying to size complementarity principles, using the purine(Pu)-pyrimidine (Py) scheme (*Figure 1.3a*), was explored by *Benner*[21], thus creating a new code known as *Artificial Expanded Genetic Information System* (AEGIS), whose complementary components perform similarly to adenine-thymine and cytosine-guanine base pairs. The challenge for AEGIS has been the possibility to implement a genetic six letter alphabet able to act as substrate for a polymerase enzyme. Indeed, a mutant polymerase developed from HIV reverse transcriptase[22] was able to amplify a DNA duplex containing a py(DAD)-pu(ADA) pair in a primitive PCR reaction.

Another approach by Kool and co-workers was to create an expanded genetic system involving benzohomologation[23] of natural nucleobases (*Figure 1.3c*). The resulting size-expanded double helices are called xDNA ("expanded DNA") and display high stacking affinity and selectivity when paired with natural counterparts (*Figure 1.3d*). Early experiments with xDNA bases in template strands of DNA using DNA polymerase I (Klenow fragment) inserted natural DNA oligonucleotides opposite to xDNA, though with lower efficiency, thus suggesting interesting applications in synthetic biology.

One of the roles of the natural phosphate backbone is to support the genetic alphabet not only by providing water solubility, but also keeping DNA strands away from folding, thus allowing appropriate pre-organization of the bases. This offered the possibility for supramolecular chemistry to use the phosphate backbone as a scaffold to organize functional molecules, taking advantage also of the well developed synthetic DNA chemistry. A very interesting example came again from Kool's laboratories, that used the phosphodiester backbone as a scaffold for arranging fluorescent aromatic compounds in a potentially stacked oligonucleotide-like arrangement[24]. The synthesis of combinatorial oligonucleotide libraries containing only few fluorescent aromatic bases allowed to create a wide range of molecules having different absorption and fluorescence properties. Just as nature uses only four bases for genetic information, this approach allowed to create molecular diversity with a limited number of bricks.

Base modifications made to improve potency and selectivity of antisense or antigene agents were extensively studied, and will be listed in a subsequent paragraph, since many of these modifications have been used also in the modification of PNA.

1.4 Peptide Nucleic Acids: Structure, Synthesis and Modifications

Peptide Nucleic acids (PNAs) were first reported by Nielsen and co-workers[25] in *1991*, but still attract much interest for their unique recognition properties. PNAs are polyamide mimics of DNA in which the deoxyribose-phosphate backbone has been replaced by units of *N*-(2-aminoethyl)glycine, while the nucleobases are linked to the main chain by methylenecarbonyl linkers (*Figure 1.4a*).

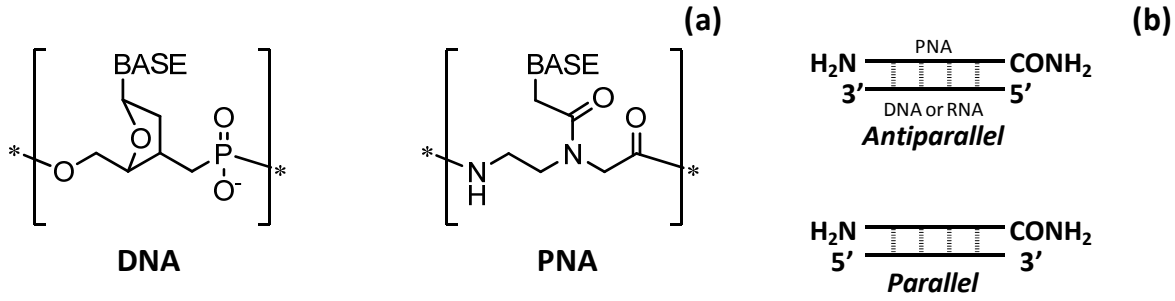


Figure 1.4. (a) DNA and PNA structures; (b) Parallel and antiparallel binding mode.

Due to their uncharged structures, PNAs are capable to bind strongly complementary DNA, and RNA (*Figure 1.4b*) giving rise to important applications in molecular biology and biotechnology. PNAs can also bind to complementary PNA strands, to form interesting PNA:PNA duplexes potentially useful in material chemistry and nanotechnology[26]. Since PNAs chemistry and properties have been reviewed in several papers[27] and in a book[28], herein we will present a short overview on synthesis, applications and structures, focusing on more the recent literature.

PNA Synthesis. PNA oligomer synthesis is usually performed on solid phase by oligomerization of opportunely protected PNA monomers, composed by the protected N-(2-aminoethyl)glycine backbone in which the secondary amino function is acylated with the nucleobase acetic acid derivatives. For each monomer two orthogonal protecting groups are needed, with temporary protection for the primary amino group and semi-permanent protection of the exocyclic amino group of the bases.

The oldest strategy, applied to the PNA synthesis of homothymine[29] and then to mixed sequences[30], employed Boc (*tert*-butyloxycarbonyl) for protecting the primary amino function, and Cbz (carbobenzyloxy) for the exocyclic amino groups of the bases and the monomers were oligomerized on a MBHA resin. Numerous efforts have been made to develop new protective group combinations, also employed for the synthesis of PNA-DNA (or RNA) chimera[31], leading to a variety of strategies such as: Dts/Cbz[32], Fmoc/Cbz[33], Fmoc/Mmt[34], Fmoc/acyl[35], Mmt/acyl[36] and Boc/Acyl[37]. However, since protected Boc/Cbz[38] and Fmoc/Bhoc monomers are commercially available, these two strategies are mainly employed for the synthesis of custom PNA oligomers. An innovative PNA synthetic strategy, designed for large scale synthesis, was patented from *Panagene, using* benzothiazole-2-sulfonyl (Bts)-protected cyclic monomers[39] (*Figure 1.5a*). These monomers are readily inserted on a PNA growing chain without pre-activation, giving elongation of oligomers; the subsequent monomer is inserted after Bts deprotection in basic media. *Hudson and Wojciechowski*[40] proposed an alternative to the Fmoc/Bhoc PNA synthesis, by replacing the Bhoc protective group of the exocyclic amine with bis-Boc derivatives (*Figure 1.5b*). Fmoc/bis-Boc[41] protected monomers were efficiently employed for the synthesis of mixed PNA sequences affording crude PNA oligomers with yields and purity comparable to those obtained with the Fmoc/Bhoc strategies, but avoiding the use of expensive reagents such as carbodiimidazole or toxic triphosgene. In an attempt to establish a synthetic strategy fully orthogonal with Fmoc protection, Bradley et al[42] reported the synthesis of 1-(4,4-dimethyl-2,6-dioxacyclohexylidene)ethyl/4-methoxytrityl (Dde/Mmt) protected PNA monomers (*Figure 1.5c*); these monomers were employed for the synthesis of Fmoc protected peptide-PNA conjugates. The use of the azido instead of the N-protected amino group and the synthesis of the *N*-(2-azidoethyl)glycine PNA monomer (*Figure 1.5d*) was reported by *Debaene and Winssinger*[43]; the azide was used as a masked amino group, which could be converted in the amino function after Staudinger's reduction by trimethylphosphine. The synthesis of some PNA-oligomers can be disturbed by the occurrence of a significant amount of truncated products, probably because on-resin aggregation hinders access during the coupling step. The use of a chaotropic salt could be a partial remedy to circumvent these problems. A different approach was used by Seitz[44] for the synthesis of homothymine PNA oligomers; inter-strand aggregation were solved by allyl protection of the oxygen atoms of the thymine base (*Figure 1.5e*).

These examples show the versatility and easy modification of the standard synthetic procedure for PNA production, often inspired by the close peptide chemistry.

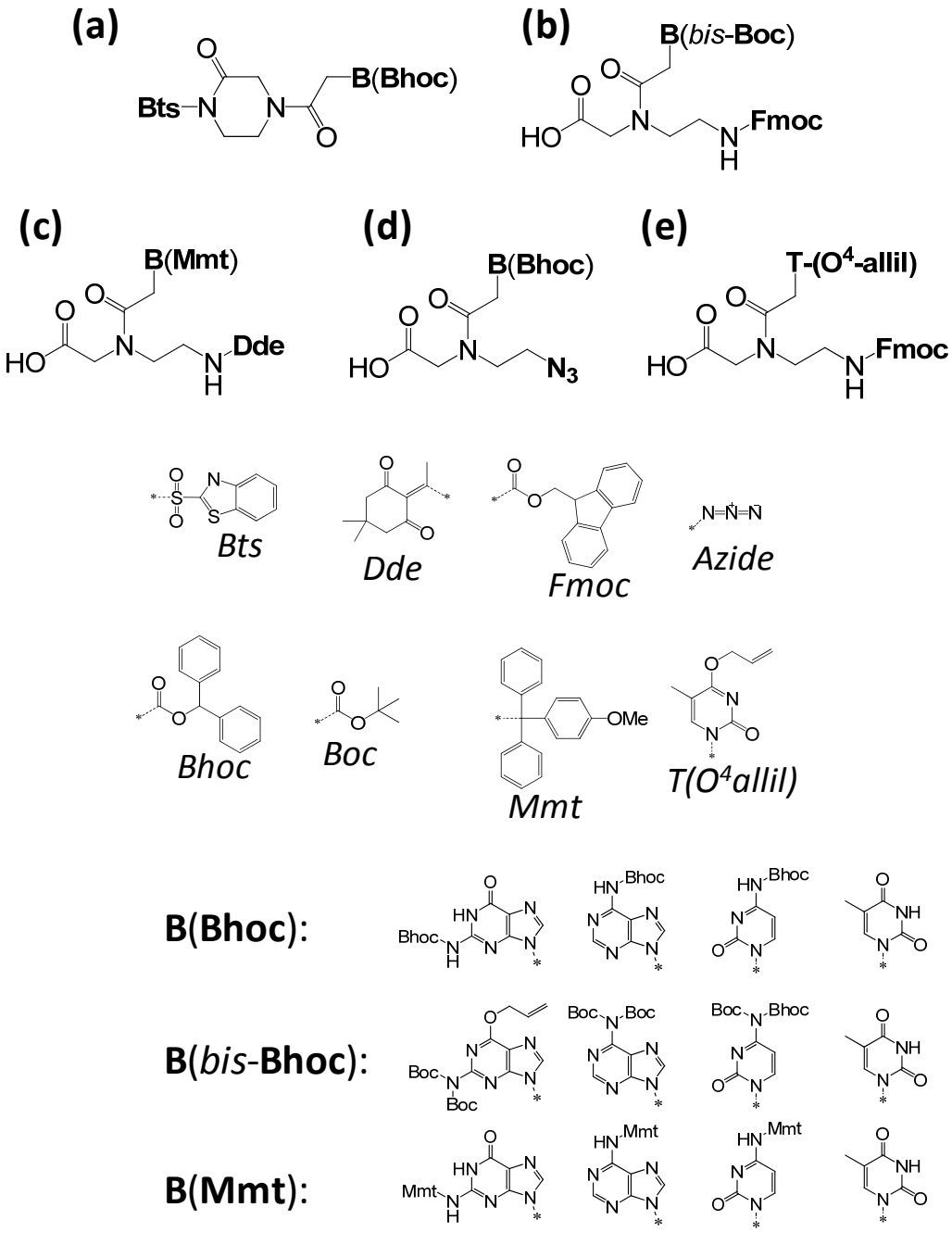


Figure 1.5. PNA synthetic strategies. (a) Bts-protected cyclic monomers. Deprotection of Bts is carried out in DIEA/benzendithiol; (b) Bis-Boc protected PNA monomers; (c) Dde/Mmt protected PNA monomers; (d) Aminoethylazido PNA monomers. (e) O^4-allil thymine Fmoc-PNA monomer. Abbreviations: benzothiazole-2-sulfonyl (Bts), 1-(4,4-dimethyl-2,6-dioxacyclohexylidene)ethyl (Dde), fluorenylmethoxycarbonyl (Fmoc), benzydriloxycarbonyl (Bhoc), 4-methoxytrityl (Mmt).Thymine containing monomers (a), (b), (c) and (d) thymine do not need any protection.

PNA binding to Nucleic Acids and Applications. PNAs hybridize more strongly (about 1°C for base pair[45]) to complementary DNA and RNA than natural oligonucleotides (ODNs). According to the sequence design, PNAs can bind DNA and RNA in parallel (5'-end facing the PNA N-term) and antiparallel (3'-end facing the PNA N-term) although, the antiparallel direction is preferred. Also the selectivity of PNA in DNA/RNA recognition is more pronounced than ODNs. Considering a 15-mer DNA sequence having a mismatch in the middle, the PNA:DNA duplex melting temperature drops of 13°C as compared to the fullmatch duplex, while for DNA:DNA the loss is only 4°C[46]. Moreover, the binding performances of PNAs are slightly affected by ionic strength[47] (due to the uncharged backbone) with a reverse behavior with respect to dsDNA duplexes. For a 10-mer DNA duplex at zero molar NaCl, the melting temperature is 20°C, and it increases of 10°C raising the salt concentration to 50 mM. The PNA:DNA duplex melting temperature is depressed of only 1°C from 0mM to 50 mM NaCl. PNAs are also stable in biological fluids[48], since they are not recognized by nucleases and proteases. All the properties listed above make PNAs very useful tools for applications in molecular biology, diagnostics and drug development. PNAs have been used in numerous applications, in connection with advanced diagnostic methods[49], for detecting specific gene sequences. Some common platforms that exploit PNA probes are: southern blotting[50], *in situ* hybridization[51], fluorescent *in situ* hybridization (FISH)[52], surface plasmon resonance[53], microarray systems[54], and affinity purification of nucleic acids[55]. Recently, combinations of nanoparticle-enhanced surface plasmon resonance imaging with PNA allowed the ultrasensitive detection of the target DNA at a femtomolar[56] concentration (corresponding to a 150 zeptomoles of total DNA amounts).

Targeting of double strand DNA with PNAs can occur through different binding modes; three of these (*Figure 1.6a-c*) require homopurine/homopyrimidine DNA targets, whereas double duplex invasion (*Figure 1.6d*) requires a DNA duplex containing at least 50% of AT (using pseudocomplementary PNA oligomers containing 2,4-diaminopurine and 4-thiouracil[57]). PNA triplex invasion complexes have sufficient stability to arrest the elongation of RNA polymerases. The formation of a triplex between T_{10} PNA and an A_{10} terminator site has been used as roadblock for arresting the transcription by RNA polymerase III[58].

Some PNA have the ability to invade mixed sequences of chromosomal DNA (linear and/or supercoiled) and thus it can be used to block transcription of DNA in mRNA[59].

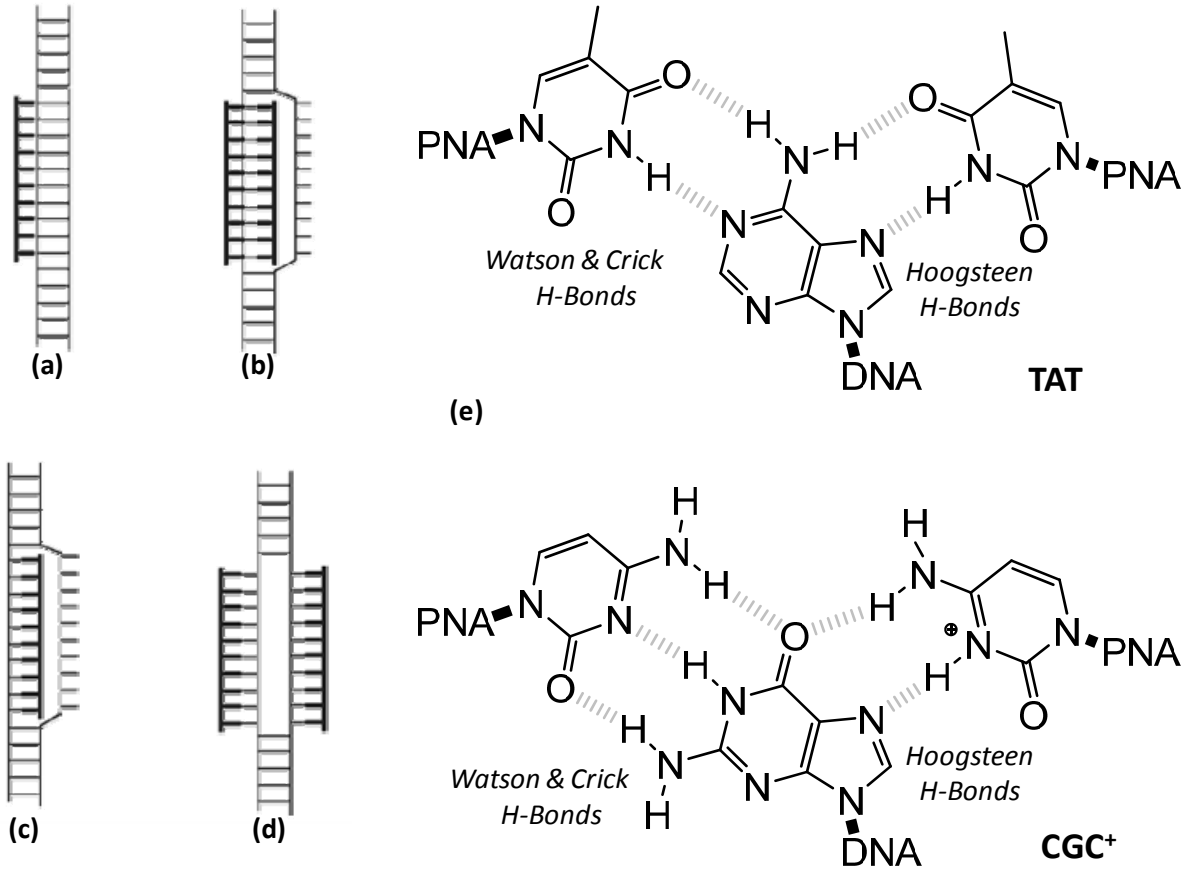


Figure 1.6. PNA binding modes toward DNA. (a) Triplex; (b) Triplex invasion; (c) Duplex Invasion; (d) Double duplex invasion. Bold line indicate PNA strand(s), thin line represent DNA duplex. (e) Watson & Crick and Hoogsteen hydrogen bonding.

Recently, an antigene peptide nucleic acid, conjugated to a nuclear localization signal peptide, was targeted against exon 2 of the MYCN gene in neuroblastoma cells; this led to the selective inhibition of MYCN transcription[60]. This finding is currently exploited to develop antigene PNA–based tumor-specific drugs for neuroblastoma and other solid tumors, for *in vivo* applications. Double duplex invasion by PNAs was used by *Komiyama et*[61] al to create new artificial restriction enzymes (ARCUT); non-symmetric double duplex invasion promoted by pseudocomplementary PNAs led to a complex with phosphate hot spot, susceptible to hydrolysis in the presence of a Ce(IV)/EDTA catalyst.

Modified PNAs. The applications of PNAs are often hampered by some drawbacks like their solubility in water, poor cellular uptake, tendency to self aggregate and direction binding in DNA/PNA recognition. Modifications of PNAs are not only needed to overcome these problems, but also to add functionalities or to increase the binding affinity when it is not sufficient for a particular application. Modifications of PNAs can be broadly divided into three classes: (i) backbone modifications[62], (ii) conjugation with small molecules[63] and (iii) nucleobase modifications[64]. Often these modification are not mutually exclusive and often can be used in combination.

Backbone modifications of PNAs are mostly aimed at improving the "preorganization", intended as the ability to adopt a more convenient conformation for minimizing the entropy loss during DNA binding. These "constrained" structures can be obtained by insertion of chiral centers, insertion of a cyclic structure in the backbone, or rigidification of the standard achiral *N*-(2-aminoethyl)glycine scaffold. In *Figure 1.7* we present a short overview of these modifications, both involving chiral or achiral backbone modifications.

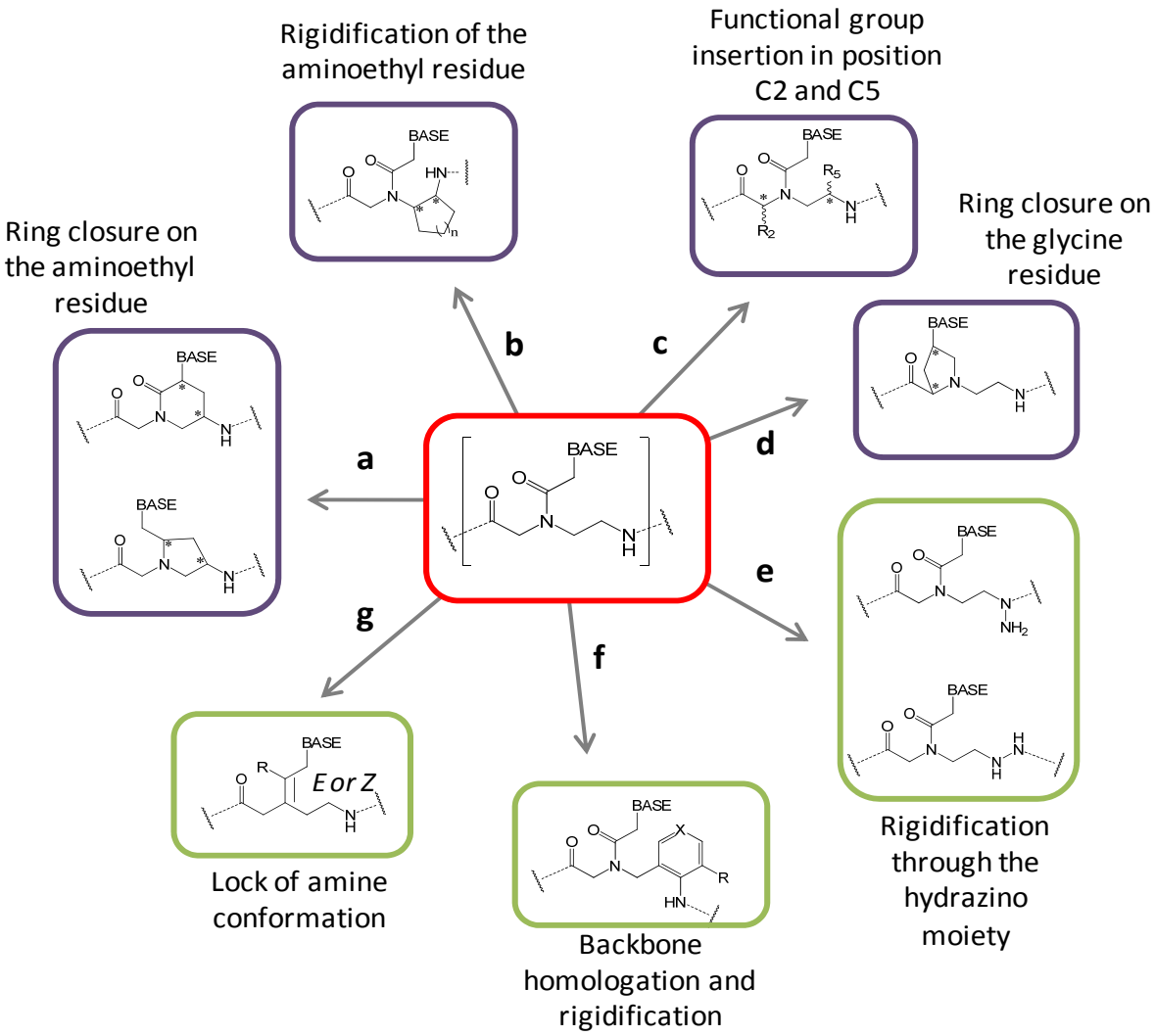


Figure 1.7. Strategies for inducing preorganization in the PNA monomers. Modifications that involve insertion of stereogenic centers are depicted In purple. For more information see the following references: (a)pyrrolidine-based PNAs[65] and piperidinone-PNA[66]; (b)cyclohexyl PNAs (chPNAs)[67] and cis-cyclopentyl PNA (cpPNA)[68]; (c)acyclic chiral PNA[69]; (d)aminoethylprolyl-PNA (aep-PNA)[70]; (e)N^{α} – *hyd* and N^{β}-*hyd* PNAs[71]; (f)aromatic PNA (APNA)[72]; (g)olefin PNA (OPA)[73].

One of the most efficient backbone modification by insertion of functional groups at position C2 (*Figure 1.8a*) is given by the "chiral box"[74], based on three adjacent C2 modified chiral monomers based on D-Lys, inserted into the middle of a PNA oligomer. This modification confers a better direction control in DNA binding (preference for antiparallel orientation) and improved mismatch recognition as compared to unmodified PNAs. On the other hand the L-Lys chiral box showed preferential binding and mismatch discrimination toward parallel DNA orientations.

Side chains of modified monomers derived from positively charged amino acids (for example lysine, or pseudoarginine, GPNA) confer an enhanced water solubility, improved cell permeability[75] and can be used also for backbone conjugation with fluorophores[76] or linking of PNA to a solid surface[77] (*Figure 1.8b-d*).

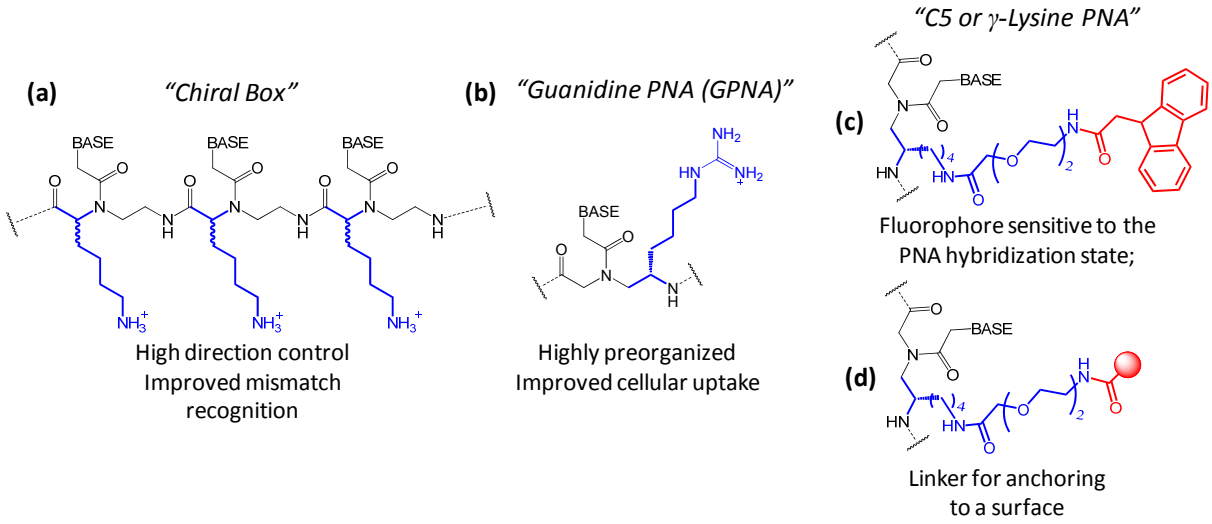


Figure 1.8. Some examples of modified chiral acyclic PNAs. (a) Chiral box; (b) GPNAs; (c) and (d) C5-Lys PNAs.

PNAs can be conjugated to small molecules at the N- and C-terminus by using standard or modified protocols for peptide synthesis. N-terminus modifications usually use the terminal free amino group of the pseudopeptide backbone. C-terminus modifications require more elaborate schemes, since they need the derivatization of the side chain of an opportune amino acids inserted at the C-terminus. In the literature, numerous example of PNA conjugates, for example with carrier peptides[78], fluorescent markers[79], intercalators[80], metal complexes[81], minor groove binders[82], lipophilic molecules[83], and polyamine[84] (*Figure 1.9*) are reported. These examples are not exhaustive, but clearly show the type of chemical diversity that can be introduced in peptide nucleic acids.

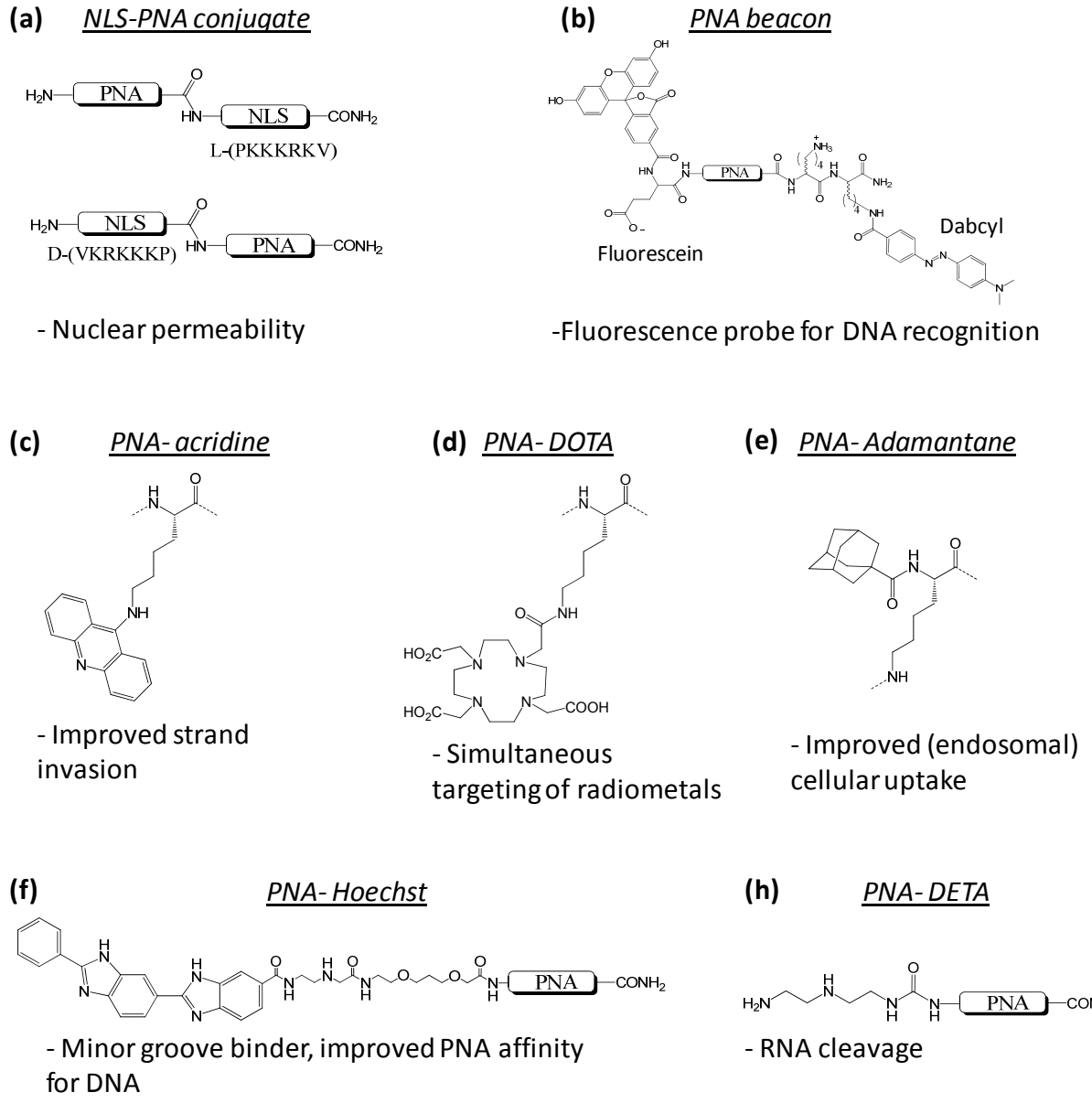


Figure 1.9. Survey on some PNA conjugates to small molecules and their applications. (a) NLS-carried peptides[78] (P:Proline, K: lysine, V: Valine, R: arginine); (b) PNA beacon[78]; (c) Acridine conjugate[79]; (d)PNA-DOTA[80]; (e) PNA-Adamantane[81]; (f)P NA-Hoechst[82]; (h) PNA-DETA[83].

Base Modifications in PNAs. "Molecular engineering" of PNAs by replacement of the standard nucleobases (adenine, guanine, cytosine and thymine) is a good strategy to improve binding properties and introduce chemical diversity. Modifications of nucleobases, according to their functions and properties, can be divided into four groups: (i) nucleobase analogues, (ii) nucleobases for triple helix formation, (iii) non-discriminating nucleobases analogues and (iv) reactive nucleobases.

Nucleobase analogues group (i) are base modifications that do not alter substantially the Watson-Crick hydrogen bond patterns; all these modifications are aimed at increasing binding affinity, selectivity or at conferring special selectivity. Larger π-surface area moieties (increasing the hydrophobic/stacking interactions) able to perform additional hydrogen bonding (e.g. binding of Hoogsteen free sites) and the presence of positive charges are all factors that positively affect DNA binding. In *Figure 1.10* some examples that employed the design principles described above are presented. The synthesis of PNA oligomers containing a cytosine analogue (using a benzothiazine-pyrimidine[85] ring, *Figure 1.10a*) was described by *Nielsen*[86]. This cytosine analogue effectively stabilized the PNA:DNA duplex by the increasing stacking interactions. *Manoharan and co-workers*[87] prepared PNA oligomers containing a phenoxazine ring analogously to the G-clamp base developed by Matteucci for oligonucleotides. While the phenoxazine replacement for cytosine led to a minor stabilization, the G-clamp led to a marked stabilization up to 20°C. The stabilization effect given by the aminoethyl side chain on the phenoxazine ring indicated the relevance of adjunct H-bonding on the Hoogsteen site of guanine, and the beneficial effect of the positive charge in binding to the negatively charged DNA. New cytosine analogues, based on a phenylpyrrolecytosine (PhpC) heterocycle, were reported by *Hudson and Wojciechowski*[88]. Single replacement of a cytosine with PhpC led to limited PNA:DNA duplex stabilization (about 2°C of the melting temperature compared to unmodified cytosine). While ortho-(aminoethoxy)phenylpyrrolecytosine (moPhpC), designed to engage guanine with an additional hydrogen bonding such as the G-Clamp, led to 10°C increase of the melting temperature. In attempts to use benzoyl protective groups for the exocyclic amino moiety of cytosine, *Nilesen and co-workers*[89] reported interesting properties of protected N^4-benzoylcytosine (C^{bz}) as cytosine surrogate. Due to the lower basicity of the N(3) nitrogen and to the impossibility to be protonated at a physiological pH, the C^{bz} showed inability to form triple helices. Bicyclic thymine analogues (bT) based on the 1,8-napthyridin-2(1*H*)-one ring system have also been prepared and evaluated by *Nielsen and coworkers*[90]. The best performance of these thymine analogues was obtained from 7-Cl-bT, bearing a chlorine substituent in a more favorable position to interact with the neighboring bases.

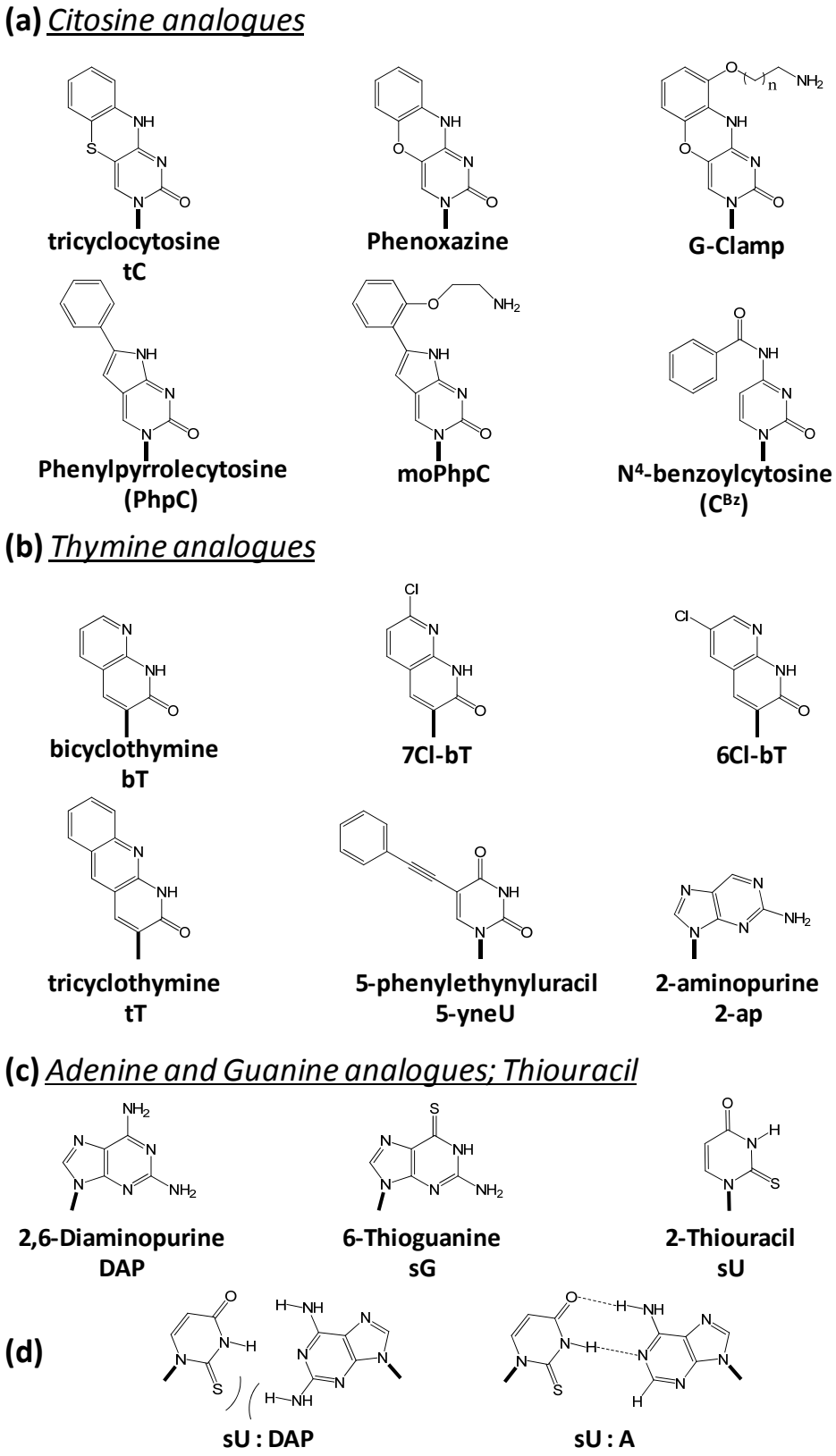


Figure 1.10. (a-c) Modified nucleobases of group (i). (d) sU does not pair with DAP, but perfectly matches with an adenine counterpart.

2-Aminopurine was reported for its interesting fluorescence response to DNA hybridization[91]. 6-Thioguanine, firstly introduced in oligonucleotide chemistry to avoid formation of secondary structures by guanine rich oligomers, was incorporated in PNA oligomers by *Nielsen and co-workers*[92] although it led to duplex destabilization. Adenine replacement by 2,6-Diaminopurine (DAP) was used both as adenine analogue and for building pseudocomplementary PNA oligomers for strand invasion[93] (in combinations with 2-thiouracil). Replacement of adenine with DAP[94] gives rise to a melting temperature gain of 4-6°C for each substitution.

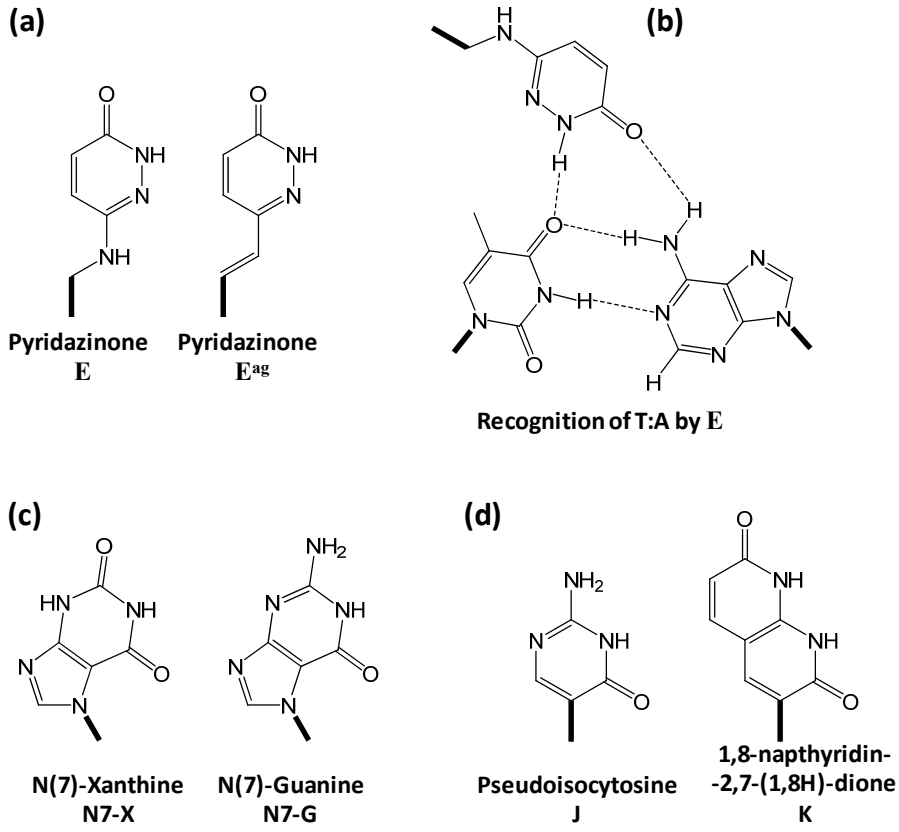


Figure 1.11. Nucleobases for triple helix formation. (a)Pyridazinone nucleobases; (b) Recognition scheme of **E**; (c)Cytosine and Thymine analogues for exclusive Hoogsteen binding; (d) mimics of protonated cytosine.

The scope to hit Hoogsteen sites of purine is an interesting strategy to develop new molecules able to perform triplex invasion. *Nielsen and co-worker*, in order to obtain nucleobases capable of thymine-adenine base pair recognition, created pyridazinone

nucleobases (*Figure 1.11a*) able to bind TA base pairs with a fair affinity in DNA duplexes (*Figure 1.11b*)[95]. One of the most important problems of mixed thymine/cytosine triplexes is their pH dependent stability, due to need to have protonated cytosine on the third strand. To this scope, bases mimicking a protonated cytosine were developed; these analogues are capable to bind via hydrogen bond the Hoogsteen sites of guanine without interacting with the Watson-Crick sites, since the hydrogen bonding pattern resulted changed. N7-modified xanthine and guanine (resembling Kool's xDNA) act as mimics of thymine and protonated cytosine to interact exclusively with Hoogsteen sites of their complementary purine in DNA strands[96,97]. These pyrimidine analogues, because of expanded sizes, are not able to pair on Watson-Crick sites due to steric hindrance, but are compatible with Hoogsteen sites. Other mimics of protonated cytosine[98] are depicted in *Figure 1.11d*.

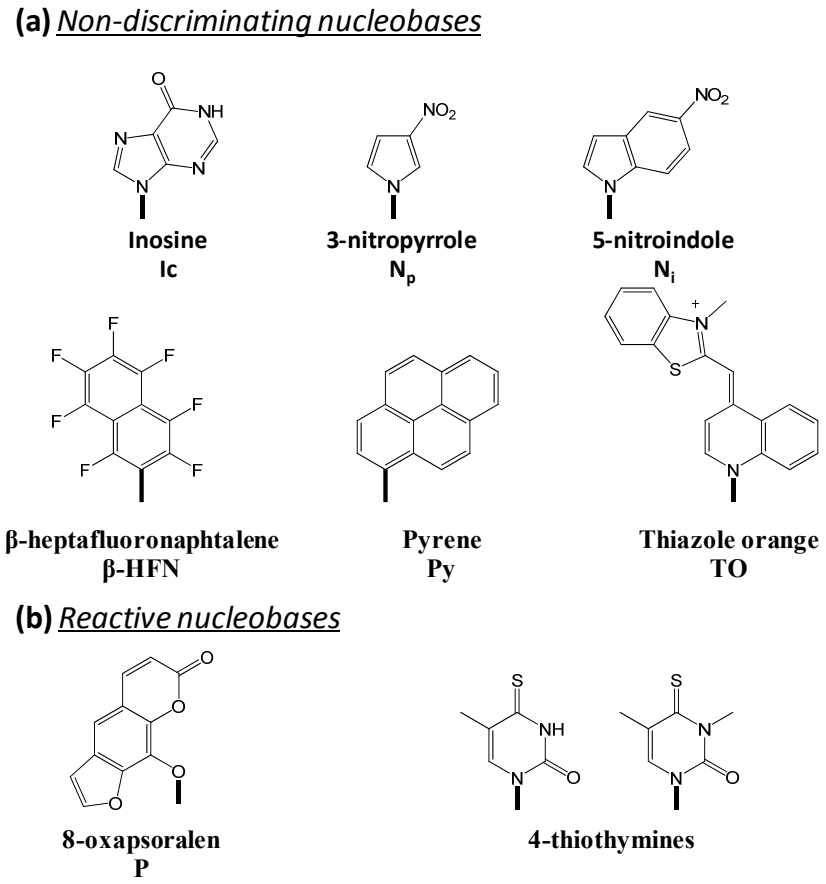


Figure 1.12. (a)Non-discriminating nucleobase analogues[99]; (b) Reactive nucleobases.

Non-discriminating base analogues were designed in principle to bind almost with the same efficiency to all four natural counterparts. This role of "universal base" can be played by aromatic or heteroaromatic residues, since they cannot form hydrogen bonding and they stack efficiently with neighboring bases. In *Figure 1.12a* some examples of "universal bases" are reported. The TO (thiazole orange) resulted very interesting since its intense fluorescence is quenched by adjacent mismatches on the complementary DNA[100] (this approach to mismatch recognition is called FIT, forced intercalation). Some example of reactive nucleobases such as 8-oxapsoralen[101], or 4-thiothymine,[102] are depicted in *Figure 1.12b*. For example PNA short oligomers containing psoralen can be useful as a tool for molecular biology and for the development of potent phototerapeutic drugs for the PUVA approach (psoralen plus UVA irradiation).

For all figures of this paragraph (*Figure 1.10 -11-12*) the heavy line in structures indicates attachment to $-CH_2C(O)N$ of the PNA backbone.

The C(5) position of uracil: a stepping stone towards the major groove of duplexes

The C(5) position of uracil is a common site to introduce modifications in nucleic acids since it places substituents protruding into the major groove without affecting the hybridization properties. For oligonucleotides this type of modification is tolerated by polymerase reactions; for this reason there are numerous example of C(5) modified uridines bearing fluorophores, reactive groups, anchoring groups, metal chelators etc.[103].

A survey of C(5) uracil modifications in PNAs is showed in *Figure 1.13*.

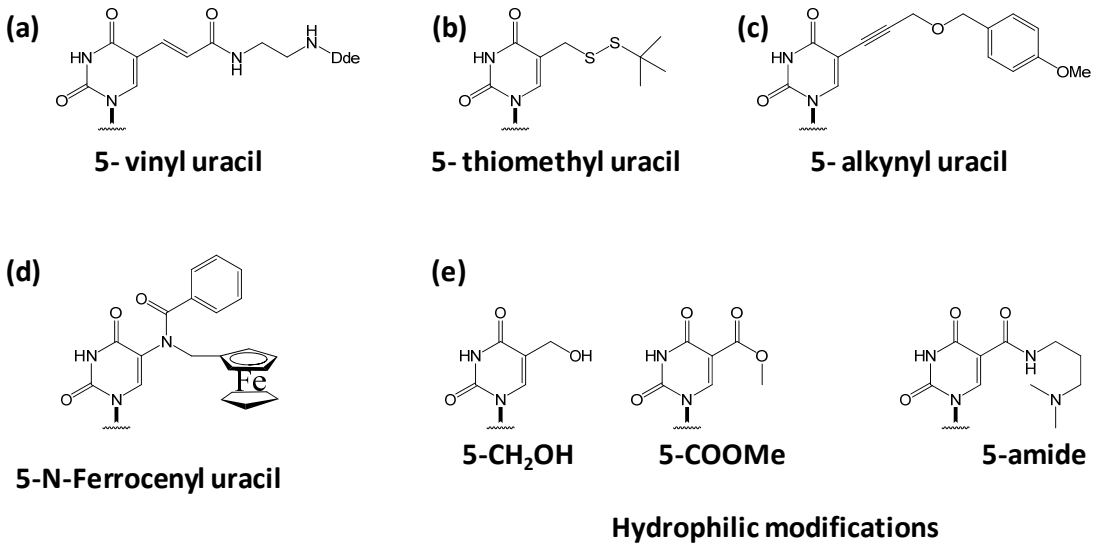


Figure 1.13. Uracil C(5) modifications inserted into a PNA scaffold. (a)5-vinyluracil[104]; (b)5-thiomethyl[105]; (c)5-alkynyl[106]; (d)5-ferrocenyluracil[107]; (e) 5-hydroxymethyluracil, 5-isoorotic acid methyl ester and 5-(N,N'-dimethylaminopropyl)isorootamide[108].

PNA-containing duplex structures. PNAs can form a variety of duplex structures that involve both parallel and antiparallel complementary PNA, DNA or RNA. Only few solid state and NMR structures are available; in *Figure 1.14* PNA:DNA, *c*PNA:DNA, PNA_2DNA and PNA:PNA structures are depicted and their features are reported in *Table 1* and compared to the DNA.

Table 1. Averaged helical parameters for duplexes involving PNA and A-DNA structures depicted in *Figure 1.14*.

Duplex	Twist (°)	Rise (Å)	Displacement (Å)	Base per turn
A-DNA/B-DNA[109]	32.7/3.6	2.6/3.4	-4.5/-0.1	11/10
PNA:DNA[110]	28.0	3.3	-3.8	13
Chiral box PNA:DNA[111]	23.2	3.5	-3.8	16
PNA_2DNA Triplex[112]	22.9	3.4	-6.8	16
antip. PNA:PNA[113]	19.8	3.2	-8.3	18

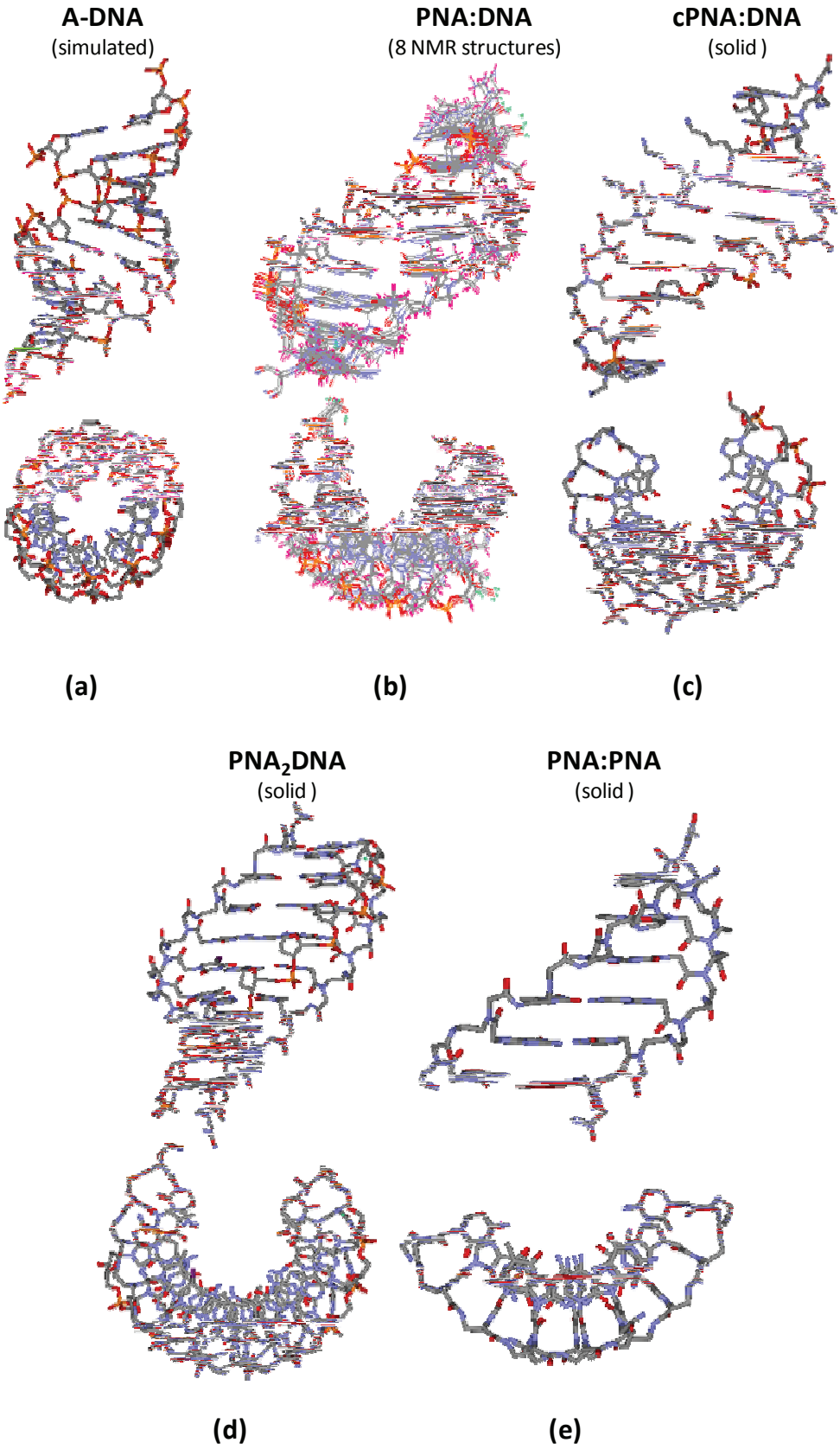


Figure 1.14. PNA containing structures compared vs A-DNA. (a) A-DNA structure (10 mer); (b) PNA:DNA superimposed NMR structures (8-mer); (c) "chiral box" PNA:DNA (10-mer); (d) triplex PNA:DNA (8-mer); (e) PNA:PNA antiparallel(6mer). Hydrogen atoms were omitted for clarity.

At a first glance, PNA involving duplexes are structurally different from A- or B- forms of DNA, and their structures are much wider due to a minor twist and major numbers of base pairs for turn. Compared to dsDNA, in the PNA-containing duplex structures the helix is less tightly wound. The antiparallel PNA:PNA duplex shows a new helix form, known as P-helix typical for PNA:PNA, while for PNA:DNA duplexes (and triplexes) the helix is much closer to the P-form rather than to the A or B forms of DNA. These structures clearly demonstrate that PNA is not as flexible as can be predicted from its linear structure; in the duplexes, amide E/Z stereoisomers present in the single strand collapse to a unique configuration. The PNA hybridization process is enthalpically driven and accompanied by a large entropy loss; a margin for improving PNA binding can therefore be searched in backbone preorganization. Indeed, preorganized L-serine γ-PNAs[114] bind more strongly to complementary DNA rather than unmodified PNAs, while D-serine γ-PNA monomers imparting a left-handed preorganization lead to a less efficient DNA recognition.

Backbone constraints, as showed from different helical parameters, dictated the helical shape. PNA analogues with a highly constrained backbone, having a high tendency to adopt β-sheet conformations like the β-homoalanyl scaffold[115], generated homo PNA duplexes with peculiar structural features. In *Figure 1.15* a solid structure of a β-homoalanyl PNA in which the highly constrained backbone forces the bases to adopt a ladder-type conformation is reported; these findings can be important as models of ladder-type conformations proposed for DNA complexes with proteins[116].

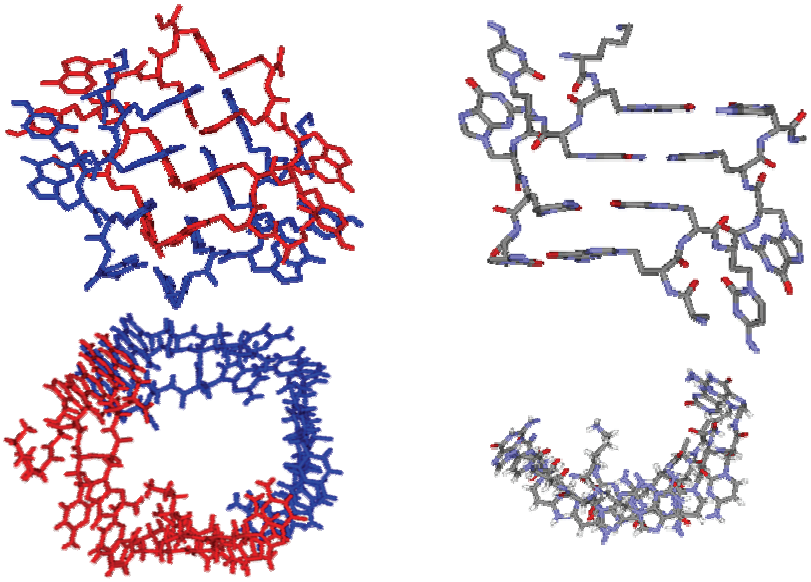


Figure 1.15. Solid state structure of β-homoalanyl-PNA[117] molecular box. (a)Top and front view; (b) Sketch of a PNA duplex portion.

Study of the helicity in PNA-containing duplexes. Insertion of chiral synthons in the peptide nucleic acids backbone is a very useful tool to modulate nucleic acid recognition. However, unmodified PNA:PNA duplexes are interesting since the P-helix is racemic in the absence of chiral inducers and can be easily modulated by chiral groups placed either at the end of the PNA strand or in the backbone of one or more monomers[118]. *Nielsen et al*[119] observed a preferential handedness in solution for PNA:PNA duplexes bearing a single L-lysine at C-terminus, as determined by circular dichroism measurements in the 220-280 nm region. While for PNA:DNA or PNA:RNA duplexes the helix is right handed, due to the predominance of the chiral centers of ribose, for PNA:PNA duplexes the assessment of handedness by circular dichroism in the region between 220-280 nm is not trivial. Thus, the exciton coupling method[120] was applied for the first time to assess absolute supramolecular helicity of nucleic acids duplexes by *Armitage and co-workers*[121].

This method exploits the minor groove binding abilities of some dyes, which are able to interact with PNA:DNA and PNA:PNA duplexes giving rise to dye aggregates able to produce an intense exciton couplet; from the sign of this couplet it is possible to infer the handedness of helix (*Figure 1.16*).

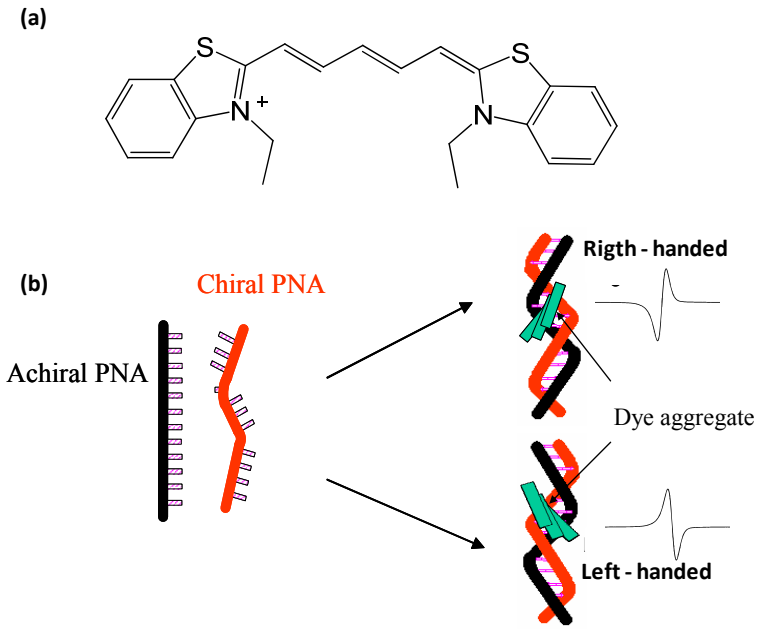


Figure 1.16. (a) Structure of $Disc_2(5)$ (diethyldicarbocyanine); (b) Speculative picture to explain the cyanine binding mode and the induced CD signal expected.

All these findings[122] indicate that PNAs are not perfect structural mimics of DNA, but that the pseudo-peptide backbone plays a fundamental role in recognition processes in contrast to what happens in others "flexible" analogues of DNA, which perform worse than PNA[4] in DNA recognition. Finally, since the PNA monomer is composed of an aminoacidic backbone, it has been proposed that PNA precursors might have been formed on the primitive earth (or in the outer space)[123]. These factors (in connection with recent findings that PNAs can be synthesized by template chemical reactions) have induced some scientists to propose that a PNA world occurred before the RNA world in the origin of life. Though this hypothesis is far from being demonstrated, it offers an interesting scenario[124] for future studies.

1.5 References

[1] Watson J., Crick F., Nature, **1953**, *171*, 737-738;

[2] Hecht SM, editor. *Bioorganic chemistry—Nucleic acids*., **1996**, Oxford, UK:Oxford University Press;

[3] (a) Stephenson M. L., Zamecnik P. C., *Proc. Natl. Acad. Sci.USA*, **1978**, *75*, 285; (b) Zamecnik P. C., Stephenson M. L., *Proc. Natl. Acad. Sci.USA*, **1978**, *75*, 280;

[4] S. A. Benner, A. M. Sismour, *Nature Review*, **2005**, *6*, 533;

[5] C. J. Leumann, Bioorg. & Med. Chem., **2002**, *10*, 841–854;

[6] Breslow R., *Acc. Chem. Res.*, **1995**, *28* (3), 146–153;

[7] S. A. Benner, A. M. Sismour, *Nature Review*, **2005**, *6*, 533 (and reference cited therein);

[8] S. A. Benner, *Nature*, **2003**, *9*, 421;

[9] Rawls R., *Chem. & Eng. News*, **2000**, *april 24*, 49-53;

[10] Elliott S.L., Brazier J., Cosstick R., Connolly B. A., *J. Mol. Biol.,* **2005**, *353*, 692–703;

[11] Bell N. M., Micklefield J., *ChemBioChem*, **2009**, *10*, 2691 – 2703;

[12] Micklefield J., *Curr. Med. Chem.* , **2001**, *8*, 1157.

