Design, Synthesis and Physicochemical Properties

of

Aromatic Peptide Nucleic Acids

By

Lee David Fader

A thesis submitted to the Faculty of Graduate Studies and Research of McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 2002

Department of Chemistry

McGill University

Montreal, Quebec, Canada

Your file Votre référence
ISBN: 0-612-88462-7
Our file Notre référence
ISBN: 0-612-88462-7

Dedicated to my grandmother, Ellen Fader, whose importance to me, contribution to my life and place in my heart cannot be expressed in words.

"I skate to where the puck is going to be, not where it is."
-Wayne Gretzky

"Try not. Do, or do not. There is no try."
-Master Yoda

Abstract

Synthetic methods were developed for the preparation of several structurally diverse aromatic peptide nucleic acid (APNA) monomers containing all four natural DNA bases. This set of building blocks was useful for the preparation of oligomers designed to evaluate the hybridization properties of novel peptide nucleic acid (PNA) analogues, which incorporate aromatic rings into their backbone. Protocols for the solid-phase synthesis of APNA-PNA chimeras and APNA homopolymers were also developed.

Thermal denaturation experiments involving APNA-PNA chimeras demonstrated that APNA backbones comprised of N-(2-aminomethylphenyl) glycine and N-(2-aminobenzyl) glycine as the backbone moiety exhibited good binding affinity for DNA and RNA. Further studies with PNA oligomers composed partly of repeating N-(2-aminobenzyl) glycine APNAs showed that these analogues displayed good sequence recognition for DNA and RNA. Furthermore, continuous tracts of APNA units were well tolerated in both the triplex and duplex binding modes. The binding of APNA modified oligomers was investigated using UV, circular dichroism spectropolarimetry (CD) and complex formation with the cyanine dye 3-ethyl-2-[5-(3-ethyl-3H-benzothiazol-2-ylidene)-penta-1,3-dienyl]-benzothiazol-3-ium iodide (**DiSC$_2$(5)**). In the latter experiments, binding of the dye to complexes formed between APNA modified oligomers and DNA or RNA suggested that the minor grooves of these complexes were not drastically different from those formed in the corresponding PNA:DNA or PNA:RNA complexes.

Finally, fully modified APNA homopolymers were prepared in order to investigate their physicochemical properties. These oligomers were found to be essentially insoluble in aqueous buffers, which prohibited study of their binding affinity to nucleic acids. However, a short homothymine oligomer was synthesized which was modified at the C- and N-terminals so that it can be dissolved in water in sufficient quantities to allow for evaluation of its recognition and binding to DNA and RNA. Complex formation was confirmed by CD and **DiSC$_2$(5)** binding experiments. These studies indicated that the APNA homopolymer bound to DNA and RNA, apparently through Watson-Crick base pairing, and formed a complex that was more stable than those formed by the corresponding homothymine DNA oligomer with complementary DNA under similar conditions.

Résumé

Des méthodes synthétiques ont été développées pour la préparation de nombreux monomères d'acides peptido-nucléiques aromatiques (APNA) de structures différentes à partir des quatre bases naturelles de l'ADN. Cette collection de monomères a été utile dans le but de préparer différents oligomères en vue d'évaluer les propriétés d'hybridations de nouveaux analogues d'acides peptido-nucléiques (APN) contenant des cycles aromatiques dans leur chaîne principale. Des protocoles pour la synthèse sur phase solide de chimères de APNA-APN et d'homopolymères de APNA ont aussi été développés.

La dénaturation thermique de molécules hybrides de APNA-APN a démontrée que les chaînes principales de ANPA constituées des monomères glycine N-(2-aminométhylphényl) et glycine N-(2-aminobenzyl) se sont avérées avoir une bonne affinité d'attachement pour l'ADN et l'ARN. Des études subséquentes sur des oligomères d'APN composé d'unités répétitives de glycine N-(2-aminobenzyl), ont montré que ces analogues faisaient preuve d'une bonne reconnaissance de séquence pour l'ADN et l'ARN. De plus, des étendues continues d'unité de APNA ont été bien toléré dans les modes d'attachement de type duplex et triplex. L'attachement d'oligomères modifiés d'APNA a été étudié par spectrophotométrie UV, spectropolarimétrie de dichroïsme circulaire (DC) et par la formation d'un complexe avec le pigment iodure de 3-éthyl-2-[5-(3-éthyl-3H-benzothiazol-2-ylidène)-penta-1,3-dienyl]-benzothiazol-3-ium (**$DiSC_2(5)$**). Dans ces dernières expériences, l'attachement du pigment aux complexes formés entre des oligomères modifiés de ANPA et d'ADN ou ARN ont suggéré que les petits sillons

de ces complexes n'étaient pas drastiquement différents de ceux formés dans les complexes APN:ADN ou APN:ARN.

