GAMMA-PNA CHEMICAL BUILDING BLOCKS: METHODOLOGY DEVELOPMENT, SYNTHESIS AND APPLICATION

Arunava Manna

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

© Copyright by Arunava Manna 2015
All rights Reserved
Department of Chemistry
Carnegie Mellon University
ABSTRACT

GAMMA PNA BUILDING BLOCKS:
METHODOLOGY DEVELOPMENT, SYNTHESIS
AND APPLICATION

By

Arunava Manna
Carnegie Mellon University
Department of Chemistry

The research described in this thesis focuses on the synthesis of chemical building blocks of an important class of nucleic acid mimic, called gamma peptide nucleic acid (γPNA). γPNA offers an attractive nucleic acid platform for biological and biotechnological applications, as molecular tools for basic research as well as possible therapeutic and diagnostic reagents. These studies were undertaken in attempts to determine the optimal route for preparing γPNA monomers on large-scale and with high optical purity, suitable for automated solid-phase oligo synthesis.

Chapter I provide a historical perspective on the field. In Chapter II we compare the optical purity of γPNA monomers prepared by various synthetic routes. Out of the various methods examined, the Mitsunobu coupling scheme provides the highest optically purity. We used $^{19}$F-NMR spectroscopic tool to determine the optical purity of the monomers. A detailed, systematic study was performed to determine the source and the extent of epimerization in each of the synthetic routes. The effects of epimerized monomers on the helical conformation as well as the binding affinity of γPNA oligomers were also investigated.
In Chapter III we describe the development of a new synthetic scheme for preparing \((R)\)-Fmoc MP\(\gamma\)PNA monomers. In order to expedite the exploration of \(\gamma\)PNA as molecular tools and reagents for basic research in biology and medicine, an automated, parallel solid-phase synthesis of \(\gamma\)PNA oligomers is necessary. However, the current Boc solid-phase synthesis (SPS) strategy is not suitable for automation. In this work we report the first synthetic methodology for synthesizing optically-pure \((R)\)-Fmoc-MP\(\gamma\)PNA monomers based on the Mitsunobu coupling strategy.

In Chapter IV we describe a new synthetic route for preparing both \((R)\)- and \((S)\)-Fmoc MP\(\gamma\)PNA monomers starting from a commercially available and relatively cheap L-Serine. Even though the Mitsunobu route to synthesizing \((R)\)-Fmoc-MP\(\gamma\)PNA monomers is a rather simple, it is not suitable for large scale synthesis. Recognition of this obstacle, we devised a simple but yet robust procedure for synthesizing optically-pure \((R)\)- and \((S)\)-Fmoc-MP\(\gamma\)PNA chemical building blocks. The synthetic flexibility and versatility of this methodology is also amendable for preparing \((S)\)-Fmoc-MP\(\gamma\)PNA monomers, using the same starting material. Owing to its recognition orthogonality with the right-handed conformer, the left-handed \(\gamma\)PNA offers an attractive and unique platform for organizing molecular self-assembly, especially for in vivo applications where enzymatic degradation and inadvertent binding hybridization of probes to the host’s genetic materials (i.e. DNA and RNA) are of concern.

The methodology development in the synthesis of \(\gamma\)PNA monomers described in this thesis could have a significant impact on the overall research of
γPNA. The findings on the optical purity of the monomers and their effects on binding affinity have important implication not only on the synthesis of γPNA building blocks but also on the development of other chiral reagents. The work described herein on the development of a new synthetic route for preparing (R)- and (S)-Fmoc MPγPNA monomers will pave the road for future development and exploration of γPNA in biology and medicine.
In the sweet memory of my mother Anima Manna, who will always be great example of strength and courage.
ACKNOWLEDGMENTS

First of all, I want to thank Professor Danith Ly for his encouragement, and constant support to do my best during the last five years. His passion, eagerness for science, provided me great amounts of motivation and determination. His high expectations made me work hard and would like to thank him for bringing out the best from me. Secondly, I want to thank Professor Newell Washburn and Professor Subha Das for their guidance, valuable feedback, and constant encouragements over the last five years. Also, I would like to thank Professor Rita Mihailescu for her support and advice. I want to thank Professor Rina Ghosh (Department of Chemistry, Jadavpur University), who introduced me to research field when I was an undergraduate student.

I want to thank our collaborators Professor James Schneider (Department of Chemical Engineering, CMU), for the project on nucleic acid detection and Professor Peter Glazer (Department of Therapeutics, Yale University), for the work on gene editing. I want to thank Professor Terry Collins and Professor Dennis Curran (Department of Chemistry, University of Pittsburgh) for allowing me to use the instrumental facilities at their laboratories.

I really appreciate the support and benevolence of my friends and colleagues: Dr. Srinivas Rapireddy, Dr. Matthew Crawford, Dr. Iulia Sacui, Dr. Raman Bahal, Dr. Suresh Gopalswamy, Dr. Dinesh Bhunia, Dr. Debasish Grahacharya and Weiche Hsieh. I thank all of our group members for their help. Thank you for being there for me. Also, I want to acknowledge all the undergraduate students that I have been interacting with during the last years, especially Allison Huang and Yidan Cong who participated in some of the projects.

I am also very grateful to the DSF Charitable Foundation for my fellowship and also to the Department of Chemistry, CMU.

I want to thank my friends here at CMU for their help whenever I needed any and special thanks to friends: Subhajit Debnath, Sumit, Debasis Dey and Souvagya for being there.

Most importantly, I want to thank my parents Biswanath Manna and Anima Manna. I am what I am due to their unconditional love and support. I pay pranam to them for their sacrifices and affections. I want to thank my beloved brother Amitava and sisters Juli and Mili for always being there for me and their unconditional love.
GAMMA-PNA CHEMICAL BUILDING BLOCKS:
METHODOLY DEVELOPMENT, SYNTHESIS AND APPLICATION

Abstract ii
Dedication v
Acknowledgments vi
Table of Contents vii
List of Tables ix
List of Figures x

CHAPTER I.  INTRODUCTION 1

I.1. Historical Perspective 1
I.2. Antisense Oligonucleotides 2
I.3. Antigene Oligonucleotides 6
I.4. Peptide Nucleic Acid 11
I.5. Thesis Overview 20
I.6. References 21

CHAPTER II.  A COMPARATIVE STUDY OF THE OPTICAL PURITY OF \( \gamma \)PNA MONOMERS BY DIFFERENT SYNTHETIC ROUTES 31

II.1. Introduction 31
II.2. Results and Discussion 32
  II.2.1. Scheme of different synthetic routes of \( \gamma \)PNA monomers 33
  II.2.2. Determination of the optical purity by \( ^{19} \)F NMR 36
  II.2.3. Detailed analysis of reductive amination route 39
  II.2.4. Effect of epimerization on conformation of PNAs 43
  II.2.5. Effect of epimerization on binding affinity of PNAs towards complementary DNA 46
II.3. Conclusion 47
II.4. Experimental section 48
II.5. References 63
II.6. Appendices 65

CHAPTER III.  SYNTHESIS OF FMOC (R) - \( \gamma \)MP PNA MONOMERS 86

III.1. Introduction 86
III.2. Design rationale 87
III.3. Results 88
III.4. Discussion 96
### III.5. Conclusion

97

### III.6. Experimental section

98

### III.7. References

109

### III.8. Appendices

112

### CHAPTER IV. A GENERAL METHOD FOR THE SYNTHESIS OF OPTICALLY PURE FMOC (R), (S) - γMP PNA MONOMERS

128

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV.1. Introduction</td>
<td>128</td>
</tr>
<tr>
<td>IV.2. Design rationale</td>
<td>129</td>
</tr>
<tr>
<td>IV.3. Results</td>
<td>130</td>
</tr>
<tr>
<td>IV.3.1. Synthesis of the (R) Fmoc γMP PNA monomers</td>
<td>131</td>
</tr>
<tr>
<td>IV.3.2. Synthesis of the (S) Fmoc γMP PNA monomers</td>
<td>136</td>
</tr>
<tr>
<td>IV.4. Discussion</td>
<td>139</td>
</tr>
<tr>
<td>IV.5. Conclusion</td>
<td>141</td>
</tr>
<tr>
<td>IV.6. Experimental section</td>
<td>142</td>
</tr>
<tr>
<td>IV.7. References</td>
<td>165</td>
</tr>
<tr>
<td>IV.8. Appendices</td>
<td>166</td>
</tr>
</tbody>
</table>

### APPLICATIONS

195

### SUMMARY

198

### FUTURE DIRECTIONS

200
CHAPTER II. A COMPARATIVE STUDY OF THE OPTICAL PURITY OF γ-PNA MONOMERS BY DIFFERENT SYNTHETIC ROUTES

Table II-1. Sequence of PNA oligomers.
LIST OF FIGURES

CHAPTER II. A COMPARATIVE STUDY OF THE OPTICAL PURITY OF γPNA MONOMERS BY DIFFERENT SYNTHETIC ROUTES

Figure II-1. 19F-NMR of MTPA-derivatized (A) D- and L-alanine starting materials, and (B) alanyl-γPNA backbones prepared from the reductive amination and Mitsunobu routes. 38

Figure II-2. 19F-NMR of MTPA-derivatized alanyl-γPNA T-monomers prepared from the reductive amination and Mitsunobu routes. 39

Figure II-3. 19F-NMR of MTPA-derivatized D- and L-alanol intermediates. 41

Figure II-4. 19F-NMR spectra of MTPA-derivatized alcohols prepared by the Swern and Parikh-Doering oxidation, and the Weinreb amide reduction route, followed by reduction of aldehydes to alcohols, Boc deprotection and coupling to MTPA-Cl. 43

Figure II-5. CD spectra of single-strand unmodified PNA and alanyl-γPNAs (Inset) UV-Vis spectra of the indicated oligomers at 90 °C showing that their concentrations were the same at 260 nm. 45

Figure II-6. UV-melting profiles of PNA-DNA hybrid duplexes. 47

CHAPTER III. SYNTHESIS OF FMOC (R) - γMP PNA MONOMERS

Figure III-1. Top: 1H NMR of pure Mitsunobu product 4; Bottom: 1H NMR of impure Mitsunobu product 4. 90

Figure III-2. Top: 1H NMR of SM 10; Bottom: 1H NMR of compound 11. 93

CHAPTER IV. A GENERAL METHOD FOR THE SYNTHESIS OF OPTICALLY PURE FMOC (R), (S) - γMP PNA MONOMERS

Figure IV-1. 1H NMR spectrum of compound 4. 132

Figure IV-2. 1H NMR spectrum of compound 6. 134

Figure IV-3. 1H NMR spectrum of compound 1a. 135
CHAPTER I
INTRODUCTION

I.1. Historical Perspective

In nature there are three classes of natural biopolymers: nucleic acids, proteins and carbohydrates. Among them, nucleic acids are the ones that provide the means for storage and transmission of genetic information. This realization has spurred a pursuit to determine the structure of DNA and elucidate the mechanism(s) by which hereditary information is stored and transmitted from one generation to the next. A major breakthrough came in 1953, when Watson, Crick, Wilkins and Franklin unveiled the structure of DNA.\(^1,2\) Through molecular modeling and with the support of Franklin’s x-ray crystal structure finding, they proposed a double helical structure in which the two complementary strands are held together in an antiparallel fashion (5’-3’/3’-5’) through hydrogen-bonding interactions between adenine (A) and thymine (T) and between cytosine (C) and guanine (G) on the opposing strands.\(^3,4\) The binding free energy arises from hydrogen-bonding as well as base-stacking, while the recognition specificity emanates from the base-pairing interactions. For the next few decades much of the research effort has been focused on determining the structures and biological functions of this and related nucleic acid biopolymers (RNA). This effort has led to the formulation of the central dogma of molecular biology, outlining the flow of genetic information from DNA to RNA to protein.\(^5\)

From these earlier studies two important properties emerged that enable DNA to perform its unique physiological functions: the specificity and reversibility of hydrogen-
bonding interactions between the complementary nucleobases in accordance with the Watson-Crick base-pairing rules (A pairs with T and C with G). With this knowledge, researchers envisioned developing synthetic oligonucleotides that can recognize the complementary DNA/RNA targets based on the same Watson-Crick base-pairing recognition principles. The focus of developing synthetic oligonucleotides has been to control gene expression, with the purpose of using them as molecular tools as well as potential therapeutic and diagnostic reagents for the treatment and detection of genetic diseases, particularly RNA because of the ease of accessibility. In parallel, there is an ongoing effort toward developing general principles for targeting double-stranded DNA (dsDNA). Over the last few decades, several such classes of synthetic oligonucleotides have been developed, largely driven by the desire to gain additional binding affinity and specificity in order to improve the efficiency and selectivity of gene targeting, and to minimize the cytotoxic effect due to off-target binding.

I.2. Antisense Oligonucleotides

Oligonucleotides, which are relatively short DNA or RNA strands (or synthetic derivatives thereof), typically in the order of 10-30 nucleotides in length, have been extensively utilized to regulate gene expression by suppressing protein synthesis through mRNA targeting in both cell culture and animal models. Such oligonucleotides are called antisense reagents. There are mainly three possible routes by which oligonucleotide molecules disrupt protein biosynthesis at the mRNA level: RNase induced RNA cleavage, inhibition of the translation initiation complex formation, and steric blockage of translation. These modes of actions depend on the nature of the oligonucleotides used and the position of the chosen target region within the mRNA.
1.2.1. First-Generation Oligonucleotides

The concept that protein synthesis could be suppressed by using exogenous nucleic acids was initially suggested by Paterson and coworkers. They showed that single stranded DNA could be used to inhibit translation of RNA. This concept was later applied to cell culture by Zamecnik and Stephenson in the late 1970s, using the nucleotide sequences from 5’- and 3’-ends of 35S RNA of Rous sarcoma virus (RSV). They synthesized a 13-mer oligonucleotide d(5’-AATGGTAAAATGG-3’), complementary to a portion of this viral sequence. When this synthetic oligonucleotide sequence was introduced into cultured fibroblast cells infected with RSV, viral production was significantly inhibited. They thus concluded that the oligonucleotide was inhibiting viral integration by hybridizing to the crucial RNA sequence and blocking protein translation.

At the same time, Tennant and Miller also reported similar findings with different systems. Subsequently, several more studies were conducted, but most of them failed to achieve the desired antisense effect. Learning from past failures, in the late 1960s Eckstein and coworkers developed the first-generation antisense reagent—phosphorothioate (Scheme 1). This was accomplished by replacing one of the non-bridging oxygens in the phosphate backbone with a sulfur atom to circumvent enzymatic degradation in the biological systems. Owing to their attractive pharmacokinetic properties, phosphorothioates have been successfully used in a large number of experiments to downregulate gene expression. However, they are not without limitations. Some of the drawbacks of phosphorothioates include low binding affinity for

![Scheme 1](image-url)

Scheme 1. Chemical structure of the first-generation phosphorothioates.
DNA as well as RNA,\textsuperscript{21} and a propensity to interact with proteins,\textsuperscript{22} preventing them from freely diffusing in the cytoplasm and nucleus and from binding to their targets.

I.2.2. Second-Generation Oligonucleotides

The problems associated with phosphorothioates were partially solved by the second-generation oligonucleotides. They consisted of RNA oligonucleotides with alkyl modifications at the 2’-position of the ribose sugar (2’-O-methyl\textsuperscript{23} and 2’-O-methoxy-ethyl RNA,\textsuperscript{24} Scheme 2). Compared to phosphorothioates, these oligonucleotides showed slightly enhanced binding affinity, antisense activities through an RNase-independent pathway and are less toxic.\textsuperscript{25,26} Because of these appealing features, they are extensively used in basic research. In order to further improve these properties, such as binding affinity, specificity, enzyme stability and pharmacokinetics, a diverse set of modified oligonucleotides has been developed, and they are often dubbed the “third-generation oligonucleotides.”

I.2.3 Third-Generation Oligonucleotides

The third-generation oligonucleotides consist of various modifications on the structure of the backbone. Replacement of the 3’-OH group of the 2’-deoxyribose ring by 3’-amino group has led to the development of phosphoramidate oligonucleotides (Scheme 3).\textsuperscript{27} Their high affinity towards complementary RNA and resistance to nucleolytic degradation makes them attractive as antisense reagents for \textit{in vivo}
applications. Another analogue is 2’-deoxy-2’-fluoro-β-D-arabino nucleic acid (FANA), where the 2’-OH group is replaced by a fluorine atom. Cyclohexyl nucleic acid (CeNA) is another example of the third-generation oligonucleotide where the furanose ring is replaced by a six-membered unsaturated ring. It exhibits high degree of conformational rigidity and enhanced RNA binding affinity while retaining the RNase H activity. An elegant modification of the pentose sugar of RNA, achieved by connecting the 2’-oxygen and the 4’carbon via a methylene linkage, has led to the development of locked nucleic acid (LNA). LNA was first developed by Wengel and Imanishi. It exhibits unusually high thermodynamic stability and recognition specificity for RNA as well as DNA. Because of its unprecedented properties, LNA is widely used in a number of biological applications including antisense, either exclusively or as LNA-DNA chimeras.

Another type of modified oligonucleotides that is water soluble and enzymatically stable in biological systems is morpholino phosphorodiamidate (MF). MF is made by replacing the ribose sugar of the oligonucleotide backbone with a morpholine moiety, and the phosphodiester linkage between the nucleotide with a phosphorodiamidate linkage. MF has excellent antisense properties including high binding affinity, antisense activity, high selectivity and enzymatic stability in biological systems. MF is extensively used in genetic studies in zebrafish, sea urchin and Xenopus embryos. However, efficient cellular delivery of MF for therapeutic applications remains a challenge.
I.3. Antigene Oligonucleotides

Another way to regulate gene expression is to target dsDNA in a sequence specific manner. This method is called antigene strategy. It exploits exogenous nucleic acid analogues that bind dsDNA and inhibit gene expression at the transcriptional level. However targeting double helical DNA is a daunting task. Firstly, in order to reach the dsDNA target, an oligonucleotide would have to penetrate both the extracellular and nuclear membrane. Secondly, the target might not be accessible because the dsDNA remains associated with DNA binding proteins or it may be packaged in a nucleosome. However, several classes of antigene reagents have been developed that can selectively inhibit gene expression, one of which is triplex binding oligonucleotides.

**Scheme 3.** Chemical structures of the third-generation oligonucleotides.
Scheme 4. Triple-helix formation in case of pyrimidine motif.

Scheme 5. Triple-helix formation in the case of purine motif.
I.3.1. Triple-Helix Forming Oligonucleotides

Triplex-forming oligonucleotides (TFOs) are major groove binding ligands that target unique DNA sequences by forming DNA triple helices by virtue of specific hydrogen bonding interactions between the TFO and the complementary strand of the duplex DNA. TFOs were shown to be effective in inhibiting the action of various DNA polymerases and hamper DNA synthesis due to the formation of triple helices in a cell free system. Subsequently such studies were also reported in living cells. TFOs prevent DNA transcription by various mechanisms depending on the location of the target site. However, triplex formation is limited to polypurine/polypyrimidine regions of dsDNA. In case of the purine motif, TFOs bind in an antiparallel fashion by forming reverse Hoogsteen hydrogen bonds with the purine strand of dsDNA. However in the case of the polypyrimidine motif, TFOs bind in a parallel fashion with respect to the polypurine strand of dsDNA. The low efficiency of TFO-induced activities in cell has been overcome by conjugating intercalators to the termini of the oligonucleotides. Even though the use of TFO is limited to the presence of polypurine/polypyrimidine sequences of the DNA target, they have a wide range of applications—for instance, in the regulation of gene expression, site-directed mutagenesis, recombinant DNA, and molecular therapy for treatment of genetic diseases such as sickle cell anemia.

I.3.2. Polyamides

Another class of DNA binding molecules includes polyamides that bind in the minor groove of DNA double helix. Polyamides, developed by Dervan and coworkers, can be designed to recognize DNA segments containing all four possible Watson–Crick base-pairs with affinities and specificities compared to those of DNA-binding
proteins. The main recognition elements are hydroxypyrrole (Hp), imidazole (Im), and pyrrole (Py). These polyamides bind in a 2:1 (PA/DNA) complex. To minimize the entropic penalty associated with such a binding evening, most polyamides are typically designed as unimolecular hairpins. The recognition domains stack with each other, which can specifically recognize both sides of the Watson-Crick base pairs. Polyamides can suppress gene transcription by interfering with the binding of natural transcription factors to DNA. Thus, they are promising molecules for gene targeting; however, their in vivo applicability remains to be established. The initial structural drawback was that the recognition sequence cannot be extended beyond five base pairs. Now the recognition length has been increased up to eleven base pairs.

I.3.3. Zinc-Finger Finding Proteins

While working on the 5S RNA genes of *Xenopus laevis*, which are transcribed by RNA polymerase III, Roeder and Brown found that the correct initiation of gene transcription requires the binding of a 40 kDa protein factor called Transcription Factor IIIA (TFIIIA). Studies of TFIIIA binding have led to the discovery of a remarkable repeating motif within the protein, which was later called zinc-finger binding motif because it contains a zinc (Zn) atom coordinated by two adjoining cysteine and two histidine residues, which folds into a “finger-like” DNA recognition module. Extensive work has been done on the characterization and structure elucidation of zinc-finger proteins. Zinc-finger proteins could bind any sequence of DNA of different lengths via this novel principle of recognition. For this reason, the zinc-finger motifs became ideal natural building blocks for the de novo design of proteins for recognizing any given sequence of DNA, for instance, for application in gene correction by
homologous recombination. However, the secondary effect as the result of off-target binding has yet to be resolved.\(^{76}\)

**I.3.4. Transcription Activator-Like Effector Nucleases (TALENs)**

Another naturally occurring protein that can recognize and bind DNA is TALE. These proteins are found in the plant pathogenic bacteria genus Xanthomonas. TALE proteins contain DNA-binding domains composed of a series of 33–35-amino-acid repeat domains, each capable of recognizing a single base-pair.\(^{77-79}\) This recent discovery of a simple modular DNA recognition code by TALE proteins has led to the explosive expansion of an alternative platform for engineering programmable DNA-binding proteins.\(^{80-82}\) Like zinc-finger proteins, modular TALE repeats are linked together to recognize contiguous DNA sequences. However, in case of TALE proteins in the construction of long arrays for targeting a single site in the genome, no re-engineering of the linkage is required as in the case of the zinc-finger proteins. Even though TALE–DNA binding repeats afford greater design flexibility than triplet-confined zinc-finger proteins, there is a targeting limitation for TALE arrays. TALE binding sites should start with a T base.\(^{77}\) Another problem that needs to be resolved is the efficient delivery of these nucleases into cells and organisms. The large size of TALENs may limit their delivery by size-restricted vectors such as recombinant adeno-associated virus (AAV), which has been shown to accommodate ZFN genes.\(^{83,84}\) These findings suggest that the development of new TALEN delivery systems will be a critical area of future research.

**I.3.5. CRISPR**

Recently targeted genome editing using engineered nucleases has been largely fueled by the advent of clustered, regularly interspaced, short palindromic repeat
(CRISPR) technology, an important new approach for generating RNA-guided nucleases, such as Cas9, with customizable specificities.\textsuperscript{85,86} In bacteria, the CRISPR system provides acquired immunity against invading foreign DNA via RNA-guided DNA cleavage.\textsuperscript{87,88} Genome editing mediated by these CRISPR-Cas9 systems has been used to rapidly, easily and efficiently modify endogenous genes in a wide variety of biomedically important cell types and in organisms that have traditionally been challenging to manipulate genetically.\textsuperscript{89} Recent work has shown that target recognition by the Cas9 protein requires a ‘seed’ sequence within the crRNA and a conserved dinucleotide-containing protospacer adjacent motif (PAM) sequence upstream of the crRNA binding region.\textsuperscript{90} Although CRISPR/Cas systems show great promise and flexibility for genetic engineering, sequence requirements within the PAM sequence may constrain some applications. Nonetheless, together these technologies promise to expand our ability to explore and alter any site within the genome and constitute a new and promising paradigm for understanding and treating genetic diseases.

I.4. Peptide Nucleic Acid

Peptide nucleic acid (PNA) is a synthetic analogue of DNA and RNA, first reported by Neilsen and co-workers in which the sugar phosphodiester backbone was replaced by achiral $N$-(2-aminoethyl)glycine units.\textsuperscript{91} A salient feature of PNA is its tight binding with DNA as well as RNA, in accordance with the Watson-Crick base-pairing rules.\textsuperscript{92} The higher binding affinity of PNA can be attributed in part to the charge-neutral pseudopeptide backbone, which mitigates electrostatic repulsion between the two opposing strands upon binding.\textsuperscript{92} Another important feature of PNA is that recognition is highly sequence-specific.\textsuperscript{93,94} An additional advantage of having a pseudopeptide
backbone is that PNA is resistant to enzymatic degradation by proteases and nucleases,\textsuperscript{95,96} and is chemically stable, as compared to the natural DNA or RNA counterpart. These appealing features make PNA an attractive tool for various biological and biomedical applications. Since its initial development more than two decades ago, extensive chemical modifications have been made to PNA in attempts to further improve its binding affinity, sequence-specificity, water solubility, cellular uptake, antisense and antigenic activities.

\begin{center}
\includegraphics[width=0.6\textwidth]{Scheme6.png}
\end{center}

\textbf{Scheme 6.} Structure of peptide nucleic acid and various backbone modifications.

\textbf{I.4.1.} \textit{α-} and \textit{β-}Backbone Modifications

Most of the structural modifications on PNA were made to the backbone as well as the nucleobases, but predominantly the former. Initial structural exploration revealed that only certain alterations of the original structure are possible without significantly affecting the binding affinity and/or sequence-specificity. Much of the initial backbone modifications were made at the \textit{α-}backbone. Neilsen reported the first chiral \textit{α}PNA in
1994, in which the glycine moiety was replaced by alanine. Subsequently, additional modifications at the α-position were made by introducing other chemical functionalities. Collectively, it was shown that chemical functionality has little effect on the hybridization properties of PNA as compared to chirality. Those functional groups derived from L-amino acids are detrimental to the binding of PNA. Introduction of a single L-α-unit can destabilize the PNA-DNA or PNA-RNA by as much as 10 °C. However, the effect of the D-isomer on the thermal stability of such heteroduplexes is minimal, at most -1 to -2 °C per unit. The next modification was made at the β-backbone. Although some cyclic PNA analogues containing a chiral centre at the β-position have been made, those with a single substituent at the β-backbone have been synthesized only recently by Sugiyama and coworkers. Their results showed that the R-isomer is detrimental to PNA binding, while the S-isomer maintains binding affinity of the original PNA design.

I.4.2. γ-Backbone Modifications

Despite the diverse structural modifications that have been made to the backbone of PNA, not much attention has been given to γ-position. The first chemical modification made at this position was first reported by Kosynkina and coworkers in 1994. Even though they synthesized the monomer, a detailed study of the hybridization properties of the corresponding oligomers was not done. Later on, Appella, Marchelli and Oliver reported the synthesis of γ-substituted lysine PNA chemical building blocks. Incorporation of these monomers into the mixed-sequence PNA oligomers has led to slight improvements in the binding affinity for DNA as well as RNA. Seitz and coworkers demonstrated the convergent synthesis of PNAs via native chemical ligation.
by building a monomeric unit containing a cysteine side-chain at the N-terminal γ-backbone. At around the same time, our group has also focused on development of γPNA. Using the concept of “conformational preorganization,” developed by Cram, Pedersen and Lehn, we envisioned the possibility of preorganizing the conformation of PNA into a right-handed helical motif—one that closely resembles that of the bound PNA-DNA duplex state. Such conformational preorganization would minimize the entropic penalty associated with molecular recognition and in turn enhance the thermodynamic stability and sequence-specificity of PNA in recognition of DNA or RNA. Our initial goal was to achieve superior hybridization properties. For this, we embarked on a systemic study aimed at gaining a better understanding of the molecular principles that direct γPNA folding.