[13] Schoning K., Scholz P., Guntha S., Wu X., Krishnamurthy R., Eschenmoser A., *Science*, **2000**, *290*, 1347-1351;

[14] Vester B., Wengel J., *Biochemistry*, **2004**, *43*, 13233-13241;

[15] Sato K., Tawarada K., Seio K., Sekine M., *Eur. J. Org. Chem.*, **2004**, 2142;

[16] Isobe H., Fujino T., Yamazaki N., Guillot-Nieckowski M., Nakamura E., *Org. Lett.*, **2008**, *10*, 3729;

[17] Appella D. H., Curr. Op. Chem.Biol., **2009**, *13*, 687–696;

[18] Ichida J.K., Horhota A., Zou K., McLaughlin L.W., Szostak J.W., Nucl. Acids Res., **2005**, *33,* 5219-5225;

[19] Ura Y., Beierle J.M., Leman L.J., Orgel L.E., Ghadiri M.R., *Science*, **2009**, *325,* 73-77;

[20] Kool E. T. , Acc. Chem. Res., **2002,** *35*, 936-943;

[21] Benner S. A., Acc. Chem. Res., **2004,** 37, 784-797;

[22] Sismour A. M., Lutz S., Park J.-H., Lutz M. J., Boyer P. L.,, *Nucleic Acids Res*., **2004**, 32(2), 728-735;

[23] Kool E. T., *Acc. Chem. Res.*, **2007,** *40*, 141-150;

[24] Gao J., Watanabe S., Kool E. T., **J. Am. Chem. Soc., 2004**, *126*, 12748-12749;

[25] Nielsen PE, Egholm M, Berg RH, Buchardt O, *Science*, **1991**, *254*, 1497-1500;

[26] Nielsen P. E., *DNA-based nanodevices (AIP conference proceedings)*, **2008**, 1062, 13-18;

[27] (a)Nielsen P.E., *Molecular Biothecnology.*, ***2004***, *26*, 233;

[28] Nielsen P. E., *Peptide Nucleic Acids: Protocosl and applications*, **2004**, *Horizon Scientific Bioscience*;

[29] Egholm M., Buchardt O, Nielsen P. E., Berg R. H., *J. Am. Chem. Soc.*, **1996**, *118*, 7049;

[30] Buchardt O, Egholm M., Berg R. H, Nielsen P. E., *Trends Biotechol*, **1993**, *8*, 53;

[31] Moggio L., Romanelli A., Gambari R., Bianchi N., Borgatti M., Fabbri E., Mancini I., Di Blasio B., Pedone C., Messere A.; *Biopolymers*, **2007**, *88(6)*, 815-822;

[32] Planas M., Bardaji E., Jensen K. J., Barany G., *J. Org. Chem*., **1999**, *64*, 6281;

[33] Thomson S. A., Josey J. A., Cadilla R., Gaul M.D., Hassman C.F., Luzzio M.J., Pipe A.J., Reed K.L., Ricca D.J., Wiethe R.W., Noble S.A., *Tetrahedron*, **1995**, *51(22)*, 6179-6194;

[34] Breiphol G., Knolle J., Langner D., O'Malley G., Ulhaman E., *Biorg. Med. Chem. Lett.*, **1996**, *6*, 665;

[35] Kovacs G., Timar Z., Kupihar Z., Kele Z., Kovacs L., *J. Chem.Soc., Perkin Trans. 1*, **2002**, *57*, 9481;

[36] Uhlman E., Will D. W., Breiphol G., Peyman A., Langnar D., Knolle J., O'Malley G. J., *Nucleosides Nucleotides*, **1997**, *16*, 603;

[37] Kofoed T., Hansen H. F., Oerum H., Koch T., *J. Pept. Sci.*, **2001**, *7*, 402;

[38] Uhlmann E., Peyman A., Breipohl G., Will D. W., *Angew. Chem. Int. Ed.*, **1998**, *37(20)*, 2797-2823;

[39] Hyunil L., Jae Hoon J., Jong Chan L., Hoon C., Yeohong Y., Sun Kee K., *Org. Lett.*, **2007**, *9(17)*, 3291-3293;

[40] Wojciechowski F., R. HE Hudson, *J. Org. Chem*., **2008**, *73*, 3807–3816;

[41] Porcheddu A., Giacomelli G., Piredda I., Carta M., Nieddu G., *Eur. J. Org. Chem.,* **2008**, 5786–5797;

[42] Bialy L., Dı´az-Mocho´n J. J., Specker E., Keinicke L., Bradley M., *Tetrahedron*, **2005**, *61*, 8295–8305;

[43] Debaene F:, Winssenger N., Org. Lett., **2003**, *5(23)*, 4445-4447;

[44] Altenbrunn F., O. Seitz, *Org. Biomol. Chem.*, **2008,** *6*, 2493–2498;

[45] Giesen U., Kleider W., Berding C., Geiger A., Ørum H., Nielsen P.E., **.** *Nucleic Acids Res,* **1998**, *26*, 5004-5006;

[46] Egholm M., Buchart O., Christensen L, Beherens C, Freier SM, Driver D.A., Berg R.H., Kim S.K., Norden B., Nielsen P.E.; *Nature*, **1993**, *365(6446)*, 566-568;

[47] Tomac S., Sarkar M.,R atilainen T., Wittung P., Nielsen P.E., Norden B., Graslund A., J. Am. Chem. Soc., **1996,** *118,* 5544-5552;

[48] Demidov, V.V., *Biochem. Pharmacol.*, **1994**, *48*, 1310;

[49]Nielsen P. E., *Curr. Opin. Biotech.* , **2001**, *12*, 16;

[50] Perry-O'Keefe H., Yao X.W., Coull J.M., Fuchs M, Egholm M., *Proc. Natl. Acad. Sci.*, **1996**, *93*, 14670;

[51] Murakami T.. *et al., J. Pathol.*, **2001**, *130*, 194;

[52] Taneja K.L., Chavez E.A., Coull J., Lansdorp P.M., *Genes Chromosomes Cancer*, **2001**, *30*, 57;

[53](a)Burgener M., Sanger M., Candrian U., *Bioconjug. Chem.*, **2000**, *11*, 749; (b) Corradini R., Feriotto G., Sforza S., Marchelli R., Gambari R*., J. Mol. Rec.*, **2004**, *17(1)*, 76-84;

[54] Calabretta A., Tedeschi T., Di Cola G., Corradini R., Sforza S., Marchelli R., *Mol. Biosys.*, **2009**, *5(11)*, 1323-1330;

[55] Orum, H. *et al.,* Biotechniques, **1995**, *19*, 472;

[56]D'Agata R., Corradini R., Grasso G., Marchelli R., Spoto G., *ChemBioChem*, **2009**, *9(13)*, 2 067-2070;

[57] Lohse J., Dahl O., Nielsen P.E., *Proc. Natl. Acad. Sci.,* **1999**, *96*, 11804-11808;

[58] Dieci G., Corradini R., Sforza S., Marchelli R., Ottonello S., *J. Biol. Chem.*, **2001**, *276(8)*, 5720-5725;

[59] K. Kaihatsu, B. A. Janowski, D. R. Corey, *Chemistry & Biology*, **2004**, *11*, 749–758;

[60] Tonelli R., Purgato S., Camerin C., Fronza R., Bologna F., Alboresi S., Franzoni M., Corradini R., Sforza S., Faccini A., Shohet J. M., Marchelli R., Pession A., *Mol. Cancer Ther.* **2005**, *4(5)*, 779–86;

[61] Yamamoto Y., Uehara A., Tomita T., Komiyama M., *Nucl. Acids Res*., **2004**, *32(19)*, e153;

[62] Corradini R., Sforza S., Tedeschi T., Totsingan F., Marchelli R., *Curr. Top. Med. Chem.*, **2007**, *7(7)*, 681-694;

[63] Z. V. Zhilina, A. J. Ziemba, S. W. Ebbinghaus, *Curr. Top. Med. Chem.,* **2005**, *5,* 1119-1131;

[64] Wojciechowski F., R. HE Hudson, *Curr. Top. Med. Chem.,* **2007**, *7,* 667-679;

[65] Puschl A., Tedeschi T., Nielsen P.E. , Org Lett, **2000**, *2*, 4161–4163;

[66] Shirude P.S., Kumar V.A., Ganesh K.N., *Tetrahedron*, **2004**, *60*, 9485–9491;

[67] Lagriffoule P., Wittung P., Eriksson M., Jensen K., Norden B., Buchardt O., Nielsen P.E., *Chem. Eur.J.*, **1997**, *3*, 912–919;

[68] (a) Govindaraju T., Madhuri V., Kumar V.A., Ganesh K.N., *J. Org.Chem.*, **2006**, *71*, 14–21; (b)Govindaraju T., Kumar V.A., Ganesh K.N., *Chem Commun.*, **2004**, 860–861;

[69] (a)Puschl A., Sforza S., Haaima G., Dahl O., Nielsen P.E., *Tetrahedron Lett .*, **1998**, *39*, 4707–4710; (b)Dueholm K.L., Petersen K.H., Jensen D.K., Egholm M., Nielsen P.E., Buchardt O., *Bioorg Med Chem Lett*, **1994**, *4*, 1077–1080; (c)Sforza S., Haaima G., Marchelli R., Nielsen P.E., *Eur. J. Org.Chem.*, **1999**, 197–204; (d)Sforza S., Corradini R., Ghirardi S., Dossena A., Marchelli R., *Eur. J. Org.Chem.*, **2000**, 2905–2913; (e)Menchise V., De Simone G., Tedeschi T., Corradini R., Sforza S., Marchelli R., Capasso D., Saviano M., Pedone C., *Proc. Natl. Acad. Sci. ,* **2003**, *100*, 12021–12026;

[70] D'Costa M., Kumar V., Ganesh K.N., *Org. Lett.*, **2001**, *3*, 1281–1284;

[71] Cerea P., Giannini C., Dall'Angelo S., Licandro E., Maioranaa E., Marchelli R., *Tetrahedron*, **2007**, *63*, 4108–4119;

[72] (a) Fader L., Myer E., Tsantrizos Y. S., *Tetrahedron*, **2004**, *60*, 2235; (b) Manea V. P., Wilson K. J., Cable J. R*., J. Am. Chem. Soc.*, **1997**, 2033;

[73] Scutz R., Cantin M., Roberts C., Greiner B., Uhlmann E., Leumann C., *Angew. Chem. Int. Ed.*, **2000**, *39*, 1250;

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

[75] Sahu B., Chenna V., Lathrop K. L., Thomas S., Zon G., Livak K. J., Ly D. H., *J. Org. Chem.*, **2009**, *74*, 1509–1516;

[76] Englund E. A., Appella D. H., **2005**, *7(16)*, 3465-3467;

[77] Totsingan F., Tedeschi T., Sforza S., Corradini R., Marchelli R., Chirality, **2009**, *21(1)*, 245-253;

[78] Kaihatsu K., Huffman K. E., Corey D. R., Biochemistry, **2004,** *43,* 14340-14347;

[79] Totsingan F., Rossi S., Corradini R., Tedeschi T., Sforza S., Juris A., Scaravelli E., Marchelli R., *Org. Biomol. Chem.*, **2008**, *6*, 1232–1237;

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

[81] (a)Gallazz F., Wang Y., Jia F., Shenoy N., Landon L. A., Hannink M., Lever S. Z., Lewis, M. *Bioconjug. Chem.,* **2003,** *14*, 1083-1095; (b)Lever S. Z., Hannink M., *Bioconjug. Chem.,* **2002,** *13*, 1176-1180;

[82] Nielsen P. E., Frederiksen K., Behrens C., *ChemBioChem,* **2005,** *6*, 66-68;

[83] (a) Ljungstrom T., Knudsen H., Nielsen P. E., *Bioconjug. Chem.,* **1999,** *10*, 965-972; (b) Mologni L., Marchesi E., Nielsen P. E., Gambacorti-Passerini C., *Cancer Res,* **2001,** *61*, 5468-5473;

[84] Verheijen J.C., Birgit A. L., Deiman M., Yeheskiely E., van der Marel G. A., van Boom J. H., *Angew. Chem. Int. Ed*., **2000**, *39(2)*, 369-372;

[85] Lin, K.-Y., Jones R. L., Matteucci M., *J. Am. Chem. Soc,* **1995**, *117*, 3873;

[86] Eldrup A. B., Nielsen B. B., Haaima G., Rasmussen H., Kastrup J. S., Christensen C., Nielsen P.E., *Eur. J. Org. Chem.,* **2001**, *9*, 1781;

[87] Rajeev K.G., Maier M. A., Lesnik E. A., Manoharan M., *Org. Lett.,* **2002**, *4*, 439;

[88] Wojciechowski F., Hudson R. H. E., *J. Am. Chem. Soc.,* **2008**, *130*, 12574–12575;

[89] Christensen L., Hansen H. F., Koch T., Nielsen P.E., *Nucleic Acids Res,* **1998**, *26*, 2735;

[90] (a) Eldrup A. B., Nielsen B. B., Haaima G., Rasmussen H., Kastrup J. S., Christensen C., Nielsen P. E., *Eur. J. Org. Chem,* **2001**, *9*, 1781; (b)Eldrup A. B., Christensen C., Haaima G., Nielsen P. E., *J. Am. Chem. Soc.,* **2002**, *124*, 3254;

[91] Gangamani B.P., Kumar V.A., Ganesh K.N., *Biochem. Biophys. Res.Comm.,* **1997**, *240*, 778;

[92] Hansen H.F.; Christensen L., Dahl O., Nielsen P.E., *Nucleos. Nucleot.,* **1999**, *18*, 5;

[93] Lohse J., Dahl O., Nielsen P.E., *Proc. Natl. Acad. Sci USA,* **1999**, *6*, 11804;

[94] Haaima G., Hansen H.F., Christensen L., Dahl O., Nielsen P.E., *Nucleic Acids Res*., **1997**, *25*, 4639;

[95] Eldrup A.B., Dahl O., Nielsen P.E., *J. Am. Chem. Soc.,* **1997**, *119*, 11116;

[96] Hudson R.H.E., Goncharenko M., Wallman A.P., Wojciechowski F., *Synlett,* **2005**, *9*, 1442;

[97] (a)D'Costa M., Kumar V.A., Ganesh K.N., *J. Org. Chem.,* **2003**, 68, 39; (b)Shirude P.S., Kumar V.A., Ganesh K.N., *Eur. J. Org. Chem.,* **2005**, *24*, 5207;

[98] (a)Egholm M., Christensen L., Dueholm, K.L., Buchardt O., Coull, J., Nielsen P.E., *Nucleic Acids Res.,* **1995**, *23*, 217; (b)Christensen C., Eldrup A.B., Haaima G., Nielsen P.E., *Bioorg. Med. Chem.,* **2002**, *12*, 3121;

[99] Challa H., Styers M.L., Woski S.A., *Org. Lett.,* **1999**, *1*, 1639;

[100] Seitz O., Bergmann F., Heindl D., *Angew. Chem. Int. Ed.,* **1999**, *8*, 2203;

[101] Okamoto A., Tanabe K., Saito I., *Org. Lett.,* **2001**, *3*, 925-927;

[102] Clivio P., Guillaume D., Adeline M-T., Hamon J., Riche C., Fourrey J-L, *J. Am. Chem. Soc.,* **1998**, *120*, 1157;

[103] Luyten L., Herdewijn P., *Eur: J. Med. Chem.*, **1998**, *33*, 515-576;

[104] Oquare B. Y., Taylor J. S., *Bioconjugate Chem.,* **2008,** *19,* 2196–2204;

[105] Jianfeng Cai, Xiaoxu Li, Taylor J. S., **2005**, *7(5)*, 751-754;

[106] Hudson R. H. E., Ge Li, Tse J., *Tetrahedron Lett.*, **2002**, *43*, 1381–1386;

[107] Gasser G., Spiccia L., J. Organomet. Chem., **2008**, *693* , 2478–2482;

[108] (a)Hudson R.H.E., Viirre R.D., Liu Y.H., Wojciechowski F., Dambenieks A.K., *Pure Appl. Chem.,* **2004**, *76*, 1591; (b)Hudson R.H.E., Wojciechowski F., *Can. J. Chem.,* **2005**, *83*, 1731; (c)Hudson R.H.E., Wojciechowski F., *Nucleos. Nucleot. Nucl.,* **2005**, *24*, 1123;

[109] Bloomfield V. A., Crothers D. M., Tinoco I., **2000**, *Nucleics Acids, Structures, properties and function, University Science Books* (Sausalito California), pp 88-91;

[110] Eriksson M., Nielsen P. E., 1996, *Nat. Struct. Biol.,* **3,** 410–413;

[111] Menchise V., De Simone G., Tedeschi T., Corradini R., Sforza S., Marchelli R., Capasso D., Saviano M., Pedone C., *Proc. Natl. Acad. Sci.*, 2**003**, *100*, 12021–12026;

[112] Betts L., Josey J. A., Veal J. M., Jordan S. R., 1995, *Science,* **270**, 1838–1841;

[113] Rasmussen H., Kastrup J. S., Nielsen J. N., Nielsen J. M., Nielsen P. E., *Nat. Struct. Biol., 1997,* **4,** 98–101; (b)Haaima G., Rasmussen, Schmidt G., Jensen D. K., Kastrup J. S., Stafshede P. W., Norden B., Buchardt O., Nielsen P. E., 1999, *New J. Chem.,* **23**, 833–840;

13. Eldrup A. B., Nielsen B. B., Haaima G., Rasmussen H., Kastrup J.,Christensen C., Nielsen P. E., **2001**, *Eur. J. Org. Chem.,* **9**, 1781–1790;

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

[115] Diederichsen U., Schmitt H. W., *Angew. Chem. Int. Ed.*, **1998**, *37(3),* 302-305;

[116] Yagil G., Sussman J. L., *EMBO J.*, **1986**, *5(7)*, 1719- 1725;

[117] Cuesta-Seijo J.A., Sheldrick G.M., Zhang J., Diederichsen U., *Crystal Structure of an alternating D-Alanyl, L-Homoalanyl PNA, PDB code 3C1P*, 2009;

[118] Wittug P., Eriksson M., Lyng R., Nielsen P. E., Norden B., *J. Am., Chem., Soc.*, **1995**, *117 (41)*, 10167-10173;

[119] Rasmussen H., Liljeforors T., Peterson B., Nielsen P. E, *J. Biomol. Struc. Dynam.*, **2004**, *21(4)*, 495-502;

[120] Berova N., Pescitelli G., Di Bari L., *Chem. Soc. Rev.*, **2007**, *36*, 914–931;

[121] Hannah K.C., Armitage B., *Acc. Chem. Res.*, **2004**, *37(11)*, 853;

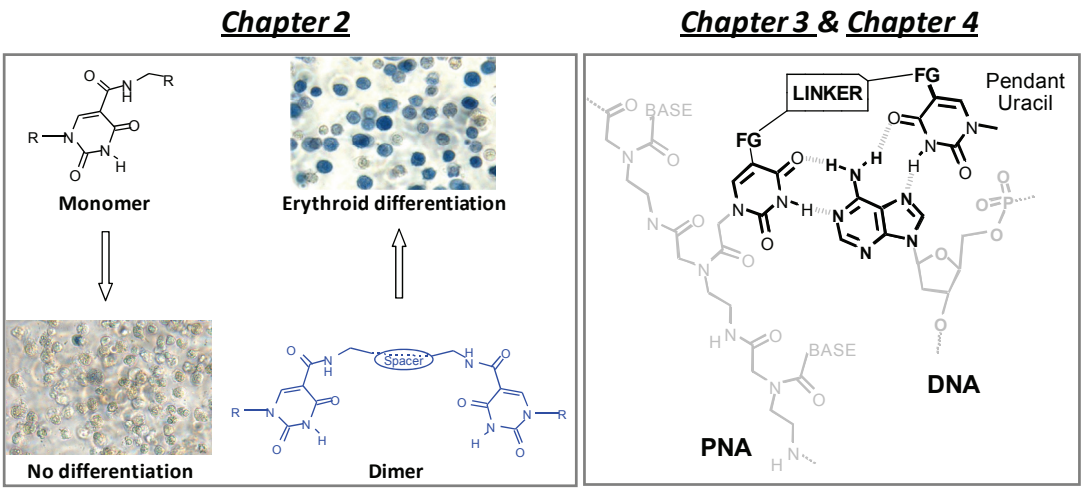
[122] Nielsen P. E., *Acc. Chem. Res.*, **1999**, *32* (7), 624–630;

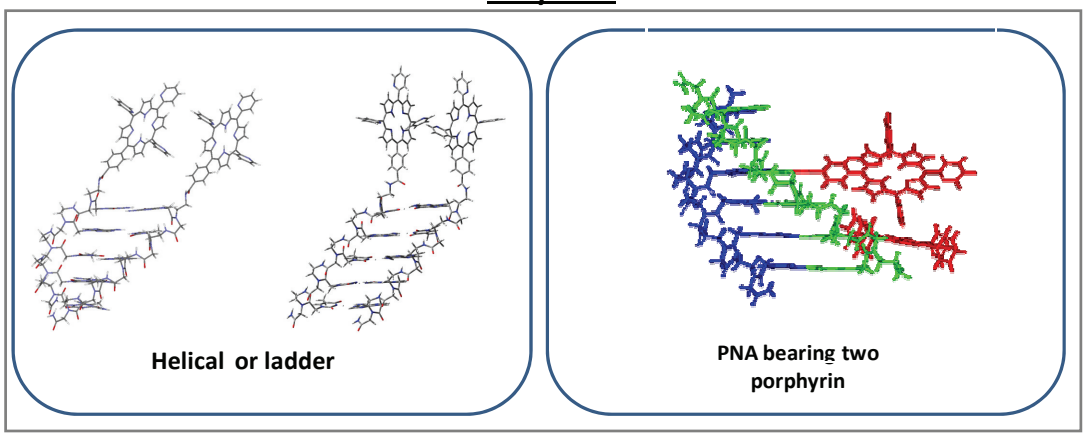
[123] Nielsen P. E., *Chem. & Biodiversity*, **2007**, *4*, 1996;

[124] Nielsen P., *Scientific American*, **2008**, *299**(6)*, 64-71.

Thesis Outlook

As pointed out from the title of this thesis we used modified uracils and porphyrins to add new function to Peptide Nucleic Acids.





As pointed out from the title of this thesis we used modified uracils and porphyrins to add new function to Peptide Nucleic Acids. The most one core idea behind this work originates from the quest of new tailored adenine receptors.

The model was designed from observation of triple helices PNA:DNA:PNA, and in order to build mimics the uracils of TAT triplets, we synthesized a series of uracil dimers. These uracil dimers showed interesting biological properties probably in connection with their ability to cooperatively interact with adenine, since monomeric reference compounds were found to be inactive. Using the same geometry we synthesized a PNA monomer containing novel modified dimeric uracils. This novel nucleobases performed better than thymine for the adenine recognition on complementary DNA, conferring better selectivity and affinity. Most of the time dedicated to this work was spent to synthesize the uracil derivatives, the challenge was to adapt uracil chemistry to the specific design ("engineering") of new PNA derivatives.

In the remaining part of the thesis we developed new tools for supramolecular conformations of PNA containing duplexes using by porphyrins conjugations. In particular we explored the parallel model and a more general model based on porphyrin embedded on C(5) of uracil. Part of this work was carried out under the supervision and in collaboration with Professor Nina Berova and Dr. Ana Petrovic at Chemistry Department of Columbia University, New York.

Alessandro Accetta

CHAPTER 2

New Uracil Dimers Showing Erythroid Differentiation Inducing Activities

2.1 Introduction

Pyrimidine derivatives are very important constituents of naturally occurring bioactive compounds, including natural nucleobases and their analogues[1] A series of pyrimidine derivatives are currently used as drugs; for example, fluorouracil has cytostatic effects and is currently used in cancer therapeutics, and azidothymidine (AZT) was the first applied drug for HIV treatment.

The potentiality of pyrimidine compounds is linked to the possibility to be used as antagonists in the biosynthetic pathways of pyrimidine nucleobases or in other important processes, by competing for the same binding sites of naturally occurring compounds. For example oxime libraries based on dimeric uracil derivatives have been proposed for the development of uracil DNA glycosylase (UNG) inhibitors.[2] 5-substituted uracil derivatives have been described as cytostatic and antiviral compounds.[3]

Another possible use of pyrimidine analogues has recently been explored in connection with the synthesis of tailor-made modified nucleobases to be included in DNA analogues, such as modified oligonucleotides[4,5] or peptide nucleic acids (PNAs).[6,7] For example, extended aromatic analogues of cytosine (named “G-clamp”) bearing extra hydrogen bonding sites has been shown to be able to strongly bind guanine by simultaneous formation of Watson-Crick and Hoogsteen base pairing.[8]

The ability to bind to specific sites of DNA is also a characteristic of many bioactive molecules able to act as antibiotics, antiproliferative or differentiating drugs. In this respect, several in vitro experimental systems are available for screening purposes. Among them, the K562 cell line, isolated and characterized by Lozzio and Lozzio from a patient with chronic

myelogenous leukemia in blast crisis,[9,10,11] has been proposed as a very useful experimental system to identify (a) antitumor compounds[12,13] and (b) inducers of erythroid differentiation and γ-globin gene expression of possible interest in the therapy of several haematological diseases, including β-thalassemia and sickle cell anaemia.[14,15,16,17,18,19,20] K562 cells exhibit a low proportion of hemoglobin-synthesizing cells under standard cell growth conditions, but are able to undergo erythroid differentiation when treated with a variety of compounds, including short fatty acids,[19] 5-azacytidine,[19] mithramycin and chromomycin,[21] cisplatin and cisplatin analogues,[17] tallimustine,[16,19] rapamycin,[22] everolimus,[23] psoralens[24] and resveratrol.[25] Following erythroid induction, a sharp increase of expression of human ε and γ globin genes is observed in K562 cells, leading to a cytoplasmic accumulation of Hb Portland ($\zeta_2\gamma_2$) and Hb Gower 1 ($\zeta_2\varepsilon_2$).[19] Several antitumor drugs were demonstrated to induce erythroid differentiation of K562 cells. Research group directed by Prof. R. Gambari (University of Ferrara) has recently demonstrated that DNA binding drugs (DBDs) exhibiting antitumor activity are powerful inducers of differentiation of K562 cells, suggesting that the expression of crucial genes involved in terminal erythroid differentiation of these cells are influenced by DBDs.[17,18,21] Several DBDs, such as tallimustine, mithramycin, cisplatin and angelicin, increase fetal haemoglobin (HbF) production in erythroid precursor cells from normal human subjects.[19] Thus, this experimental cell system appears to be suitable for the screening of molecules able to inhibit cell growth by acting on the activation of terminal differentiation pathways.

In a general project aimed at the synthesis of oligonucleotide analogues, in particular PNA, with modifications able to improve their binding activity,[26,27,28] we have designed uracil dimers connected with a spacer through the 5-position (*Figure 2.1*). The design was performed by considering as a model the very stable TAT triplet found in PNA:DNA:PNA triplex crystal structure.[29] Since these modified nucleobases could also be considered as potential drugs *per se,* they were subjected to a screening process for ability to induce erythroid differentiation and to exhibit cytotoxicity, with positive results for some of the tested compounds. In the present work we describe the synthesis of these uracil dimers and the results obtained in the evaluation of their differentiating properties, which allowed to define a new class of drug candidates, together with a rationale of the possible connection between structure and activity based on their design.

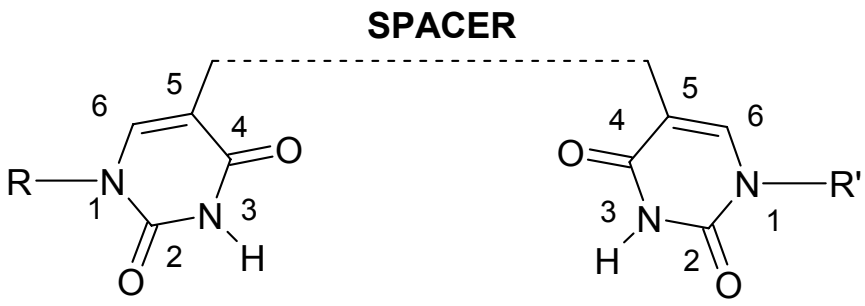


Figure 2.1. General scheme of uracil derivatives described in the present study.

2.2 Synthesis and Biological activity of 5-carboxamido Uracil dimers

Design. The general design of the uracil dimers is reported in *Figure 2.1*. Alkylation of uracil N1 is necessary either to introduce lipophilic groups or for linking this compounds to the backbone of nucleosides, nucleotides or nucleotide analogs. In all cases the N3 hydrogen was preserved, in order to keep the same hydrogen-bond pairing scheme presented by uracil. As a first part of this project, we designed a series of dimers linked through amide bonds, since molecular modelling indicated that the hydrogen bonding moieties could be preserved using this kind of modification. Since the compounds were designed in order to mimic ad reinforce the binding ability of uracil, we checked if the modification introduced is able to induce severe changes in the uracil potential binding sites, since electronegative substituents can shift the tautomeric lactamic-lactimic equilibrium.[30] The most dramatic change in the structure was the introduction of a carboxyamide moiety in the 5-position of uracil; thus, we tested the ability of these derivatives to be in the same tautomeric forms, and hence with the same recognition pattern by hydrogen bonds, by means of semiempirical calculation performed on the model compound **Mod-1** (*Figure 2.2*).[31]

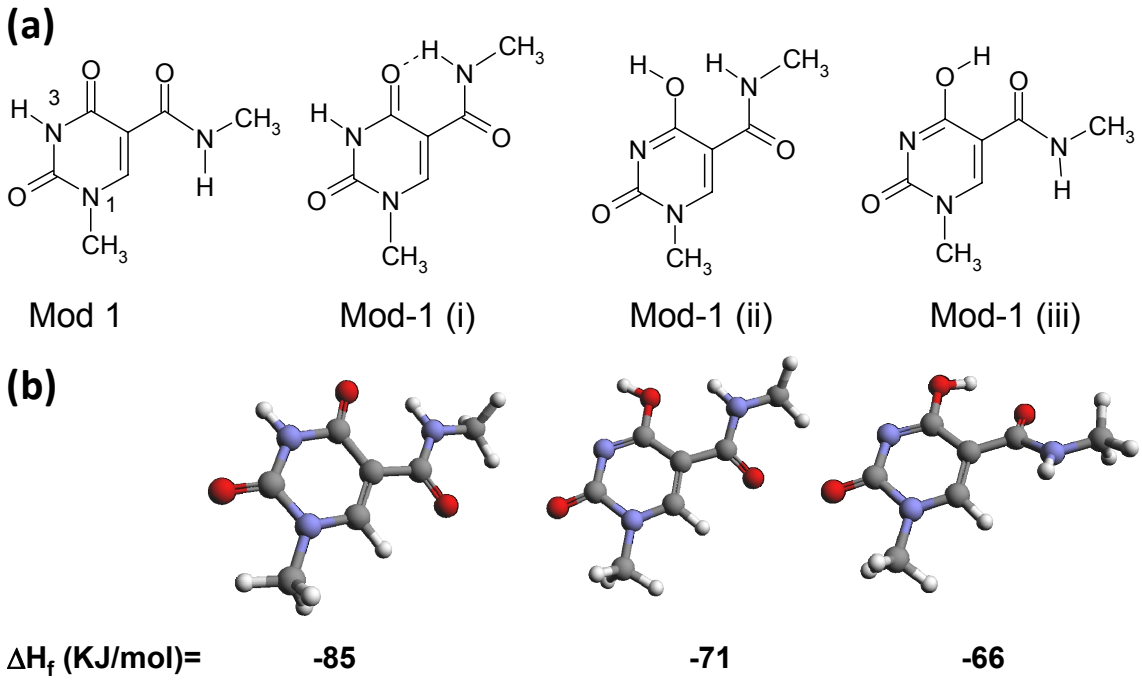
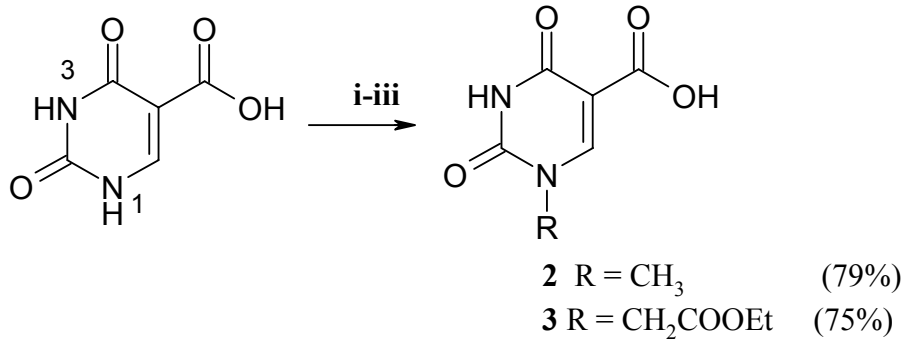


Figure 2.2. (a) Tautomeric forms of uracil-5-carboxyamide moiety considered for calculations; (b) energy-minimized structures calculated using the AM1 method, and corresponding energies in $KJmole^{-1}$.

The structure of **Mod-1** was first minimized by molecular mechanics using universal force field (UFF), and then subjected to energy minimization by a semiempirical method using Austin Model 1 (AM1) method. The same procedure was applied to the tautomer **Mod-1(ii)**. The most stable structures obtained in both cases are reported in *Figure 2.2* together with the corresponding enthalpies of formation. The most stable structure **Mod-1 (i)** has the amide hydrogen pointing towards the C4 carbonyl oxygen, thus giving rise to a intramolecular hydrogen bond. Both the structures **Mod-1(ii)** and an alternative conformation **Mod-1(iii)** show higher energies. Therefore, the 5-carboxyamide compounds are prone to exhibit the same recognition pattern of uracil as originally designed.

Synthesis. In our retrosynthetic design, we used as starting material either isoorotic acid (5-uracilcarboxylic acid) or thymine. As a reference compound, isoorotic acid methyl ester (**1**) was synthesized by reaction with $SOCl_2$ in methanol. Several derivatives were obtained with the purpose of varying the group at N1, the type (monomeric or dimeric) of the amine residue and the size and rigidity of the spacer, in order to produce chemical diversity.

Direct alkylation of isoorotic acid with reactive substrates, such as methyl or allyl halides lead to dialkylated products at both nitrogen atoms (results not shown); therefore, regioselective mono-alkylation of uracil was performed using temporary protection of the carboxylic and carbonyl oxygens with trimethylsilyl groups, through reaction with hexamethyldisilazane (HMDS) in a 3:1 excess and in the presence of trimethylchlorosilane (TMS-Cl) (*Figure 2.3*).



2 R = CH_3 (79%)
3 R = CH_2COOEt (75%)

Figure 2.3. Synthesis of N1-alkylated isoorotic acid derivatives via regioselctive alkylation of 5-orotic acid; (i) HMDS,TMS-Cl, reflux 4h; (ii) CH_3I or ethyl bromoacetate in excess, reflux,18h; (iii) H_2O/CH_3COOH, rt,20 min.

This strategy not only allows for monoalkylation, since the positively charged pyrimidinium intermediate prevents further attack, but also direct the alkylation towards the less hindered N1 position. Confirmation of the position of the substituent in compounds **2** and **3** was obtained by NOE effects between the alkyl group and the CH(6) of uracil as measured by 1D and 2D NOESY spectra (*Figure 2.4*).

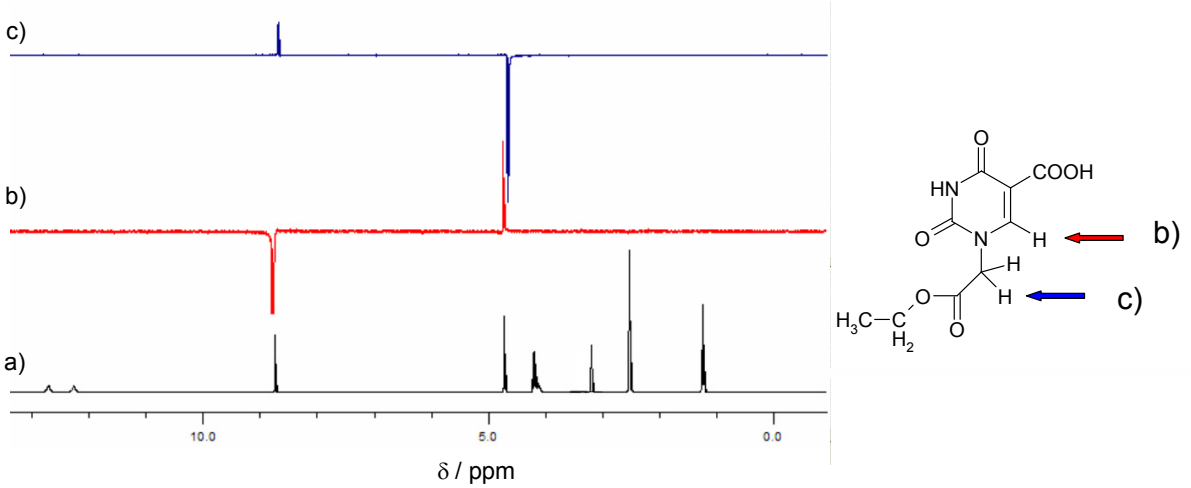


Figure 2.4. 1D NOESY spectra (DMSO d_6, 300 Mhz) of the compound 3, showing the regioselective alkylation at N1. (a) Normal spectrum; (b) selective 1D NOE obtained by selective pulse on CH(6); (c) selective 1D NOE obtained by selective pulse on the CH_2 signal. Pulse program: selnoegp, mixing time: 0.5 s; NOE signals are plotted as positive.

Reaction of thymine with more hindered long chain primary haloalkanes lead to selective monoalkylation at N1 (*Figure 2.5*), thus affording the subtrate **4** suitable for the synthesis of uracil derivatives with a C_8 alkyl chain at N1. Oxidation of methyl group of **4** with $K_2S_2O_8$ in the presence of copper(II) leads to the uracil-5-carboxaldehyde which was then oxidized by reaction with sodium chlorite to the corresponding carboxylic acid **5** bearing a C8 alkyl chain at N1 position.

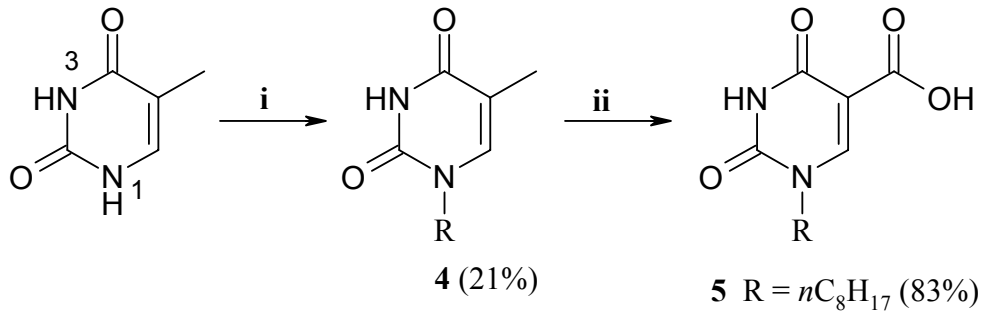


Figure 2.5. Synthesis of N1-alkylated isoorotic acid derivatives via oxidation of thymine derivatives. (i) [Br-nC_8H_{17}, NaH, DMF, 80 °C,4h]; (ii) [(a) 2,6-lutidine, $K_2S_2O_8$, $CuSO_4$, H_2O/AcCN, 80°C, 1,5h; (b) $NaClO_2$, NaH_2PO_4; t-BuOH/THF, rt, 24h].

According to our model, the 2,7-di(aminomethyl)naphthalene (**8**) was one good candidate as linking group for allowing cooperative binding to adenine. Therefore this compound was synthesized using substitution of the corresponding bromide (**6**) with tritylamine, followed by acidic solvolysis of the tritylated amine (**7**) (*Figure 2.6*).[32]

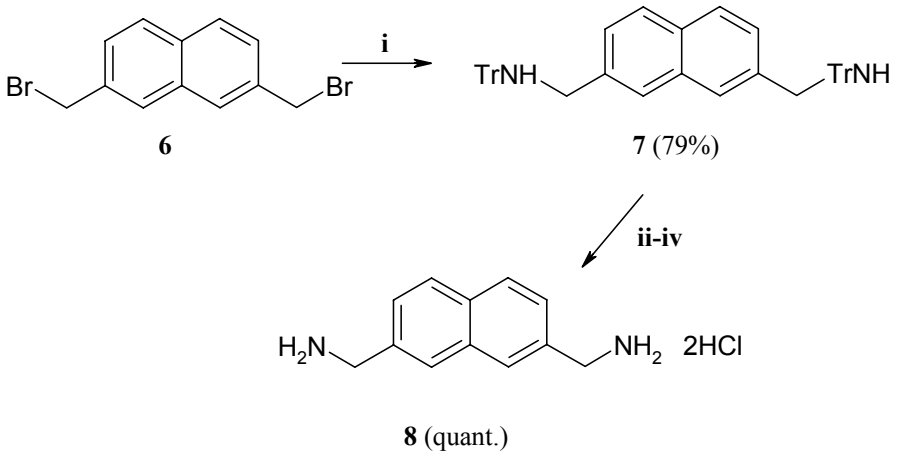


Figure 2.6. Synthesis of 2,6-dimethylaminonaphthalene. (i) [$TrNH_2$, AcCN, 50 °C, 72h]; (ii) TFA, DCM, rt, 35 min; (iii) MeOH, rt, 2h; (iv) HCl 1M in MeOH.

Several other commercially available diamines were used as linkers for the reaction with the carboxylic derivatives, butylamine was used for generating reference monomeric compound **10** (*Figure 2.7*) and benzylamine was used for generating compound **12**, both containing a uracil moiety with a carboxyamide group at C-5 (*Figure 2.7*).

A first series of derivatives, containing either a methyl or an ethoxycarbonylmethyl groups at N1 was synthesized using reaction with thionyl chloride to generate the corresponding acyl chloride, followed by reaction with the corresponding amine in pyridine (*Figure 2.7a*).

Since the yields obtained with this method were not optimal (25-30%), mainly due to loss of product during work-up, a second series of derivatives was obtained by activation of the carboxylic moiety of **5** with fluoride using 2,4,6-trifluoro-1,3,5-triazine as fluorinating agent, thus providing the stable intermediate **11** which could be isolated.

Subsequent reaction of the acyl fluoride **11** with the corresponding diamine or monoamine in acetonitrile gave the compounds **12-17** (*Figure 2.7b*). Yields were in the range 37-66%, with lowest value for the short, rigid *m*-xylylene bridge.

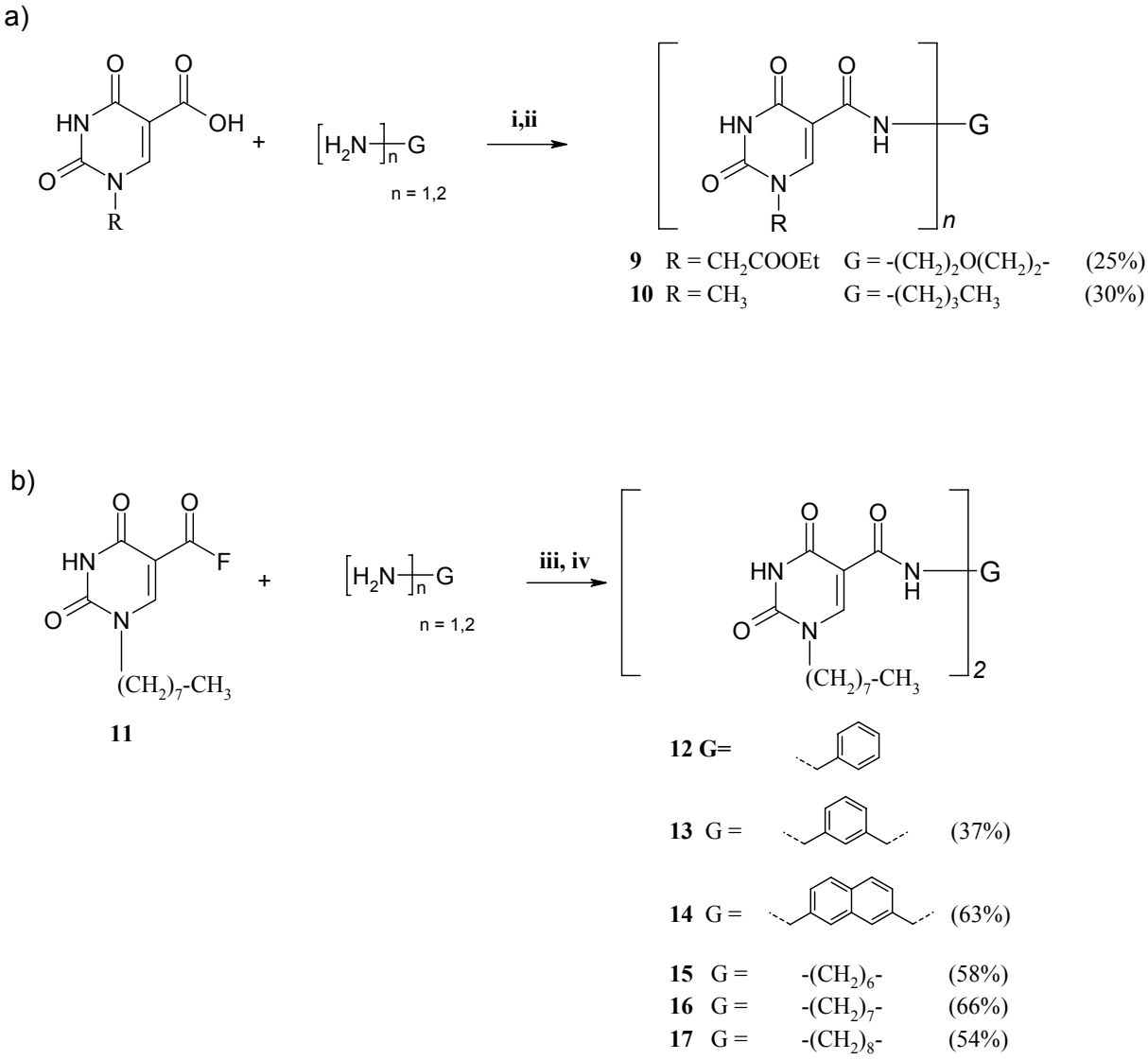


Figure 2.7. Synthesis of monomeric and dimeric uracil derivatives. a) synthesis of 12-13 via acyl chloride (i) $SOCl_2$, DMF, 70-80 °C, 2h; (ii) $(H_2N)_n$-G, Py, 2h. b) 19-23 via acyl fluoride. (iii) H_2N-G-NH_2 o H_2N-G-NH_2 2HCl, DIEA, AcCN, 80 °C, 7h; (iv) HCl 1M, 0.5 h].

Adenine binding. Dimers depicted above, bearing two *n*-octyl chains on N(1), were designed for adenine complexation in organic apolar solvent for evaluation of their binding ability in order to select more appropriate diamine spacer. Unfortunately for the more interesting compound **13** and **14** designed for having better preorganization, they were found to be poorly soluble in organic solvent like chloroform or toluene. However more flexible compound **15**, **16** and **17** showed enough solubility for preliminary studies. In particular compound **16**, having a 1,7-eptadiamine diamine spacer, was studied by ^{1}H-NMR in presence of increasing amounts of 9-ethyladenine (*Figure 2.8*).

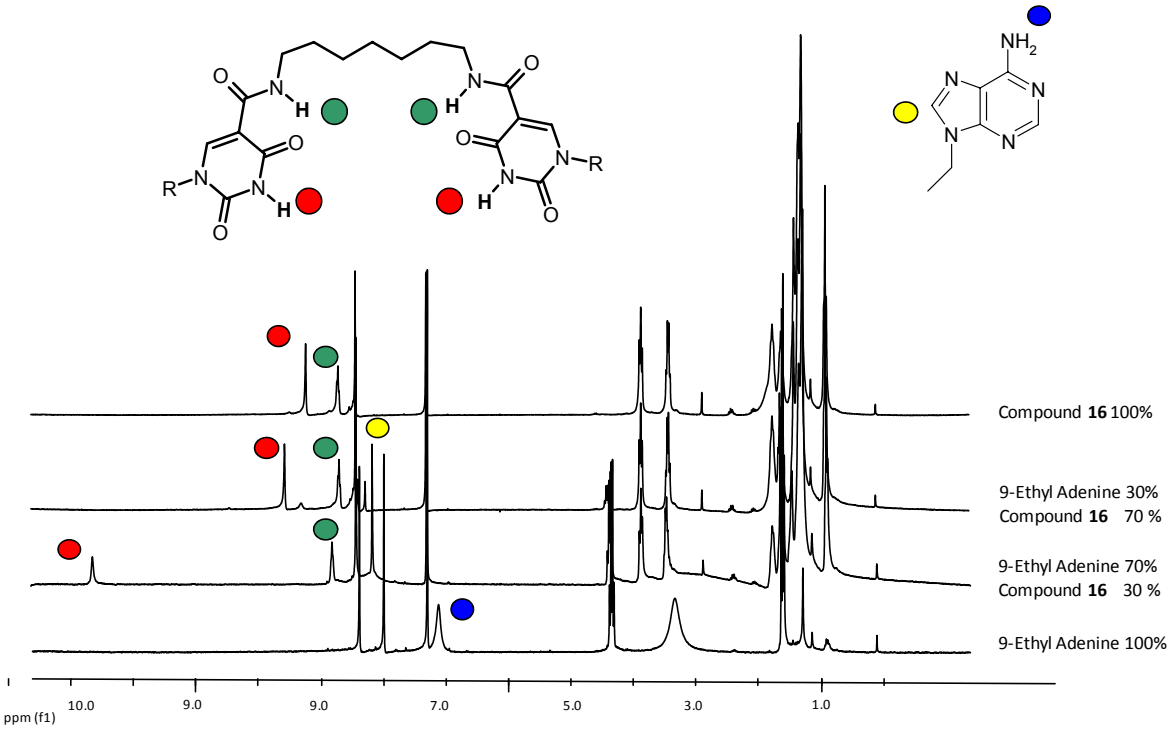


Figure 2.8. ^{1}H-NMR spectra of compound **16** (host) and **9-ethyladenine** (guest) at different ratios. The sum of concentration of host and guest was kept constant to 10 mM. Spectra were recorded at 300MHz, 25°C in $CDCl_3$.

Since the complexation process is fast on NMR timescale, we followed ^{1}H-NMR shift of NH proton involved in complexation by evaluation of different guest/host ratio. Starting from 100% of host (compound **16**) and progressively increasing guest/host ratio a very large downfield shift of uracils NH proton was observed (*Figure 2.8,* red circle), while the position of NH amide signal of the host (green circle) remained almost constant.

This indicates that only NH of uracils interact with 9-ethyladenine, in line with the proposed model of interaction. In order to evaluate the ability to form complexes of the rigid linked dimers we used a more polar chloroform/DMSO solvent mixture. In this mixture we observed a limited shift of uracils NH, due to more competitive environment, however more pronounced downfield shift were observed for compound **14** (data not shown) compared to other compounds.

In literature it is reported that uracils derivatives are potentially biologically active compound. The above described library of small uracil dimers was subjected to screening of biological properties and some compounds were eventually found to act as antiproliferative and erythroid differentiation inducing drugs. The biological studies reported in the next section were therefore carried out by the group of *Prof. Roberto Gambari* of *University of Ferrara*.

Screening of antiproliferative and differentiation inducing properties First for all on the synthesized molecules the effects on cell proliferation were determined. To this aim, K562 cells were cultured in the presence of increasing concentrations of compounds and cell number/ml was determined after 3, 4 and 5 days. These time points were selected because between day 3 and 5 untreated control K562 cells are on the log phase of cell growth. A representative experiment is shown in *Figure 2.9*, in which the effects on cell growth (panels A and C) and erythroid differentiation (panels B and D) of compounds **9** and **14** are compared. Compound **9** displays very low growth inhibiting properties, while the IC_{50} for compound **14** was about 500 μM. Erythroid differentiation of the compounds under investigation was studied by determining the proportion of benzidine-positive (hemoglobin containing) cells. As clearly evident by *Figure 2.9* (B and D), compound **14** was found to stimulate erythroid differentiation, while compound **9** was not active (*Figure 2.9*, A and B).

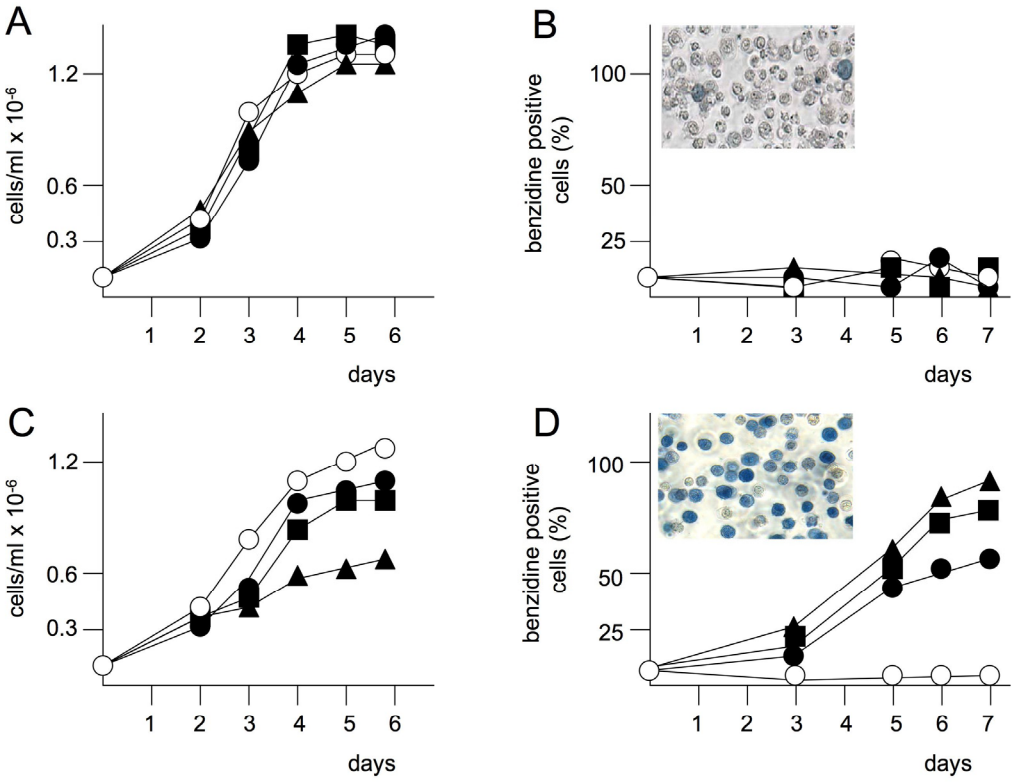


Figure 2.9. Effects of compound 9 and 14 on in vitro proliferation (A and C) and erythroid differentiation (B and D) of K562 cells. Cells were cultured in the absence (○) or in the presence of increasing amounts (●, 400 µM, ■, 600 µM, ▲, 800 µM) of compound 9 (A and B) and compound 14 (C and D). After the indicated length of time cell number/ml and the proportion of benzidine-positive cells were determined. Representative images of benzidine staining assay of 6 days K562 cell cultures treated with 600 µM compound 9 and 14 are shown in the relative inserts. The results of three independent experiments are reported in *Table 2.1*.

Table2.1 indicates the antiproliferative effects (IC_{50} values) and the erythroid induction ability (% of benzidine-positive cells) of all the tested compounds. The best erythroid induction ability was displayed by compound **14**. The data shown in *Table 2.1* were obtained using concentrations of compounds approaching those giving 50% of inhibition of cell growth (these concentrations were chosen to better compare the potential erythroid inducting activity in experimental conditions leading, for most of the compounds tested, to similar effects on cell proliferation rate).

In addition, it should be noted that compound **14** was able to induce erythroid differentiation of K562 cells after 6 days cell culture even if added at concentrations lower than that shown in *Table 2.1* (an average of 52 ± 4.5 % of benzidine-positive cells was obtained in four independent experiments with 400 μM compound **14**).

Table 2.1. Antiproliferative effects (IC_{50}) and percentage of benzidine-positive cells (%B+) after treatment with the various compounds and concentration used (C).

Compound n.	Antiproliferative effect IC_{50} (μM)	% of benzidine positive cells after 6 days	concentration (μM)
1	>800	5 ± 3.4	800
5	247 ± 33	30 ± 5.8	300
9	>800	1 ± 0.8	800
10	536 ± 45	5 ± 2.3	800
12	75 ± 7.3	5 ± 3.3	100
13	247 ± 23	40 ± 8.4	200
14	517 ± 63	78 ± 7.3	600
15	600 ± 85	40 ± 5.5	600
16	220 ± 35	50 ± 6.8	600
17	420 ± 93	45 ± 3.5	600

Results are presented as average ± SD of three independent experiments performed). The IC50 was calculated as the concentration of compounds necessary to decrease cell number (after 4 days culture period) at 50% of the values obtained in control untreated K562 cell cultures. The % of benzidine-positive (hemoglobin-containing) cells was determined after 6 days induction period at concentrations of the tested compounds indicated in the right column.

Under these experimental conditions no inhibition of cell growth was detectable. In *Table 2.2* the effect of compound **14** is compared with those of other erythroid inducers of K562 cells: the very high induction level suggests that **14** is indeed a very active inducer with activity comparable to other previously reported compounds in this cell line.

Table 2.2. Effects of compound 14 on *in vitro* growth and erythroid differentiation of human leukemic K562 cells.

Compound	concentration	Erythroid induction[a] (% of benzidine-positive cells)
compound 14	600 μM	78 ± 7.3
Ara-C	500 nM	78.3 ± 4.5
mithramycin	100 nM	86.8 ± 8.3
rapamycin	1.0 mM	75.5 ± 7.5
butyric acid	2.0 mM	32.5 ± 3.4

[a]Results are presented as average ± SD (three independent experiments performed) of the % benzidine-positive (hemoglobin-containing) cells after 6 days induction period at the indicated concentrations of the tested compounds.

The antiproliferative effect of compound **14** was obtained using other tumor cell lines as cellular targets, including the rhabdomyosarcoma RD[33] and the two breast carcinoma MCF-7[34] and MDA-MB-231[35] cell lines (*Table 2.3*). Low antiproliferative activity was found when the human cystic fibrosis bronchial IB3-1 cell line[36,37] was used (*Table 2.3*). As far as the effects on K562 cell growth displayed by all the compounds tested (*Table 2.1*), some compounds (for instance compound **1** and **9**) were found to be not effective in inhibiting cell growth under the experimental conditions employed; all the other compounds were found to be moderately efficient in inhibiting *in vitro* cell growth of treated K562 cells. As shown for compound **14** *(Table 2.3*), the antiproliferative activity of the compounds here studied is not restricted to K562 cells (data not shown).

Table 2.3. Effects of compound 14 on *in vitro* growth of human cell lines.

Cell line	Phenotype/origin	Antiproliferative effect[a] (IC_{50})
K562	chronic myelogenous leukemia (CML)	517 ± 63
RD	rhabdomyosarcoma	76.6 ± 9.3
MDA	breast cancer	171.2 ± 27.5
MCF7	Breast cancer	197.4 ± 32.4
IB3-1	Bronchial epithelial cell line (cystic fibrosis)	505 ± 58

[a]Results are presented as average ± SD (three independent experiments performed) of concentration needed to obtain 50% inhibition of cell growth after a 4 cell culture period.

Our results suggest that, among the compounds studied, some are efficient to induce erythroid differentiation without exhibiting strong antiproliferative activity (i.e. compound **14**); conversely, some other compounds (i.e. compound **12**) strongly inhibit cell growth of K562 cells without inducing differentiation (*Figure 2.10*).

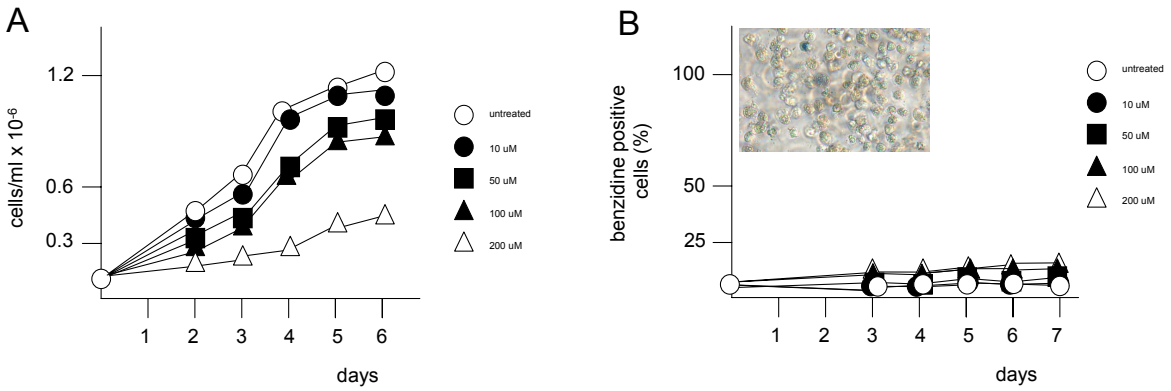
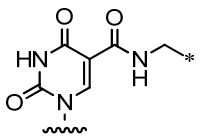


Figure 2.10. Effects of compound 12 on in vitro proliferation (A and erythroid differentiation (B) of K562 cells. Cells were cultured in the absence (○) or in the presence of increasing amounts (●, 10 µM, ■, 50 µM, ▲, 100 µM and △ 200 µM) of compound 12. After the indicated length of time cell number/ml and the proportion of benzidine-positive cells were determined. Representative image of benzidine staining assay of 6 days K562 cell cultures treated with 100 µM compound 12 are shown in the insert.

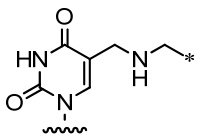
When the experiments were conducted at concentration higher than that reported in *Table 2.1*, compound **12** was confirmed to be inactive in inducing differentiation (analysis was conducted at 200, 400 and 600 µM without any effects in stimulating the increase of the proportion of benzidine positive K562 cells). Further experiments employing proteomic and transcriptomic analyses are necessary in order to understand the interplay between effects of this class of molecules on cell growth and erythroid differentiation.

2.3 Uracil 5-methylamino derivatives

The biological property of uracils dimers showed above were found to be interesting and opened the quest for more efficient dimers. One of the major problem of "first generation" dimers depicted above stands in their solubility. We designed the synthesis of new class of dimers by substitution of amide linker on C(5) with more flexible methylamino linker (*Figure 2.11*); this modification should depress part of self-aggregation of dimers in favour of solubility since secondary amine is protonate in physiological conditions. Synthesis of these dimers were designed by reaction of 5-formyl-uracils with different diamine linker through reductive amination reaction.



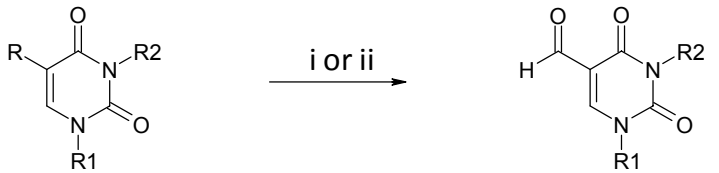
5-Carboxamido Uracil



5-methylamino Uracil

Figure 2.11. 5-Carboxamido-Uracil and 5-methylamino-Uracil.

The synthesis of 5-formyl-uracils was carried out by oxidation of N(1)-alkylated thymine or by formylation under Vielsmeyer-Haack[38] conditions of N(1)-alkylated-N(3)-benzoylated uracil (*Figure 2.12*).



4 R: *Me*, R1: *n-octyl*, R2:*H*

18 R: *H*, R1: *n-pentyl*, R2: 

19 R1: *n-octyl*, R2:*H* (32%)

20 R1: *n-pentyl*, R2:  (44%)

Figure 2.12. Synthesis route for 5-formyluracil. (i) [2,6-lutidine, MeCN, $CuSO_4$, $K_2S_2O_8$, H_2O, rt., 80°C, 2h]. (ii) [$POCl_3$, DMF, 80°C, overnigth].

In particular the Vielsmeier-Haack formylation of compound **18** afforded the N(3)-protected 5-formyluracils in good yield. The reactivity under reductive amination conditions of these 5-formyl uracil substrates is ongoing.

2.4 Conclusion

In the present work we have demonstrated for the first time that uracil derivatives bearing a N1-alkyl chain and a 5-carboxyamide linker can bind adenine substrates through a rationally designed hydrogen-bonding interactions. These molecules were found to act as a new class of erythroid differentiation inducers, and that dimeric derivatives with suitable spacers have the best performing characteristics: low cytoxicity and higher differentiating ability. The dependence of the induction of increase of hemoglobin-containing cells upon both the length and rigidity of the linking moiety suggests a cooperative mechanism. Furthermore, the best results were obtained with the compound bearing a naphthalene linker, which avoids collapse of the uracil moieties, indicating that a possible recognition of complementary functionalities (such as adenine derivatives) could be implicated in the production of biological properties. Further studies are needed to evaluate the biological mechanisms implicated in this process.