Finalement, des homopolymères entièrement modifiés de APNA ont été préparés dans le but d'étudier leurs propriétés physico-chimiques. Ces oligomères se sont trouvés être essentiellement insoluble dans les solutions tampons aqueuses ce qui empêchait l'étude de leur affinité à l'attachement aux acides nucléiques. Cependant, un homooligomère court de thymine, modifié en position terminal C et N a été synthétisé afin de le rendre suffisamment soluble dans l'eau pour pouvoir évaluer sa reconnaissance et son attachement à l'ADN et l'ARN. La formation de complexe a été confirmé par DC et par complexation à **DiSC$_2$(5)**. Ces études ont indiquées que l'homopolymère d'ANPA s'attache à l'ADN et à l'ARN, apparemment par pairage de base de type Watson-Crick, et forment un complexes plus stables que ceux formés par les oligomères d'ADN de homotymine avec l'ADN complémentaire dans des conditions similaires.

Acknowledgements

I would like to thank my Ph.D. supervisor, Professor Youla S. Tsantrizos. From the very beginning, your interest in my future and dedication to my development as a scientist have been tireless and unconditional. I believe your trust and encouragement have paved the way for my growth and success as a chemist. Furthermore, your ambition and enthusiasm for science in general has been inspirational to say the least. Your support for me over the years has been, and will always be, greatly appreciated. Thank you for giving me the opportunity to work in your laboratory. It has truly been a pleasure.

I would like to thank the Walter C. Sumner Foundation, the Natural Sciences and Engineering Research Council of Canada and the J.W. McConnell McGill Major Foundation for support in the form of Post-Graduate Fellowships. I would also like to thank Boehringer Ingelheim Cananda for generous financial support and for making their infrastructure available to me during the course of my graduate studies.

I would like to extend sincere thanks to Professor James L. Gleason for his support during the course of my graduate studies. Professor Gleason fostered in me a love of organic chemistry from an alternative perspective I might not have otherwise seen. The opportunity to participate in his weekly group meetings provided me with a stimulating and rich experience that I am very grateful to have been a part of. Your support and encouragement has meant a great deal to me as a chemist and as a person and will never be forgotten.

I would like to thank Professor Masad J. Damha for extending the same trust, respect and kindness granted to his own graduate students to me and for making me feel

at home in his laboratory. Your guidance and encouragement through conversation, group meetings and impromptu one-on-one brainstorming sessions has been greatly appreciated. Your support over the years has been unconditional and for this I am indebted to you.

I would like to thank Mom and Dave, Dad and Agnes, Terri and Nan. Your love, understanding, encouragement and support have provided me with essential strength for this long journey away from home. You have all had a part in making me who I am today and I know I could not have made it this far without you. Thank you.

I would like to thank all of the usual suspects of Otto Maass Chemistry who have made my stay at McGill the one of the best times of my life. In particular, I thank: members of the best ball-hockey team ever assembled, Studley Hungwell/Rybaks Rejects; all three generations of the best ice-hockey team ever assembled, Co-Ed. Naked/Ex-Men/Jr Hitmen; all members of the best co-ed broomball team ever assembled, the Maassholes; easily the best co-ed softball team to step on the reservoir, the Maassholes/Moore Beer Fermi; all members of the Otto Maass Golfers Association; and finally, all members of the Otto Maass Weightlifting and Squash Club. Special mention to repeat team captains Owen Terreau and Shane Pawsey: thanks for doing all the work so we didn't have to. Throughout my time at McGill, the special release of pent up frustration provided by "athletics" has kept me sane. No Mugs.... but no straightjackets either.

I would like to extend a warm thanks to my colleagues in OMC 207 and H-1145-1. Over the years, the diverse individuals I have worked with have always provided a stimulating and friendly work environment. Special mention goes to the fellow members

of team APNA... for we have all been in the trenches together. Mike, John, Jackie and Eddie: thanks for everything.

A special thanks is also extended to François Ravenelle for translation of the abstract. Thanks is also made to Dr. Graham McGibbon (Boehringer Ingelheim Canada) for assistance with Mass Spectral analysis.

Last, but by no means least, I would like to thank my wife, Stephanie Warner. Throughout the course of my B.Sc. and Ph.D. I have done hundreds of reactions of all description, but I would like to thank you for your part in creating the strongest and most meaningful bond I will ever make. Without your unconditional love, support, encouragement, respect and guidance, I would never have come this far. Since our days at Dalhousie, you have inspired me to be better and try harder at everything I do. There are no words to express the magnitude of my appreciation for EVERYTHING you have done for me, for us. Thank you.