I.4.2a. Chemical Synthesis

In most cases, the γPNA monomers were prepared by the reductive amination route starting from Boc/Fmoc protected amino acids. One of the key intermediates for such a synthetic route is Boc/Fmoc protected amino aldehydes. Published reports show that these amino aldehydes are susceptible to epimerization. Due to the presence of adjacent electron-withdrawing carbonyl group, the α-proton is relatively acidic and susceptible to deprotonation by base. During the course of the reaction such amino aldehydes have been shown to undergo epimerization. To minimize the possibility for epimerization, the aminoaldehyde intermediates are often kept cold and used immediately without purification. Despite such a precaution, epimerization could still occur. Following this procedure, γPNA monomers containing Ala, Ser, Lys and Cys side-chains were prepared. The submonomeric approach adopted
for the synthesis of α-monomers has been applied in the synthesis of γPNA building blocks.\textsuperscript{115} Recently, Winssinger and coworkers\textsuperscript{124} reported a new strategy using Boc/Mitt protected groups in the preparation of Fmoc-serine-containing γPNA monomers. We have also attempted the reductive amination route in gaining access to the γPNA monomers. However, we have discovered that despite the careful consideration in handling the aminodehyde intermediates and in optimizing the subsequent reaction steps, epimerization still occurred to a certain extent. This has prompted us to explore the Mitsunobu coupling reaction in the preparation of the γ-backbone, which bypasses the formation of aminoaldehyde intermediates.\textsuperscript{125-127} Since the α-proton of the aminoalcohol is inert to deprotonation by base, this method yielded optically superior monomers.

One of the major drawbacks of PNA is water solubility. Due to the charge-neutral backbone, PNA is only moderately soluble in water; further, it has a propensity to aggregate and adhere to surfaces and other macromolecules in a nonspecific manner. Such an inherent property makes it difficult in the handling of PNA and limits its utility in biology and medicine due to the concern for cytotoxicity. We envisioned that by attaching the miniPEG (diethylene glycol) side-chain at the γ-backbone, this would improve its water solubility and biocompatibility due to the hydrophylic nature of such a chemical moeity. This design consideration has led to the development of (R)-MP-γPNA, with the monomers prepared via the Mitsunobu route.\textsuperscript{126}

**I.4.2b. Conformational Preorgnaization**

The effect of γ-backbone on the conformation of PNA was determined by CD experiments. As expected, no exciton coupling pattern was observed for unmodified PNA in the nucleobase absorption range of (220-300) nm because of the random-coil, globular
nature of such a molecule. On the other hand, we observed a distinct biphasic exciton coupling pattern for γPNA, characteristic of a right-handed helix. Through a detailed spectroscopic study of the serine and alanine-based γPNA, it was shown that a single chiral unit is sufficient to induce helical preorganization in PNA, in line with the concept of “sergeants and soldier” developed by Greene and coworkers to explain polymer folding. We showed that randomly folded, single-stranded PNAs can be preorganized into either a right-handed or left-handed helical motif by installing an appropriate stereogenic center at the γ-backbone. This helical induction is unidirectional, occurring from C to N terminus. It is sterically driven and stabilized by base-stacking. γPNA derived from L-amino acid adopted a right-handed helix, while that derived from a D-amino acid adopted a left-handed helix (unpublished data); however, only the right-handed helical γPNA is able to hybridize to DNA or RNA with high affinity and sequence-specificity. Next, we carried out a detailed multinuclear and multidimensional NMR analysis and MD simulations to gain insight into the structure of γPNA. The NMR data revealed that the P-form helical structure adopted by the aeg-PNA duplexes is prefered by the single strand γPNA. During the course of this study, we learned that the γ-position proves to be a robust jumping point for incorporation of new functional group. Thus, we incorporated different chemical groups at the γ-backbone and characterized their conformations and hybridization properties. Our data revealed that the γ-position can accommodate a variety of functional groups; however, unlike the modifications made at the α-backbone which result in destabilization of the duplex, incorporation of the γPNA building block leads to a significant improvement in the thermal stability of the duplex.
I.4.2c. Binding Affinity

Due to conformational preorganization in the backbone, we observed a significant enhancement in the binding affinity of γPNA for DNA and RNA as compared to the unmodified PNA.\textsuperscript{120,141,142} On average the $T_m$ of a PNA-DNA duplex increased by 3 °C and that of a PNA-RNA by ~2 °C for every γ-chiral unit incorporated into the PNA oligomer.\textsuperscript{120} The thermal stability gain is irrespective of the spacing arrangement, consecutive or alternating. We also determined the thermodynamic parameters for the various γPNA-DNA and γPNA-RNA duplexes using Van’t Hoff analysis to gain a better understanding of the binding free energy contribution.\textsuperscript{126} The Gibbs binding free energy was found to steadily increase with increasing number of γPNA building blocks incorporated into the PNA oligomer. We also examined the kinetics of hybridization by Surface Plasma Resonance (SPR).\textsuperscript{126} A comparative study between the unmodified PNA and γPNA revealed that the slower dissociation rate constant gives rise to the higher binding affinity of γPNA towards the complementary DNA and RNA strands.

I.4.2d. Sequence-Specificity

Achieving high sequence selectivity is one of the most salient requirements for molecular recognition. PNA has been shown to possess high sequence selectivity for DNA as well as RNA.\textsuperscript{92} It is well known that binding affinity and sequence selectivity are inversely related. Having established that γPNA adopts a right-handed helical motif closely mimicking that of the bound γPNA-DNA duplex state, we set out to determine the sequence selection of γPNA. Our UV-melting experiments revealed that for a 10mer sequence, a single-base mismatch destabilized the γPNA-DNA duplex by 17 to 21 °C and
\(\gamma\text{PNA-RNA}\) by 16 to 20 °C, depending on the mismatch pair.\textsuperscript{126} This level of sequence selectivity is higher than that of PNA. As expected, our results showed that with conformational preorganization in the backbone, \(\gamma\text{PNA}\) recognizes and binds DNA and RNA with higher affinity and sequence selectivity than the achiral counterpart. We attribute the improvements in binding affinity and sequence selectivity to the reduction in the entropic penalty and backbone rigidity of \(\gamma\text{PNA}\).

\textbf{I.4.2e. Strand Invasion}

PNA has been shown to be capable of invading double helical DNA, with recognition occurring through Watson-Crick base-pairing.\textsuperscript{128-131} However, the drawback of the original (achiral) design is that DNA strand invasion is restricted to mostly homopurine and homopyrimidine targets. The inability of PNA to invade mixed-sequence DNA has been attributed in large part to the lack of binding free energy. In order to overcome this energetic barrier, researchers have adopted a double-duplex invasion strategy.\textsuperscript{132} However, due to the inevitable result of sequence complementarity, such a strategy is limited in scope. With the additional binding free energy gained from backbone preorganization, it was thought that \(\gamma\text{PNA}\) might be able to invade mixed-sequence DNA. To address this question, our group synthesized a series of \(\gamma\text{PNAs}\) and examined DNA strand invasion by gel electrophoresis.\textsuperscript{133} Our result showed that \(\gamma\text{PNAs}\) can indeed invade double helical B-form DNA (B-DNA), with recognition occurring through direct Watson-Crick base-pairing.\textsuperscript{133} The mechanism of strand invasion was confirmed by chemical and enzymatic footprintings. However, we noted that DNA strand invasion was less efficient at high salt concentration [10 mM sodium phosphate, 2 mM MgCl\(_2\), and 150 mM KCl (pH 7.4)].\textsuperscript{134,135} This was expected because the stability of the
DNA double helix increases with increasing salt concentration due to the reduction in electrostatic repulsion in the backbone, and as the result the base-pair breathing rate is reduced. This leads to a slower invasion kinetic and less efficient in the formation of invasion complex. More recently our group has demonstrated that with longer γPNA oligomers, 13-20 nucleotides in length, DNA strand invasion can be accomplished with any sequence at physiological temperature. To the best of our knowledge this is the only class of nucleic acids developed to date that is capable of invading any sequence of double helical B-DNA.

I.4.2f. Cell-Permeability

α-GPNA, where a guanidine group is attached at the alpha-backbone, holds considerable promise as an antisense agent because of its cell-permeability, but the high cost associated with the production of the required monomers has prevented it from finding widespread applications because of the need for large scale production. This has led our group to develop the next-generation γGPNA, employing the same Misunobu route as described previously. With the guanidinium group installed at the γ-backbone, γGPNA is just as efficient in traversing the cell membrane. Further, it binds DNA or RNA with higher affinity and sequence selectivity than the precursor αGPNA, not to mention the significantly lower monomer production cost, since they are prepared from L-amino acid. In addition to getting into live cells, γGPNA can be systemically delivered into intact organisms (zebrafish, sea urchins and mice), as demonstrated in a recent series of studies by our group, and is capable of selectively knocking down gene expression in mice. Another group has shown that an 18-mer γGPNA can knock down miR-210. The same group also reported the synthesis of a modified PNA having a nuclear localization
sequence (NLS) incorporated at the termini.\textsuperscript{149} This modified PNA was internalized into RH30 cells and localized in the nuclei. Based on the importin-mediated transportation mechanism of NLS peptide, they proposed that the modified PNA can interact with importin. Recently, Ganesh and colleagues have developed chiral PNAs called \textit{am}-PNAs,\textsuperscript{150,151} with the cationic aminomethyl groups separately attached at the \(\alpha\)- and \(\gamma\)-backbone. These \textit{am}-PNAs have been shown to cross the extracellular membrane of HeLa cells and localize in the nucleus. The order of the cellular uptake efficiency was determined to be \(\gamma-(S)>\alpha-(R)>\alpha-(S)\). Overall, the \(\gamma\)-backbone provides a versatile jumping point for introducing chemical functionalities, such as the guanidinium group, as mentioned above for improving cellular uptake.

I.5. Thesis Overview

The research work described in this thesis focuses on methodology development and synthesis of \(\gamma\)PNA monomers. These synthetic methodologies were developed in an attempt to synthesize optically pure \((R)\)- and \((S)\)-\(\gamma\)PNA monomers. As stated in the above sections, \(\gamma\)PNA has great potential for biological and biomedical applications, as well as molecular engineering, but the bottle neck has been in the synthesis of the chemical building blocks in an economical and scalable way. Chapter II describes a comparative study of the different methods used to synthesize optically-pure \(\gamma\)PNA monomers and the effect of epimerization on the properties of \(\gamma\)PNA. We explored different synthetic routes to optimize the conditions for preparing \(\gamma\)PNA monomers with optimal chemical yield and enantiomeric purity. In Chapter III, we describe the synthetic route for preparing \((R)\)-Fmoc-MP-\(\gamma\)PNA monomers for the first time. Initially we focused on Boc solid-phase chemistry for the synthesis of \(\gamma\)PNA. However, due to the harsh cleavage conditions
employed in solid-phase synthesis, it is not compatible with automated synthesizers as well as a number of chemical groups. This drawback has motivated us to develop a new synthetic route for preparing optically-pure Fmoc monomers using Mitsunobu coupling method. Chapter IV reports a general synthetic strategy for making both optically-pure \((R)\)- and \((S)\)-Fmoc-MP-\(\gamma\)PNA monomers in an economical, robust, versatile and scalable way. The route described in Chapter III was subsequently found to be not suitable for large scale monomer synthesis due to difficulty in removing the triphosphine oxide byproduct generated in the Mitsunobu reaction. One of the salient features of this route is that the synthesis of both isomers starts with the same commercially available and relatively cheap L-serine. The methodology development described in this thesis could have a significant impact on the research and development of \(\gamma\)PNA as molecular tools for regulating gene expression, as well as diagnostic and therapeutic reagents for the detection and treatment of genetic diseases. The development of Fmoc-MP-\(\gamma\)PNA monomers now render them suitable for automated solid-phase synthesis, which in turn will facilitate the parallel production of oligomers.

I.6. References


130. Smolina, I.V.; Demidov, V.V.; Soldatenkov1, V.A.; Chasovskikh, S.G.; Frank-Kamenetskii, M.D. End invasion of peptide nucleic acids (PNAs) with mixed-base composition into linear DNAduplexes. Nucleic Acids Res. 2005, 33, e146.


CHAPTER II

SYNTHESIS OF OPTICALLY-PURE γPNA MONOMERS: A COMPARATIVE STUDY

II.1. Introduction

Achieving tight and specific binding has been a major goal of molecular design for recognition and self-assembly. However, it is generally difficult to realize one without compromising the other due to their inverse relationship. Nature has devised a creative solution to this problem by stitching individual chemical building blocks, such as amino acids, nucleotides, and nucleosides, into linear chains, which then self-organize into compacted three-dimensional structures. Such conformational pre-organization, whereby the substrates adopt the conformations of the bound state prior to recognition, would not only improve the binding free energy, by minimizing the entropic cost, but also increase the rigidity of the system, by making it less accommodating to structural changes—hence the enhancements in binding affinity and specificity. The concept of conformational pre-organization has long been recognized by Pedersen, Cram, and Lehn, and successfully applied in the design of small-molecules for guest–host interactions. However, such a concept has been less successful with the design of larger macromolecular and oligomeric systems due to their many degrees of freedom—although some progress has been made.

As mentioned in Chapter I, Ly group has recently shown that PNA can be pre-organized into a right-handed helix by installing an appropriate stereogenic centre at the gamma backbone. Such a conformational motif hybridizes to DNA as well as
RNA with high affinity and sequence selectivity, and capable of invading any sequence of dsDNA. During the course of this study we noted that such properties are achieved only if the chemical building blocks (and consequently the oligomer themselves) are optically-pure. The efficiency of strand invasion was found to be strongly dependent on the optical purity of the monomers. The presence of even a minute amount of the epimerized mixture can have a profound adverse effect on strand invasion. It was therefore necessary to have optically-pure monomers, if the corresponding γPNA oligomers were to be able to invade dsDNA (or dsRNA) and be successfully employed in antigene applications. In this chapter, we report a systematic study, exploring different synthetic routes for preparing optically-pure γPNA monomers and examining the effects of epimerization on the binding properties of PNA.

II.2. Results and Discussion

So far most of the chiral γPNA monomers were prepared through the reductive amination route, starting with either Boc- or Fmoc-protected amino acids. The key intermediate involved in this reductive amination route is Boc- and Fmoc-protected amino aldehydes. It is well known in the literature that these intermediates undergo epimerization, even under mild experimental conditions such as rapid chromatography on silica gel. One way to suppress epimerization has generally relied on the immediate use of cold, unpurified amino aldehyde solutions. Even when such is precautionary measure is adopted, epimerization could still occur. In order to address the issue of optical purity of γPNA monomers, we explored two synthetic routes in the preparation of γPNA backbone: (i) reductive amination with the α-amino
group separately protected with Boc- and N-(9-(9-phenylfluorenyl)) (PhFl), and (ii) Mitsunobu reaction with Boc-protection. The sterically bulky PhFl-protecting group was chosen because it has been shown by Rapoport\textsuperscript{21,22} to be configurationally stable under reductive amination conditions. The sterically bulky PhFl group shields the alpha proton of the amino aldehyde from deprotonation by the base, and thus its optical purity is retained. Similarly, the Mitsunobu reaction was ascertained because of its superior enantioselectivity.\textsuperscript{11} In the latter case, instead of going through the configurationally unstable aminoaldehyde, the protected amino acid is immediately converted to aminoalcohol, the enantioselective chemical transformation for which has already been established.\textsuperscript{6}

II.2.1. Synthesis of γPNA monomers

γPNA monomers were prepared following the two synthetic routes shown in Scheme 1, starting from the same batch of commercially available and relatively cheap L-alanine 1 (99.6% ee). In route (i), L-alanine 1 was protected with Boc and PhFl, separately. Conversion of 2a and 2b to Weinreb amides 3a and 3b, followed by reduction with LiAlH\textsubscript{4} yielded the aldehydes. Condensation of 4a-i (denoting the Weinreb amide reduction route) and 4b with glycine ethyl ester via reductive amination produced the desired alanyl-γPNA backbones 5a-i and 5b, respectively. In route (ii), Boc-L-alanine 2a was converted to aminoalcohol 6 under a mild condition.\textsuperscript{23} Treatment of 6 with NBS-Gly-OEt under the Mitsunobu condition followed by deprotection with thiophenol/K\textsubscript{2}CO\textsubscript{3} yielded the backbone intermediate 5a-ii (denoting the Mitsunobu route). All three alanyl-γPNA backbones 5a-i, 5a-ii and 5b were
coupled to the thymine nucleobase using DCC/DhbtOH and then hydrolyzed with 2N
NaOH to give the desired monomers 10a-i, 10a-ii and 10b, respectively, (Scheme 2).
Detailed information about the synthesis and characterization of the monomers are
provided in the Experimental section of this chapter.
Scheme 1. Synthesis of alanyl-\(\gamma\)PNA backbones. Reagents and conditions: (a) (Boc)\(_2\)O, NaHCO\(_3\), 1,4-dioxane/water, rt; (b) (CH\(_3\))\(_3\)SiCl, CHCl\(_3\)/CH\(_3\)CN, reflux, then Et\(_3\)N, Pb(NO\(_3\))\(_2\), 9-bromo-9-phenylfluorene, rt; (c) NMM, ClCO\(_2\)iBu, DCM, -15 °C, CH\(_3\)NHOCH\(_3\)•HCl, rt; (d) DCC, DhbtOH, CH\(_3\)NHOCH\(_3\)•HCl, DMF, rt; (e) LiAlH\(_4\), THF, 0 °C; (f) glycine ethyl ester, MeOH, 4 °C, then acetic acid, NaBH\(_3\)CN; (g) i) NMM, ClCO\(_2\)iBu, DME, -20 °C, ii) NaBH\(_4\), H\(_2\)O, -20 °C; (h) o-NBS-Gly-OEt, TPP, DIAD, THF, 0 °C \(\rightarrow\) rt; (i) PhSH, K\(_2\)CO\(_3\), rt.
Scheme 2. Synthesis of alanyl-\(\gamma\)PNA thymine monomers. Reagents and conditions: (a) T-CH\(_2\)CO\(_2\)H, DCC, DhbtOH, DMF, 50 °C; (b) 2N NaOH:THF (1:1), 0 °C.

II.2.2. Determination of optical-purity by \(^{19}\)F NMR

Next, we determined the optical-purity of the alanyl-\(\gamma\)PNA backbones 5a-i, 5a-ii and 5b and the corresponding monomers 10a-i, 10a-ii and 10b using fluorine NMR, after coupling to Mosher’s reagent. Generally there are two ways of determining the optical purity of a compound: chiral HPLC and chemical derivatization. We adopted the chemical method for determining the optical-purity since it is a relatively economical route. In this instant, the protecting groups (Boc and PhFl) were removed and the resulting amino groups were coupled to R-(+)-\(\alpha\)-Methoxy \(\alpha\)-(trifluoromethyl)phenylacetyl chloride (MTPA-Cl) (Scheme 3) for \(^{19}\)F-NMR analysis. The other diastereomers, D-alanine-derived \(\gamma\)PNA backbone and thymine monomer, were also prepared for comparison using the Mitsunobu protocol.
Scheme 3. Synthesis of MTPA derivatives. Reagents and conditions: (a) 5% m-cresol, TFA, DCM, rt (Boc removal); (b) H2-Pd/C, MeOH, rt (PhFl removal); (c) MTPA-Cl, DIPEA, DCM 0 °C→rt; (d) MTPA-Cl, NaHCO3, acetone/H2O (1/1).

19F-NMR analysis of the backbone 11a-i (Boc-protected) obtained from reductive amination showed roughly 5% epimerization (Figure II-1B), whereas that protected with PhFl (11b) and Boc-protected 11a-ii prepared via the Mitsunobu route showed excellent ee values (>99.5%, the minimum detection limit of 19F-NMR).24 The starting materials were determined to be optically-pure, as shown in Figure II-1A. It was also noted that the degree of epimerization (~5%) for 11a-i was carried from the backbone to the final monomer (12a-i, Figure II-2). This indicates that epimerization occurred during the reductive amination step involving the configurationally unstable amino aldehyde intermediate. On the other hand, the reductive amination product containing PhFl-protecting group gave enantiopure backbone because the steric bulkiness of PhFl prevents deprotonation of the α-proton in aldehyde 4b. The Mitsunobu route proved to be comparatively better than the reductive amination route in yielding optically-pure γPNA monomers. 19F-NMR of MTPA-derived monomers 12a-i, 12a-ii and 12b showed two peaks due to the existence of two rotamers (major
and minor), which occur as the result of rotational restriction around the amide nitrogen that connects the backbone to the nucleobase.

![Figure II-1](image)

**Figure II-1.** $^{19}$F-NMR of MTPA-derivatized (A) D- and L-alanine starting materials, and (B) alanyl-$\gamma$PNA backbones prepared from the reductive amination (11a-i and 11b) and Mitsunobu (11a-ii) routes.
Figure II-2. $^{19}$F-NMR of MTPA-derivatized alanyl-$\gamma$PNA T-monomers prepared from the reductive amination (12a-i and 12b) and Mitsunobu (12a-ii) routes.

II.2.3. Detailed analysis of reductive-amination route

Since reductive amination is a two-step process, involving reduction of Weinreb amide (or oxidation of alcohol) to aldehyde and reductive coupling of aldehyde with amine, we decoupled the two reaction steps to determine the origin and the extent of their epimerization separately. For this study, we first synthesized optically-pure Boc-protected aminoalcohols, for both D- and L-isomers, using a known procedure.\textsuperscript{23} $^{19}$F-NMR data showed that they were optically-pure (Figure II-3). During the reduction of the Weinreb amide to aldehyde as shown in Scheme1, we used LiAlH$_4$ reagent. Upon completion of the reaction we performed an acid workup. We speculated that this chemical transformation step might be a predominant source of
epimerization, as the result of prolonged exposure of the aldehyde to low pH aqueous solution in the workup. Thus, we examined two additional reactions, Swern\textsuperscript{25} and Parikh-Doering,\textsuperscript{26} that have been shown to be milder and less prone to epimerization in the oxidation of alcohol to aldehyde (Scheme 4). Initially following the standard protocols of Swern oxidation of alcohol using 9 equivalents of triethylamine, we observed a considerable degree of epimerization of the resulting backbone. This could be due to the excessive use of base in the reaction. We thus carried out a series of reactions with varying ratios of base to alcohol, and found that 3 equivalents of base (Et\textsubscript{3}N) gave the highest chemical yield (>90%) and optical purity within a one-hour period. A detailed experimental procedure is given at the end of this chapter. Similarly, for Parikh-Doering oxidation, we optimized the reaction conditions as described in the experimental section. By employing three different methods starting from 3a and 6, we synthesized compounds 13a-i, a-ii, a-iii. After removal of the Boc-protecting group, we coupled the amines to Mosher’s reagent and analyzed the resulting products (14a-i, a-ii, a-iii) by \textsuperscript{19}F-NMR. Under these conditions, the optical purities of the aldehydes produced by the three methods, reduction of Weinreb amide by LiAlH\textsubscript{4}, Swern and Parikh-Doering oxidation, were similar (~ 1.5%) to one another, as shown in (Figure II-4). Irrespective of the methods employed in the synthesis of aldehyde, a small amount of epimerization was observed for each even under mild reaction conditions. This result shows that the origin of epimerization is in the formation of aldehyde. Following a similar characterization protocol, we next selected one of the aldehyde products with a known ee value, in this case the product of the Swern oxidation (4a-ii), due to the convenience in scaling, and subjected it to an optimized reductive amination
condition to get the backbone. $^{19}$F-NMR spectrum of the resulting backbone product following Boc-removal and coupling to Mosher’s reagent showed nearly doubling of the degree of epimerization. These results together show that under these optimized conditions, each reaction step (alcohol to aldehyde and aldehyde to secondary amine) contributed a similar degree of epimerization (~1.5% each).

**Figure II-3.** $^{19}$F-NMR of MTPA-derivatized D- and L-alanol intermediates.
Scheme 4. Comparison of the oxidation methods. Reagents and conditions: (a) LiAlH$_4$, THF, 0 °C; (b) Swern [(COCl)$_2$, DMSO, Et$_3$N, DCM, -78 °C]; (c) Parikh-Doering [Py•SO$_3$, DMSO/DCM (1/1), Et$_3$N, 0 °C→rt]; (d) NaBH$_3$CN, MeOH, 0 °C→rt; (e) 5% m-cresol, TFA, DCM, rt; (f) MTPA-Cl, DIPEA, DCM 0 °C→rt; (g) glycine ethyl ester, MeOH, 4 °C, then acetic acid, NaBH$_3$CN.
II.2.4. Effect of epimerization on the conformation of PNA

First, we investigated the effect of epimerization of monomers on the conformation of PNA. For this study we synthesized a series of oligomers. They were prepared from two separate batches of monomers: those obtained via the reductive amination (10a-i) and Mitsunobu (10a-ii) routes. The sequences of different PNA oligomers employed in this study are given in Table II-1.
Circular dichroism (CD) experiments were performed. CD spectra were recorded at room temperature for all oligomers as single strands (Figure II-5). As expected, no noticeable CD signals were observed for the unmodified PNA1 in the nucleobase absorption regions (200-350 nm) in accordance with the literature. This is because PNA, as individual strands, generally do not have well-defined conformations. On the other hand, a pronounced biphasic exciton coupling pattern characteristic of a right-handed helix was observed for PNA2a, 2b, 3a and 3b, with maxima at 220 and 260 nm and minima at 245 and 290 nm. An interesting finding was that the amplitudes of the CD signals of PNA2a at 220 and 260 nm are smaller than that of PNA2b. This can be attributed to the fact that PNA2a was prepared from the epimerized building block (10a-i) while PNA2a was synthesized from optically-pure monomers via the Mitsunobu route. Since the epimerized monomer contained only ~ 95% of the desired (S)-diastereomer, which adopts a right-handed helix, and the other 5% being an (R)-diastereomer, which adopts a left-handed helix, its incorporation into PNA2a would

Parenthesis: The monomers used in the oligomer synthesis indicated in bold.

<table>
<thead>
<tr>
<th>PNA</th>
<th>Sequence (N to C terminus)</th>
<th>M.W.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>H-GCATGTTTGA-\textsuperscript{-1}Lys-NH\textsubscript{2}</td>
<td>2886</td>
</tr>
<tr>
<td>2a (10a-i)</td>
<td>H-GCATGTTTGA-\textsuperscript{-1}Lys-NH\textsubscript{2}</td>
<td>2900</td>
</tr>
<tr>
<td>2b (10a-ii)</td>
<td>H-GCATGTTTGA-\textsuperscript{-1}Lys-NH\textsubscript{2}</td>
<td>2900</td>
</tr>
<tr>
<td>3a (10a-i)</td>
<td>H-GCATGTTTTGA-\textsuperscript{-1}Lys-NH\textsubscript{2}</td>
<td>2942</td>
</tr>
<tr>
<td>3b (10a-ii)</td>
<td>H-GCATGTTTTGA-\textsuperscript{-1}Lys-NH\textsubscript{2}</td>
<td>2942</td>
</tr>
</tbody>
</table>

T = Alanyl-\gamma PNA T-monomeric unit; 10a-i: prepared via reductive amination route, 10a-ii: Mitsunobu route.
cancel each other to an extent that leads to a decrease in the overall amplitude of the CD signal. A similar pattern was observed for PNA3a, with four epimerized building blocks, as compared to that of PNA3b which contained optically-pure monomers, with the latter exhibiting greater signal strengths. We ruled out the differences in the CD amplitudes as the result of variations in the concentrations, because the UV-absorption profiles of the oligomers were identical at 90 °C, the temperature at which all the strands were completely denatured, indicating that they were of the same concentration. Based on these observations it can be concluded that epimerization affects the overall conformational preorganization of γPNA.

Figure II-5. CD spectra of single-strand unmodified PNA (PNA1) and alanyl-γPNAs (PNA2a, 2b, 3a and 3b). All samples were prepared at 5 μM strand concentration in 10 mM NaPi buffer. CD spectra were recorded at 22 °C. (Inset) UV-Vis spectra of the indicated oligomers at 90 °C showing that their concentrations were the same at 260 nm.
II.2.5. Effect of epimerization on binding affinity of PNA towards complementary DNA

Next, we studied the effect of epimerization on the binding property of PNA. We determined the thermal stability of the same set of oligomers upon hybridization to a complementary (antiparallel) DNA strand using variable-temperature UV-spectroscopy (Figure II-6). Roughly a 3 °C decrease in the $T_m$ was observed for the PNA2a-DNA hybrid duplex as compared to that of PNA2b-DNA, due to epimerization in the monomer (10a-i) employed in the synthesis of PNA2a. These results are in agreement with our recent findings for a different project in which we showed that the left-handed γPNAs are unable to hybridize to DNA or RNA due to conformational mismatch, despite the sequence complementarity—which explains the destabilizing effect. Similarly, the epimerized PNA3a-DNA duplex showed a 6 °C decrease in the $T_m$ as compared to the enantiomerically-pure PNA3b-DNA duplex. The differences in the thermal stability of the PNA-DNA duplexes are consistent with the CD data. Thus, epimerization of the monomer leads to decreased binding affinity of the γPNA towards the complementary DNA.
Figure II-6. UV-melting profiles of PNA-DNA hybrid duplexes. The concentration of each strand was 5 µM prepared in 10 mM NaPi buffer (pH 7.4). Both the heating and the cooling runs were recorded, with both runs showing nearly identical profiles. The $T_m$s were determined by taking the first derivative of the melting curves.