This findings can be the starting point for the quest of more effective and specific drugs for the induction of terminal erythroid differentiation, ultimately leading to new insights in the treatment of neoplastic diseases with molecules acting by inducing differentiation rather than simply by exerting cytotoxic effects. In addition, these molecules might be of interest for the experimental treatment of β-thalassemic erythroid cells for which the induction of γ-globin mRNA could be very beneficial.[19,21] In this respect it has been demonstrated that inducers of K562 erythroid differentiation are often able to induce fetal hemoglobin production in erythroid cells isolated from β-thalassemia patients.[19]

2.5 Experimental Section

General. Chemicals were purchased from Aldrich, Fluka, Acros or Merck and have been used without further purification. TLC was performed on aluminium sheets coated with silca gel F_{254} 0.2 mm. Flash chromatography was carried out under nitrogen pressure using Merck silca gel 60H. NMR spectra were measured on Brucker AC 300 and on a Varian INOVA 600 instruments. Melting points were determined on Gallenkamp melting apparatus and are uncorrected. FT-IR spectra were recorded on Nicolet FT-IR. Mass spectra were recorded on Waters Acquity Mass Spectrometer.

Molecular Mechanics Calculations. Starting structures were generated using the sketch module of ArgusLAB. Geometry optimization was carried out using UFF (universal force field). Electrostatic terms were treated with simple cutoff (10 – 8 Å) and structures were optimized using BFGS (Broyden-Fletcher-Golfarb-Shanno) algorithm.

Semiempirical Calculations. Geometry optimization was carried out using AM1 (NDDO) module of ArgusLAB, BFGS algorithm with SCF convergence of 1.5936e-013 au for energy and STO 3G was used as basis set.[31]

Methyl 2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxylate (1). A round bottomed flask containing 50 mL of MeOH was cooled at 10 °C and $SOCl_2$ (9.62 mmol, 0.70 mL) was added carefully. After 5 minutes isoorotic acid was added and the mixture was heated and refluxed for 18 hours, then the solvent was removed, and the white powder obtained was washed many times with MeOH, suspended in hexane, and filtered to yield 1.10 g of **1** (*Yield* >99%).

(1) Dec. T > 250 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) *δ(ppm)* 3.68 (s, 3H), 8.13 (s, 1H), 11.31 (s, 1H), 11.60 (s, 1H). (75 Mhz, DMSO-d_6, 25°C) δ ppm) 50.6, 99.8, 155.2, 157.1, 161.7, 164.5. FT-IR (KBr pellet) ν (cm^{-1}) 1781, 1738, 1620. MS (ESI+): *m/z* [MH^+] 171.0, [MNa^+] 192.9. Anal. ($C_6H_6N_2O_4$) C: calcd. 42.36,; found, 41.77, H, N: calcd, 16.47; found, 15.64.

1-methyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxylic acid (2). Isoorotic acid (1.83 mmol, 300 mg) was suspended in hexamethydisilazane (HMDS) (11.50 mmol, 2.5 mL) and trimethylchlorosinalne (TMS-Cl) (1.10 mmol, 0.10 mL) was added. The mixture was refluxed in a closed tube at 120°C for 4 hours (until the mixture appeared colourless); the reaction temperature was raised to r.t. then iodomethane (30.0 mmol, 1.85 mL) was added; the mixture was then heated to 50°C and kept overnight at the same temperature. The solvent was then evaporated and the residue was stirred with 3 mL of ice water and 3 mL of glacial acetic acid for 20 min; the precipitate formed was collected by filtration, and washed with cold water and ethyl acetate to afford 246 mg of a yellow pale solid corresponding to the product *(Yield* = 79%).

(2) *M.p.* 258.6 – 259.2°C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) δ (ppm) 3.38 (s, 3H), 8.67 (s, 1H), 12.19 (s, 1H), 12.67 (s, 1H) . ^{13}C-NMR (75 Mhz, DMSO-d_6, 25 °C) δ(ppm) 36.2, 100.7, 150.0, 154.2, 163.3, 164.8. FT-IR (KBr pellet) ν (cm^{-1}) 1752, 1728, 1704, 1618. MS (ESI+): *m/z* [MH$^+$] 170.0, [MK$^+$] 209.5. Anal. ($C_6H_6N_2O_4$) C: calcd, 42.36; found, 39.61, H: calcd. 3.55; found, 4.01, N.

1-(2-ethoxy-2-oxoethyl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxylic acid (3). Compound **3** was synthesized as described above for compound **2**, using ethyl bromoacetate as alkylating agent. *(Yield* = 75%). *M.p.* 184.2 – 187.2 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) δ (ppm) 1.21 (t, 3H, *J* = 7.3 Hz), 4.16 (q, 2H, *J* = 7.3 Hz), 4.69 (s, 2H), 8.71 (s, 1H), 12.24 (s, 1H), 12.67 (s, 1H) . ^{13}C-NMR (75 Mhz, DMSO-d_6, 25 °C) δ (ppm) 13.8, 49.2, 61.3, 102.2, 149.7, 153.5, 163.1, 163.3, 167.4. FT-IR (KBr pellet) ν (cm^{-1}) 1793, 1740, 1712, 1625. MS (ESI+): *m/z* [MH$^+$] 243.0, [MNa$^+$] 265.0. Anal. ($C_9H_{10}N_2O_6$) C: calcd, 44.63; found, 43.91, H: calcd. 4.16; found, 4.90, N: calcd., 11.57; found, 10.54.

5-methyl-1-octylpyrimidine-2,4(1H,3H)-dione (4). Compound **4** was synthesized according to a literature procedure.[39] *M.p.* 111.9 – 113.5 °C (*Yield* = 22%). ^{1}H-NMR (300 Mhz, $CDCl_3$, 25 °C) δ (ppm) 0.82(t, 3H, *J* = 6.9 Hz), 1.21 – 1.25 (m, 10H), 1.61 (q, 2H, *J* = 6.5 Hz), 3.64 (t, 2H, *J* = 6.5 Hz), 6.96 (s, 1H), 10.09 (s, 1H). ^{13}C-NMR (75 Mhz, $CDCl_3$, 25 °C) δ (ppm) 12.2, 13.9, 22.5, 26.3, 29.0, 29.1, 31.6, 48.4, 110.4, 140.4, 151.1, 164.8. FT-IR (KBr pellet) ν (cm^{-1}) 3160, 3067, 3029, 2955, 2926, 2854, 1692, 1653. MS (ESI+): *m/z* [MH$^+$] 239.2, [MNa$^+$] 261.3, [MK$^+$] 277.2. Anal. ($C_{13}H_{22}N_2O_2$) C, H, N.

1,2,3,4-tetrahydro-1-octyl-2,4-dioxopyrimidine-5-carboxylic acid (5) . A solution of **4** (2.55 g, 10.2 mmoles) and 2,6-lutidine (4.3 mL) in acetonitrile (40 mL) was added to a stirring solution of $K_2S_2O_8$ (5.75 g, 21 mmoles) and $CuSO_4$ (0.66 g, 4.2 mmoles) in water (40mL). The reaction was heated at 80°C for 2 hours, then the mixture was dried and the residue was partitioned between ethyl acetate and saturated aqueous solution of EDTA. The ethyl acetate was removed from the organic extract and the resulting yellow oil was dissolved in 100 mL of t-BuOH – THF – *i*-butene (6:3:1) mixture. To this mixture a solution of $NaClO_2$ (12 g, 105 mmoles) and $Na_2H_2PO_4$ monohydrate (7.5 g, 52.5 mmoles) in water (25 mL) was added dropwise over a period of 30 minutes, then the reaction was stirred at room temperature overnight. The solvent was removed and the residue partitioned between saturated aqueous $KHSO_4$ and ethyl acetate. The organic phase was evaporated and the residue was partitioned between methylene chloride and aqueous NaOH (2 M). The aqueous extract was neutralized with concentrated HCl at pH 3, and a pale-yellow solid precipitated. The solid was collected by filtration and washed with abundant cold water, and was recrystallized from acetone – water to obtain 1.54 g (83%) of desired product.

(5)*M.p.* 127.2 – 128.1 °C. ^{1}H-NMR (300 Mhz, $CDCl_3$, 25°C) δ (ppm) 0.88 (t, 3H, J = 6.9 Hz), 1.26 – 1.33 (m, 10H), 1.70 – 1.77 (m, 2H), 3.88 (t, 2H, J = 7.7 Hz), 8.46 (s, 1H), 9.40 (s, 1H), 12.16 (s-broad, 1H). ^{13}C-NMR (75 Mhz, $CDCl_3$, 25°C) δ (ppm) 14.0, 22.5, 26.2, 28.9, 31.6, 50.5, 102.0, 149.0, 152.7, 162.9, 165.0. IR (KBr pellet) ν (cm^{-1}) 3461, 3420, 3179, 3056, 2956, 2924, 2855, 1744, 1701, 1668. MS (ESI+): *m/z* [MNa^+] 291.2. Anal. ($C_{13}H_{20}N_2O_4$) C, H, N.

N,N'-[naphthalene-2,7-diyldi(methylene)]bis(1,1,1-triphenylmethanamine) (7). 2,7-dibromonaphthalene (1.92 mmol, 610 mg) and tritylamine (7.71 mmol, 2.0 g) were dissolved in 30 mL of acetonitrile and heated at 50 °C for 96 hours; then the solvent was removed and the residue was partitioned between saturated $NaHCO_3$ and DCM. The organic layer, was dried with Na_2SO_4, filtered, and the solvent was evaporated. The solid obtained was purified by flash chromatography on silca gel using a hexane – dichloromethane 6/4 mixture, to afford 990 mg of **7** as a white solid. (*Yield* = 79%). *M.p.* 221.6 – 223.3 °C.

(7)^{1}H-NMR (300 Mhz, $CDCl_3$, 25 °C) δ (ppm) 1.28 (s, 2H), 3.51 (s, 4H), 7.21 – 7.26 (m, 6H), 7.31 – 7.36 (m, 12H), 7.46 (dd, 2H, J_1 = 8.4 Hz, J_2 = 1.3 Hz), 7.61–7.64 (m, 12H), 7.79 (d, 2H, J = 8.4 Hz), 7.88 (d, 2H, J = 1.3 Hz). ^{13}C-NMR (75 MHz, $CDCl_3$, 25 °C) δ (ppm) 48.0, 71.0, 125.7,

126.2, 126.3, 126.3, 127.6, 127.9, 128.6, 131.7, 133.5, 138.6, 146.0. FT-IR (KBr pellet) ν (cm^{-1}) 3315, 3057, 3019, 2850, 1488, 1448. MS (ESI+): *m/z* [M-2Tr-NH_3^+] 170.1, [Tr^+] 243.2. Anal. ($C_{50}H_{42}N_2$) C: calcd, 89.51; found, 88.85; H, N.

Naphthalene-2,7-diyldimethanamine hydrochloride (8). Compound **7** (1.48 mmol, 960 mg) was dissolved in a mixture of 10 mL of TFA and 7 mL of DCM (a intense yellow colour appears immediately after the addition of TFA) and stirred at room temperature for 40 minutes, then 10 mL of MeOH was added and the mixture was stirred for 1 hour until it became colourless. The solvent mixture obtained was evaporated and the resulting oil was dissolved in the minimum amount of 2M HCl in MeOH and stored at -20°C for 1 hour. A white precipitate appeared, which was separated by centrifugation and washed with diethyl ether, to afford 365 mg of the corresponding diamine chloridrate. (*Yield* = 96%). *M.p.*

(8) Dec. T > 285 °C. ^{1}H-NMR (300 Mhz, D_2O, 25 °C) *δ* (ppm) 4.33 (s, 4H), 7.59 (dd, 2H, J_1 = 8.5 Hz, J_2 = 1.7 Hz), 7.98 (d, 2H, *J* = 1.7 Hz), 8.02 (d, 2H, *J* = 8.4 Hz). ^{13}C-NMR (75 Mhz, D_2O, 25°C) *δ* (ppm) 46.5, 130.2, 131.6, 132.3, 134.5, 136.1, 136.2. FT-IR (KBr pellet) ν (cm^{-1}) 3125, 3033, 3013, 2997, 2965, 2867, 1599, 1495, 1480. MS (ESI+): *m/z* [MH^+-NH_3] 170.0, [MH^+] 187.1. Anal. (free amine) ($C_{12}H_{14}N_2$) C: calcd, 77.38; found, 76.77, H, N.

Diethyl 2,2'-{oxybis[ethane-2,1-diyliminocarbonyl(2,4-dioxo-3,4-dihydropyrimidine-5,1-diyl)]}diacetate (9). 3 mL of DMF dry, compound **3** (0.88 mmol, 150 mg) and $SOCl_2$ (1.80 mmol, 130 µL) were heated to 80 °C for 2 hours under N_2 atmosphere, then 2-(aminoetoxy)etanammine (0.44 mmol, 45 mg) and 3 mL of pyridine dry were added. After 2 hours the solvent was removed under reduced pressure and the resulting oil was suspended in 10 mL of water. A yellow pale precipitate appeared, and was collected and washed with methanol and ethyl ether to afford 45 mg of **9**. (*Yield* = 25%).

(9) Dec. T > 285°C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) *δ* (ppm) 3.35 (s, 6H), 3.41 (q, 4H, *J* = 6.0 Hz), 3.50 (t, 4H, *J* = 6.0 Hz), 8.45 (s, 2H), 8.84 (t, 2H, *J* = 6.0 Hz), 11.77 (s, 2H) . ^{13}C-NMR (75 Mhz, DMSO-d_6, 25 °C) *δ* (ppm) 35.9, 38.2, 67.7, 104.0, 150.3, 151.6, 161.7, 163.5. FT-IR (KBr pellet) ν (cm^{-1}) 1730, 1696, 1611, 1636. MS (ESI+): *m/z* [MH^+] 409.1, [MNa^+] 431.1. Anal. ($C_{22}H_{28}N_6O_{11}$) C: calcd, 47.83; found, 45.92; H; N: calcd, 15.21; found, 14.79.

N-butyl-1-methyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (10). 3 mL of DMF dry, compound **2** (2.35 mmol, 400 mg) and $SOCl_2$ (4.70 mmol, 341 µL) were heated to 80 °C for 2 hours under N_2 atmosphere, then butylamine (4.70 mmol, 470 µL) and 6 mL of pyridine dry were added. After 2 hours the solvent was removed under reduced pressure and the resulting yellow oil was suspended in 4 mL of water and acidified to pH 2. The precipitate collected by filtration was washed with NaOH (1M), abundant water, MeOH and diethyl ether to afford 157 mg of a **10** as a white solid (*Yield* = 30%).

(10)*M.p.* 244.2 – 245.2 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) δ (ppm) 0.88 (t, 3H, J = 7.2 Hz), 1.29 (m, 2H, J = 6.9 Hz), 1.45 (m, 2H, J = 6.9 Hz), 3.25 (q, 2H, J = 6.3 Hz), 3.36 (s, 3H), 8.46 (s, 1H), 8.72 (t, 1H, J = 5.6 Hz), 11.80 (s, 1H). ^{13}C-NMR (75 Mhz, DMSO-d_6, 25°C) δ (ppm) 13.5, 19.4, 31.1, 35.9, 37.8, 104.1, 150.3, 151.6, 161.5, 163.7. FT-IR (KBr pellet) ν (cm^{-1}) 3441, 3306, 3170, 3047, 2951, 2874, 2835, 1738, 1683, 1615. MS (ESI+): *m/z* [MH^+] 226.2. Anal. ($C_{10}H_{15}N_3O_3$) C, H, N.

1,2,3,4-tetrahydro-1-octyl-2,4-dioxopyrimidine-5-carbonyl fluoride (11) . To a solution of compound **5** (1.5 g, 5.57 mmoles) in 30 mL of dry dichloromethane, cyanuric fluoride (2.9 mL, 33.42 mmoles) was added dropwise, followed by addition of 0.6 mL of dry pyridine. The reaction was kept under argon and stirred at room temperature overnight. The reaction mixture was extracted with water and the organic layer was dried over Na_2SO_4 and filtered on a sintered glass (G5) funnel. The solvent was removed to yield a yellow-brown solid. Recrystallization from chloroform – hexane afforded 1.12 g (75%) of **11**.

(11)*M.p.* 125.6 – 127.3 °C. ^{1}H-NMR (300 Mhz, $CDCl_3$, 25 °C) δ (ppm) 0.87 (t, 3H, J = 6.9 Hz), 1.27 – 1.39 (m, 10H), 1.72 – 1.76 (m, 2H), 3.86 (t, 2H, J = 7.2 Hz), 8.28 (s,1H), 9.14 (s, 1H). ^{13}C-NMR (75 Mhz, $CDCl_3$, 25 °C) δ (ppm) 14.0, 22.5, 26.2, 28.9, 29.0, 29.6, 31.6, 50.5, 99.5 (d, $^2J_{CF}$ = 61.6 Hz), 149.0, 152.1 (d, $^1J_{CF}$ = 333.0 Hz), 154.6, 158.2. FT-IR (KBr pellet) ν (cm^{-1}) 3447, 3182, 3053, 2959, 2923, 2854, 1840, 1820, 1800, 1726, 1696, 1619. MS (ESI+): *m/z* [MNa^+] 293.2, [MK^+] 309.2. Anal. ($C_{13}H_{19}FN_2O_3$) C: calcd, 57.77; found, 56.85; H, N: calcd, 10.36; found, 11.37.

General procedure for compounds from (12) to (17) . A solution of **11** (0.37 mmoles) and DIEA (1.85 mmoles) in acetonitrile (6 mL) was stirred at room temperature for five minutes, then the diamine or the corresponding chloridrate (0.185 mmoles) was added and

the reaction was refluxed for 7 hours. The reaction mixture was poured into 20 mL of 1 M aqueous HCl and a yellow-brown precipitate appeared. The solid was collected by filtration and washed with water to afford the product.

N-benzyl-1-octyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide (12). Compound **12** was synthesized according the general procedure described above using an tenfold excess of benzylamine. (*Yield* = 88%) . *M.p.* 180.3 – 181.6 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) δ (ppm) 0.85 (t, 3H, *J* = 6.3 Hz), 1.21– 1.29 (m, 10H), 1.56 – 1.60 (m, 2H), 3.80 (t, 2H, *J* = 7.2 Hz), 4.48 (d, 2H, *J* = 6.0 Hz), 7.24 – 7.35 (m, 5H), 8.50 (s, 1H), 9.13 (t, 1H, *J* = 6.0 Hz), 11.83 (s, 1H). ^{13}C-NMR (75 Mhz, DMSO-d_6, 25 °C) δ (ppm) 13.9, 22.0, 25.6, 28.3, 28.5, 31.1, 42.0, 48.4, 99.5, 104.4, 126.8, 127.2, 128.3, 139.2, 150.0, 151.0, 161.7, 163.5. FT-IR (KBr pellet) ν (cm^{-1}) 3302, 3169, 3115, 3045, 2955, 2922, 2853, 1727, 1678, 1632, 1607. MS (ESI+): *m/z* [MH^+] 358.3, [M_2H^+] 715.6. Anal. ($C_{20}H_{27}N_3O_3$) C: calcd, 67.20; found, 65.47, H, N.

***N,N'*-phenylene-1,3-diylbismethylene(1-octyl-2,4-dioxo-1,2,3,4 tetrahydropyrimidine -5-carboxamide) (13).** (*Yield* = 37%) . *M.p.* Dec. T > 210 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) δ (ppm) 0.85 (t, 6H, *J* = 6.6 Hz), 1.19 – 1.30 (m, 20H), 1.55 – 1.64 (m, 4H), 3.81 (t, 4H, *J* = 7.5 Hz), 4.46 (d, 2H, 6.0 Hz), 7.16 – 7.31 (m, 4H), 8.50 (s, 2H), 9.12 (t, 2H, *J* = 6 Hz), 11.84 (s, 2H). ^{13}C-NMR (75 Mhz, DMSO-d_6, 25 °C) δ (ppm) 13.9, 22.0, 25.6, 28.3, 28.5, 31.1, 42.0, 48.42, 104.3, 125.7, 128.5, 139.3, 149.7, 151.1, 153.4, 161.7, 163.1. FT-IR (KBr pellet) ν (cm^{-1}) 3299, 3235, 3015, 2956, 2922, 2852, 2813, 1731, 1693, 1637, 1608. HRMS (MALDI+): *m/z* [$M+Na^+$] calcd for $C_{34}H_{48}N_6NaO_6^+$: 659.3533, found 659.3961. Anal. ($C_{34}H_{48}N_6O_6$) C: calcd, 64.13; found, 62.42, H, N.

***N,N'*-naphtalene-2,7-diylbismethylene(1-octyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide) (14).** (*Yield* = 66%) . *M.p.* Dec. T > 280 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) δ (ppm) 0.85 (t, 6H, *J* = 6.9 Hz), 1.17 – 1.30 (m, 20H), 1.55 – 1.64 (m, 4H), 3.81 (t, 4H, *J* = 6.6 Hz), 4.64 (d, 2H, 5.4 Hz), 7.42 (d, 2H, *J* = 8.4 Hz), 7.71 (s, 2H), 7.85 (d, 2H, *J* = 8.4 Hz), 9.22 (t, 2H *J* = 5.4 Hz), 11.85 (s, 2H). ^{13}C-NMR (150 MHz, DMSO-d_6, 90 °C) δ (ppm) 13.1, 21.3, 25.1, 27.77, 27.79, 27.84, 41.9, 48.0, 104.4, 124.9, 125.2, 127.2, 130.9, 132.5, 136.6, 150.3, 157.6, 158.9, 161.4. FT-IR (KBr pellet) ν (cm^{-1}) 3297, 3128, 3010, 2956, 2920, 2851, 2809, 1728, 1696, 1636, 1608. HRMS (MALDI+): *m/z* [$M+Na^+$] calcd for $C_{38}H_{50}N_6NaO_6$ 709.3690, found 709.4341. Anal. ($C_{38}H_{50}N_6O_6 \cdot H_2O$) C, H, N.

***N,N'*-hexane-1,6-diylbis(1-octyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide) (15).** (*Yield* = 58%) . *M.p.* Dec. T > 237 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) ⍰⍰(ppm) 0.84 (t, 6H, *J* = 6.5 Hz), 1.20 – 1.28 (m, 24H), 1.42 – 1.49 (m, 4H), 1.54 – 1.61 (m, 4H), 3.25 (m, 4H, *J* = 5.7 Hz), 3.80 (t, 4H, *J* = 7.1 Hz), 8.45 (s, 2H), 8.74 (t, 2H, *J* = 5.7 Hz), 11.83 (s, 2H). ^{13}C-NMR (75 Mhz, DMSO-d_6, 25 °C) δ (ppm) 13.8, 21.9, 25.5, 25.9, 28.2, 28.4, 28.8, 28.9, 31.0, 38.1, 48.3, 104.4, 149.9, 150.6, 161.4, 163.5. FT-IR (KBr pellet) ν (cm^{-1}) 3298, 3241, 3013, 2956, 2921, 2853, 2813, 1729, 1697, 1635, 1608. HRMS (ESI+): *m/z* [M+Na^+] calccd for $C_{32}H_{52}N_6NaO_6^+$: 639.3846, found 639.3835. Anal. ($C_{32}H_{52}N_6O_6$) C: calcd, 62.31; found, 61.28; H, N.

***N,N'*-eptane-1,7-diylbis(1-octyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide) (16).** (*Yield* = 66%) . *M.p.* 226.4 – 229.2 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) δ (ppm) 0.85 (t, 6H, *J* = 6.8 Hz), 1.21 – 1.29 (m, 26H), 1.42 – 1.48 (m, 4H), 1.54 – 1.60 (m, 4H), 3.24 (m, 4H, *J* = 6.4 Hz), 3.79 (t, 4H, *J* = 7.1 Hz), 8.45 (s, 2H), 8.73 (t, 2H, *J* = 6.4 Hz), 11.82 (s, 2H). ^{13}C-NMR (75 Mhz, DMSO-d_6, 25 °C) δ (ppm) 13.8, 21.9, 25.5, 26.2, 28.2, 28.4, 28.9, 31.0, 38.1, 104.4, 149.9, 150.6, 161.4, 163.5. FT-IR (KBr pellet) ν (cm^{-1}) 3296, 3135, 3011, 2956, 2921, 2853, 2814, 1727, 1697, 1636, 1608. HRMS (ESI+): *m/z* [M+Na^+] calccd for $C_{33}H_{54}N_6NaO_6^+$: 653.4003, found 653.4008. Anal. ($C_{33}H_{54}N_6O_6$) calcd, 62.83; found, 61.54; C, H, N.

***N,N'*-octane-1,8-diylbis(1-octyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide) (17).** (*Yield* = 54%) . *M.p.* 225.2 – 227.8 °C. ^{1}H-NMR (300 Mhz, DMSO-d_6, 25 °C) δ (ppm) 0.84 (t, 6H, *J* = 6.8 Hz), 1.21 – 1.28 (m, 28H), 1.42 – 1.50 (m, 4H), 1.54 – 1.60 (m, 4H), 3.23 (m, 4H, *J* = 6.1 Hz), 3.80 (t, 4H, *J* = 6.8 Hz), 8.45 (s, 2H), 8.73 (t, 2H, *J* = 6.4 Hz), 11.82 (s, 2H). ^{13}C-NMR (75 Mhz, DMSO-d_6, 25 °C) δ (ppm) 13.9, 22.0, 25.6, 26.3, 28.3, 28.5, 28.7, 29.0, 29.1, 31.1, 38.2, 48.4, 104.5, 150.0, 150.7, 161.5, 163.6. FT-IR (KBr pellet) ν (cm^{-1}) 3299, 3132, 3013, 2956, 2919, 2851, 2813, 1731, 1696, 1635, 1608. HRMS (ESI+): *m/z* [M+Na^+] calccd for $C_{34}H_{56}N_6NaO_6^+$: 667.4159, found 667.4171. Anal. ($C_{32}H_{52}N_6O_6$) C: calcd, 63.33; found, 62.11, H, N.

Biological activity. The human chronic myelogenous leukemia K562 (9), rhabdomyosarcoma RD,[33] breast cancer MCF7,[34] and MDA-MB-231,[35] cystic fibrosis bronchial IB3-1[36,37] were maintained in a humidified atmosphere of 5% CO_2/air in RPMI 1640

medium (Sigma, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (FBS; Celbio, MI, Italy), 50 Units/mL penicillin and 50 μg/mL streptomycin.[19] In order to determine the ability of the tested compounds to inhibit cell growth and to induce erythroid differentiation, K562 cells (30,000 cells/mL) were cultured in the absence or in the presence of the indicated concentrations of compounds and the cell number/mL determined with a ZF Coulter Counter (Counter Electronics, Hialeah, FL, USA) at different days from the culture set-up. In order to verify possible effects on erythroid differentiation, the proportion of benzidine-positive K562 cells was determined and compared to the values obtained employing other known inducers of erythroid differentiation, including cytosine arabinoside (Ara-C),[19] mithramycin,[21] rapamycin [22] and butyric acid.[19]

3-Benzoyl-1-pentyl-1H-pyrimidine-2,4-dione (18). N(3)-benzoiluracil (2,5 g, 11.6 mmol) and K_2CO_3 (2.13g, 11.6 mmol) were dispersed in 40mL of MeCN and stirred for 5 min, then Iodopentane (2.1 mL, 15.1 mmol) was added and the mixture stirred for 2 h. The reaction mixture was dried and the crude solid partitioned between AcOEt and sat. NaCl, and washed with aqueous sodium thiosulphate The organic layer was anhydrified with sodium sulfate and dried to afford 3.3 g (Yield: up to 99%) of pure product.

(18) ^{1}H-NMR (300 Mhz, $CDCl_3$, 25 °C) δ (ppm): 0.9 (t, 3H, *J*=6.8Hz), 1.23-1.39 (m, 4H), 1.69 (q, 2H, *J*=7 Hz), 3.74 (t, 2H, *J*=7.4Hz), 5.79 (d, 1H, *J*=7.9Hz), 7.25 (d, 1H, *J*=7.9Hz), 7.49 (dt, *J*=7.8Hz, 1.6Hz), 7.64 (dt, 1H, *J*=7.4Hz, 1.1), 7.92 (dt, 2H, *J*=8.2Hz, 1.6). ^{13}C-NMR (75 Mhz, $CDCl_3$, 25 °C) δ (ppm): 13.7, 21.9, 28.2, 28.3, 48.8, 101.3, 129.0, 130.1, 131.4, 134.9, 144.9, 149.6, 162.4, 169.1. FT-IR (NaCl disc) ν (cm^{-1}): 2957-2861, 1748, 1705, 1663.

1-Octyl-2,4-dioxo-1,2,3,4-tetrahydro-pyrimidine-5-carbaldehyde (19). A solution of **4** (3 g, 12.6 mmoles) and 2,6-lutidine (5.4 mL) in acetonitrile (50 mL) was added to a stirring solution of $K_2S_2O_8$ (6.8 g, 25.2 mmoles) and $CuSO_4$ (0.8 g, 5 mmoles) in water (50mL). The reaction was heated at 80°C for 2 hours, then the mixture was dried and the residue was partitioned between ethyl acetate and saturated aqueous solution of EDTA. Organic layer was dried and crude product purified by flash chromatography on silica gel (AcOEt/Hexane 4/6) to afford 1.02 g (Yield:33%) of pure compound.

(19) ^{1}H-NMR (300 Mhz, $CDCl_3$, 25 °C) δ (ppm): 0.87 (t, 3H, *J*= 6.9Hz), 1.26-1.32 (m, 10H), 1.73 (q, 2H, *J*= 7.4Hz), 3.84 (t, 2H, *J*= 7.4Hz), 8.09 (s, 1H), 9.32 (s, 1H), 10.01 (s, 1H). ^{13}C-NMR (75 Mhz, $CDCl_3$, 25 °C) δ (ppm):14.0, 22.5, 26.3, 28.9, 29.0, 31.6, 50.1, 110.9, 149.2, 162.1, 186.1. FT-IR (NaCl disc) ν (cm^{-1}): 2953-2850, 1731, 1677.

3-Benzoyl-2,4-dioxo-1-pentyl-1,2,3,4-tetrahydro-pyrimidine-5-carbaldehyde (20). A round bottomed flask containing 10 mL of $POCl_3$ were added dropwise at 0°C 5 mL of DMF followed by addition of compound **18** (5 g, 17.4 mmol). Reaction mixture was kept under nitrogen and stirred at 80°C for 24h. Solvent was removed from reaction mixture, and crude partitioned between AcOEt and sat. $NaHCO_3$. Organic layer was dried and solid purified by flash column chromatography (AcoEt/Hexane 4/6) to afford 2.4 g (Yield: 44%) of pure product.

(20)^{1}H-NMR (300 Mhz, $CDCl_3$, 25 °C) δ (ppm): 0.89 (t, 3H, *J*=7.0Hz), 1.18-1.42 (m, 4H), 1.75 (quint., 2H, *J*=7.4Hz), 3.84 (t, 2H, *J*=7.4Hz), 7.51 (t, 2H, *J*=7.3Hz), 7.67 (t, 1H, *J*=7.3Hz), 7.92 (d, 2H, *J*=7.4Hz), 8.17 (s, 1H), 9.96 (s, 1H). ^{13}C-NMR (75 Mhz, $CDCl_3$, 25 °C) δ (ppm): 13.7, 22.0, 28.3, , 28.7, 50.4, 110.7, 129.3, 130.4, 134.2, 135.4, 148.8, 161.2, 167.7, 185.6. FT-IR (NaCl disc) ν (cm^{-1}): 2958-2931, 1754, 1670.

NOTE: Part of this work has been published on this article: Accetta A., Corradini R., Sforza S., Tedeschi T., Brognara E., Borgatti M., Gambari R., Marchelli R., *J.Med. Chem.*, **2009**, *52*(*1*), 87-94). Some Pictures were reproduced from this source and are copyrighted from ACS.

2.6 References

[1] Lagoja I.M., *Chem. Biodivers.* **2005**, *2*, 1-50;

[2] Jiang Y.L.; Krosky D.J.; Seiple L.; Stivers, J.T., *J. Am. Chem. Soc.* **2005**, *127*, 17412-17420;

[3] Gazivoda, T.; Raic´-Malic´, S.; Marjanovic´, M.; Kralj, M.; Pavelic´, K.; Balzarini, J.; De Clercqc E.; Mintas, M., *Bioorg. Med. Chem.* **2007**, *15,* 749–758;

[4] Benner, S.A., *Acc. Chem. Res.* **2004**, *37*, 784-797;

[5] Krueger, A.T.; Lu, H., Lee, A.H.F.; Kool, E.T., *Acc. Chem. Res.* **2007,** *40*, 141-150;

[6] Nielsen PE.**,** *Pharmacol. Toxicol,.* **2000,** *86*, 3-7;

[7] Wojciechowski, F.; Hudson, R.H.E., *Current Topics in Medicinal Chemistry* **2007**, *7*, 667-679;

[8] Flanagan W. M.; Wolf J. J.; Olson P., Grant D., Lin K. Y.; Wagner R.W.; Matteucci M. D., *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3513–3518;

[9] Lozzio, C.B.; Lozzio B.B., *Blood* **1975,** *45*, 321-34;

[10] Rutherford T.; Clegg, J.B.; Higgs, D.R.; Jones, R.W.; Thompson, J.; Weatherall, D.J., *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 348-52;

[11] Rutherford, T.R.; Clegg, J.B.; Weatherall, D.J., *Nature* **1979**, *280*,164-165;

[12] Lampronti, I.; Bianchi, N.; Zuccato, C.; Medici, A.; Bergamini, P.; Gambari, R., *Bioorg. Med. Chem.* **2006**, *14,* 5204-10;

[13] Lampronti, I.; Martello, D.; Bianchi, N.; Borgatti, M.; Lambertini, E.; Piva, R.; Jabbar, S.; Choudhuri, M.S.; Khan, M.T.; Gambari, R., *Phytomedicine* **2003**,*10*, 1300-1308;

[14] Osti, F.; Corradini, F.G.; Hanau, S.; Matteuzzi, M.; Gambari, R., *Haematologica* **1997**, *82*, 395-401

[15] Olivieri, N. F., *Semin Hematol* **1996**, *33*, 24-42

[16] Bianchi, N.; Chiarabelli, C.; Borgatti, M.; Mischiati, C.; Fibach, E.; Gambari, R., *Br J Haematol* **2001**, *113*, 951-61

[17] Bianchi, N.; Onagro, F.; Chiarabelli, C.; Gualandi, L.; Mischiati, C.; Bergamini, P.; Gambari, R., *Biochem Pharmacol* **2000**, *60*, 31-40;

[18] Bianchi, N.; Osti, F.; Rutigliano, C.; Corradini, F.G.; Borsetti, E.; Tomassetti, M.; Mischiati, C.; Feriotto, G.; Gambari, R., *Br J Haematol* **1999**, *104*, 258-65;

[19] Gambari, R.; Fibach, E. , *Curr Med Chem* **2007**, *14*, 199-212;

[20] Chiarabelli, C.; Bianchi, N.; Borgatti, M.; Prus, E.; Fibach, E.; Gambari, R. , *Haematologica* **2003**, *88*, 826-827;

[21] Fibach, E.; Bianchi, N.; Borgatti, M.; Prus, E.; Gambari, R., *Blood* **2003**, *102*, 1276-1281;

[22] Fibach, E.; Bianchi, N.; Borgatti, M.; Zuccato, C.; Finotti, A.; Lampronti, I.; Prus, E.; Mischiati, C.; Gambari, R., *Eur J Haematol.* **2006**, *77*, 437-41;

[23] Zuccato, C.; Bianchi, N.; Borgatti, M.; Lampronti, I.; Massei, F.; Favre, C.; Gambari, R. , *Acta Haematol.* **2007**, *117*,168-176;

[24] Lampronti, I.; Bianchi, N.; Borgatti, M.; Fibach, E.; Prus, E.; Gambari, R. Accumulation of gamma-globin mRNA in human erythroid cells treated with angelicin. *Eur J Haematol.* **2003**, *71*, 189-195;

[25] Bianchi, N.; Zuccato, C.; Lampronti, I.; Borgatti, M.; Gambari, R., *Evidence-based Complementary and Alternative Medicine (eCAM)*, doi:10.1093/ecam/nem139;

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

[27] Sforza, S.; Tedeschi, T.; Corradini, R.; Marchelli,R., *Eur. J. Org. Chem.* **2007**, 5879–5885;

[28] Corradini, R.; Sforza, S.; Tedeschi T.; Totsingan F.; Marchelli R., *Curr. Topics Med. Chem.,* **2007**, *7*, 681-694;

[29] Betts, L.; Josey, J.A.; Veal, J.M.; Jordan, S.R., *Science* **1995**, *270*, 1838–1841;

[30] Kokko, J.P.; Mandell, L.; Goldstein J., *J. Am. Chem. Soc.* **1962**, *84*, 1042-1047;

[31] (a) ArgusLab 4.0, Mark A. Thompson, Planaria Software LLC, Seattle, WA; (b) Thompson, M..A.; Zerner M.C., *J. Am. Chem. Soc.*, **1991**, *113*, 8210-8215; (c) Dewar, M.J.S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J.J.P, . *J. Am. Chem. Soc.*, **1985**, *107*, 3902-3909;

[32] Theodorou, V.; Ragoussis, V.; Strongilos, A. ; Zelepos, E. Eleftheriou, A.; Dimitriou, M.,*Tetrahedron Lett.* **2005**, *46,* 1357–1360;

[33] A. Feriotto G, Finotti A, Breveglieri G, Treves S, Zorzato F, Gambari R: , *FEBS J*, **2007**, *274*: 4476-4490;

[34] B. Lambertini E, Piva R, Khan MT, Lampronti I, Bianchi N, Borgatti M, et al., *Int J Oncol.*, **2004**; *24*, 419-23;

[35] C. Penolazzi L, Zennaro M, Lambertini E, Tavanti E, Torreggiani E, Gambari R, Piva R, *Mol Pharmacol.*, **2007**, *71(6),* 1457-62;

[36] D. Borgatti M, Bezzerri V, Mancini I, Nicolis E, Dechecchi MC, Lampronti I, Rizzotti P, *Biochem Biophys Res Commun*, **2007**, *357*, 977-83;

[37] E. Borgatti, M.; Bezzerri, V.; Mancini, I.; Nicolis, E.; Dechecchi, MC.; Rizzotti, P.; Lampronti, I.; Rizzotti, P.; Gambari, R.; Cabrini, G., *Minerva Biotecnologica*, **2008**, *20*, 79-84;

[38] Vilsmeier, A.; Haack, A., *Ber.* **1927**, *60*, 119;

[39] Agryropoulou, E.; Zachariadou, C., *J. Heterocyclic Chem.* **2005**, *42*, 1135-1142.

PAGE LEFT INTENTIONALLY BLANK

CHAPTER 3

Design and Synthesis of "Tailor Made" Modified Uracil derivatives for Selective Adenine Recognition

3.1 Introduction

Synthetic oligonucleotides and analogues[1] (among which peptide nucleic acids[2] is one of the most important classes) are widely used in research as tools for biochemistry, molecular biology, gene therapy and diagnostics. Modifications of the natural nucleobases represent an excellent way to enhance recognition properties or to confer new functions to these DNA mimics. In literature many examples of modified nucleobase derivatives with special properties are reported. For example the replacement of natural bases with universal base analogues like nitroazole and deoxyinosine led to obtain probes that bind indifferently to the four natural bases with little discrimination among them[3]. Kool *et al*[4] reported benzohomologation of natural nucleotides to create a new size expanded DNA (xDNA, yDNA and yyDNA). Merging the four natural with the new expanded bases allowed to create a new genetic set of unnatural eight bases[5]. Matteucci and co-workers[6] reported tricyclic analogues of cytosine (based on the phenoxazine ring), termed G-clamp (*Figure 3.1a*), able to form additional H-bonds on Hoogsteen sites of complementary guanine and to stack more efficiently with neighboring bases due to its enhanced aromatic surface. Replacement of cytosine by G-clamp was found to stabilize nucleic acid duplexes (up to +18°C of melting temperature) and to confer more selectivity and potency to antisense oligonucleotides.

Despite the remarkable hybridization properties of PNA (widely described in the introduction chapter), numerous chemical modification of the aminoethyl glycine backbone[7] and nucleobases have been reported. PNA base modifications[8] were inspired from modified nucleotides or often new modifications were proposed. Manoharan *et al*[9] reported the synthesis and properties of PNA oligomers containing G-clamp modifications improving the recognition properties as for oligonucleotides. Moreover, G-clamp modifications inserted in

chiral γ-PNA oligomers were found to promote more efficiently strand invasion in mixed B-DNA sequences.[10] An interesting PNA base modification was presented by Hudson and Wojciechowski[11]; they reported a phenyl pyrrole-cytosine (*Figure 3.1b*) as cytosine substitute that increases the melting temperature of the duplex with complementary DNA up to 10°C and showing fluorescence dependence from the hybridization state. For adenine complexation only one noteworthy example was described by Nielsen and co-workers, who reported a series of substituted 1,8-naphthyridines[12] as thymine analogues for PNA oligomers (*Figure 3.1c*). Though naphthyridines have a larger aromatic core, the insertion of these modifications led to a limited stability increases compared to thymine. Only 2-4 degrees of melting temperature were gained upon introduction of this modification. The quest for modified bases for selective and efficient adenine complexation is no less important than guanine complexation. Often mutations that transform healthy gene in oncogene are due to replacement of GC base pair with AT[13]; for example the expression of mutated K-Ras gene, where codon 12 GGT is mutated in GAT, leads to a mutated gene which is present in a large percentage of all pancreatic adenocarcinomas[14]. Therefore PNA probes able to discriminate wild-type genes from mutated ones are promising tools for diagnostic purposes or for gene therapy.

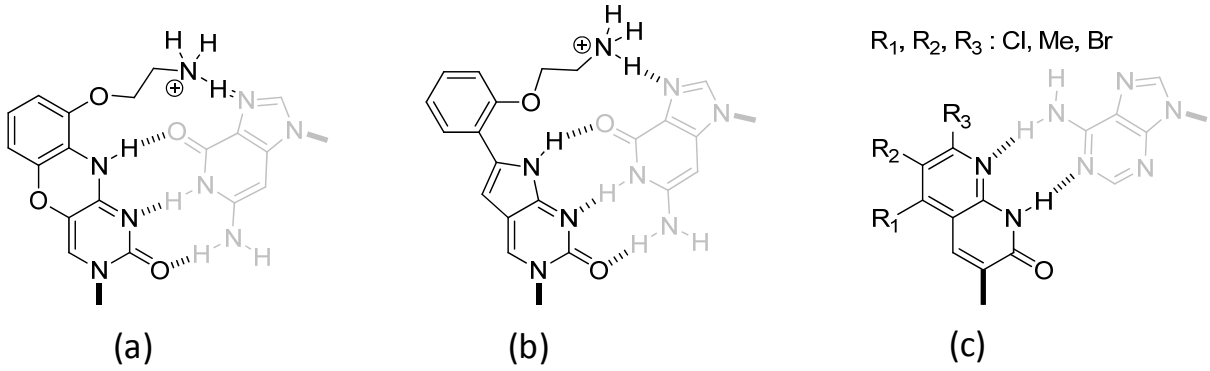


Figure 3.1. (a)G-clamp, (b)Phenylpyrrolecytosine and (c) 1,8 – naphthyridines.

In this chapter we present our contribution to the development of novel selective and potent modified nucleobases for adenine recognition; in particular, herein, the design and the monomer synthesis of some tailor made modified uracil derivatives is showed.

3.2 Adenine Templated Base Design

Nucleic acids structures are stabilized mainly by hydrogen bonding and base stacking roughly by the same contributions, which were estimated for the DNA terminal base pair to be ~ 1 Kcal/mol[15]. These terms could be different for duplex involving PNAs; however, design principles used to improve nucleotides binding abilities were shown also suitable to improve PNA binding properties. The strength of the stacking contributions can be enhanced by extending the aromatic surface of bases or increasing the base polarizability by insertion of opportune aromatic moieties, without disrupting the ability of the base to form hydrogen bonding[16]. Electrostatic effects can be exploited to gain stabilization by insertion of positive charged moieties able to interact with the negative charges of DNA. However, it is difficult to increase the hydrogen bonding contributions strength without affecting the polarizability of the base (linked to the stacking energy), but it is possible to increase their number by adding moieties able to interact with free Hoogsteen sites on purine.

PNA in certain circumstances can form a 2:1 complex that involves the formation of a PNA:DNA:PNA triple helix; a homo-purine strand of DNA can be targeted by two homo-pyrimidine PNAs that bind through both Watson and Crick and Hoogsteen sites of DNA (*Figure 3.2a*). From solid state structure[17] it is evident that PNA_2DNA triplexes are stabilized by hydrogen bonding and by Van der Waals contacts of the pyrimidine with neighbouring purine on DNA (*Figure 3.2b*). Looking inside the TAT triplet of PNA_2DNA, we found that the distance between the methyl groups of thymine was roughly the same along the triplex (*Figure 3.2c*). On the base of this observation, we designed a novel modified base sketch by linking two uracil moieties in order to be able to mimic a TAT triplette; the result is a modified “uracil dimer” that binds adenine on both Watson and Crick and Hoogsteen sites, with the additional pendant uracil able to establish stacking interactions with the neighbouring bases on the complementary DNA strand (*Figure 3.3a*). In order to replace the geometry and the distance of TAT triplets, we chose an amide group on the C(5) position and 2,7-naphthalene diamine as linker. As showed in the previous chapter, the amide group was found not to shift the tautomerism of uracil, and the geometry of the linking group is preorganized by H bonding between the amide NH and the carbonyl oxygen at uracil C(4)[18].

The naphthalene linker was chosen as a rigid spacer in order to avoid collapse of the uracil moieties; the steric hindrance of this group should not be a severe drawback, based on models, since it would be located in the major groove of the PNA:DNA duplex (*Figure 3.3a*).

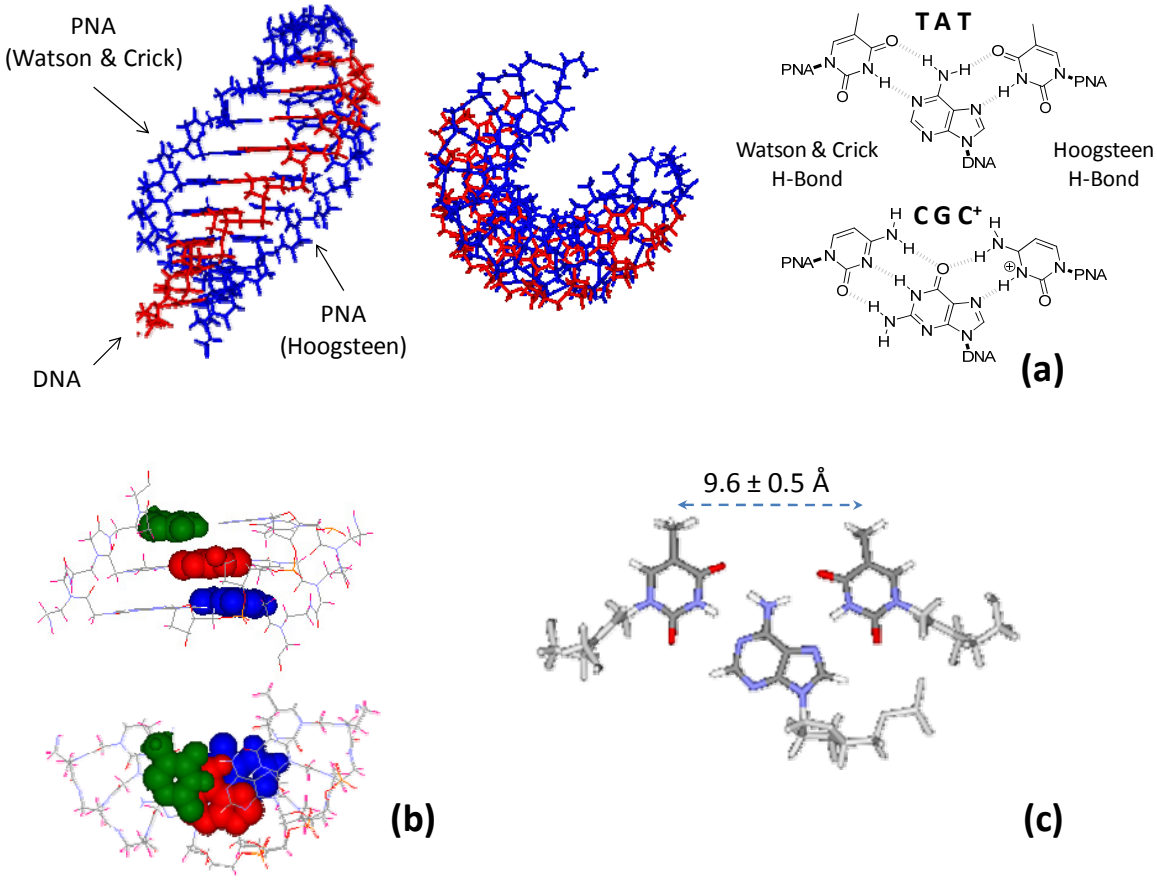


Figure 3.2. (a) PNA_2DNA triple helix; (b) Van der Walls contact between purine on DNA (red) and neighboring pyrimidine on Watson and Crick strand (green) and Hoogsteen (blue); (c) T A T triplette and distance between methyl groups of thymine along the entire structure (its averaged value is 9.6 ± 0.5 Å).

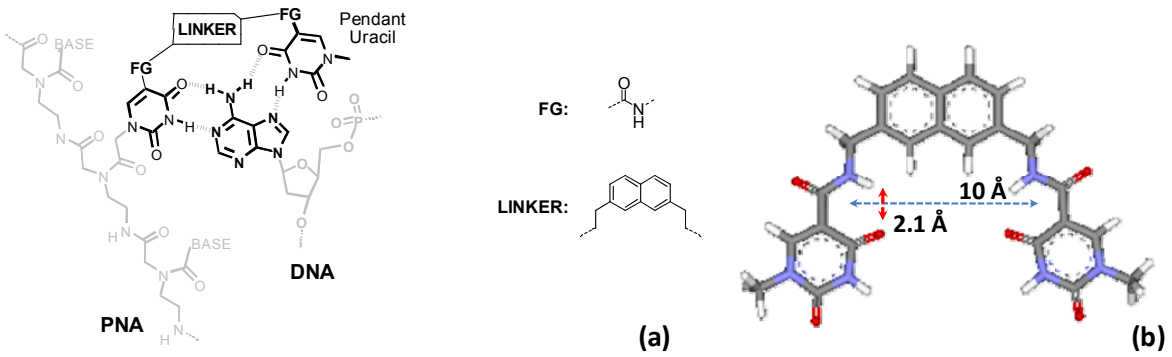


Figure 3.3. (a) "Tailor made " modified uracil for adenine recognition ; (b) Optimized model of modified nucleobases mimicking the geometry of TAT triplette.

3.3 Synthesis of 5-carboxamidomethyl-naphthalene derivatives

Based on the design depicted above, we planned the synthesis of three PNA monomers: Napht-2, the modified uracil dimer, Napht-1, bearing only the naphthalene linker, as reference, and Napht-3,bearing the naphthl linker with an amino group, as showed in *Figure 3.4*.

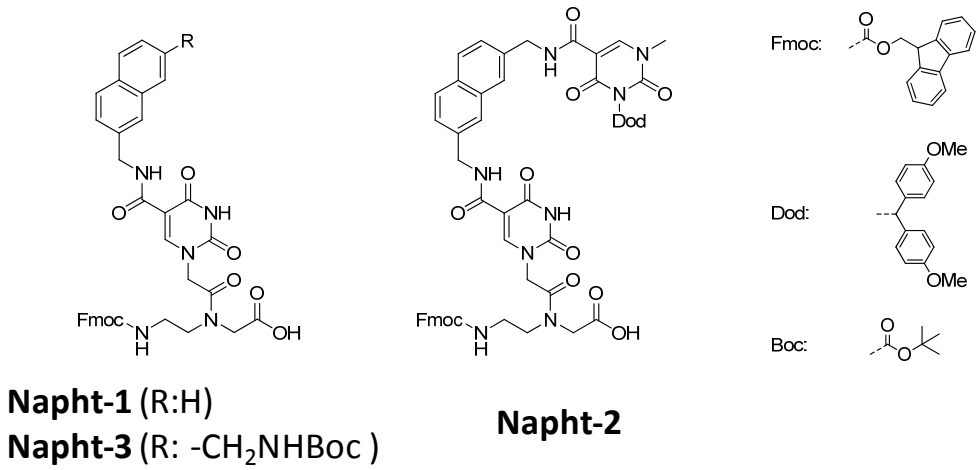


Figure 3.4. Napth-1, Napht-2 and Napht-3 modified Fmoc PNA monomers.

We chose to synthesize Fmoc monomers since they are compatible with both Fmoc/Bhoc and Boc/Cbz PNA oligomeric synthesis and because the Fmoc group (Fluorenylmethyloxycarbonyl) is more lipophilic than Boc (t-Butyloxycarbonyl), ensuring a better solubility of the monomer derivatives in organic solvents. Solubility in organic media due to self-association was a very critical issue for dimeric uracil derivatives[19]. Therefore, in order to limit self aggregation of dimers and enhance their solubility in organic solvents, we inserted a protective group on N(3) nitrogen of uracils, thus removing their ability to form hydrogen bonding. In order to have a protective group cleavable in acidic media and resistant under basic conditions (carbamate-based protective groups on N(3) of uracil are labile in alkaline media) 4,4'-dimethoxybenzydril (Dod)[20] was inserted on the N(3) nitrogen of pendant uracil (*Figure 3.4*). Synthons needed for the synthesis of Napht-2 are depicted in *Figure 3.5.*

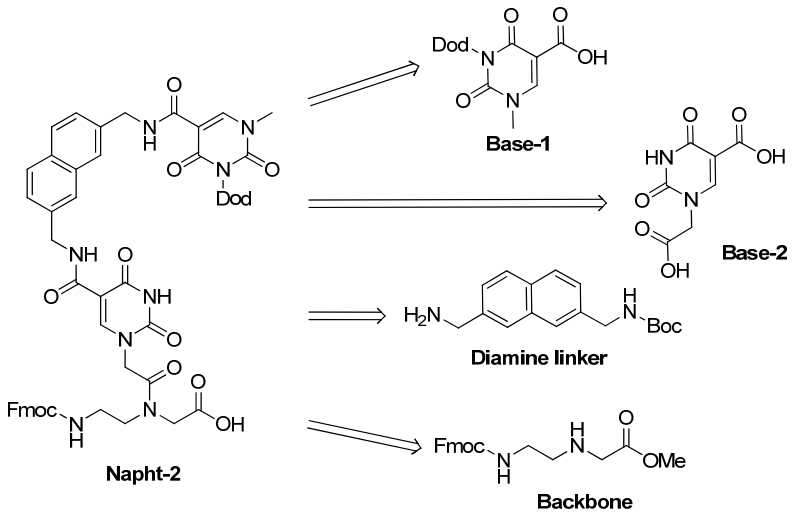
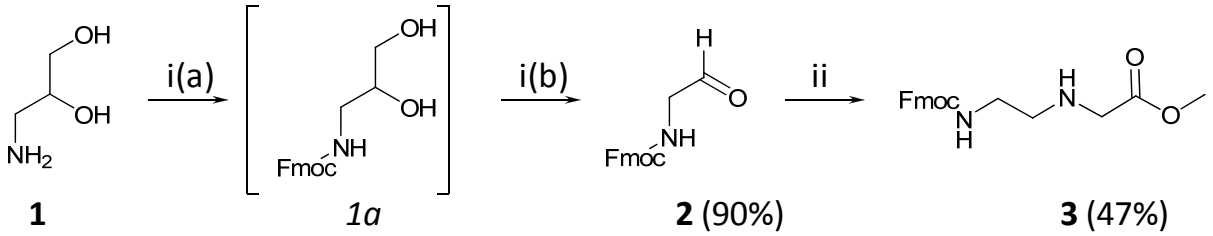


Figure 3.5. Retrosynthesis of the Napht-2 derivative. Some syntons are useful also in Napth-1 and Napth-3 synthesis.

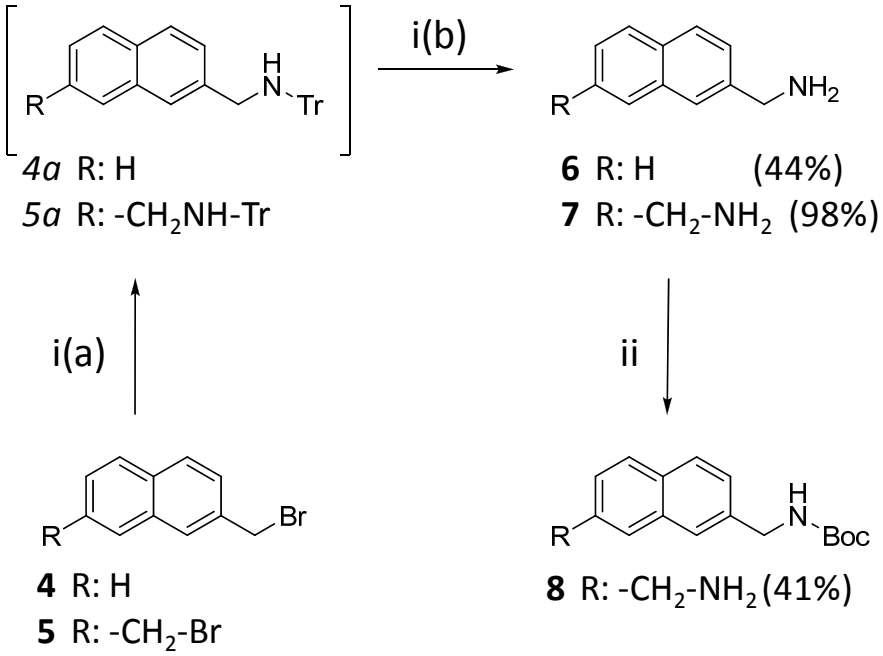
The synthesis of Fmoc-aminoethylglycine backbone **3** was carried out by reductive amination of glycine methyl ester with fmoc-aminoacetaldehyde **2** (*Scheme 3.1*) which was prepared according to a literature[21] procedure by oxidation of Fmoc protected aminopropanediol.



1 *1a* **2** (90%) **3** (47%)

Scheme 3.1. Synthesis of the Fmoc-aeg-OMe backbone. (i) [(a) Fmoc-NOSu, Na_2CO_3, dioxane, water, 75 min, rt; (b)KIO_4, acetone, water, 12h, rt]. (ii) [GlyOMe, $NaBH_3CN$, MeOH, 6h, 0°C to rt].

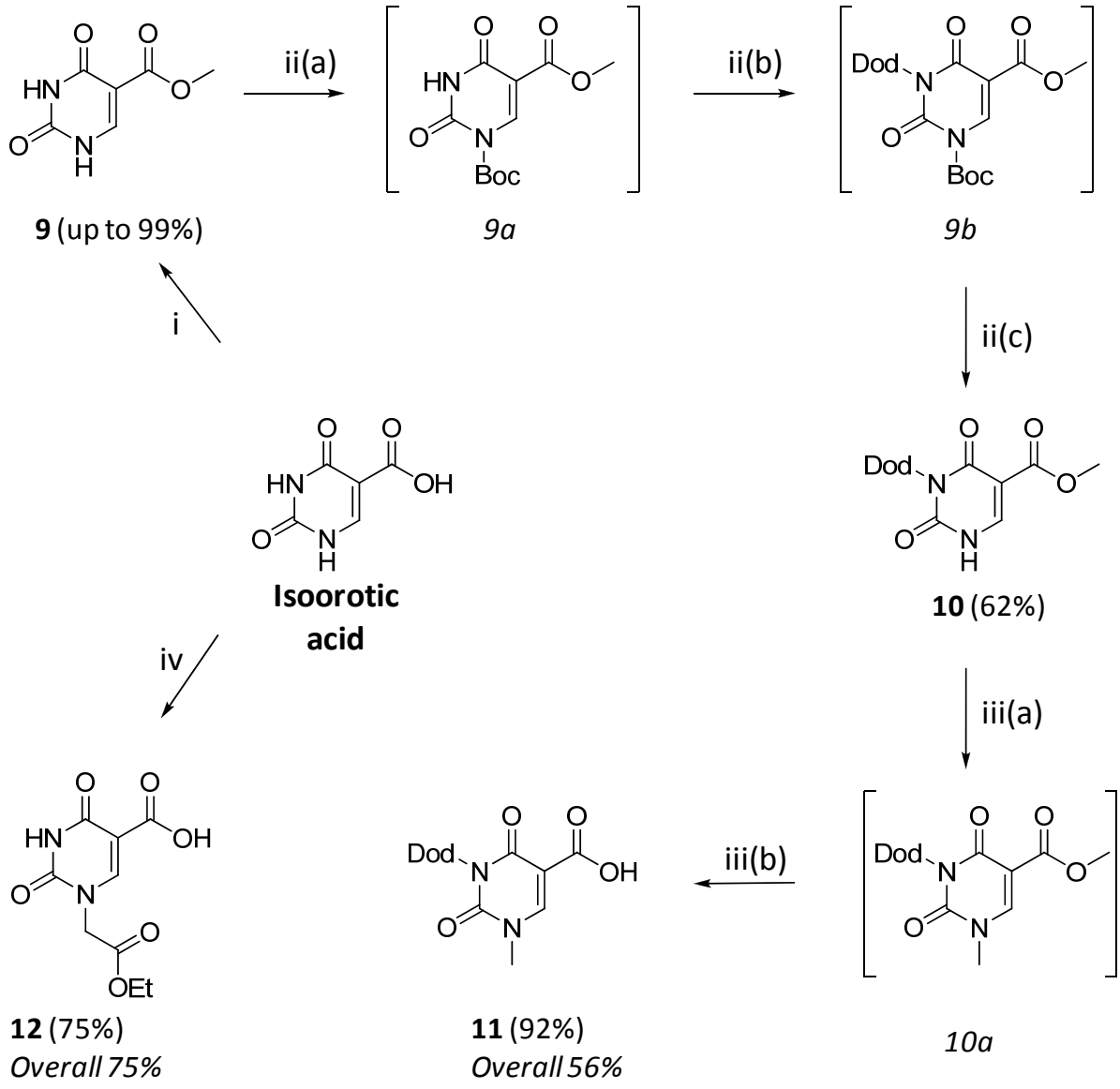
The synthesis of the diamine **7** and the monoamine **6** was carried out using alkyl bromide derivatives by nucleophilic substitution with protected ammonia in the form of tritylamine[22]; then N-trityl group was removed from intermediates *4a* and *5a* by acidic cleavage. The synthesis of **7** was previously reported in *Chapter 2*. In that procedure at least 2 equivalents tritylamine were used for each of bromide, one as source of ammonia and the other one as a base to scavenge the HBr. Synthesis presented in *Scheme 3.2* uses tritylamine only as a source of ammonia while a basic environment is ensured by potassium carbonate; in this way we were able to decrease the consumption of the expensive tritylamine. Compound **6** was obtained with a yield about 50% of that observed for **7**, due to a not optimized workup, though the reactivity of bromide **4** and **5** was found to be comparable.



Scheme 3.2. Synthesis of the monoprotected diamine linker and of aminomethylnaphtalene (Tr: trityl). (i) [(a)Tr-NH_2, K_2CO_3, MeCN, 96h, 55°C; (b) TFA, DCM, MeOH, 40 min, rt]. (ii) [BocOPh, EtOH, overnight, reflux].

The monoprotected diamine **8** was synthesized by reaction of **7** with the very mild Boc donor *tert*-butyl phenyl carbonate (BocOPh)[23]. Due to its low reactivity, BocOPh was used in a 1/1 ratio with the diamine, leading to a low quantity of the doubly Boc-protected side product and a better yield compared to that obtained with the more reactive Boc anhydride.