TABLE OF CONTENTS

List of Abbreviations

Å	angstrom
°C	degrees Celsius
Δ	delta, change in …
λ_{max}	UV maxima
μmol	micromole
μL	microlitre
δ	chemical shift in ppm
A	adenosine or PNA adenine residue
Abs	absorbance
Ac	acetyl
Ac_2O	acetic anhydride
AcOH	acetic acid
Ad	adenine
Ad^{Cbz}	N-(benzyloxycarbonyl)adenine
Alloc	allyloxycarbonyl
Alloc-Cl	allylchloroformate
APNA	aromatic peptide nucleic acid
atm	atmosphere
B	base
BnOH	benzyl alcohol
Boc	*tert*-Butoxycarbonyl

Boc_2O	di-*tert*-butyl dicarbonate
BOP-Cl	
Bz_2O_2	Benzoyl peroxide
C	cytidine or PNA cytosine residue
Cbz	benzyloxycarbonyl
Cbz-Cl	benzylchloroformate
CDN	canadian dollar
Cy	cytosine
Cy^{Cbz}	N-(benzyloxycarbonyl)cytosine
CD	circular dichroism
d	doublet
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane
dd	doublet of doublets
DECA	diethylcyclohexylamine
$DiSC_2(5)$	3-ethyl-2-[5-(3-ethyl-3H-benzothiazol-2-ylidene)-penta-1,3-dienyl]-benzothiazol-3-ium iodide
DIPEA	diisopropylethylamine
DMAP	dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DMSO-d_6	perdeuterated dimethylsulfoxide
DNA	deoxyribonucleic acid

EDC	1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride
EDTA	ethylene diamine tetraacetate, disodium salt
Et	ethyl
Et_2O	diethyl ether
EtOAc	ethyl acetate
EtOH	ethanol
equ.	equivalent
Fmoc	9-fluorenylmethoxycarbonyl
Fmoc-Cl	9-fluorenylmethyl chloroformate
Fmoc-ONSu	N-(9-Fluorenylmethoxycarbonyloxy) succinimide
G	guanosine or PNA guanine reside
g	gram
Gu	guanine
Gu^{Cbz}	N-(benzyloxycarbonyl)guanine
h	hour
HATU	O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HPLC	high pressure liquid chromatography
hυ	UV irradiation
J	J-coupling constant

K	degrees Kelvin
K	equilibrium constant
Lys	lysine, lysyl
m	multiplet
M	molar, molarity
ma	major
MBHA	methylbenzhydrylamine
MCPI	N-methyl-2-chloropyridinium iodide
Me	methyl
MEM	2-methoxyethoxymethyl
MeOH	methanol
mi	minor
mg	milligram
MHz	megahertz
min	minutes
mL	milliliter
mmol	millimole
mM	millimolar
MsCl	methanesulfonyl chloride, mesyl chloride
mol	mole
mol%	percent by mole, mole percent
mRNA	messenger RNA
NaOAc	sodium acetate

NBS	N-bromosuccinamide
nm	nanometer
NMO	N-methylmorpholine-N-oxide
NMR	nuclear magnetic resonance
nOe	nuclear Overhauser enhancement
OPA	olefinic peptide nucleic acids
OPNA	oxy- peptide nucleic acids
Pfp	pentafluorophenyl
PG	protecting group
pH	$-\log[H^+]$
poly dA	polydeoxyriboadenylate
poly rA	polyriboadenylate
poly rU	polyribothymidylate
PHONA	phosphono peptide nucleic acids
PNA	peptide nucleic acid
ppm	parts per million
PS	polystyrene
pyr	pyridine
q	quartet
RNA	ribonucleic acids
RNase H	ribonuclease H
rRNA	ribosomal RNA
RP	reverse phase

RT	room temperature
S	Svedberg unit
SPPS	solid-phase peptide synthesis
T	thymidine or PNA thymine residue
t	triplet
TBAF	tertabutylammonium fluoride
tBu	*tert*-butyl
tBuOH	2-methyl-2-propanol
TDS	*tert*-butyldiphenylsilyl
TDS-Cl	chloro-*tert*-butyldiphenylsilylsilane
Th	thymine
TEA	triethylamine
TFA	trifluoroacetic acid
TFMSA	trifluoromethanesulfonic acid, triflic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
T_m	thermal melting temperature
tRNA	transfer RNA
UV	ultra violet
VT	variable temperature

List of Tables

List of Figures

List of Schemes

CHAPTER 1:

INTRODUCTION

1.1 - The Central Dogma of Molecular Biology.[1,2,3]

From the days of the earliest developments in molecular and cellular biology more than 130 years ago, science has steadily revealed the secrets behind the complex cellular machinery involved in all living organisms. As the importance of nucleic acids in molecular biology became evident, elucidation of the structure and function of these fascinating biopolymers attracted immense interest over the decades following their discovery. Indeed, as more is learned about the function of DNA and RNA, more questions are raised concerning the relationship between the information stored in genomic DNA and RNA and the protein molecules whose synthesis is encoded within that information. With the recent mapping of the entire human genome[4] and the advent of proteomics,[5] this trend will undoubtedly continue at an unprecedented pace. Science is at the dawn of a new era in understanding cell function and human diseases related to our genetic material.