II.3. Conclusion

In summary, we have shown that depending on the size of the protecting group (Boc vs. PhFl) and the synthetic route (reductive amination vs. Mitsunobu) through which the backbone intermediates were prepared, optically-pure $\gamma$PNA monomers could be obtained. Utilization of PhFl-protection via reductive amination and Boc-protection via the Mitsunobu route yielded optically-pure backbone intermediates and the corresponding $\gamma$PNA monomers, while Boc-protection via reductive amination resulted in considerable epimerization. Further studies revealed that reduction of the Weinreb amide or oxidation of the alcohol to aldehyde via the Swern or Parikh-Doering method contributes about half to the overall epimerization, while the other half comes from the reductive amination of aldehyde with glycine to the backbone.
Incorporation of the epimerized monomers into oligomers resulted in conformational heterogeneity and reduced the degree of conformational preorganization and binding affinity of PNA towards complementary DNA strand, as demonstrated in the CD and UV-thermal melting measurements. When the oligomer was comprised entirely of the epimerized building blocks (A, C, G, T), a dramatic reduction in the $T_m$ (15-20 °C, unpublished data) was observed. Obtaining optically-pure $\gamma$PNA monomers is therefore critical to achieving ultrahigh binding affinity and effectively being able to invade double helical B-DNA or RNA (data not shown).

II.4. Experimental section

Material and Methods. All commercial reagents were used without further purification. Solvents were dried by standard methods and distilled freshly prior to use. All chemicals were purchased from Aldrich except for L-alanine, which was purchased from Novabiochem. All Boc/Z-protected PNA monomers were purchased from Applied Biosystems. $^1$H-, $^{13}$C- and $^{19}$F-NMR spectra were recorded on a Bruker Avance AV-300 NMR spectrometer using standard Bruker software. Flash chromatography was performed using standard grade silica gel from Sorbent Technologies. TLC was performed with silica gel 60 F-254 precoated plates from Sorbent Technologies. MALDI-TOF experiments were performed on a PerSeptive Biosystems Voyager STR MALDI-TOF mass spectrometer using a 10mg/ml solution of $\alpha$-hydroxycinnamic acid in ACN-water (1:1) with 0.1% TFA. Mass spectra were recorded on a Finnigan LCQ ESI/APCI ion trap mass spectrometer by electrospray ionization. CD experiments were performed on a Jasco J-715 spectropolarimeter that
equipped with a thermoelectrically-controlled single-cell holder. UV-Vis measurements were taken on a Varian Cary 300 Bio spectrophotometer equipped with a thermoelectrically controlled multi-cell holder.

**General procedures**

**Oligomer synthesis**

All PNA oligomers were synthesized on methylbenzhydryl amine (MBHA) resin according to the published procedures. All oligomers were cleaved from the resin using m-cresol: thioanisole:TFMSA:TFA (1:1:2:6) cocktail. The resulting mixtures were precipitated with ether, purified by RP-HPLC and characterized by MALDI-TOF mass spectrometry. Concentrations of the PNA oligomers were determined from the OD at 260 nm recorded at 90°C, using the following extinction coefficient: T = 8,600 M⁻¹cm⁻¹, A = 13,700 M⁻¹cm⁻¹, C = 6,600 M⁻¹cm⁻¹, and G = 11,700 M⁻¹cm⁻¹. The typical yields for these oligomers are in the 27-37% range, after purification.

**UV-Melting of PNA-DNA hybrid duplexes**

All samples were prepared in buffer containing 10 mM sodium phosphate (pH 7.4). UV-Vis absorbance at 260 nm was recorded at the rate of 1°C/min for both the heating (20-95 °C) and cooling (95-20 °C) runs. The TmS were determined by taking the first derivative of the heating profiles (both the heating and cooling profiles yielded similar results, with little hysteresis).

**Circular Dichroism (CD)**

All samples were prepared in buffer containing 10 mM sodium phosphate (pH 7.4) at 5 μM strand concentration each. All samples were annealed prior to recording the CD spectra, by heating to 90 °C for 5 min followed by gradual cooling to room
temperature. All spectra represent an average of at least ten scans collected at the rate of 100 nm/min between 200-320 nm in a 1-cm path-length cuvette at 22 ºC. CD spectrum from buffer solution was subtracted from the sample spectra, which were then smoothed via an eight-point adjacent averaging algorithm.

**Monomer Synthesis**

*N-tert-Butoxycarbonyl Alanine (2a):* L-alanine 1 (10 g, 112.24 mmol) was dissolved in water and sodium bicarbonate (2 equiv., 224.45 mmol, 18.85 g) was added with stirring. The resulting solution was cooled to 5 ºC and Boc anhydride (1.5 equiv., 168.36 mmol, 36.75 g) was added slowly along with 1,4-dioxane (also cooled). The resulting mixture was stirred at 0 ºC for 1 h and allowed to warm to room temperature overnight. Water was then added and the aqueous layer was extracted twice with EtOAc (200 mL). The organic layer was back extracted twice with saturated sodium bicarbonate solution. The combined aqueous layers were acidified to pH 1 with 10% HCl, and extracted 3 times with EtOAc (100 mL). The combined organic layers were dried (sodium sulfate) and concentrated *in vacuo*. The resulting residue was purified by silica gel chromatography. $^1$H NMR (300 MHz, DMSO-d$_6$): $\delta$ 1.21 (d, $J = 7.32$ Hz, 3H), $\delta$ 1.37 (s, 9H), 3.92 (q, $J = 7.42$ Hz, 1H), 7.05 (d, $J = 7.54$ Hz, 1H), 12.03 (brs, 1H). ESI-MS (positive mode): m/z 190.20 (M$_{calc}$ for C$_8$H$_{15}$NO$_4$: 189.21).

*N-tert-Butoxycarbonyl Alanine Weinreb amide (3a):* To a stirred solution of 2a (10 g, 52.88 mmol) in anhydrous DCM (100 mL) was added N-methylmorpholine (2 equiv., 105.76 mmol, 10.7 g) at -15 ºC (dry ice/methanol) under argon atmosphere. Isobutylchloroformate (1.05 equiv., 55.52 mmol, 7.20 mL) was added drop-wise. After stirring for 15 min at the same temperature N-methyl,O-methylhydroxylamine

50
hydrochloride (1.1 equiv., 58.2 mmol, 5.6g) was added and the ice bath was removed. The reaction mixture was allowed to stir at rt for 14 h. The solution was diluted with DCM (200 mL) and washed with 10% aq. KHSO₄ (100 mL), saturated NaHCO₃ solution (2 x 50 mL) and saturated NaCl solution. The organic layer was dried over Na₂SO₄ and evaporated under reduced pressure. The residue was purified by silica gel column chromatography to afford 3a as a white solid. Yield 90%, Rₚ: 0.6 (3:2, EtOAc: hexane). H NMR (300 MHz, CDCl₃): δ 1.28 (d, J = 6.91 Hz, 3H), 1.41 (s, 9H), 3.18 (s, 3H), 3.74 (s, 3H), 4.65 (m, 1H), 5.22 (brs, 1H). C NMR (75 MHz, CDCl₃): δ 18.0, 27.8, 31.6, 45.9, 61.0, 78.8, 154.6, 173.1. ESI-MS (positive mode): m/z 233.7 (M calc for C₁₀H₂₀N₂O₄: 232.8).

N-tert-Butoxycarbonyl Alanal (4a-i): A stirred solution of 3a (5 g, 21.4 mmol) in dry THF (100 mL) was cooled to 0 °C under argon atmosphere. Lithium aluminum hydride (2.0 equiv, 42.8 mmol, 1.63 g) was added portion-wise and after 1 h, the reaction mixture was quenched with 10% KHSO₄ solution (100 mL) and allowed to stir for 10 minutes. The solution was then extracted with ethyl acetate (3x100 mL). The combined organic layers were washed with 1M HCl (2 x 50 mL), saturated NaHCO₃ (2 x 50 mL) and saturated NaCl (2x50 mL). The organic layer was dried on Na₂SO₄ and evaporated under reduced pressure to afford white solid which was used in the next step without further purification. Yield 90%, Rₚ: 0.8 (3:2, EtOAc: Hexane). H NMR (300 MHz, CDCl₃): δ 1.33 (d, J = 7.42 Hz, 3H), 1.45 (s, 9H), 4.22 (m, 1H), 5.09 (brs, 1H), 9.56(s, 1H). C NMR (75 MHz, CDCl₃): δ 14.3, 27.8, 55.0, 79.5, 154.8, 199.2. ESI-MS (positive mode): m/z 174.2 (M calc for C₈H₁₅NO₃: 173.2).
**N-tert-butoxycarbonyl-1-(methyl)-aminoethylglycine ethyl ester (5a-i):** A stirred solution of 4a-i (1.11 g, 6.45 mmol) in dry MeOH (50 mL) was cooled to 0 °C under argon atmosphere. In a separate round bottom flask, ethylglycinate hydrochloride (2.1 equiv, 13.54 mmol, 1.89 g) was dissolved in MeOH (5 mL) and DIPEA (2.1 equiv, 13.56 mmol, 2.36 mL). The flasks were then mixed and stirred at 4 °C for 4 h at which point acetic acid (1.2 mL, 20 mmol) was added, followed by NaBH₃CN (0.62 g, 10 mmol) and the solution was allowed to stir for another 30 min. 10% NaHCO₃ solution (100 mL) was then added to the reaction mixture and extracted it with EtOAc (3 x 30 mL). The combined organic layers were washed with brine and dried over anhydrous Na₂SO₄ and filtered. The solvent was evaporated under reduced pressure and the resulting residue was purified by column chromatography to afford 5a-i as viscous liquid. Yield 68%, Rf = 0.75 (20:1, EtOAc: EtOH). ¹H NMR (300 MHz, CDCl₃): δ 1.15 (d, J = 6.6 Hz, 3H), 1.29 (t, J = 7.1 Hz, 3H), 1.46 (s, 9H), 2.64 (d, J = 5.9 Hz, 2H), 3.41(ABq, J = 17.4 Hz, 2H), 3.74 (m, 1H), 4.21 (q, J = 7.1 Hz, 2H), 4.79 (brs,1H). ¹³C NMR (75 MHz, CDCl₃): δ 13.6, 18.4, 27.8, 45.6, 50.3, 53.9, 60.1, 78.4, 155.0, 171.9. ESI-MS (positive mode): m/z 261 (Mcalc for C₁₂H₂₄N₂O₄: 260.33).

**N-(9-(9-Phenylfluorenyl))-alanine (2b).** To a stirred suspension of L-alanine (18g, 0.2 mol) in 600 mL of CHCl₃/CH₃CN (5/1) in a Morton flask was added TMS-Cl (1 equiv., 0.2 mol, 26 mL) at room temperature under nitrogen atmosphere. This mixture was heated under reflux for 2 h and then allowed to cool to room temperature. Addition of Et₃N (20 equiv., 4 mol, 56 mL) at a rate sufficient to maintain gentle reflux was followed by the addition of Pb(NO₃)₂ (0.12 mol, 40g) and a solution of 9-bromo-9-phenylfluorene (1 equiv., 0.2 mol, 64.2g) in CHCl₃, (200 mL). The resulting
mixture was vigorously stirred for 48 h at room temperature, and then excess MeOH (0.5 mol) was added. Filtration followed by evaporation gave a residue which was partitioned between Et₂O (1 L) and pre-cooled 5% aqueous citric acid (1 L). The organic phase was washed with 1N NaOH (2 x 400 mL) and H₂O (2 x 200 mL), and the combined aqueous layers were washed with Et₂O (400 mL), cooled to 0 °C, and neutralized with glacial AcOH. The precipitated product was extracted with Et₂O (4 x 300 mL). The combined organic layers were washed with water (200 mL), dried over anhydrous Na₂SO₄ and evaporated to give compound 2b as a light yellow foam (55g, 84%), which was used without further purification. Careful crystallization of 2b from EtOAc/hexane produced a white solid: ¹H NMR (300 MHz, CDCl₃): δ 1.18 (d, J = 7.16 Hz, 3H), 2.72 (q, J = 7.16 Hz, 1H), 6.14 (brs, 1H), 7.21-7.50 (m, 11H), 7.70-7.79 (m, 2H). ESI-MS (positive mode): m/z 329.9 (M⁺ calc for C₂₂H₁₉NO₂: 329.14).

N-(9-(9-Phenylfluorenyl))-alanine Weinreb amide (3b): To a stirred solution of 2b (10g, 30.35 mmol) in dry DCM:DMF (3 :1, 20 mL) were added DCC (1.10 equiv., 33.4 mmol, 6.89g) and DhbtOH (1.1 equiv., 33.4 mmol, 5.4g) under N₂ atmosphere at 0 °C. The stirring was continued for 1 h at the same temperature. In a separate flask the mixture of N-methyl-O-methyl hydroxylamine hydrochloride salt (1.1 equiv., 33.4 mmol, 3.21g) and DIPEA (1.1 equiv., 33.4 mmol, 5.81 mL) in 5 mL DCM: DMF (3:1), cooled to 0 °C was added and the ice bath was removed. The reaction mixture was stirred at rt for 14 h. Following evaporation of the solvent, the residue was partitioned between ethyl acetate (100 mL) and saturated NaHCO₃ solution (100 mL). The organic layer was washed with 10 % KHSO₄ (3 x 50 mL), 10 % NaHCO₃ (3 x 50 mL), brine (100 mL) and dried over Na₂SO₄. The solvent was removed under reduced
pressure and purified by column chromatography to afford 3b as pale yellow solid. Yield 73%, Rf: 0.6 (3:2, EtOAc:hexane). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 1.06 (d, $J = 6.93$ Hz, 3H), 2.7-3.1 (m, 1H), 2.85 (s, 3H), 2.88 (s, 3H), 3.3 (brs, 1H), 7.1-7.6 (m, 11H), 7.63-7.75 (m, 2H). $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 21.3, 31.4, 47.6, 59.7, 72.7, 118.9, 119, 124, 125.5, 126.2, 127.5, 127.6, 139.3, 140.6, 144.3, 148.9, 149.7, 176.5. ESI-MS (positive mode): m/z 372.8 (M$_{\text{calc}}$ for C$_{24}$H$_{24}$N$_2$O$_2$: 372.1).

$N$-(9-(9-Phenylfluorenyl))-alanal (4b): The target compound was made from compound 3b by following the synthesis procedure of 4a. Yield 80%, Rf: 0.2 (3:2, EtOAc: Hexane). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 0.9 (d, $J = 7.27$ Hz, 3H), 2.57 (qd, $J = 1.68, 7.27$ Hz, 1H), 2.7 (bs, 1H), 7.05-7.45 (m, 11H), 7.56-7.65 (m, 2H), 9.17 (d, $J = 1.68$ Hz, 1H).

$N$-(9-(9-Phenylfluorenyl))-1-(methyl)-aminoethyl glycine ethyl ester (5b): The corresponding backbone 5b was synthesized from the aldehyde 4b via reductive amination as described in the synthesis of 5a. Yield 76%, Rf: 0.25 (3:2, EtOAc: hexane). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 0.61 (d, $J = 5.9$ Hz, 3H), 1.27 (t, $J = 7.13$ Hz, 3H), 2.10 (bs, 1H), 2.2-2.4 (m, 3H), 3.16 (ABq, $J = 17.32$ Hz, 2H), 4.18 (q, $J = 7.13$ Hz, 2H), 7.12-7.28 (m, 5H), 7.32-7.39 (4H, m), 7.39-7.47 (m, 2H), 7.66-7.72 (m, 2H). $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$13.7, 20.9, 47.2, 50.3, 55.7, 60, 72.4, 119.3, 124.7, 125, 125.6, 126.4, 127.1, 127.5, 127.7, 139.7, 140, 145.2, 149.6, 150.4, 172.

$N$-tert-butyloxycarbonyl alanol (6): To a cold (-20 °C) stirred solution of Boc-alanine 2a (5g, 26.45 mmol) in anhydrous 1,2-dimethoxyethane (DME) (30 mL) were successively added N-methylmorpholine (1 equiv., 26.45 mmol, 2.9 mL) and isobutylchloroformate (1 equiv., 26.45 mmol, 3.46 mL) under nitrogen atmosphere.
After 10 min, the precipitated white solid was removed by filtration and washed with DME (3 x 10 mL). The filtrate and DME washings were combined in a 500 mL round bottom flask and cooled to -20 °C. A solution of sodium borohydride (1.5 equiv., 39.6 mmol, 1.5 g) in water (15 mL) in portion wise, followed by water (500 mL) was added. The reaction mixture was extracted with ethyl acetate (3 x 200 mL) and the combined organic layers were washed with brine, dried over anhydrous Na₂SO₄, evaporated to dryness, and then purified by silica gel column chromatography. Yield 90%, R_f: 0.5 (3:2 EtOAc: hexane). ¹H NMR (300 MHz, CDCl₃): δ 1.15 (d, J = 6.6 Hz, 3H), 1.46 (s, 9H), 2.59 (brs, 1H), 3.49 (dd, J = 6.1, 10.7 Hz, 1H), 3.62 (dd, J = 4.0, 10.7 Hz, 1H), 3.78 (m, 1H), 4.58 (brs, 1H); ESI-MS (positive mode): m/z 176 (M_calc for C₈H₁₇NO₃: 175.23)

[(2-tert-Butoxycarbonylamino-propyl)-(2-nitrobenzene sulfonyl)amino]-acetic acid ethyl ester (8): Compound 6 (3.15g, 18 mmol), o-NBS-GlyOEt (0.95 equiv., 17.1 mmol, 4.7g) and triphenylphosphine (1 equiv., 18 mmol, 4.73g) were dissolved in freshly distilled anhydrous THF (100 mL). The solution was stirred at 0 °C under nitrogen atmosphere, and diisopropylazodicarboxylate (DIAD 3.54 mL, 18 mmol) was then added drop-wise over the course of 20 min. The reaction mixture was stirred overnight at room temperature. The solvent was evaporated to dryness and the oily residue was re-dissolved in diethyl ether (25 mL) and cooled to 0 °C. The next day the precipitate was filtered off and the solvent evaporated in vacuo. The remaining residue was purified by silica gel column chromatography to give pale yellow oil. Yield 78%, R_f: 0.5 (2:3, EtOAc:hexane). ¹H NMR (300 MHz, CDCl₃): δ 1.14 (m, 6H), 1.40 & 1.42 (2xs, 9H), 3.3 & 3.37 (2xd, J = 5.4 Hz, 1H), 3.50 (m, 1H), 3.87 (m, 1H), 4.04 (m,
(2-tert-Butoxycarbonylamino-propylamino)-acetic acid ethyl ester (5a-ii): To a stirred solution of 8 (3g, 6.73 mmol) in acetonitrile (75 mL) were added potassium carbonate (1.5 equiv., 10.09 mmol, 1.38g) and thiophenol (2 equiv., 13.5 mmol, 1.60 mL) at room temperature. The reaction mixture was stirred overnight. The solvent was evaporated and the residue was partitioned between ethyl acetate (100 mL) and water (50 mL). The ethyl acetate layer was extracted with 10% aqueous citric acid (3 x 100 mL). To the combined aqueous layers solid potassium carbonate was added portion-wise in high excess (CAUTION! A strong evolution of gas was observed). The aqueous mixture was then extracted with ethyl acetate (3 x 75 mL). The combined ethyl acetate layers were washed with brine and dried over Na$_2$SO$_4$. The solvent was evaporated and the residue was purified by silica gel column chromatography. Yield 65%, $R_f$: 0.75 (1:20, EtOH:EtOAc). $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 1.15 (d, $J$ = 6.6 Hz, 3H), 1.29 (t, $J$ = 7.1 Hz, 3H), 1.46 (s, 9H), 2.64 (d, $J$ = 5.9 Hz, 2H), 3.41(ABq, $J$ = 17.4 Hz, 2H), 3.74 (m, 1H), 4.21 (q, $J$ = 7.1 Hz, 2H), 4.79 (brs, 1H). ESI-MS (positive mode): m/z 261 ($M_{calc}$ for C$_{12}$H$_{24}$N$_2$O$_4$: 260.33).

Boc $\gamma$-ala PNA thymine monomer ethyl ester (9a-i and 9a-ii): To a stirred solution of thymine acetic acid (T-CH$_2$CO$_2$H, 0.96g, 5.28 mmol) in dry DMF (20 ml) were added DCC (1 equiv., 5.28 mmol, 1.08g) and DhbtOH (1 equiv., 5.28 mmol, 0.85g) under argon atmosphere. The stirring was continued for 1 h at room
temperature. Compound 5a-i (1.14g, 4.4 mmol) in dry DMF (5 mL) was added and the reaction was stirred at 50 °C for 14 h. Following evaporation of the solvent, the residue was partitioned between EtOAc (100 mL) and saturated NaHCO₃ solution (100 mL). The organic layer was washed with 10 % KHSO₄ (3 x 50 mL), 10 % NaHCO₃ (3 x 50 mL), brine (100 mL) and dried over Na₂SO₄. The solvent was removed under reduced pressure and purified by silica gel column chromatography to afford 9a-i as off white foamy solid. The same procedure was adopted to get 9a-ii as well from 5a-ii. Yield 60%, R_f = 0.75 (20:1, EtOAc:EtOH). ¹H NMR (300 MHz, DMSO-d₆, 2 rotamers): δ 0.95 & 1.08 (2xd, J = 6.67 Hz, 3H), 1.17 & 1.24 (2xt, J = 7.15 Hz, 3H), 1.36 & 1.37 (2xs, 9H), 1.75 (s, 3H), 3.02-3.31 (m, 3H), 3.46-3.78 (m, 1H), 3.93-4.21 (m, 3H), 4.3 & 4.46 (2xs, 1H), 4.67 (ABq, J = 16.4 Hz, 1H), 6.67 & 6.85 (brs, 1H), 7.23(s, 1H), 11.24 & 11.26 (2xs, 1H). ¹³C NMR (75 MHz, DMSO-d₆): δ 11.8, 13.9, 17.8, 18.2, 24.4, 25.2, 28.1, 33.2, 44.5, 47.6, 48.1, 49.1, 51.7, 52.4, 60.3, 61.0, 77.6, 77.9, 108.1, 148.1, 150.9, 155.1, 164.3, 167.4, 167.8, 168.7, 169.1. ESI-MS (negative mode): m/z 869 (Mcalc for C₄₃H₅₀N₈O₁₂: 870.9).

**Boc-γ-ala PNA thymine monomer (10a-i and 10a-ii):** To a solution of 9a-i (1.3g, 1.49 mmol) in THF (20 mL) was added drop-wise 2N sodium hydroxide (20 mL) at 0 °C. After stirring for 30 min at the same temperature, water (100 mL) was added and the pH was adjusted to 4 at which point precipitate was formed. Filtered the precipitate, washed the solid with cold water and dried under vacuum to obtain the required title compound 10a-i (1.17g) Yield: 95%, R_f = 0.1 (EtOAc:EtOH = 4:1). The same procedure was adopted for the synthesis of compound 10a-ii using 9a-ii as starting material. ¹H NMR (300 MHz, DMSO-d₆, 2 rotamers): δ 0.96 & 1.08 (2xd, J =
6.9 Hz, 3H), 1.37 & 1.38 (2xs, 9H), 1.75 (s, 3H), 3.01-3.46 (m, 2H), 3.65-4.05 (m, 3H), 4.45 (s, 1H), 4.69 (ABq, J = 16.7 Hz, 1H), 6.7 & 6.88 (brs, 1H), 7.24 (s, 1H), 11.26 (s, 1H). \(^{13}\)C NMR (75 MHz, DMSO-d\(_6\)): \(\delta\) 11.8, 17.8, 18.2, 28.7, 44.4, 44.6, 47.6, 48.0, 50.1, 51.9, 52.2, 77.6, 77.9, 108.0, 141.8, 150.9, 155.1, 164.3, 167.1, 167.8, 170.5, 171.1; ESI-MS (positive mode): m/z 848.8 (M\(_{\text{calc}}\) for C\(_{41}\)H\(_{46}\)N\(_8\)O\(_{12}\): 847.9).

**PhFl-γ-ala PNA thymine monomer ethyl ester (9b):** The aforementioned procedure for compounds 9a-i and 9a-ii was used for the synthesis of 9b as white solid from the starting material 5b. Yield 75%, \(R_f\): 0.7 (80:20, EtOAc:EtOH). \(^1\)H NMR (300 MHz, CDCl\(_3\), 2 rotamers): \(\delta\) 0.75 & 0.86 (2xd, \(J = 6.48\) Hz, 3H), 1.22 & 1.30 (2xt, \(J = 7.15\) Hz, 3H), 1.89 (2xs, 3H), 2.15 & 2.7 (2xq, \(J = 6.48\) Hz), 2.9-3.25 (m, 2H), 3.6 & 3.75 (ABq, \(J = 17.29\) Hz, 2H), 4.10 & 4.16 (2xq, \(J = 7.15\) Hz, 2H), 4.25 & 4.30 (2xABq, \(J = 16.33\) Hz, 2H), 6.82-6.89 (2xq, \(J = 1.15\) Hz), 7.16-7.4 (m, 11H), 7.64-7.76 (m, 2H), 8.08 (s, 1H). \(^{13}\)C NMR (75 MHz, CDCl\(_3\)): \(\delta\) 11.8, 13.9, 20, 20.4, 20.7, 47.1, 47.7, 53.8, 54.9, 59.6, 60.3, 61, 72.8, 107.9, 120, 125.3, 125.9, 126.7, 127.6, 128, 139.7, 140, 141.9, 145.3, 145.5, 149.5, 150.8, 164.3, 167.2, 167.4, 168.4, 168.9. ESI-MS (positive mode): m/z 567.1 (M\(_{\text{calc}}\) for C\(_{33}\)H\(_{34}\)N\(_4\)O\(_5\):566.2)

**PhFl-γ-ala PNA thymine monomer (10b):** The aforementioned procedure for compounds 10a-i and 10a-ii was used for the synthesis of 10b as white solid from the starting material 9b. Yield 80%, \(R_f\): 0.6 (80:20, EtOAc:EtOH). \(^1\)H NMR (300 MHz, DMSO-d\(_6\), 2 rotamers): \(\delta\) 0.52 & 0.67 (2xd, \(J = 6.01\) Hz, 3H), 1.72 & 1.74 (2xs, 3H), 2.19 (2xq, \(J = 6.01\) Hz, 1H), 2.8-3.3 (m, 2H), 3.3-4.0 (m, 2H), 7.10-7.55 (m, 12H), 7.7-7.9 (m, 2H), 11.20 (1bs, 1H). \(^{13}\)C NMR (75 MHz, CDCl\(_3\)): \(\delta\) 11.8, 14, 21, 19.9, 47.8, 49.1, 54.6, 72.8, 107.9, 120.2, 125.5, 125.9, 126.8, 128, 139.8, 141.9, 150.8,
164.3, 167.1, 170, 170.3, 171.9. ESI-MS (positive mode): m/z 539.1 (M_{calc} for C_{31}H_{30}N_{4}O_{5}:538.2)

**Synthesis of MTPA-derivative of γ-Alanine-PNA backbones 11a-i and 11a-ii:** To the stirred solution of Boc protected PNA backbones 5a-i and 5a-ii (0.5g, 1.92 mmol) was added 5% m-cresol in TFA (7 mL) in DCM at 0 °C. The reaction mixture was stirred for 1 h at rt. Solvents was evaporated under reduced pressure and the resulting residue was triturated with diethyl ether (3 x 5 mL). The residue was dried under high vacuum for several hours, and then dissolved in anhydrous DCM (5 mL) at room temperature under nitrogen atmosphere. DIPEA (2.5 equiv., 4.6 mmol, 0.8 mL) and MTPA-Cl (1.5 equiv., 2.8 mmol, 0.6g) were added to the dissolved solution at 0 °C and the reaction mixture was stirred for 14 h at rt. The solution was diluted with DCM (50 mL) and washed with water (50 mL), saturated NaHCO₃ (30 mL) and saturated NaCl solutions (30 mL). The organic layer was dried over anhydrous Na₂SO₄, filtered and concentrated in vacuo and the residue was purified by silica gel column chromatography to get the title compound 11a-i and 11a-ii. Yield 70%, Rf: 0.75 (1:20, EtOH: EtOAc). $^1$H NMR (300 MHz, DMSO-d$_6$): 1.10 (d, J = 6.7 Hz, 3H), 1.18 (t, J = 7.1 Hz, 3H), 2.55 & 2.59 (2xd, J = 7.0 Hz, 2H), 3.27(ABq, J = 17.2 Hz, 2H), 3.31 (brs, 1H), 3.40 & 3.41 (2xs, 3H), 3.98 (m, 1H), 4.08 (q, J = 7.1 & 14.3 Hz, 2H), 7.43 (m, 3H), 7.55 (m, 2H), 8.09 (d, J = 8.7 Hz, 1H); $^{13}$C NMR (75 MHz, DMSO-d$_6$): 13.9, 17.9, 44.4, 49.7, 52.7, 54.6, 60.1, 127.3, 128.2, 129.4, 132.9, 165.1, 172.