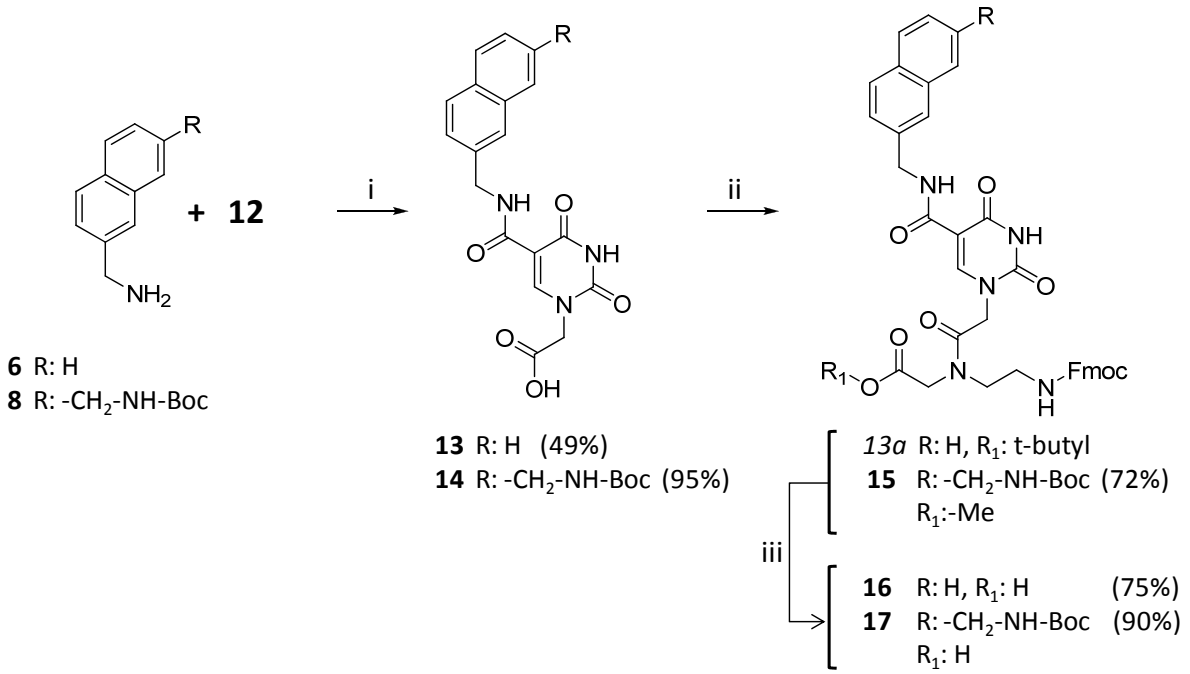
The synthesis of the two synthons Base-1 and Base-2 is depicted on *Scheme 3.3*. Compound **12** was obtained with a good yield from one pot reaction that involves alkylation of the fully silylated isoorotic acid; the regioselectivity of this reaction is discussed in details in the previous chapter. Synthesis of **11** was more complex than that of **12,** and was carried out using the same starting material. Fisher's esterification of isoorotic acid in refluxing methanolic HCl gave **9** quantitatively.



Scheme 3.3. Synthesis of Base-1 and Base-2 synthons. (i) [$SOCl_2$, MeOH, overnight]. (ii) [(a) Boc_2O, DMAP, MeCN, 5h, rt; (b) Dod-Cl, NaH, DMF, overnight, 0°C to rt; (c) K_2CO_3, MeOH, 2h, rt]. (iii) [(a)MeI, K_2CO_3, MeCN, 3h, rt; (b) 1M NaOH, MeOH, 2h, rt]. (iv) [(a)HMDS, TMS-Cl, 4h, reflux; (b) Br-CH_2COOEt, 18h, reflux; (c) water, AcOH, 20 min, rt].

The direct alkylation of nitrogen N(3) of compound **9** with Dod-Cl (4,4'-dimethoxybenzhydryl chloride) was not possible, since the reactivity of hindered alkylating agents is directed preferentially to the N(1) position. This problem was readily circumvented by temporary protection[24] of **9** on N(1) with the Boc group, which was then removed after alkylation of N(3) with Dod-Cl. Enhanced reactivity N(1)-Boc on the intermediate *9b*, due to the electron withdrawing effect of the uracil ring, allowed its removal under alkaline conditions to afford **10** with a fairly good yield.

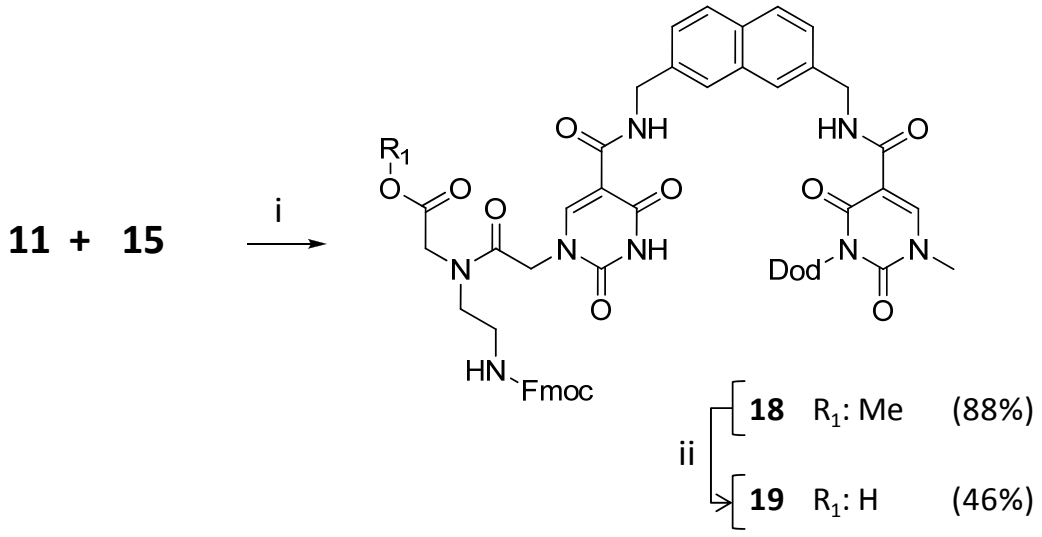
Derivative **11** was obtained from **10** after methylation and hydrolysis of the methyl ester moiety with a good yield (92%). From isoorotic acid to derivative **11** the overall yield was 52%; thus this method seems to be suitable for the selective synthesis of N(1) alkylated derivatives of isoorotic acid.



Scheme 3.4. Synthesis of Napht-1 (**16**) and Napht-3 (**17**). (i) [(a) HBTU, DIEA, DMF, 4h, rt; (b) 2M NaOH, THF, 90 min, rt]. (ii) [Fmoc-aeg-OR, DCC, DhBtOH, DIEA, DMF, overnight, 0°C to rt]. (iii) [(a) TFA, DCM *for **16***; (b) $Ba(OH)_2$, THF, 1.5h, rt *for **17***].

Amines **6** and **8** were then coupled with HBTU activated derivatives of **12** to give the corresponding amides, which, after hydrolysis of the ethyl ester, allowed to obtain **13** and **14**. The yield of **13** was lower than **14** due to a not fully optimized workup (as in the case of **6** and **8**). Derivative **14** was coupled with Fmoc-aminoethylglycine methyl ester **3** using DCC/DhBtOH condensing agents to afford **15** with a good yield. Derivative **15** was used for the synthesis of Napht-3 (**17**) and further elaborated to Napht-2. The hydrolysis of methyl ester of **15** to afford **17** was carried out using $Ba(OH)_2$ in THF[25], carefully controlling the reaction time in order to avoid undesired Fmoc deprotection.

For the synthesis of Napht-1 we employed a commercially available Fmoc-aminoethylglycine *tert*-butyl ester, which, after coupling with the activated derivative **13** and acidic cleavage of the *tert*-butyl ester, gave **16** in good yield.



Scheme 3.5. Synthesis of Napht-2. (i) [(a) TFA, DCM, 1h, 0°C to rt; (b) HBTU, DIEA, DMF, 4h, rt]. (ii) [$Ba(OH)_2$, THF, 1.5h, rt].

After Boc deprotection of the aminomethyl moiety of **15** and coupling with **11** using HBTU, the methyl ester **18** was obtained in high yield. Napht-2(**19**) was obtained from the methyl ester hydrolysis of **18** using the same conditions employed for **17**.

3.4 Synthesis of the 5-azidomethyluracil PNA monomer

To explore the properties of new derivatives bearing a methylamido moiety instead of a carboxamidomethyl (*Figure 3.6a*) we designed a synthetic pathway for the preparation of these oligomers bearing a "reversed amide" that envisages the 5-azidomethyluracil PNA monomer as key intermediate (*Figure 3.6b*). Details of the solid phase oligomeric synthesis are presented in the next chapter; herein only the synthesis of 5-azidomethyl derivatives will be discussed.

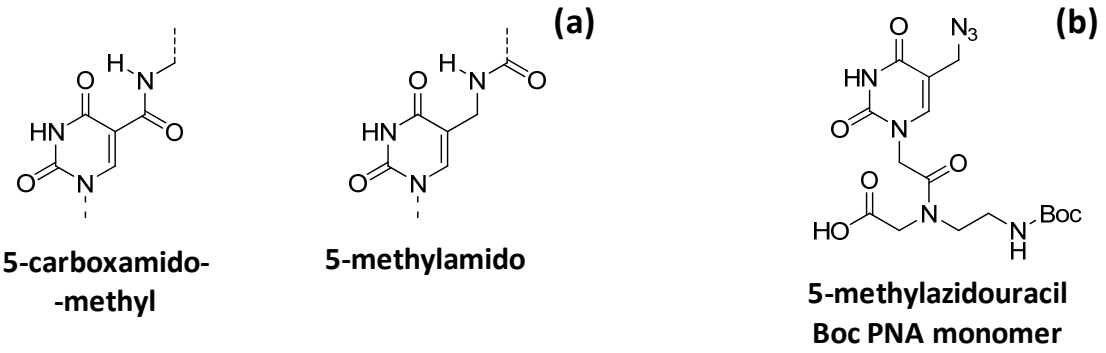
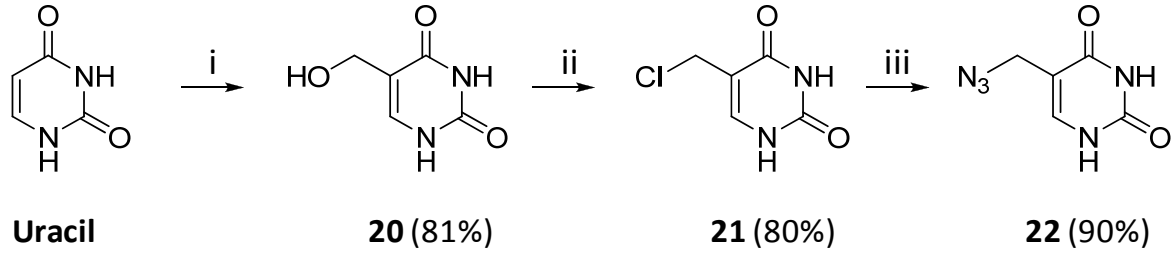


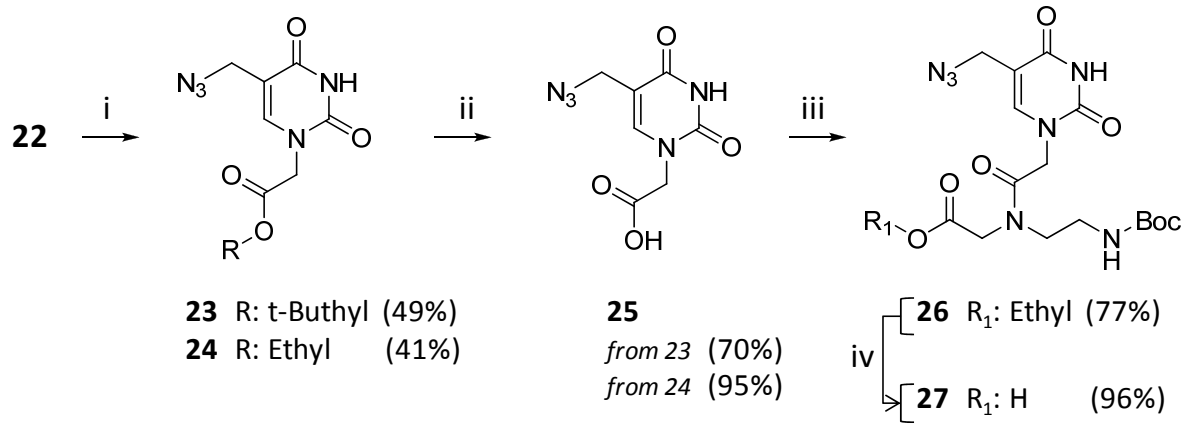
Figure 3.6. (a) 5-carboxamidomethyl and 5-methylamido models; (b) Key intermediate 5-methylazidouracil Boc PNA monomer.

The synthesis of 5-methylazidouracil PNA monomer is described below in two phases: synthesis of 5-methylazidouracil (*Scheme 3.6*) and its incorporation into Boc-aminoethylglycine backbone (*Scheme 3.7*).



Scheme 3.6. Synthesis of 5-azidomethyluracil. (i) [CH_2O, Et_3N, water, 60°C, overnigth]. (ii) [HCl 37%, 4H, rt]. (iii) [NaN_3, DMF, 0°C, 1h].

The synthesis of 5-methylazidouracil started from the regioselective hydroxymethylation of C(5) of uracil to afford 5-hydroxymethyluracil[26] **20**, that was transformed to the alkyl chloride **21** using concentrated HCl. Nucleophilic substitution of **21** with sodium azide gives the desired product **22**. This last reaction was carried out at low temperature (0°C) in order to avoid self reaction of **21** on its N(1) or N(3), which would form polymeric compounds, thus lowering the yield.



Scheme 3.7. Synthesis of 5-methylazidouracil Boc-PNA monomer. (i) [$BrCH_2COOR$, K_2CO_3, DMF, rt, overnight]. (ii) [(a) TFA/DCM; (b) NaOH, water/MeOH *for**24***]. (iii) [DCC, DhBtOH, DIEA, DMF]. (iv) [NaOH, water/MeOH].

5-methyluracil was then regioselectively alkylated on N(1) by bromoacetic acid esters to give **23** and **24**, leading to only 49% yield due to formation of the bis-alkylated side product. Ester deprotection of **23** and **24** using appropriate conditions (basic hydrolysis for **24** and acid cleavage for **23**) leads to compound **25**. 5-methylazidouracil acetic acid **25** was activated with DCC/DhBtOH and linked to the Boc-aeg-OEt backbone to afford the ethyl ester PNA monomer **26**. The hydrolysis of ethyl ester of **26** afforded the desired PNA monomer ready to use for the solid phase synthesis of oligomers.

3.5 Conclusion

In this chapter the synthesis of a new series of PNA derivatives modified on the C(5) position of uracil was discussed. While modifications at position C(5) of uridine in oligonucleotides are well known, for modifications on C(5) positions of uracil there are only few reports. Synthetic routes presented here expand the possibility to use 5-carboxamidomethyl and 5-methylamido moieties as linker between uracil C(5) positions and a large variety of groups, fluorophores or reactive moieties. The synthesis of oligomers containing these modified monomers and their properties will be discussed in the next chapter.

3.6 Experimental Section

General. Reagents were purchased from Sigma-Aldrich, Fluka, Merck, Carlo Erba, TCI Europe and used without further purifications. TLC were run on Merck 5554 silica 60 aluminium sheets. Column chromatography were performed as flash chromatography on Merck 9885 silica 60 (0.040 – 0.063) under 0.1-0.2 bar of air/nitrogen pressure. NMR spectra were obtained on a Bruker 300 MHz AC or Advance spectrometers. δ values are expressed in ppm. FT-IR were recorder on Thermo Nicolet 5700 using KBr pellets. ESI-MS spectra were recorded on a Waters Acquity instrument.

Fmoc-aminoacetaldheyde (2). Aminopropanediol (1.43g, 15.7 mmol) was dissolved in 96 mL of 4% acqueous solution of Na_2CO_3 and stirred at 0°C for 5min. followed by the addition of Fmoc-N-hydroxysuccinimmide (5.73g, 15.8 mmol) in 96 mL of dioxane. The mixture was allowed to cool to room temperature and stirred for further 75min., then neutralized with 300 mL of sat. $KHSO_4$ and extracted with EtOAc (400 mL). The organic layer after washings (with 200 mL sat. Na_2CO_3 and 200 mL sat. NaCl) was dried and the white powder obtained was dissolved in 350 mL of 1/1 water/acetone followed by the addition of KIO_4 (8.0 g, 34.54 mmol) and stirred for 1 hour. The mixture was then filtered on a G5 funnel, distilled to remove acetone and the remaining mixture was extracted with ethyl acetate many times. Organic layer was dried and the residue was purified by flash column chromatography on silca gel using AcOEt as eluent to afford 3.98 g (Yield: 90%) of **1**.

(2) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm):9.65 (s, 1H, CHO), 7.77 (d, 2H, *J*=7.4 Hz, aromatic CH), 7.60 (d, 2H, *J*=7.3 Hz, aromatic CH), 7.41 (dd, 2H, *J*=7.7 Hz,1.3Hz, aromatic CH), 7.32 (dd, 2H, *J*=7.5 Hz,1.3Hz, aromatic CH), 5.47 (s, 1H, NHFmoc), 4.43 (d, 2H, *J*=6.9 Hz, CH_2), 4.23 (t, 1H, *J*=6.9 Hz, CH Fmoc), 4.15 (d, 1H, *J*=5 Hz, CH_2NH_2). Full characterization is reported in ref.21.

Fmoc-aminoethylglycine-OMe (3). Glycine-OMe hydrochloride (1.49g, 11.75 mmol) and **1** (3.0g, 10.68 mmol) were dissolved in 100 mL of MeOH, then $NaBH_3CN$ (1.55g, 23.5 mmol) and AcOH (4.03 mL, 70.5 mmol) were added at 0°C, and the mixture stirred at room temperature for 5 h. Methanol was removed and the residue partitioned between 200 mL of sat. $KHSO_4$ and (extracted four times) 200 mL of AcOEt. The combined organic layer was

dried and residue purified by flash column chromatography on silica gel with AcOEt/MeOH 95/5 as eluent to afford 1.77 g (Yield: 47%) of the product.

(3) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 7.88 (d, 2H, *J*=7.5 Hz, aromatic CH Fmoc), 7.69 (d, 2H, *J*=7.2 Hz, aromatic CH Fmoc), 7.41 (t, 2H, *J*=7.2Hz, aromatic CH Fmoc), 7.33 (t, 2H, *J*=7.5Hz, aromatic CH Fmoc), 7.25 (t, 1H, *J*=5.7Hz, NH-Fmoc), 4.29 (d, 2H, *J*=6.6Hz, CH_2 Fmoc), 4.21 (t, 2H, *J*=6.7Hz, CH Fmoc), 3.61 (s, 3H, -COOMe), 3.36 (s, 2H, -CH_2-COOMe), 3.05 (q, 2H, *J*=6.0Hz, Fmoc-NH-CH_2-CH_2), 2.57 (t, 2H, *J*=6.3Hz, Fmoc-NH-CH_2-CH_2-NH), 2.21 (s (broad), 1H, CH_2-NH-CH_2-). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 172.6, 156.1, 143.8, 140.7, 127.5, 127.0, 125.1, 120.0, 65.2, 51.1, 49.8, 48.2, 46.7, 40.3.

2-aminomethylnapthalene (6). Tritylamine (2,16 g, 8.16 mmol) and K_2CO_3 (0.75 g, 5.44 mmol) were dissolved in 50 mL of dry MeCN, then 1.26 g (5.44 mmol) of 2-bromomethylnapthalene was added. The reaction mixture was stirred at 50°C under nitrogen for 96h then, after removing solvent, the crude material was partitioned between sat. $NaHCO_3$ (50 mL) and AcOEt (50 mL). The organic phase was dried with Na_2SO_4 and evaporated under reduced pressure to give a crude product, which was purified by flash chromatography on silca gel using Hexane/AcOEt 95/5 as eluent (R_f:0.48) to afford 1.46 g (Yield: 71%) of intermediate *4a*. *^{1}H-NMR* (*300 Mhz, 25°C, $CDCl_3$*) δ(ppm): 7.86-7.79 (m, 4H, aromatic CH), 7.62-7.58 (m, 6H, aromatic CH), 7.52-7.44 (m, 3H, aromatic CH), 7.35-7.29 (m, 6H, aromatic CH), 7.25-7.19 (m, 3H, aromatic CH), 3.50 (s, 2H, Ar-CH_2-NH), 1.97 (s (broad), 1H, NH). *^{13}C-NMR* (75 *Mhz, 25°C, $CDCl_3$*) δ(ppm): 146.0, 138.5, 133.5, 132.6, 128.6, 127.92, 127.85, 127.67, 127.62, 126.6, 126.4, 125.95, 125.92, 125.4, 71.0, 48.0. Intermediate *4a* (1.42 g, 3.55 mmol) was dissolved in 18 mL of DCM and 25 of TFA and stirred at room temperature (an intense yellow color appered immediately after TFA addition). After 40 min, 10 mL of MeOH were added and the mixture stirred at room temperature until yellow color disappeared (about 1h). The reaction mixture was evaporated and the crude material dissolved in 100 mL of 1 M HCl and washed two times with DCM, then the aqueous phase was then neutralized with NaOH 2M and extracted with DCM. The organic layer was dried to give 0.35 g of white powder (Yield of trytyl deprotection:62%, overall yield: 44%).

(6) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 7.88-7.81 (m, 4H, aromatic CH), 7.52-7.42 (m, 3H, aromatic CH), 3.88 (s, 2H, Ar-CH_2-NH_2), 1.93 (s (broad), 2H, Ar-CH_2-NH_2). *^{13}C-*

NMR (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 141.9, 133.0, 131.8, 127.40 - 127.39, 126.1, 125.8, 125.1, 124.5, 45.7.

2,6-diaminomethylnapthalene (7). Dibromonaphthalene (1g, 3.12 mmol), trytylamine (1.9g, 7.0 mmol) and K_2CO_3 (0.98g, 7.0 mmol) were dissolved in 45 mL of dry MeCN and stirred at 50°C in a sealed tube for 96h.Reaction mixture was dried and solid partitioned between sat. $NaHCO_3$ and AcOEt. The organic layer was anhydrified over Na_2SO_4 and dried. Crude intermediate *5a* was dissolved in 5 mL of TFA and 3,5 mL of DCM and stirred at room temp. for 40 min, during which the mixture turn to orange-yellow; 3,5 mL of MeOH were then added and the mixture stirred until it turned again colourless. The reaction mixture was dried to obtain a yellow-pale solid, that was partitioned between HCl 2M and DCM (washed three times). The acidic phase pH was adjusted to pH 1 using concentrated NaOH, and then extracted many times with DCM. The organic layer after anhydrification over Na_2SO_4, was filtered and evaporated to give 575 mg (Yield: 99%) of the desired product.

(7) Full characterization is reported in *Chapter 2*.

N-boc-2,6-diaminomethylnapthalene (8). Compound **7** (420 mg, 2.26 mmol) and Boc-OPh (0.507 mL, 2.71 mmol) were dissolved in 1,5 mL of absolute EtOH and stirred at 75°C for 22h, then the reaction was partitioned between water (100 mL) and DCM (100 mL). The organic phase after drying gave a crude product that was purified by flash chromatography on silica gel using DCM/MeOH 96/4 (1% NH_3) to afford 301 mg (Yield: 47%) of pure product. TLC (Silica gel, AcOEt/MeOH 7/3 + 1% NH_3) R_f: 0.49 in MeOH 4/6DCM. (Note: Eluent containing ammonia were dried with Na_2SO_4 prior to use).

(8) *^{1}H-NMR* (*300 Mhz, 25°C, $CDCl_3$*) δ(ppm):7.79 (d,2H, *J*=8.4Hz, aromatic CH), 7.69 (d,2H, *J*=9.0 Hz, aromatic CH), 7.43-7.36 (m, 2H, aromatic CH), 4.93 (s (broad), 1H, NHBoc), 4.47 (d, 2H, *J*=5.4 Hz, CH_2NHBoc), 4.03 (s, 2H, CH_2NH_2), 1.65 (s, 2H, CH_2NH_2), 1.48 (s, 9H, t-butyl). *^{13}C-NMR* (75 *Mhz, 25°C, $CDCl_3$*) δ(ppm): 141.1, 136.6, 133.4, 131.7, 128.1, 127.9, 125.7, 125.4, 124.9, 77.2, 46.5, 44.8, 29.6, 29.4.

For synthesis and characterization of **5-carboxymethyl-uracil (9)** see *Chapter 2*.

N(3)-(4,4'-dimetoxybenzidryl)-5-carboxyuracil methyl ester (10). Compound **9** (1.07g, 6.26 mmol) and Boc_2O (1.50g, 6.7 mmol) were dissolved in 40 mL of dry MeCN and stirred at room temperature under dry argon atmosphere. After 5h the solvent was distilled-off under reduced pressure and the yellow-pale oil obtained was re-dissolved in 20 mL of dry DMF. NaH (450 mg, 8.45 mmol) and Dod-Cl (1.84g, 7.00 mmol) were added at 0° C under argon and stirred at room temperature for 15h. After reaction completion, the mixture was partitioned between $KHSO_4$ 0.1M (200 mL) and AcOEt (200 mL), and the organic layer washed three times with water, and finally with sat. NaCl. The organic layer was dried under reduced pressure to afford an yellow oil that was dissolved in 80 ml of MeOH. To this solution 420 mg of K_2CO_3 were added and the mixture was stirred at room temperature. for 2h. The solvent was then removed and the residue partitioned between $KHSO_4$ 0.1M and EtOAc. The organic layer was dried and the crude product submitted to flash chromatography on silica gel using AcOEt/Hexane 7/3 as eluent to afford 2.18 g (Yield: 62%) of 10 as white foam. TLC (Silica gel AcOEt/Hexane R_f:0.3).

(10) *^{1}H-NMR* (*300 Mhz, 25°C, $CDCl_3$*) δ(ppm): 9.97 (d, 1H, *J*=6.0Hz, NH in 1), 8.08 (d, 1H, *J*=6.0Hz, CH in 6), 7.29 (d,4H, *J*=9.0Hz, aromatic CH), 7.28 (s, 1H, Ar_2CH-N), 6.82 (d,4H, *J*=9.0Hz, aromatic CH), 3.83 (s, 3H, -COOMe), 3.77 (s, 6H, Ar-OMe). *^{13}C-NMR* (75 *Mhz, 25°C, $CDCl_3$*) δ(ppm): 163.5, 159.5, 158.8, 151.2, 146.8, 129.9, 129.8, 113.5, 104.5, 58.3, 55.2, 52.3.

N(1)-Methyl-N(3)-(4,4'-dimetoxybenzidryl)-5-carboxyUracil (11). Compound **10** (0.69g, 1.75 mmol), and K_2CO_3 were dissolved in 5mL of MeCN and stirred at room temperature for 5 min followed by the addition of MeI (0.166 mL, 2.63 mmol). The reaction mixture was stirred for 3h and monitored by TLC on silca gel (DCM/THF 9/1) until it was fully converted to *6a* (Rf(*6a*): 0.63, Rf(10):0.45). The reaction was then dried and the residue dissolved in 3 mL of MeOH and 3 mL of aqueous NaOH 3M and stirred for further 3h until complete hydrolysis of methyl ester (monitored by on silca gel using DCM/THF 9/1, Rf(*6a*): 0.63, Rf(6): 0.41). The reaction mixture was acidified using $KHSO_4$ 0.5M to about pH 3, extracted with AcOEt and treated with Na_2SO_4. The organic layer was evaporated to afford a white foam. Cold recrystallization from chloroform and hexane afforded 0.64 g of desidered product **6** (Yield: 92%).

(11) *^{1}H-NMR* (*300 Mhz, 25°C, $CDCl_3$*) δ(ppm): 12.59 (s,broad, 1H, COOH), 8.43 (s, 1H, CH in 6), 7.30 (d,4H, *J*=8.4Hz, aromatic CH), 7.29 (s, 1H, Ar_2CH-N), 6.87 (d,4H, *J*=9.0Hz, aromatic CH), 3.81 (s, 6H, Ar-OMe), 3.51 (s, 3H,N-CH_3).

For synthesis and characterization of **N(1)-ethoxycarbonylmethyl-5-carboxylic acid (12)** see *Chapter 2*.

5-(2-methylaminonapthyl-N-carbonyl)- N(1)-methylenecarboxylate-Uracil (13). Compound **12** (110 mg, 0.45 mmol), HBTU (172 mg, 0.45 mmol) and 0.23 mL of DIEA were dissolved in 2 mL of dry DMF and stirred at 0°C for 5 min and further 5 min at room temperature, then 156 mg (0.59 mmol) of **6** dissolved in 2 mL of dry DMF and further 0.23 mL of DIEA were added. The reaction mixture was stirred at room temperature overnight (The reaction was monitored by TLC on silca gel and AcOEt as eluent, R_f of intermediate : 0.72) and then partitioned between sat. $KHSO_4$ and AcOEt. The organic layer was washed with saturated $NaHCO_3$, saturated NaCl and dried. The residue was dissolved in 20 mL of MeOH and 20 mL of NaOH 2M and stirred at room temperature for 3h. After complete conversion (reaction monitored by TLC on same condition displayed above showed Rf:0 for product) the reaction mixture was acidified to pH 4 (using concentrated HCl) and centrifuged to obtain 79 mg (Yield: 49%) of a white solid.

(13) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 13.30(s, 1H, COOH), 12.04 (s, 1H, NH uracil), 9.22 (t, 1H, *J*=6.0Hz, NH amide), 8.63 (s, 1H, CH in 6 uracil), 7.90-7.86 (m, 3H, CH napht.), 7.78 (m, 1H, CH napht), 7.52-7.45 (m, 2H, CH napht.), 4.67 (d, 2H, *J*=6.0Hz, CH_2 napht.), 4.63 (s, 2H, CH_2C(O)N linker). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 169.1, 163.5, 161.7, 151.8, 150.0, 136.8, 132.8, 132.0, 128.0, 127.5, 126.2, 125.8, 125.6, 125.3, 104.6, 49.3, 42.2.

5-(2,7-dimethylaminonaphtyl-N-Boc-N-carbonyl)-N(1)-methylenecarboxylate-Uracil (14). Compound **12** (530 mg, 2.18 mmol), HBTU (790 mg, 2.10 mmol) and 1.12 mL of DIEA were dissolved in 10 mL of dry DMF and stirred at 0°C for 5 min and further 5 min at room temperature, then 520 mg (1.82 mmol) of **8** dissolved in 10 mL of dry DMF and further 1.12 mL of DIEA were added. The reaction mixture was stirred at room temperature overnight (the reaction was monitored by TLC on silca gel and AcOEt as eluent, R_f of intermediate : 0.68) and then partitioned between 0.1M $KHSO_4$ and AcOEt. The organic layer was washed

with saturated $NaHCO_3$, saturated NaCl and dried. The residue was dissolved in 25 mL of MeOH and 25 mL of NaOH 2M and stirred at room temperature for 3h. After complete conversion (reaction monitored by TLC under the same conditions displayed above showed Rf about 0 for product) the reaction mixture was adjusted to pH 7 and partitioned between sat. $KHSO_4$ 1M and AcOEt. The organic layer was then dried on Na_2SO_4 and the solvent removed under reduced pressure to obtain 810 mg of a white solid (Yield: 95%).

(14) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 13.29 (s (broad), 1H, COOH), 12.04 (s, 1H, NH uracil), 9.21 (t, 1H, *J*=6.0Hz, NH amide), 8.63 (s, 1H, CH(6) uracil), 7.84 (dd, 2H, *J*=8.5Hz, 1.3Hz, aromatic CH napht.), 7.69 (d, 2H, *J*=8.5Hz), 7.73 (s, 1H, aromatic CH napht.), 7.66 (s, 1H, aromatic CH napht.), 7.49 (t, 1H, *J*=6.1Hz, NH-Boc), 7.42 (dd, 1H, *J*=8.5Hz, 1.5Hz, aromatic CH napht.), 7.37 (dd, 1H, *J*=8.4Hz, 1.4Hz, aromatic CH napht.), 4.66 (d, 2H, *J*=5.9Hz, CH_2 naptht.), 4.62 (s, 2H, CH_2-COOH), 4.27 (d, 2H, *J*=6.1Hz, CH_2 naptht.), 1.40 (s, 9H, t-butyl). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm):169.1, 163.5, 161.6, 155.7, 151.7, 149.9, 137.9, 136.9, 132.6, 131.0, 127.7, 127.5, 125.3, 125.0, 124.7, 104.5, 77.7, 49.2, 43.4, 42.1, 28.1.

5-(2,7-dimethylaminonaphtyl-N-Boc-N-carbonyl)-Uracil PNA monomer-OMe (15). Compound **14** (294mg, 0.63 mmol), DCC (128mg, 0.62 mmol) and DhBtOH (112 mg, 0.69 mmol) were dissolved in 6 mL of DMF and stirred at 0°C for 30 min and for further 30 min at room temperature. Compound **3** (270 mg, 0.76 mmol) and 0.130 mL of DIEA were added, and the mixture stirred overnight. Reaction was stopped by adding 30 mL of AcOEt and filtered to remove DCU (dicyclohexylurea). Filtrate was further diluted with the same amount of AcOEt and washed with water (three times), sat. $NaKHSO_4$ (three times), sat. $NaHCO_3$ (three times) and once with saturated NaCl. The organic layer was anhydrified on Na_2SO_4 and evaporated to afford crude product as white foam. Gradient elution chromatography on silica gel (AcOEt/Hexane 8/2 (R_f:0.19) to AcOEt) afforded 370 mg (Yield: 72%) of pure product. .

(15) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 11.96 (s, 1H, NH uracil), 9.23 (t, 1H, *J*=5.9Hz, NH amide), 8.48(M) and 8.39 (m) (s, 1H, CH uracil), 7.88 (d, 2H, *J*=7.5Hz, aromatic CH fmoc), 7.84 (d, 2H, *J*=8.4Hz, aromatic CH napht.), 7.72 (s, 1H, aromatic CH napht.), 7.68 (d, 2H, *J*=7.5Hz, aromatic CH fmoc), 7.66 (s, 1H, aromatic CH napht.), 7.48 (t, 1H, *J*=6.3Hz, NH-Boc), 7.45-7.36 (m, 4H, 2H aromatic CH fmoc and 2H aromatic CH napht.), 7.32 (dd, 2H,

J=7.5Hz, 1.2Hz, aromatic CH fmoc), 7.27 (t, 1H, *J*=5.7Hz, NH-Fmoc) 4.95(M) and 4.75 (m) (s, 2H, CH_2-C(0)N linker), 4.66 (d, 2H, *J*=5.7Hz, CH_2 naptht.), 4.36-4.21 (m, 5H, CH_2 and CH Fmoc, CH_2 napht.), 4.08, 3.72(M) and 3.61(m) (s, 2H, CH_2 backbone), 3.63(s, 3H, COOMe), 3.44(M) and 3.04(m) (t, 2H, *J*(M)=6.3, *J*(m)=5.7, CH_2 backbone), 3.27(M) and 3.12(m) (q, 2H, *J*(M)=5.4, *J*(m)=6.0, CH_2 backbone), 1.40 (s, 9H, t-butyl). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 169.6(m) and 169.3(M), 167.2(m) and 166.9(M), 163.6, 161.7, 156.3, 155.8, 152.0(M) and 151.9(m), 150.0, 143.8, 140.6, 138.0, 137.0, 132.7, 131.1, 127.7, 127.5, 127.0, 125.4, 125.1, 124.8, 120.0, 104.5, 77.7, 65.4(M) and 65.3(m), 52.2, 51.7, 48.4, 47.7, 46.9, 46.6, 43.5, 42.2, 37.8, 28.2.

5-(2-methylaminonapthyl-N-carbonyl)- Uracil PNA monomer-OH (16). Compound **13** (60mg, 0.17 mmol), EDC (33mg, 0.168 mmol) and DhBtOH (29 mg, 0.17 mmol) were dissolved in 1 mL of DMF and stirred at 0°C for 30 min and for further 30 min at room temp.. Fmoc-aminoethylglycine-*tert*-butylester (142 mg, 0.63 mmol) and 73 μL of DIEA were added, and the mixture stirred overnight. Reaction was stopped by adding 40 mL of AcOEt and washed with water (three times), 0.1M $KHSO_4$ (three times), sat. $NaHCO_3$ (three times) and once with sat. NaCl**.** The organic layer was evaporated and residue dispersed in a mixture composed of 1.2mL of DCM, 0.6 mL of TFA and 0.2 mL of triethylsilane (TES) and mixture stirred at 0°C for 3h. Reaction mixture was dried and followed by precipitation using hexane to afford 72 mg (Yield: 75%) of pure product.

(16) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 12.0 (s, 1H, NH uracil), 9.23 (t, 1H, *J*=6.0Hz, NH amide), 8.49(M) and 8.37 (m) (s, 1H, CH uracil), 7.89-7.87 (m, 5H, aromatic CH fmoc and CH napht.), 7.78-7.67 (m, 3H, aromatic CH fmoc and CH napht.), 7.50-7.38 (m, 6H, aromatic CH fmoc and CH napht. and NH-Fmoc), 7.32 (t, 2H, *J*=7.4Hz, aromatic CH fmoc), 4.94(M) and 4.75 (m) (s, 2H, CH_2-C(0)N linker), 4.67 (d, 2H, *J*=6.0Hz, CH_2 naptht.), 4.35-4.20 (m, 5H, CH_2 and CH Fmoc, CH_2 napht.), 3.99 (s, 2H, CH_2 backbone), 3.44-3.26 and 3.14-3.09 (m, 4H, CH_2 CH_2 backbone). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 170.6(m) and 170.2(M), 167.1(m) and 166.7(M), 163.6, 161.7, 156.3, 156.0, 152.1(M) and 151.9(m), 150.0, 143.8, 140.6, 136.8, 132.8, 132.0, 128.0, 127.51-127.47, 127.0, 126.1, 125.8, 125.6, 125.3, 125.1, 120.0, 104.4, 65.4(M) and 65.3(m), 48.9(m) and 48.4(M), 48.6(m) and 48.7(M), 46.9, 46.8, 46.6, 42.2.

5-(2,7-dimethylaminonaphtyl-N-Boc-N-amidyl)-Uracil PNA monomer-OH (17). Compound **15** (68 mg, 0.083 mmol) was dissolved in 4 mL of a mixture 1/1 water/THF, then $Ba(OH)_2$ $8H_2O$ (40 mg, 0.125 mmol) for 35-40 min controlling the reaction process through TLC (silica gel, AcOEt/MeOH 7/3, R_f:0.12). Reaction mixture was acidified to pH 3-4 and partitioned between 0.1M $KHSO_4$ and AcOEt. Organic layer, after washing with sat. NaCl and anhydrification on Na_2SO_4, afforded 61 mg (Yield: 90%) of the PNA monomer product used for PNA oligomer synthesis without further purification.

(17) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm):12.8 (s, 1H, COOH), 12.0 (s, 1H, NH uracil), 9.22 (t, 1H, *J*=5.8Hz, NH amide), 8.48 (M) and 8.37 (m) (s, 1H, CH(6) uracil), 7.88 (d, 2H, *J*=7.5Hz, aromatic CH fmoc), 7.84 (d, 2H, *J*=8.5Hz, aromatic CH napht.), 7.72-7.66 (m, 4H, aromatic CH napht. and CH fmoc), 7.48 (t, 1H, *J*=6.1 Hz, NH-Boc), 7.43-7.29 (m, 7H, 4H aromatic CH , aromatic CH napht. and NH-Fmoc), 4.94(M) and 4.74 (m) (s, 2H, CH_2-C(0)N linker), 4.66 (d, 2H, *J*=5.5Hz, CH_2 naptht.), 4.35-4.19 (m, 5H, CH_2 and CH Fmoc, CH_2 napht.), 3.98 (s, 2H, s, CH_2 gly backbone), 3.42-3.26 (m, 4H, CH_2-CH_2 backbone), 1.40 (s, 9H, t-butyl).

5-[2,7-dimethylaminonaphtyl-N-carbonyl-(N(1)-methyl-N(3)-dod-Uracil)-N-carbonyl]-Uracil PNA monomer-OMe (18). Compound **15** (231mg, 0.282 mmol) was dissolved in 3mL of DCM and 3 mL of TFA at 0°C, and stirred for 1h at room temp. The reaction mixture was dried affording boc deprotected product of compound **15**. At the same time compound **11** (134 mg, 0.338 mmol), HBTU (116 mg, 0.305 mmol) and 0.12 mL of DIEA were dissolved in 3 mL of dry DMF and stirred at 0°C for 5 min and further 5 min at room temperature, then deprotected **15** previously prepared (dissolved in 3ml af DMF) was added. After stirring for 4h, reaction was diluted with AcOEt and washed with water (three times), 0.1M $KHSO_4$ (three times), sat. $NaHCO_3$ (three times) and sat. NaCl. Organic layer was dried and crude product purified by flash chromatography on silica gel (AcOEt/MeOH 97/3, R_f:0.24) to afford 301 mg (Yield:88%) of pure product.

(18) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 11.99 (s, 1H, NH uracil), 9.21 (t, 1H, *J*=5.4Hz, NH amide), 9.17 (t, 1H, *J*=6.3Hz, NH amide), 8.63 (1H, s, CH uracil), 8.47(M) and 8.37 (m) (s, 1H, CH uracil), 7.88 (d, 2H, *J*=7.2Hz, aromatic CH fmoc), 7.83 (d, 2H, *J*=8.4Hz, aromatic CH napht.), 7.73 (s, 1H, aromatic CH napht.), 7.71 (s, 1H, aromatic CH napht.), 7.68 (d, 2H, *J*=7.8Hz, aromatic CH fmoc), 7.45 (t, 1H, *J*=5.7 Hz, NH-Fmoc), 7.44-7.37 (m, 4H, 2H aromatic

CH fmoc and 2H aromatic CH napht.), 7.31 (t, 2H, *J*=7.5Hz, aromatic CH fmoc), 7.20 (dt, 4H, *J*=8.8Hz, 2.9Hz, aromatic CH dod), 7.17 (s, 1H, $CHAr_2$ dod), 6.86 (dt, 4H, *J*=8.8Hz, 2.9Hz, CH dod), 4.94(M) and 4.75 (m) (s, 2H, CH_2-C(0)N linker), 4.66 (d, 2H, *J*=6.6Hz, CH_2 naptht.), 4.62 (d, 2H, *J*=6.3Hz, CH_2 naptht.), 4.35-4.18 (m, 3H, CH_2 and CH Fmoc), 4.07(M) and 3.96(m) (s, 2H, CH_2 backbone), 3.72 (s, 6H, OMe dod), 3.63(s, 3H, COOMe), 3.44(M) (t, 2H, *J*(M)=6.7, CH_2 backbone), 3.40 (s, 3H, Me uracil), 3.27(M) and 3.11(m) (q, 2H, *J*(M)=6.0, *J*(m)=5.7, CH_2 backbone). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 169.6(M) and 169.3(m), 167.2(m) and 166.9(M), 163.6, 162.6, 161.7, 158.2, 156.3, 156.0, 152.0(M) and 151.9(m), 151.1, 150.0, 143.8, 140.6, 137.2, 137.0, 132.7, 131.1, 129.9, 129.5, 127.74, 127.69, 127.5, 127.0, 125.8, 125.5, 125.1, 125.0, 120.0, 113.3, 104.5, 103.6, 65.4(M) and 65.3(m), 59.6, 54.9, 52.2, 51.7, 48.7, 48.4, 47.7, 46.9, 46.6, 42.3, 42.2, 37.3.

5-[2,7-dimethylaminonaphtyl-N-carbonyl-(N(1)-methyl-N(3)-dod-Uracil)-N-carbonyl]-Uracil PNA monomer-OH (19). Hydrolisis of methyl ester was carried out using the same procedure previously described for **17**.(Yield:46%) TLC on silica gel, AcOEt/MeOH 7/3, R_f:0.17.

(19) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 12.70 (s, 1H, COOH), 11.99 (s, 1H, NH uracil), 9.20-9.17 (m, 2H,NH amide), 8.63 (1H, s, CH(6) uracil), 8.47(M) and 8.37 (m) (s, 1H, CH uracil), 7.87 (d, 2H, *J*=7.4Hz, aromatic CH fmoc), 7.83 (d, 2H, *J*=8.6Hz, aromatic CH napht.), 7.73-7.67 (m, 4H, aromatic CH napht and fmoc), 7.45-7.37 (m, 4H, 2H aromatic CH fmoc, 2H aromatic CH napht and NH-Fmoc), 7.31 (t, 2H, *J*=7.4Hz, aromatic CH fmoc), 7.22-7.12 (m, 5H, 4H CH dod and 1H Ar_2CH), 6.90-6.84 (m, 4H, CH dod), 4.93(M) and 4.74 (m) (s, 2H, CH_2-C(0)N linker), 4.66-4.60 (m, 4H, CH_2 naptht.), 4.34-4.19 (m, 3H, CH_2 and CH Fmoc), 3.99(M) (s, 2H, CH_2 backbone), 3.72 (s, 6H, OMe dod), 3.40 (s, 3H, Me uracil), 3.40-3.27 (m, 4H, CH_2-CH_2 backbone).

5-hydroxymethyluracil (20). Uracil (10g, 90 mmol), para-formaldehyde(8.5g, 283 mmol) and 17,5 mL of triethylamine were dispersed in 200 mL of water and stirred at 60°C in a closed shlenk overnight. Reaction mixture was then concentrated (CAUTION!) to obtain a yellow oil that was re-crystalizzed by water/ethanol to afford 10.3g (Yield 81%) of desired product.

(20) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 11.05 (s, 1H, NH in 1), 10.71 (s, 1H, NH in 3), 7.24 (s, 1H, CH in 6), 4.85 (s, 1H, OH), 4.10 (s, 2H, CH_2). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 163.8, 151.3, 138.2, 112.7, 55.8. FT-IR (KBr pellet) ν (cm^{-1}): 3369, 3189, 3038, 2865, 1704, 1672. ESI(-)-MS(MeOH) m/z for: $[M-H]^-$ Calcd.141.1 Found 141.0. Elem. Anal. (%): C Calcd 42.26, Found 42.24, H Calcd. 4.26 Found 4.35, N Calcd 19.33 Found 19.41.

5-chloromethyluracil (21). Compound **20** (6.7g, 47.6 mmol) was dissolved in 25mL of HCl 37% w/w and stirred at room temp.. Reaction mixture remained clear for few minutes until product started to precipitate. After 4 h the precipitate was collected by filtration from the reaction mixture and dried under vacuum/P_2O_5 system to afford desired product (6.1g, Yield:80%).

(21) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 11.27 (s, 1H, NH in 1), 11.04 (s, 1H, NH in 3), 7.73 (s, 1H, CH in 6), 4.40 (s, 2H, CH_2). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 162.8, 151.0, 142.4, 108.8, 39.9. FT-IR (KBr pellet) ν (cm^{-1}): 3305, 3125, 3044, 2829, 1759, 1697, 1656, 1178, 738. ESI(-)-MS(MeOH) m/z for: $[M-H(^{35}Cl)]^-$ Calcd.159.0 Found 159.0. Elem. Anal. (%): C Calcd 37.40, Found 37.48, H Calcd. 3.14 Found 3.20, N Calcd 17.45 Found 16.94.

5-azidomethyluracil (22). NaN_3 (2.65g, 40.8 mmol) was dissolved in 25 mL of dry DMF, then a solution of compound **21** (5.9g, 36.6 mmol) in 80mL of dry DMF was added dropwise at 0°C. Reaction was stirred at 0°C for 2h, then 1 mL of concentrated HCl was added carefully even at 0°C (*WARNING: this step is hazardous for HN_3 development, please be carefull. Reaction scale carried out here is safe and its scale-up could be riskfull*). DMF was collect by distillation under reduced pressure and the solid residue after washings with cold water afforded 5.5g of the desired product (Yield: 90%). TLC (MeOH): R_f 0.79.

(22) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*) δ(ppm): 11.27 (s, 1H, NH in 1), 11.17 (s, 1H, NH in 3), 7.65 (s, 1H, CH in 6), 4.02 (s, 2H, CH_2). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 163.8, 151.0, 141.9, 106.4, 46.3. FT-IR (KBr pellet) ν (cm^{-1}): 3219, 3079, 3026, 2825, 2132 (stretching N_3), 1757, 1689, 1667. ESI(-)-MS(MeOH) m/z for: $[M-H]^-$ Calcd. 166.0 Found 166.0. Elem. Anal. (%): C Calcd 35.93, Found 35.53, H Calcd. 3.02 Found 2.95, N Calcd 41.90 Found 41.51.

5-azidomethyl-N(1)-carboxymethyl-t-butyl ester (23). Compound **22** (510 mg, 3.1 mmol) and K_2CO_3 (422mg, 3.1 mmol) were dispersed in 5mL of dry DMF at 0°C, then one equivalent of tert-butylbromoacetate solution (0.493 mL in 1mL of DMF) was added dropwise. After raising temperature to room temp. the mixture was stirred overnight. DMF was then removed from reaction the mixture and the solid residue was partitioned between AcOEt and water. The organic layer was dried and crude product was purified by silica gel chromatography using AcOEt/Hexane 5/3 as eluent (R_f: 0.27) to afford 420 mg of pure product (Yield:49%).

(23) *^{1}H-NMR* (*300 Mhz, 25°C, $CDCl_3$*) δ(ppm): 9.07 (s, 1H, NH in 3), 7.17 (s, 1H, CH in 6), 4.38 (s, 2H, CH_2N_3), 4.17 (s, 2H, N- CH_2-CO), 1.49 (s, 9H, t-butyl). *^{13}C-NMR* (75 *Mhz, 25°C, $CDCl_3$*) δ(ppm): 166.2, 162.4, 150.2, 142.7, 109.7, 83.8, 49.4, 46.9, 28.0. FT-IR (KBr pellet) ν (cm^{-1}): 3178, 3054, 2998, 2824, 2117 (stretching N_3), 1736, 1705. ESI(-)-MS(MeOH) m/z for: $[M-H)]^-$ Calcd. 280.3 Found 280.2. Elem. Anal. (%): C Calcd 46.97, Found 47.05, H Calcd. 5.38 Found 5.33, N Calcd 24.90 Found 24.56.

5-azidomethyl-N(1)-carboxymethyl-ethyl ester (24). Compound **22** (1.0g mg, 6.1 mmol) and K_2CO_3 (839mg, 6.1 mmol) were dispersed in 10mL of dry DMF at 0°C, then one equivalent of ethylbromoacetate solution (0.493 mL in 1mL of DMF) was added dropwise. After raising temperature to room temp. the mixture was stirred overnight. DMF was then removed from the reaction mixture and the solid residue was partitioned between AcOEt and water. The organic layer was dried and crude product was purified by silica gel chromatography using AcOEt/Hexane 8/2 as eluent (R_f: 0.30) to afford 630 mg of pure product (Yield:41%).

(24) *^{1}H-NMR* (*300 Mhz, 25°C, $CDCl_3$*) δ(ppm): 8.83 (s, 1H, NH in 3), 7.17 (s, 1H, CH in 6), 4.48 (s, 2H, CH_2N_3), 4.26 (q, *J*=7.1Hz, 2H, OCH_2-Me), 4.18 (s, 2H, N-CH_2-CO), 1.31 (t, 3H, *J*=7.1Hz, OCH_2-CH_3). *^{13}C-NMR* (75 *Mhz, 25°C, $CDCl_3$*) δ(ppm): 167.1, 162.4, 150.2, 150.2, 110.0, 62.4, 48.8, 46.9, 14.0. FT-IR (KBr pellet) ν (cm^{-1}): 3160, 3036, 2841, 2109 (stretching N_3), 1739, 1701. ESI(-)-MS(MeOH) m/z for: $[M-H)]^-$ Calcd. 252.2 Found 252.2. Elem. Anal. (%): C Calcd 42.69, Found 42.82, H Calcd. 4.38 Found 4.28, N Calcd 27.66 Found 27.37.

5-azidomethyl-N(1)-carboxymethyl acid (25). *Route a* (*From compound 23)* Compound **23** (386 mg, 1.36 mmol) was dissolved in 5mL of DCM, then 3mL of TFA were added at 0°C. Mixture was stirred for 2h, and then evaporated to give a yellow oil that was partitioned between water and ethyl acetate. The organic layer after anhydrification with Na_2SO_4 was dried to afford 216 mg of pure product (Yield:70%). *Route b* (*From compound 24)* Compound **24** (1.15g, 4.54 mmol) was dissolved in 20 mL of aqueous NaOH 1M/MeOH and stirred for 1h. MeOH was removed by distillation from reaction mixture and aqueous residue was acidified until pH3 with HCl and partitioned with AcOEt. Organic layer after anhydrification with Na_2SO_4 and drying afforded 932 mg of pure product (Yield: 95%).

(25) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO)* δ(ppm): 13.21 (s, 1H, COOH), 11.65 (s, 1H, NH (3) uracil), 7.83 (s, 1H, CH in 6), 4.43 (s, 2H, CH_2N_3), 4.06 (s, 2H, NCH_2-COOH). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 169.3, 163.3, 150.6, 145.3, 107.3, 48.6, 46.5. FT-IR (KBr pellet) ν (cm^{-1}): 3412, 3002, 2825, 2108 (stretching N_3), 1751, 1700, 1691, 1482. ESI(-)-MS(MeOH) m/z for: $[M-H)]^-$ Calcd. 224.2 Found 224.3. Elem. Anal. (%): C Calcd 37.34, Found 37.45, H Calcd. 3.13 Found 3.32, N Calcd 31.10 Found 30.24.

5-azidomethyluracil-PNA-OEt monomer (26). Compound **25** (0.7 g, 3.1 mmol), DhBtOH (0.4 g, 2.5 mmol) and DCC(0.63 g, 3.05 mmol) were dissolved in 10 mL of DMF and stirred at 0°C for 30 min and for further 30 min at room temp. Formation of active esters was evidenced by formation of a DCU precipitate. After activation, Boc-aminoethylglicine ethyl ester (0.93 g, 3.3 mmol) and 0.8 mL of DIEA were added, and the reaction mixture was stirred overnight, then filtered, diluted with 100 mL of AcOEt and washed with water (3 times), sat. $NaHCO_3$ (3 times), $KHSO_4$ (3 times) and NaCl (once); the organic layer, after anhydrification with Na_2SO_4 and drying under reduced pressure to obtain the crude product. Purification was carried out by flash chromatography on silica gel using AcOEt as eluent (R_f:0.33) affording to 1.06g (Yield: 77%) of pure compound.

(26) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO)* δ(ppm): 11.59 (s, 1H, NH in 3), 7.66 (m) and 7.64 (M) (s, 1H, CH in 6), 6.94(M) and 6.76 (m) (t, 1H, J(M)=5.5Hz , J(m)=5.3Hz, NH- Boc), 4.72(M) and 4.54(m) (s, 2H, N-CH_2-CO base linker), 4.25-3.95 (m, 6H, CH_2N_3, O-CH_2-Me, N-CH_2-CO-backbone), 3.41 (M) and 3.31 (m) (t, 2H, *J(M)*=6.5Hz, *J(m)*=6.9Hz, N-CH_2-CH_2-NBoc-), 3.17(M) and 3.02(m) (q, 2H, *J(M)*=6.0Hz, *J(m)*=6.2Hz, N CH_2-CH_2- NBoc), 1.38(M) and 1.37(m) (s, 9H,

t-butyl), 1.24(M) and 1.17(m) (t, 3H, *J(M)*=7.1Hz, *J(m)*=7.1Hz, OCH_2-CH_3). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 169.1(m) and 168.8 (M), 167.2(m) and 167.0(M), 163.3, 155.6, 150.5, 145.4, 107.1, 77.9(M) and 77.6(m), 61.0 (m) and 60.4 (M), 48.8(m) and 47.7(M), 48.0(m) and 47.6(M), 46.7, 46.5, 46.4, 38.0(M) and 37.5(m), 28.0, 13.9. FT-IR (KBr pellet) ν (cm^{-1}): 3329, 3210, 3065, 2981, 2934, 2851, 2108 (stretching N_3), 1750-1600, 1466. ESI(-)-MS(MeOH) m/z for: $[M-H)]^-$ Calcd. 452.2 Found 452.6. Elem. Anal. (%): C Calcd 47.78, Found 48.22, H Calcd. 6.00 Found 6.29, N Calcd 21.62 Found 20.20.

5-azidomethyluracil-PNA-OH-monomer (27). Compound **26** (1.09g, 2.40 mmol) was dissolved in aqueous NaOH 1M (10mL) and MeOH (10mL) and stirred at room temp. for 90 min.. The reaction mixture was neutralized with HCl and after MeOH evaporation under reduced pressure was diluted with sat. $KHSO_4$. The acidic layer was extracted with AcOEt three times and the organic solution, after anhydrification on Na_2SO_4, was dried to afford 980 mg (Yield:96%) of pure compound. TLC (silica gel, AcOEt/MeOH 6/4), R_f: 0.27.

(27) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO)* δ(ppm): 12.80 (s, 1H, COOH), 11.59(M) and 11.57(m) (s, 1H, NH in 3), 7.67 (M) and 7.64 (m) (s, 1H, CH in 6), 6.94(M) and 6.77 (m) (t, 1H, J(M)=5.4Hz , J(m)=5.2Hz, N- Boc), 4.71(M) and 4.53(m) (s, 2H, CH_2N_3), 4.20(m) and 4.06(M) (s, 2H, N-CH_2-CO base linker), 4.01(m) and 3.97(M) (s, 2H, N-CH_2-CO- backbone), 3.39(M) and 3.30(m) (t, 2H, J*(M)*=6.5Hz, *J(m)*=6.9Hz, N-CH_2-CH_2-NBoc), 3.16(M) and 3.02(m) (q, 2H, *J(M)*=5.7Hz, *J(m)*=5.7Hz, N CH_2-CH_2- NBoc), 1.38(M) and 1.37 (m) (s, 9H, t-butyl). *^{13}C-NMR* (75 *Mhz, 25°C, d_6-DMSO*) δ(ppm): 170.6 (m) and 170.2(M), 167.1(m) and 166.7(M), 163.3, 155.6, 150.5, 145.6(m) and 145.5(M), 77.9(M) and 77.6(m), 48.8(m) and 47.7(M), 48.0(m) and 47.4(M), 47.0-46.0, 37.9(M) and 37.4(m), 28.0. FT-IR (KBr pellet) ν (cm^{-1}): 3360, 3178, 3061, 2980, 2935, 2853, 2109 (stretching N_3), 1800-1600, 1472. ESI(+)-MS(MeOH) m/z for: $[M+H)]^+$ Calcd. 426.4 Found 426.3.

M and m respectively for Major and minor rotamers.

3.7 References

[1] Leumann CJ, *Bioorg. & Med. Chem.,* **2002,** *10 (4),* 841-854;

[2] Porcheddu A, Giacomelli G, *Curr. Med. Chem.,***2005***, 12 (22)*, 2561-2599;

[3] Loakes D., *Nucl. Acid. Res.*, **2001**, *29(12)*, 2437-2447;

[4] Krueger AT, Lu HG, Lee AHF, Kool E. T., *Acc. Chem. Res.*, **2007**, *40 (2)*, 141-150;

[5] Lynch SR, Liu HB, Gao JM, Kool E.T., *J. Am. Chem. Soc*., **2006**, *128 (45)*, 14704-14711;

[6] Flanagan W. M., Wolf. J.J., Olson P., Grant D., Lin K.Y., Wagner R.W., Matteucci M.D., *Proc. Nat. Acc. Sci.*, **1999**, *96(7)*, 3513-3518;

[7] Corradini R., Sforza S., Tedeschi T., Totsingan F., Marchelli R., *Curr. Top. Med. Chem.*, **2007**, *7(7)*, 681-694;

[8] Wojciechowski F, Hudson RHE , **2007**, *Curr. Top. Med. Chem.*, *7(7)*, 667-679;

[9] K. G. Rajeev, M. A. Maier, E. A. Lesnik, M. Manoharan, *Org. Lett.*, **2002**, *4*(*25*), 4395-4398;

[10] V. Chenna, S. Rapireddy, B. Sahu, C. Ausin, E. Pedroso, D. H. Ly, *ChemBioChem*, **2008**, *9*, 2388 – 2391;

[11] F. Wojciechowski, R. H. E. Hudson, *JACS*, **2008**, *130*, 12574–12575;

[12] A. B. Eldrup, C. Christensen, G. Haaima, P. E. Nielsen, *JACS*, **124**, 3254-3262;

[13] C. Greenman1 et al, *Nature*, **2007**, *446*, 153;

[14] S. Cogoi, A. Codognotto, V. Rapozzi, N. Meeuwenoord, G. van der Marel, L. E. Xodo, *Biochemistry*, **2005**, *44*, 10510-10519;

[15] S. M. Freier, N. Sugimoto, A. Sinclair, D. Alkema, T. Nielson, R. Kierzek, M. H. Caruthers, D. H. Turner, *Biochemistry*, **1986**, *25*, 3214;

[16] E. T. Kool, *Chem. Rev.*, **1997,** 97, 1473-1487;

[17] L. Betts, A. Josey, J.M. Veal, S.R. Jordan, *Science,* **1995**, *270*, 1838-1841;

[18] E. Czarina, M. Juan, J. Kondo, T. Kurihara, T. Ito, Y. Ueno, A. Matsuda, A. Take´naka, *Nucl. Acid. Res*., **2007**, 35(6), 1969–1977;

[19] Accetta A., Corradini R., Sforza S., Tedeschi T., Brognara E., Borgatti M, Gambari R., Marchelli R., *J. Med. Chem.*, **2009**, *52(1)*, 87-94; See also *CHAPTER 2* for further details;

[20] Jonsson D., A. Unden, *Tetrahedron Lett.*, **2002**, *43*, 3125-3128;

[21] T.Tedeschi, S.Sforza, F.Maffei, R.Corradini, R.Marchelli; *Tetrahedron Letters,* **2008***,* Vol.: 49(33), 4958-4961;

[22] Theodorou V., Ragoussis V., Strongilos A., Zelepos E., Eleftheriou A., Dimitriou M., *Tetrahedron Lett.* **2005**, *46,* 1357–1360;

[23] (a) G. Bram, *Tetrahedron Lett.*, **1973**, 469; (b) U. Ragnarsson, *Acta Chem. Scand.*,**1972**, *26*, 2550; (c)Pittelkow M., Lewinsky R., Christensen J. B., Org. Synth., 2007, 84, 209-214;

[24] M. F. Jacobsen, M. M. Knudsen, K. V. Gothelf, *J. Org. Chem.*, **2006**, *71*, 9183-9190;

[25]Corradini R., Sforza S., Dossena A., Palla G., Rocchi R., Filira F., Nastri F., Marchelli R., *J. Chem. Soc., Perkin Trans.*, **2001**, *20*, 2690-2696;

[26] J. Cai, X.Li, J. S. Taylor, *Org. Lett.*, **2005**, *7(5)*, 751-754.