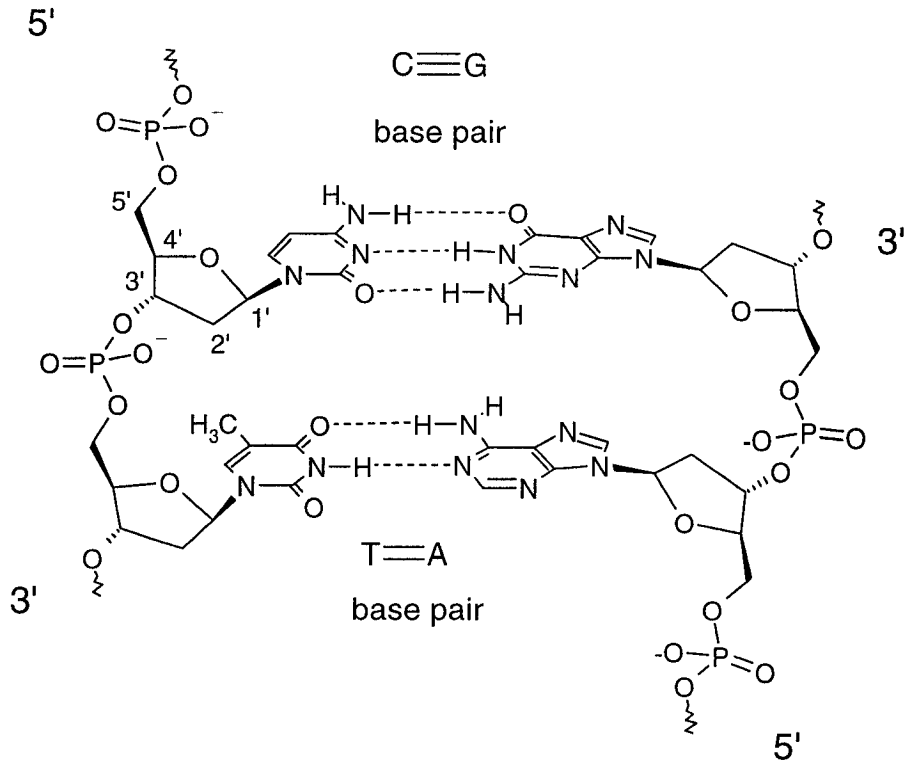


Figure 1.1: Basic Structure of Double Stranded DNA.

In the 1950's, the landmark elucidation of the three dimensional structure of the double helix DNA by Watson and Crick shed light on several unique features of this biopolymer whose importance will become clear throughout this chapter.[6] DNA molecules, which are the primary carriers of the genetic information, are made up of repeating 2-deoxyribose residues that are linked together by phosphodiester moieties (Figure 1.1). At C1', each pentose ring carries one of the four heterocyclic bases and it is the sequential arrangement of these bases along the polymer that nature uses to store the genetic information for every cell function. DNA molecules exist primarily as double stranded complexes between two complementary DNA sequences. The sequence recognition inherent to the complex comes from "base pairing" between the heterocyclic bases through an array of hydrogen bonds. In this way, adenine (A) pairs with thymine

(T) through a set of two hydrogen bonds and guanine (G) pairs with cytosine (C) through three hydrogen bonds. In each base pair, the purine and pyrimidine bases are co-planar and stacked on top of each other creating the hydrophobic core of the helix (Figure 1.2). The helix itself is right-handed and the two strands are oriented in an antiparallel sense, with the 5'-end of one strand pairing up with the 3'-end of its complement. The two strands wind around the core of the helix, creating two grooves of differing dimensions, known as the major groove and minor groove, and a hydrophilic outer surface by virtue of the negatively charged phosphodiester linkages. The rise of the helix is approximately 34 Å (about 10-11 bases per turn) and in nature, there are two predominate types of helices known, the A-form and B-form helices.[3] The differences between the A-form and B-form helixes are in the dimensions of the minor and major groove, the helical rise and the conformation of the ribose units. Double helices of DNA are typically B-form, while those involving one or more RNA strands tend to be A-form.[3]