**Synthesis of MTPA-derivative of γ-Alanine-PNA backbone 11b:** Compound 5b (0.5g, 1.92 mmol) was dissolved in anhydrous methanol (10 mL) at room temperature under nitrogen atmosphere. 10% Pd/C (150 mg) was added carefully under nitrogen
atmosphere. The resulting reaction mixture was stirred for 2 h under hydrogen atmosphere. After the disappearance of starting material as judged by TLC, the reaction mixture was filtered through celite pad and washed with methanol. The filtrate was concentrated in vacuo and the deprotected-free amine was directly taken to the next step without further purification. Amine (0.3g, 1.87 mmol) was dissolved in anhydrous DCM (5 mL) at room temperature under nitrogen atmosphere. DIPEA (1.5 equiv., 2.8 mmol, 0.45 mL) and MTPA-Cl (1.5 equiv., 2.8 mmol, 0.5 g) were added at 0 °C to the solution. The reaction mixture was stirred at room temperature for 14 h under nitrogen atmosphere and quenched by adding water (30 mL). The reaction mixture was extracted with dichloromethane and the organic layer was washed with water (25 mL), saturated NaHCO₃ (30 mL), saturated NaCl (30 mL) solutions and dried over anhydrous Na₂SO₄. The solution was filtered, and the solvent was removed under reduced pressure and the resulting material was subjected to silica gel column chromatography to obtain the required compound 12b. Yield: 70%, Rₜ: 0.75 (1:20, EtOH:EtOAc).

**Synthesis of MTPA-derivative of γ-Alanine PNA Thymine monomers, 12a-i and 12a-ii:** Boc-group in compound 10a-i and 10a-ii was deprotected using 5% m-cresol in TFA. Deprotected-amine (0.5g, 0.6 mmol) was dissolved in a mixture of water and acetone (1:1, 10 mL) at room temperature. Solid NaHCO₃ (3 equiv., 1.8 mmol, 0.15 g) and MTPA-Cl (1.5 equiv., 0.9 mmol, 0.18g) were added and the reaction mixture was stirred for 14 h at rt. Acetone was removed in vacuo and the resulting reaction mixture was diluted with water and extracted with ethyl acetate (2 x 20 mL). Combined organic layers were washed with saturated NaCl (20 mL) solution, dried
over anhydrous Na$_2$SO$_4$ and filtered. The solvent was concentrated under reduced pressure and the residue was purified by column chromatography to afford MTPA derivatives 12a-i and 12a-ii. Yield 65%, R$_f$: 0.1 (1:20, EtOH:EtOAc). $^1$H NMR (300 MHz, DMSO-d$_6$): 1.06 (d, $J$ = 6.5 Hz, 3H), 1.16-1.2 (2xt, $J$ = 7.0 Hz, 3H), 1.72 & 1.74 (2xs, 3H), 3.11-3.36 (m, 2H), 3.34 & 3.45 (2xs, 3H), 3.78 (ABq, $J$ = 18.8 Hz, 1H), 3.91 (ABq, $J$ = 18.8 Hz, 1H), 4.09-4.28 (m, 1H), 4.29-4.54 (m, 1H), 4.59 (ABq, $J$ = 16.8 Hz, 1H), 7.09 & 7.14 (2xs,1H), 7.38-7.57 (m, 5H), 8.23 & 8.38 (2xd, $J$ = 8.5 Hz, 1H), 11.27 (brs, 1H). $^{13}$C NMR (75 MHz, DMSO-d$_6$): δ 11.8, 17.4, 44.1, 47.6, 51.6, 52.2, 54.9, 83.5, 107.9, 126.9, 128.2, 129.3, 133.4, 142.1, 151, 164.4, 164.9, 168.4, 171.3.

**Synthesis of MTPA-derivative of γ-Alanine PNA Thymine monomers, 12b:**

Compound 10b was hydrogenated using the protocol described in the synthesis of compound 11b and the resulting amine was converted to compound 12b by using a procedure mentioned in the synthesis of compounds 12a-i and 12a-ii.

**N-tert-Butoxycarbonyl Alanal (4a-ii) (Swern oxidation route):** Oxaly chloride (2 equiv., 5.70 mmol, 0.48 mL) was added to dry DCM (5 ml) taken in a two-neck round bottom flask equipped with a stirrer under argon atmosphere. The flask was cooled to -78 °C. To it was added DMSO (4 equiv., 11.40 mmol, 0.80 mL) drop-wise and stirring was continued for 30 min at -78 °C. Next, the alcohol 6 (0.5g, 2.85 mmol) dissolved in dry DCM (5mL) was added to the reaction mixture drop-wise and stirred for further 1h. Then Et$_3$N (3 equiv., 8.55 mmol, 1.19 mL) was added with stirring at -78 °C. After 5 min the cooling bath was removed and the reaction mixture was allowed to warm to room temperature. Upon completion of the reaction as checked by TLC, the reaction was
quenched by NH$_4$Cl solution drop-wise. The organic layer was then separated. The aqueous phase was re-extracted with CH$_2$Cl$_2$ (20ml) and the organic layers were combined and washed with 10% citric acid solution, and brine. The organic layer was dried over anhydrous Na$_2$SO$_4$ and evaporated to give the crude product. The crude product was carried forward to the next step immediately.

**N-tert-Butoxycarbonyl Alanal (4a-iii) (Parikh-Doering oxidation route):** To a stirred solution of alcohol 6 (0.5g, 2.85 mmol) in CH$_2$Cl$_2$/DMSO (1:1, 10.0 mL), SO$_3$•py (4 equiv., 11.40 mmol, 1.8g) and Et$_3$N (3 equiv., 8.55 mmol, 1.19 mL) were added at 0 °C. The resulting mixture was then stirred for 1 h at room temperature before it was quenched with H$_2$O (10 mL) and extracted with DCM (2 × 10 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$ and concentrated to afford the crude product which was carried forward to the next step immediately without any purification.

**Synthesis of MTPA-derivative of alanol 14a-i, a-ii, a-iii:** To the stirred solution of Boc protected alanol 13a-i, 13a-ii and 13a-iii (0.5g, 2.85 mmol) was added 5% m-cresol in TFA (7 mL) in DCM at 0 °C. The reaction mixture was stirred for 1 h at rt. Solvents was evaporated under reduced pressure and the resulting residue was column purified to get the corresponding amine product. This amine (0.05g, 0.66 mmol) was then dissolved in anhydrous DCM (5 mL) at room temperature under nitrogen atmosphere. DIPEA (2.5 equiv., 1.66 mmol, 0.3 mL) and MTPA-Cl (1.3 equiv., 0.86 mmol, 0.25g) were added to the dissolved solution at 0 °C and the reaction mixture was stirred for 14 h at rt. The solution was diluted with DCM (10 mL) and washed with water (10 mL), saturated NaHCO$_3$ (10 mL) and saturated NaCl solutions (10 mL). The organic layer was dried over anhydrous Na$_2$SO$_4$, filtered and concentrated in vacuo and the residue was purified
by silica gel column chromatography to get the title compound **14a-i**, **14a-ii** and **14a-iii**.

Yield 80%, $R_f$: 0.5 (40:60, EtOAc: hexane). $^1$H NMR (300 MHz, CDCl$_3$): 1.23 (d, J = 6.8 Hz, 3H), 2.58 (brs, 1H), 3.38 (s, 3H), 3.52(m, 1H), 3.62 (m, 1H), 4.11 (m, 1H), 6.98 (d, J = 7.07 Hz, 1H), 7.39-7.56 (m, 5H); $^{13}$C NMR (75 MHz, CDCl$_3$): 16.7, 28.3, 47.6, 54.8, 66.3, 120.4, 122.7, 125.0, 127.3, 127.8, 128.6, 129.5,132.4,166.7.

II.5. References


II.6. Appendices

*Figure 1A.* $^1$H NMR Spectrum of compound 2a (CDCl$_3$, 300 MHz)

*Figure 2A.* $^{13}$C NMR Spectrum of compound 2a (DMSO-d$_6$, 75 MHz)
Figure 3A. $^1$H NMR Spectrum of compound 3a (CDCl$_3$, 300 MHz)

Figure 4A. $^{13}$C NMR Spectrum of compound 3a (CDCl$_3$, 75 MHz)
Figure 5A. $^1$H NMR Spectrum of compound 4a-i (CDCl$_3$, 300 MHz)

Figure 6A. $^{13}$C NMR Spectrum of compound 4a-i (CDCl$_3$, 75 MHz)
**Figure 7A.** $^1$H NMR Spectrum of compound 5a-i (CDCl$_3$, 300 MHz)

**Figure 8A.** $^{13}$C NMR Spectrum of compound 5a-i (CDCl$_3$, 75 MHz)
Figure 9A. $^1$H NMR Spectrum of compound 2b (CDCl$_3$, 300 MHz)

Figure 10A. $^{13}$C NMR Spectrum of compound 2b (DMSO-d$_6$, 75 MHz)
**Figure 11A.** $^1$H NMR Spectrum of compound 3b (CDCl$_3$, 300 MHz)

![H NMR Spectrum of compound 3b](image)

**Figure 12A.** $^{13}$C NMR Spectrum of compound 3b (CDCl$_3$, 75 MHz)

![13C NMR Spectrum of compound 3b](image)
Figure 13A. $^1$H NMR Spectrum of compound 4b (CDCl$_3$, 300 MHz)
Figure 14A. $^1$H NMR Spectrum of compound 5b (CDCl$_3$, 300 MHz)

Figure 15A. $^{13}$C NMR Spectrum of compound 5b (CDCl$_3$, 75 MHz)
Figure 16A. $^1$H NMR Spectrum of compound 6 (CDCl$_3$, 300 MHz)

Figure 17A. $^{13}$C NMR Spectrum of compound 6 (CDCl$_3$, 75 MHz)
**Figure 18A.** $^1$H NMR of compound 8 (CDCl$_3$, 300 MHz)

* Double the integrated peaks value to obtain the title compound 8

**Figure 19A.** $^{13}$C NMR of compound 8 (CDCl$_3$, 75 MHz)
Figure 20A. $^1$H NMR Spectrum of compound 9a-i (DMSO-d$_6$, 300 MHz)

Figure 21A. $^{13}$C NMR Spectrum of compound 9a-i (DMSO-d$_6$, 75 MHz)
Figure 22A. $^1$H NMR Spectrum of compound 10a-i (DMSO-d$_6$, 300 MHz)

Figure 23A. $^{13}$C NMR Spectrum of compound 10a-i (DMSO-d$_6$, 75 MHz)
Figure 24A. $^1$H NMR Spectrum of compound 9b (CDCl$_3$, 300 MHz)

Figure 25A. $^{13}$C NMR Spectrum of compound 9b (DMSO-d$_6$, 75 MHz)
Figure 26A. $^1$H NMR Spectrum of compound 10b (DMSO-d$_6$, 300 MHz)

* Double the integrated peaks value to get the compound 10b correct NMR value

Figure 27A. $^{13}$C NMR Spectrum of compound 10b (DMSO-d$_6$, 75 MHz)
Figure 28A. $^1$H NMR Spectrum of compound 11a-i, 11a-ii (DMSO-d$_6$, 300 MHz)

Figure 29A. $^{13}$C NMR Spectrum of compound 11a-i, 11a-ii (DMSO-d$_6$, 75 MHz)
**Figure 30A.** $^1$H NMR Spectrum of compound 12a-i, 12a-ii (DMSO-d$_6$, 300 MHz)

![NMR Spectrum](image)

**Figure 31A.** $^{13}$C NMR Spectrum of compound 12a-i, 12a-ii (DMSO-d$_6$, 75 MHz)

![NMR Spectrum](image)
Figure 32A. MALDI-TOF Spectrum of oligomer, PNA2a
Figure 33A. MALDI-TOF Spectrum of oligomer, PNA2b
Figure 34A. MALDI-TOF Spectrum of oligomer, PNA3a
Figure 35A. MALDI-TOF Spectrum of oligomer, PNA3b
Figure 36A. $^1$H NMR Spectrum of compound 14a-i (CDCl$_3$, 300 MHz)

Figure 37A. $^{13}$C NMR Spectrum of compound 14a-i (CDCl$_3$, 75 MHz)
III.1. Introduction

The ease of synthesis and the flexibility in structural modification of PNA, especially in the backbone, make it an attractive nucleic acid platform for further chemical derivatization. We thus concentrated much of our effort on making chemical modifications in the backbone of PNA in order to further improve its hybridization properties, biocompatibility, and cellular uptake, so that it could be effectively used for biological and biomedical research, and possibly therapeutic applications. Of the three achiral centers in the PNA backbone, the gamma-position was the least explored.\(^1\) Because of the scarcity in the chemical modification at this position, and also because of our promising initial modeling work, we have made a concerted effort to explore the effects of chirality and chemical functionality on the conformation and hybridization properties of PNA. Our group has demonstrated that installation of a (S)-Me group at the γ-backbone resulted in a right-handed helical induction of PNA,\(^2\) and the resulting molecular “foldamer” exhibits unusually strong binding affinity and sequence-specificity for DNA as well as RNA\(^3\). Further, we showed that γPNA of an appropriate length, 13-20 nucleotides, can invade any sequence of double helical B-DNA at physiological temperature, with recognition occurring through direct Watson-Crick base-pairing.\(^4\) Revelation of such appealing properties has garnered considerable interest in PNA in a variety of applications, from regulation to diagnostics and therapeutics to molecular
engineering, and thus demands larger-scale monomer production and parallel oligomer synthesis.\textsuperscript{5} However, the current Boc-chemistry approach, typically employed by our group and others,\textsuperscript{6} is not compatible for solid-phase synthesis automation due to the harsh TFMSA cleavage conditions employed that are known to rapidly corrode the instrument. Further, the harsh cleavage condition has proven to be problematic for attaching acid-sensitive dyes and other chemical moieties to PNA during solid-phase synthesis. To circumvent this problem, we have devised a new synthetic strategy to prepare Fmoc-monomers, enabling the synthesis of γPNA oligomers through automation and incorporation of a wide-range of acid-sensitive chemical moieties. Although there are reports on the synthesis of Fmoc-γPNA monomers bearing different functional groups at the gamma-backbone, this is the first on the synthesis of (R)-Fmoc-MPγPNA monomers.

III.2. Design rationale

In designing the synthetic scheme for preparing Fmoc-γPNA monomers, we adhered to two requirements. The first is optical purity. In this regard, we have decided to follow the Mitsunobu route in the preparation of the backbone intermediate because it has been shown to produce optically-pure monomers, as demonstrated in Chapter II. Due to
the lability of the Fmoc-protecting group in the presence of a strong base, which is required in the initial alkylation step, we have decided to delay the Fmoc-protection until a later stage. These decisions have led to the development of a synthesis strategy shown in Scheme 1 for the preparation of (R)-MPγPNA monomers.

III.3. Results

III.3.1. Monomer synthesis

The new synthetic route has enabled the preparation of (R)-Fmoc-MPγPNA monomers (Scheme 1) from a commercially available and relatively cheap Boc-L-serine, as a starting material. (R)-Fmoc-MPγPNA monomers containing all four natural nucleobases (A, C, G, T) were synthesized according to the procedures outlined in Scheme 2. In the first step, alkylation of Boc-L-serine was performed with side-chain 12. Instead of using methyl ether as a protecting group for the miniPEG-sidechain, we have switched to t-butyl because it can be readily removed under a relatively mild acidic condition (TFA), as compared to methyl ether which requires a harsh (TFMSA) cocktail. The first alkylation step was carried out under an optimized reaction condition developed in our group to get optically-pure product. Formation of the mixed anhydride of the alkylated product 2, followed by reduction with sodium borohydride yielded serinol 3. Temperature maintenance and slow addition of sodium borohydride dissolved in water were found to be crucial in obtaining high reaction yield. In this reaction, borohydride complex is formed as a byproduct. Formation of such byproduct could be minimized by drop-wise addition of the substrate at -15 °C.

After removing the impurity we coupled 3 to N-(2-nitrophenylsulfonamido) glycine methyl acetate, which was prepared according to a published procedure,
resulting in the formation of compound 4. In the process we have discovered that the Mitsunobu reaction produced many byproducts. The similarity in the polarity of these compounds made the product purification by column chromatography highly tedious. For small scale synthesis, column purification can be accomplished without much effort; however, for large-scale synthesis, purification has proven to be extremely difficult, and as a result we could not isolate the product cleanly. The triphenylphosphine oxide byproduct bled in the column, and after several rounds of purification, we still could not obtain the desired product with high degree of purity. In search of a more effective purification strategy, we have discovered that selective washing of the reaction mixture with hexane and ether removed nearly 80% of the impurities and the remaining (solid) material was of sufficient purity for next reaction step.
Figure III-1. Top: $^1$H NMR of pure Mitsunobu product 4; Bottom: $^1$H NMR of impure Mitsunobu product 4.

Figure III-1A shows the $^1$H NMR of the purified Mitsunobu product 4 performed on a small-scale. However, when the same reaction was performed on a large-scale (Figure III-1B), purification became excruciatingly tedious. Pure product could not be isolated. Comparison of the two spectra reveals that there are additional peaks besides that of the product. One set of protons correspond to the DIAD byproduct generated in the reaction (structure shown in the spectrum). The peak at δ 1.1 ppm could be assigned to protons $n$ while those at δ 5.1 ppm to protons $o$ of the DIAD byproduct. Another set of protons in the spectrum correspond to the triphenylphosphine oxide byproduct. The peaks at δ 7.3 (m, 10H), 7.9 (m, 2H), 8.1 ppm correspond to the aromatic protons of the triphenylphosphine oxide. Thus, on a large-scale synthesis, even after several rounds of purification we could not obtain pure Mitsunobu product 4. Nonetheless, with selective
washing with hexane and ether we were able to obtain roughly 80% purity and this product mixture was subjected to the next step.

Next, the Boc-protecting group was removed in the presence of 4 N HCl. There have been several reports for selective removal of Boc-protecting group in the presence of tert-butyl; however, many of them require special conditions such as 4 N HCl/dioxane\textsuperscript{9} and 4 N HCl/EtOAc\textsuperscript{10} under anhydrous condition. We have attempted to develop a simpler and more economical method for selective removal of Boc. While the method works well for small-scale synthesis, providing nearly quantitative yield of the desired product, it is less selective for large-scale synthesis, resulting in the removal of Boc as well as tert-butyl group. In order to plow through the reaction scheme, while at the same time obtaining pure product with selective Boc removal, we have decided on a brute force approach by carrying out the reaction in several small aliquots.

Protection of the resulting primary amine 5 with Fmoc-OSU in DCM at 0 °C yielded compound 6. With the Fmoc-protection in place, we next removed the 2-nitrobenzenesulfonyl group by treatment with thiophenol in acetonitrile in the presence of Cs\textsubscript{2}CO\textsubscript{3} at room temperature. Under this condition we found that a small but acceptable amount of Fmoc-protecting group was concomitantly removed. Since we have found that the backbone intermediate (compound 7) was not stable over a prolonged period, we proceeded immediately to the next step after column purification. Initially we had tried DCC/DhbtOH coupling, but due to the formation of many byproducts in this reaction, we obtained only a modest yield (~70%) of the desired ester product. To improve the reaction yield, we standardized the reaction condition using HBTU as a coupling reagent. Under this condition, we obtained over 80% in coupling efficiency for all four
nucleobases (A, C, G, T). Subsequent hydrolysis of the resulting esters gave the desired monomers 1a-d. The Boc-protected nucleobases (A, C, G, T) were synthesized according to the established procedures.\textsuperscript{11}

\begin{center}
\textbf{Scheme 2. Synthesis of R-MP-γPNA monomers, Reagents and conditions:} (a) NaH, DMF, RI, 0 °C-rt; (b) NMM, isobutyl chloroformate, DME, -15 °C, 1h, NaBH\(_4\), 1h; (c) TPP, 2-nitrobenzenesulfonyl glycine methyl ester, DIAD, THF, 0 °C -rt; (d) 4N HCl, THF:H\(_2\)O (1:1,v/v) 0 °C -rt; (e) Fmoc-OSU, DIPEA, DCM, 0 °C -rt, 12h; (f) Cs\(_2\)CO\(_3\), PhSH, ACN, 0 °C -rt, ; (g) B-CH\(_2\)CO\(_2\)H, NMM, HBTU, DMF, rt; (h) 2N NaOH, THF, 0 °C.
\end{center}

**III.3.2. Synthesis of miniPEG-sidechain**

Compound 12 was synthesized from the starting material 10 according to the procedures outlined in **Scheme 3**. We observed a complete conversion of the chloro- to iodo-derivative using 5 equivalents of NaI, refluxing in acetone for 12 hours. This was confirmed by NMR. However for large scale synthesis, due to the use of a large excess of NaI, the handling of such reaction was difficult. Through a series of reaction optimization
steps, we have discovered that 3 equivalents of NaI were sufficient for complete conversion of the chloro to iodo compound. The R_f values for both compounds 10 and 11 were the same under various TLC conditions examined. Thus, the conversion from the chloro to iodo derivative could not be determined by TLC and had to be confirmed by NMR. The spectra of compounds 10 and 11 are shown in Figure III-2.

Scheme 3. Synthesis of the side-chain. Reagents and conditions: (a) NaI, Acetone, reflux; (b) Boc₂O, Mg(ClO₄)₂, DCM, 0 °C-rt.
In the NMR spectrum of SM 10 we observed two sets of protons: one in the range of δ (3.70-3.75) ppm assigned to protons c & d; while the others in the range of δ (3.55-3.65) ppm assigned to protons a & b. During this chemical transformation it was found that one set of protons (c & d) remain unchanged while the other had shifted considerably. Due to the shielding effect of iodo in comparison to chloro functionality, protons a shifted to an up-field region of δ 3.25 ppm while protons b became slightly shielded, δ (3.55-3.60) ppm. Thus, the NMR analysis clearly demonstrates the successful conversion from SM 10 to the corresponding iodo derivative.

The primary alcohol 11 was then protected with tert-butyl group by refluxing with Boc-anhydride in the presence of magnesium perchlorate in DCM. The last reaction step was tricky to handle, especially for large-scale synthesis, due to the highly exothermic nature of this chemical transformation and a possible run-away reaction. We addressed this issue by first carrying out the reaction in an ice-bath and then gradually allowed the reaction to warm to room temperature and stirred at the same temperature.

**Figure III-2.** Top: $^1$H NMR of SM 10; Bottom: $^1$H NMR of compound 11.
overnight. Such a condition gave the same chemical yield but circumvented the concern for run-away reaction. We have also carried the reaction in reverse order, whereby the primary alcohol was first protected with tert-butyl group, followed by conversion of terminal chloro to iodo compound. This sequence of reactions however did not work well. We found that the tert-butyl group fell off during the chloro-to-iodo exchange as the result of a large excess of NaI. While it was not effective in the preparation of our desired sidechain, we have learned that this is a rather mild and neutral condition for removing tert-butyl, in contrast to the traditional methods which employ relatively strong acidic conditions. For our work, the miniPEG-sidechain was synthesized according to Scheme 3.

Most of the existing methodologies for the protection of alcohol with tert-butyl group require harsh conditions based on the reaction of alcohol with isobutylene in the presence of a strong acidic catalyst. The method discussed here provides a mild and robust route for the incorporation of a tert-butyl group. A plausible mechanism for this protection step was suggested by Bartoli and coworkers.

![Scheme 4](image-url)  
**Scheme 4.** A plausible mechanism for the formation of tert-butyl ether.
It was found that Boc$_2$O decomposed in the presence of Mg(ClO$_4$)$_2$ to give CO$_2$, t-BuOH, and isobutylene. Based on this observation, the formation of a t-butyl carbocation intermediate was assumed. One plausible scenario is that the carbocation is captured by the alcohol present in the reaction mixture to give the t-butyl ether. This hypothesis was ruled out because of the lack of evidence for electrophilic substitution reaction of t-butyl carbocation with activated Phenyl rings. However, the fact that Boc$_2$O decomposes in the presence of Mg(ClO$_4$) explained the use of excess of Boc$_2$O to drive the reaction to completion. The possibility of a carbamate formation was also ruled out. It was reported in the literature that during the synthesis of RO-Boc from R-OH and Boc$_2$O in presence of DMAP, a mixed dicarbonate intermediate 14 was formed. Thus based on this finding, Bartoli and coworkers synthesized the t-butyl cyclopentyl dicarbonate as a mixture with the corresponding carbonate. When this inseparable mixture was treated with Mg(ClO$_4$)$_2$, formation of the corresponding t-butyl ether was detected. This study gave rise to the hypothesis that initially a mixed dicarbonate intermediate is formed. A plausible mechanism for the next step is that the metal catalyst co-ordinates with the lone pair of electrons of the carbonyl groups to form a six membered transition state. This facilitates the decomposition to the tert-butyl ether and CO$_2$ by a concerted mechanism. Studies are underway in the Bartoli lab to further validate the mechanism of this reaction.

### III.4. Discussion

The significance of this work lies in two aspects. The first is the optical purity of the monomers. So far most of Fmoc-γPNA backbones were prepared using the reductive amination route. Optically active aminoaldehydes are typically prepared by reduction of
Weinreb amides. Reductive amination of these chiral aldehydes gave the \( \gamma \)-modified PNA backbone. However, the susceptibility of these aminoaldehydes to epimerisation greatly limits the scope of this methodology to obtain optically-pure monomers.\(^{14}\) Winssinger et al. reported a protecting group combination (Mtt/Boc) orthogonal to Fmoc-based solid-phase synthesis.\(^{15}\) Appella et al.\(^{16}\) reported PNA oligomers bearing Lys side-chain for attaching various functionalities, while Oliver et al.\(^{17}\) developed cysteine-based \( \gamma \)PNA monomers for native chemical PNA ligation. Using the same methodology various Fmoc-\( \gamma \)PNA monomers were synthesized.\(^{18}\) Due to the concern for epimerization, as discussed in Chapter II, we opted for the Mitsunobu synthetic route to get optically-pure (\( R \))-Fmoc MP\( \gamma \)PNA monomers following the steps shown in Scheme 2 & 3. The second aspect is the incorporation of the MiniPEG side-chain at the gamma backbone. In our previous study we showed that the incorporation of the MiniPEG sidechain greatly enhanced the water solubility and biocompatibility of the \( \gamma \)PNA oligomer, which in turn minimizes off-target binding.\(^{7,19}\) It was done by alkylation of the carboxylic acid under basic condition in the first step. Using the existing methodology for synthesizing Fmoc-\( \gamma \)PNA monomers (Weinreb amide route), the MiniPEG side-chain cannot be introduced into the \( \gamma \)PNA backbone because Fmoc is susceptible to cleavage under basic condition. Thus, the synthetic methodology described in this Chapter provides a simple and robust means for synthesizing \( (R) \)-Fmoc-MP\( \gamma \)PNA monomers.

**III.5. Conclusion**

The enormous potential of PNA for biological and biomedical research has motivated scientists to develop new PNA analogues with superior properties. A large number of modified PNAs have been obtained to date.\(^{5}\) We have developed \( \gamma \)PNA by
introducing a substituent at the gamma backbone. This simple strategy has proven fruitful in improving DNA/RNA binding affinity and sequence selectivity. γPNA is the only class of nucleic acids that has been shown to be capable of invading any sequence of B-DNA under stimulated physiological buffer condition. GPNAs provide a solution to the poor cellular uptake of unmodified PNA. Considering the application of γPNAs for therapeutic purposes we need to expedite the basic research on issues such as toxicity, pharmacokinetics and pharmacodynamics. For this, a straightforward synthetic protocol for preparing monomers that is suitable for automation and can accommodate a wide range of chemical moieties is necessary. In this Chapter, we report the first synthetic methodology for preparing (R)-Fmoc-MPγPNA monomers. In a simple, economical, and robust synthetic procedure following the Mitsunobu route optically-pure (R)-FmocMP γPNA monomers were synthesized. Based on this methodology, sufficient quantities of γPNA can be prepared for biological and animal experiments in the near future.

However, for a larger scale (10-100 g), synthesis of the monomers has proven to be problematic owing to the difficulty in the purification of Mitsunobu intermediate and in the selective removal of Boc protecting group in the presence of tert-butyl group. Even though the methodology described here is a rather simple, robust and economic one, additional work needs to be done to improve the feasibility of the existing one or develop a new route for scaling-up the synthesis of γPNA monomers. In conclusion, the work describes herein paves the way for future development and exploration of γPNA in biology and medicine.