PAGE LEFT INTENTIONALLY BLANK

CHAPTER 4

Synthesis and Properties of PNA Oligomers Based on Engineered Uracil Derivatives

4.1 Introduction

Chemical modification of Peptide Nucleic Acids (PNA) is being studied extensively as an approach for the development of improved probes for nucleic acids recognition[1] and for therapy based on gene regulation[2]. Due to commercial availability of protected monomers, the synthesis of unmodified PNA oligomers can be carried out by standard solid phase protocols for peptides, mainly using Boc/Cbz and Fmoc/Bhoc strategies[3] (*Figure 4.1*), coupling agents (such as HBTU or HATU) and polystyrene resin as solid support with methylbenzhydryl amine (MBHA) or Rink amide handles.

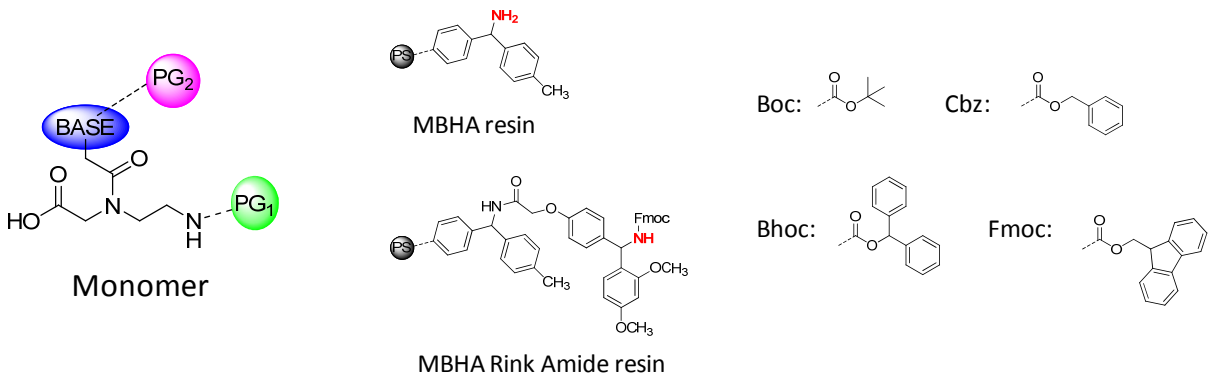


Figure 4.1. Protective groups used in Boc/Cbz and Fmoc/Bhoc strategies and resins for standard PNA oligomeric synthesis.

Table 4.1. Conditions for Boc/Cbz and Fmoc/Bhoc strategies for standard PNA oligomeric synthesis.

Strategy	Resin	PG_1	PG_2	Deprotection	Cleavage
Boc/Cbz	MBHA	Boc	Cbz	TFA	TFMSA
Fmoc/Bhoc	MBHA*	Fmoc	Bhoc	Piperidine	TFA

*MBHA rink amide

The synthesis of modified PNA oligomers often requires optimization and slightly modification of the standard procedures in order to ensure better yields and purity. Herein we present the synthesis and hybridization properties of PNA oligomers containing the modified uracil monomers (*Figure 4.2*) described in the previous *Chapter 3*. The synthesis was performed using different strategies: Boc/Cbz chemistry for *Mod-1*, *Mod-2* and *Mod-4*, and Fmoc/Bhoc chemistry for *Mod-3.*

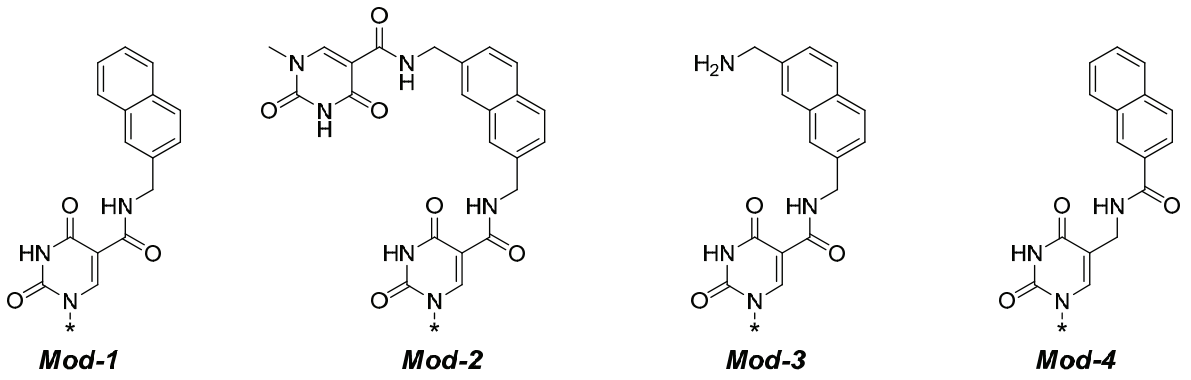


Figure 4.2. Modification incorporated into PNA oligomers employed in this study.

Table 4.2. Modified uracil PNA oligomers synthesized in this study. Sequence S1: H-GTAGA***X***CACT-gly-NH_2; sequence S2: H-TCCT***X***CACT-gly-NH_2; sequence S3: H-TCC***YX***CACT-gly-NH_2. ***X*** and ***Y*** indicate the position of modified monomers depicted in *Figure 4.2*.

Seq.\Mod.	Mod-1	Mod-2	Mod-3	Mod-4	Thymine
S1	PR211	PR210	PR198	-	PR5
S2	PR213	PR212	-	PRAAUN3	PR15
S3	***Y***=Mod-1, ***X***=Mod2 PR214		-	-	-

The sequences used in the present studies are reported in *Table 4.2* and correspond to a reference PNA sequence (S1) for which a large set of melting temperature data are available either for unmodified and modified PNA, and to a specific tract of the cystic fibrosis (CF) gene containing a point mutation (corresponding to the base indicated with X in *Table 4.2*) related to a pathological status. Only one (or, in S3 sequence, two) modified monomer was introduced and the properties of the corresponding PNA was compared with those of the unmodified PNA in order to evaluate the effect of structural modifications. The combinations of modifications and sequences considered are reported in *Table 4.2*.

4.2 Synthesis and Recognition Properties of PNA bearing C(5)-Carboxamido Modified Uracil Derivatives

Synthesis and preliminary study of the modified H-GTAGA*T*CACT-NH_2 (S1) sequence. The H-GTAGATCACT-NH_2 sequence[4] (S1) was chosen as a model to test potency and selectivity of novel uracil derivatives (*Figure 4.2*), since this has been thoroughly studied by *Nielsen and coworkers*[4a], by our group[4b] and by many research laboratories[4c] as model sequence in order to compare the effect of modification of the PNA structure on thermal stability. The synthesis of PNA oligomers containing *Mod-1* and *Mod-2* was performed by hybrid Boc/Fmoc chemistry. Unmodified commercially available Boc-(Z)PNA monomers were used, while modified monomers were Fmoc protected at the aminoethyl nitrogen. In *Figure 4.3* the synthesis of the *Mod-2* oligomer (PR210) is reported; the modified monomer was inserted in the middle of the sequence. The Fmoc-modified monomer *Mod-2* was coupled on the growing oligomer H-CACT-NH_2 using HBTU as activating agents (*step A*) and after Fmoc deprotection, the oligomer was elongated using Boc monomers (*step B and C*). Boc deprotection of the first PNA monomer inserted after the modified monomer (*Figure 4.3*, *step C*) induced also the removal of the Dod (4,4'-dimethoxybenzhydryl) group on the uracil derivative; therefore it was carried out using a deprotection mixture containing triisopropylsilane (TIS) in order to better scavenge carbocation formed in the Dod deprotection. The crude product was obtained by cleavage (*step D*). The same strategy

depicted below was employed for the synthesis of derivatives containing *Mod-1* (PR211). The synthesis of the oligomer containing *Mod-3* (PR198) was performed using the Fmoc chemistry, since in the modified PNA monomer the amino group linked to naphthalene was orthogonally Boc protected.

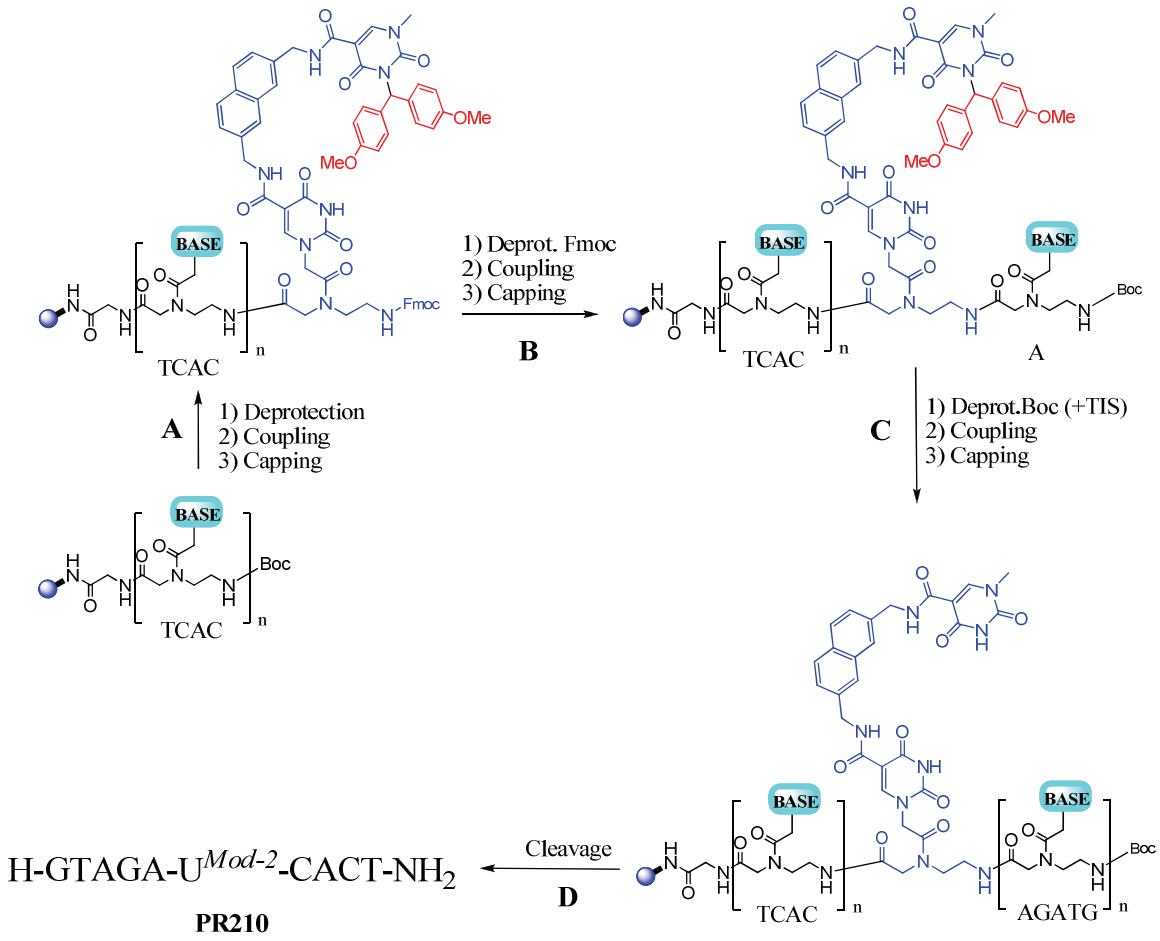


Figure 4.3. Mixed Fmoc/Boc strategy employed for the synthesis of PR210. Detailed reaction conditions are reported in the Experimental Section.

The synthesis of PR210 led to low yield (6% after purification); analysis of the reaction mixture by HPLC-MS revealed that truncated products were present, caused by incomplete coupling of the modified monomer, due to its high steric hindrance (*Figure 4.4*). Only one side product was fully identified as the capped truncated product. A second by-product, with a MW corresponding to 152 mass unit higher than that the truncated [H-CACT-(Gly)-NH_2] sequence, could be due to coupling of N(1)-methyl--5-carboxyuracil present in traces the

modified monomer or to transamidation transfer of pendant uracil from the activated monomer. No acetylated product was detected after the base insertion; this indicates that steric hindrance is critical only during the insertion of modified monomer, while it does not affect the subsequent monomer insertion.

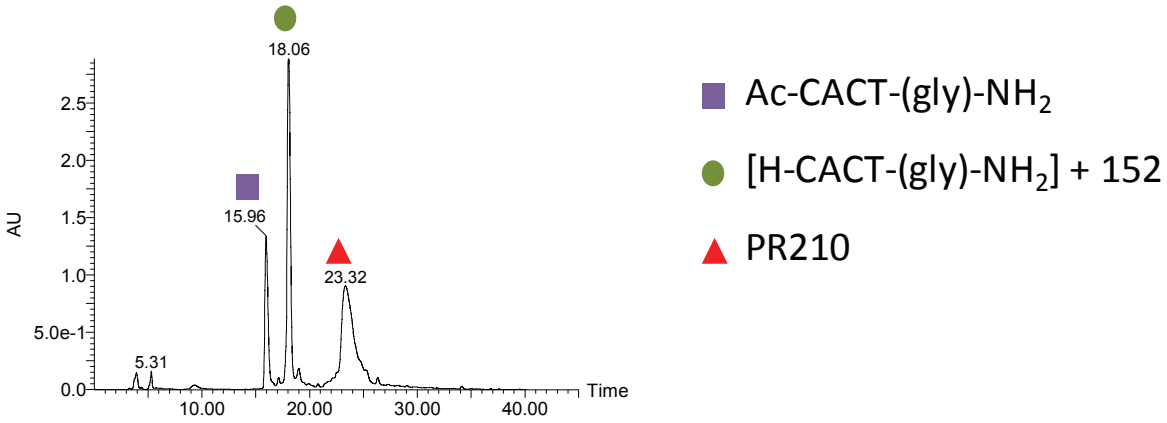


Figure 4.4. HPLC-UV (260 nm detector) profile of crude PR210.

In *Figure 4.5* HPLC-UV/MS profile and MS spectrum of pure PR210 are reported. The analysis of the reaction mixture of PR211 (H-GTAGAU$^{Mod\text{-}1}$CACT-Gly-NH_2) did not reveal truncated products; this can be due to the smaller steric hindrance of the monomer bearing only the naphthalene linker (*Figure 4.6*), and led to a higher yield of the purified product (12%).

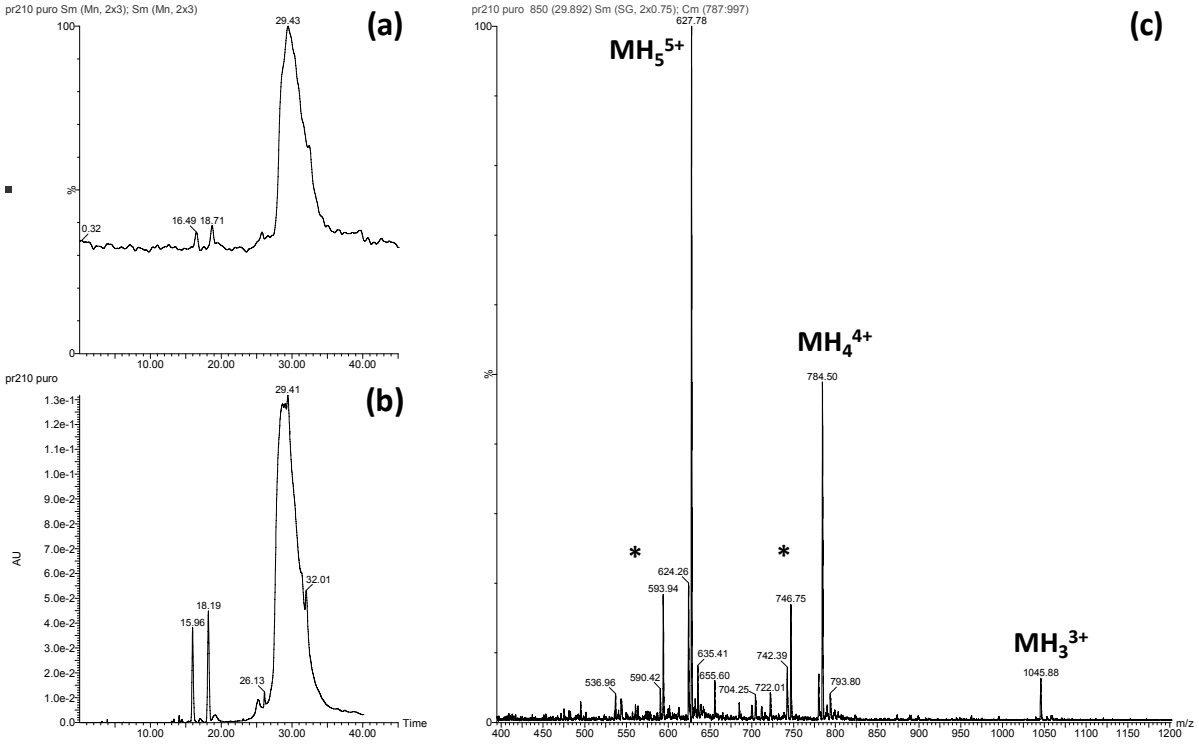


Figure 4.5. HPLC-UV-ESI-MS profile (*a* for MS and *b* for UV) and (c) MS-spectra at 29.4 min of pure PR210 (H-GTAGAU$^{Mod\text{-}2}$CACT-(gly)-NH_2). Conditions are reported in Experimental Section. Peaks marked (*) are multicharge fragments due to fragmentation of the naphthalene benzylic position of the modified nucleobase.

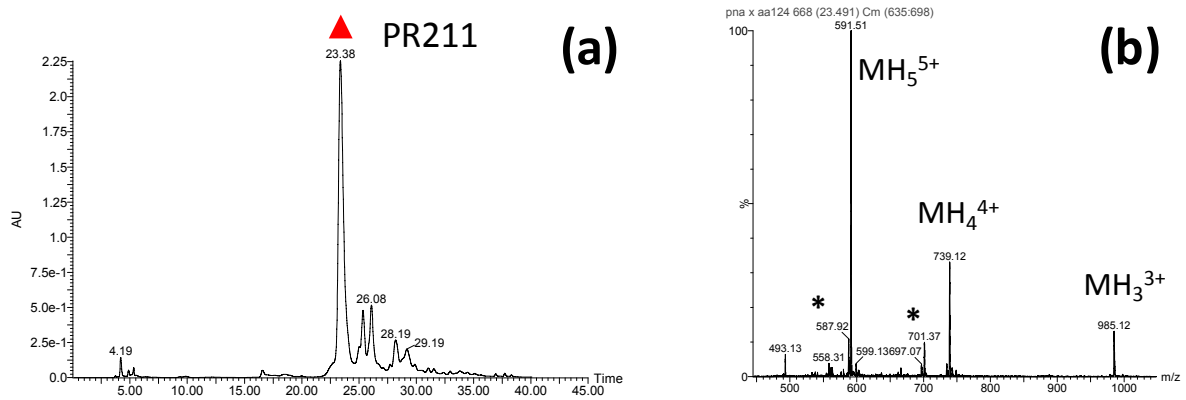


Figure 4.6. HPLC-UV profile (a) and ESI-MS spectrum (b) at 23.4 min of crude PR211. Peaks marked (*) are multicharge fragment due to fragmentation of the naphthalene benzylic position of modified nucleobase.

The stability of the PNA:DNA duplexes formed by the modified PNA oligomers PR210 and PR211 with the full-match DNA and single-mismatch DNA with guanine, replacing the central adenine facing the modified nucleobases, is reported in *Table 4.3*. The results were compared with those obtained using the corresponding unmodified PNA (PR5).

Table 4.3. Melting temperature of modified PNA oligomers. *Conditions:* [PNA]=5µM, [DNA]=5µM, buffer PBS pH 7, (1mM EDTA, 10mM phosphate and 100mM NaCl) Heating rate 1°C/min, monitored wavelength 260nm.

Entry	***H-GTAGA-X-CACT-gly-NH_2 3'-CATCT-Y-GTGA-5'***	***T_m(°C) Y=A Fullmatch***	***T_m(°C) Y=G Mismatch***	***ΔT_m(°C)***
PR5	X: *Thymine*	50	40	10
PR211	X: *Mod-1*	50	40	10
PR198	X:*Mod-3*	52	42	10
PR210	X: *Mod-2*	56	55	1

The insertion of *Mod-1* did not lead to any change of stability and selectivity respect to thymine. In *Mod-3* a positive charge on naphthalene led a small stability enhancement (2°C) with no selectivity improvements. *Mod-3* and *Mod-1* seemingly did not led to dramatic changes respect to unmodified thymine oligomers (PR5). Insertion of *Mod-2* apparently brought an enhancement of stability (fullmatch +6°C compared to thymine) with dramatic loss in selectivity (only +1°C). This anomalous behavior was therefore studied more in detail using circular dichroism. Since PR210 is achiral, while the PNA:DNA duplex is chiral, we could follow by CD only the transition of the PR210 PNA:DNA (fullmatch) duplex; no transition near 56°C was observed at the CD channel (*Figure 4.7*). This means that the melting profile observed by UV do not involve complementary DNA, and is therefore due to self-association. Melting at 42°C found by the CD experiments could be ascribable to partial duplex formation involving DNA and PR210.

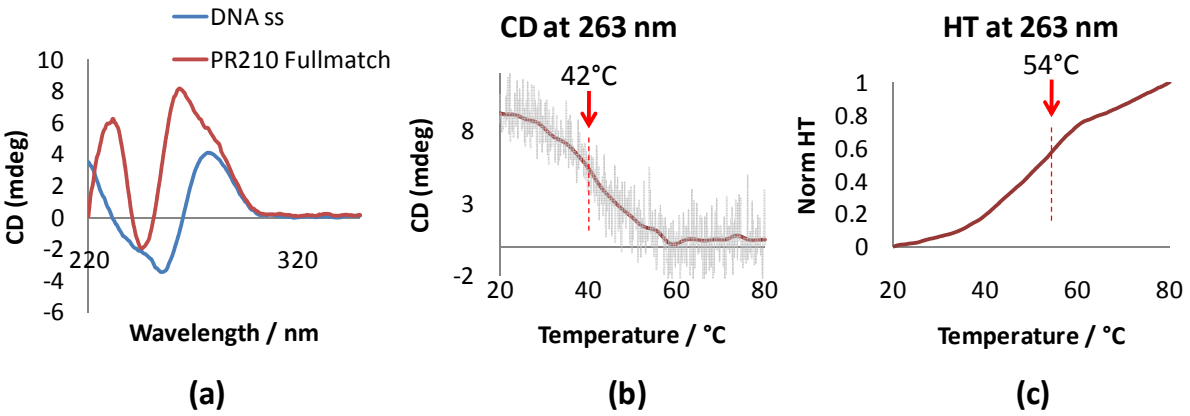


Figure 4.7. (a) CD spectra of PR210 duplex fullmatch and DNA alone. Melting curves were followed at 263nm. (b) CD curves showed an inflection point at 42°C, while (c) HT curves at 54°C. Conditions as showed in *Table 4.3*.

The self-associated dimer of PR210 has therefore a melting temperature higher than its fullmached duplex with DNA.

Self Aggregation of H-GTAGA*X*CACT-NH_2. In order to better understand the self-melting process, we measured the melting curves of the single-strand PNA. Melting curves of PR210 (*Figure 4.8a*) showed similar hypercromicity and melting temperature for fullmatch, mismatch and self melting, suggesting that in all three cases the same self-melting transition is observed.

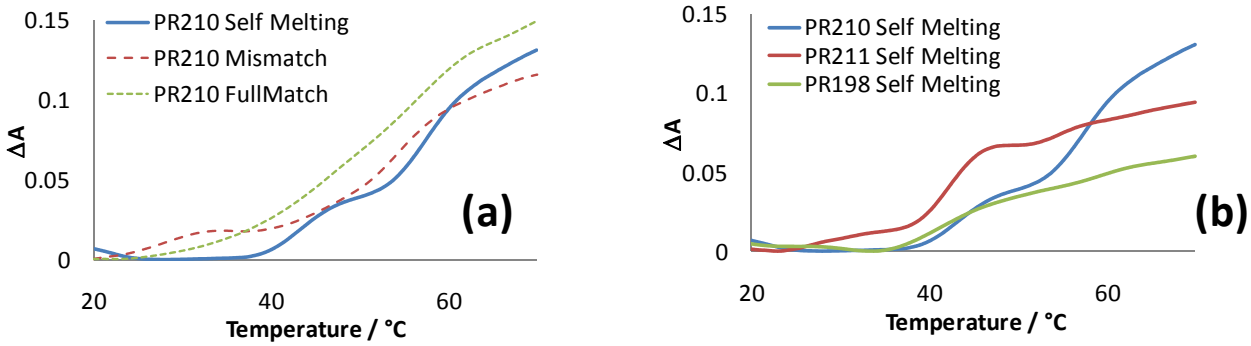


Figure 4.8. (a) Fullmatch (56°C), mismatch (55°C) and self melting (58°C) curves for PR210; (b) Self melting curves for PR210 (58°C), Pr211 (42°C) andPR198 (41°C). PNA was used at a concentration of 5μM in self melting experiments. Conditions as showed in *Table 4.3*.

Switching from PR210 to PR211 (*Figure 4.8b*) the self melting drops of 16°C, indicating a crucial role of pendant uracil of *Mod-2* (*Figure 4.9*). PR198 showed the same self melting value of PR211.

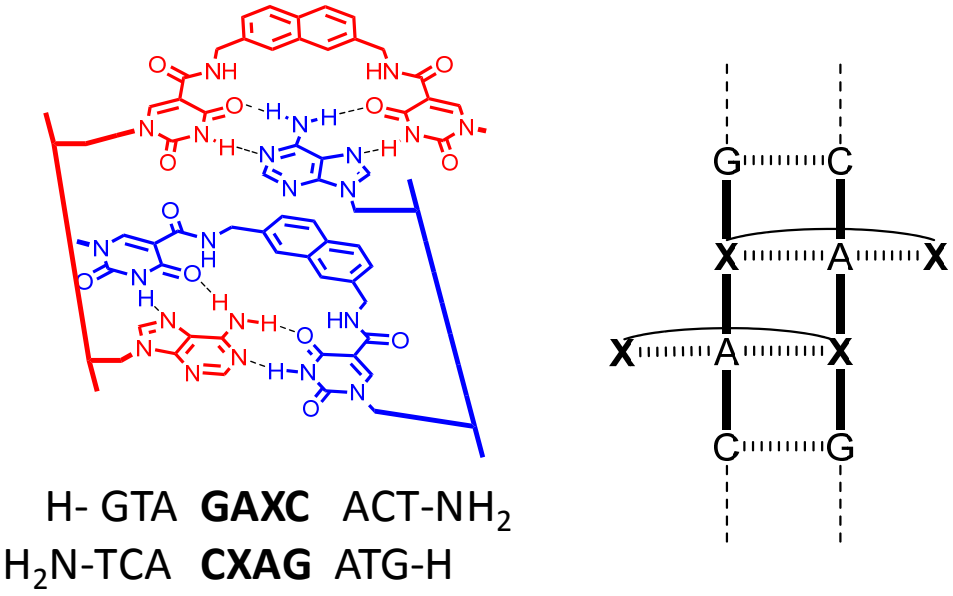


Figure 4.9. Schematic model of PR210 self aggregation. The standard sequence is self-complementary for the central portions GTAC. The stabilization of PR210 self aggregate is due to additional hydrogen bonding given by pendant uracils, which enforces the stacking interaction between the interstrand naphthyl groups.

All self melting curves of *Figure 4.8* showed more than one transition, probably due to further self pairing schemes of single strand PNA oligomers that we did not investigate.

Confirmation that self aggregation was occurring, was given by further experiments. Electrospray mass spectrometry was used to investigate PNA containing duplexes[5] by observing their multicharge ions by direct infusion of a solution containing the duplex in neutral buffer. Choosing appropriate conditions of potentials it is possible to observe supramolecular complexes that exist in solution. Self aggregation of the PR210 duplex was observed by ESI mass spectrometry (*Figure 4.10a*). The presence of the multicharge ion $(M_2H_5)^{5+}$ proved the existence of dimeric PR210.

The HPLC profile of the PNA PR210 was also affected by self-association; while the peak profile at 25°C was broad, by increasing the temperature to 45°C it became sharp (*Figure 4.10b*).

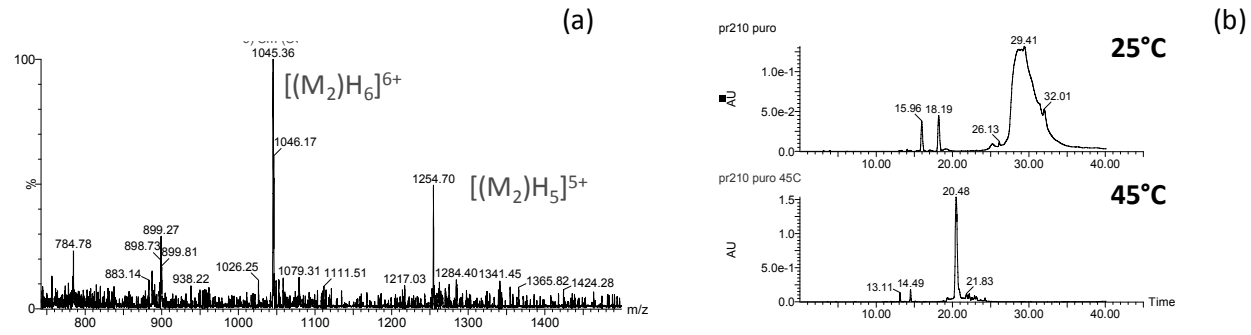


Figure 4.10. (a) Mass spectra of PR210 obtained for direct infusion. Peaks at 1254.7 m/z is diagnostic of the self aggregate duplex. *Conditions:* [PNA]=10μM, Buffer $HCOONH_4$ 10 mM at pH 7, ES+, Source Temperature 60°C, Desolvation temperature 100°C, Capillary Voltage 3.30 kV, Cone 40 V; (b) HPLC profile of PR210 at 25°C and 45°C. Conditions are the same for both, except the temperature, and are reported in the Experimental Section.

The very tight self-association of the PR210 PNA can easily be rationalized considering that the central tract containing the *Mod-2* modified monomer has a stretch of four self-complementary bases. This induces the formation of PNA:PNA duplexes, responsible for the self-association observed in all cases. This association is particularly strong in the case of PR210, since the formation of two adjacent $U^{Mod\text{-}1}$-A base pair is highly cooperative, due to the occurrence of Watson-Crick and Hoogsteen hydrogen bonding, as well as stacking interactions, which are not present in the other cases, as schematically described in *Figure 4.9*.

The self aggregation of PR210 does not allow formation of the PR210 duplex with DNA since the thermal stability of the self aggregate is higher than that of the complex with DNA. Hybridization of PR210 with complementary PNA gave a duplex more stable than self aggregate, having melting temperature of 62°C, which was 3°C less stable of the unmodified PNA:PNA antiparallel duplex with the same sequence[6]. The lower stability of the former can be attributed to the competing self-aggregation process.

Therefore, the model sequence GTAGATCACT does not allow to infer the role of the second uracil moiety in DNA binding. A different sequence was therefore chosen for this purpose.

Cystic Fibrosis-related sequence H-TCCT*X*CACT-NH_2. In order to perform a comparison of the PNA modification avoiding interference by self-association, the H-TCCT*X*CACT-NH_2 sequence was chosen (S2). This PNA -9mer sequence with *X*=T is complementary and able to recognize a single point mutation M-W1282X in the human cystic fibrosis (CFTR) gene recognized as one of the most common mutations leading to the cystic fibrosis disease[7]. Synthesis of PNA oligomers containing the *Mod-1* (PR213) and *Mod-2* (PR212) modifications was carried out using the same conditions employed for PR210 and PR211. In *Figure 4.11* the HPLC and MS-ESI spectra of the pure oligomers are reported.

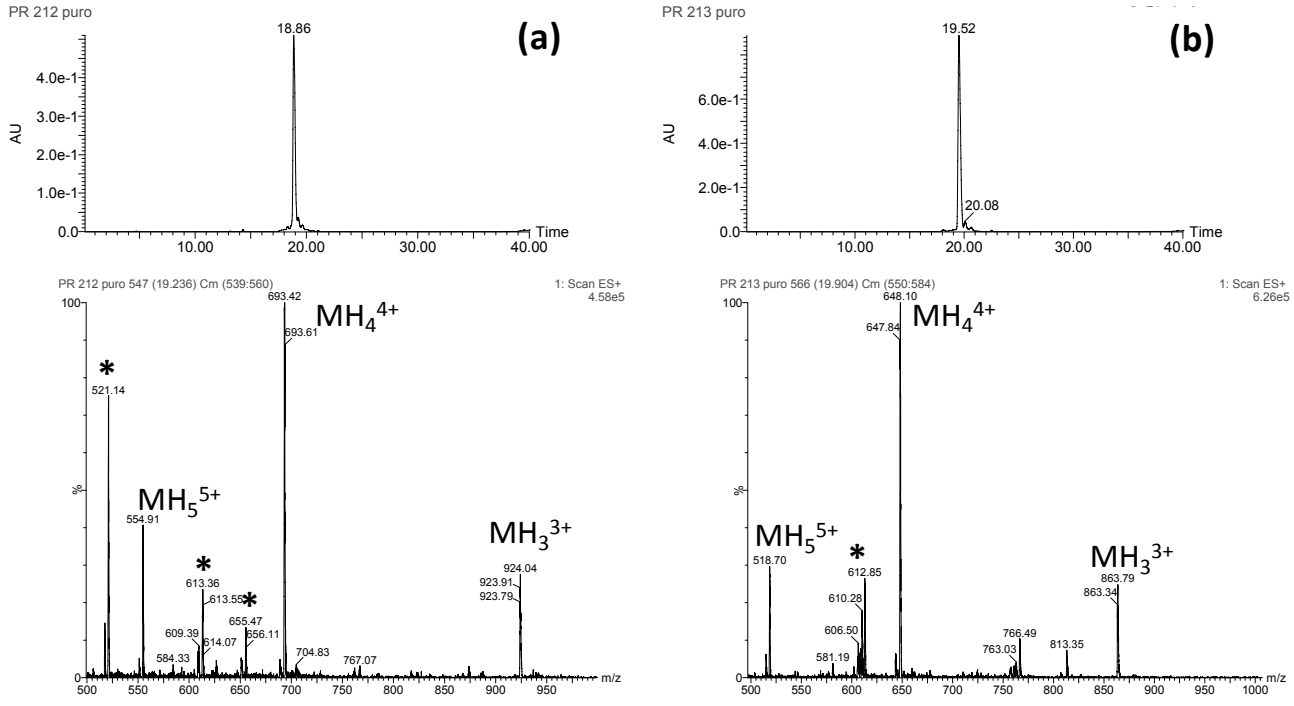


Figure 4.11. HPLC-UV profiles (above) and corresponding ESI-MS spectra (below) of pure PR212 (a) and PR213 (b). Peaks marked (*) are multicharge fragments due to breaking of naphthalene benzylic position of modified nucleobase. The yields of pure compounds were 37% for the PR213 and 8.3% for the PR212.

These new sequences bearing *Mod-1* and *Mod-2* did not show self association, as reported in *Figure 4.12*; hypercromicity of PR212 and PR213 is negligible (compared to PR210 and PR212) without showing any cooperative melting transition.

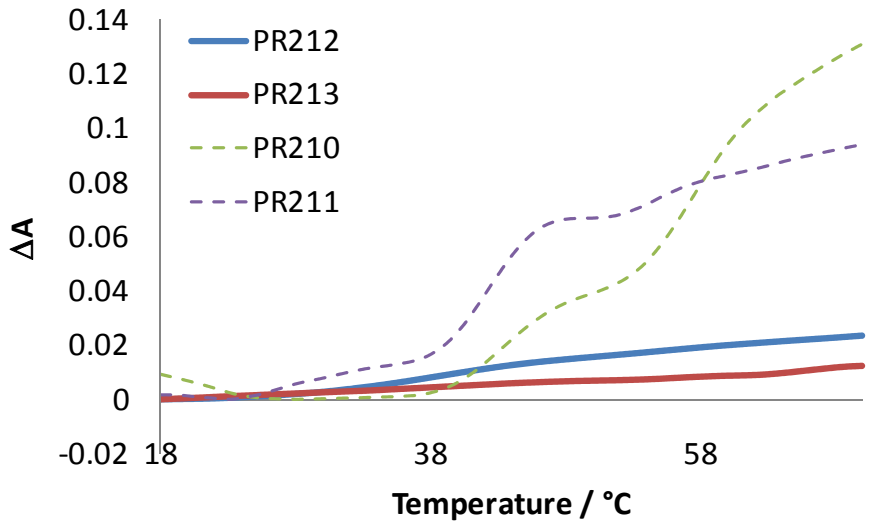


Figure 4.121. Self melting curves for PR212 and PR213 (continuum lines) compared to PR210 and PR211 (dashed lines). PNA concentration was 5μM. Conditions as showed in *Table 4.3*.

Melting curves involving new PNA oligomers were obtained by monitoring the UV signal at 260nm, and by circular dichroism (CD). T_m and ΔT_m data for UV, CD and HT channels are presented in *Table 4.4;* all three series of data showed good agreement in terms of ΔT_m. In *Figure 4.13* the CD spectra for PNA studied are reported.

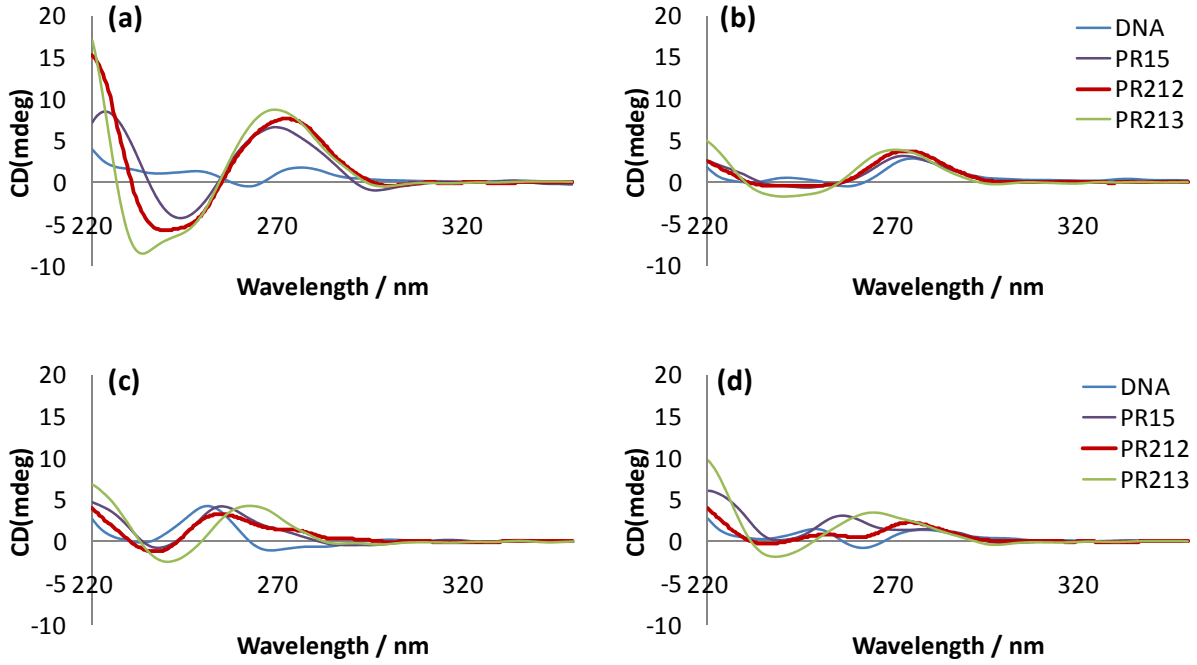


Figure 4.13. CD spectra of PNA:DNA duplexes studied in *Table 4.4*. The DNA sequence contained (a) Adenine, (b) Cytosine, (c) Gaunine and (d) Thymine facing the modified nucleobase. Spectra were recorded at 10°C, [PNA]=5µM, [DNA]=5µM, buffer PBS pH 7, (1mM EDTA, 10mM phosphate and 100mM NaCl), optical path 1cm.

Substitution of thymine by *Mod-1* led to slight decrease of the thermal stability (-1,5°C), suggesting a destabilizing effect of the carboxyamide and naphthalene moieties. Insertion of *Mod-2* (complete uracil dimer) increased the thermal stability of 2°C compared to unmodified PNA. From these data it is evident that the additional uracil moiety present in *Mod-2* is able to interact with Hoogsteen sites of adenine located on the complementary DNA strand, and by comparing of the *Mod-1* and *Mod-2* we can see that the additional interactions increase the melting temperature of 3.4°C. In order to evaluate the sequence-selectivity, we measured the ΔT_m between fullmatch and mismatch (G, C, T) DNA; substitution of thymine with *Mod-1* led to a decrease of selectivity; on the contrary, insertion of *Mod-2* showed a higher selectivity than the unmodified PNA 15 (these results are graphically summarized in *Figure 4.14*).

Table 4.4. Melting temperatures of the unmodified (PR 15) and modified PNA oligomers (PR 213 PR212). Melting curves were recorded monitoring the UV absorbance at 260 nm (1st series), CD (2nd series) and the HT-channel of the CD spectrum (3rd series) Melting temperatures are given by average of three experiments. *Conditions:* [PNA]=5µM, [DNA]=5µM, buffer PBS pH 7, (1mM EDTA, 10mM phosphate and 100mM NaCl). Heating rate 1°C/min from 10°C to 70°C.

Entry	***H-GTAGA-X-CACT-gly-NH_2 3'-CATCT-Y-GTGA-5'***		***T_m(°C) Y=A***	***T_m(°C) Y=G***	***ΔT_m***	***T_m(°C) Y=C***	***ΔT_m***	***T_m(°C) Y=T***	***ΔT_m***
PR15	*X: Thymine*	***UV***	**34.3±0.2**	**24.5±0.1**	**9.8**	**22.7±0.1**	**11.6**	**24.0±0.2**	**10.3**
		CD	36	26	10	22	14	25	11
		HT	36	26	10	24	12	25	11
		λ(nm)	*267*	*265*		*266*		*260*	
PR213	*X: Mod-1*	*UV*	**32.8±0.2**	**26.8±0.2**	**6.0**	**23.2±0.4**	**9.6**	**24.7±0.3**	**8.1**
		CD	35	26	9	27	8	27	8
		HT	34	27	7	26	9	26	8
		λ(nm)	*266*	*266*		*266*		*264*	
PR212	*X:Mod-2*	*UV*	**36.2±0.2**	**23.9±0.3**	**12.3**	**20.8±0.8**	**15.4**	**21.9±0.2**	**14.3**
		CD	36	21	15	22	14	24	12
		HT	36	24	12	23	13	23	13
		λ(nm)	*267*	*268*		*266*		*266*	

Note: the wavelength was chosen in order to gain maximum sensibility and it was determined by comparison of CD spectra of DNA and PNA:DNA duplex.

Selectivity of PR212 is more clear by observation of melting curves and of their first derivatives (*Figure 4.15)*; for PR212 the sigmoidal shape of the melting curve is less defined and its first derivative more broad, indicating a poor cooperative transition, probably due to a not well formed duplex structure with complementary DNA. Also CD spectra showed a difference between PR212 and PR213 for mismatch guanine (*Figure 4.13c*).

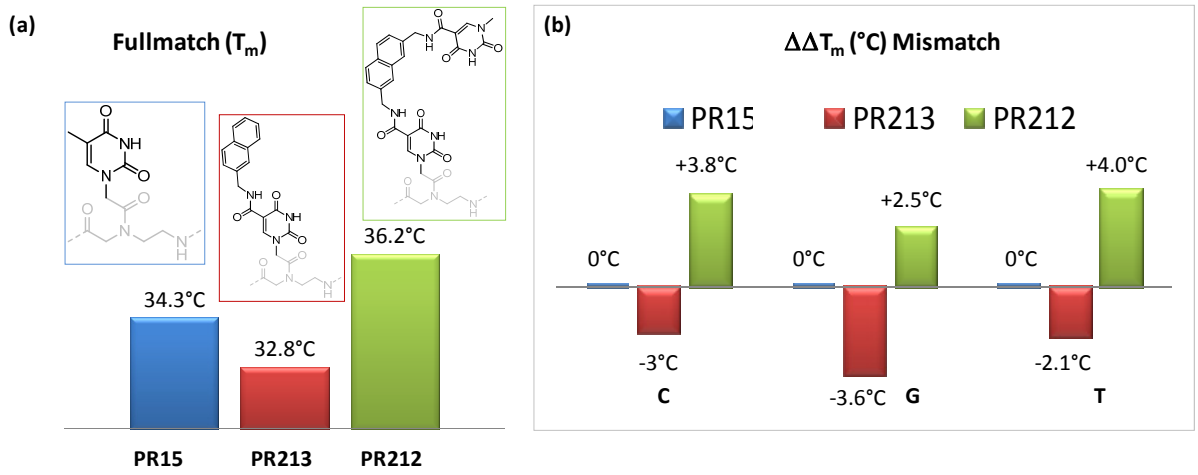


Figure 4.14. Graphical summary of *Mod-2* (PR213) and *Mod-1* (PR212) effects. *Mod-2* increases potency and selectivity. (a)Fullmatch T_m (°C); (b) Selectivity respect unmodified PR15 is expressed as $\Delta\Delta T_m$(°C) that was calculated as the difference between T_m (PR15) - T_m (PR213 or PR212) for each mismatch.

Focusing on PR212 and PR213, we can conclude that interactions of the pendant uracil in *Mod-2* gives binding interactions in the case of adenine, while it only induces destabilizing effects when one of the three mismatch (G, C, T) bases are placed in present in the place of adenine probably due to steric hindrance. This result strongly supports the formation of both Watson-Crick and Hoogsteen-hydrogen bonding for the modified dimeric uracil moiety of PR212.

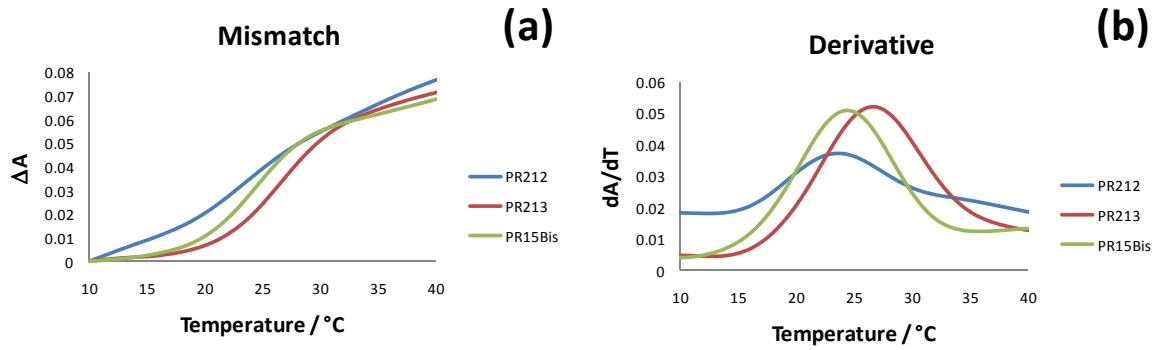


Figure 4.15. Melting curves of PR212, PR213 and PR15 (a) and first derivatives (b) relative of PNA:DNA with a guanine mismatch.

PR212 and PR213 duplexes with DNA were further investigated using exciton coupling method with the cyanine dye $Disc_2(5)$[8] (*Figure 4.16*).

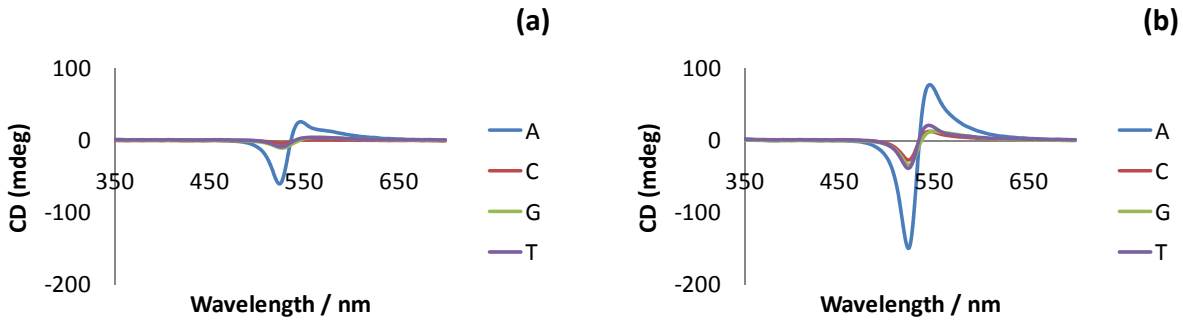


Figure 4.16 Induced CD spectra in the presence of PNA:DNA duplexes and $Disc_2(5)$. (a) PR212 and (b) PR213. Spectra were recorded at 20°C, [Disc2(5)]=25µM, [PNA]=5µM, [DNA]=5µM, buffer PBS pH 7, (1mM EDTA, 10mM phosphate and 100mM NaCl), optical path 2mm).

It is interesting to note that the PR212:DNA fullmatch duplex showed a signal three fold lower than that of PR213, though the CD signal at 220-280 nm is comparable for both PR213:DNA and PR212:DNA duplexes. Since the cyanine binding site has been proved to be within the minor groove and modifications can affect the intensity of the induced CD signal,[9] these data suggest that the strength of the *Mod-2* interaction with complementary adenine could introduce some distortion in the PNA:DNA duplex, thus making the minor groove less accessible to the dye.

PNA bearing two modified monomers. Since the naphthalene linker on uracil was able to stabilize partially self-complementary PNA oligomers, probably due to naphthalene – naphthalene stacking interactions, we designed a new PNA probe where the central position was occupied by *Mod-2* and the adjacent position by *Mod-1* PNA monomers. Our purpose was to create additional favourable energy contributions for the *Mod-2* monomer by stacking interaction with the neighbouring naphthalene moiety of *Mod-1,* in order to improve recognition abilities towards complementary DNA.

The synthesis of a new PNA (PR214) was therefore carried out using the same strategies described above; this product was obtained in low yield (2.8% purified material), due to difficulties in the coupling step of *Mod-1* hindered by the presence of *Mod-2*. The structure and the relative HPLC-UV and ESI-MS of PR214 are reported in *Figure 4.17*.

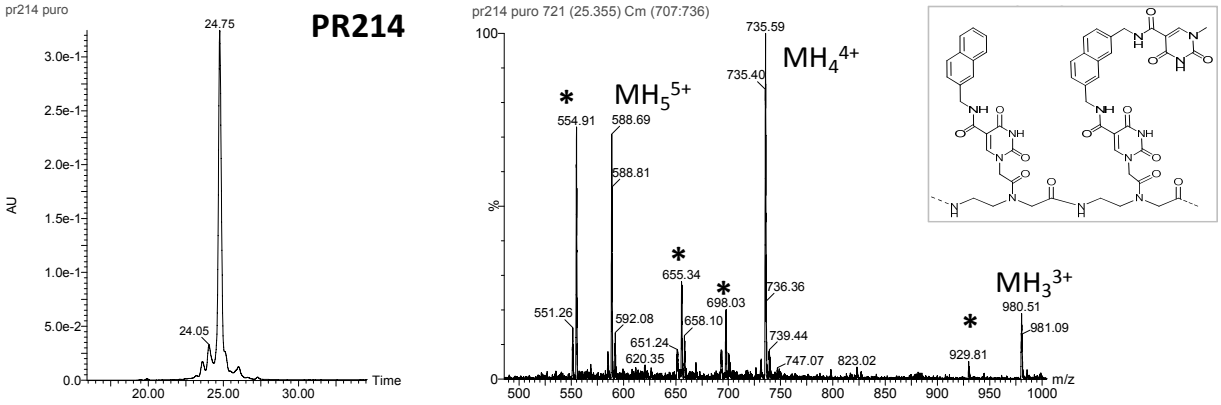


Figure 4.17. HPLC-UV profiles and ESI-MS spectra of PR214 after purification. Peaks marked (*) are multicharge fragments due to fragmentation of the naphthalene benzylic position of both modified nucleobases.

Recognition abilities of PR 214 were then tested toward complementary and mismatched DNA (*Table 4.5*). CD spectra for fullmatch and guanine-bearing mismatch were also recorded (*Figure 4.18*). Although the spectra obtained for mismatches are not easily explained for its unusual shape, we can hypothesize by comparing CD spectra of PR214:DNA_G with DNA alone that this duplex is partially formed.

Table 4.5. Melting temperature of PR214. For further details see *Table 4.4*.

Entry	*H-GTAGA-X-CACT-gly-NH$_2$* *3'-CATC-Z-Y-GTGA-5'*		*T_m(°C)* *Y=A*	*T_m(°C)* *Y=G*	*ΔT_m*	*T_m(°C)* *Y=C*	*ΔT_m*	*T_m(°C)* *Y=T*	*ΔT_m*
PR114	*X: Mod-2* *Z: Mod-1*	***UV***	**37.3±0.2**	**26.7±1**	**10.6**	**n.a.**	**-**	**n.a.**	**-**
		CD	38	29	9	27	11	29	9
		HT	38	29	9	26	12	30	8
		λ(nm)	*275*	*270*		*273*		*278*	

The fullmatch PNA:DNA duplex resulted to be slightly more stable (+1 °C), but with worse selectivity than PR212 (*Mod-2*) and PR15 (unmodified).

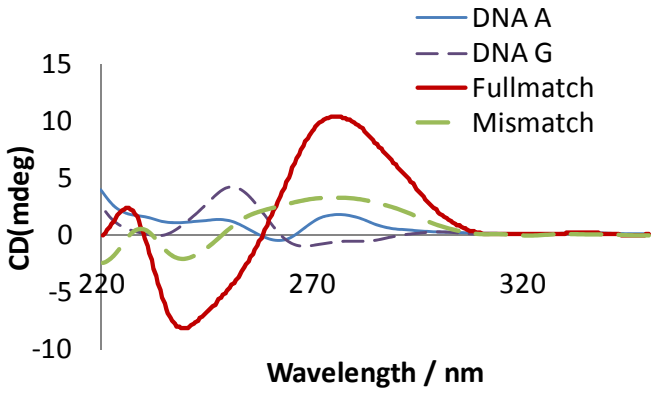


Figure 4.18. CD spectra of PR214 fullmatch (A) and mismatch (G). For further details see *Figure 4.13*.

4.3 Synthesis of PNA with a reversed Amide: C(5) Methylamido Uracil derivatives

The PNA oligomer PR212 containing *Mod-2* showed interesting recognition properties towards complementary DNA. The amide group directly linked to the C(5) position of uracil (*Figure 4.19a*) was tolerated and did not lead to a dramatic drops in stability of the PNA:DNA duplex. However, we explored the possibility to use derivatives with a substituent showing electronic effects more similar to those of thymine by simply reversing the amide (*Mod-4*) on the C(5) position of uracil (*Figure 4.19b*).

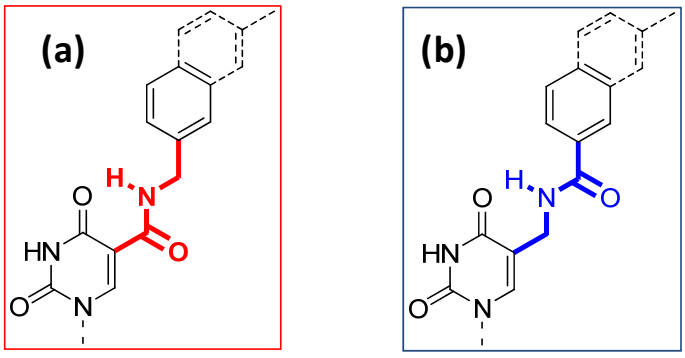


Figure 4.19.- (a)Carboxamido derivatives; (b) Reversed Amide, methylamido derivatives.

PNA oligomers containing *Mod-4* were synthesized using an innovative strategy by means of an azide modified precursor monomer of uracil which could be further modified by solid phase synthesis. This strategy can allow faster and simpler synthesis of oligomers, since it avoids the necessity to synthesize PNA monomers containing hindered and complex modified bases. Furthermore, the modification of the base on solid phase could avoid problems of steric hindrance in coupling steps, since the modified base is built step by step. The 5-(azidomethyl)uracil-PNA-monomer (*Figure 4.20*) was chosen as the common precursor monomer.

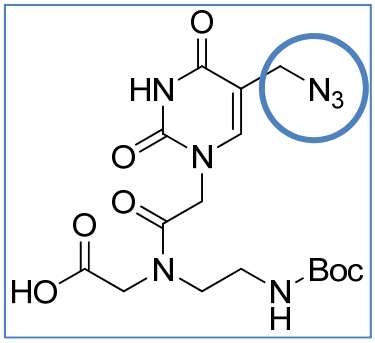


Figure 4.20. 5-(azidomethyl)-PNA monomer. The azide group was chosen as hidden and protected amino group.

The synthesis of 5-(azidomethyl)-PNA monomer is reported in *Chapter 3*. The azide group was chosen because it is easily transformed into an amino group under mild conditions, such as those used in Staudinger's reduction[10], which can be performed also on solid phase; the amino group thus obtained can be used in reactions with activated carboxylic acids to form a new amide bond on the C(5) position of uracil. To test the validity of this strategy, we synthesized the PNA oligomer PNA AAUN3 with sequence S2 (*Table 4.1*) since S1 was found to be prone to self-association. The synthesis is reported in *Figure 4.21*. The synthesis of the modified base (*Mod-4*) on solid phase was carried out in three steps (*Figure 21*, *step b, c and d*). During *step b* the azidomethyl-PNA monomer was attached to the growing chain of PNA (in this step DIC/DhBtOH was found to be a more effective coupling agent than HBTU), then the azide was reduced by trimethylphosphine (*step c*) and the resulting amine was by reacted with activated 2-naptalene carboxylic acid (*step d*). For sake of simplicity in steps *c* and *d* of *Figure 4.21* we reported that the azido group was reduced to amino group and then coupled with the activated carboxylic acids. Actually in literature it is reported that also iminophosphoranes (derived from the Staudinger reaction of azide with phosphine) are able to react with activated carboxylic acids[11]. Also in our case the mechanism likely involved reaction of the trimethyl azaylide with the activated carboxylic acid, since the analogous reduction with tributyl phosphine failed, probably due to steric hindrance in the tributyl azaylide intermediate. In *step e* the synthesis of the remaining part of PNA and cleavage from the resin were carried out and afforded the crude PNA AAUN3 (*Figure 4.22a*) in fairly good yield (12%) after purification (*Figure 4.22b*).

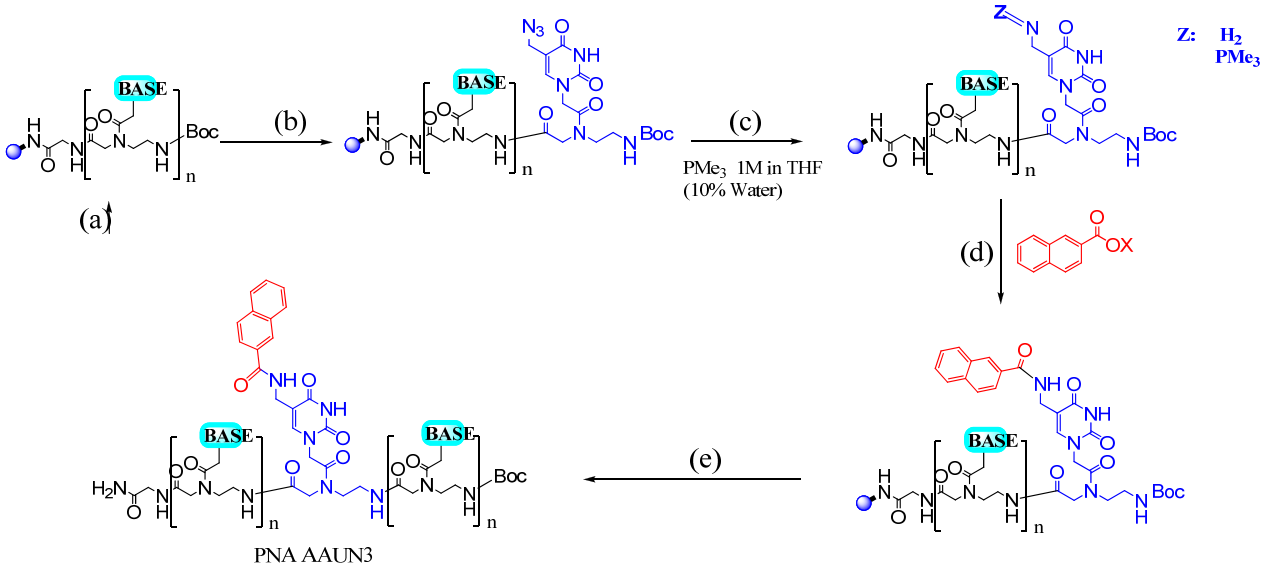


Figure 4.21. Synthesis of PNA AAUN3 (H-TCCTU$^{Mod\text{-}4}$CACT-gly-NH_2). (a)Oligomeric synthesis sequence H-CACT-gly-NH_2; (b)deprotection of the Boc group and coupling of the azide modified monomer using diisopropyl carbodiimide (DIC) / 3-Hydroxy-1,2,3-benzotriazin-4(3*H*)-on**e** (DhBtOH) as activating agents; (c) azide reduction by trimethyl phosphine; (d) coupling of naphthalene-2-carboxilic acids activated by diisopropyl carbodiimide (DIC) / 3-Hydroxy-1,2,3-benzotriazin-4(3*H*)-on**e** (DhBtOH), followed by capping with Ac_2O; (e) cleavage from the resin with TFMSA/TFA. For further details see the Experimental Section.

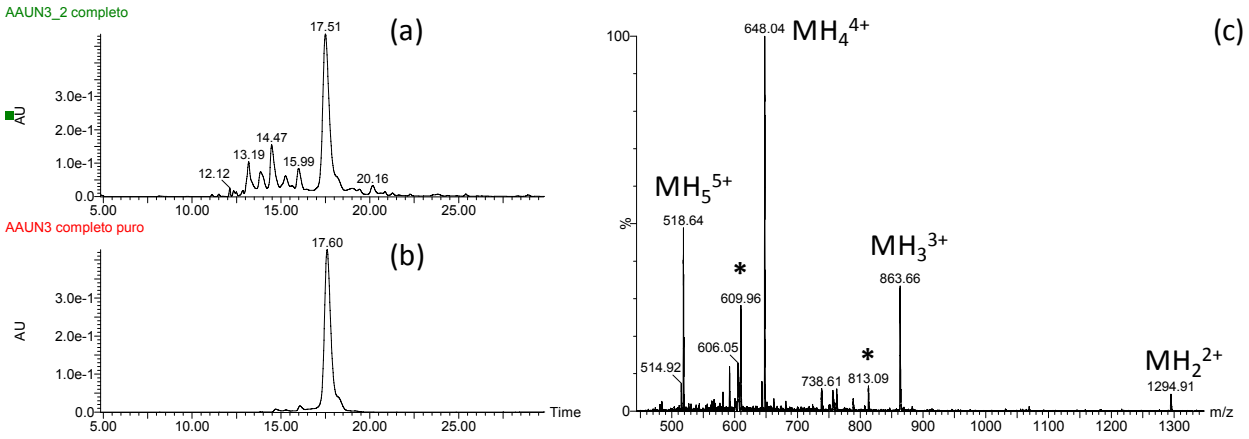


Figure 4.22. HPLC-UV spectra of crude PR AAUN3 (a), after purification (b) and ESI-MS spectra (c) of the relative PR AAUN3 peak. Peaks marked (*) are multicharge fragments due to fragmentation at the benzylic position C(5) position of uracil.

The recognition abilities of PR AAUN3 have been evaluated by measuring the melting temperatures with complementary and mismatched DNA, as showed in *Table 4.6* . CD spectra were also measures and are reported in *Figure 4.23*.