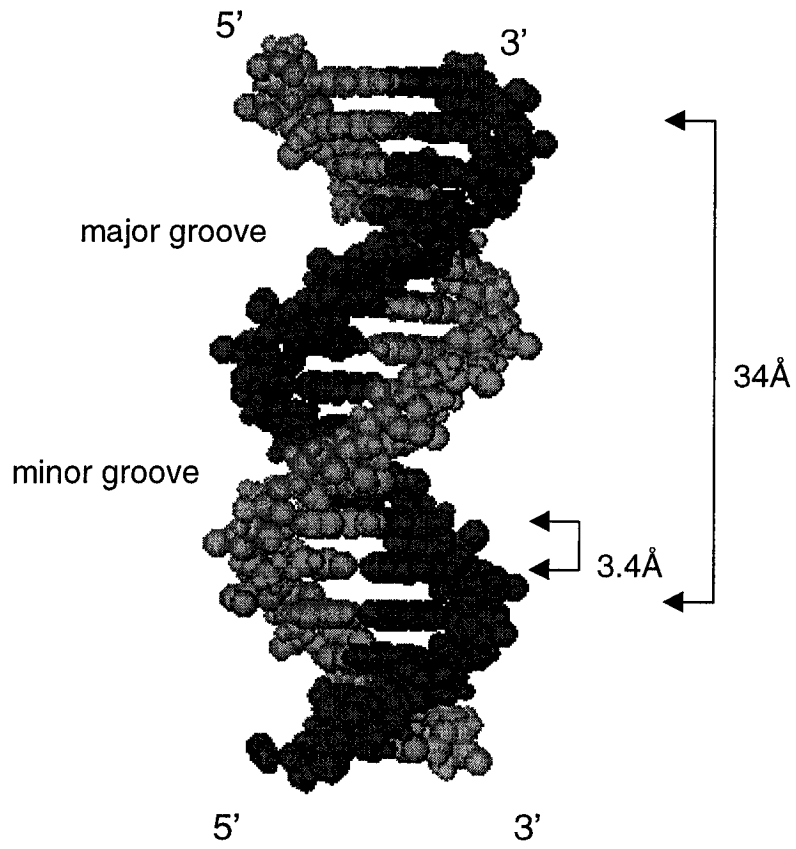


Figure 1.2: The DNA B-Form Double Helix.[7]

From the surface, the central dogma of molecular biology is deceptively simple (Figure 1.3). In eukaryotes, during the process of *transcription*, the base sequence of double stranded DNA of a particular gene is "read" and a complementary RNA molecule is synthesized. Subsequently, in the process of *translation*, the RNA molecule is moved from the nucleus to the cytosol where it is then "read" in steps of 3 nucleotides (codons) to direct the biosynthesis of a specific protein molecule. However, the underlying complexity of these processes is bustling with interactions between DNA, RNA and protein molecules. This interplay is mediated by specific recognition events, which serve to initiate, regulate and terminate different biochemical pathways. Apart from the

fundamental goal of understanding how these pathways operate and interact with one another, probing these recognition events could potentially lead to novel therapeutic intervention in biological systems that have been compromised by disease or infection.

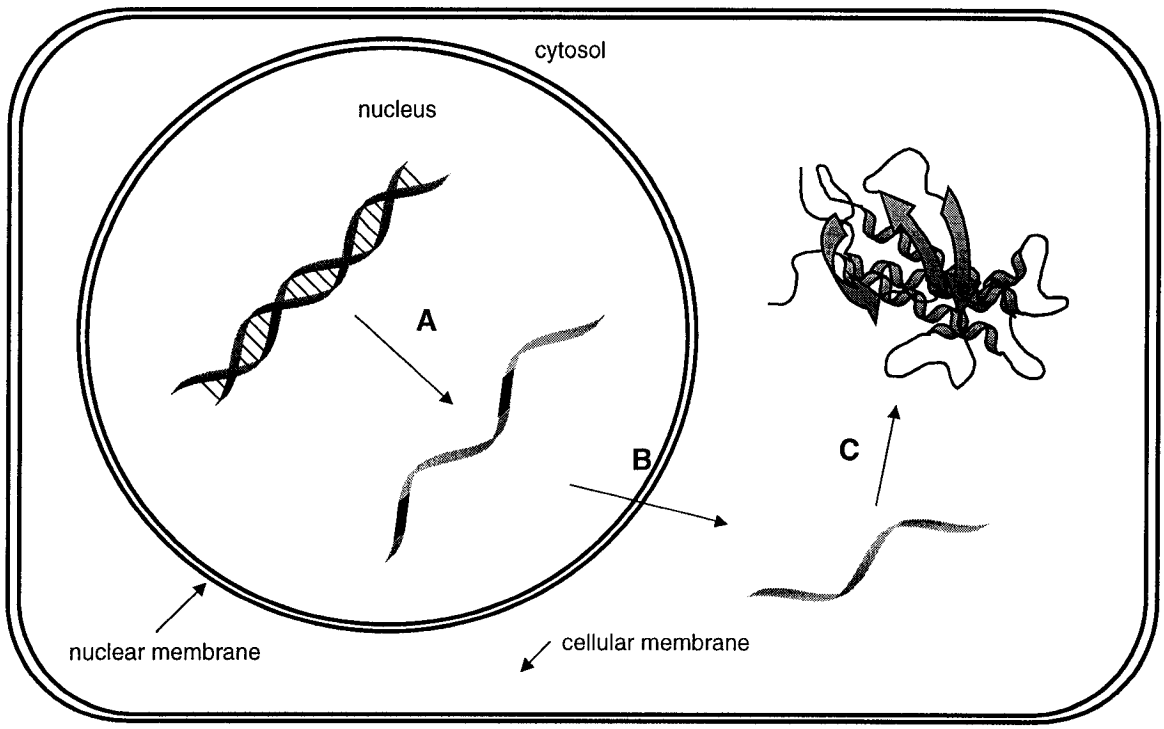


Figure 1.3 The Central Dogma of Molecular Biology: A) Transcription; B) RNA processing and transport to cytosol; C) Translation