III.6. Experimental section
Materials and Methods. All commercial reagents were used without further purification. Solvents were dried by standard methods and freshly distilled prior to use. $^{1}$H-, $^{13}$C-NMR spectra were recorded on a Bruker Avance AV-300 NMR spectrometers using standard Bruker software. Flash chromatography was performed using standard grade silica gel (60A, 65x250mesh) from Sorbent Technologies. TLC was performed with Silica XHL Plates (w/UV254, glass backed, 250µm, 20x20cm) from Sorbent Technologies. Mass spectra were recorded on a Finnigan LCQ ESI/APCI ion trap mass spectrometer by electrospray ionization.

O-(2-(2-(tert-butoxy)ethoxy)ethyl)-N-(tert-butoxycarbonyl)-L-serine (2): Anhydrous DMF (300 mL) was added to icebath-cooled NaH (60%, 2.5 equiv, 1227 mmol, 49 g) taken in a flame dried 3 neck rb flask and stirred for 10 min. L Boc-serine (1 equiv, 491 mmol, 100 g) dissolved in dry DMF (200 mL) was added to the reaction mixture via an addition funnel. The reaction was then continued to stir at 0 °C for ~ 2h until the bubbling ceased. Next, the side chain 12 (1.5 equiv, 736 mmol, 199 g) dissolved in dry DMF (100 mL) was added to the reaction mixture drop-wise via an addition funnel, and the reaction was allowed to warm to room temperature overnight. TLC was checked (EtOAc/EtOH: 20/1; Rf: 0.3). DMF was removed under reduced pressure at 40 °C. 200 mL ice-cold water was added and the pH of the mixture was adjusted to ~4. It was then extracted with EtOAc (3x 300 mL). The combined organic layers were washed with 1M HCl (100 mL), brine (300 mL) and dried over anhydrous Na$_2$SO$_4$. Removal of the solvent under reduced pressure yielded the crude product 2, which was purified by column chromatography to give the pure product in 70% yield.

$^1$H NMR (300 MHz, DMSO-$d_6$): $\delta$ 1.12 (s, 9H), $\delta$ 1.38 (s, 9H), 3.29-3.69 (m, 10H),
4.12 (m, 1H), 6.84 (d, \(J=8.15\) Hz, 1H); \(^{13}\)C NMR (75 MHz, DMSO-\(d_6\)): \(\delta\) 27.7, 28.5, 54.3, 61.1, 70.1, 70.3, 70.5, 70.9, 72.7, 78.6, 155.7, 172.4; HRMS (ESI TOF) \(M\) calcd for \(C_{16}H_{31}NO_7\) \([M+H]^+\) 349.2101, found 350.2171.

**Tert-butyl \((R)-(1-(2-(2-(tert-butoxy) ethoxy) ethoxy)-3-hydroxypropan-2-yl) carbamate (3):** To the cold, stirring solution of compound 2 (1 equiv, 286 mmol, 100g) in DME (400 mL), NMM (1.1 equiv, 315 mmol, 35 mL) was added drop-wise over the course of 30 min. After this, isobutyl chloroformate (1.1 equiv, 315 mmol, 42 mL) was added drop-wise over 30 min and the reaction was then continued to stir in ice-bath for 1h. The reaction mixture was then filtered off, and the solid washed with DME (2x 50 mL). The combined filtrate was cooled in an ice-bath. To it, NaBH\(_4\) (2 equiv, 572 mmol, 21.6 g) dissolved in 100 mL water was added portion wise and the mixture was continued to further stir for 1 h at 0 °C. TLC was checked (EtOAc/Hexane: 50/50; \(R_f\): 0.3). The solvent was then removed under reduced pressure. EtOAc (500 mL) was added to the crude mixture and washed with saturated NaHCO\(_3\) (300 mL) and then brine (300 mL). The organic layer was dried over anhydrous Na\(_2\)SO\(_4\). Removal of the solvent gave the crude mixture, which on column purification afforded the desired product 3 as viscous liquid. \(^1\)H NMR (300 MHz, DMSO-\(d_6\)): \(\delta\) 1.12 (s, 9H), \(\delta\) 1.37 (s, 9H), 3.30-3.51 (m, 13H), 4.57 (m, 1H), 6.44 (d, \(J=8.07\) Hz, 1H); \(^{13}\)C NMR (75 MHz, DMSO-\(d_6\)): \(\delta\) 27.7, 28.6, 52.5, 61.1, 61.2, 70.2, 70.3, 70.9, 72.7, 78.0, 155.6; HRMS (ESI TOF) \(M\) calcd for \(C_{16}H_{33}NO_6\) \([M+H]^+\) 335.2380 found 336.2381.

**Methyl \((R)-11-((tert-butoxycarbonyl) amino)-2, 2-dimethyl-13-((2-nitrophenyl) sylfonyl)-3, 6, 9-trioxa-13-azapentadecan-15-oate (4):** Compound 3 (1.1 equiv, 298
mmol, 100g) was dissolved in dry THF (500 mL) and stirred in an ice-bath for 10 min under inert atmosphere. To it, TPP (1.1 equiv, 298 mmol, 58 mL) and methyl ((2-nitrophenyl)sulfonyl)glycinate (1 equiv, 271 mmol, 74.3 g) were added and stirred for 5 min. DIAD (1.1 equiv, 298 mmol, 58 mL) was then added to the reaction mixture drop-wise at 0 °C and then the reaction was allowed to warm to room temperature overnight. TLC was checked (EtOAc/Hexane: 50/50; Rf: 0.4). THF was removed under reduced pressure. The residue was dissolved in EtOAc (200 mL) and washed with saturated solution of NaHCO₃, brine (2 x 50 mL) and dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure to get the crude mixture which on column purification yielded the desired pure product as a viscous yellow liquid. For the small scale synthesis, the crude product was column purified to get the pure product. But for the large scale, after removing THF under reduced pressure, hexane (250 mL) and diethyl ether (250 mL) were added to the residue, swirled to mix and stored at -40 °C overnight. The solvent was decanted and the above process was repeated thrice. After getting rid of most of the impurities, the mixture was extracted with EtOAc. Following the similar work up as discussed above, the product thus obtained was carried to the next step without any further column purification. ¹H NMR (300 MHz, CDCl₃): δ 1.14 (s, 9H), δ 1.41 (s, 9H), 3.09 (d, 1H, J=6.8 Hz), 3.24 (s, 1H), 3.49-3.78 (m, 13H), 3.81 (s, 1H), 7.61-7.70 (m, 2H), 7.74-7.83 (m, 2H); ¹³C NMR (75 MHz, CDCl₃): δ 27.4, 28.3, 46.4, 51.5, 53.0, 53.4, 54.2, 61.1, 70.3, 70.5, 71.1, 73.2, 76.8, 124.7, 125.5, 130.7, 131.5, 132.6, 133.0, 133.9, 156.0, 168.2; HRMS (ESI TOF) M calcd for C₂₅H₄₁N₃O₁₁S[M+H]+ 591.2462 found 592.1822.
Methyl (R)-11-amino-2, 2-dimethyl-13-((2-nitrophenyl)sulfonyl)-3, 6, 9-trioxa-13-azapentadecan-15-oate (5): Compound 4 (1 equiv, 149 mmol, 88 g) was dissolved in THF (187 mL) and stirred at 0 °C. 187 mL of 4N HCl (10 equiv) was added to the reaction mixture drop-wise via an addition funnel. The reaction was then allowed to warm to room temperature. The progress of the reaction was monitored by TLC (DCM/MeOH: 90/10; Rf: 0.5). After the completion of the reaction, it was given hexane (3x 100 mL) wash to remove the non-polar impurities. The aqueous layer was cooled to 0 °C and neutralised with 187 mL of 4N NaOH solution drop-wise. 187 ml of saturated NaHCO₃ solution was added to the reaction mixture and was then extracted with EtOAc (3x 300 mL). The combined organic layer was dried over anhydrous Na₂SO₄. After concentration of the solvent, the crude compound was purified by column chromatography to yield the desired product as a foamy off white material. ¹H NMR (300 MHz, CDCl₃): δ 1.21 (s, 9H), 3.30-4.08 (m, 17H), 4.2-4.4 (m, 1H), 6.94 (1H), 7.52-7.81(m,3H), 8.01-8.9(m,1H); ¹³C NMR (75 MHz, CDCl₃): δ 27.4, 44.1, 47.9, 51.3, 61.1, 70.5, 70.8, 71.2, 71.5, 76.8, 77.0, 77.3, 124.4, 124.5, 130.9, 131.3, 131.9, 134.3, 165.2; HRMS (ESI TOF) M calcd for C₂₀H₃₅N₃O₉S [M+H]⁺ 491.1938 found 492.1922.

Methyl(R)-11-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-2,2-dimethyl-13-((2-nitrophenyl)sulfonyl)-3,6,9-trioxa-13-azapentadecan-15-oate (6): To the solution of compound 5 (1 equiv, 17.3 mmol, 8.5 g) in dry DCM (200 mL) at 0 °C was added DIPEA (1.2 equiv, 20.1 mmol, 3.6 mL) drop-wise and stirred for 15 min. Fmoc-OSU (1 equiv, 17.3 mmol, 5.8 g) was then added to the reaction mixture and continued to stir overnight at room temperature. TLC was checked (EtOAc/Hexane: 80/20; Rf: 0.5).
After the completion of the reaction, the organic layer was washed with 10% citric acid (3x 300 mL), saturated NaHCO₃ (3x 100 mL) and brine (300 mL). The DCM layer was dried over anhydrous Na₂SO₄, removed under reduced pressure to give the crude product, which on column purification afforded compound 6 as a white, foamy, crusty material. ¹H NMR (300 MHz, CDCl₃): δ 1.1 (s, 9H), 3.29-3.54 (m, 15H), 3.81 (m, 1H), 4.16-4.3 (m, 5H), 7.21 (d, J=8.07 Hz, 1H), 7.32 (t, J=7.2 Hz, 2H), 7.41 (t, J=6.9 Hz, 2H), 7.68 (d, J=7.1 Hz, 2H), 7.8-8.1 (m, 6H); ¹³C NMR (75 MHz, CDCl₃): δ 27.5, 47.1, 49.1, 49.3, 52.2, 61.1, 66.8, 70.0, 70.4, 71.2, 73.1, 119.9, 124.1, 125.2, 127.1, 127.6, 128.5, 131.0, 131.7, 132.1, 133.6, 141.2, 143.9, 156.5, 169.5; HRMS (ESI TOF) M calcd for C₃₅H₄₃N₃O₁₁S[M+H]⁺ 713.2618 found 714.2690.

Methyl (R)-(2-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-3-((2-(2-(tert butoxy) ethoxy) ethoxy) propyl) glycinate (7): Compound 6 (1 equiv, 7 mmol, 5g) was dissolved in ACN (20 mL) and stirred at 0 °C under an inert atmosphere. To this reaction mixture, Cs₂CO₃ (2.2 equiv, 15.42 mmol, 5.02 g) was added and after 5 min thiophenol (1.1 equiv, 7.7 mmol, 78 mL) was added drop-wise to the stirring solution at 0 °C. Once the reaction was completed as determined by TLC (EtOAc/Hexane: 80/20; Rf: 0.1), the reaction mixture was filtered and the solvent was removed under reduced pressure. EtOAc (50 mL) was added and then washed with saturated NaHCO₃ (2x 30 mL) and brine (30 mL). The organic layer after drying over anhydrous Na₂SO₄ was removed under reduced pressure to give the crude product, which was column purified to give the titled compound 7 as a clear viscous liquid. ¹H NMR (300 MHz, CDCl₃): δ 1.11 (s, 9H), 3.36-3.62 (m, 15H), 4.1-4.4 (m, 3H), 7.12 (d, J=8.07 Hz, 1H), 7.34 (t, J=6.9 Hz, 2H), 7.41 (t, J=7.1 Hz, 2H), 7.71 (d, J=7.2 Hz, 2H), 7.88 (d, J=6.8 Hz
$^1$H; $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 27.4, 47.2 48.2, 48.4, 49.8, 50.3, 61.0, 66.8, 70.0, 70.1, 70.5, 71.6, 71.7, 123.4, 123.4, 125.6, 127.5, 128.1, 141.3, 141.4, 143.7, 143.9, 156.3, 168.5; HRMS (ESI TOF) M calcd for C$_{29}$H$_{40}$N$_2$O$_7$[M+H]$^+$ 527.2404 found 528.2806.

**Methyl (R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-2, 2-dimethyl-13-(2-(5-methyl-2,4-dioxo-3, 4-dihydropyrimidin-1(2H)-yl)acetyl)-3, 6, 9-trioxa-13-azapentadecan-15-oate (8a):** To a stirred solution of thymine acetic acid (1.2 equiv, 5.12 mmol, 1g) in dry DMF (10 mL) under an inert atmosphere was added HBTU (1.2 equiv, 5.12 mmol, 1.9g) and stirred for 10min. After that, NMM (2.2 equiv, 10.24mmol, 1.2mL) was added drop-wise to the rb and the mixture was further stirred for 30 min. At this point, compound 7(1equiv, 4.26mmol, 2.25g) dissolved in dry DMF (10 mL) was added to the reaction mixture and continued to stir overnight. TLC was checked to confirm the completion of the reaction. DMF was removed under reduced pressure. To it H2O (20 mL) was added and the compound was extracted with EtOAc (2x 25 mL). The organic layer was washed with 10% citric acid (2x 10 mL), saturated NaHCO3 (2x 10 mL) and brine (15 mL). The EtOAc layer was dried over anhydrous Na2SO4, removed under reduced pressure to give the crude product, which on column purification afforded compound 8a as a white, foamy, crusty material: yield 82%. 1H NMR (300 MHz, DMSO-d6) major rotamer: $\delta$ 1.12 (s, 9H), 1.12-1.1.32 (m, 2H), 1.71 (s, 3H), 3.2-3.61 (m,10H), 3.62 (s, 3H), 3.7 (m, 2H), 4.0-4.7 (m,9H), 7.2-7.47 (m, 4H), 7.52 (t, J=6.4 Hz, 1H), 7.7 (d, J=7.2 Hz, 2H), 7.9(d, J=6.9 Hz, 2H), 7.96(s,1H); $^{13}$C NMR (75 MHz, DMSO-d6): $\delta$ 12.3, 28.6, 38.7, 48.1, 48.3, 49.4, 52.3, 52.8, 69.6, 70.1, 70.2, 70.3, 71.0, 72.5, 78.6, 108.5,
108.6, 120.5, 120.6, 127.9, 128.6, 141.2, 142.4, 151.4, 155.8, 164.8, 167.8, 168.4, 170.7, 171.7; HRMS (ESI TOF) M calcd for C36H46N4O10 [M+H]+ 694.3214 found 695.3221.

Methyl (R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-((2-((tert-butoxycarbonyl) amino)-2-oxopyrimidin-1(2H)-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxo-13-azapentadecan-15-oate (8b): Cytosine ester 8b was prepared, purified, and characterised the same way as described for 8a. $^1$H NMR (300 MHz, DMSO-$d_6$) major rotamer: $\delta$ 1.09 (s, 9H), 1.43 (s, 9H), 3.3-3.6 (m, 11H), 3.62 (s, 3H), 3.71 (m, 1H), 3.86-4.4 (m, 6H), 4.59-4.88 (m, 2H), 7.2 (d, $J=8.07$ Hz, 1H), 7.32 (t, $J=6.4$ Hz, 2H), 7.43 (t, $J=6.7$ Hz, 2H), 7.7 (d, $J=7.07$ Hz, 2H), 7.84 (t, $J=7.2$ Hz, 1H), 7.89 (d, $J=7.1$ Hz, 2H), 10.28 (s, 1H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 27.4, 28.0, 38.6, 47.1, 48.5, 48.9, 49.6, 50.0, 52.3, 52.8, 61.0, 66.7, 66.8, 68.8, 70.2, 70.6, 70.6, 71.0, 73.2, 119.9, 125.1, 126.5, 127.1, 127.7, 141.3, 141.3, 143.7, 143.9, 149.7, 156.3, 163.0, 167.6, 168.5, 169.6; HRMS (ESI TOF) M calcd for C$_{40}$H$_{53}$N$_5$O$_{11}$ [M+H]$^{+}$ 779.3742 found 780.3748.

Methyl (R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-((2-((tert-butoxycarbonyl) amino)-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxo-13-azapentadecan-15-oate (8c): Adenine ester 8c was prepared, purified, and characterised the same way as described for 8a. $^1$H NMR (300 MHz, CDCl$_3$) major rotamer: $\delta$ 1.21 (s, 9H), 1.42 (s, 9H), 3.3-3.84 (m, 15H), 3.9-4.5 (m, 6H), 4.9-5.4 (m, 2H), 7.28 (m, 2H), 7.37 (q, $J=6.8$ Hz, 2H), 7.55 (m, 2H), 7.76 (m, 2H), 8.05 (m, 1H), 8.68 (s, 1H); $^{13}$C NMR (75 MHz, DMSO-$d_6$): $\delta$ 27.4, 27.9, 47.2, 48.6, 49.1, 50.3, 52.3, 61.0, 66.8, 70.1, 71.1, 71.6, 73.1, 83.7, 120.3, 124.8, 125, 127.0, 127.1, 127.8, 139.6, 141.3, 143.5, 143.6, 147.6,
149.0, 155.8, 156.3, 167.1, 169.5; HRMS (ESI TOF) M calcd for C₄₁H₅₃N₇O₁₀[M+H]⁺ 803.3854 found 804.3882.

Methyl (R)-11-((((9H-fluoren-9-yl)methoxy) carbonyl) amino)-13-(2-(2-((tert-butoxycarbonyl) amino)-6-oxo-1, 6-dihydro-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oate (8d): Guanine ester 8d was prepared, purified, and characterised the same way as described for 8a. ¹H NMR (300 MHz, DMSO-ｄ₆) major rotamer: δ 1.08 (s, 9H), 1.42 (s,9H), 3.23-3.57 (m, 9H), 3.61 (s,3H), 3.7-3.8 (m,2H), 4.0-4.5(m, 5H), 7.25 (t, J=7.1 Hz, 1H), 7.3(m,2H), 7.4(m,2H), 7.52-7.72(m,2H), 7.8 (d, J=8.07 Hz, 2H), 7.86-7.92 (m,2H), 10.6 (s,1H), 11.3(s,1H); ¹³C NMR (75 MHz, CDCl₃): δ 27.4, 27.9, 43.6, 47.2 48.7, 49.0, 50.2, 52.3, 61.0, 66.7, 69.6, 70.2, 70.7, 71.1, 73.2, 83.8, 120.0, 124.8, 124.9, 127.0, 127.1, 127.8, 139.6, 141.3, 141.3, 143.5, 143.6, 147.6, 149.0, 155.7, 156.3, 167.1, 169.5; HRMS (ESI TOF) M calcd for C₄₁H₅₃N₇O₁₁[M+H]⁺ 819.3803 found 820.3821.

(R)-11-((((9H-fluoren-9-yl) methoxy) carbonyl) amino)-2, 2-dimethyl-13-(2-(5-methyl-2, 4-dioxo-3, 4-dihydropyrimidin-1(2H)-yl) acetyl)-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (1a): To the stirred solution of 8a (0.5 g, 0.72 mmol, 1equiv) in THF (5 mL) 2N NaOH (10 equiv, 7.2 mmol, 0.3g) was added drop-wise at 0 °C. After 10 min, the reaction was neutralized with 1N HCl (12 equiv), resulting in the formation of a cloudy white solution. It was extracted with EtOAc (3x 10 mL) and washed with brine (5 mL). The organic layer was dried over anhydrous Na₂SO₄, filtered and removed under reduced pressure. To the crude oily material a small amount of DCM was added, swirled to mix, and then crashed with diethyl ether to get white sticky solid. The solvent
was decanted, and the crude was column purified to yield the pure product as a white crystalline material. $^1$H NMR (300 MHz, DMSO-$d_6$): δ 1.12 (s, 9H), 1.72 (s, 3H), 3.21-3.6 (m, 12H), 3.7-4.1 (m, 3H), 4.11-4.4 (m, 2H), 4.45 (s, 1H), 4.6 (s, 1H), 7.25 (s, 1H), 7.33 (m, 2H), 7.42 (t, $J$=7.1 Hz, 2H), 7.7 (m, 2H), 7.9 (d, $J$=6.9 Hz, 2H), 11.3 (s, 1H); $^{13}$C NMR (75 MHz, CDCl$_3$): δ 12.3, 27.9, 47.2, 48.2, 48.4, 49.8, 50.3, 65.8, 66.0, 70.0, 70.1, 70.3, 70.4, 71.7, 78.0, 108.4, 120.5, 125.6, 127.5, 128.1, 141.2, 144.3, 151.4, 156.3, 164.8, 165, 168, 170.7; HRMS (ESI TOF) M calcd for C$_{35}$H$_{44}$N$_4$O$_{10}$[M+H]$^+$ 680.3057 found 681.3102.

(R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-((2-((tert-butoxycarbonyl) amino)-2-oxopyrimidin-1(2H)-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (1b): Cytosine monomer 1b was prepared, purified, and characterised the same way as described for 1a. $^1$H NMR (300 MHz, DMSO-$d_6$): δ 1.1 (s, 9H), 1.4 (s, 9H), 3.21-3.6 (m, 11H), 3.7-4.4 (m, 7H), 4.61-4.86 (m, 2H), 6.93 (m, 1H), 7.32 (t, $J$=7.1 Hz, 2H), 7.41 (t, $J$=6.9 Hz, 2H), 7.7 (d, $J$=6.5 Hz, 2H), 7.82 (m, 1H), 7.87 (d, $J$=6.8 Hz, 2H); $^{13}$C NMR (75 MHz, CDCl$_3$): δ 27.4, 28.2, 47.2, 48.4, 49.9, 50.4, 65.8, 66.0, 70.0, 70.1, 70.3, 70.4, 71.6, 71.7, 81.3, 94.3, 120.5, 125.7, 127.5, 128.0, 141.2, 144.3, 150.8, 150.9, 152.6, 155.3, 156.3, 167.9, 170.7, 171.1; HRMS (ESI TOF) M calcd for C$_{39}$H$_{51}$N$_5$O$_{11}$[M+H]$^+$ 765.3585 found 766.4117.

(R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-((2-((tert-butoxycarbonyl) amino)-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (1c): Adenine monomer 1c was prepared, purified, and characterised the same way as described for 1a. $^1$H NMR (300 MHz, DMSO-$d_6$): δ 1.11
(s, 9H), 1.5 (s, 9H), 3.38-3.7 (m, 11H), 3.9-4.4 (m, 6H), 7.32 (m, 2H), 7.41 (m, 2H), 7.7 (m, 2H), 7.88 (d, , J=7.1 Hz, 2H), 8.25 (s, 1H), 8.5 (s, 1H), 10.612 (s, 1H); $^{13}$C NMR (75 MHz, CDCl$_3$): δ 27.7, 28.6, 47.2, 48.4, 48.5, 48.9, 49.8, 50.3, 66.0, 70.0, 70.1, 70.12, 70.4, 71.0, 71.6, 71.7, 80.4, 83.8, 120.5, 123.4, 123.5, 125.6, 127.5, 128.1, 141.2, 144.3, 144.4, 145.3, 150.1, 151.6, 151.8, 152.7, 156.4, 167.4, 170.6; HRMS (ESI TOF) M calcd for C$_{40}$H$_{51}$N$_7$O$_{10}$[M+H]$^+$ 789.3697 found 790.3721.

(R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(2-((tert-butoxycarbonyl) amino)-6-oxo-1, 6-dihydro-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (1d): Guanine monomer 1d was prepared, purified, and characterised the same way as described for 1a. $^1$H NMR (300 MHz, DMSO-$d_6$): δ 1.13 (s, 9H), 1.41 (s, 9H), 2.9-3.2 (m, 2H), 3.2-3.8 (m, 11H), 3.91-4.41 (m, 6H), 4.9 (s, 1H), 5.1 (m, 1H), 7.21-7.45 (m, 4H), 7.57-7.73 (m, 2H), 7.78 (m, 1H), 7.86 (d, J=6.8Hz, 2H), 11.33 (s, 1H); $^{13}$C NMR (75 MHz, CDCl$_3$): δ 27.4, 28.1, 44.2, 47.2 48.5, 48.8, 49.9, 50.9, 66.0, 70.0, 70.1, 70.5, 71.6, 71.7, 82.9, 83.0, 119.2, 120.5, 120.6, 125.5, 125.6, 127.5, 127.9, 128.0, 128.1, 140.6, 140.8, 141.1, 141.2, 144.3, 147.9, 149.9, 154.0, 155.4, 156.6, 167.4, 170.7; HRMS (ESI TOF) M calcd for C$_{40}$H$_{51}$N$_7$O$_{11}$[M+H]$^+$ 805.3647 found 806.3651.

2-(2-iodoethoxy) ethan-1-ol (11): To the solution of the SM10 (1 equiv, 401.5 mmol, 50g) dissolved in acetone (250 mL), NaI (3 equiv, 1.2 mol, 180.5 g) was added and then refluxed overnight while stirring vigorously. The progress of the reaction was monitored by TLC (EtOAc/Hexane: 10/90; R$_f$: 0.4). After the completion of the reaction, the solid was filtered off, acetone was removed under reduced pressure; H$_2$O was added and then
extracted with EtOAc (2x 250 mL). The organic layer was then washed with saturated solution of Na$_2$S$_2$O$_5$ (150 mL), brine (100 mL), and dried over anhydrous Na$_2$SO$_4$. Concentration of the organic layer then gave the titled compound as colourless, viscous oil. $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 3.25 (t, 2H), 3.58 (m, 2H), 3.68-3.76 (m, 4H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 3.2, 61.6, 68.6, 71.5; (ESI TOF) M calcd for C$_4$H$_9$IO$_2$ [M+H]$^+$ 215.9647 found 216.9807.

2-(2-(2-iodoethoxy) ethoxy)-2-methylpropane (12): Compound 11 (1 equiv, 277.7 mmol, 60g) was dissolved in dry DCM (250 mL) and cooled to 0 °C. To it were added Mg(ClO$_4$)$_2$ (0.1 equiv, 27.7 mmol, 6.2g) and Boc$_2$O (2.3 equiv, 638.8 mmol, 139.5g) at 0 °C and continued to stir for 10 min. The reaction mixture was then allowed to warm to room temperature overnight. TLC was checked to monitor the reaction (EtOAc/Hexane: 10/90; $R_f$: 0.8). The reaction was then quenched by the addition of H$_2$O (50 mL). The organic layer was then washed with saturated solution of NaHCO$_3$ (2×100 mL), brine (50 mL), and dried over anhydrous Na$_2$SO$_4$. Concentration of the organic layer then gave the crude product which on column purification gave the titled compound as yellow, viscous oil. 87% yield. $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 3.25 (t, 2H), 3.49-3.54 (m, 2H), 3.58-3.64 (m, 2H), 3.76 (t, 2H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 3.1, 27.5, 61.2, 70.8, 72.0, 73.0; HRMS (ESI TOF) M calcd for C$_4$H$_9$IO$_2$ [M+H]$^+$ 272.0273 found 273.2836.