Table 4.6. Melting temperatures of PR AAUN3 duplexes. For further details see *Table 4.4.*

Entry	***H-GTAGA-X-CACT-gly-NH$_2$ 3'-CATCT-Y-GTGA-5'***		***T_m(°C) Y=A***	***T_m(°C) Y=G***	***ΔT_m***	***T_m(°C) Y=C***	***ΔT_m***	***T_m(°C) Y=T***	***ΔT_m***
PR--AAUN3	*X: Mod-4*	***UV***	**28.8**	**20.1**	**8.7**	**17.8**	**11**	**19.1**	**9.7**
		CD	33	19	14	19	14	20	13
		HT	30	22	8	18	12	20	10
		λ(nm)	*267*	*267*		*275*		*263*	

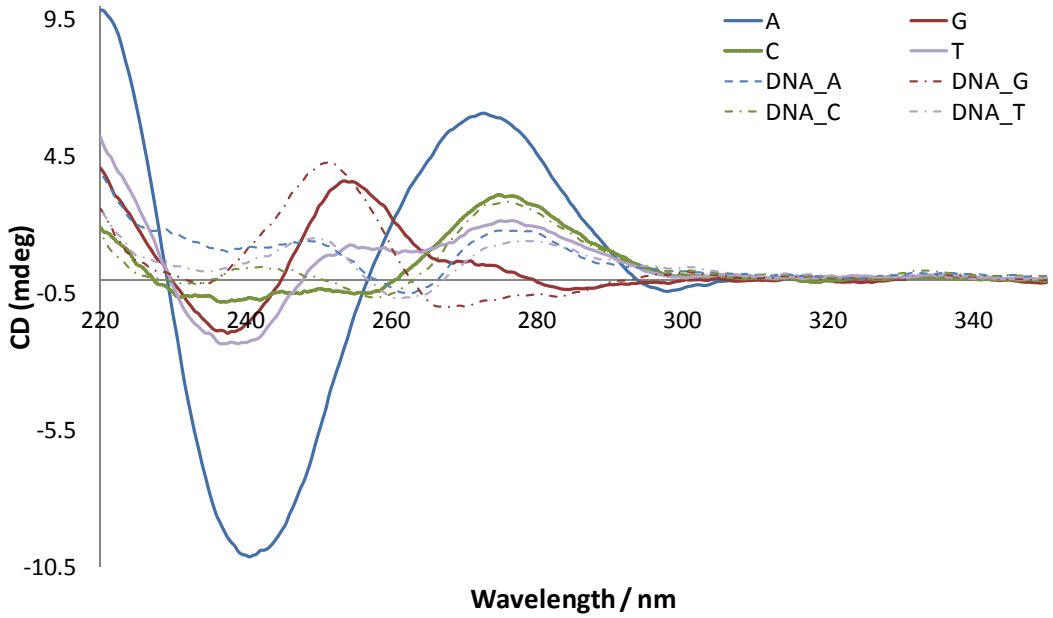


Figure 4.23. CD spectra of PR AAUN3 for fullmatch and mismatched duplexes. For further details see *Figure 4.13.*

CD spectra of *Figure 4.23* showed that at 20°C only fullmatch duplex is formed.

In *Figure 4.24* the results obtained with *Mod-4* containing PNA (PR AAUN3) are compared with data obtained with unmodified (PR15) and the *Mod-1* containing PR213 which differs only in the direction of the amide bond connecting the naphthalene moiety to uracil.

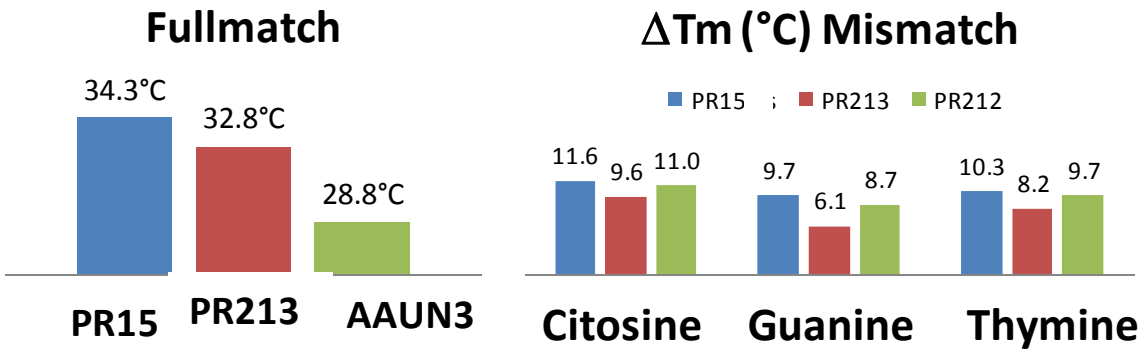


Figure 4.24. Comparison of PR AAUN3 data with PR15 (thymine) and PR213 (*Mod-1*).

Reverse amide (*Mod-4*), as clearly showed in *Figure 4.24*, is detrimental on fullmatch duplex stability. Reversing amide of *Mod-1* in *Mod-4* led to a drop of thermal stability of about 4°C; but on the other hand, the selectivity resulted higher than that of *Mod-1*, and similar to the unmodified PNA PR15.

Fluorescence studies. 2-napthalene carboxylic acid show weak fluorescence and this property is retained when it is linked to PNA through an amide bond (PNA AAUN3, *Mod-4*). Upon excitation at 278 nm, a fluorescence emission centered at 368 nm was observed for single stranded PR AAUN3; this signal was found to be sensitive to the hybridization state of PNA (*Figure 4.25*).

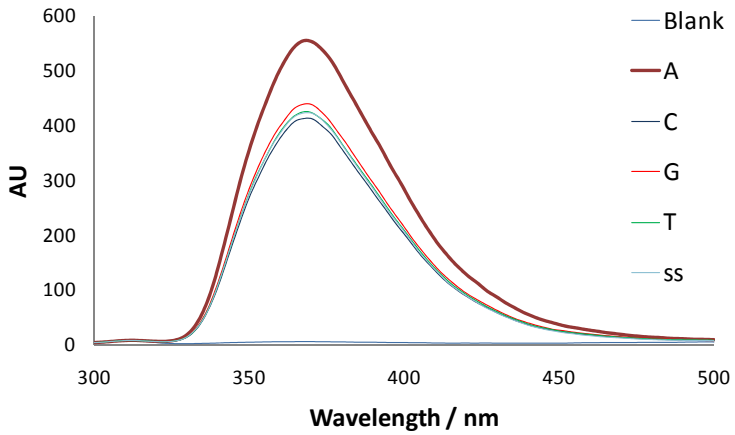


Figure 4.25. Fluorescence of PR AAUN3 fullmatch, mismatch (G, C, T) and single strand (ss). Excitation at 278 nm, T=20°C. [PNA]=5µM, [DNA]=5µM, buffer PBS pH 7, (1mM EDTA, 10mM phosphate and 100mM NaCl).

Fluorescence of PR AAUN3 is about 25% higher in fullmatch duplex, while it was almost the same for single strand and mismatched duplexes. This behavior was found to be quite similar to oligonucleotide probes bearing a pyrene fluorophore on the C(5) position of uracil[12] and can be used in the selective sensing of full-match DNA in the presence of singly mismatched sequences.

4.4 Conclusion and Perspectives

The overall trend of stability and selectivity of modified PNA oligomers studied are presented graphically in *Figure 4.26*.

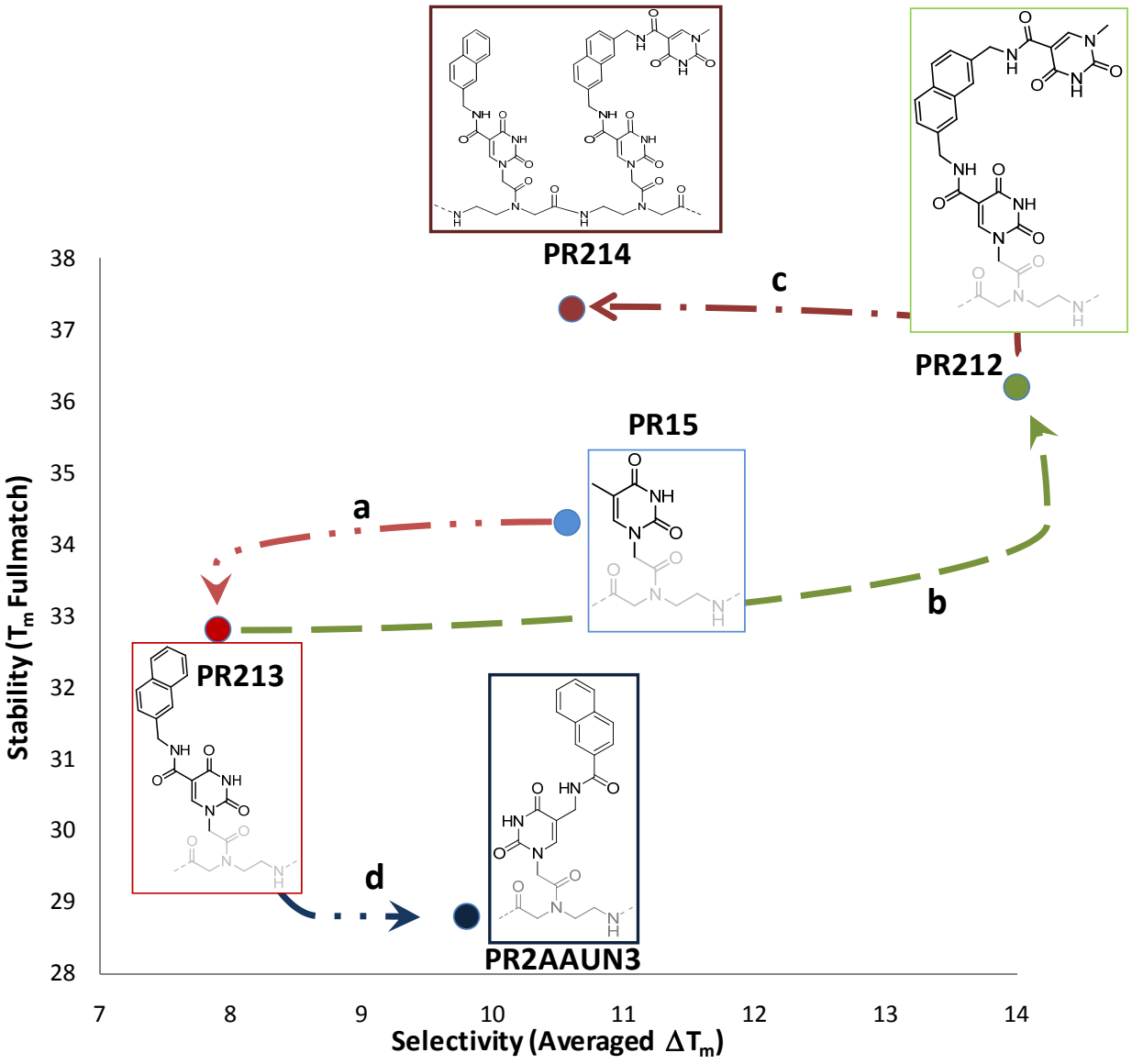


Figure 4.26. Schematic illustration of modified PNA oligomer properties. Stability is defined as melting temperature of the fullmatch duplex (Adenine) with DNA, while selectivity is calculated as the average of ΔT_m for mismatched duplexes (Citosine, Guanine and Thymine).

When replacing methyl group of thymine (PR15), with 5-carboxamidomethylnapthalene (PR213) we observed a drop of stability and selectivity (*Figure 4.26*, *path a*); this negative gap is restored with *Mod-2* in PR212, with even improved properties superior to thymine (*Figure 4.26, path b*). Therefore, the present results clearly indicate that it is possible to perform recognition of the adenine through both Watson-Crick and Hoogsteen hydrogen bonds and that the spacer used, rationally designed, is sufficiently long and rigid to allow to mimic the TAT triplet pattern; this has the effect of increasing both affinity and selectivity of the PNA itself. However, self-association should be carefully considered when using this type of modification.

These findings are not only important for improving of the stability and selectivity of PNA oligomers, but it is important since these engineered PNAs were specifically designed in order to precisely interact with specific sites on complementary DNA. This model of "tailor made" modified base (*Figure 4.27*) could allow us to further modify the "Hoogsteeen" uracil moiety with functional groups, for examples with moieties able to perform hydrolytic cleavage of the phosphate esters of DNA or RNA, thus producing a sequence-specific artificial nuclease.

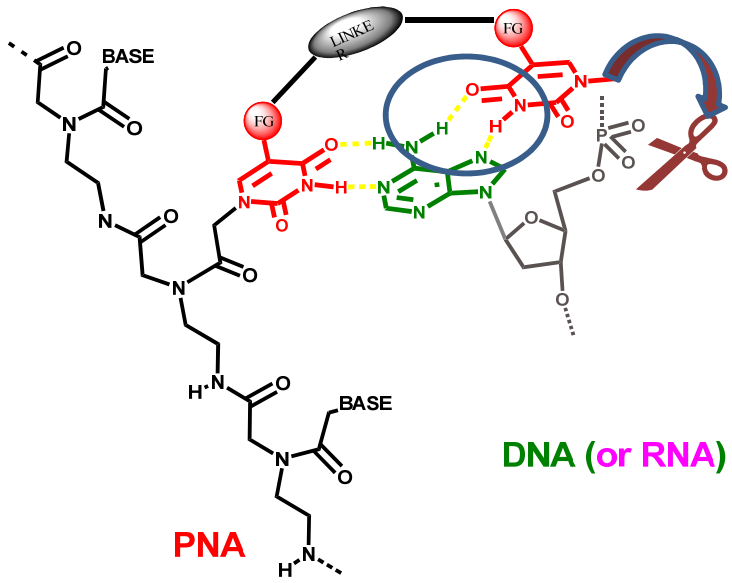


Figure 4.27. Schematic model of potential applications of tailor a made nucleobase.

Inversion of the amide group of *Mod-1* to *Mod-4* resulted to be detrimental for stability, but not for selectivity; however, in this case the different fluorescence response of naphthalene in PNA PR AAUN3 with full-match and mismatched DNA opens the possibility to develop fluorescent probes for DNA recognition. The possibility to perform fast synthesis of PNA libraries bearing different fluorophores, using intermediate azide derivatives, could be interesting for applications in sensing and diagnostics.[13] This strategy can also be used for the introduction of other more efficient fluorophores.

4.5 Experimental Part

General Information. Reagents and solvents were purchased from Sigma-Aldrich, Fluka, Applied Biosystem, ASM research, Carlo Erba and used without further purification. N,N-dimethylformamide (DMF) for solid phase synthesis was dried over 4Å molecular sieves before use.

PNA Oligomers Synthesis. PNA oliogmers were synthesized by Fmoc (only used for PR198) or Boc strategy and modifications of standard protocol were reported. The synthesis was performed manually using plastic vessel equipped with polyethylene frit, cap and valve. MBHA and rink amide MBHA resins were respectively used for Boc and Fmoc strategy; resins were downloaded before use to 0.2 mmol/g of active sites using Fmoc-gly-OH or first monomer; unreacted amino group were capped by 1/1 mixture of DMF/Ac_2O. Exact resin loading was determined by measuring UV absorbance at 290 nm of small amounts of resin dispersed in a mixture 1/4 piperidine/DMF. Loading expressed in mmol/g was given from following formula $1.68/Abs(290nm) \times mg(resin)$. **Fmoc Strategy**. For each monomer elongation we employed the following module: a) Deprotection Fmoc in piperidine/DMF 20/80 for 8 min 2 times. b) Monomer activation with 5 eq (respect resin active sites) of PNA-monomer, 4.5 eq HBTU, 10eq DIEA for 2 min. c) Coupling 45 min. d) Capping with DMF/DIEA/Ac_2O 89/6/5 1 min for 2 times. e) DIEA 10% washings 2 times for 2 min. After completion the PNA was cleaved in TFA/m-cresol 95/5 2 times for 1 h, then dried on nitrogen flow, precipitated with ethyl ether and washed three times. **Boc Strategy**. For each monomer elongation we employed the following module: a) Deprotection Boc in TFA/m-cresol for 4 min 2 times. b) Monomer activation with 5 eq (respect resin active sites) of PNA-monomer, 4.5 eq HBTU, 10eq DIEA for 2 min. c) Coupling 45 min. d) Capping with DMF/Pyridine/Ac_2O 25/25/1 1 min for 2 times. e) Piperidine 10% washings 2 times for 2 min. After completion the PNA was cleaved in TFA/TFMSA/Tioanisole/m-cresol 6/2/1/1 1 h for 2 times, then dried on nitrogen flow, precipitated with ethyl ether and washed three times. **Oligomers purification.** Crude PNA oligomers were purified by reversed phase with UV detection at 260nm. Semipreparative column C18 (5μm, 250 x 10 mm, Jupiter Phenomenex, 300 Å) was utilized and oligomers purified by gradient eluition using water + 0.1% TFA (solvent A) and acetonitrile + 0.1% TFA (solvent B), Flow 4 mL/min. For each oligomers is specified retention time and gradient used for purification. Resulting pure PNA

oligomers were characterized by RP-HPLC-UV-ESI-MS, Conditions: Phenomenex C_{18} 250 mm length, I.D. 5 mm, particle size 25 µM, Gradient eluition 0-5 min 100% Water (0.2 %HCOOH), 5-35 min to 60/40 Water (0.2% HCOOH) – MeCN (0.2% HCOOH).For each PNA listed below were reported gradient eluitions conditions expressed as variation of solvent A and solventB.

PR 198 (H-GTAGA-U^{Mod-3}CACT-NH$_2$). Synthesized by Fmoc chemistry. Modified Fmoc-U^{Mod-3}-OH monomer was inserted using the same coupling condition used for standard monomer. Purification: 0-5 min 100%A, 5-35 min to 60%A/40%B, rt 20.2 min. MS (ESI+) m/z for $C_{120}H_{146}N_{60}O_{31}$ (2924.9): $[MH_4]^{4+}$ Calcd. 732.2 Found 732.7, $[MH_5]^{5+}$ Calcd. 586.0 Found 586.1, $[MH_6]^{6+}$ Calcd. 485.8 Found 485.8.

PR 210 (H-GTAGA-U^{Mod-2}CACT-gly-NH$_2$). Non –modified sequence was synthesized by Boc chemistry. Activation of modified Fmoc-U^{Mod-2}-OH was carried out by stirring(3.3 eq) of monomer with HBTU (3.1eq) and DIEA (6.6 eq) for 2 min, coupled with growing chain for 1h and followed by capping (DMF/DIEA/Ac_2O 89/6/5 1 min for 2 times). After Fmoc deprotection of modified nucleobase (piperidine/DMF 20/80 for 8 min 2 times), the oligomer was completed using standard condition for Boc synthesis as described above, except for Boc deprotection (was added 5% of triisopropylsilane at standard TFA/m-cresol 95/5 mixture). Purification: 0-5 min 100%A, 5-35 min to 60%A/40%B, rt 22.5 min. (Yield: 5.9%). MS (ESI+) m/z for $C_{128}H_{153}N_{63}O_{35}$ (3134.1): $[MH_3]^{3+}$ Calcd. 1045.7 Found 1045.8, $[MH_4]^{4+}$ Calcd. 784.5 Found 784.3, $[MH_5]^{5+}$ Calcd. 627.8 Found 627.7.

PR 211 (H-GTAGA-U^{Mod-1}CACT-gly-NH$_2$). Synthesis as described for PR210. (Yield: 12.1%). Purification: 0-5 min 90%A/10%B, 5-35 min to 60%A/40%B, rt 17.3 min. MS (ESI+) m/z for $C_{121}H_{146}N_{60}O_{32}$ (2952.9): $[MH_5]^{5+}$ Calcd. 591.6 Found 591.6, $[MH_4]^{4+}$ Calcd. 739.2 Found 739.2, $[MH_5]^{5+}$ Calcd. 985.3 Found 985.3.

PR 212 (H-TCCT-U^{Mod-2}CACT-gly-NH$_2$). Synthesis as described for PR210. Purification: 0-5 min 95%A/5%B, 5-35 min to 60%A/40%B, rt 19.8 min. (Yield: 8.3%). MS (ESI+) m/z for $C_{115}H_{141}N_{49}O_{35}$ (2769.7): $[MH_2]^{2+}$ Calcd. 1385.9 Found 1385.9, $[MH_3]^{3+}$ Calcd. 924.2 Found 924.0, $[MH_4]^{4+}$ Calcd. 693.4 Found 693.4, $[MH_5]^{5+}$ Calcd. 554.9 Found 554.9.

PR 213 (H-TCCT-U^{Mod-1}CACT-gly-NH$_2$). Synthesis as described for PR210. Purification: 0-5 min 90%A/10%B, 5-25 min to 70%A/30%B, rt 18.5 min. (Yield: 36.9%). MS (ESI+) m/z for

$C_{108}H_{134}N_{46}O_{32}$ (2588.6): $[MH_2]^{2+}$ Calcd. 1295.3 Found 1295.2, $[MH_3]^{3+}$ Calcd. 863.7 Found 863.8, $[MH_4]^{4+}$ Calcd. 648.2 Found 648.1, $[MH_5]^{5+}$ Calcd. 518.7 Found 518.7.

PR 214 (H-TCC-U$^{Mod\text{-}1}$U$^{Mod\text{-}2}$CACT-gly-NH$_2$). Synthesis as described for PR210. (Yield: 2.8%). Purification: 0-5 min 95%A/5%B, 5-35 min to 76%A/24%B, rt 26.6 min. MS (ESI+) m/z for $C_{126}H_{148}N_{50}O_{36}$ (2938.9): $[MH_2]^{2+}$ Calcd. 1470.5 Found 1471.0, $[MH_3]^{3+}$ Calcd. 980.6 Found 980.5, $[MH_4]^{4+}$ Calcd. 735.7 Found 735.6, $[MH_5]^{5+}$ Calcd. 588.8 Found 588.7.

PR AAUN3 (H-TCCT-U$^{Mod\text{-}4}$CACT-gly-NH$_2$). Non –modified sequence was synthesized by Boc chemistry. *Insertion of PNA modification.* 5-(methylazido)-PNA monomer (22 mg, 0.050 mmol) was activated by DIC (0.068 mL, 0.045mmol) and DhBtOH (8.2mg, 0.050mmol) in 0.8 mL of DMF for 15 min. Coupling of activated monomer was carried out overnigth. Azide moieties was reduced by using 1M of PMe_3 + 10% of water, under nitrogen, for 90 min and repeated two times. 2-napthoic acid (11mg. 0.060mmpl) was activated by DIC(0.0077 mL, 0.050 mmol) and DhBtOH (9.8 mg, 0.060 mmol) for 15 minute and added to reduced azide. After overnigth coupling of 2-naphtoic acid, resin was capped by Ac_2O and washed many times with THF, DCM and DMF. Synthesis of remaining part of oligomers was done as usually by using Boc chemistry. Purification: 0-5 min 95%A/5%B, 5-25 min to 72%A/28%B, rt 19.4 min. (Yield: 12.0%). MS (ESI+) m/z for $C_{108}H_{134}N_{46}O_{32}$ (2588.6): $[MH_2]^{2+}$ Calcd. 1295.3 Found 1294.9, $[MH_3]^{3+}$ Calcd. 863.7 Found 863.5, $[MH_4]^{4+}$ Calcd. 648.2 Found 647.8, $[MH_5]^{5+}$ Calcd. 518.7 Found 518.6.

PR 15 (H-TCCT-T-CACT-gly-NH$_2$). Purification: 0-5 min 95%A/5%B, 5-20 min to 76%A/24%B, rt 18.5 min. Synthesized by standard boc conditions. (Yield: 25 %). MS (ESI+) m/z for $C_{97}H_{127}N_{45}O_{31}$ (2419.4): $[MH_2]^{2+}$ Calcd. 1210.7 Found 1210.4, $[MH_3]^{3+}$ Calcd. 807.5 Found 807.3, $[MH_4]^{4+}$ Calcd. 605.9 Found 605.7, $[MH_5]^{5+}$ Calcd. 484.9 Found 484.8.

Melting experiments. DNA oligonucleotides (HPLC grade) were purchased from Thermo electron and used without further purification. Concentration of stock PNA solutions were determined by UV measurement at 260 nm using the following extinction coefficients ($M^{-1}cm^{-1}$): PR198 (112100), PR210 (120700), PR211 (112100), PR5 (74500), PR212 (83100), PR213 (74500) and PRAAUN3 (74500). Sample were prepared dissolving an equal amounts of stock solutions of PNA and DNA in appropriate buffer and incubated for 5 min at 90°C prior to measure. Thermal denaturation profiles (Abs vs T) for duplex were recorded using

UV/Vis Lambda Bio 20 Spectrophotometer equipped with a Peltier temperature controlling system. Heating rate was set to 1°C/min. Melting curves were smoothed using Savitsky Golay algorithm. T_m values were obtained from minima of first derivative of curves.

CD melting experiments. Sample were prepared as described for UV melting experiment. Spectra and Curves were recorded using JASCO J710 Spectropolarimeter equipped with Peltier temperature control. Wavelength was chosen in order to gain maximum sensibility comparing spectra of duplex with DNA alone.

Fluorescence Experiments. Fluorescence measurement were performed on Luminescence Spectrometer LS 55. Sample were prepared as described above for UV and CD studies.

4.6 References

[1] Corradini R., Sforza S., Tedeschi T., Marchelli R., *Chirality*, **2007**, *19(4)*, pp 269-294;

[2] Tonelli R, Purgato S, Camerin C, Fronza R, Bologna F, Alboresi S, Franzoni M, Corradini R, Sforza S, Faccini A, Shohet JM, Marchelli R, Pession A, *Mol. Canc. Ther.*, 2005, 4(5) 779 786;

[3] Peptide Nucleic Acids: Synthesis and Protocol, **2004**, Nielsen P. E., Horizon Bioscience;

[4] Model sequence H-GTAGATCACT-NH_2 is generally recognized for PNA as a model sequence to test chemical modification. Herein some representative examples about modified PNA being this sequence are reported. (a) Sforza S., Haaima G., Marchelli R., Nielsen P. E., *Eur. J. Org. Chem*, **1999**, 197, and related papers; (b) Sforza S., Corradini R., Ghirardi, S., Dossena, A., Marchelli R., *Eur. J . Org. Chem.*, **2000**, 2905; (c) See *Chapt.1*;

[5] Sforza, S.; Tedeschi, T.; Corradini, R.; Dossena, A.; Marchelli, R., *Chem Comm.,* **2003**, 1102-1103;

[6] Wittung P., Eriksson M., Lyng R., Nielsen P. E., Norden B., *J. Am. Chem. Soc.*, **1995**, *117(41)*, 10167;

[7] Corradini R., Feriotto G., Sforza S., Marchelli R., Gambari R., *J. Mol. Rec.*, **2004**, *17*, 76;

[8] See *Chapter 1* for further details;

[9] Dilek I., Madrid M., Singh R., Urrea C.P., Armitage B.A.., *J. Am. Chem. Soc.*, **2005**, *127(10)*, 3339-3345;

[10] Staudinger H., Meyer J., *Helv. Chim. Acta,* **1919**, *2*, 635;

[11] (a) Debaene F, Winssinger N, *Org. Lett.*, **2003**, *23(5)*, 4445-4447; (b) Serwa R., Wilkening I., Del Signore G., Muehlberg M., Claussnitzer I., Weise C., Gerrits M, Hackenberger C., *Angew. Chem. Int. Ed.*,**2009**, *48(44)*, 8234-8239;

[12] A. Okamotoa, Y. Saito, I. Saito, J. *Photochem. and Photobiol. C: Photochem. Rev*., **2005**,. *6*, 108–122;

[13] Kaihatsu K., Janowski B. A., Corey D. R., *Chem. & Biol.*, **2004**, *11*, 749–758.

CHAPTER 5

Porphyrins – PNAs Conjugates: Tools to Study and to Control Helicity in PNA Containing Duplexes

5.1 Introduction

Natural porphyrin complexes are very important biological molecules. Tetrapyrrolic macrocycles are involved in a wide number of biological processes ranging from oxygen transport to photosynthesis and catalysis. For example iron porphyrin residues contribute to oxygen transport and storage, electron transport in cytochromes and oxygen activation in cytocrome P450. All of these processes are possible since a protein matrix surround active porphyrin catalytic centre. Nature uses a wide range of biosynthetic pathways to build and assemble this complex huge structures synthetically impossible to be produced in a laboratory. In order to overcome this problem chemists have studied synthetic porphyrins[1] and their complexes in order to mimic natural function and to create new structures able to perform functions such as molecular binding[2], reaction catalysis[3], energy and electron transfers[4], and light harvesting[5]. Assembly of porphyrin subunits can be made using covalent chemistry[1] or more attractive noncovalent synthesis[6]. Porphyrins can be used as supramolecular synthons since they can be modified with variety of auxiliary moieties, such as H-bonding motifs or ligands with predefined geometries for the synthesis of self-assembled architectures. Moreover porphyrins show stability under a wide range of temperature and pH, and show interesting optical properties easily tuned by inserting exocyclic substituents and by choice of metal ion chelated in the centre.

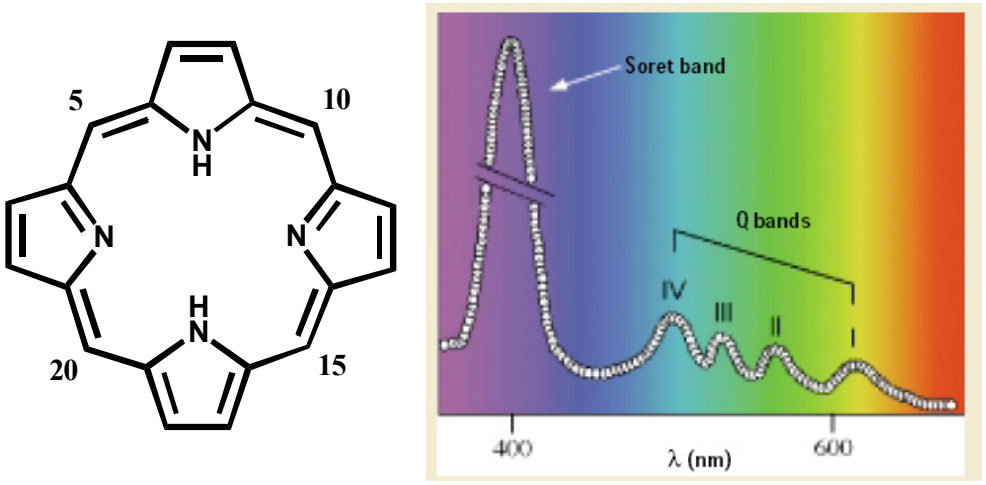


Figure 5.1. Porphyrin macrocycle (*meso* position are numbered) and its electronic absorption spectra. Porphyrins have very strong absorption bands around 400 nm (Soret band) with absorptivities on the order of 10^5 M^{-1} cm^{-1} and several Q-band between 500 and 650 nm with ε 10-20 times less.

In the literature numerous examples of supramolecular functional porphyrin architectures are reported [7]. For example H-bonding moieties directly attached to the *meso* positions (*Figure 5.1*) allow formation of rigid structures with different discrete geometries or 2D and 3D network. Electrostatic interactions (between cationic and anionic tetra-arylporphyrins) that hold the porphyrins aggregated was exploited by *Purrello* and *coworkers* to create chiral switchable memories by pH tuning of the size of porphyrin aggregates . These aggregates are directed by undetectable persistent chiral seed[8]. The unique properties of the porphyrin chromophore can also be utilized in circular dichroism (CD) spectroscopy due to its very large electric transition moment which can give rise to exciton coupling effects over a long distance. Reporter groups based on porphyrins[9] are very efficient in structural studies and in absolute configuration assignment[10]. Dimeric zinc porphyrin tweezers that bind derivatized chiral amines and alcohols have been employed for the determination of absolute stereochemistry, based on the induced exciton-coupled bisegnate CD spectra which is due to preferred porphyrin helicity (generated by the stereogenic centre of the substrate[11]). Covalently linked porphyrins are also suitable to probe chirality in large molecules like cholesterol and steroids[12]. The above mentioned examples show how wide is nowadays the application of porphyrins, moreover porphyrin conjugates can be very useful in the study of nucleic acid structure and chemistry.

Porphyrins (either covalently linked to DNA or not) have been used as tools for structural studies of DNA[13] and DNA has been used as scaffold to organize porphyrins at nanoscale[14]. *Balaz et al*[15] reported a cationic zinc porphyrin as chiroptical probe for Z-DNA exploiting different coordination ability of guanine N^7 in B- or Z- DNA toward zinc porphyrin. Different organization of anionic porphyrins around B- or Z- DNA mediated by spermine have been employed to organize porphyrins on nanoscale in order to obtain functional devices as logic gates[16]. Covalently linked porphyrins on 5'- position of an oligonucleotide have been used for the detection of salt induced B- to Z- transition by monitoring the Soret band circular dichroism exciton couplet signal at 420 nm[17]. *Stultz et al*[18] used double stranded B-DNA as a scaffold to build zipper porphyrin arrays; in this case the porphyrins were linked on the C(5) position of uracil to create an helical stacked array, which gave rise to energy transfer processes and duplex stabilization.

Peptide Nucleic Acids are widely used in DNA recognition for their unique properties. The interest toward PNA is not only focused in DNA recognition, but also in the possibility to form homo-duplexes in which the helix handedness can be controlled by insertion of opportune chiral centers[19]. While the structural features of antiparallel PNA duplexes have been extensively studied, parallel duplexes are still not fully characterized. In this chapter we explored the possibility to use covalently attached porphyrins to peptide nucleic acids as reporter to study supramolecular chirality.

Part of this work was carried out under the supervision and in collaboration with Professor Nina Berova and Dr. Ana Petrovic at *Columbia University*, New York.

5.2 Parallel Duplexes of PNA-Porphyrin Conjugates: Helicity and Helical Control

The use of nucleic acids in bottom-up approach allow building of new 2D or 3D nanoscopically controlled system[20]. Hybridization of complementary nucleic acids tethered to other components provides a specific pathway to assemble different component on nanoscale or acting as template to place reactive moieties in space proximity[21]. PNA were also efficiently incorporated into a DNA double crossover to create 2D hybrid DNA/PNA nanostructure. The different number of base pair per turn of PNA:DNA (antiparallel) double strand lead to a different 2D structure[22]. Although not extensively explored, the possibility of substitution of DNA with PNA can also improve stability and durability of nanostructures, since duplex containing PNAs (PNA:PNA and PNA:DNA) are more stable than ds DNA, are less affected by environmental factor like ionic strength, and are not subjected to chemical or enzymatic degradation.

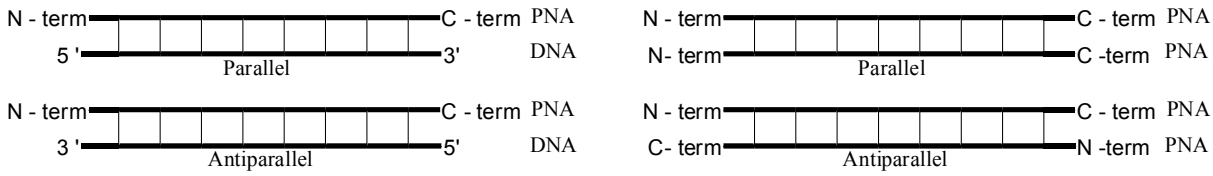


Figure 5.2. Parallel and Antiparallel orientation in PNAs containing duplexes.

While antiparallel PNA containing helices are well known in literature, for parallel PNA duplexes only few data are available[23]. Parallel duplexes are less stable than antiparallel. For example on a standard PNA 10-mer sequence H-GTAGATCACT-NH_2 switching from antiparallel PNA-PNA double strand to parallel double strand there is a drop of 18°C in thermal stability. The induction of helicity in antiparallel PNA:PNA duplexes can be obtained using a terminal aminoacid as "seeding unit" by incorporation of stereogenic centers in the PNA backbone. Both these approaches have been extensively studied and a rationale on the effect of stereochemistry on PNA:PNA helix handedness have been established.

However, for parallel PNA:PNA duplexes, the helicity and the conformational properties are far from being understood, mainly because of the lack of extensive experimental data and modeling studies. However, by modification of PNA at N-terminus, in antiparallel PNA:PNA the two N-terminal moieties would result on the opposite ends of duplex, while in the parallel they are on the same side. Therefore, tethering molecular object at N-term on PNAs and assembling in a parallel PNA:PNA duplex allows close positioning of this linked object (*Figure 5.2*).

General project. In order to better understand the behaviour of PNA parallel duplexes and their structure we explored the possibility to use covalently attached porphyrins to the N-terminus of PNA as CD reporter of supramolecular chirality (*Figure 5.3*). Furthermore, helicity and helix handedness of parallel duplexes could be in principle either probed or regulated by the interaction of the conjugated moieties. L-Lysine was chosen as terminal amino acids since it is known to be an efficient chiral inducers in antiparallel PNA:PNA duplexes[24].

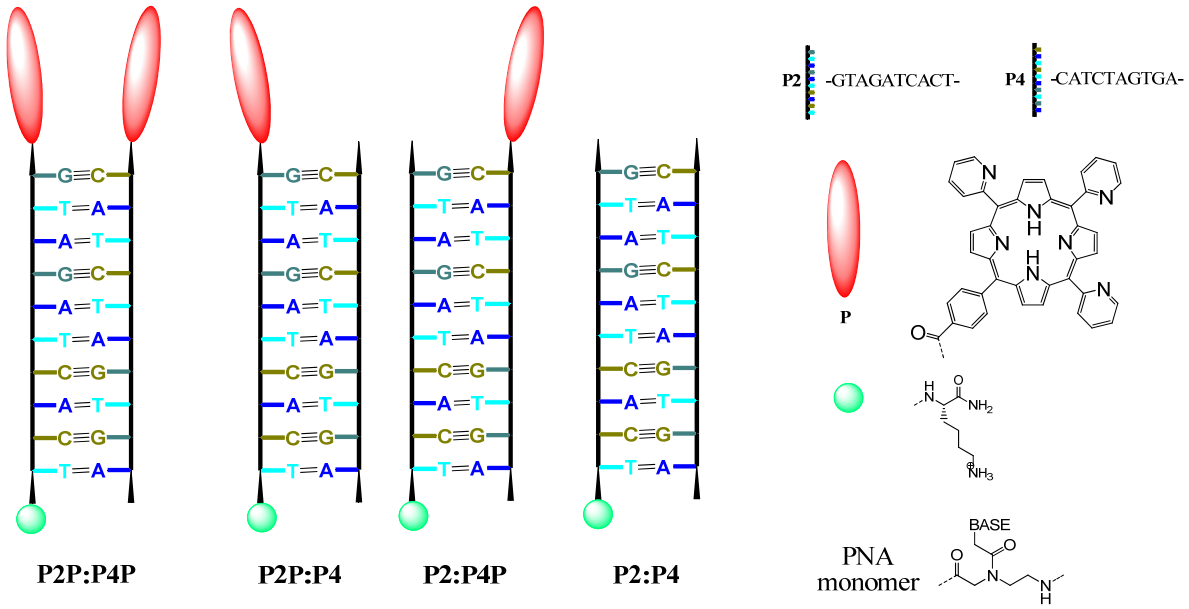
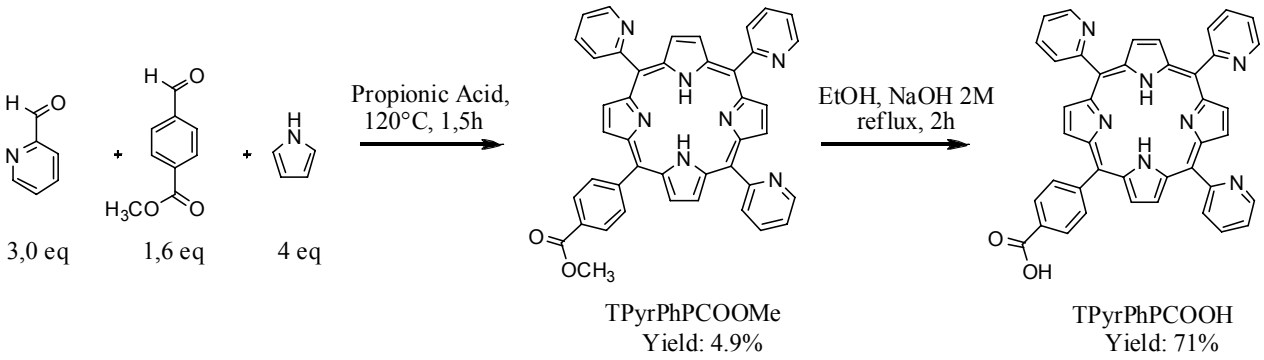


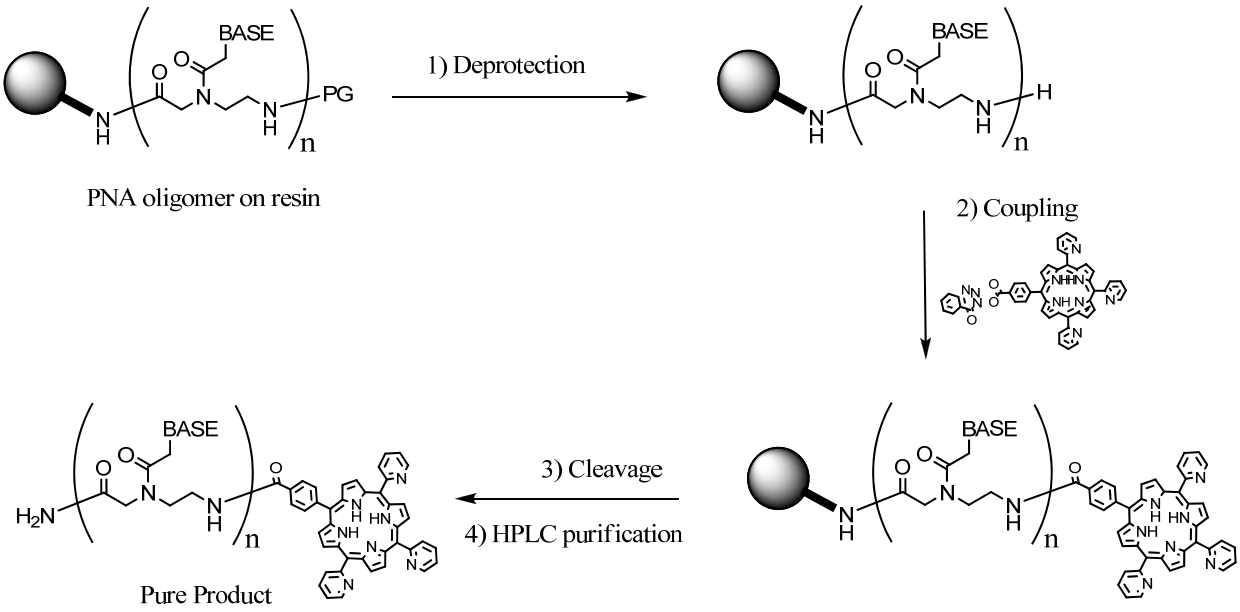
Figure 5.3. PNA – Porphyrin parallel duplexes used in the present study. Sequences are designed to match only in parallel fashion.

Synthesis of PNA-porphyrin conjugated oligomers. To the best of our knowledge only one report deals with a manganese porphyrin PNA conjugate[25] for applications in DNA targeting and cleavage. In that work the cationic manganese porphyrin was conjugated to PNA oligomer in homogeneous phase using carbodiimidazole/HOBt coupling reagents with 80% yield of porphyrin conjugation on purified PNA. In our work we use a naked *meso*-tris(pyridyl)phenylporphyrin carboxylic acid (TPyrPhPCOOH) which was synthetized according to a procedure reported in the literature[26] (*Scheme 5.1*).



Scheme 5.1. Synthesis of TPyrPhPCOOH.

PNA sequences were synthetized on solid phase using standard protocol for PNA (both Fmoc and Boc chemistry were employed). Trispyridylphenylporphyrin was conjugated on deprotected N-terminal amino group of PNA oligomers on solid phase using diisopropylcarbodiimide (DIC) and 3-hydroxy-1,2,3-benzotriazin-4(3*H*)-one (DhBtOH) coupling reagents. Porphyrin conjugate oligomers were then cleaved from resin and purified by reverse phase HPLC (*Scheme 5.2*).



Scheme 5.2. Synthesis of PNA – porphyrin conjugate. *Boc strategy* (PG: Boc): 1)TFA/m-cresol 95:5, 2x2 min; 2) Porphyrin preactivated with DIC/DhBtOH, 10 eq overnight; 3) TFA/TFMSA/m-cresol/tioanisole 6/2/1/1 2 x 1h, then drying and precipitation by Et_2O. *Fmoc strategy* (PG: Fmoc): 1)Piperidine 20%, 2 x 8 min; 2) Porphyrin preactivated with DIC/DhBtOH, 10 eq overnight; 3) TFA/ m-cresol 95/5 2 x 1h, then drying and precipitation by Et_2O.

Unexpectedly, after cleavage and purification of TPyrPhPCONH-GTAGATCACT-(gly)-NH_2 a side product with molecular mass + 61 m/z (*Figure 5.4*) was isolated together with the desired product. As shown in *Figure 5.4*, the side product has the same retention time of desired product suggesting complexation of ions pollutant on the porphyrin bound to PNA. Metal contamination during the cleavage step of porphyrin oligonucleotide derivative was also reported by *Mammana et al*[27]. They ascribed this contamination to traces of copper(II) in ammonia used for cleavage of DNA-conjugate from resin and to a catalytic effect of DNA in the metallation process[28]. They also assigned contamination to copper (II) after UV-visible spectra and Q-band measurement of contaminated product. In our case the separation of contaminated product for UV characterization was difficult, but the M+61 molecular weight (*Figure 5.5*) suggested copper contamination (with loss of two hydrogen atoms from porphyrin).

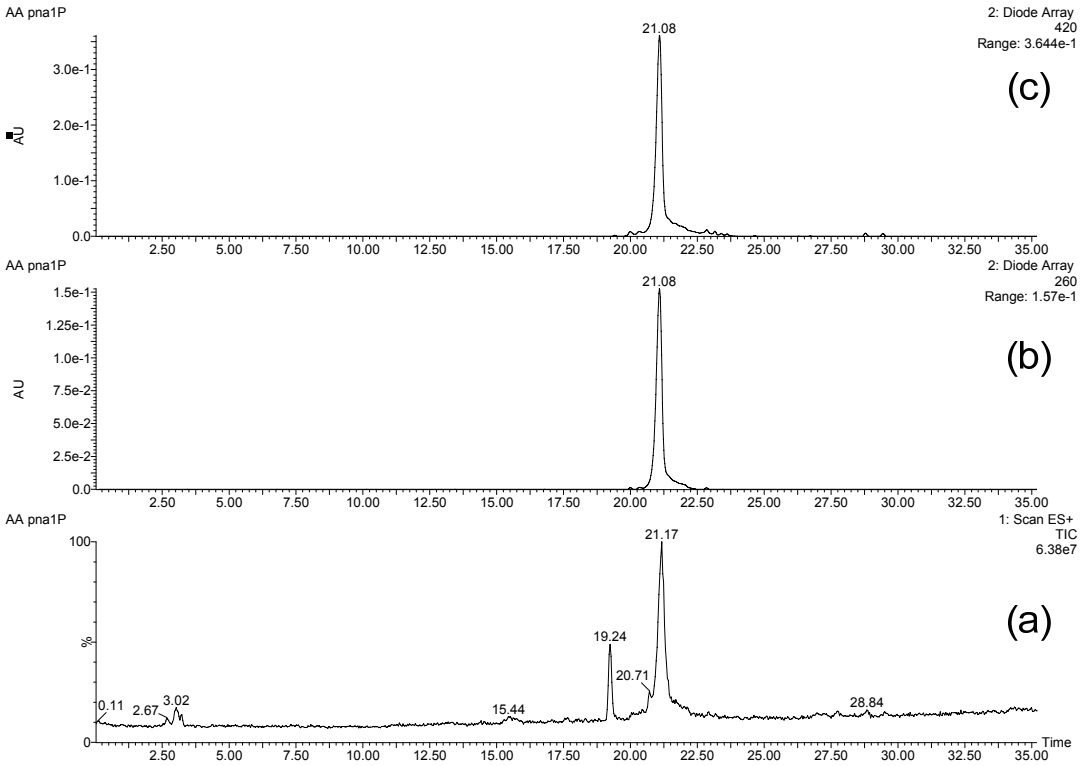


Figure 5.4. HPLC-MS-UV of product P1P (TPyrPhPCONH-GTAGATCACT-NH_2). (a) ESI-MS profile, (b) UV 260 nm and (c) visible 420 nm. (Conditions: column Phenomenex C_{18} 250 mm length, I.D. 5 mm, particle size 25 µM. Gradient elution: 0-5 min 100% Water - 0.1 %TFA, 5-35 min to 60/40 Water - 0.1% TFA / MeCN- 0.1% TFA).

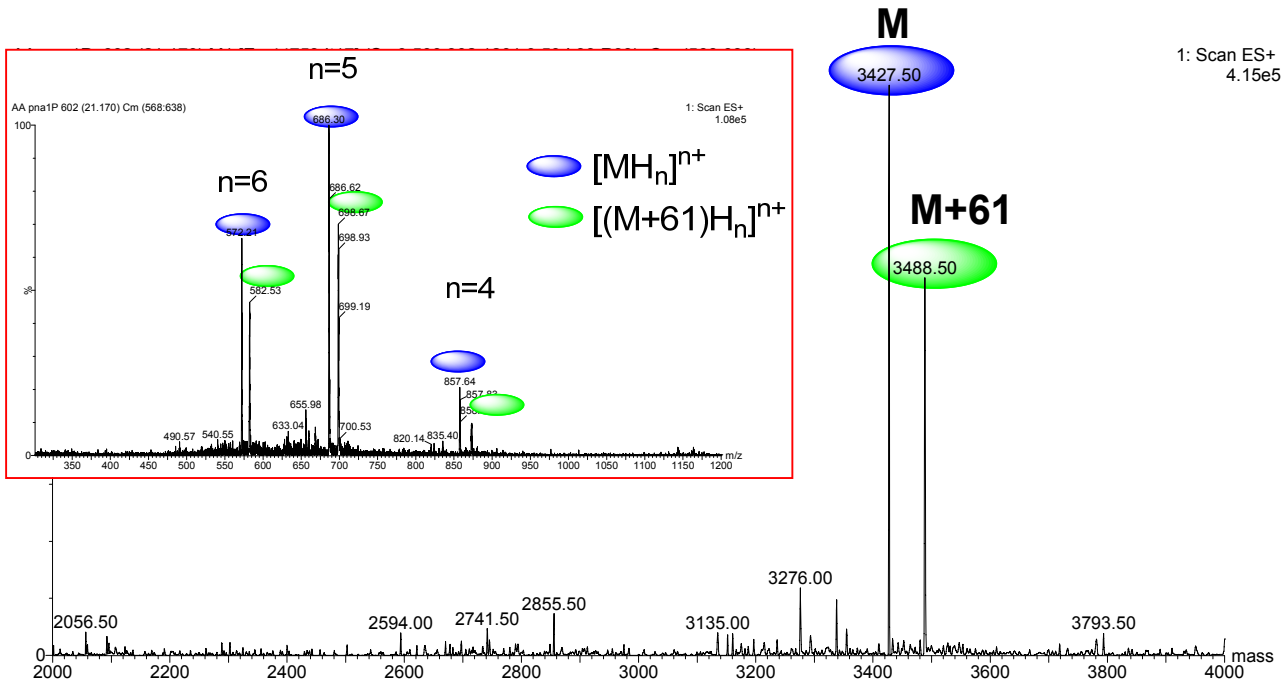


Figure 5.5. ESI-mass spectrum (red inset) and reconstructed spectrum relative to the peak at 21.08 min of chromatogram showed in *Figure 5.4*.

Exact mass measurements confirmed copper contaminations plus traces of nickel. Contaminations probably arises from ethyl ether used in washings of crude PNA that contains appreciable amounts of Ni(II) and Cu(II). Probably metals are first bound to PNA (coordination ability toward copper has been reported[29]), and catalytically inserted into the porphyrin. Since metallation process is slow, purification of crude PNA immediately after cleavage allowed to isolate free base porphyrin – PNA conjugates (*Figure 5.6*).

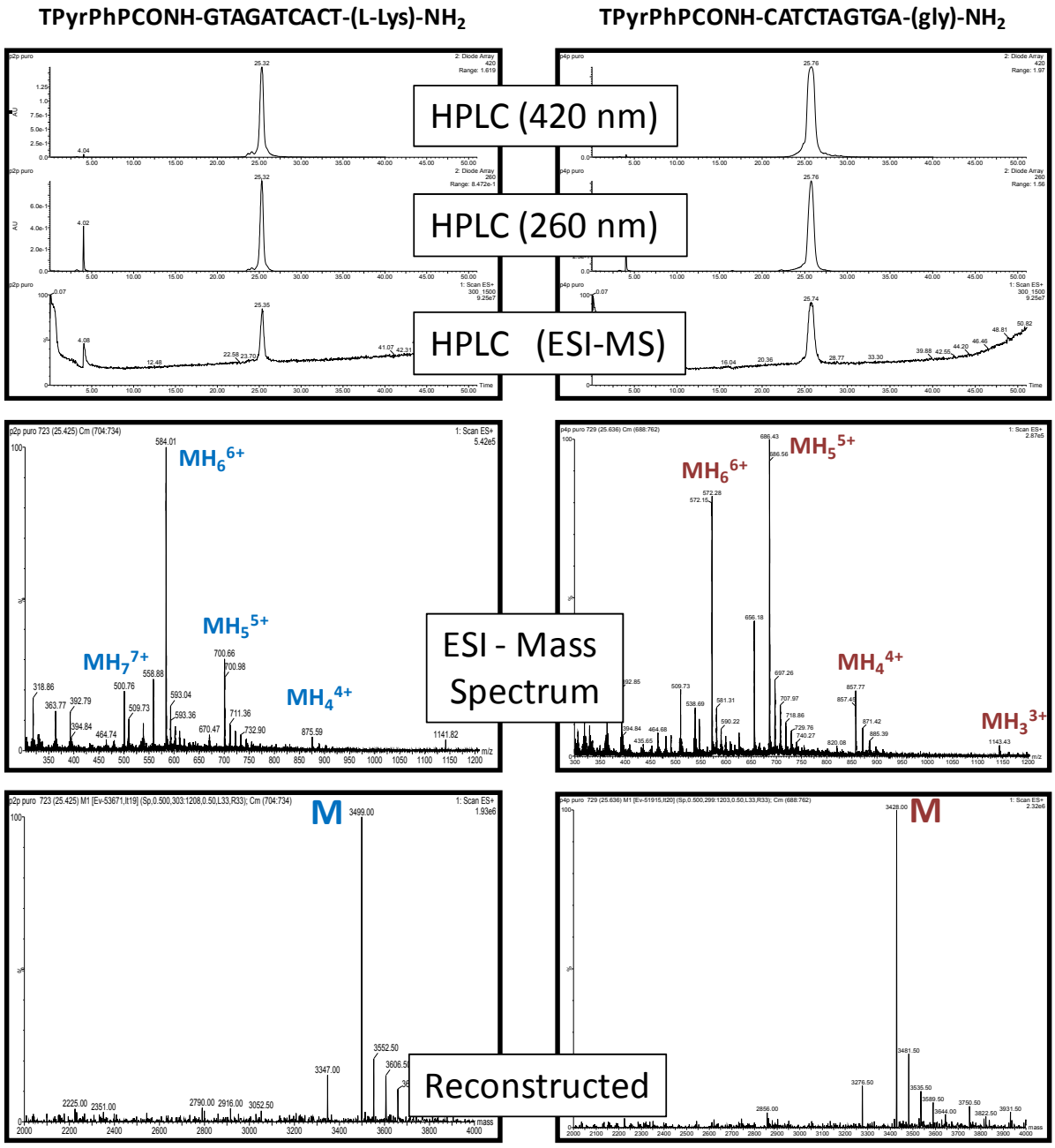


Figure 5.6. HPLC-ESI-MS-UV profile and related spectra of TPyrPhPCONH-GTAGATCACT-(L-Lys)-NH_2 (P2P) and TPyrPhPCONH-CATCTAGTGA-(gly)-NH_2 (P4P).

Yield of pure product were 3.9% for P2P and 13.7% for P4P. This difference is mainly due to different synthetic strategy employed for PNA oligomers, P2 synthesized by Fmoc strategy was obtained with 8,1% yield while was 33.1% for P4 (using Boc chemistry). This is the first example of synthesis of free base PNA-porphyrin conjugate on solid phase. UV-visible spectra of pure product are depicted in *Figure 5.7*.

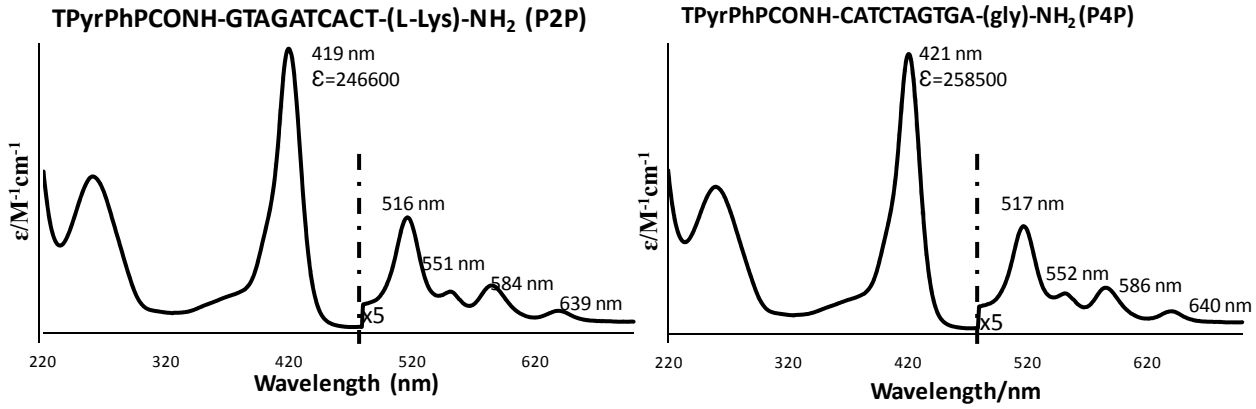


Figure 5.7. UV-Visible spectra of derivative P2P and P4P in water. Q-Band pattern is in agreement with free porphyrin[30].

Circular Dichroism (CD) studies of Parallel PNA:PNA duplexes. The ability of the PNA-porphyrin conjugates to form helical structures was studied by CD spectroscopy. Two different media were chosen: (a)glycerol/water 35/65 buffered at pH 7 with 10 mM phosphate; (b) water buffered at pH 4 by 50 mM of ammonium formate. These particular conditions were selected in order to avoid problems of solubility with porphyrin conjugates.

In a previous study on PNA:PNA duplexes bearing an asymmetric α-carbon embedded in the backbone (see *Figure 1.7c* of *Chapter 1*) it was shown that using the same configuration of the stereogenic center a helical inversion was observed between parallel and antiparallel structure[22]. This was rationalized on the basis of retro-inverso concept for peptide[31]. We found that for L-Lys attached to C-terminal induced for both parallel and antiparallel the same duplex helical sense. This was proved to be left-handed for both P2:P4 parallel and antiparallel duplexes by induced CD exciton chirality of a cyanine dye (*Figure 5.8*)[32], which is known to interact with dsPNA forming chiral aggregates.

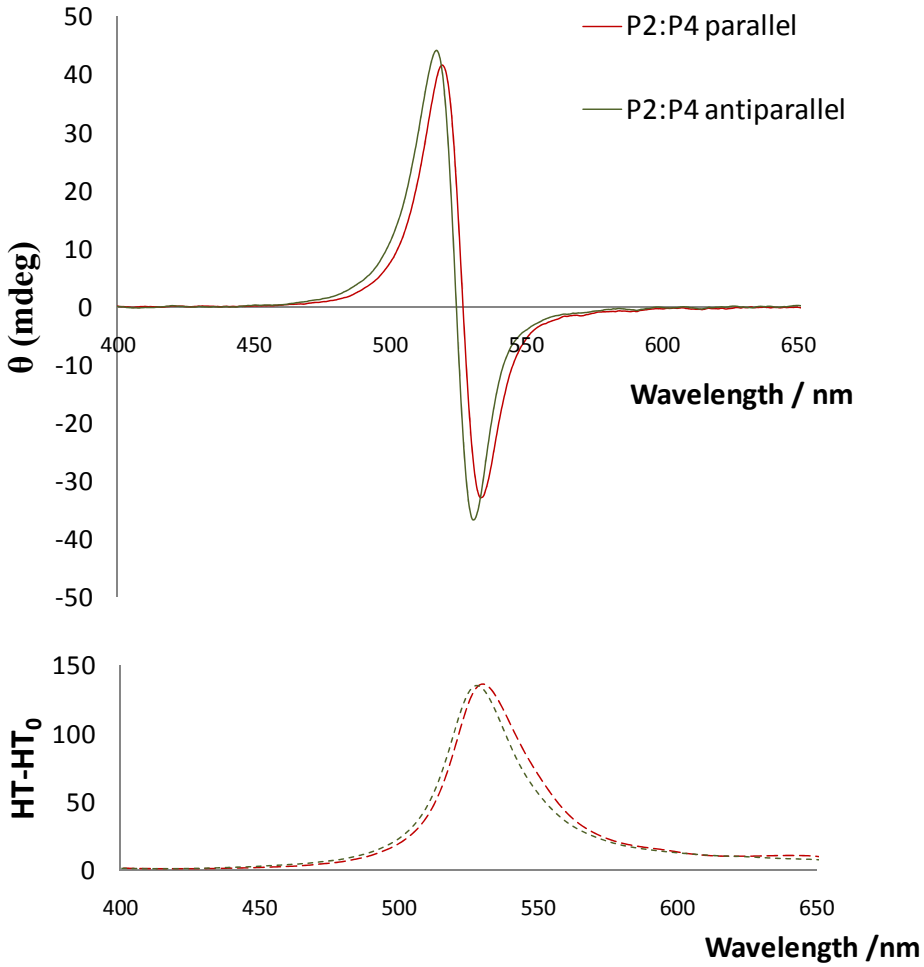


Figure 5.8. Parallel and antiparallel duplexes induced CD in presence of cyanine dye. (Conditions: [PNA]=5μM, [$Disc_2$(5)]=25μM, Buffer $HCOONH_4$ 50 mM pH 4, 25°C, optical path 0.2 cm).

The CD signal in the range from 240-280 of P2:P4 parallel duplex (*Figure 5.9*) looks similar to that of the antiparallel duplex based on L-lysine published by *Nielsen* and c*o-workers*[23,33] with an expected decrease of thermal stability. These findings indicate that also PNA:PNA parallel duplex assume an helical conformation such as in antiparallel duplex to maximize stacking interactions with the adjacent bases. Moreover parallel duplex showed a preferential left-handed helicity in presence of a terminal L-amino acid.