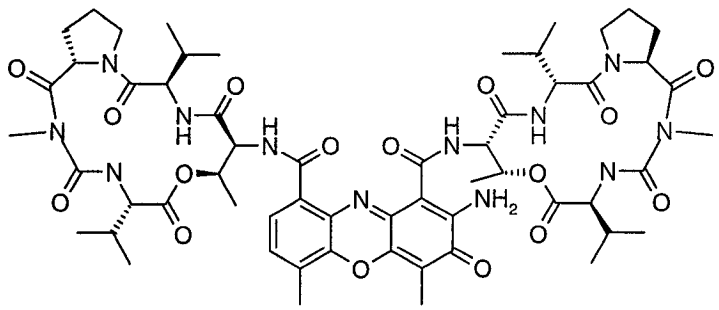
1.2 - Nucleic Acids as Drug Targets.[8,9]

As drug targets, DNA and RNA display many potential sites for drug interaction. In the ideal case, a drug would interact with DNA in a sequence selective way and allow for regulation of the transcription of a *specific gene* associated with the biological system of interest. While most of an organism's DNA is widely diffuse throughout the nucleus during the interphase portion of the cells life cycle, some parts of the double helical

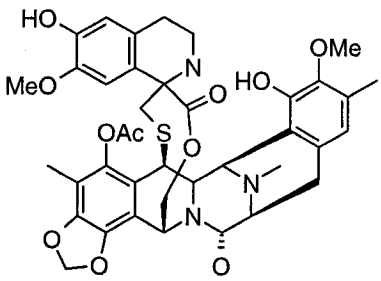
molecules are associated with protein molecules and so may not be readily accessed. Furthermore, since the DNA of eukaryotic cells is localized within the cell nucleus, it is important that the drug is able to traverse both the cellular and nuclear membranes. There are currently numerous drugs either in clinical use or in clinical trials whose therapeutic effect rely upon their interaction with double stranded DNA. Generally, these can be grouped into two classes: the minor groove binders, examples of which include actinomycin D (**1**),[10] ecteinascidin 743 (**2**)[11] and distamycin (**3**);[12] and neocarzinostatin chromaphore (**4**), [13] and the major groove binders such as adriamycin (**5**), [14] mitoxantrone (**6**)[15] and nogalamycin (**7**).[16]

Once a DNA molecule is transcribed to its complementary RNA, enzymatic processing or modification of the RNA is required to form the functional messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA) molecules. These biopolymers become associated with a plethora of different proteins en route to their intended function. In contrast to DNA, the function of RNA is far more than just to encode information for protein biosynthesis. Some types of RNA molecules exist which become part of enzymes or other cellular machinery, without ever serving as a template for protein biosynthesis. Examples include the 5S, 5.8S and 28S rRNAs of the 60S ribosomal subunit and the 18S RNA of the 40S ribosomal subunit. Other examples include tRNA molecules, which are critical to translation, and also the RNA portion of the telomerase holoenzyme complex, which is involved in telomere synthesis in certain cell types.[17] Typically, RNA molecules have a unique three-dimensional topology, involving many double stranded and hairpin loops, which are created by complementary base pairing at different regions of the RNA strand. Thus, drugs targeted to RNA could

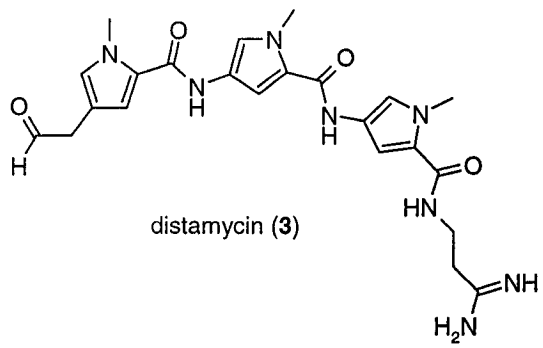
disrupt translation by interacting with mRNA, or could interfere with protein-RNA recognition events leading to a disruption of other important biological pathways.[10]



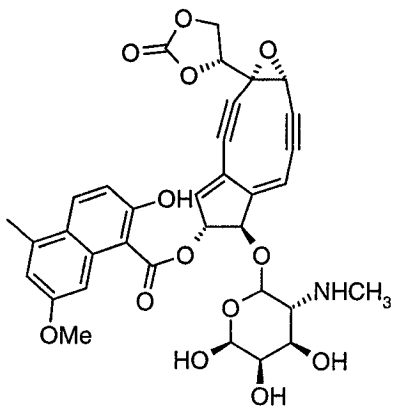
actinomycin D (**1**)