III.7. References


III.8. Appendices

*Figure 1A.* $^1$H NMR spectrum of compound 11 (CDCl$_3$, 300 MHz)

*Figure 2A.* $^{13}$C NMR spectrum of compound 11 (CDCl$_3$, 75 MHz)
**Figure 3A.** $^1$H NMR spectrum of compound 12 (CDCl$_3$, 300 MHz)

**Figure 4A.** $^{13}$C NMR spectrum of compound 12 (CDCl$_3$, 75 MHz)
Figure 5A. $^1$H NMR spectrum of compound 2 (DMSO-$d_6$, 300 MHz)

Figure 6A. $^{13}$C NMR spectrum of compound 2 (DMSO-$d_6$, 75 MHz)
Figure 7A. $^1$H NMR spectrum of compound 3 (DMSO-$d_6$, 300MHz)

Figure 8A. $^{13}$C NMR spectrum of compound 3 (DMSO-$d_6$, 75 MHz)
Figure 9A. $^1$H NMR spectrum of compound 4 (CDCl$_3$, 300MHz)

Figure 10A. $^{13}$C NMR spectrum of compound 4 (CDCl$_3$, 75 MHz)
Figure 11A. $^1$H NMR spectrum of compound 5 (CDCl$_3$, 300MHz)

Figure 12A. $^{13}$C NMR spectrum of compound 5 (CDCl$_3$, 75MHz)
Figure 13A. $^1$H NMR spectrum of compound 6 (DMSO-$d_6$, 300MHz)

Figure 14A. $^{13}$C NMR spectrum of compound 6 (CDCl$_3$, 75MHz)
Figure 15A. $^1$H NMR spectrum of compound 7 (DMSO-$d_6$, 300MHz)

Figure 16A. $^{13}$C NMR spectrum of compound 7 (CDCl$_3$, 75MHz)
Figure 17A. $^1$H NMR spectrum of compound 8a (DMSO-$d_6$, 300MHz)

Figure 18A. $^{13}$C NMR spectrum of compound 8a (CDCl$_3$, 75MHz)
Figure 19A. $^1$H NMR spectrum of compound 8b (DMSO-$d_6$, 300MHz)

Figure 20A. $^{13}$C NMR spectrum of compound 8b (CDCl$_3$, 75MHz)
**Figure 21A.** $^1$H NMR spectrum of compound 8c (CDCl$_3$, 300MHz)

![Figure 21A](image)

**Figure 22A.** $^{13}$C NMR spectrum of compound 8c (CDCl$_3$, 300MHz)

![Figure 22A](image)
Figure 23A. $^1$H NMR spectrum of compound 8d (DMSO-$d_6$, 300MHz)

Figure 24A. $^{13}$C NMR spectrum of compound 8d (CDCl$_3$, 75MHz)
Figure 25A. $^1$H NMR spectrum of compound 1a (DMSO-$d_6$, 300 MHz)

Figure 26A. $^{13}$C NMR spectrum of compound 1a (DMSO-$d_6$, 75 MHz)
Figure 27A. $^1$H NMR spectrum of compound 1b (DMSO-$d_6$, 300MHz)

Figure 28A. $^{13}$C NMR spectrum of compound 1b (DMSO-$d_6$, 75MHz)
Figure 29A. $^1$H NMR spectrum of compound 1c (DMSO-$d_6$, 300MHz)

Figure 30A. $^{13}$C NMR spectrum of compound 1c (DMSO-$d_6$, 75MHz)
Figure 31A. $^1$H NMR spectrum of compound 9d (DMSO-d$_6$, 300MHz)

Figure 32A. $^{13}$C NMR spectrum of compound 1d (DMSO-d$_6$, 75MHz)
CHAPTER IV

GENERAL METHOD FOR PREPARING OPTICALLY PURE (R)-AND (S)-FMOC-MPγPNA MONOMERS

IV.1. Introduction

As discussed in Chapter III, we have attempted to develop (R)-Fmoc-MPγPNA monomers in order to facilitate the automated synthesis of γPNA oligomer. Up to this point we have described a straightforward and economical methodology to synthesize Fmoc-MPγPNA monomers. Even though we could prepare such monomers, there were certain issues that need to be resolved. The previously developed methodology is not suitable for large-scale synthesis. It employed Mitsunobu coupling as a key step in the preparation of the backbone intermediate. In this Mitsunobu coupling step several byproducts were generated, one of them was triphenylphosphine oxide (TPPO). TPPO was difficult to separate from the reaction mixtures by means of chromatography. Even though subsequently we were able to remove the majority of it by selective washing with hexane and ether, the overall process rendered it unsuitable for large-scale monomer production. Another roadblock was in the selective removal of tert-butyl in the presence of Boc-protecting group. Both protecting groups were inadvertently removed during the deprotection step in the synthesis scale-up. These factors have motivated us to develop a synthetic strategy to overcome such shortcomings so that we could
synthesize optically-pure \((R)\)-Fmoc-M\(\gamma\)PNA monomers in a scalable, economical, and robust manner.

More recently we have shown that the right-handed (RH) and left-handed (LH) \(\gamma\)PNA are orthogonal to each other in recognition. Because of the mismatch in the helical sense, the two conformers are unable to hybridize to each other despite the sequence complementarity. The helical sense of \(\gamma\)PNA, \textit{i.e.} whether it adopts a right-handed or left-handed helical motif, is determined by the stereochemistry at the gamma backbone.\(^1\) \(\gamma\)PNAs prepared from \(L\)-amino acids adopt a right-handed helix, while those prepared from \(D\)-amino acids adopt a left-handed helix; however, only the right-handed helical \(\gamma\)PNAs hybridize to complementary DNA and RNA with high affinity and sequence selectivity. On the other hand, unmodified PNA can recognize both the RH-\(\gamma\)PNA and LH-\(\gamma\)PNA. Thus, utilizing this unique property of \(\gamma\)PNA, we envisioned developing a \(\gamma\)PNA platform for biotechnological applications and material science, along with our ongoing efforts to develop \(\gamma\)PNA for antisense and antigene technology. With this goal in mind, we sought to develop a general strategy for preparing both \((R)\)- and \((S)\)-Fmoc-M\(\gamma\)PNA monomers, starting from the same commercially available and relative cheap chemical reagent.

**IV.2. Rational design**

While designing the new synthetic scheme we kept a few things in mind. Firstly, to maintain optical purity we avoided the aminoaldehyde route,
and instead decided to reduce the amino acid to an alcohol in the initial step to avoid the possibility for epimerization. Secondly, based on the findings described in Chapter II, we used a sterically-bulky amino protecting group to shield the alpha proton from deprotonation by the base that would lead to epimerization. Thirdly, in order to avoid tedious purification of the alkylated serine product, we deferred the alkylation step to a later stage. Fourthly, we devised the scheme in such a way so that the two enantiomers are prepared from the same starting material. Lastly, the Fmoc-protecting group was installed towards the end of the reaction scheme to avoid the possibility of it being inadvertently removed in the earlier reaction steps. Taking these points into consideration, we devised the following schemes for preparing (R)- and (S)-Fmoc-MPγPNA monomers (Scheme 1).

**Scheme 1.** Chemical structure of Fmoc MP-containing γPNA units.

IV.3. Results

IV.3.1. Monomer synthesis
Scheme 2. Synthesis of (R)-MPγPNA monomers. Reagents and conditions: (a) K₂CO₃, AcN, BnBr, rt; (b) DHP, PPTS, DCM, 50 °C; (c) LAH, THF, 0 °C; (d) Et₃N, MsCl, DCM, 0 °C; (e) NaN₃, DMF, 80 °C; (f) pTSA, MeOH, 80 °C; (g) NaH, THF, 0 °C, 10 min, 50 °C, 2h, add compound 17, 50 °C; (h) Zn/NH₄Cl, EtOH/H₂O (3:1), 0 °C-rt; (i) Et₃N, CH₂BrCOOCH₃, DCM, 0 °C-rt; (j) B-CH₂CO₂H, NMM, DCC, DhbtOH, DMF, 50 °C; (k) H₂-Pd(OH)₂/C, MeOH, rt; (l) Fmoc-OSU, DIPEA, THF, 0 °C-rt; (m) 2N NaOH, THF, 0 °C.

IV.3.1a. (R)-Fmoc-MPγPNA monomers

(R)-Fmoc-MPγPNA monomers containing all four natural nucleobases (A, C, G, T) were synthesized according to the procedures outlined in Scheme 2 & 3. In the first step, benzylation of L-serine methyl ester 3 with benzyl bromide in THF/H₂O afforded the intermediate 3-I in a quantitative yield. We
wanted to protect the primary alcohol in such a way so that it can be easily cleaved under a mild acidic condition. Thus, considering the simplicity of the reaction and the ease of removal, we protected the alcohol group with THP. This was accomplished by refluxing the alcohol with DHP in the presence of a catalytic amount of PPTS under reflux condition. Subsequent reduction of the methyl ester with LAH afforded compound 4. No purification was needed for the first three steps, with each step producing nearly quantitative yield and only trace amounts of impurity.

Figure IV-1. $^1$H NMR spectrum of compound 4.

Figure IV-1 shows the proton NMR spectrum of the intermediate 4. Compound 4 is a key reactive intermediate for the synthesis of (R)- as well as (S)-Fmoc-MPγPNA monomers. Signals in the range of $\delta$ (1.4-2.0) ppm are
characteristic of protons $e$, $f$, $g$, and $h$ of the THP ring. A peak at $\delta$ 3.16 ppm indicates the presence of proton $b$, while the peak at $\delta$ 4.6 indicates that of proton $d$. The two benzylic protons $i$ and $j$ appear as a doublet in the region of 3.87-4.0 ppm. The 10 aromatic protons are accounted for in the region (7.2-7.6) ppm.

Next, the alcohol was mesylated by reacting compound 4 in DCM in the presence of MsCl and triethylamine. After acid/base workup, which was determined to be of sufficient chemical purity, the crude product mixture was treated with sodium azide under heating condition, yielding the desired product 4-II in an overall yield of 90%. Removal of THP with PTSA in MeOH under reflux condition yielded compound 5. Again, no column purification was necessary for compounds 4-I & 4-II, since they were relatively pure and used as crude mixtures in the subsequent steps.

Alkylation of compound 5 with sidechain 17 (Scheme 4) was performed under an optimized reaction condition. The backbone 6 was synthesized in a two-step process from compound 5: reduction of azide to amine by Zn/NH$_4$CL in EtOH/H$_2$O and the subsequent reaction of primary amine with methyl bromoacetate. The last reaction was performed under an optimized condition in order to minimize the production of a dialkylated byproduct and increase the overall reaction yield.
Figure IV-2. $^1$H NMR spectrum of compound 6.

The backbone 6 is an important intermediate in this reaction scheme. In the proton NMR the characteristic peaks of the compound can be found. The presence of a shielded, singlet peak accounting for 9 protons confirms the presence of the tert-butyl group, while the peak at $\delta 3.8$ (s, 3H) indicates protons $a$ of the methyl ester functional group. In addition to the aromatic protons, the peaks at $\delta 3.9$ (q, 4H) corresponding to protons $b$ and $c$ confirm the presence of the benzylic group. The protons $d$-$k$ corresponding to the MiniPEG side-chain can be seen clustered together in the region of $\delta$ (3.38-3.72) ppm.

$A^{\text{Boc}}$, $T$ and $G^{\text{Boc}}$ nucleobases were coupled to the benzyl backbone 6 using the standard DCC/DhbtOH coupling procedure. The nucleobases were
synthesized following the reported procedures. The monomer containing C\textsuperscript{Boc} nucleobase was prepared under a different condition, as described in the next section (Scheme 3), due to its incompatibility to the hydrogenation reaction. The resulting A, T and G esters were subjected to hydrogenation, performed in a Parr Hydrogenation apparatus using Pd(OH)\textsubscript{2}/C catalyst in dry MeOH under an optimized condition to remove the Bn-protecting group. Protection of the resulting primary amine with Fmoc-OSu yielded the (R)-Fmoc-MP\textsubscript{γ}PNA intermediates 8 (A, T, G). Hydrolysis of the ester afforded the corresponding (R)-Fmoc-MP\textsubscript{γ}PNA monomers 1a-c.

Figure IV-3. \textsuperscript{1}H NMR spectrum of compound 1a.

A representative proton NMR of (R)-Fmoc-MP\textsubscript{γ}PNA monomer is shown in Figure IV-3. The integral and splitting pattern of the peaks at δ (7.2-
7.9) ppm region indicates the presence of protons \( f, g, h \) and \( i \) Fmoc-functional group. The peaks at \( \delta 1.1 \) (s, 9H) and \( \delta 1.72 \) (s, 3H) correspond to the tert-butyl protons \( o \) and methyl proton of thymine, \( q \) respectively. The characteristic peaks of the protons of MiniPEG side-chain, \( j-n \) can also be seen in the region of \( \delta \) (3.2-3.7) ppm.

![Scheme 3](image)

**Scheme 3.** Synthesis of (R)-MPγPNA C monomer. Reagents and conditions: (a) \( \text{H}_2-\text{Pd(OH)}_2/C, \text{MeOH, rt} \); (b) Fmoc-OSU, DIPEA, THF, 0 °C-rt; (c) \( \text{C-CH}_2\text{CO}_2\text{H}, \text{NMM,DCC, DhbtOH, DMF, 50 °C} \); (d) 2N NaOH, THF, 0 °C.

It has been shown in the literature that hydrogenation resulted in a partial reduction of the cytosine nucleobase.\(^{2}\) To circumvent this challenge, we undertook different route to get to the \( \text{C}^{\text{Boc}} \)-monomer. First, the benzyl backbone 6 was subjected to a hydrogenation reaction. After removal of the Benzyl group, the primary amine was selectively protected with Fmoc using a known procedure.\(^{3}\) \( \text{C}^{\text{Boc}} \) nucleobase was coupled to the resulting Fmoc-protected backbone. Hydrolysis of the methyl ester backbone yielded the desired \( (R) \)-Fmoc-MPγPNA cytosine monomer 1d.

**IV.3.1b. (S)-Fmoc-MPγPNA monomers**

\( (S) \)-Fmoc-MPγPNA monomers were synthesized according to **Scheme 4**. The divergent in the synthesis of \( (R) \)- and \( (S) \)-Fmoc-MPγPNA monomers
start with compound 4. Alkylation of 4 with sidechain 17 under the optimized condition resulted in the inversion of stereochemistry at the gamma backbone. Removal of the THP-protecting group yielded compound 10. Preparation of intermediate 11 was carried out under identical condition as described above, mesylation followed by azidation. Reduction of azide to amine followed by alkylation with methyl bromoacetate gave backbone 12. Coupling of nucleobases to the backbone was performed under identical condition as described for the other enantiomer, yielding esters 13. Hydrogenation reaction to remove the Benzyl group, followed by Fmoc protection and subsequent hydrolysis of the ester group afforded the (S)-Fmoc-MPγPNA (A, T, G) monomers 2a-c in good yields. (S)-Fmoc-MPγPNA C-monomer 2d was synthesized under identical conditions as outlined in Scheme 5.
Scheme 4. Synthesis of (S)-MPγPNA monomers. Reagents and conditions: (a) NaH, THF, 0 °C, 10 min, 50 °C, 2h, add compound 12, 50 °C; (b) pTSA, MeOH, 80 °C; (c) Et3N, MsCl, DCM, 0°C; (d) NaN3, DMF, 80 °C; (e) Zn/NH4Cl, EtOH:H2O (3:1), 0 °C-rt; (f) Et3N, CH2BrCOOCH3, DCM, 0 °C-rt; (g) B-CH2CO2H, NMM, DCC, DhbtOH, DMF, 50 °C; (h) H2-Pd(OH)2/C, MeOH, rt; (i) Fmoc-OSU, DIPEA, THF, 0 °C-rt; (j) 2N NaOH, THF, 0 °C.

Scheme 5. Synthesis of (S)-MPγPNA C monomer. Reagents and conditions: (a) H2-Pd(OH)2/C, MeOH, rt; (b) Fmoc-OSU, DIPEA, THF, 0 °C-rt; (c) C-CH2CO2H, NMM, DCC, DhbtOH, DMF, 50 °C; (d) 2N NaOH, THF, 0 °C.

IV.3.2. Sidechain

The compound 17 was chosen for attachment as the sidechain to the gamma backbone. The rationale for using tert-butyl as protecting group is the same as discussed in Chapter III. In this case, we have chosen tosyl as a leaving group rather than the iodo because it is more stable under heating condition, whereas the latter has a tendency to undergo an elimination reaction. Changing the leaving group to tosyl stabilized the sidechain under an identical reaction condition and the yield of the alkylation step was found to be significantly improved. The desired sidechain 17 was synthesized from the commercially available starting material 15 in a two-step process. In the first step, monotosylation of the diol was done according to the established procedure to give compound 16. Using the optimized reaction condition
developed in our group, Boc-protection was performed on the terminal alcohol of 16 to afford the desired product 17 in good yield.

Scheme 6. Synthesis of Side-chain. Reagents and conditions: (a) Et$_3$N, Ts-Cl, DCM, 0 °C-rt; (b) Mg(ClO$_4$)$_2$, Boc anhydride, DCM, 0 °C-rt.

IV.4. Discussion

The synthetic schemes described in this Chapter provide a novel way of synthesizing both (R)- and (S)-Fmoc-MP$_\gamma$PNA monomers in large-scale. The first key point of this route is optical purity. As discussed in Chapter III, the existing methodologies for the preparation of Fmoc-γPNA monomers involve aminoaldehydes as key intermediates. Due to the susceptibility of the aldehydes to epimerization, the reductive amination route is not suitable for synthesizing optically-pure monomers. In this scheme, we convert the acid to an alcohol, thus bypassing the racemisation prone aldehyde intermediate. To minimize the chance of epimerisation we used dibenzyl protecting group of the amino moiety. The sterically bulky dibenzyl groups shield the alpha proton from deprotonation, preventing epimerisation from taking place. The monomers synthesized via these routes are optically-pure. The second salient feature of this route is scalability. As discussed earlier, the Mitsunobu route employed in the synthesis of the (R)-Fmoc-MP$_\gamma$PNA monomers is not suitable for large-scale synthesis. Due to a delicate in the handling of the
aminoaldehydes the reductive amination route is also not suitable for large-scale synthesis of the monomers. The new synthetic route circumvents all these obstacles. The steps employed in this reaction scheme are relatively simple and the intermediates formed are stable and could be stored over a prolonged period. At the beginning of the scheme, most of the steps gave nearly quantitative yield of the products. This enabled us to carry forward the crude materials to the next step without column purification. The large scale handling (100 g) of the reactions did not become an issue. The column purification steps are also straightforward which make this route highly suitable for large-scale synthesis of the monomers. The third important feature is the versatility of the synthetic scheme. Starting from the same commercially available cheap material, we can synthesize both enantiomers in a relatively straightforward series of steps. Compound 4 is a key intermediate from which the path to (R)- and (S)-Fmoc-MPγPNA monomers diverges. Another interesting aspect of this route lies in the incorporation of the MiniPEG side-chain at gamma PNA backbone. In the Mitsunobu route (Chapter III), we could not alkylate the carboxylic acid with the MiniPEG side-chain on a large-scale since prolonged exposure of Boc-L-Ser to a strong base (NaH) leads to epimerization. Chromatographic purification was also tedious because of the poor separation between the starting material and the alkylated product. Again as discussed in Chapter III, the reductive amination route involving Fmoc-amino acid as a starting material is incompatible with the alkylating condition. However, in this case alkylation of compound 5 with the MiniPEG side-chain
can be achieved on a large-scale with satisfactory yield and there is no issue with epimerization. Thus, this synthetic route is robust, synthetically flexible, versatile, economical, and scalable for preparing ($R$)- and ($S$)-Fmoc-$\gamma$PNA monomers. One issue with the current synthetic scheme however is the total number of steps. Even though it involves many steps, most of them are straightforward, give nearly quantitative yields and do not require chromatographic separation. If necessary, the alcohol group could be converted directly to azide bypassing the mesylation step, thereby shortens the overall reaction scheme by one step.

**IV.5. Conclusion**

Since the invention of PNA, due to its appealing features, there has been a considerable effort to further improve its properties from both the fundamental science and practical application standpoints. Synthetic chemists have extensively modified its structure. Most of the work on backbone modifications of PNA has attempted, with mixed success, to improve the binding affinity and sequence selectivity of Watson-Crick recognition of DNA and RNA. One such modification which shows the most promise is $\gamma$PNA. The ease and flexibility of synthesis, along with superior hybridization properties make $\gamma$PNAs an attractive nucleic acid platform for various biological and medical applications—as molecular tools as well as therapeutic and diagnostic reagents. In order to further explore the potentials of $\gamma$PNA for *in vivo* applications and expedite research on this front, a synthetic route to synthesize ($R$)-Fmoc-$\gamma$PNA monomers on a large-scale, compatible with
automated PNA synthesis, is essential. In this Chapter we report a simple yet robust procedure for making optically pure \((R)\)- and \((S)\)-Fmoc-\(\gamma\)PNA monomers in an economical and scalable manner. The synthetic methodology described here can afford the monomers in large quantities and a cost-effective way. The flexibility and versatility of this methodology is also amendable to the synthesis of \((S)\)-Fmoc-\(\gamma\)PNA monomers, starting from the same \((L)\)-Serine. Owing to its recognition orthogonality with the RH conformer, the LH-\(\gamma\)PNA provides a unique platform for molecular self-assembly, especially for in vivo applications. In conclusion, in this Chapter we report the synthesis of optically-pure \((R)\)- and \((S)\)-Fmoc-\(\gamma\)PNA monomers in a simple, robust, flexible, economical, and scalable way.

**IV.6. Experimental section**

**Materials and Methods.** All commercial reagents were used without further purification. Solvents were dried by standard methods and freshly distilled prior to use. \(^1\)H-, \(^{13}\)C-NMR spectra were recorded on a Bruker Avance AV-300 NMR spectrometers using standard Bruker software. Flash chromatography was performed using standard grade silica gel (60A, 65x250mesh) from Sorbent Technologies. TLC was performed with Silica XHL Plates (w/UV254, glass backed, 250\(\mu\)m, 20x20cm) from Sorbent Technologies. Mass spectra were recorded on a Finnigan LCQ ESI/APCI ion trap mass spectrometer by electrospray ionization.
Methyl dibenzyl-L-serinate (3-I): L-serine methyl ester hydrochloride (100g, 643 mmol,1equiv) was dissolved in 2L of acetonitrile and stirred with a mechanical stirrer for 30 min until lumps of the starting material dispersed into fine particles. Then K₂CO₃ (2.2 equiv, 2.12 mmol, 294 g) was added and stirring was continued for further 30 min. After this, BnBr (2 equiv, 1.286 mol, 152.7 mL) was added via an addition funnel drop wise and the mixture was stirred at room temperature overnight. The completion of the reaction was monitored by TLC (EtOAc/Hexane: 20/80; Rₚ: 0.3). Then the mixture was filtered and the solvent was removed under reduced pressure at 40 °C. The crude mixture thus obtained was then column purified to afford the desired product (170 g, 89%) as viscous liquid. ¹H NMR (300 MHz, CDCl₃): δ 3.60 (t, J=6.8 Hz, 1H), 3.71 (d, J=7.2 Hz, 2H), 3.78 (dd, J=6.2 Hz, J=6.4 Hz 2H), 3.83 (s,3H), 3.94 (d, J=7.1 Hz, 2H), 7.2-7.5 (m, 10H); ¹³C NMR (75 MHz, CDCl₃): δ 51.4, 54.8, 59.3, 61.8, 127.4, 128.5, 129.0, 138.7, 171.7; HRMS (ESI TOF) M calcd for C₁₈H₂₁NO₃ [M+H]⁺ 299.1521 found 300.1527.

Methyl N, N-dibenzyl-O-(tetrahydro-2H-pyran-2-yl)-L-serinate (3-II): To the solution of compound 3-I (100g, 334 mmol,1equiv) in dry DCM (1L) was added PPTS (0.2 equiv, 66.8 mmol, 17.1 g) and 2,3-DHP (1.5 equiv, 501 mmol, 46 mL). The reaction mixture was then refluxed at 50 °C overnight under an inert atmosphere. After the reaction was over as monitored by TLC (EtOAc/Hexane: 20/80; Rₚ: 0.6) it was quenched by the addition of H₂O and the organic phase was washed with sat.NaHCO₃ (3x 300 mL), 10% citric acid
(3x 300 mL) and then brine (300 mL). The organic phase was dried over anhydrous Na₂SO₄ and then the solvent was removed under reduced pressure to give the crude product which was purified by column chromatography to yield the desired product 3-II as a colorless viscous liquid (123g, 96%). ¹H NMR (300 MHz, CDCl₃): δ 1.41-1.9 (m, 7H), 3.5 (m, 1H), 3.68-4.1 (m, 10H), 4.62 (m,2H), 7.2-7.5 (m, 10H); ¹³C NMR (75 MHz, CDCl₃): δ 18.87,19.2, 25.4, 30.4, 51.2, 55.3, 60.9, 61.0, 61.1, 61.5, 62.0, 66.3, 66.6, 98.5,98.9, 126.9, 127.0, 128.2, 128.7, 139.5, 139.6, 171.9; HRMS (ESI TOF) M calcd for C₂₃H₂₉NO₄ [M+H]⁺ 383.2097 found 384.2088.

(2R)-2-(dibenzylamino)-3-((tetrahydro-2H-pyran-2-yl) oxy) propan-1-ol (4): In an rb flask LiAlH₄ was taken (1.1 equiv, 287 mmol, 10.9g) and dry THF (100 mL) was added. The reaction mixture was then cooled in an ice bath and was stirred for 10 min. Compound 3-II dissolved in dry THF in a separate rb flask was transferred to an addition funnel. It was then added drop wise to the cooled solution of LAH at 0 °C and continued to stir for 1 h under an inert atmosphere. Upon completion of the reaction as monitored by TLC (EtOAc/Hexane: 20/80; Rf: 0.3) the reaction was quenched by addition of saturated Na₂SO₄ drop wise at 0 °C until the mixture solidified. Then EtOAc was added and the mixture was stirred at room temperature overnight. The organic layer was then separated, washed with brine and then dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure and then the crude mixture was column purified to give the titled compound 4 as a clear
viscous liquid (84g, 94%). $^1$H NMR (300 MHz, CDCl$_3$): δ 1.5-1.98 (m, 7H), 3.1 (s, 1H), 3.18 (m, 1H), 3.41-3.9 (m, 10H), 4.64 (m, 1H), 7.2-7.42 (m, 10H); $^{13}$C NMR (75 MHz, CDCl$_3$): δ 19.3, 19.7, 25.4, 30.6, 54.1, 58.1, 58.2, 59.6, 59.7, 62.2, 64.8, 65.1, 98.9, 99.3, 127.1, 128.4, 128.9, 129.0, 139.4; HRMS (ESI TOF) M calcd for C$_{22}$H$_{29}$NO$_3$ [M+H]$^+$ 355.2147 found 356.2151.

(2S)-2-(dibenzylamino)-3-((tetrahydro-2H-pyran-2-yl) oxy) propyl methanesulfonate (4-I): Compound 4 (100g, 293 mmol, 1 equiv) was dissolved in dry DCM (350 mL) and was stirred in an ice bath for 10 min under an inert atmosphere. To this mixture Et$_3$N (1.2 equiv, 352 mmol, 49 mL) was added dropwise via an addition funnel and was continued to stirring for 10 min. Then MsCl (1.2 equiv, 352 mmol, 27 mL) was added to the reaction mixture at 0 °C and reaction was continued for 5h. TLC was checked (EtOAc/Hexane: 20/80; R$_f$: 0.7). After completion, the reaction was then quenched by H$_2$O and the organic phase was washed with saturated NaHCO$_3$ (2x100 mL), brine (100 mL) and then dried over anhydrous Na$_2$SO$_4$. After the removal of the solvent under reduced pressure the crude product 4-I (120g, 98%) was obtained.

(2R)-1-azido-N, N-dibenzyl-3-((tetrahydro-2H-pyran-2-yl) oxy) propan-2-amine (4-II): To the stirred solution of 4-I (100g, 231 mmol, 1 equiv) in dry DMF (500 mL), NaN$_3$ (2 equiv, 462 mmol, 30 g) was added. The reaction mixture was then heated at 85 °C for 6h. The progress of the reaction was
monitored by TLC (EtOAc/Hexane: 20/80; R_f: 0.65). After the reaction was completed, DMF was removed under reduced pressure; EtOAc (500mL) was added to it and then extracted with sat. NaHCO_3 (3x300 mL), brine (300mL). The organic layer after drying over anhydrous Na_2SO_4 was removed under reduced pressure and then the crude was column purified to afford the pure product 4-II (86 g, 98%) as a colorless viscous liquid. \( ^1H \) NMR (300 MHz, CDCl_3): \( \delta \) 1.3-1.9 (m, 7H), 3.15 (m, 1H), 3.32-4.07 (m,10H), 4.64 (m,1H),7.2-7.51 (m, 10H); \( ^{13}C \) NMR (75 MHz, CDCl_3): \( \delta \) 19.2, 19.3, 25.4, 30.3, 50.5, 50.6, 54.3, 56.9, 59.6, 59.1, 62.1, 66.3, 66.6, 99.0, 127.01, 128.3, 128.6, 128.7, 128.9, 138.9; HRMS (ESI TOF) M calcd for C_{22}H_{28}N_{4}O_{2} [M+H]^+ 380.2212 found 381.2314.