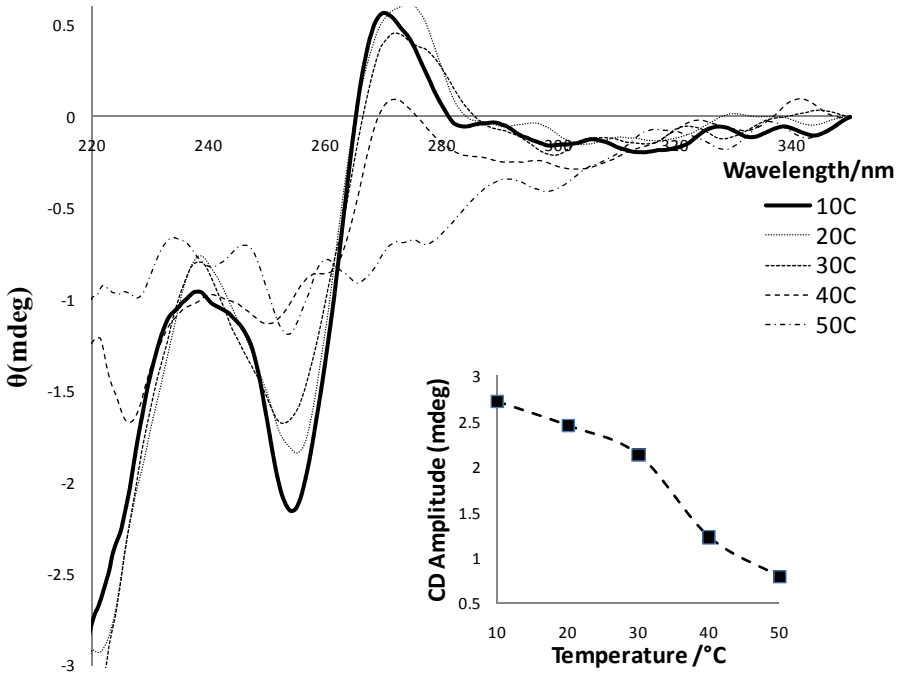


Figure 5.9. CD spectra at various temperature of P2:P4 parallel duplex from 220 nm to 350 nm. $T_{1/2}$(CD) is between 40°C and 30°C. (Conditions: Water/Glycerol-65/35 10mM Phosphate, [PNA]=5μM, optical path 1 cm).

The CD spectrum in the region of 240-280 nm of duplex P2:P4P (*Figure 5.10*) is almost the same as that of P2:P4. This means that insertion of one porphyrin at N-end of one strand did not lead to major conformational changes. Soret band of P2:P4P, as shown in UV absorption, is narrow and centered on 419 nm and does not lead to any signal in the CD spectrum. Insertion of two porphyrins on parallel duplex at N-terminus of each strands leads to dramatic changes in CD and absorption spectra (*Figure 5.11*). The bisignate signal in the region of 240 – 280 nm disappeared and no couplet of the two porphyrins was detected at Soret's band wavelength. As shown from HT channel Soret band in P2P:P4P is broad and red shifted compared to P2:P4P. Since the signal at 240 - 280 nm is weak, we decided to investigate duplexes using cyanine dye as previously done in the measurements showed in *Figure 5.9*. P2:P4 and P2:P4P showed intense bisignate signal as expected, while P2P:P4P displayed a very low bisignate signal, that is ten times less than P2:P4, which is different from chiral single strand P2 (*Figure 5.12*).

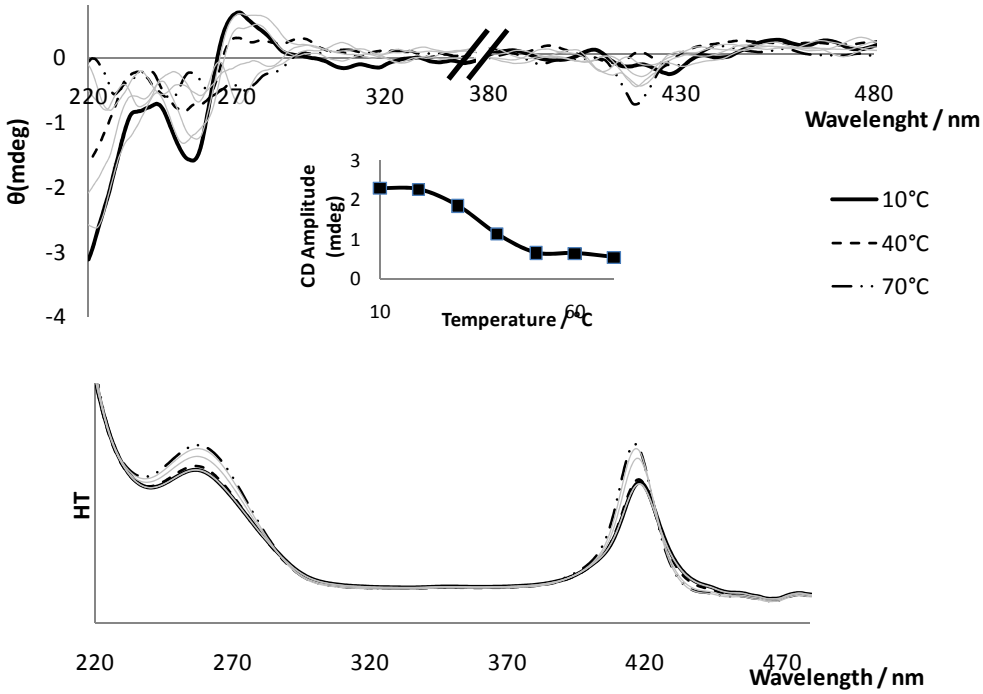


Figure 5.10. CD spectra at various temperature of P2:P4P duplex from 220 nm to 480 nm. (Conditions: Water/Glycerol-65/35 10mM Phosphate, [PNA]=5μM, optical path 1 cm). Grey lines indicate spectra for 20°C, 30°C, 50°C and 60°C.

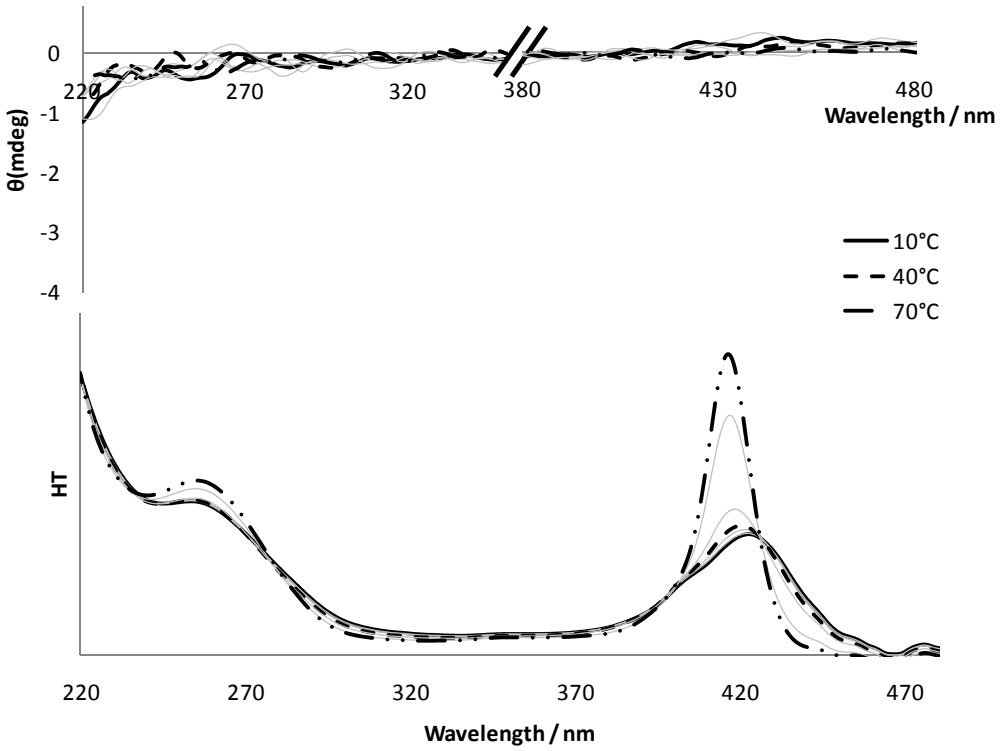


Figure 5.11. CD spectra at various temperature of P2P:P4P from 220 nm to 480 nm. (Conditions: Water/Glycerol-65/35 10mM Phosphate, [PNA]=5μM, optical path 1 cm). Grey lines indicate spectra for 20°C, 30°C, 50°C and 60°C.

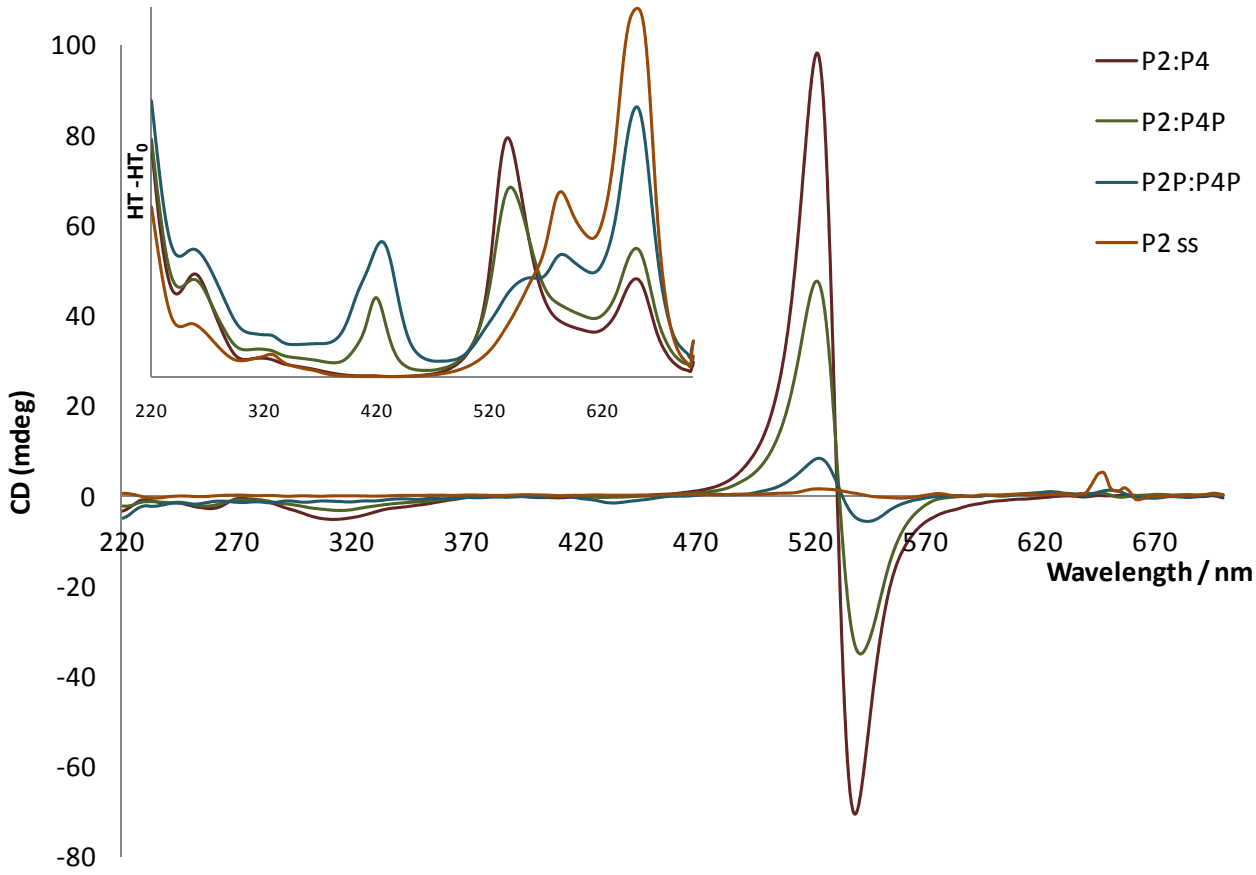


Figure 5.12. CD and absorption (inset) spectra of duplex P2:P4, P2P:P4 and P2P:P4P in the presence of the cyanine dye Disc2(5). Unbound dye absorbs at 650 nm, unbound dye dimers at 584 nm and bound dimers at 535 nm. Bisignate signals arise from couplets of two bound dye dimers in the PNA:PNA minor groove. Spectra were recorded at 10°C, using phosphate buffer in water/glycerol mixture 65/35 at pH 7, [ds]=5µM, [$Disc_2$(5)]=25 µM.

Measurements at pH 4 in water in the 240-280 nm region displayed a very low intensity compared to the same data in glycerol; this solvent effect is easily rationalized by the fact that polyols[34] and osmolytes[35] can stabilize native structure of some protein, as also observed for PNA duplexes in a previous PhD thesis work[36]. As showed by glycerol/water measurement, also in water at pH = 4 no exciton chirality was observed at Soret's band of P2P:P4P.

Duplex P2P:P4P at pH 4 showed a very low signal comparable with chiral single strand P2P (*Figure 5.13a*). In order to understand if the signal observed was really related to heteroduplex P2P:P4P or to a self aggregation of chiral strands, induced CD spectra were recorded at 50°C (*Figure 5.13b*).

Only P2P:P4P showed a bisignate signal centered on cyanine dye absorption of its bound dimers. These data proved that at 50°C for P2P:P4P a short helical segment is still preserved; thus we can conclude that the signal of P2P:P4P at 20°C is due to the presence of a duplex, while that of P2P is due to self-association of the single strands.

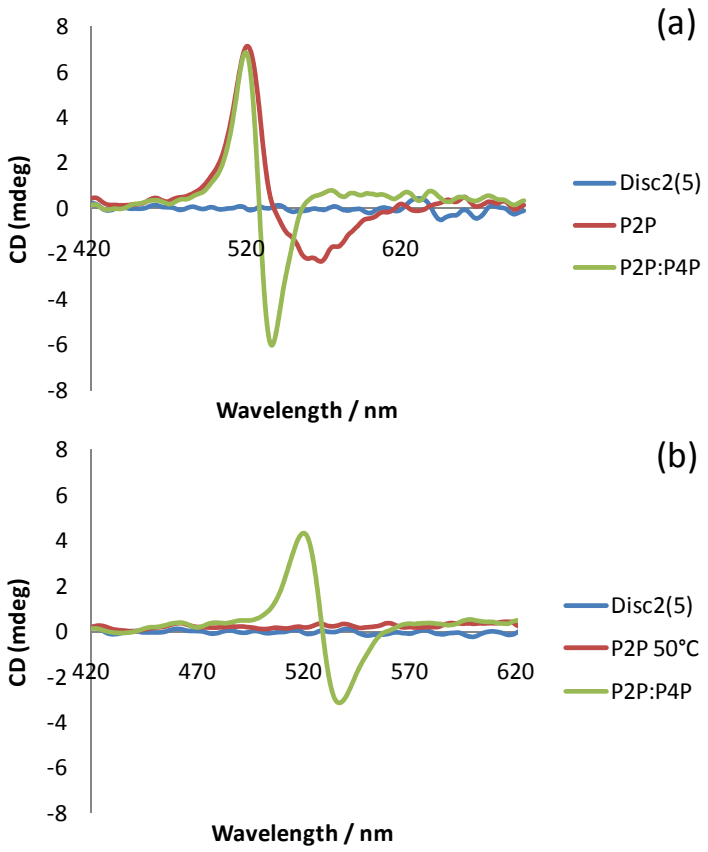


Figure 5.13. Induced CD spectra of P2P and P2P:P4P with $Disc_2(5)$. (a) Spectra recorded at 20°C; (b) Spectra recorded at 50°C. Conditions: Buffer $HCOONH_4$ pH 4, [PNA]=5μM, [$Disc_2(5)$]=25 μM, optical path 5 mm.

UV-Visible thermal denaturation studies. From CD studies it was established that in P2P:P4P the interaction between porphyrins dramatically change the duplex structure and CD spectra. In order to support P2P:P4P duplex formation and to study the influence of porphyrin on thermal stability of the duplex, we performed UV melting measurements of the various PNA.

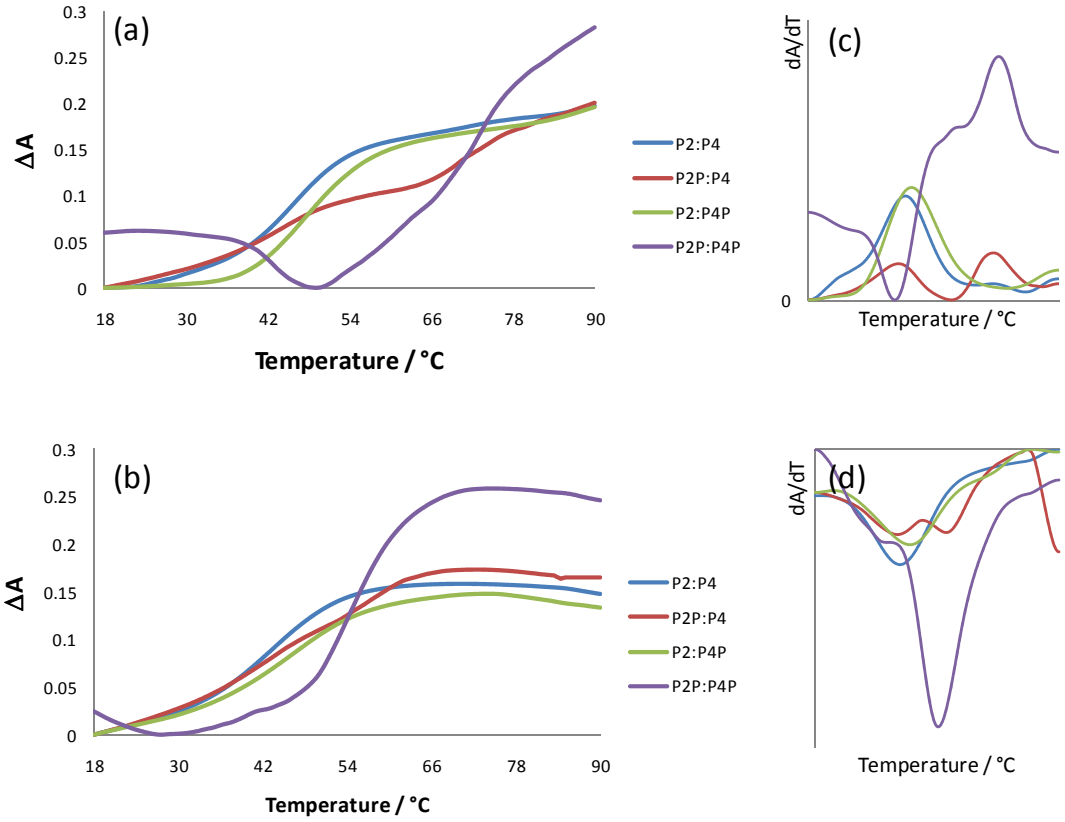


Figure 5.14. Melting (a) and annealing (b) curves of PNA duplexes, and first derivatives (c) and (d) of the curves reported in (b and (c) respectively. Conditions: buffer phosphate pH 7 in water/glycerol mixture 65/35, [ds]=5μM, optical path 10 mm, λ=260 nm. Melting was recorded by heating at 1°C/min from 18°C to 90°C, while annealing by following reverse process by cooling at 1°C/min rate from 90°C to 18°C.

In *Figure 5.14* melting and annealing curves of PNA duplexes are depicted and the corresponding melting temperature (T_m) and annealing temperature (T_a) values are shown in *Table 5.1*. From these data it is clear that insertion of one porphyrin on the parallel PNA:PNA duplex does not lead to dramatic changes in thermal stability if compared with the unmodified duplex P2:P4. It is interesting to comment on the slight difference of stability between P2P:P4 and P2:P4P. In P2P:P4 porphyrin is linked near a guanine monomer , while in P2:P4P near cytosine monomer and guanine is on the opposite strand. Guanine is recognized to stack efficiently with large aromatic system[37] leading stabilization of P2:P4P of few degrees. This capping effect is well known for DNA conjugate with large aromatics[38] and even porphyrins appended to DNA can stabilize non – Watson and Crick adenine guanine

base pair[39]. On the contrary, in P2P:P4 porphyrin labile interactions with complementary cytosine do not lead to duplex stabilization.

Table 5.1. Melting and annealing temperature values relative to curves showed in *Figure 15*.

PNA:PNA	$T_{melting}$ (°C)	$T_{annealing}$ (°C)
P2:P4	46	43
P2P:P4	43	43
P2:P4P	48	46
P2P:P4P	Multiple maxima from 53°C to 72°C	55

P2P:P4 duplex showed two less pronounced inflection at 67°C (melting) and 61°C (annealing).

P2P:P4P duplex stability is seriously affected by porphyrins that stabilize duplex struxture as showed by data in *Table 5.1*. Melting and annealing in P2P:P4P showed different shape and T_m values; while melting did not show a clear inflection point and first derivative maxima (72°C is the most intense positive maxima found in first derivative), annealing showed a very clear inflection point at 55°C. For all duplex we observed a difference between melting and annealing temperature, that is almost negligible for P2:P4 and for P2:P4P, while became more pronounced for P2P:P4P. To confirm cooperative association of P2P and P4P and to exclude aspecific porphyrin - porphyrin interaction between duplexes or homo-strand we repeated melting experiment halving duplex concentration to 2.5μM. Again melting temperature was higher (69°C) than annealing temperature which was the same as in the previous 5μM experiment (*Figure 5.15*). From the values of *Table 5.1* it is evident that small differences among melting and annealing temperatures were observed for P2:P4, P2P:P4 and P2:P4P. These data indicate that in melting and annealing processes of P2:P4, P2P:P4 and P2:P4P show slow kinetics in the annealing process or in the disassembly of the PNA duplex. However, this difference is very high in the case of P2P:P4P duplex (up to 20°C of difference were observed). In literature it is reported that disassembly of discrete porphyrin

aggregates has slow kinetics[40].Probably a kinetic effect of porphyrins – porphyrin dimers dissociation in P2P:P4P duplex shift the melting temperature of the duplex.

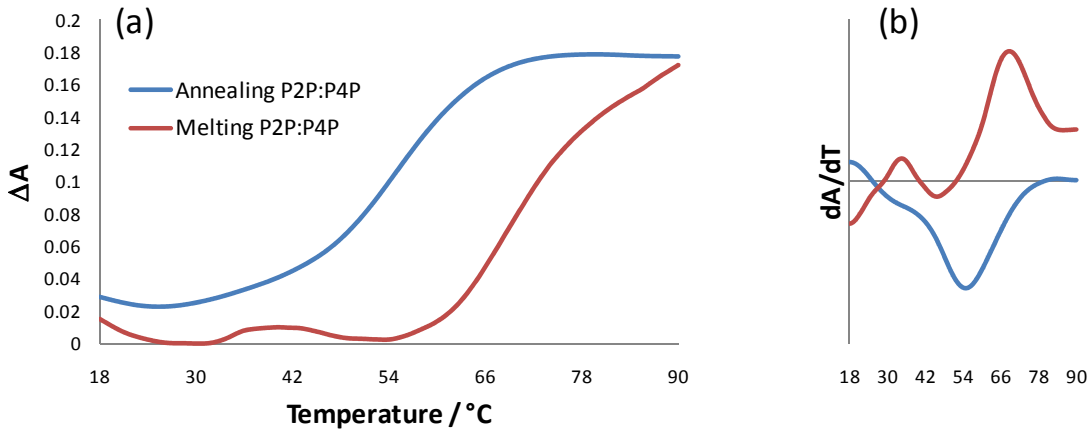


Figure 5.15. Melting and Annealing (a) and derivatives (b) curves of P2P:P4P at 2,5µM of duplex concentration. Conditions are the same of *Figure 14* except duplex concentration. Melting point is 69°C while annealing point is at 55°C. Hypercromicity as calculated in *table 1* is about 13% for both.

In P2P:P4P duplex the porphyrin –porphyrin interaction stabilized duplex at least of 10°C, based on annealing value. Further evidence of cooperative assembly of P2P and P4P in a parallel duplex came from melting experiment of the single strand P4P at 10 µM that did not show any cooperative melting or annealing curves compared to P2P:P4P (*Figure 5.16*).

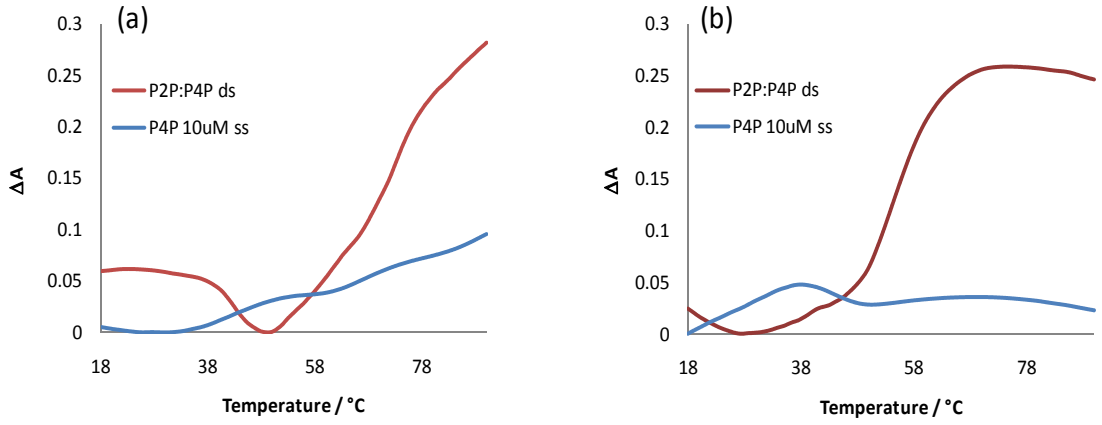


Figure 5.I6. Melting (a) and annealing (b) curves of P2P:P4P duplex 5µM (red line) and P4 10µM (blue line). Conditions are the same of *Figure 5.14*.

Soret's band shifts and shape are very diagnostic of porphyrin aggregation[41]. As pointed out before, while insertion of one porphyrin (P2P:P4 and P2:P4P duplexes) did not lead appreciable changes, in P2P:P4P the Soret's band is broad and red shifted. In *Figure 5.17* variable temperature spectra of P2:P4P and P2P:P4P in the visible region are reported. P2:P4P showed a narrow Soret's band centered at 419 nm with a limited bathochromic shift when temperature was decreased. On the contrary, by raising temperature the changes of Soret's band of P2P:P4P are more pronounced. In particular we observed a small red shift of about 4 nm, associated with an intense hypocromicity (decrease in intensity of the band). This behavior is consistent with a porphyrin-porphyrin interaction in the parallel P2P:P4P duplex. The red shift of few nanometer and hypocromicity[42] are compatible with an offset face to face geometry[43]

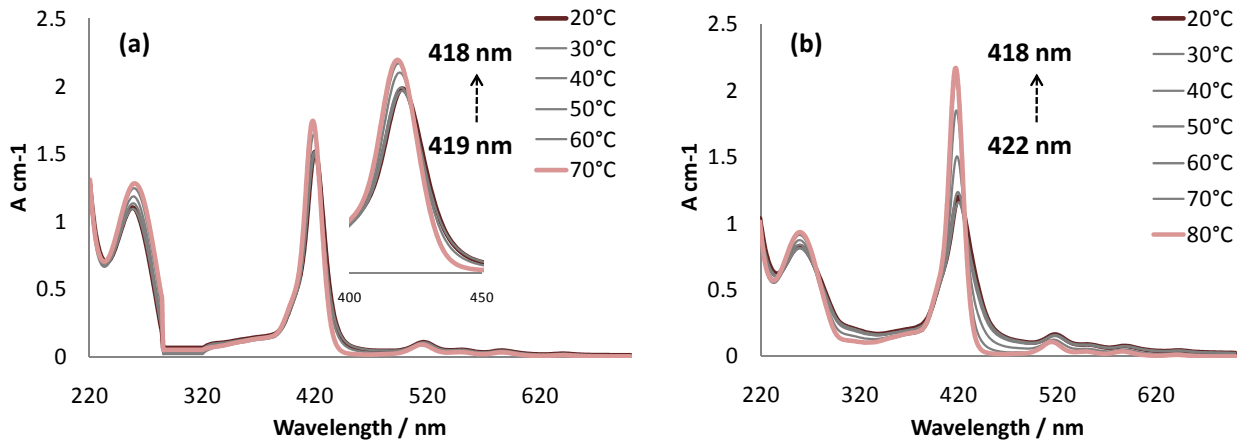


Figure 5.17. UV-Visible spectra of P2:P4P (a) and P2P:P4P (b) collected at variable temperature heating from 20°C to 80°C. Conditions as the same of *Figure 5.14*.

Melting curves obtained in the visible region at 418 nm for P2P:P4P duplex showed a sigmoidal shape, with inflection points close to those observed in the UV melting and annealing curves.

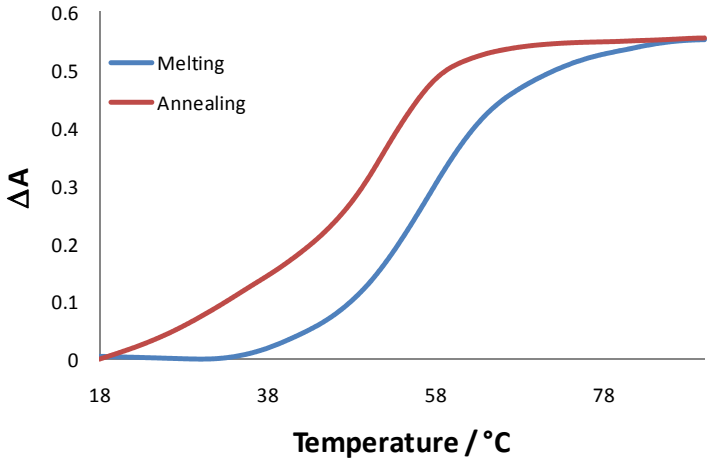


Figure 5.18. Melting and annealing curves at 418 nm. (Conditions as *Figure 14, optical path 5 mm*). Inflection for melting curves is 57°C and 53°C for annealing curves.

Stabilization of parallel duplex by porphyrin was also observed in water at pH= 4. In *Figure 5.19* melting and annealing curves of duplexes P2:P4 and P2P:P4P are depicted.

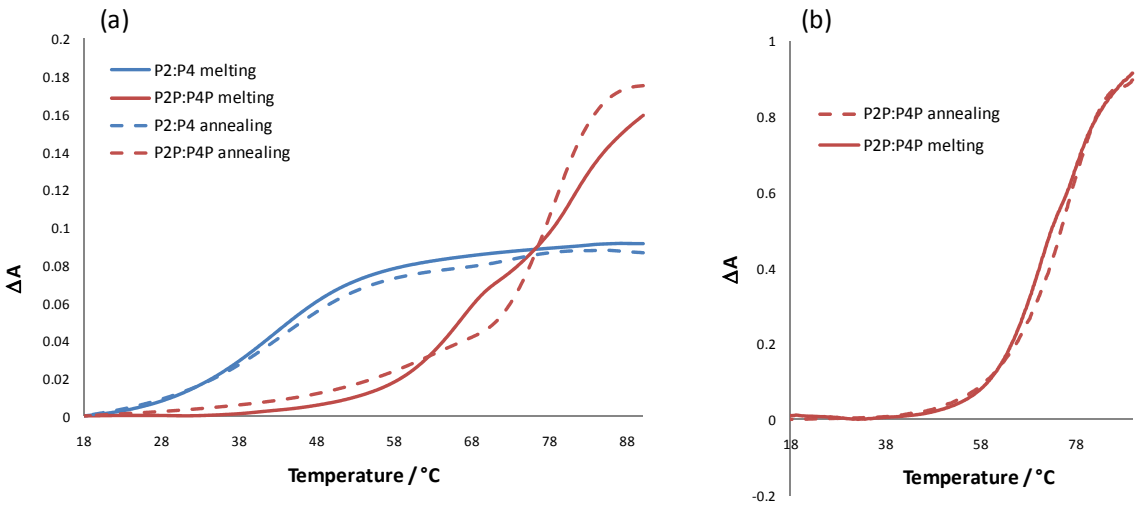


Figure 5.19. (a) Melting and Annealing curves for P2:P4 and P2P:P4P duplexes recorded at 260 nm. Conditions: Buffer at pH 4 (50 mM $HCOONH_4$), [Duplex]=5μM, heating rate 1°C/min, optical path 5 mm. (b) Melting and Annealing curves for P2P:P4P duplex recorded at 418 nm.

For the P2:P4 duplex melting and annealing values were both at 44°C. At both wavelength P2P:P4P melting profile showed a double transition (67°C/80°C for *Figure 19a* and 71°C/78°C for *Figure 5.19b*), while it became single in the annealing (78°C for *Figure 5.19a* and 76 °C for *Figure 5.19b*) probably due to slow melting/hybridization kinetic as discussed before (*Figure 5.20*).

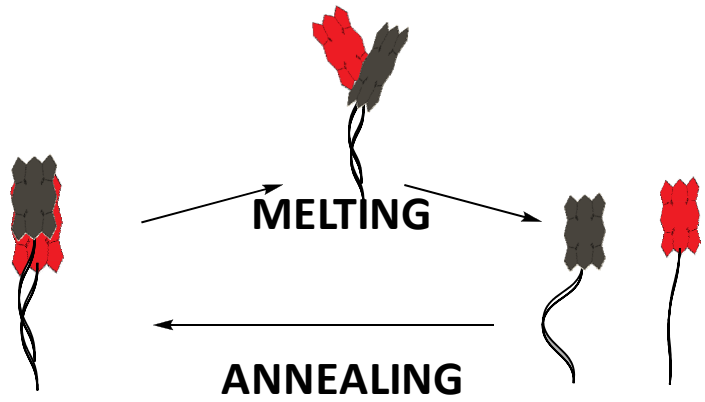


Figure 5.20. Possible explanation of different melting and annealing values of P2P:P4P. Melting of PNA part occurred only after loosening of porphyrin interactions. Annealing is cooperative since porphyrin-porphyrin interaction and pairing of two strand showed one transition.

A global analysis of melting data suggests that porphyrins strongly affect the behavior of PNA parallel duplex. The first issue is about hysteresis between melting and annealing curves. Hysteresis is negligible for P2:P4 at least of 3°C less for annealing values. This could be arise from dielectric constant modification of media that relax stacking interaction; in water at pH 4 very little hysteresis was observed confirming this interpretation of the solvent effect. More pronounced differences between curves occurred for P2P:P4P. Since melting and annealing values were found to be bigger for water, this solvent effect can be due to loosening of stacking interaction by glycerol, thus supporting the hypothesis of porphyrin – porphyrin interaction. Curves for P2P:P4P showed also double transition for melting while became single in the annealing; the same curve at 418 nm showed a melting value near the first transition at 260 nm. This confirm porphyrin effect in shape of melting curves and suggest that while porphyrin – porphyrin is starting to loosen, the PNA double helix is still paired.

Melting of porphyrin (values at 418nm similar to first transition at 260nm) is then followed by melting of PNA; the final results is an apparent very high increase of the melting temperature of PNA due to porphyrin slow kinetics of dissociation that delay melting of PNA part. Annealing of P2P:P4P showed a smaller value since porphyrin – porphyrin interaction is cooperative with annealing of PNA.

Hybridization effect on Fluorescence and Soret's band Shape. In order to further probe the porphyrin-porphyrin interactions, fluorescence measurements on porphyrin containing single-strand PNA and PNA:PNA duplexes were carried out. We observed changes in Soret's band and fluorescence of single strand PNA-conjugate upon hybridization (*Figure 5.21*).

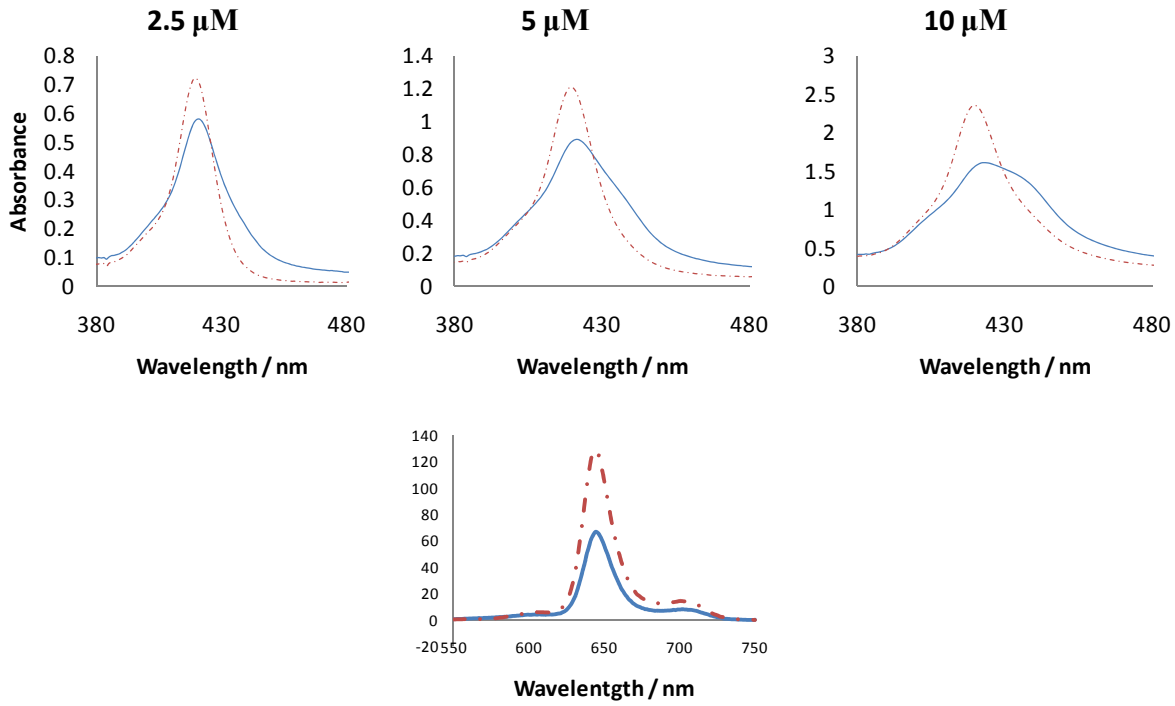


Figure 5.21. Soret's band shape and fluorescence of P4P single strand (blue line) and P2:P4P duplex (red dashed line) at different concentrations. Conditions as showed before in *Figure 14*. Fluorescence excitation at 419 nm. Absorbance is normalized for 1 cm of optical path. Fluorescence intensity is expressed as arbitrary unit. For 5µM concentration relative quantum yield of P4 toward P2:P4P is 76%.

P4P single strand at 2.5 μM, 5 μM and 10μM respectively, showed a broadening in Soret 's band shape compared to P2:P4P duplex at same concentration. P4P single strand showed fluorescence at 645 nm if excited at the maximum of Soret's band at 419 nm. Upon hybridization with complementary strand P2 to form P2:P4P duplex, the fluorescence intensity was increased (*Figure 5.21*). In aqueous solution both PNA single strand and porphyrins have tendency to self-aggregate[44]. P4P alone has great tendency to self aggregation leading broadening of Soret's band and reduced fluorescence emission. Upon hybridization, porphyrin linked to P2:P4P duplex is more hindered and less prone to aggregation. The same effect was observed also in the comparison between P2P and P2P:P4 duplex.

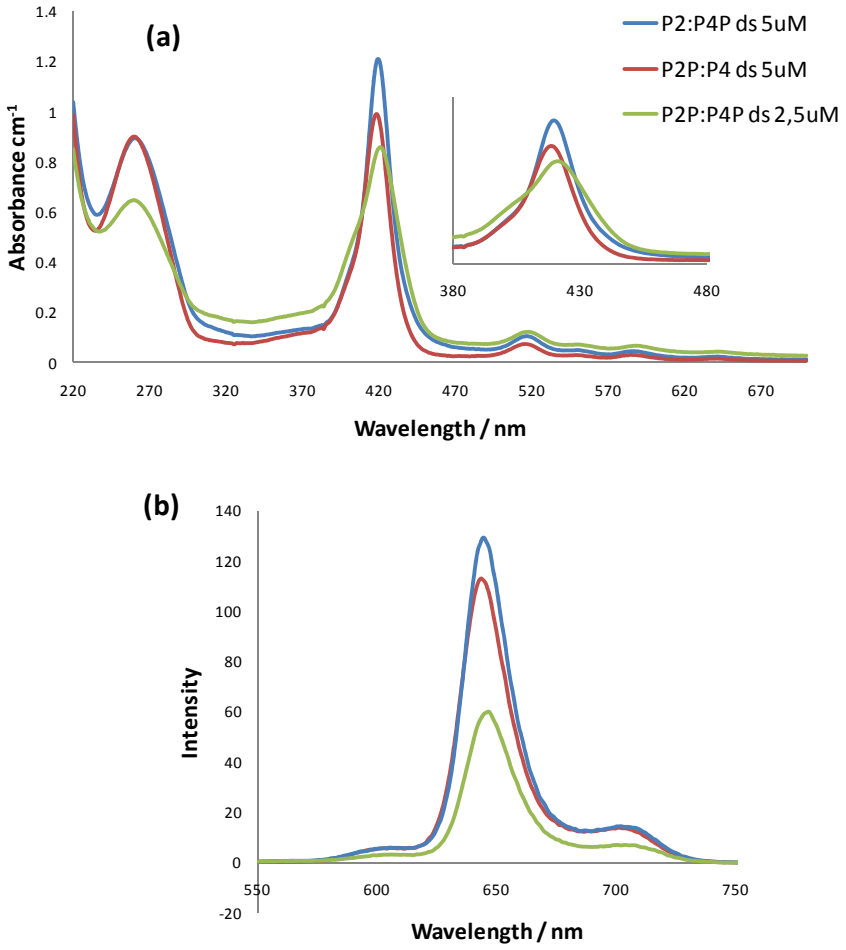


Figure 5.22. (a) Absorption and (b)Fluorescence spectra of P2:P4P, P2P:P4 at 5 μM and P2P:P4P at 2,5 μM. Conditions as in *Figure 5.14*. Relative quantum yield (reported as 100% for P2P:P4) are 90% for P2:P4P and 62% for P2P:P4P.

In order to avoid concentration effects in the porphyrin fluorescence, we compared P2P:P4P duplex at concentration half than P2P:P4 and P2P:P4P, therefore keeping constant the porphyrin concentration (*Figure 5.22*).

Fluorescence of P2P:P4P resulted quenched compared to P2P:P4 and P2:P4P. This self quenching arise from porphyrin – porphyrin interaction[45,46] in parallel duplex. Fluorescence and Soret's band shape of PNA-porphyrin conjugates are both affected by aggregation and structure of duplex. Single strands bearing porphyrin have tendency to self – aggregate that decrease upon hybridization, while hybridization of complementary strands bearing porphyrins lead to a self assembled porphyrin dimer with a parallel duplex scaffold.

Porphyrins effect on PNA parallel duplex. Transition from Helical to Ladder structure. In summary we observed that parallel P2:P4 duplex is helical as demonstrated by CD experiments both in the UV region and by interaction with a cyanine dye. Insertion of one porphyrin does not lead any substantial change in helical structure and stability, while drastic changes occurred in P2P:P4P, where we observed a loss of helicity, leading to disappearance of the CD band in the UV region and very weak induced circular dichroism with cyanine dye. The formation of the P2P:P4P duplex was demonstrated by UV melting measurements and its increased stability, when compared to P2:P4, P2P:P4 and P2:P4P, pointed out a stabilizing effect of the porphyrin-porphyrin interaction. The lack of the typical CD signature of helical structure imply a modification of the duplex induced by terminal porphyrin- porphyrin interaction. Thus, the natural preference of PNA:PNA duplexes for helical conformations, as revealed by the formation of P-Helix[47] in the solid state, is considerably deformed by the overwhelming effect of porphyrin-porphyrin interaction, which induces a constraint to the rest of the duplex. The results indicate that at least a segment of the P2P:P4P duplex is in an achiral conformation, though the hypochromic effect and hence base stacking is still observed, thus suggesting the formation of a ladder-type segment of the duplex.

This switch from helical form of the PNA:PNA duplex to a ladder type structure resembles the transition occurring in DNA in stretching experiments at about 60 pN of applied force.

A transition to a ladder-type structure S-DNA was proposed on the basis of experimental data and MD calculation[48]. The existence of this hypothetical structure has been subsequently questioned by other authors, on the basis of a thermodynamic model of stretching. However, overstretched states of DNA have been proposed to play an important role in several biological processes such as DNA recombination.

Ladder-type, base-paired duplexes have been described for DNA analogs by Eschenmoser et al in the case of pyranosyl-RNA and by Diederichsen *et al*[49] for poly-alanyl peptides containing nucleobase linked at β-position. Crystal structure of the latter derivatives showed a ladder base arrangement of nucleobase imposed from rigid β-sheet-like conformation of peptide backbone[50]. However, in these examples the ladder type structure was a consequence of the preferred conformation of the monomers, and no transition from helical to ladder-type structure was observed.

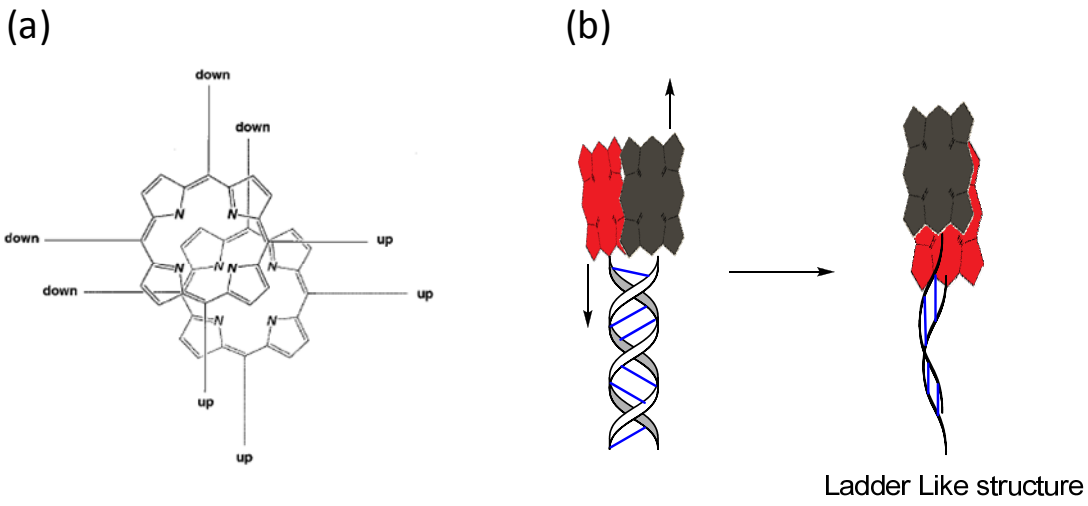


Figure 5.23. (a) Offset geometry for porphyrin – porphyrin interaction. (b) Model describing the effect of porphyrin – porphyrin interaction on the parallel duplex P2P:P4P.

In our case, the PNA:PNA parallel duplex can be modulated to a different conformation by the tension created by the two terminal porphyrins. Evidences for porphyrin – porphyrin interaction in P2P:P4P are in agreement with experimental data for offset face to face arrangement (*Figure 5.23a*). This porphyrin – porphyrin interaction in the parallel duplex could be the driving force that induces ladder like structure.

We believe that porphyrins induced distortion in parallel duplex resemble the tension applied to DNA; in other words, the constraint due to porphyrin dimer formation acts as a force that pull up one strand and pull down the second strand to induce conformational change (*Figure 5.23b*).

It is therefore interesting to study the effect of porphyrin-porphyrin interaction on the helical structure of parallel PNA:PNA duplexes by molecular modeling.

5.3 Molecular Modelling Studies on parallel PNA:PNA: Helical and Ladder Structure.

Our efforts to predict the most probable structure of the parallel PNA duplex have two-fold aspects: first, to infer conformational similarities/differences between parallel and antiparallel PNA:PNA structures and second, to open opportunity for exploring the effect of porphyrin conjugation to parallel PNA. In the search for a credible structure, we have selected a pool of four structurally-known helices (B-DNA, A-DNA, antiparallel PNA-DNA, antiparallel PNA:PNA), whose base-pair motif has served as a template for generating a starting geometry of the parallel PNA decamer. The correspondence between DNA and PNA based moieties is provided schematically in *Figure 5.24*. Regardless of whether starting from DNA, DNA/PNA or antiparallel PNA templates, the backbone strands of parallel PNA models have been conformed to accommodate the exact relative arrangement among 10 base-pairs (hydrogen bonding and stacking interactions) associated with the four known helices. A series of restrained minimization steps focusing on backbone-strands, while keeping the nucleic-bases harmonically constrained, has provided four initial structures. Since molecular dynamics (MD) simulation is regarded as a powerful method for obtaining structural and dynamic (degree of conformational flexibility) information of a macromolecular system, we have relied on this molecular mechanics approach in the quest for the most probable structure of the parallel PNA. By evaluating and comparing the energy of conformers that exhibit the highest residence time within MD trajectories, we have obtained a structure of parallel PNA which herein we propose as probable.

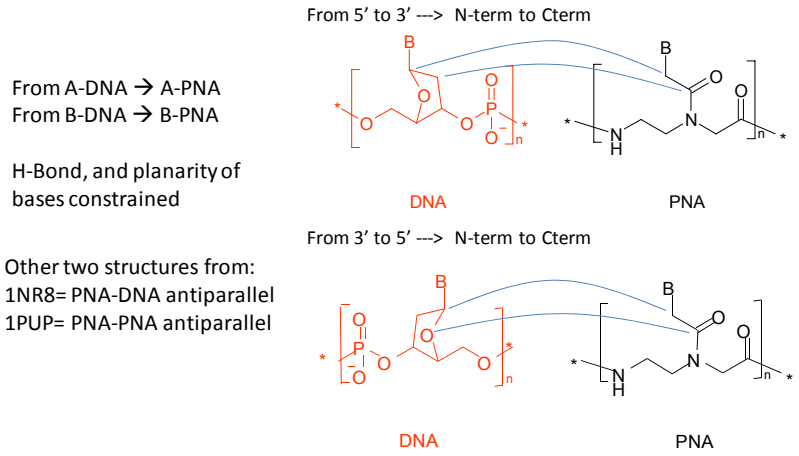


Figure 5.24. Correspondence between atom in DNA monomer modified in PNA monomer for both strands of parallel PNA duplex. In 1NR8 (PNA:DNA) DNA strands was changed from 2' to 5'.

Molecular Dynamics (MD) Methodology. In order to remove local distortion due to manual assembly procedure, the initial parallel PNA duplexes were relaxed by 5 ns harmonically constrained MD (FC=100 kcal/mol/Å^2) using OPLS-2005 force field. The four structures obtained as final geometries of the 5 ns trajectory have been designated as B-PNA (from B-DNA), A-PNA (from A-DNA), 1NR8-PNA (from PNA:DNA 1NR8-X-ray crystal structure in PDB), 1PUP-PNA (from PNA:PNA 1PUP-X-ray crystal structure in PDB). However, OPLS-2005 force field is not properly parameterized for obtaining the potential stable structure of the parallel PNA duplex. Since PNA is not regarded as a commonly encountered biomolecule, the conformational parameters (bonds, angles, and dihedrals) of PNA residues are not yet available within force-fields of standard molecular mechanics and molecular dynamics software packages. All MD simulations were done by employing the Desmond software package[51], by implementing custom-made Amber99 force field with atom-types and parameterization in accordance to publication by *Orozco and coworkers*[52]. An orthorhombic explicit water box was added, using TIP3P model, with solvent boundaries set at 10 Å from the surface of PNA. The MD algorithm built within Desmond module is regarded as highly efficient, partly because it automatically performs a prerequisite alternating minimization steps of the PNA model-structure and the explicit solvent.

The unconstrained MD simulations via customized Amber99 force field have been conducted for 3ns, using NPT ensembles with equilibration temperature set to 300K.

The four resulting trajectories obtained for A-PNA, B-PNA, 1NR8-PNA and 1PUP-PNA are provided in Figure 5.25. MD has sampled conformations along the 3 ns dynamics trajectory at a constant interval of 100 ps.The average geometries deduced from the stable 1ns intervals of the four MD trajectories were energy minimized in vacuum with 10 steps of the steepest descent. The obtained energies have been compared to infer the most likely prevalent structure of the parallel PNA duplex.

MD data-analysis and discussion. The criteria for assessing the conformational data and inferring the most probable structure was based on the MD-trajectory analysis (Figure 5.25) and energy comparison of the average output-structures (Figure 5.26). As it can be seen from Figure 5.25, the 3ns time evolution of the root mean square deviations (RMSDs in Å) of the overall PNA structure with respect to the starting conformation for each MD trajectory showed that all PNA trajectories were equilibrated after approximately 1 ns. The steadiest RMSD within 1ns time-interval has been identified and conformations surveyed within this region of the trajectory have been used for generating the average structure, representative of each MD run (Figure 5.26). Furthermore, the progression of RMSD values with respect to the average structure from each run is also considered in lower panel of Figure 5.25 and one can note the most steady trend is exhibited by 1NR8-PNA and 1PUP-PNA models. This data provides preliminary insight into the most reasonable models for parallel PNA duplex.

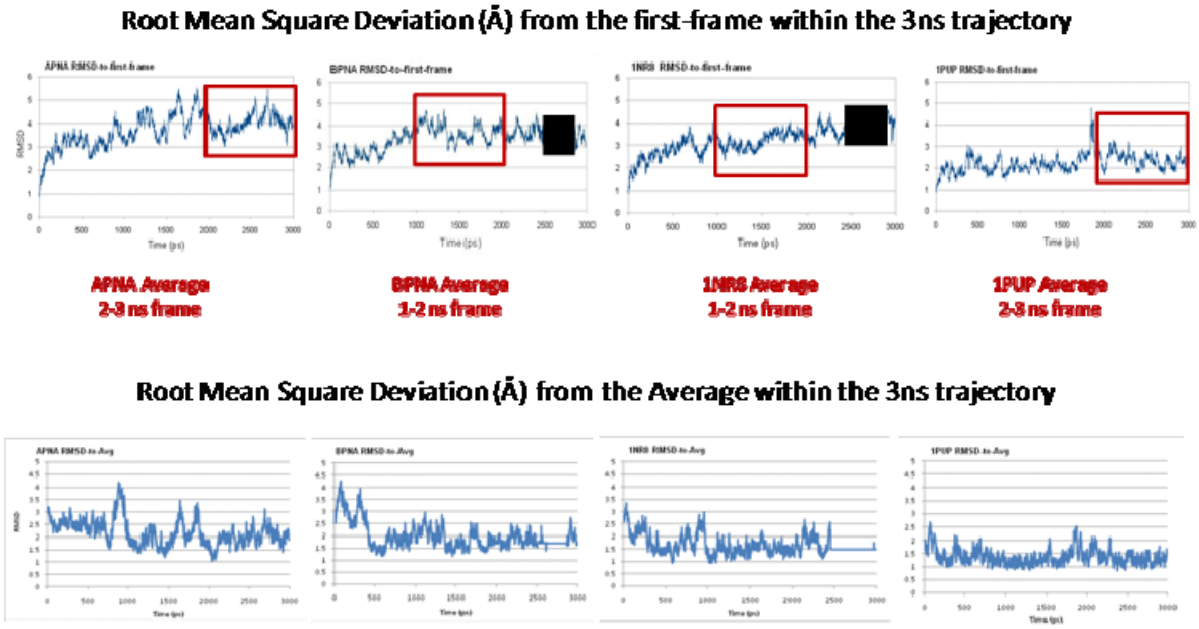


Figure 5.25. RMDS (Root Mean Square Deviation) from the first frame and from the average and trajectories portion chosen for representative structure.

From MD runs and subsequent energy calculations we found that 1NR8-PNA and 1PUP-PNA trajectories are stable and their averaged structures are very close looking with respect to both conformation and energy. An additional 5ns unconstrained MD applied on parallel 1NR8-PNA has lead to modification of this conformation to more closely resemble 1PUP-PNA. On the other hand, 1PUP-PNA was found to be stable upon further 5ns of unconstrained MD run (RMDS within 2Å). On the ground of these data it can be suggested that parallel PNA can exist in helical form, similar to 1PUP-PNA model. With respect to the antiparallel PNA, parallel PNA is similar in conformation yet it seems to exhibit an increased number of base-pairs per turn. Specifically, our parallel PNA structure showed 22 bases per turn, instead of 18 which is characteristic[53.] for antiparallel PNA.

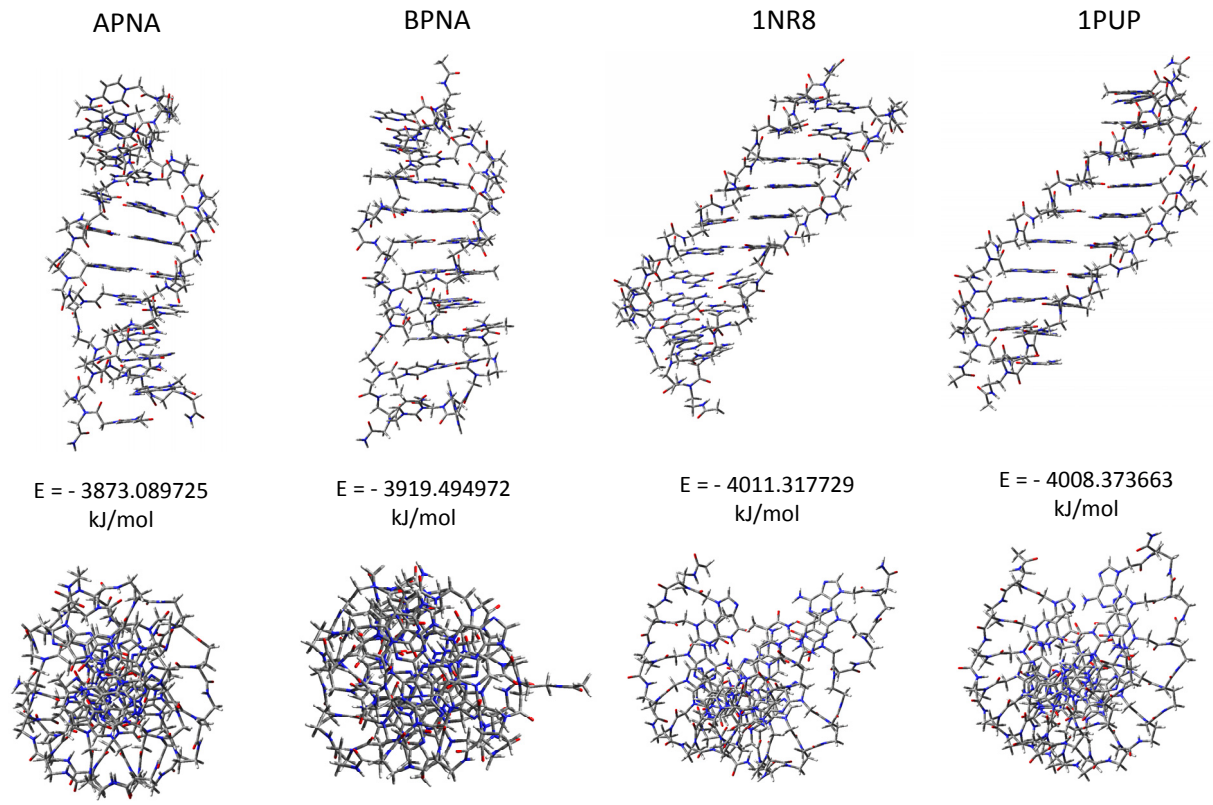


Figure 5.26. Averaged structures for trajectories portions selected in *Figure 5.25* and spot energy in vacuum.

Molecular Modelling of Bis-porphyrin/PNA Conjugate (Helix vs. Ladder Structure)

As reported earlier, our empirically-based interpretation of porphyrin effect on parallel PNA duplex was rationalized by assuming a helical distortion in a ladder like structure. Our objective herein was to provide evidence for validity of this assumption via molecular modelling. The challenge in performing such study, however, is the lack of universal force filed that would be able to simultaneously parameterize both PNA and porphyrin moieties. Although custom made Amber 99 is suitable for PNA, it does not include proper parameters for porphyrins. The suitable parametrization for porphyrins, yet lack in parameterization of PNA, similarly applied to OPLS-2005 force-field. Our strategy was then to resort to lower-level quantum mechanics to assess the structure and compute energy of P2P:P4P conjugate.

Starting from the assumption that upon covalent attachment of porphyrins, parallel PNA become notably less helical, our first step was to unwind and hence ladderize the 1PUP helix. The idea was to capture the conformational features of the helical twist so that we can distinguish between helical and ladder PNA-forms and at the same time optimize the size of

the system for quantum mechanical calculations. Therefore, the model-structure has been truncated from decamer to pentamer. Unwinding of the 1PUP helix and linearization in order to form a ladder-type parallel PNA was achieved by performing a series of minimization steps by placing a high force constant (FC= 1000 kcal/mol/Å^2) constraints on a set of dihedral angles, angles, and distance parameters, as indicated in *Figure 5.27*. Subsequently, the porphyrins have then been linked to the N-terms of the helical-PNA and ladder-PNA to explore their relative arrangement and stability.

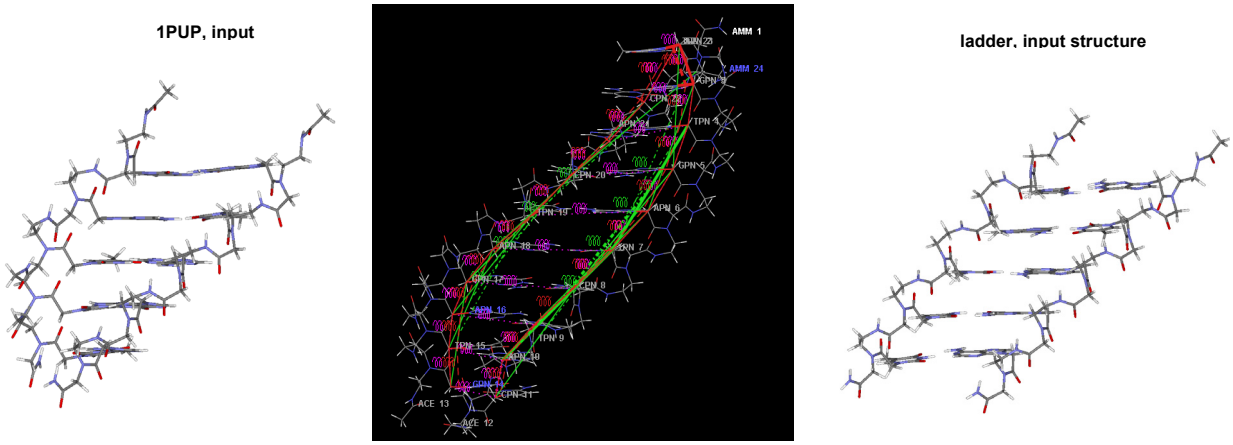


Figure 5.27. The left- and right-most panels respectively visualize the conformations of the helical 1PUP and its ladderized analogue. In the black box angles, dihedrals, and distances constrained are underlined by green and red lines. Basically dihedral angle between –CH_2-N(N^1 base) with its neighbors base are constrained to value close to 0°, and distance between base kept around 3.5Å.

The conformational influence that porphyrins can exert on the parallel PNA scaffold and vice versa, has been explored through the following specific aims:

A) Resorting to the Metropolis Monte Carlo (MC/OPLS-2005) conformational search to explore the preferred orientation of the bis-porphyrins under harmonic constraint of the helical PNA duplex;

B) Carry out the same search as in part A), except with PNA duplex transformed and harmonically constrained to a ladder-like scaffold;

C) Performing 200 hours of DFT-based energy minimization with B3LYP/STO-3G level of theory on 1) helical and ladderized form of parallel PNA and on 2) their bis-porphyrin analogues. The DFT runs have been performed on several starting geometries: representative low-energy conformations obtained from molecular mechanics modeling carried out under parts A) & B); also, PNA/bis-porphyrin conjugates that have been initially built were manually positioned, yet with identical arrangement of the two porphyrins relative to the helical and ladder PNA scaffolds.