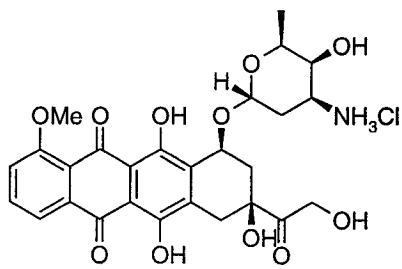
ecteinascidin 743 (**2**)



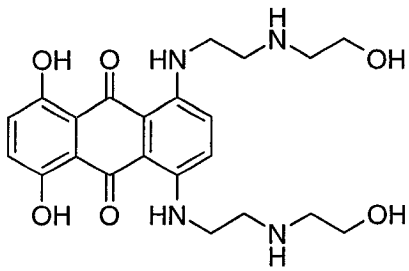
distamycin (**3**)



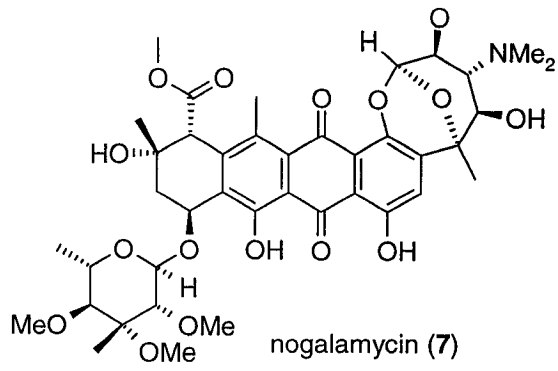
neocarzinostatin chromaphore (**4**)



adriamycin (**5**)



mitoxanthrone (**6**)



nogalamycin (**7**)

1.3 - Nucleic Acids as Therapeutic Tools.

The potential use of nucleic acids in drug development is immense, due to the important roles they play in cell biology. However, unmodified nucleic acids are not appropriate for most clinical applications since they are readily degraded in biological fluids and due to their negatively charged phosphodiester backbone, they do not readily penetrate cell membranes. Mimicking DNA or RNA structure is a proven strategy for interfering with biological pathways *in vitro* and could potentially be extended to pharmacological intervention. There are four ways in which DNA/RNA or their analogues have been implemented for clinical use: 1) nucleoside analogues as enzyme inhibitors;[18] 2) modified oligonucleotides as antisense/antigene agents (Section 1.4);[19] 3) modified oligonucleotides as enzyme inhibitors;[20] and 4) oligonucleotides as diagnostic tools.[21] While this thesis will focus primarily on development of an antisense oligonucleotide analogue, it is important to point out that the molecules that are described could potentially be used in the other applications mentioned.

1.4 - The Antisense and Antigene Strategies for Inhibition of Gene Expression

In 1978, Zamecnik and Stephenson[22] first demonstrated the therapeutic use of synthetic oligonucleotides in what became known as the antisense strategy. While many drugs target particular proteins in order to exert their therapeutic effect, antisense intervention occurs upstream in the flow of genetic information at the stage of protein biosynthesis (Figure 1.4). After processing of the *pre-m*RNA transcript and transport of the mature mRNA into the cytosol, a synthetic oligonucleotide binds in a sequence selective manner to the target mRNA and physically blocks the ribosomal machinery, inhibiting the biosynthesis of protein. Alternatively, it has been suggested that the ideal antisense molecule would also elicit RNase H, an endonuclease responsible for the degradation of the RNA strand of a RNA:DNA duplex.[19] This would essentially reduce the amount of the antisense drug required since the target is degraded, resulting in a mechanism of irreversible inhibition, and the drug is made available for further action.

In addition to interaction with mRNA, leading to antisense down regulation of a gene, synthetic oligonucleotides could also potentially interact with microbial nucleic acid-processing enzymes (such as helicase or DNA/RNA-dependent polymerase), thus blocking essential steps in the microbial life cycle. Synthetic oligonucleotides that are complementary to the template RNA in human telomerase have also been used to inhibit telomere elongation in cell extracts of the melanoma cell line SK-Mel-28.[23]