(R)-3-azido-2-(dibenzylamino)propan-1-ol (5): Compound 4II (100g, 263 mmol, 1 equiv) and PTSA (0.2 equiv, 52.6 mmol, 10g) dissolved in MeOH (500 mL) in a 2L-rb flask was heated at 80 °C overnight. TLC was checked, (EtOAc/Hexane: 20/80; R_f: 0.3). Then MeOH was removed under reduced pressure and the crude oil product was dissolved in EtOAc (500 mL). The organic layer was washed with 10% citric acid (3x 250 mL), sat. NaHCO_3 (3x 300 mL), and then brine (300 mL). After drying the organic phase over Na_2SO_4, the solvent was removed and a column was run using EtOAc/Hexane (20/80) as an eluent to afford compound 5 (75g, %) as a thick liquid. \( ^1H \) NMR (300 MHz, CDCl_3): \( \delta \) 3.08 (m, 1H), 3.35 (m, 1H), 3.54-3.74 (m,6H), 3.9 (d, J=6.8 Hz, 1H),7.2-7.5 (m, 12H); \( ^{13}C \) NMR (75 MHz, CDCl_3): \( \delta \) 49.0, 56.7,
(R)-1-azido-N,N-dibenzyl-3-(2-(2-(tert-butoxy)ethoxy)ethoxy)propan-2-amine (5-I): To the cooled stirring solution of compound 5 (50g, 168.9 mmol, 1 equiv) in dry THF (300 mL) was added NaH (2.5 equiv, 422.2 mmol, 16.8g) in portions and continued to stir for 30 mins at 0 °C under an inert atmosphere. Then, the reaction mixture was heated at 50 °C for 2h. After that, the reaction mixture was cooled to room temperature. At this point, side-chain compound 12 (1.2 equiv, 202.7 mmol, 64 g) dissolved in dry THF (100 mL) was added to the rb flask and heated at 50 °C for 2h. TLC was checked, (EtOAc/Hexane: 20/80; Rf: 0.5). The solvent was removed under reduced pressure and the crude mixture was re-dissolved in EtOAc (200 mL) and washed successively with sat. NaHCO₃ (3x 100 mL), and then brine (100 mL). After drying the organic phase over anhydrous Na₂SO₄, the solvent was removed under reduced pressure and then column purified to yield compound 5-I as a viscous liquid. ¹H NMR (300 MHz, CDCl₃): δ 1.12(s, 9H), 3.2 (brs, 2H), 3.4-3.82 (m,15H), 7.17-7.41 (m, 11H);¹³C NMR (75 MHz, CDCl₃): δ 27.5, 50.4, 54.8, 56.9, 59.1, 61.1, 69.4, 70.6, 70.7, 73.0, 126.9, 128.2, 128.6, 128.9, 139.8; HRMS (ESI TOF) M calcd for C₂₅H₃₆N₂O₃ [M+H]⁺440.2787 found 441.2785.
(R)-N^2, N^2-dibenzyl-3-(2-(2-(tert-butoxy) ethoxy) ethoxy) propane-1, 2-diamine (5-II): Compound 5-I (20g, 45 mmol, 1 equiv) was dissolved in 160 mL of EtOH: H_2O (3:1). To this stirring solution Zn (2.5 equiv, 113.6 mmol, 7.4 g) and NH_4Cl (3 equiv, 135 mmol, 7.2 g) were added at 0 °C. The reaction mixture was allowed to warm to room temperature. The progress of the reaction was monitored by TLC, (MeOH/ DCM: 10/90; R_f: 0.3). After the completion of the reaction, EtOH was removed under reduced pressure and the crude mixture was then extracted with EtOAc (2x 100 mL). The combined organic layers were washed with brine (50 mL) and then dried over anhydrous Na_2SO_4. After removal of the solvent under reduced pressure column chromatography was done to get the pure product 5-II as a viscous liquid. ^1H NMR (300 MHz, CDCl_3): δ 1.2(s, 9H), 2.9 (m, 2H), 3.42-3.9 (m,18H), 7.18-7.4 (m, 11H); ^13C NMR (75 MHz, CDCl_3): δ 27.5, 49.1, 54.2, 58.4, 61.1, 70.2, 70.3, 70.5, 71.1, 71.1, 73.5, 127.3, 128.5, 129.0, 129.1, 129.4, 130.1, 138.0, 139.0; HRMS (ESI TOF) M calcd for C_{25}H_{38}N_2O_3 [M+H]^+ 414.2882 found 415.2834.

Methyl (R)-(3-(2-(2-(tert-butoxy) ethoxy) ethoxy)-2-(dibenzylamino)propyl)glycinate (6): To the solution of 5-II (25g, 60 mmol, 1 equiv) in dry DCM (200 mL) was added Et_3N (1 equiv, 60 mmol, 8.0 mL) drop-wise at 0 °C and continued to stir for 30 min under an inert atmosphere. Then, methyl bromoacetate (1.1 equiv, 66 mmol, 6.3 mL) was added drop-wise at 0 °C very slowly and the reaction mixture was allowed to
warm to room temperature. After 3h, TLC was checked, (MeOH/DCM: 5/95; 
Rf: 0.4). The reaction was then quenched by the addition of water and the 
organic layer was extracted with 10% citric acid (100 mL), brine (100 mL) 
and dried over anhydrous Na₂SO₄. The solvent was removed under reduced 
pressure and then purified by column chromatography to give the backbone 6 
as a yellow viscous liquid. ¹H NMR (300 MHz, CDCl₃): δ 1.14(s, 9H), 3.18 
(m, 2H), 3.37-3.74 (m,15H), 3.78 (s, 3H), 3.96 (q, 4H), 7.2-7.5 (m, 11H);¹³C 
NMR (75 MHz, CDCl₃): δ 14.1, 27.4, 46.2, 52.9, 55.7, 59.0, 60.3, 61.0, 68.7, 
68.9, 70.0, 70.5, 71.1, 72.5, 73.1, 99.8, 117.5, 127.8, 127.9, 128.5, 128.6, 
128.9, 129.3, 129.4, 134.9, 168.7, 171.1; HRMS (ESI TOF) M calcd for 
C₂₈H₄₂N₂O₅ [M+H]⁺486.3094 found 487.3121.

**Methyl (R)-11-(dibenzylamino)-2,2-dimethyl-13-(2-(5-methyl-2,4-dioxo- 
3,4-dihydropyrimidin-1(2H)-yl)acetyl)-3,6,9-trioxa-13-azapentadecan-15-
oate (7a):** To the stirred solution of Thymine acetic acid (1.2 equiv, 1.68 g, 
9.13 mmol) in dry DMF (15 mL), DCC (1.2 equiv, 9.13 mmol, 1.88 g) and 
DhbtOH (1.2 equiv, 9.13 mmol, 1.49 g) were added at room temperature 
under an inert atmosphere. The resulting mixture was stirred at room 
temperature for 1h. Then the backbone 6 (3.7 g, 7.6 mmol, 1equiv) dissolved 
in dry DMF (15 mL) was added to the above reaction mixture and the reaction 
was heated at 50 °C overnight. TLC was checked, (EtOAc/EtOH: 20/1; Rf: 
0.6). DMF was removed under reduced pressure and the solid residue was 
dissolved in EtOAc (50 mL). The organic layer was washed with saturated
NaHCO₃ solution, brine (25 mL) and then dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure and the crude was purified by column chromatography to yield the desired product as a crusty solid material. 

¹H NMR (300 MHz, CDCl₃): δ 1.13 (s, 9H), 1.2-1.35 (m, 1H), 1.57 (s, 3H), 2.9-3.2 (m, 1H), 3.23-3.62 (m, 8H), 3.62-3.64 (m, 5H), 3.64-4.1 (m, 7H), 4.1-4.7 (m, 4H), 6.7-6.97 (m, 1H), 7.15-7.5 (m, 13H); ¹³C NMR (75 MHz, CDCl₃): δ 12.3, 27.4, 47.8, 48.1, 52.1, 52.6, 53.8, 54.2, 54.9, 56.4, 61.1, 64.1, 67.5, 69.8, 70.6, 71.2, 110.1, 111.1, 127.2, 128.2, 128.5, 128.7, 128.9, 129.0, 139.5, 141.2, 150.6, 164, 167.4, 169.3; HRMS (ESI TOF) M calcd for C₃₅H₄₈N₄O₈[M+H]⁺ 652.3472 found 653.3387.

Methyl (R)-13-(2-(2-((tert-butoxycarbonyl) amino)-6-oxo-1, 6-dihydro-9H-purin-9-yl) acetyl)-11-(dibenzylamino)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oate (7b): Guanine ester 7b was prepared, purified, and characterized the same way as described for 7a. ¹H NMR (300 MHz, CDCl₃) major rotamer: δ 1.12 (s, 9H), 1.51 (s, 9H), 3.12 (m, 1H), 3.21-3.62 (m, 8H), 3.71 (s, 3H), 3.72-4.2 (m, 8H), 4.5-4.7 (m, 1H), 4.87-4.9 (m, 1H), 7.1-7.4 (m, 12H), 7.52 (s, 1H), 7.74 (m, 1H); ¹³C NMR (75 MHz, CDCl₃): δ 27.4, 38.6, 47.2, 48.5, 48.9, 49.6, 50.0, 52.3, 52.8, 61.0, 66.7, 66.8, 68.9, 70.2, 70.7, 71.0, 119.9, 127.2, 128.2, 128.5, 128.6, 128.67, 128.9, 129, 139.5, 141.2, 156.3, 163.0, 167.6, 168.5, 169.6; HRMS (ESI TOF) M calcd for C₄₀H₅₅N₇O₉[M+H]⁺ 777.4061 found 778.4103.
Methyl \((R)-13-(2-(6-((\text{tert-butoxycarbonyl})\text{ amino})-9H\text{-purin-9-yl})\text{ acetyl})-11-(\text{dibenzylamino})-2, 2\text{-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oate}\) (7c): Adenine ester 7c was prepared, purified, and characterized the same way as described for 7a. $^1$H NMR (300 MHz, CDCl$_3$) major rotamer: $\delta$ 1.13 (s, 9H), 1.49 (s, 9H), 3.1 (s, 1H), 3.1-4.12 (m, 20H), 4.13 (m,1H), 4.5-4.92 (m, 2H), 7.1-7.4 (m, 12H), 7.56 (s, 1H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 27.4, 47.2, 48.4, 49.8, 50.3, 52.1, 52.6, 66.0, 70.04, 70.1, 70.3, 70.4,71.7, 110.1, 11.1, 127.2,128.2, 12 168.4, 170.7,171.1; HRMS (ESI TOF) M calcd for C$_{40}$H$_{55}$N$_7$O$_8$[M+H]$^+$ 761.4112 found 762.4133.

Methyl \((R)-11-(((9H\text{-fluoren-9-yl})\text{methoxy})\text{carbonyl})\text{amino})-2,2\text{dimethyl-13-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl})\text{acetyl})-3,6,9- trioxa-13-azapentadecan-15-oate}\) (8a): Compound 7a (1g) taken in a Parr hydrogenator reaction vessel was dissolved in dry MeOH (10 mL). To it was added Pd(OH)$_2$/C catalyst (800 mg). The hydrogenation reaction was performed at 50 psi in a Parr hydrogenator at room temperature overnight. TLC was checked, (EtOAc/EtOH: 80/20; $R_f$: 0.2). The reaction was filtered off. The residue was washed with methanol. The combined filtrate was removed under reduced pressure and the crude product obtained was used directly in the next reaction step. The free amine 7-Ia (500 mg, 1.05 mmol, 1 equiv) was dissolved in dry DCM (10 mL) and stirred under an inert atmosphere at 0 °C. To it was added DIPEA (1.2 equiv, 1.26 mmol, 0.25 mL) drop-wise and continued to stir for 15 min. Fmoc-OSU (1 equiv, 1.05 mmol,
0.35 g) was then added to the reaction mixture and continued to stir overnight at room temperature. TLC was checked (EtOAc/EtOH: 20/1; Rf: 0.3). After the completion of the reaction, the organic layer was washed with 10% citric acid (10 mL), saturated NaHCO₃ (10 mL) and brine (10 mL). The DCM layer was dried over anhydrous Na₂SO₄, removed under reduced pressure to give the crude product, which on column purification afforded compound 8a as a white, foamy, crusty material.¹H NMR (300 MHz, DMSO-d₆) major rotamer: δ 1.11 (s, 9H), 1.26 (m, 2H), 1.73 (s, 3H), 3.4-3.67 (m,10H), 3.69 (s, 3H), 3.71 (m, 2H), 4.0-4.52 (m,6H), 4.6-5.2(m, 2H), 7.32 (t, J=6.2 Hz, 2H), 7.43 (t, J=6.5 Hz, 2H), 7.7(d, J=7.01 Hz, 2H), 7.84 (t, J=7.1 Hz, 1H), 7.89 (d, J=6.8 Hz, 2H); ¹³C NMR (75 MHz, DMSO-d₆): δ 12.3, 28.6, 38.7, 48.1, 48.3, 49.4, 52.3, 52.8, 69.6, 70.1, 70.2, 70.3, 71.0, 72.5, 78.6, 108.5, 108.6, 120.5, 120.6, 127.9, 128.6,141.2,142.4, 151.4, 155.8, 164.8, 167.8, 168.4, 170.7, 171.7; HRMS (ESI TOF) M calcd for C₃₆H₄₆N₄O₁₀[M+H]⁺ 694.3 found 695.3.

Methyl (R)-11-((((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(2-((tert-butoxycarbonyl) amino)-6-oxo-1, 6-dihydro-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxo-13-azapentadecan-15-oate (8b): Guanine ester 8b was prepared, purified, and characterised the same way as described for 8a. ¹H NMR (300 MHz, DMSO-d₆) major rotamer: δ 1.1 (s, 9H), 1.41 (s,9H), 3.2-3.81 (m, 14H), 3.9 -4.5(m, 5H), 4.9-5.3 (m, 2H), 7.23 (t, J=7.1 Hz, 1H), 7.3(m,2H), 7.41(m,2H), 7.5-7.7 (m,2H), 7.8 (d, J=7.7 Hz, 2H), 7.86-7.92 (m,2H), 10.7 (s,1H), 11.3(s,1H); ¹³C NMR (75 MHz, CDCl₃): δ 27.4, 27.9, 43.6, 47.2 48.7, 49.0, 50.2, 52.3, 61.0, 66.7, 69.6, 70.2, 70.7, 71.1, 73.2, 83.8,
120.0, 124.8, 124.9, 127.0, 127.1, 127.8, 139.6, 141.3, 141.3, 143.5, 143.6, 147.6, 149.0, 155.7, 156.3, 167.1, 169.5; HRMS (ESI TOF) M calcd for C_{41}H_{53}N_{7}O_{11}[M+H]^+ 819.3 found 820.3.

Methyl (R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(6-((tert-butoxycarbonyl) amino)-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oate (8c): Adenine ester 8c was prepared, purified, and characterised the same way as described for 8a. $^1$H NMR (300 MHz, CDCl$_3$) major rotamer: $\delta$ 1.12 (s, 9H), 1.38 (s,9H), 3.3-3.72 (m, 10H), 3.81 (s, 3H), 3.94 (m, 2H), 4.1-4.42 (m,6H), 5-5.5 (m, 2H), 6.28 (m, 2H), 7.35 (q, $J=6.8$ Hz, 2H), 7.5 (m, 2H), 7.72 (m, 2H), 8.1 (m, 1H), 8.3-8.7 (m,1H); $^{13}$C NMR (75 MHz, DMSO-$d_6$): $\delta$ 27.4, 47.2 48.2, 48.4, 49.6, 50.3, 61.0, 66.8, 70.0, 70.1, 70.5, 71.5, 71.7, 123.4, 123.4, 125.6, 127.5, 128.1, 141.3, 141,156.3, 168.5; HRMS (ESI TOF) M calcd for C_{41}H_{53}N_{7}O_{10}[M+H]^+ 803.3 found 804.3.

(R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-2, 2-dimethyl-13-(2-(5-methyl-2, 4-dioxo-3, 4-dihydropyrimidin-1(2H)-yl) acetyl)-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (1a): To the stirred solution of 8a (0.5 g, 0.72 mmol, 1equiv) in THF (5 mL) 2N NaOH (10 equiv, 7.2 mmol, 0.3g) was added drop-wise at 0 °C. After 10 min, the reaction was neutralized with 1N HCl (12 equiv), resulting in the formation of a cloudy white solution. It was extracted with EtOAc (3x 10 mL) and washed with brine (5 mL). The organic layer was dried over anhydrous Na$_2$SO$_4$, filtered and removed under
reduced pressure. To the crude oily material a small amount of DCM was added, swirled to mix, and then crashed with diethyl ether to get white sticky solid. The solvent was decanted, and the crude was column purified to yield the pure product as a white crystalline material. $^1$H NMR (300 MHz, DMSO-$d_6$): $\delta$ 1.12 (s, 9H), 1.72 (s, 3H), 3.21-3.6 (m, 12H), 3.7-4.1 (m, 3H), 4.11-4.4 (m, 2H), 4.45 (s, 1H), 4.6 (s, 1H), 7.25 (s, 1H), 7.33 (m, 2H), 7.42 (t, $J$=7.1 Hz, 2H), 7.7 (m, 2H), 7.9 (d, $J$=6.9 Hz, 2H), 11.3 (s, 1H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 12.3, 27.9, 47.2, 48.2, 48.4, 49.8, 50.3, 65.8, 66.0, 70.0, 70.1, 70.3, 70.4, 71.7, 78.0, 108.4, 120.5, 125.6, 127.5, 128.1, 141.2, 144.3, 151.4, 156.3, 164.8, 165, 168, 170.7; HRMS (ESI TOF) M calcd for C$_{35}$H$_{44}$N$_4$O$_{10}$[M+H]$^+$ 680.3 found 681.3.

(R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(2-((tert-butoxycarbonyl) amino)-6-oxo-1, 6-dihydro-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (1b): Guanine monomer 1b was prepared, purified, and characterised the same way as described for 1a. $^1$H NMR (300 MHz, DMSO-$d_6$): $\delta$ 1.13 (s, 9H), 1.41 (s, 9H), 2.9-3.2 (m, 2H), 3.2-3.8 (m, 11H), 3.91-4.41 (m, 6H), 4.9 (s, 1H), 5.1 (m, 1H), 7.21-7.45 (m, 4H), 7.57-7.73 (m, 2H), 7.78 (m, 1H), 7.86 (d, $J$=6.8 Hz, 2H), 11.33 (s, 1H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 27.4, 28.1, 44.2, 47.2, 48.5, 48.8, 49.9, 50.9, 66.0, 70.0, 70.1, 70.5, 71.6, 71.7, 82.9, 83.0, 119.2, 120.5, 120.6, 125.5, 125.6, 127.5, 127.9, 128.0, 128.1, 140.6, 140.8, 141.1, 141.2, 144.3, 147.9, 149.9, 154.0, 155.4, 156.6, 167.4, 170.7; HRMS (ESI TOF) M calcd for C$_{40}$H$_{51}$N$_7$O$_{11}$[M+H]$^+$ 805.3 found 806.3.
(R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(6-((tert-butoxycarbonyl) amino)-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (1c): Adenine monomer 1c was prepared, purified, and characterised the same way as described for 1a.  

$^1$H NMR (300 MHz, DMSO-$d_6$): $\delta$ 1.11 (s, 9H), 1.5 (s,9H), 3.38-3.7 (m, 11H), 3.9-4.4 (m, 6H), 7.32 (m, 2H), 7.41 (m,2H), 7.7 (m,2H), 7.88 (d, $J$=7.1 Hz, 2H), 8.25 (s, 1H), 8.5 (s, 1H), 10.612 (s,1H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 27.7, 28.6, 47.2, 48.4, 48.5, 48.9, 49.8, 50.3, 66.0, 70.0, 70.1, 70.12, 70.4, 71.0, 71.6, 71.7, 80.4, 83.8, 120.5, 123.4, 123.5, 125.6, 127.5, 128.1, 141.2, 144.3, 144.4, 145.3, 150.1, 151.6, 151.8, 152.7, 156.2, 156.4, 167.4, 170.6; HRMS (ESI TOF) M calcd for C$_{40}$H$_{51}$N$_7$O$_{10}$[M+H]$^+$ 789.3 found 790.3.

Methyl (R)-(2-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-3-(2-(tert-butoxy ethoxy) ethoxy) propyl) glycinate (9): Compound 6 (2g, 4.1 mmol, 1 equiv) taken in a Parr hydrogenator reaction vessel was dissolved in dry MeOH (25 mL). To it was added Pd(OH)$_2$/C catalyst (2g). The hydrogenation reaction was performed at 50 psi in a Parr hydrogenator at room temperature overnight. TLC was checked, (EtOAc/EtOH: 80/20; R$_f$: 0.1). The reaction was filtered off. The residue was washed with methanol. The combined filtrate was removed under reduced pressure and the crude product obtained was used directly in the next reaction step. The resulting free amine (1g, 3.2 mmol, 1 equiv) and Fmoc-OSU (1.1 equiv, 3.6 mmol, 1.2 g) were taken in a rb and dissolved in dry THF (10 mL). The suspension was stirred at 0 °C. After 5 min DIPEA (2.5 equiv, 8.15 mmol, 1.4 mL) dissolved
in dry THF (5 mL) was added drop-wise over 15 min at 0 °C. The reaction was continued to stir for 1h at 0 °C and then allowed to warm to room temperature. TLC was checked, (EtOAc/Hexane: 80/20; Rf: 0.3). The reaction was quenched by H2O. The compound was extracted with EtOAc (20 mL) and washed with saturated NaHCO3 (20 mL), and brine (20 mL), dried over anhydrous Na2SO4, and concentrated in vacuo. The crude mixture was column purified to get the desired pure product as a colorless viscous liquid. 1H NMR (300 MHz, CDCl3): δ 1.11 (s, 9H), 3.36-3.62 (m, 15H), 4.1-4.4 (m, 3H), 7.12 (d, J=8.07 Hz, 1H), 7.34(t, J=6.9Hz,2H), 7.41(t, J=7.1 Hz, 2H), 7.71(d, J=7.2 Hz, 2H), 7.88(d, J=6.8 Hz 2H); HRMS (ESI TOF) M calcd for C29H40N2O7[M+H]+ 527.2 found 528.2.

Methyl (R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(4-((tert-butoxycarbonyl) amino)-2-oxopyrimidin-1(2H)-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oate (9-I): Cytosine ester 9-I was prepared, purified, and characterised the same way as described for 7a. 1H NMR (300 MHz, DMSO-d6) major rotamer: δ 1.1 (s, 9H), 1.41 (s,9H), 3.3-3.6 (m, 11H), 3.63 (s,3H), 3.71 (m,1H), 3.8-4.3 (m, 6H), 4.54-4.86 (m,2H), 6.9(m, 1H), 7.25 (t, J=7.3 Hz, 1H), 7.32 (m, 2H), 7.42 (m, 2H), 7.54-7.7(m, 2H), 7.79 (d, J=7.1 Hz, 1H), 7.89 (m, 2H), 10.28 (s,1H); 13C NMR (75 MHz, CDCl3): δ 27.4, 28.0, 38.6, 47.1, 48.5, 48.9, 49.6, 50.0, 52.3, 52.8, 61.0, 66.7, 66.8, 68.8, 70.2, 70.6, 70.6, 71.0,73.2, 119.9, 125.1, 126.5, 127.1,127.7, 141.3, 141.3, 143.7, 143.9, 149.7, 156.3,163.0, 167.6, 168.5, 169.6; HRMS (ESI TOF) M calcd for C40H48N2O11[M+H]+ 779.3 found 780.3.
(R)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(4-((tert-butoxycarbonyl) amino)-2-oxopyrimidin-1(2H)-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (1d): Cytosine monomer 1d was prepared, purified, and characterised the same way as described for 1a. 

\[1H\text{ NMR (300 MHz, DMSO-}d_6\text{): } \delta 1.1 \text{ (s, 9H), 1.4 (s,9H), 3.21-3.6 (m, 11H), 3.7-4.4 (m,7H), 4.61-4.86 (m,2H), 6.93 (m, 1H), 7.32 (t, } J=7.1 \text{ Hz, 2H), 7.41(t, } J=6.9 \text{ Hz, 2H), 7.7 (d, } J=6.5 \text{ Hz, 2H), 7.82 (m,1H), 7.87 (d, } J=6.8 \text{ Hz, 2H);}

\[13C\text{ NMR (75 MHz, CDCl}_3\text{): } \delta 27.4, 28.2, 47.2 48.4, 49.9, 50.4, 65.8, 66.0, 70.0, 70.1, 70.3, 70.4, 71.6, 71.7, 81.3, 94.3, 120.5, 125.7, 127.5, 128.0, 141.2, 144.3, 150.8, 150.9, 152.6, 155.3, 156.3, 167.9, 170.7, 171.1; \text{ HRMS (ESI TOF) } M \text{ calcd for C}_{39}H_{51}N_5O_{11}[M+H]^+ 765.3 found 766.3.}

(2R)-N, N-dibenzyl-1-(2-(2-(tert-butoxy) ethoxy) ethoxy)-3-((tetrahydro-2H-pyran-2-yl) oxy) propan-2-amine (4'): To the solution of compound 4 (50 g, 140.8 mmol, 1 equiv) dissolved in dry THF (250 mL) at 0 °C, NaH (2.5 equiv, 352 mmol, 14 g) was added in portion-wise. The reaction was stirred at 0 °C for 10 min and then heated at 50 °C for 2h. After that the reaction mixture was cooled to room temperature. To it side-chain 17 (1.2 equiv, 168.9 mmol, 53.3 g) dissolved in dry THF (50 mL) was added drop-wise and after stirring at room temperature for 5 min was heated at 50 °C for 12h. TLC was checked, (EtOAc/Hexane: 10/90; } Rf: 0.3). The solvent was removed under reduced pressure and the crude mixture was re-dissolved in EtOAc (100 mL) and washed successively with saturated solution of NaHCO$_3$ (2x 50 mL), and then brine (50 mL). After drying the organic phase over anhydrous Na$_2$SO$_4$. 

157
the solvent was removed under reduced pressure and then column purified to yield above titled compound as a viscous liquid. $^1$H NMR (300 MHz, CDCl$_3$): δ 1.2 (s, 9H), 1.4-1.9 (m, 7H), 3.21 (m, 1H), 3.43-3.71 (m, 5H), 3.8-4.16 (m, 9H), 4.2-4.4 (m, 2H), 4.6 (m, 2H), 5.0 (m, 2H), 7.13-7.51 (m, 11H); $^{13}$C NMR (75 MHz, CDCl$_3$): δ 19.7, 21.3, 25.4, 27.5, 30.6, 54.7, 55.7, 55.8, 61.7, 61.9, 62.8, 65.4, 65.7, 67.2, 67.8, 98.7, 98.9, 126.8, 128.1, 128.5, 128.6, 128.9, 140.1; HRMS (ESI TOF) M calcd for C$_{30}$H$_{45}$NO$_5$[M+H]$^+$ 499.3298 found 500.3289.

**(S)-3-(2-(2-(tert-butoxy) ethoxy) ethoxy)-2-(dibenzylamino) propan-1-ol (10):** To the stirred solution of 4’ (41g, 82.1 mmol, 1 equiv) in Methanol (250 mL), pTSA (0.5 equiv, 410.8 mmol, 7.8 g) was added and the reaction mixture was heated at 80 °C overnight. TLC was checked, (EtOAc/Hexane: 20/80; R$_f$: 0.1). Then MeOH was removed under reduced pressure and the crude oil product was dissolved in EtOAc (200 mL). The organic layer was washed with 10% citric acid (2x 100 mL), saturated NaHCO$_3$ (100 mL), and then brine (50 mL). After drying the organic phase over Na$_2$SO$_4$, the solvent was removed and a column was run using EtOAc/Hexane (20/80) as an eluent to afford compound 10 as a viscous liquid. $^1$H NMR (300 MHz, CDCl$_3$): δ 1.2 (s, 9H), 2.73 (s, 2H), 3.12 (m, 1H), 3.52-3.71 (m, 13H), 3.75-3.82 (m, 1H), 3.87 (d, J=7.1 Hz, 2H), 7.1-7.5 (m, 11H); $^{13}$C NMR (75 MHz, CDCl$_3$): δ 27.5, 54.2, 58.1, 59.8, 61.2, 69.2, 70.6, 70.7, 71.3, 76.8, 77.0, 77.3, 126.4, 127.1, 128.4, 128.9, 139.6; HRMS (ESI TOF) M calcd for C$_{25}$H$_{37}$NO$_4$[M+H]$^+$ 415.2723 found 416.2814.
(S)-1-azido-N, N-dibenzyl-3-(2-(2-(tert-butoxy) ethoxy) ethoxy) propan-2-amine (11): Compound 10 (50g, 120 mmol, 1 equiv) was dissolved in dry DCM (250 mL) and was stirred in an ice bath for 10 min under an inert atmosphere. To this mixture Et₃N (1.2 equiv, 144.5 mmol, 20 mL) was added drop-wise via an addition funnel and was continued to stir for 10 min. Then MsCl (1.2 equiv, 144.5 mmol, 11.5 mL) was added to the reaction mixture at 0 °C and reaction was continued for 3h. TLC was checked (EtOAc/Hexane: 40/60; Rf: 0.7). After completion, the reaction was then quenched by H₂O and the organic phase was washed with saturated NaHCO₃ (100 mL), brine (100 mL) and then dried over anhydrous Na₂SO₄. After the removal of the solvent under reduced pressure the crude product 10-I was obtained.