More details concerning computational methodology are given in the subsequent paragraphs.

Monte Carlo (MC) Molecular Mechanics Methodology. An extended MC conformational search based on OPLS-2005 force field (MC/OPLS-2005) was performed on PNA/bis-porphyrin conjugate assuming that the helical conformation remain intact, by harmonically constraining PNA segment (FC=100 kcal/mol/Å^2), while allowing full relaxation of bis-porphyrin units. The system has been subjected to two search methods: a) Monte Carlo torsional sampling (*MCMM*) and b) Systematic torsional sampling (*SPMC*).

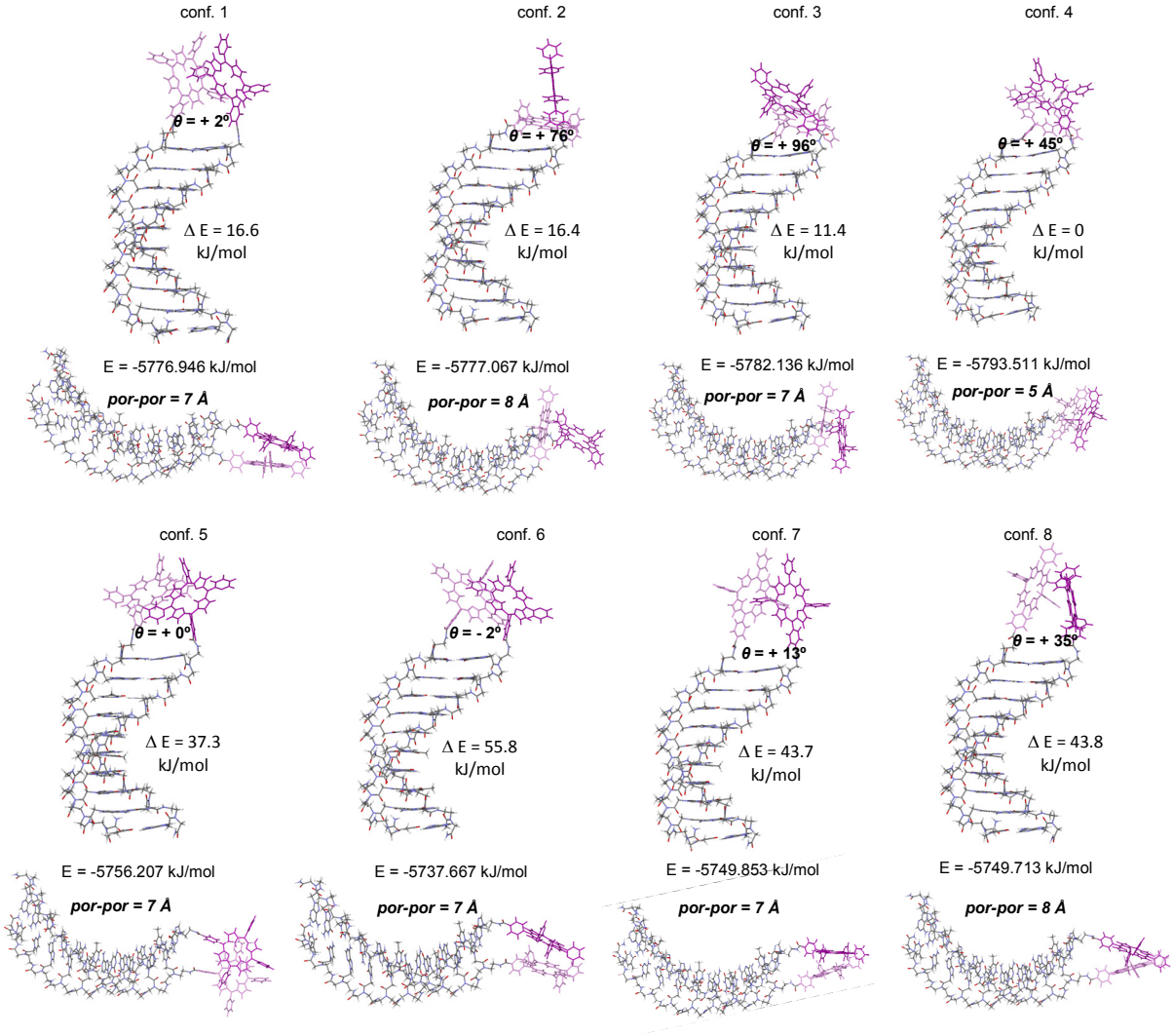


Figure 5.28. Two structural views, energies (E) and relative energies (ΔE) of the most stable helical conformations within 20 KJ/mole.

Upon running a series of consecutive searches each involving 10,000 steps and 20kJ/mol energy window, the 8 conformations (*Figure 5.28*) have been identified as representative orientations of bis-porphyrin units around the helical, parallel 1PUP scaffold.

The listed energies of the eight conformers have been determined by spot-energy calculation based on the OPLS-2005 force-field. The conformer with the lowest energy is labeled as *conf. 4* (*Figure 5.28*) and it is characterized by the shortest interporphyrin distance (5Å) among the eight structures and with the interporphyrin helicity, ϑ = +45° (ϑ = dihedral angle in 5-15 direction; the interporphyrin distance has been determined as an average of the two distances between NH moieties within the porphyrin cores).

We can infer that the π-π stacking stabilizes the relative orientation of the porphyrins and the right-handed helical scaffold induces the positive helical arrangement of the porphyrins. However, since OPLS-2005 is not parameterized for the PNA segment of the conjugate, the obtained data serve only as a preliminary indicator of the orientation that porphyrins are likely to adopt when covalently linked to the PNA helix. From Figure 5.28 we can note that, on average, porphyrins adopt co-facial arrangement with a positive interporphyrin twist.

The ladder/bis-porphyrin conjugate underwent the same type of the extended MC/OPLS-2005 search, yet this time with harmonic constraint placed on the ladder segment of the PNA. The conformational search has resulted in 7 representative conformations depicted in *Figure 5.29*.

As mentioned earlier, the lack of parameterization for the PNA segment within the OPLS-2005 force field requires only descriptive analysis of the prevailing conjugate conformations. We can see from Figure 5.29 that similarly to the case of helical scaffold, porphyrins on average have a tendency for a co-facial arrangement, yet unlike in the case of helical-PNA, here with a negligible interporphyrin helicity (ϑ). The low energy geometries, such as conf. 1 (*Figure 5.29*), are subsequently used as the starting structures for quantum mechanical energy minimization that can accurately treat both PNA and porphyrins segments of the conjugate. Therefore, MC/OPLS-2005 based molecular mechanics simulations serve to provide the reasonable input structures for quantum runs.

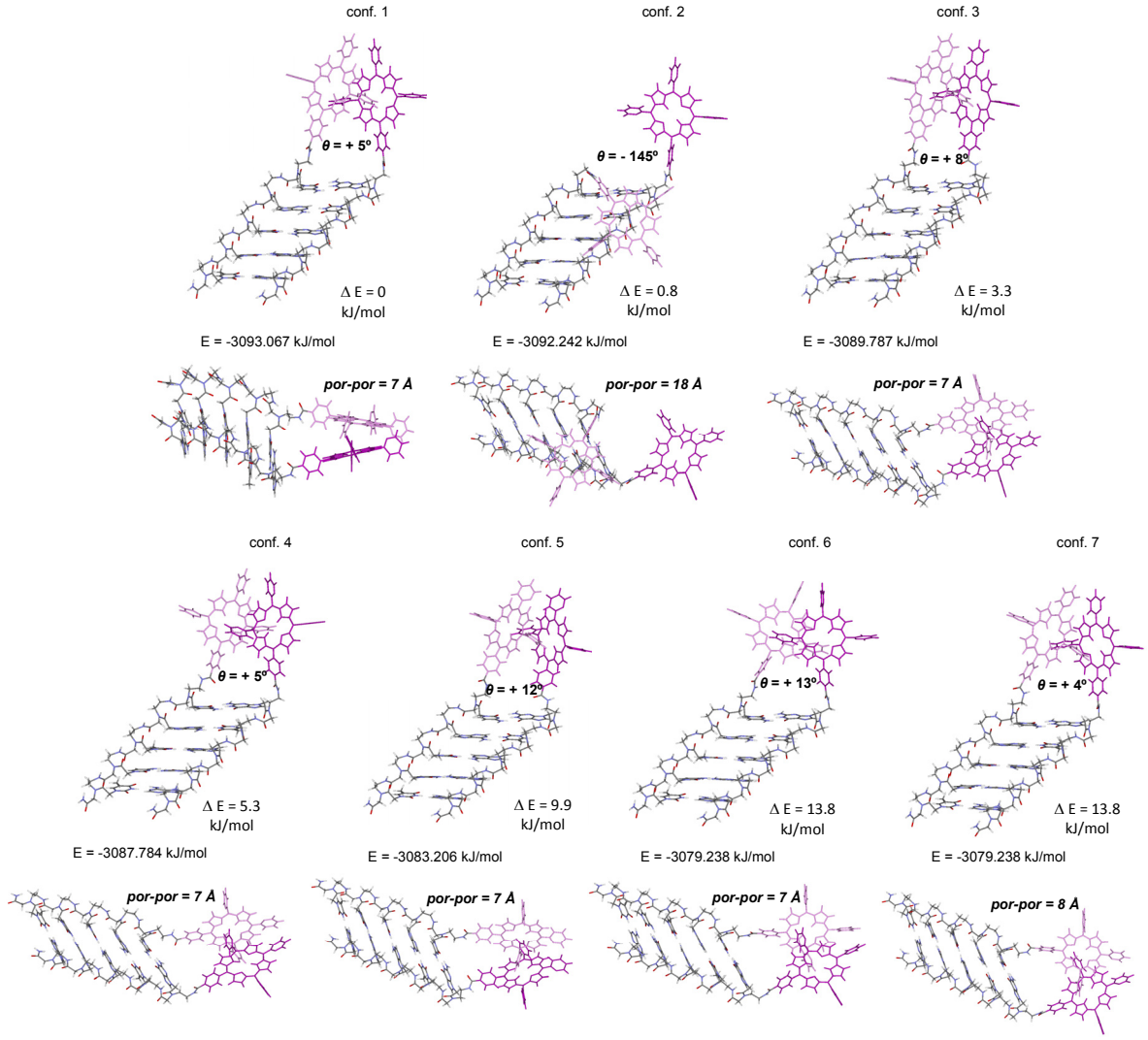


Figure 5.29. Two structural views, energies (E) and relative energies (Δ E) of the most stable ladder-type conformations within 20 KJ/mole.

Quantum Mechanics (QM) Methodology. The DFT energy minimizations have been first executed on the PNA templates lacking the porphyrins, in order to validate that the proposed helical structure is more stable than the ladder one and to estimate the energy difference. The quantum mechanical energy minimization was accomplished with DFT/B3LYP/STO-3G level of theory, using Gaussian03 software. The minimum level basis set was chosen in order to optimize the computational cost, yet still provide a DFT treatment. Considering the size of the system, our goal was not to obtain a fully optimized geometry, but to obtain the energy and the corresponding structural features after the entire geometry becomes relaxed, yet stable (strains are removed). Therefore, the DFT minimization was subjected to 200 hours calculation time.

Despite the fact that the minimization did not converge to full geometry optimization, the minimization algorithm allowed to approach the convergence as evidenced in plots of root mean square (RMS) gradient versus minimization steps as well as electronic energy versus minimization steps (*Figure 5.30*). The same type of calculations and analysis have been performed on the bis-porphyrin/PNA conjugates with both ladder and helical (1PUP) scaffolds of the parallel PNA.

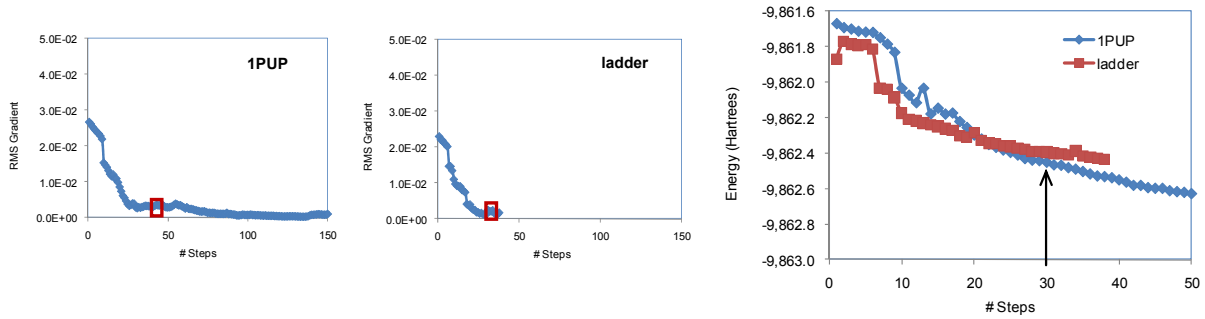


Figure 5.30. RMDS gradient vs. minimization steps (panels on the left) and Energy vs. minimization step (panel on the right)

Data-analysis and discussion (Ladder vs. Helix). *Figure 5.31* depicts the helical and ladder form of truncated 5-mer parallel PNA duplex along with their energies. We can conclude from this data that the helical PNA is energetically more stable than the ladder analogue by ~ 37 kcal/mol. On the other hand we can see from Figure 5.32 that the ladder PNA/bis-porphyrin conjugate is energetically more stable than the helical analogue by ~ 98 kcal/mol. In terms of the relative arrangement of the porphyrins, the interporphyrin distance is ~ 12.4 Å in case of the helical-PNA (1PUP-2P), while in the case of ladder-PNA (ladder-2P) the distance ranges from ~ 10.7 Å to ~ 18 Å. The most stable conformer of the bis-porphyrin ladder-PNA conjugate, nevertheless, is the one with interporphyrin distance of 10.7 Å. Although the experimentally observed fluoresence quencing is typically reported for interporphyrin distance of 4-5 Å, we cannot exclude the posiibility that such effect could also be elicited by porphyrins-porphyrin distance of ~ 10.7 Å. Nevertheless, this speculations are still under investigation.

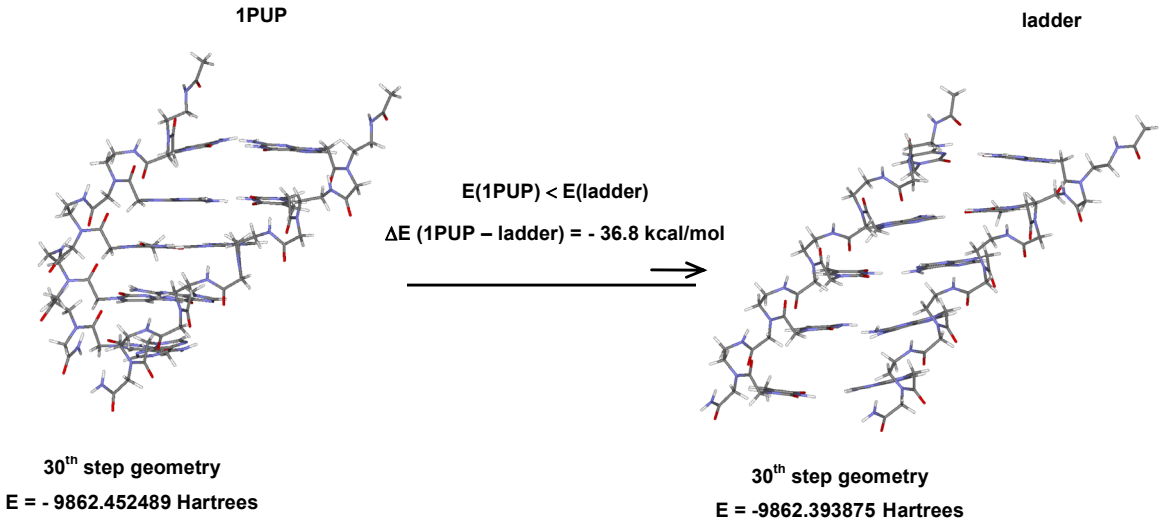


Figure 5.31. Structures and energy for helical and ladderized 5-mer parallel duplex.

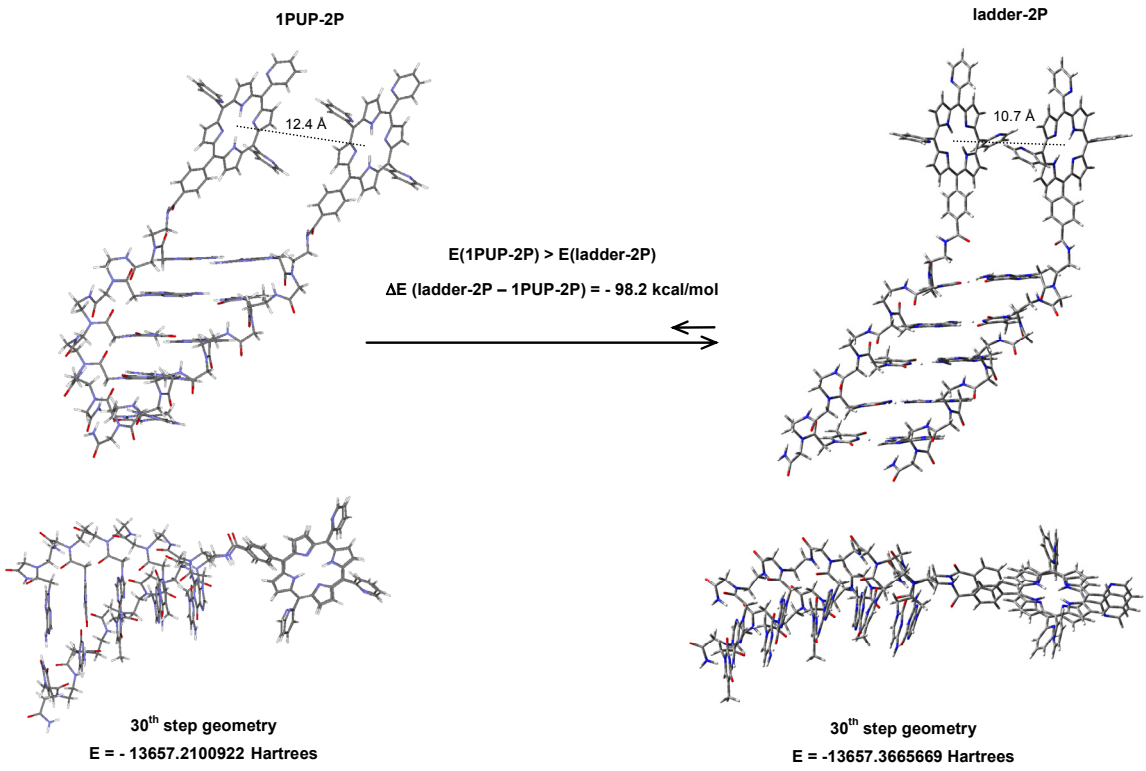


Figure 5.32. Two structural views and energies for helical and ladderized 5-mer of bis-porphyrin/PNA parallel duplex. The more stable conformation is represented by a ladderized PNA and a cofacial arrangement of two porphyrins.

Conclusion. Bis-porphyrin/PNA conjugate at N-term have shown very peculiar structural features. Although the parallel PNA duplex has a preferred helical structure, the introduction of a face-to face porphyrin-porphyrin interaction strongly affects the duplex conformation inducing a dramatic change. To further explore and understand the effect of porphyrins, additional quantum mechanical calculations are still underway (optimization of additional bis-porphyrin/PNA geometries).

The possibility to control conformation of duplex by conjugation with large aromatic molecules is interesting as a tool to control helicity in nanoscopic objects. Moreover, the proposed ladder-like structure for P2P:P4P is an important finding since can be used as a model for DNA ladder and S-DNA.

5.4 Incorporation of Uracil-Porphyrin Conjugate into PNA

The In order to produce a more general tool to probe supramolecular chirality and conformation of PNA containing duplexes we designed a new PNA containing two porphyrins, both on the same strand (*Figure 33*a), attached by a rigid linker through C (5) position of two uracil moieties (*Figure 33b*).

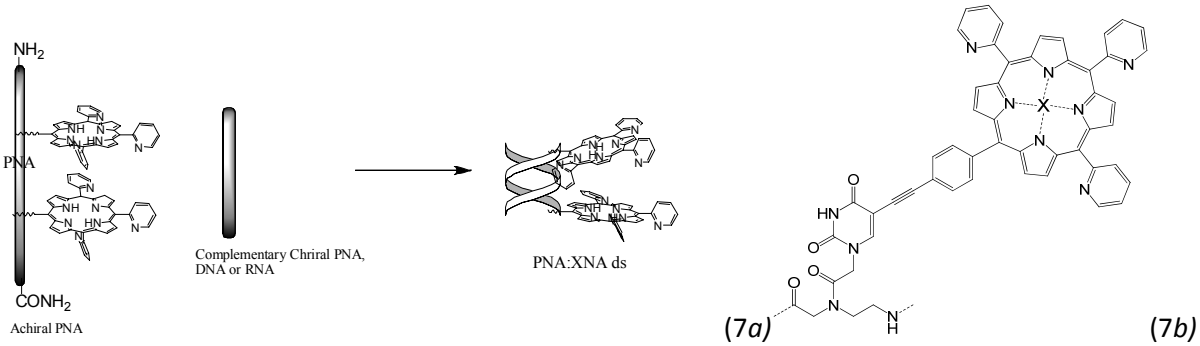
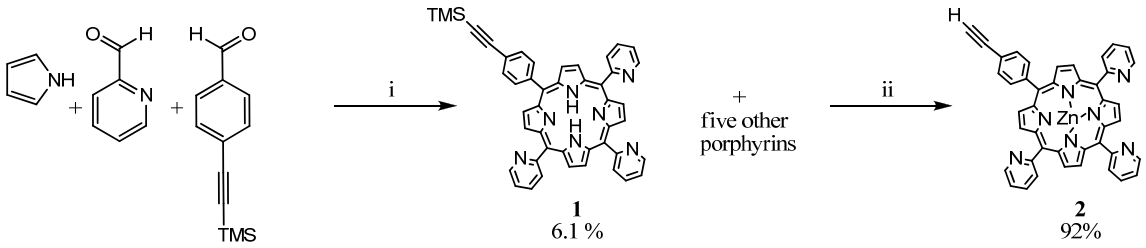


Figure 5.33. Design of PNA probe bearing two porphyrins on the same strand (X: 2H or Zn).

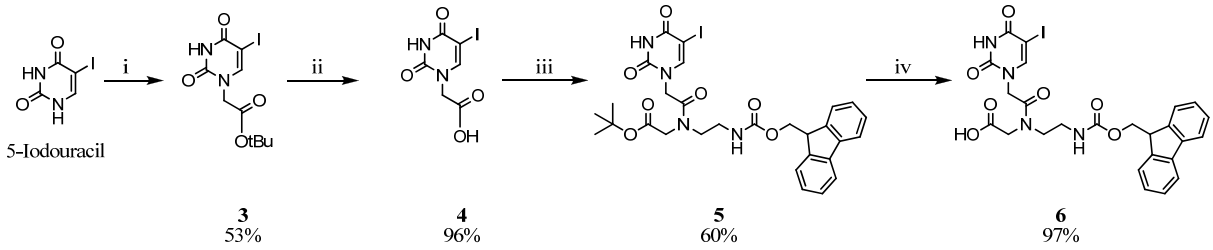
The orientation of porphyrins forced in fixed position by PNA:XNA duplex formation (XNA: chiral PNA, DNA, RNA) should reflect conformation of the entire duplex structure. Our aim is to produce a diagnostic exciton coupled CD at wavelength of Soret's band of porphyrins which can be used as a probe of the PNA conformation. This model has some advantage compared to parallel model showed previously: a) porphyrins can be placed by sequence design at distance suitable to avoid stacking interaction; b) the model is applicable to PNA:PNA, PNA:DNA, and PNA:DNA duplexes both parallel and antiparallel and can led a direct comparison of different structure if the same sequences are used. The synthesis of the probe was carried out following two different strategies (showed in the next pages), both using Sonogashira coupling between 5-iodouracil moieties and porphyrin. We chose to use a porphyrin containing three pyridyl rings instead of three phenyl to enhance water solubility of final derivatives. The porphyrin was synthetized according to the method of *Adler,* by refluxing pyrrole and corresponding aldehyde in propionic acid, followed by zincation and deprotection of alkyne protective group (*Scheme 5.3*).

Metallation of porphyrin with Zn is important for the next reaction. Naked porphyrin could interfere in Sonogashira coupling sequestrating metal catalyst (Palladium and Copper). In literature there is no report of Sonogashira reaction carried out with naked porphyrin.



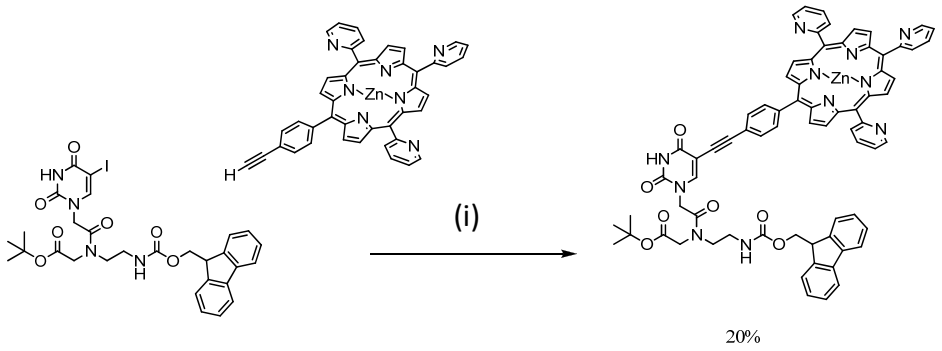
Scheme 5.3. (i) [Propionic Acid, reflux 4h]. (ii)[(a)Et_3N, $ZnCl_2$, rt, 1 day. (b) K_2CO_3, MeOH - $CHCl_3$, 1day].

Two strategies were employed to synthetize PNA bearing two porphyrin, one is based on the synthesis of PNA monomer containing porphyrin to use in standard PNA solid phase synthesis by Fmoc chemistry, and the second on the synthesis of PNA containing two iodouracil moieties on which Sonogashira coupling can be performed during solid phase synthesis. Both strategies need a 5-iodouracil PNA monomer as key intermediate (synthesis depicted in *Scheme 5.4*). The synthesis of key derivatives **4** and **5** started from iodouracil. followed by alkylation on N(1) position, deprotection of tert-butyl ester and coupling with Fmoc-aminoethylglycine-OtBu to give the protected monomer. Deprotection of **4** by acidic treatment afforded the compound **5** suitable for preparation of PNA containing two iodouracil moieties.



Scheme 5.4. (i) [$BrCH_2COOtBu$, K_2CO_3, DMF]; (ii)[TFA, DCM, 0°C]. (iii) [4, DCC, HOBt, DIEA 30 min then Fmoc-aeg-OtBu overnigth]. (iv)[TFA, DCM, 0°C].

The derivative **5** was employed in Sonogashira coupling with derivative **2** (*Scheme 5.5*).



Scheme 5.5. Sonogashira Coupling. (i)[$Pd(PPh_3)_4$, CuI, Et_3N, DMF].

The yield was pretty low: this result is due to Fmoc deprotection during the reaction, probably by TEA (triethylenammine). The deprotected product is also difficult to isolate, and thus cannot be reprotected by Fmoc. It is well known in literature that 5-alkynyl uracil are prone to intramolecular cyclization involving the triple bond and carbonyl in position 4 in presence of transition metals. This was avoided treating with saturated EDTA the reaction after coupling with porphyrin. The 1H-NMR spectrum of product is showed in F*igure 5.34*.

The Boc-U^I-OH PNA monomer was also synthesized in order to avoid Fmoc deprotection during Sonogashira coupling and to prepare entire monomer containing a porphyrin moiety on C(5) position.

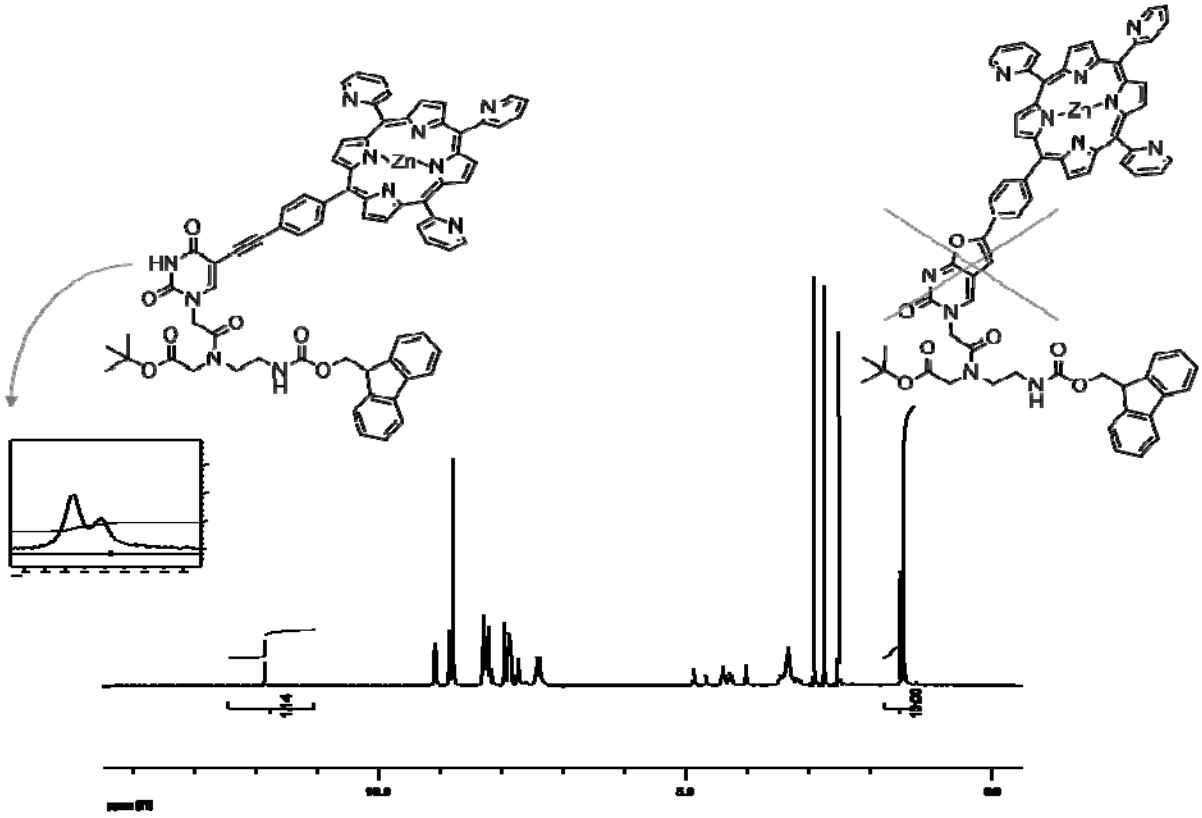


Figure 5.34. ^{1}H-NMR in DMSO of derivative 6. The presence of NH (splitted for the presence of two rotamers) in position 3 of uracil in right integration ratio with other protons exclude formation of cyclic product. Two molecule of DMF for one of compound are also present, probably attached to Zn porphyrin.

Since the synthesis of the Fmoc-monomer gave low yield, we explored the possibility to attach porphyrin on PNA directly by solid phase synthesis (*Figure 5.35*). Thus we synthesized Fmoc-AU^{I}CU^{I}AC-gly-(L-Lys)-NH_2 using PNA synthesis and Sonogashira coupling in solid phase, under the same reaction condition of S*cheme 5.5*.

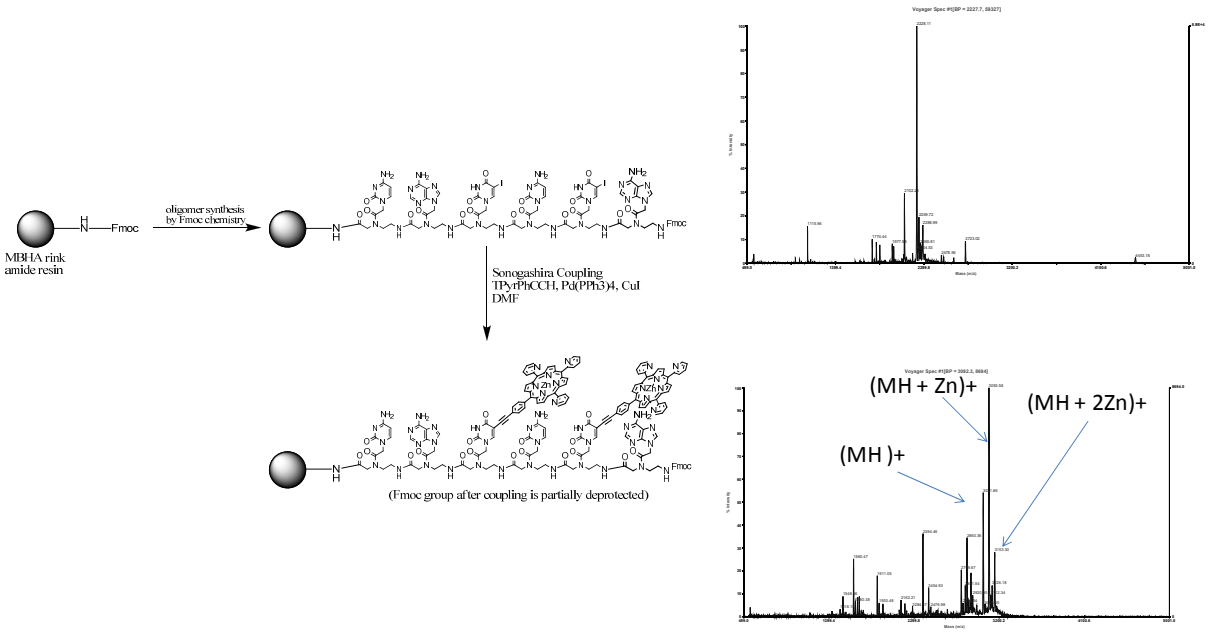


Figure 5.35. Synthesis of PNA bearing two porphyrins by Sonogashira on solid phase. Experimental conditions are showed in the experimental section.

5.5 Experimental Section

General Information. Reagents were purchased from Sigma-Aldrich, ASM Research, Fluka, Applied Biosystem and used without further purification. Instrumentation used are reported in Appendix 1. UV experiments were carried out on Perkin-Elmer (Norkalk, CT, USA) λ Bio 20 spectrophotometer equipped with temperature control system. Circular dichroism experiment were carried out using JASCO (Tokyo, Japan) J715 (equipped with Peltier temperature control system) and J810 spectropolarimeters. NMR were recorded using Bruker Avance 300 spectrometer.

Synthesis of TPyrPPCOOH. Synthesis of TPyrPPhPCOOH was carried out according to literature procedure[25].Yield: 4.9%. *^{1}H-NMR* (*300 Mhz, 25°C, $CDCl_3$*): (ppm) 9.16-9.13 (m, 3H, CH), 8.88-8.82, (m, 8H, CH pyrrole), 8.45 (d, 2H, *J*=8.4Hz, CH aryl), 8.30 (d, 2H, *J*=8.1Hz, CH aryl), 8.29-8.21 (m, 3H, CH pyridil), 8.11 (dt, 3H, *J*=7.5Hz, 1.8Hz, CH pyridil), 7.72 (m, 3H, CH pyridil), 4.11 (s, 1H, COOMe), -2.80 (s, 2H, NH core). ESI-MS(+) for $C_{43}H_{29}N_7O_2$ *m/z*: [MH^+] *Calcd.6*76.7, *Found* 676.6. Hydrolysis of methyl ester was carried out according literature procedure cited.[25] ^{1}H-NMR of the product perfectly matched.

PNA-Synthesis. Synthesis of PNA oligomer (Boc or Fmoc strategy) has been carried out in a fritted plastic syringe. MBHA rink-amide resin was used. Crude PNAs were purified by using RP-HPLC with C_{18} silica column, by gradient eluition from A (water + 0.1% TFA) to B (acetonitrile + 0.1% TFA) on Semipreparative Jupiter Phenomenex C_{18}. **Fmoc Strategy**. For each monomer elongation we employed the following module: a) Deprotection Fmoc in piperidine/DMF 20/80 for 8 min 2 times. b) Monomer activation with 5 eq (respect resin active sites) of PNA-monomer, 4.5 eq HBTU, 10 eq DIEA for 2 min. c) Coupling 45 min. d) Capping with DMF/DIEA/Ac_2O 89/6/5 1 min for 2 times. e) DIEA 10% washings 2 times for 2 min. After completion the PNA was cleaved in TFA/m-cresol 95/5 2 times for 1 h, then dried under a nitrogen flow, precipitated with ethyl ether and washed three times. **Boc Strategy**. For each monomer elongation we employed the following module: a) Deprotection Boc in TFA/m-cresol for 4 min 2 times. b) Monomer activation with 5 eq (based on resin active sites) of PNA-monomer, 4.5 eq HBTU, 10 eq DIEA for 2 min. c) Coupling 45 min. d) Capping with DMF/Pyridine/Ac_2O 25/25/1 1 min for 2 times. e) Piperidine 10% washings 2 times for 2 min. After completion the PNA was cleaved in TFA/TFMSA/Tioanisole/m-cresol 6/2/1/1 1

h for 2 times, then dried on nitrogen flow, precipitated with ethyl ether and washed three times.

Porphyrin conjugation. Porphyrin was attached to the deprotected terminal end of PNA using the following conditions: TPyrPhPCOOH (10 eq), PyBOP (9,5 eq), DIEA (20 eq) activation for 5 minutes, and coupling overnight.

P1P (TPyrPPC(O)NH-GTAGATCACT-gly-NH_2) MS (MALDI+) m/z for $C_{152}H_{164}N_{66}O_{32}$: $[MH^+]$ Calcd. 3424, Found 3421.Yield: not available.

P2 (H-GTAGATCACT-(L-Lys)-NH_2). Gradient eluition 0-5 min 100%A, 5-35min 60%A/40%B, ret. time 17.1 min. MS (ESI+) m/z for $C_{114}H_{148}N_{60}O_{31}$ (2854.8): $[MH_3]^{3+}$ Calcd. 952.6, Found 952.3, $[MH_4]^{4+}$ Calcd. 714.7, Found 714.6, $[MH_5]^{5+}$ Calcd. 572.0, Found 571.9, $[MH_6]^{6+}$ Calcd. 476.8, Found 476.7. Yield: 12%.

P2P (TPyrPPC(O)NH-GTAGATCACT-(L-Lys)-NH_2). Gradient eluition 0-5 min 95%A/5%B, 5-20min-50%A/50%B, ret. time 17.0 min. MS (ESI+) m/z for $C_{156}H_{173}N_{67}O_{32}$ (3498.8): $[MH_4]^{4+}$ Calcd. 875.7, Found 875.3, $[MH_5]^{5+}$ Calcd. 700.7, Found 700.6, $[MH_6]^{6+}$ Calcd. 584.1, Found 584.0, $[MH_7]^{7+}$ Calcd. 500.8, Found 500.8. Yield: 3.9%.

P4 (H-CATCTAGTGA-(gly)-NH_2). Gradient eluition 0-5 min 100%A, 5-35min 60%A/40%B, ret. time 17.2 min. Yield: 33%. MS (ESI+) m/z for $C_{110}H_{139}N_{59}O_{31}$ (2783.7): $[MH_4]^{4+}$ Calcd. 696.9, Found 696.7, $[MH_5]^{5+}$ Calcd. 557.7, Found 557.5, $[MH_6]^{6+}$ Calcd. 464.9, Found 464.6.

P4P (TPyrPPC(O)-NH-CATCTAGTGA-(gly)-NH_2). Gradient eluition 0-5 min 95%A/5%B, 5-25min-35%A/65%B, ret. time 16.8 min. Yield: 13.2%. MS (ESI+) m/z for $C_{152}H_{164}N_{66}O_{32}$ (3427.4): $[MH_3]^{3+}$ Calcd. 1143.5, Found 1143.4, $[MH_4]^{4+}$ Calcd. 857.9, Found 857.8, $[MH_5]^{5+}$ Calcd. 686.5, Found 686.4, $[MH_6]^{6+}$ Calcd. 572.2, Found 572.3.

UV and CD measurement. Sample were prepared for dilution of appropriate amounts of PNA stock solution with desired buffer. Concentration of PNA stock solutions were determined by measuring absorbance at 260 nm and using the following extinction coefficient ε ($M^{-1}cm^{-1}$): P2 (103500), P2P (135500), P4 (103500) and P4P (135500). Sample in glycerol/water 35/65 were prepared mixing stock solution of aqueous glycerol 51.55% w/w, phosphate buffer 40 mM, required amount of PNA stock solution and water (as needed to

reach desired volume). Sample in 50 mM $HCOONH_4$ buffer were prepared mixing buffer (100mM) with PNA sock solutions and water to desired final volume. All sample were incubated at 0°C and equilibrated at room temperature at least 4-5 h before measurement.

Tripyridyl-4-trimethylsilylethynilphenyl porphyrin (1). 2-Pyridinecarboxaldehyde (5.8 mL, 60.1 mmol) and 4-trimethylsililethynil benzaldehyde (4.2g, 20.0 mmmol) were dissolved in 500 mL of pre-dried propionic acid and refluxed for 20 min, then fresh distilled pyrrole (5.5 mL, 80.1 mmol) was added dropwise. After addition, the dark mixture was refluxed in the dark for 4 hours. Then, the propionic acid was distilled-off from the reaction. The TLC of the dark residue using ethyl acetate as a eluent showed six different red spot; the compound with rf of 0.34 was isolated using flash chromatography on silca gel to afford 710 mg (6.1%) of a purple solid. *TLC* on silca gel, eluent DCM – MeOH 20/1, *Rf:* 0.36.

(1) *^{1}H-NMR* (*400 Mhz, 25°C, $CDCl_3$*): (ppm) 9.12 (d, 3H, *J*=5.2 Hz, CH pyridil), 8.85 (d, 5H, *J*=6.4 Hz, CH pyrrole), 8.22 – 8.19 (m, 3H, CH pyrrole), 8.16 (d, 2H, *J*=8.0 Hz, CH phenyl), 8.06 (dd, 3H, *J*=7.6 Hz, 1.6 Hz, CH pyridil), 7.87 (d, 2H, *J*=8.0 Hz, CH phenyl), 7.70 – 7.76 (m, 3H, CH pyridil), 0.38 (s, 9H, Si(Me)$_3$), -2.8 (s, 2H, NH). *^{13}C-NMR* (*75 Mhz, 25°C, $CDCl_3$*): (ppm) 160.5, 148.6, 142.3 ???, 134.8, 134.5, 130.4, 131.1, 130.4, 130.3, 122.7, 122.4, 120.0, 118.8, 118.4, 104.9, 95.5, 0.1. HRMS (FAB$^+$) for $C_{46}H_{35}N_7Si$ *m/z*: [MH$^+$] *Calcd.* 714.2801, *Found* 714.2797. UV-Vis (MeOH) 417 nm (232500 $M^{-1}cm^{-1}$).

Tripyridyl-4-ethynilphenyl zinc porphyrin (2). Compound **1** (710 mg, 0.995 mmol) was dissolved in 80 mL of chloroform and 40 mL of methanol, then $Zn(OAc)_2$ (351 mg, 1.91 mmol) with few drop of Et_3N were added and the mixture was stirred at rt, under argon, for 24 h. The mixture was dried and filtered on a short silica gel column using DCM/MeOH 95/5 as eluent. The metallated product, after purification and drying, was dissolved in 100 mL of saturated K_2CO_3 in $CHCl_3$ / MeOH 3/2 and stirred at rt for 36 h. After this time, the reaction was dried and the residue separated by chromatography on silca gel using DCM/MeOH 95/5 to obtain 616 mg (92%) of a purple solid.

(2) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*): (ppm) 9.07 (d, 3H, *J*=4.5 Hz, CH pyridyl), 8.83 - 8.81 (m, 2H, CH pyrrole), 8.78-8.76 (m, *6H*, CH pyrrole), 8.30 – 8.19 (m, 8H, CH pyridil and phenyl), 7.92 (d, 2H, 7.8 Hz, CH phenyl), 7.86 – 7.82 (m, 3H, CH pyridil), 4.46 (s, 1H, CCH). *^{13}C-NMR* (*75 Mhz, 25°C, d_6-DMSO*): (ppm) 160.8, 149.2, 149.1, 148.9, 148.7, 148.0, 143.2, 134.8,

134.4, 131.5, 129.9, 129.6, 122.5, 120.9, 119.7, 119.4, 119.1, 114.0, 83.6, 81.8 69.6. HRMS (FAB^+) for $C_{43}H_{27}N_7Zn$ *m/z*: [MH^+] *Calcd.* 704.1541, *Found* 704.1519.

5-Iodo-*N(1)*-carboxymethyluracil *tert*-butyl ester (3). Iodouracil (2.0 g, 8.24 mmol) was dissolved in 20 mL of dry DMF and K_2CO_3 (1.16 g, 8.24 mmol) added. The reaction mixture was cooled at 0°C, then *t*-butyl bromoacetate (1.24 mL, 8.24 mmol) was added dropwise over 30 minutes at 0°C, and then stirred at rt overnight. The mixture was partitioned between ethyl acetate (100 mL) and water (100 mL), and the organic phase was washed four more times with water and once with saturated NaCl. Anhydrification and drying of organic phase, followed by recrystalizzation in ethyl acetate gave the desidered product (1.56 g, 53%) as a white solid. *TLC* on silca gel, eluent Ethyl acetate – hexane 1/1, *Rf:* 0.56.

(3) ^{1}*H-NMR* (*300 Mhz, 25°C, d_6-DMSO*): (ppm) 11.76, (s, 1H, NH), 8.20 (s, 1H, CH), 4.38 (s, 2H, CH_2), 1.42 (s, 9H, *tBu*). ^{13}*C-NMR* (*75 Mhz, 25°C, d_6-DMSO*): (ppm) 167.4, 161.5, 151.1, 150.6, 82.4, 68.4, 49.6, 28.1 HRMS (FAB^+) for $C_{10}H_{13}IN_2O_4$ *m/z*: [MH^+] *Calcd.* 352.9998, *Found* 352.9995.

5-Iodo-*N(1)*-carboxymethyluracil (4). Compound **3** was dissolved in 20 mL of TFA/DCM 1/1 at 0°C and stirred for 1 hour at rt. The solvent was removed under reduced pressure, and the resulting oil dissolved in DCM and evaporated several times in order to remove TFA. Then the oil was dissolved in the minimum amount of DCM and precipitated with hexane to give 1.31 g (96%) of product as a white solid.

(4) ^{1}*H-NMR* (*300 Mhz, 25°C, d_6-DMSO*): (ppm) 13.13, (s, 1H, *COOH*), 11.76 (s, 1H, *NH in 3*), 8.21 (s, 1H, *CH in 6*), 4.41 (s, 2H, *CH_2 in 1*). ^{13}*C-NMR* (*75 Mhz, 25°C, d_6-DMSO*): (ppm) 169.4, 161.1, 150.8, 150.3, 68.0, 48.6. HRMS (FAB^+) for $C_6H_5IN_2O$ *m/z*: [MH^+] *Calcd.* 293.9372, *Found* 296.9374.

Fmoc-5-Iodouracil-PNA monomer-OtBu (5). Compound **4** (0.71 g, 2.4 mmol), HOBt (0.32 g, 2.4 mmol) and DCC (0.48 g, 2.35 mmol) were dissolved at 0°C in 10 mL of DMF, then stirred for 2 hours at rt (after few minutes DCU precipitation is observed). Then, a solution of Fmoc-aeg-OtBu (0.80 g, 1.85 mmol) in 5 mL of DMF and 0.82 mL of DIEA were added, and the mixture stirred a rt for 5 hours. The reaction was mixed with 15 mL of EtOAc and filtered

in order to remove solid DCU. The filtrate after diluition with further 50 mL of EtOAc, was extracted twice with $NaHSO_4$ 0.1 M, with saturated $NaHCO_3$ and saturated. NaCl solutions. The organic layer after distillation gave a yellow oil that was purified on silica gel using EtOAc/Hexane 7/3 to afford 0.75 g (60%) of white foam.

(5) *^{1}H-NMR* (*300 MHz, 25°C, d_6-DMSO*): (ppm) 11.71(M) and 11.69(m) (s, 1H), 8.05(M) and 7.98(m) (s, 1H), 7.89 (d, 2H, *J*=7.2 Hz), 7.68 (d, 2H, *J*=7.5 Hz), 7.41 (t, 2H, *J*=7.5 Hz), 7.33 (t, 2H, *J*=7.2 Hz), 7.23 (t, 2H, *J*=6.0 Hz), 4.70 (M) and 4.51 (m) (s, 2H), 4.34(M) and 4.31(m) (t, 2H, *J*=6.6 Hz), 4.23 (t, 1H, *J*=6.3 Hz), 4.16 (m) and 3.94 (M)(s, 2H), 3.39(M) and 3.33(m) (t, 2H, *J*=6.3 Hz), 3.24(M) and 3.10(m) (q, 2H, *J*=6.3 Hz), 1.54(m) and 1.39(M) (s, 9H). *^{13}C-NMR* (*75 Mhz, 25°C, d_6-DMSO*): (ppm) 168.7, 168.4, 167.9(M) and 167.5(m), 161.5, 156.8(M) and 156.6(m), 151.1(M) and 151.0(m), 144.3, 141.2, 128.1, 127.5, 125.6, 120.6, 82.4(M) and 81.4(m), 68.2, 65.9(M) and 65.8(m), 60.2, 55.4, 49.3, 48.4(M) and 48.3(m), 47.4(M) and 47.2(m), 28.14(m) and 28.07(M). HRMS (FAB^+) for $C_{29}H_{32}O_7N_4I$ *m/z*: [MH^+] *Calcd.* 675.1316, *Found* 675.1324

Fmoc-5-Iodouracil-PNA monomer (6). Compound **6** was dissolved at 0°C in 20 mL of a mixture of DCM/TFA 1/1 and stirred for 1 hour, the solvent removed by distillation and the resulting oil dissolved with DCM and dried several times until complete elimination of TFA. The oil dissolved in DCM was precipitated with cold hexane to obtain 1.23 g (99%) of yellow-pale solid.

(6) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*): (ppm) 11.72(M) and 11.70(m) (s, 1H), 8.05(M) and 7.98(m) (s, 1H), 7.89 (d, 2H, *J*= 7.2Hz), 7.68 (d, 2H, *J*= 7.2Hz), 7.41 (t, 2H, *J*= 7.2Hz), 7.32 (t, 2H, *J*= 7.2Hz), 7.26 (t, 1H, *J*= 5.4 Hz), 4.71(M) and 4.53(m) (s, 2H), 4.35 – 4.19 (m, 5H), 3.40 – 3.09 (m, 4H). HRMS (FAB^+) for $C_{25}H_{24}O_7N_4I$ *m/z*: [MH^+] *Calcd.* 619.0690, *Found* 619.0726.

Fmoc-5-(tripyridil-4-ethynilphenyl zinc porphyrin)uracil-PNA monomer-OtBu (7). Compound **5** (0.064 mmol, 43 mg) and **2** (0.042 mmol, 30 mg) were dissolved in 2 mL of deoxygenate DMF, then CuI (0.014 mmol, 2,8 mg) and $Pd(PPh_3)_4$ (0.0075 mmol, 9 mg) were added. The resulting mixture after addition of Et_3N (12 µL) was stirred for 2,5 hours. The reaction mixture was partitioned between 10 mL of DCM and 10mL of saturated acqueous EDTA. The organic layer, after anhydrification with Na_2SO_4 was evaporated and the residue

purified by flash chromatography from DCM to DCM/MeOH 9/1. Pure product was obtained with a yield of 20%. *TLC* on silca gel, eluent DCM – MeOH 9/1, *Rf:* 0.30.

(7) *^{1}H-NMR* (*300 Mhz, 25°C, d_6-DMSO*): (ppm) 11.86(M) and 11.84(m) (s, 1H, NH uracil), 9.07 (d, 2H, *J*=3.6 Hz, CH pyridyl), 8.83 (t, 2H, *J*=4.8 Hz), 8.79 – 8.77 (m, 8H, CH pyrrole), 8.31-8.15 (m, 9H), 7.91-7.82 (m, 7H), 7.74-7.70 (m, 2H), 7.45 -7.29 (m, 5H), 4.85 (M) and 4.67(m) (s, 2H, CH_2 base linker), 4.45-4.33 (m, 2H), 4.29-4.23 (m, 2H), 3.49 – 3.38 (m, 2H), 3.21 -3.12 (m, 2H), 1.50(m) and 1.50(M) (s, 9H, tBu). MS (FAB^+) for $C_{72}H_{55}N_{12}O_7Zn$ *m/z*: [MH^+] *Calcd.* 1250.3, *Found* 1250.7.

Sonogashira Coupling on solid phase. PNA oligomers containing iodo uracil derivatives was synthesized according Fmoc strategy described above using compound **6**. 15 mg of MBHA rink amide resin (0.0035 mmol) preloaded with sequence Fmoc-A-U^I-C-U^I-AC-gly-(D-Lys)-gly-NH_2 were swollen for 2h in DCM, washed with dry DMF, then a solution of porphyrin and catalyst (composed by **2** (0,105 mmol), $Pd(PPh_3)_4$ (0,007 mmol), CuI (0.014 mmol), Et_3N (0.007 mmol) in 1 mL of DMF dry) was added and the resin shaked for 18h. Reaction and reagents additions were conducted under argon atmosphere, and solvents were purged with argon, in order to avoid contact with oxygen. After reaction the resin was washed many times with DMF, DCM and MeOH until filtrate remained colourless. Cleavage of resin was carried out using same protocol described above.

5.6 References

[1] A. K. Burrell, D.Officer, P. G. Plieger, D. C. W. Reid, *Chem. Rev.*, **2001**, *101*, 2751-2796;

[2] Zheng, J.-Y., Konishi, K., Aida, T., *Tetrahedron,* **1997**, *53*, 9115;

[3] Collman, J. P.; Wagenknecht, P. S.; Hutchison, J. E. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1537;

[4] (a) Gust, D.; Moore, T. A.; Moore, A. L., *Acc. Chem. Res.* **1993**, *26*,198; (b) Wasielewski, M. R. *Chem. Rev.* **1992**, *92*, 435;

[5] Gust, D., *Nature* **1997**, *386*, 21;

[6] I. Beletskaya, V. S. Tyurin, A. Yu. Tsivadze, R. Guilard, C. Stern, *Chem. Rev.*, **2009,** *109,* 1659–1713;

[7] Drain, C.M., Varotto, A., Radivojevic, I., *Chem. Rev*., **2009,** *109,* 1630–1658;

[8] A. Mammana, A. D'Urso, R. Lauceri, R. Purrello, *JACS*, **2007**, *129(26)*, 8062;

[9] X. Huang, K. Nakanishi, N. Berova, Chirality, **2000**, *12*, 237-255;

[10] N. Berova, G. Pescitelli, A. G. Petrovic, G. Proni, *Chem. Comm*., **2009**, *5958-5980;*

[11] T. Kurtan, N. Nesnas, Y. Q. Li, X. Huang, K. Nakanishi, N. Berova, *JACS*, **2001,** *123,* 5962-5973;

[12] S. Matile, N. Berova, K. Nakanishi, *JACS,* **1995**, *117*, 7021-7022;

[13] M. Balaz, K. Bitsch-Jensen, A. Mammana, G. A. Ellestad, K. Nakanishi, N. Berova, *Pure Appl. Chem.*, **2007**, *Vol. 79*(4), 801–809;

[14] R. Varghese, H. A. Wagenknecht, *Chem. Comm*., **2009**, 2615–2624;

[15] M. Balaz, M. De Napoli, A. E. Holmes, A. Mammana, K. Nakanishi, N. Berova, R. Purrello, *Angew. Chem. Int. Ed*., **2005**, *44*, 4006 –4009;

[16] A. D'Urso, A. Mammana, M. Balaz, A. E. Holmes, N. Berova, R. Lauceri, R. Purrello, *JACS*, **2009**, *131*, 2046–2047;

[17] M. Balaz, B. C. Li, J. D. Steinkruger, G. A. Ellestad, K. Nakanishi, N. Berova, *Org. Biomol. Chem.*, **2006**, *4*, 1865–1867;

[18] T. Nguyen, A. Brewer, E. Stulz, *Angew. Chem. Int. Ed.*, **2009**, *48*, 1974 –1977;

[19] In-depth information are reported in Introductive chapter;

[20] J. J. Storhoff, C. A. Mirkin, *Chem. Rev.,* **1999,** *99,* 1849-1862;

[21] Pei-Sze Ng, D. E. Bergstrom, *Nano Letters*, **2005**, *5(1)*, 107-111;

[22] P. S. Lukeman, A. C. Mittal, N. C. Seeman, *Chem. Comm.*, 2004, 1694 – 1695;

[23] T. Tedeschi, S. Sforza, A. Dossena, R. Corradini, R. Marchelli, *Chirality*, **2005**, iss. *17,* p. 196–204;

[24] P. Wittung, M. Eriksson, R. Lyng, P. E. Nielsen, B. Norden, *JACS*, **1995**, *117*(*41*), 10167-10163;

[25] P. Bigey, S. H.Sonnichsen, B. Meunier, P. E. Nielsen, *Bioconjugate Chem.*, **1997,** *8*, 267-270;

[26] M. Balaz, A. E. Holmes, M. Benedetti, G. Proni, N. Berova, *Bioorg. & Med. Chem.*, **2005**, *13*, 2413–2421;

[27] A. Mammana, T. Asakawa, K. Bitch-Jensen, A. Wolfe, S. Chaturantabut, Y. Otani, X. Li , Z. Li, K. Nakanishi, M. Balaz, G. A. Ellestad, N. Berova, *Bioorg. & Med. Chem.*, *16,* **2008**, 6544–6551;

[28] Yingfu Li, D. Sen, *Nat. Struct. Biol.*, **1996**, *3(9)*, 743-747;

[29] J. Szyrwiel, P. Młynarz, H. Kozłowski, M. Taddei, *J. Chem. Soc*., *Dalton Trans*., **1998**, p. 1263–1264;

[30] G. D. Dorough,J. R. Miller, F. M. Huennekens, *JACS*, **1951**, *73*, 4315;

[31] M. Goodman, M. Chorev, *Acc. Chem. Res.,* **1979,** *12,* 1 – 7;

[32] J. O. Smith, D. A. Olson, B. A. Armitage, JACS, **1999**, *121*, 2686 – 2695; See also introductive chapter and references cited therein for more information on exciton coupling methods.

[33] S. Sforza, G. Haaima, R. Marchelli, P. E. Nielsen, *Eur. J. Org. Chem.* **1999**, 197-204;

[34] G.Xie, S. N. Timasheff, *Protein Science*, **1997**, *621,* 1-221;

[35] T. Arakawa, S. N. Timasheff, *Biophysical J.*, **1985**, *47*, 411-414;

[36] F. Totsingan, *Synthesis and Application of PNA and Modified PNA in Nanobiotechnology*, **2007**, PhD thesis, University of Parma;

[37] Y. Ohya, K. Yabuki, T. Ouchi*, Supramolec. Chem.*, **2003**, *15* (*2*), 149–154;

[38] S. Narayanan, J. Gall, C. Richert, *Nucleic Acids Res.*, **2004**, *32* (*9*), 2901-2911;

[39] M. Balaz, B. C. Li, S. Jockusch, G. A. Ellestad, N. Berova, *Angew. Chem. Int. Ed.*, **2006**, *45*, 3530 –3533;

[40] A. Mammana, A. D'Urso, R. Lauceri, R. Purrello, *JACS*, **2007**, *129*, 8062-8063;

[41] A. Satake, Y. Kobuke, *Org. Biomol. Chem.*, **2007**, *5*, 1679–1691;

[42] K. Kano, K. Fukuda, H. Wakami, R. Nishiyabu, R. F. Pasternack, *JACS*, **2000,** *122,* 7494-7502;

[43] C.A. Hunter, J. K. M. Sanders, *J. Am. Chem.* Soc.,**1990**, *112(14),5525;*

[44] B. D. Gildea0, S. Casey, J. MacNeill, H. Perry-O'Keefe, D .Sarensen, J. M. Coull, Tetrahedron Lett., **1998**, 39, 7255-7258;

[45] E. Masayuki Endo, M. Fujitsuka, T. Majima, *Tetrahedron* , **2008**, *64*, 1839-1846;

[46] P. Leighton, J. Cowan, R. J. Abraham, J.K. M. Sanders, **1998**, *J. Org. Chem*., *53(4)*, 733-740;

[47] Rasmussen H, Sandholm J, Nat. Struct. Biol.**, 1997**, 4(2), 98-101;

[48] M. W. Konrad, J. Bolonick, JACS, **1996,** 118, 10989-10994;

[49] U. Diederichsen, ***Angew. Chem. Int. Ed.*** *Engl.* **1996.** *35(4)*, 445-448;

[50] J. A. Cuesta-Seijo, G. M. Sheldrick, J.Zhang, U. Diederichsen, *Protein data Bank*, *Code 3C1P*;

[51] Desmond developed by D. E. Shaw Inc., New York NY 10036, (USA), 2008;

[52] G. C. Shields, C. A. Laughton, M. Orozco, *J. Am. Chem. Soc.*, **1998**, *120*, 5895-5904;

[53] Protein Data Bank, Structure 1PUP.

Acknowledgements

First of all I would like to express my gratitude to Prof. Roberto Corradini and Prof. Rosangela Marchelli for giving me the opportunity to work in their research group and for their helping during the writing up of this thesis. Especially, I would to thank Prof. R. Corradini for his constant helping and support during this three years, moreover for the immensurable knowledge that I learned from him.

I am thankful to Prof. S. Sforza and Dr. T. Tedeschi for their helpful discussion and to all colleagues of 50's and neighbor's labs.

I greatly would to thank Prof. Nina Berova and Prof. Koji Nakanishi for allowing me to joint their group and to work in one of the most exciting environment. I am very thankful to Dr Ana Petrovic for her scientific support during my time at Columbia.

Prof. R. Gambari and his group are also acknowledged for their experiments on uracil dimers.

PAGE LEFT INTENTIONALLY BLANK

Bio-CV

Personal Information:

Alessandro Accetta - *date of birth* 1983/01/13 - *place of birth* Canosa di Puglia (BT) – *address* Via Contrada Pantanelle n.138, Canosa di Puglia(BA) – *current address* Via Manzoni 2, Parma.

Educations:

- 2007 – today: PhD student under supervision of Prof. R. Marchelli and Prof. R. Corradini
- Master Degree in Chemistry *cum laude* at the University of Parma in July 2006;
- Bachelor degree in Chemistry cum laude at the University of Parma in October 2004;
- High School diploma, votation 100/100, in July 2001 at IILE (Canosa di Puglia, Bari).

Contact:
alessandro.accetta@gmail.com

Contributions:

Pubblications:

- A. Accetta, R. Corradini, S. Sforza, T. Tedeschi, E. Brognara, M. Borgatti, R. Gambari, R. Marchelli; ***New Uracil Dimers Showing Erytroid Differentiation Inducing Activities***; *J. Med. Chem.,* **2009**, 52(1), 87-94;
- Roberto Corradini, Tullia Tedeschi, Stefano Sforza, Alessandro Calabretta, Alessandro Accetta, Alessandro Tonelli, Rosangela Marchelli. ***Peptide nucleic acids (PNAs) and modified PNAs for DNA targeting*****,** *Minerva Biotechnologica***, 2009,** *20*(*4*), pp 87-94.
- Six poster communications and one oral communication in National and International symposium.