To the stirred solution of 10-I (50g, 100 mmol, 1 equiv) in dry DMF (250 mL), NaN₃ (2 equiv, 200 mmol, 13 g) was added. The reaction mixture was then heated at 80 °C for 5h. The progress of the reaction was monitored by TLC (EtOAc/Hexane: 40/60; Rf: 0.6). After the reaction was completed, DMF was removed under reduced pressure; EtOAc (250mL) was added to it and then extracted with saturated NaHCO₃ (2x 50 mL), brine (100mL). The organic layer after drying over anhydrous Na₂SO₄ was removed under reduced pressure and then the crude was column purified to afford the pure product 11 as a colourless viscous liquid. ¹H NMR (300 MHz, CDCl₃): δ 1.2 (s, 9H), 2.78-3.27 (m, 3H), 3.4-3.9 (m, 16H), 7.1-7.5 (m, 12H); ¹³C NMR (75 MHz, CDCl₃): δ 27.5, 50.4, 54.8, 56.9, 59.2, 61.1, 61.2, 70.5, 70.6, 70.7, 71.3, 72.4,
126.9, 127.1, 127.9, 128.2, 128.3, 128.6, 128.9, 138.9, 139.8; HRMS (ESI TOF) M calcd for C_{25}H_{36}N_{4}O_{3}[M+H]^+ 440.2787 found 441.2831.

(S)-N^2, N^2-dibenzyl-3-(2-(2-(tert-butoxy) ethoxy) ethoxy) propane-1, 2-diamine (11-I): The titled compound was synthesized from SM 11 by following the same protocol as described for 5-II. The crude product was not purified but taken to the next step.

Methyl (S)-(3-(2-(2-(tert-butoxy) ethoxy) ethoxy)-2-(dibenzylamino) propyl) glycinate (12): Compound 12 was synthesized in the same way as in the case of compound 6. ^1H NMR (300 MHz, CDCl3): δ 1.21(s, 9H), 3.11 (m, 2H), 3.3-3.9 (m,16H), 3.9-4.3 (m, 3H), 7.1-7.51 (m, 11H); ^13C NMR (75 MHz, CDCl3): δ 14.1, 27.4, 46.2, 52.9, 55.7, 59.0, 60.3, 61.0, 68.7, 68.9, 70.0, 70.5, 71.1, 72.5, 73.1, 100, 115.2, 127.8, 127.9, 128.5, 128.6, 128.8, 128.9, 129.3, 129.4, 134.9, 168.7, 171.1; HRMS (ESI TOF) M calcd for C_{28}H_{42}N_{2}O_{5} [M+H]^+ 486.3 found 487.3.

Methyl (S)-11-(dibenzylamino)-2,2-dimethyl-13-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetyl)-3,6,9-trioxa-13-azapentadecan-15-oate (13a): To the stirred solution of Thymine acetic acid (1.2 equiv, 1.7 g, 9.13 mmol) in dry DMF (15 mL), DCC (1.2 equiv, 9.13 mmol, 1.88 g) and DhbtOH (1.2 equiv, 9.13 mmol, 1.49 g) were added at room temperature under an inert atmosphere. The resulting mixture was stirred at room temperature for 1h. Then the backbone 12 (3.7 g, 7.6 mmol, 1equiv) dissolved in dry DMF (15 mL) was added to the above reaction mixture and the reaction
was heated at 50 °C overnight. TLC was checked, (EtOAc/EtOH: 20/1; Rf: 0.6). DMF was removed under reduced pressure and the solid residue was dissolved in EtOAc (50 mL). The organic layer was washed with saturated NaHCO₃ solution, brine (25 mL) and then dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure and the crude was purified by column chromatography to yield the desired product as a crusty solid material.

1H NMR (300 MHz, CDCl₃): δ 1.11(s, 9H), 1.2-1.35 (m, 1H), 1.52 (s, 3H), 2.9-3.2 (m, 1H), 3.23-3.62 (m, 8H), 3.62-3.64 (m, 5H), 3.64-4.1 (m, 7H), 4.1-4.7 (m, 4H), 6.7-6.97 (m, 1H), 7.15-7.5 (m, 13H); 13C NMR (75 MHz, CDCl₃): δ 27.5, 49.1, 54.2, 58.4, 61.1, 70.2, 70.3, 70.5, 71.1, 71.1, 73.5, 127.3, 128.5, 129.0, 129.1, 129.4, 130.1, 138.0, 139.0; HRMS (ESI TOF) M calcd for C₃₅H₄₈N₄O₈[M+H]⁺652.34 found 653.3.

**Methyl (S)-13-2-(2-((tert-butoxycarbonyl) amino)-6-oxo-1, 6-dihydro-9H-purin-9-yl) acetyl)-11-(dibenzyllamino)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oate (13b):** Guanine ester 13b was prepared, purified, and characterized the same way as described for 13a.

**Methyl (S)-13-2-(6-((tert-butoxycarbonyl) amino)-9H-purin-9-yl) acetyl)-11-(dibenzyllamino)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oate (13c):** Adenine ester 13c was prepared, purified, and characterized the same way as described for 13a.
**trioxa-13-azapentadecan-15-oic acid (2a):** Compound 13a was subjected to hydrogenation followed by the Fmoc protection and then purified, characterized in the same way as described for 8a. The ester was then hydrolyzed, purified and characterized by the exact procedure as described for 1a. 

$^1$H NMR (300 MHz, DMSO-$d_6$): $\delta$ 1.2 (s, 9H), 1.8 (s, 3H), 2.8-3.5 (m, 17H), 3.51-3.92 (m, 3H), 4.2-4.33 (m, 2H), 4.34-4.6 (m, 1H), 7.28 (s, 1H), 7.32 (t, $J$=6.8 Hz, 2H), 7.41 (t, $J$=6.9 Hz, 2H), 7.68 (d, $J$=6.2 Hz, 2H), 7.9 (d, $J$=6.5 Hz, 2H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 12.3, 28.2, 47.2, 48.4, 49.9, 50.4, 65.8, 66.0, 70.0, 70, 70.3, 71.6, 71.7, 81.3, 94.3, 120.5, 125.7, 127.5, 128.1, 141.2, 144.3, 150.1, 152.6, 155.5, 156.3, 163.7, 167.9, 170.7, 171.1; HRMS (ESI TOF) M calcd for C$_{35}$H$_{44}$N$_4$O$_{10}$[M+H]$^+$ 680.3 found 681.3.

**(S)-11-((((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(2-((tert-butoxycarbonyl) amino)-6-oxo-1, 6-dihydro-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (2b):** Guanine monomer 2b was prepared, purified, and characterised the same way as described for 2a. $^1$H NMR (300 MHz, DMSO-$d_6$): $\delta$ 1.16 (s, 9H), 1.45 (s, 9H), 3.0-3.6 (m, 11H), 3.9 (m, 1H), 4.1-4.5 (m, 4H), 5.01 (d, $J$= 6.9 Hz, 1H), 6.8 (s, 1H), 7.27-7.36 (m, 2H), 7.37-7.7.47 (m, 2H), 7.67 (t, $J$=7.1 Hz, 2H), 7.81 (m, 1H), 7.88 (d, $J$= 6.4 Hz, 2H); $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 27.5, 28.1, 47.2 48.5, 48.8, 50.9, 65.9, 66.0, 70.0, 70.3, 71.1, 71.6, 71.7, 83, 120.5, 125.5, 125.6, 127.5, 128, 128.1, 140.8, 141.2, 144.0, 144.3, 147.9, 149.9, 154, 155.4, 166.6, 167.4, 170.7; HRMS (ESI TOF) M calcd for C$_{40}$H$_{51}$N$_7$O$_{11}$[M+H]$^+$ 805.3 found 806.3.
(S)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-((tert-butoxycarbonyl) amino)-9H-purin-9-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (2c): Adenine monomer 2c was prepared, purified, and characterised the same way as described for 2a. \[1^H\] NMR (300 MHz, DMSO-\(d_6\)): \(\delta 1.17\) (s, 9H), 1.56 (s,9H), 3.05-3.7 (m, 13H), 4.1 (m, 1H), 4.1-4.4 (m, 4H), 5.1 (d, \(J= 6.7\) Hz, 2H), 6.82 (s, 1H), 7.37 (m, 2H), 7.53 (m, 2H), 7.7 (t, \(J= 7.1\) Hz, 2H), 7.9 (d, , \(J=7.3\) Hz, 2H), 8.31 (s, 1H), 8.5 (d, \(J=6.1\) Hz, 1H), 10.9 (s,1H); \[13^C\] NMR (75 MHz, CDCl\(_3\)): \(\delta 27.4, 28.2, 47.2, 48.4, 49.8, 50.3, 70.02, 70.04, 70.07, 70.1, 70.3, 70.4, 71.6, 71.7, 80.5, 120.5, 123.4, 125.6, 127.5, 128.1, 141.2, 144.3, 145.3, 151.6, 151.8, 152.7, 154.3, 156.3, 163.7, 167.9, 170.7, 171.1; HRMS (ESI TOF) M calcd for C\(_{40}\)H\(_{51}\)N\(_7\)O\(_{10}\)[M+H]\(^+\) 789.3 found 790.3.

(S)-11-(((9H-fluoren-9-yl) methoxy) carbonyl) amino)-13-(2-(4-((tert-butoxycarbonyl) amino)-2-oxopyrimidin-1(2H)-yl) acetyl)-2, 2-dimethyl-3, 6, 9-trioxa-13-azapentadecan-15-oic acid (2d): Cytosine monomer 1d was prepared, purified, and characterised the same way as described for 1a. \[1^H\] NMR (300 MHz, DMSO-\(d_6\)): \(\delta 1.12\) (s, 9H), 1.41 (s,9H), 3.01-3.6 (m, 16H), 3.97 (s,1H), 4.11-4.4 (m, 3H), 4.7 (d, \(J=7.1\) Hz, 2H), 6.8 (s, 1H), 6.93 (m, 1H), 7.27-7.45 (m, 4H), 7.68 (m, 2H), 7.83-7.91(m, 3H), 11.0 (s, 1H); \[13^C\] NMR (75 MHz, CDCl\(_3\)): \(\delta 27.5, 28.2, 47.2, 48.4, 49.9, 50.4, 65.8, 66.0, 70.0, 70.05, 70.3, 71.6, 71.7, 81.3, 120.5, 125.7, 127.5, 128.0, 141.2, 144.3, 150.8, 150.9, 152.6, 155.2, 156.3, 163.7, 167.9, 170.7, 171.1; HRMS (ESI TOF) M calcd for C\(_{39}\)H\(_{51}\)N\(_5\)O\(_{11}\)[M+H]\(^+\) 765.3 found 766.3.
2-(2-hydroxyethoxy) ethyl-4-methylbenzenesulfonate (16): To the diethylene glycol (100 mL, 1.05 mol, 1 equiv) dissolved in dry DCM (300 mL) Et₃N (0.5 equiv, 525 mmol, 73 mL) was added dropwise at 0 °C. The reaction mixture was stirred at 0 °C for 30 min. Then tosyl chloride (0.25 equiv, 0.26 mmol, 50 g) dissolved in dry DCM (100 mL) was added drop-wise and then the reaction was allowed to warm to room temperature. After 4h, the TLC was checked, (EtOAc/EtOH: 20/1; Rf: 0.7). The reaction was quenched by the addition of water, and extracted with saturated NaHCO₃ solution. The organic layer was washed with 10% citric acid solution (100 mL), brine (100 mL) and then dried over anhydrous Na₂SO₄. DCM was removed under reduced pressure and the reaction mixture was re-dissolved in Methanol and put in – 20 °C refrigerator overnight. The solid was filtered off and then was washed with cold methanol. The filtrate on concentration yielded the desired product as a colorless viscous liquid. ¹H NMR (300 MHz, CDCl₃): δ 2.51 (s, 3H), 3.12 (t, J=6.2 Hz, 2H), 3.62 (m, 4H), 4.11 (m, 2H), 7.31 (d, J=7.1 Hz, 2H), 7.78 (d, J=7.0 Hz, 2H); ¹³C NMR (75 MHz, CDCl₃): δ 21.6, 60.3, 68.1, 69.3, 71.8, 127.9, 129.9, 132.9, 144.9; HRMS (ESI TOF) M calcd for C₁₁H₁₆O₅S [M+H]+ 260.0718 found 261.0723.

(2-(2-(tert-butoxy) ethoxy) ethyl)-4-methylbenzenesulfonate (17): To the stirring solution of compound 16 (100g, 384.6 mmol, 1 equiv) in dry DCM (300 mL) at 0 °C, was added boc anhydride (2.3 equiv, 884.6 mmol, 192.5 g) and Magnesium perchlorate (0.1 equiv, 38.46 mmol, 8.4 g) and the mixture was allowed to warm to room temperature overnight. After the completion of
the reaction, monitored by TLC (EtOAc/Hexane:50/50;R_t:0.8) the reaction was quenched by water, extracted with saturated NaHCO₃ solution, washed with brine (100 mL), and dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure to get the crude mixture. The crude mixture on column purification yielded the desired product 17 as a colorless viscous liquid. ¹H NMR (300 MHz, CDCl₃): δ 1.12 (s, 9H), 2.4 (s, 3H), 3.42 (m, 2H), 3.5 (m,2H), 3.68 (m, 2H), 4.14 (m, 2H), 7.3 (d, , J=6.9 Hz, 2H), 7.78 (d, , J=7.0 Hz, 2H); ¹³C NMR (75 MHz, CDCl₃): δ 21.5, 27.4, 61.1, 68.6, 69.3, 71.3, 72.9, 127.9, 129.7, 133.0, 144.7; HRMS (ESI TOF) M calcd for C₁₅H₂₄O₅S [M+H]^+ 316.1344 found 317.1329.

IV.7. References


IV.8. Appendices

Figure 1A. $^1$H NMR spectrum of compound 16 (CDCl$_3$, 300MHz)

![Figure 1A](image)

Figure 2A. $^{13}$C NMR spectrum of compound 16 (CDCl$_3$, 75MHz)

![Figure 2A](image)
Figure 3A. $^1$H NMR spectrum of compound 17 (CDCl$_3$, 300MHz)

Figure 4A. $^{13}$C NMR spectrum of compound 17 (CDCl$_3$, 75MHz)
**Figure 5A.** $^1$H NMR spectrum of compound 3-I (CDCl$_3$, 300MHz)

**Figure 6A.** $^{13}$C NMR spectrum of compound 3-I (CDCl$_3$, 75MHz)
Figure 7A. $^1$H NMR spectrum of compound 3-II (CDCl$_3$, 300MHz)

Figure 8A. $^{13}$C NMR spectrum of compound 3-II (CDCl$_3$, 75MHz)
Figure 9A. $^1$H NMR spectrum of compound 4 (CDCl$_3$, 300MHz)

Figure 10A. $^{13}$C NMR spectrum of compound 4 (CDCl$_3$, 75MHz)
Figure 11A. $^1$H NMR spectrum of compound 4-I (CDCl$_3$, 300MHz)

Figure 12A. $^{13}$C NMR spectrum of compound 4-I (CDCl$_3$, 75MHz)
Figure 13A. $^1$H NMR spectrum of compound 5 (CDCl$_3$, 300MHz)

Figure 14A. $^{13}$C NMR spectrum of compound 5 (CDCl$_3$, 75MHz)
Figure 15A. $^1$H NMR spectrum of compound 5-I (CDCl$_3$,300MHz)

Figure 16A. $^{13}$C NMR spectrum of compound 5-I (CDCl$_3$,75MHz)
Figure 17A. $^1$H NMR spectrum of compound 5-II (CDCl$_3$,300MHz)

Figure 18A. $^{13}$C NMR spectrum of compound 5-II (CDCl$_3$,75MHz)
Figure 19A. $^1$H NMR spectrum of compound 6 (CDCl$_3$, 300MHz)

Figure 20A. $^{13}$C NMR spectrum of compound 6 (CDCl$_3$, 75MHz)
Figure 21A. $^1$H NMR spectrum of compound 7a (CDCl$_3$, 300MHz)

Figure 22A. $^{13}$C NMR spectrum of compound 7a (CDCl$_3$, 75MHz)
Figure 23A. $^1$H NMR spectrum of compound 7b (CDCl$_3$, 300MHz)

Figure 24A. $^{13}$C NMR spectrum of compound 7b (CDCl$_3$, 75MHz)
Figure 25A. $^1$H NMR spectrum of compound 7c (CDCl$_3$,300MHz)

Figure 26A. $^1$H NMR spectrum of compound 7c (DMSO-$d_6$,300MHz)
Figure 27A. $^1$H NMR spectrum of compound 8a (CDCl$_3$, 300MHz)

Figure 28A. $^{13}$C NMR spectrum of compound 8a (DMSO-$d_6$, 75MHz)
Figure 29A. $^1$H NMR spectrum of compound 8b (DMSO-$d_6$, 300MHz)

Figure 30A. $^{13}$C NMR spectrum of compound 8b (CDCl$_3$, 75MHz)
**Figure 31A.** $^1$H NMR spectrum of compound 8c (CDCl$_3$,300MHz)

![H NMR spectrum of compound 8c](image)

**Figure 32A.** $^{13}$C NMR spectrum of compound 8c (CDCl$_3$,75MHz)

![13C NMR spectrum of compound 8c](image)
Figure 33A. $^1$H NMR spectrum of compound 9-I (DMSO-$d_6$, 300MHz)

Figure 34A. $^{13}$C NMR spectrum of compound 9-I (CDCl$_3$, 75MHz)
Figure 35A. $^1$H NMR spectrum of compound 1a (DMSO-$d_6$, 300MHz)

Figure 36A. $^{13}$C NMR spectrum of compound 1a (DMSO-$d_6$, 75MHz)
Figure 37A. $^1$H NMR spectrum of compound 1b (DMSO-$d_6$, 300MHz)

Figure 38A. $^{13}$C NMR spectrum of compound 1b (DMSO-$d_6$, 75MHz)
**Figure 39A.** $^1$H NMR spectrum of compound 1c (DMSO-$d_6$,300MHz)

**Figure 40A.** $^{13}$C NMR spectrum of compound 1c (DMSO-$d_6$,75MHz)
**Figure 41A.** $^1$H NMR spectrum of compound 1d (DMSO-$d_6$, 300MHz)

**Figure 42A.** $^{13}$C NMR spectrum of compound 1d (DMSO-$d_6$, 75MHz)
Figure 43A. $^1$H NMR spectrum of compound 4' (CDCl$_3$, 300MHz)

Figure 44A. $^{13}$C NMR spectrum of compound 4’ (CDCl$_3$, 75MHz)
Figure 45A. $^1$H NMR spectrum of compound 10 (CDCl$_3$, 300MHz)

Figure 46A. $^{13}$C NMR spectrum of compound 10 (CDCl$_3$, 75MHz)
Figure 47A. $^1$H NMR spectrum of compound 11 (CDCl$_3$, 300MHz)

Figure 48A. $^{13}$C NMR spectrum of compound 11 (CDCl$_3$, 75MHz)
Figure 49A. $^1$H NMR spectrum of compound 12 (CDCl$_3$, 300MHz)

Figure 50A. $^{13}$C NMR spectrum of compound 12 (CDCl$_3$, 75MHz)
Figure 51A. $^1$H NMR spectrum of compound 2a (DMSO-$d_6$, 300MHz)

Figure 52A. $^{13}$C NMR spectrum of compound 2a (DMSO-$d_6$, 75MHz)
Figure 53A. $^1$H NMR spectrum of compound 2b (DMSO-$d_6$, 300MHz)

Figure 54A. $^{13}$C NMR spectrum of compound 2b (DMSO-$d_6$, 75MHz)
**Figure 55A.** $^1$H NMR spectrum of compound 2c (DMSO-$d_6$, 300MHz)

**Figure 56A.** $^{13}$C NMR spectrum of compound 2c (DMSO-$d_6$, 75MHz)
Figure 57A. $^1$H NMR spectrum of compound 2d (DMSO-$d_6$, 300MHz)

Figure 58A. $^{13}$C NMR spectrum of compound 2d (DMSO-$d_6$, 75MHz)
APPLICATIONS

The optically pure MP-γPNA monomers were prepared as discussed in the previous chapters and the oligomers were synthesized for various applications. Herein, we discuss in brief two such applications.

1. High Affinity Sandwich Hybridization assay for Rapid Detection of Short Nucleic Acid Targets

   To date several enzymatic methods have been developed for detection of nucleic acid analytes. In order to improve probe-target identification over the background noise, sandwich hybridization assay (an enzyme-free method) was developed as a direct means for detection of DNA or RNA target. However, currently there are many challenges that need to be addressed in order to make this technique more effective. One of the shortcomings of the commonly employed sandwich hybridization approach is that it cannot be applied for short nucleic acid target detection due to the insufficient binding affinity of natural DNA probes for the complementary DNA or RNA target. Thus, in order to gain additional binding affinity MP-γPNA amphiphile was employed. The exceptional high affinity of MP-γPNA enables a stable hybridization of a second DNA probe to the other portion of the nucleic acid analyte being interrogated. By using MP-γPNA and DNA as probes we present a method that is capable of rapid and stable sandwich hybridization detection of 22 nucleotides DNA and RNA targets. A combination of 8mer MP-γPNA and 14mer DNA was found to be optimal for detection of these relative short nucleic acid targets. Capillary electrophoresis (CE) method was adopted as a simple way to determine RNA/DNA on the basis of differences in the mobilities of the sandwich complexes, and the unbound DNA and MP-γPNA components.
The mobility of DNA/RNA is relatively fast while the sandwich complex has a lower mobility in comparison to the unbound PNA probe. In order to increase the difference in the mobility between the latter two, a large nonionic micelle (C18 alkyl group) was attached to the MP-γPNA to form a γPNA amphiphile. Due to the interaction of the n-alkane-modified γPNA with the capillary electrophoresis running buffer containing nonionic surfactant, the bound and unbound components can be easily discriminated. Fluorophores were attached to each probe to facilitate the detection of the targets. This method was found to be stable and capable of single base mismatch discrimination. This assay provides a simple yet efficient way for detection of short nucleic acid targets with high stability and selectivity. This method could be further optimized for high-throughput detection of microRNA or other biologically relevant nucleic acid targets of similar size.

2. c-Kit Drives Elevated PNA-mediated Gene Editing and Increased DNA Repair in Hematopoietic Stem/Progenitor Cell Populations

Several classes of antigenic reagents have been developed, as discussed in the introduction chapter, with the ultimate goal of using them to treat genetic diseases. One aspect our research has focused on the development of a general strategy for correcting genetic mutation in hematopoietic stem/progenitor cells (HSPCs). Owing to its high binding affinity and sequence selectivity for DNA and its ability to invade any sequence of double helical B-DNA, MP-γPNA was used for site directed triplex formation. The delivery of any antigenic reagent has been a concerning issue. In this case, MP-γPNA was delivered via polymer nanoparticles (NPs). MP-γPNA was found to substantially increase gene editing in HSPCs both ex vivo and in vivo via DNA recombinant technology mechanism in a transgenic mouse model with β-globin/green fluorescent protein (GFP) fusion transgene. MP-γPNA
mediated gene editing resulted in an increase in the level of homology-dependent DNA repair (HDR) in c-Kit-positive cells. It was shown that treatment of HSPCs \textit{ex vivo} or treatment of mice \textit{in vivo} with stem cell factor (SCF), the c-Kit ligand, further increased \(\gamma\text{PNA}\)-mediated gene editing with frequencies over 15\% \textit{ex vivo} and up to 1\% \textit{in vivo}, with no off-target hits being observed. The encouraging results from these studies indicate that MP-\(\gamma\text{PNA}\) mediated gene editing is a promising strategy for \textit{in vivo} applications. In the future, further optimization of the structure of \(\gamma\text{PNA}\) might yield better cellular uptake and gene editing efficiency.
SUMMARY

In this thesis, we discuss the methodology development, synthesis of the MP-γPNA monomers, and the utility of such molecules in nucleic acid detection and gene editing. In Chapter II, we report a systematic study examining the various synthetic routes (reductive amination, Swern oxidation, Parikh-Doering oxidation, and Mitsunobu coupling) for preparing optically-pure γPNA monomers. From this comparative study we found that the reductive amination route was prone to epimerization even under mild experimental conditions. The extent of epimerization could be minimized by utilizing a bulky protecting group such as PhFl; however, it is difficult to remove in the subsequent oligomer synthesis step. On the other hand, we discovered that the Mitsunobu route produced optically superior products using standard carbamate-protecting groups. The adverse effect of epimerization of the monomers on the binding affinity of the oligomers was also studied. In chapter III, we report for the first time the synthesis of (R)-Fmoc-MiniPEG-γPNA monomers. In order to automate the γPNA oligomer synthesis and avoid the harsh cleaving conditions employed in Boc-chemistry, we developed a new synthetic route for synthesizing Fmoc-based optically-pure γPNA monomers. Based on this finding, we developed a scheme for preparing (R)-Fmoc-MP-γPNA monomers utilizing Mitsunobu coupling to obtain optically-pure monomers. In Chapter IV, we discussed a general method for synthesizing both optically-pure (R) & (S)-Fmoc-MiniPEG-γPNA monomers from a relatively cheap starting material (L-Ser). The Mitsunobu route is not suitable for large scale synthesis of such monomers due to the difficulty in purification, and the fact that the (S)-monomers are prepared from a relatively expensive starting material (D-Ser). To address these issues, we developed a general method for
synthesizing both enantiomers from the same commercially available. The newly developed synthetic method is robust, flexible, economical, scalable, and versatile. In Chapter V, we briefly outlined two applications of γPNA in the detection of relatively short nucleic acid analytes such as microRNA and in gene editing.
FUTURE DIRECTIONS

The main focus of our research program has been on the development of γPNA as an antisense and antigene reagent. In pursuit of this goal, the recognition of DNA with high binding affinity and sequence-specificity and the ability of MP-γPNA to invade any sequence of dsDNA have been demonstrated. The future direction of our research can be divided into two parts.

1. Targeting RNA with γPNA

The treatment of most infectious diseases has been a great challenge because of the extraordinary replicative and mutational capacity of these infectious pathogens. The resistances of the pathogens to small molecule drugs and the alarming rate at which they are spreading have compelled researchers to develop new strategies to combat these evolving threats. Over the last few years it has been established that RNA perform various functions in cell and they can also fold into complex three dimensional shapes like proteins. Thus they can be biologically relevant targets for the potential treatment of infectious diseases. For this reason, much effort is going on in our group to develop a new set of nucleic acid recognition elements called Janus Bases—that are capable of forming H-bonding interactions with both strands of the target RNA. We rationalized that a combination of conformationally preorganized γPNA backbone and new recognition elements should enable the development of relatively short antisense agents for recognition of secondary and tertiary structures of RNA. The Janus Bases are similar to the small molecules with respect to their sizes while they behave as antisense/antigene agents because of their recognition of RNA targets based on H-bonding. However, unlike
small molecules the design of the Janus Bases can be modified to accommodate any evolving new target sequence of the pathogens. These γPNA antisense agents being small in size can be produced in large quantities, taken up by cells easily and will be more sequence specific towards the RNA targets. The future direction of our group is to synthesize these modified nucleobases, couple to the γPNA backbone to get the monomers and then assess their recognition capability.


Delivery of antisense/antigene reagents into cells has been a major challenge. In order to improve the cellular uptake of PNA our group had previously developed αGPNA by attaching the guanidine group at the alpha-backbone. Due to the high cost in the preparation of monomers and the relatively low binding affinities of the oligomers, we had developed a second-generation GPNA whereby we installed the guanidinium group at the gamma-backbone. In this case, the monomers are prepared from a relatively cheap L-Lys, with the corresponding oligomers exhibiting significantly stronger binding affinity and sequence specificity for both DNA and RNA. One aspect of our work was to improve the water solubility and compatibility of such class of molecules, so that they could be used to regulate gene expression in vivo. As such, we have developed a third-generation GPNA with the guanidinium group attached to the miniPEG-side-chain. For temporal and spatial control of the delivery, new designs of the backbone are being considered and much effort is going on the synthesis of the monomers. These third-generation GPNAs will be then explored as antisense and antigene reagents in cell culture and in animal models in the near future.