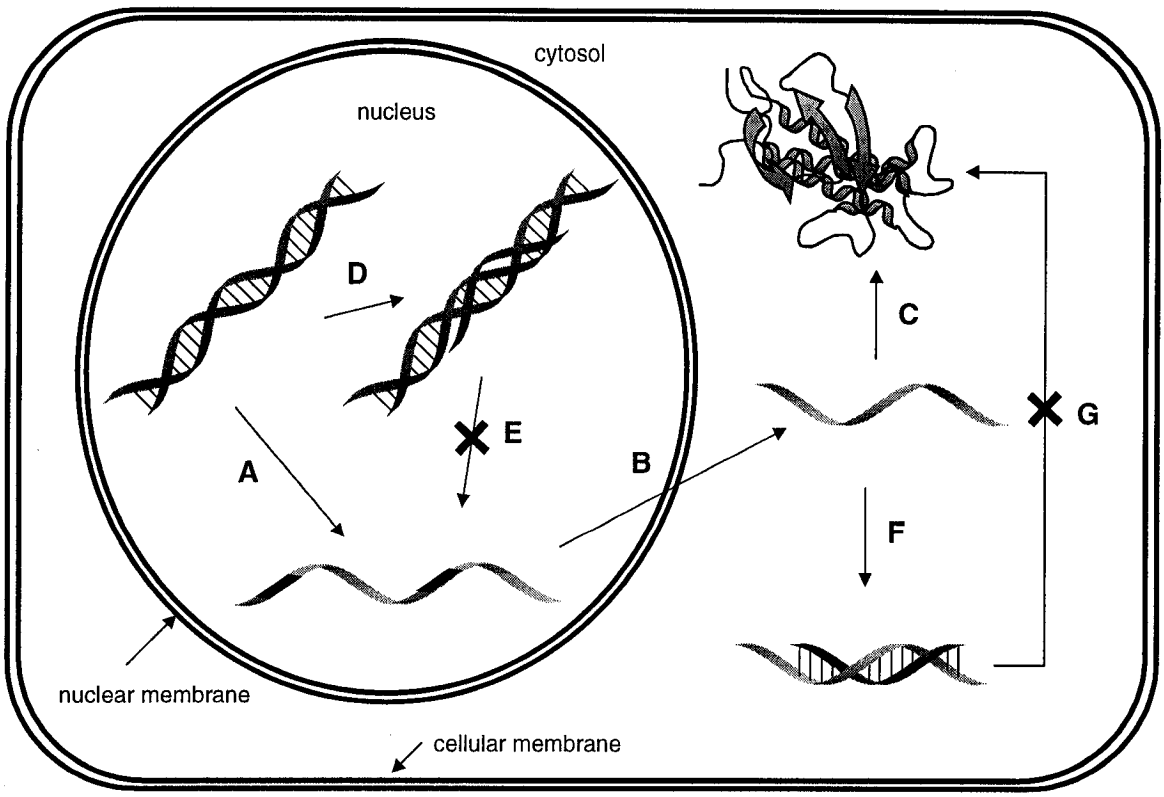
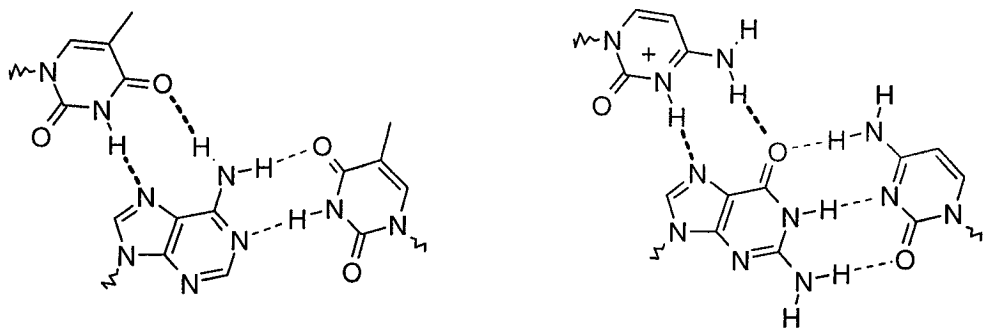


Figure 1.4: The Antisense and Antigene Strategies: A) Transcription; B) RNA processing followed by transport to cytosol; C) Translation; D) Hybridization of Synthetic Antigene Oligonucleotide; E) Inhibition of Transcription; F) Hybridization of Synthetic Antisense Oligonucleotide; G) Inhibition of Translation.

An alternate mode of inhibition of gene expression was proposed more recently which became known as the antigene strategy.[24] In this approach, a synthetic oligonucleotide binds to the major groove of double stranded DNA in a sequence selective manner. By blocking the major groove, this event then prohibits binding of proteins and/or enzymes involved in transcription, effectively shutting down production of the primary RNA transcript. However, in this approach a different type of "base pairing", known as Höogsteen (or reverse Höogsteen) base pairing, mediates the sequence recognition (Figure 1.5). Currently, recognition of the central base of the triplex is limited to recognition of adenine and guanine and so the sequences that can be

targeted can only be polypurine tracts. Furthermore, recognition of guanine requires N3 of cytosine in the third strand to be protonated, necessitating a pH < 6. Although some progress has been made in recognition of thymine[25] and guanine[26] by modified nucleobases, the antigene approach has not yet reached general applicability.



T*A-T triplet C+*G-C triplet

Figure 1.5: Hydrogen Bonding in Triplex Formation; Höogsteen (---), Watson-Crick (---) hydrogen bonding.

1.5 - Oligonucleotide Modification.

An enormous amount of effort has been made to develop structurally modified nucleosides that retain the recognition properties of DNA or RNA while overcoming problems associated with natural nucleic acids. There are essentially three ways in which the general structure of DNA can be modified (Figure 1.6): 1) modification of the nucleobase, 2) modification of the phosphodiester linkage, or 3) modification of the sugar moiety.

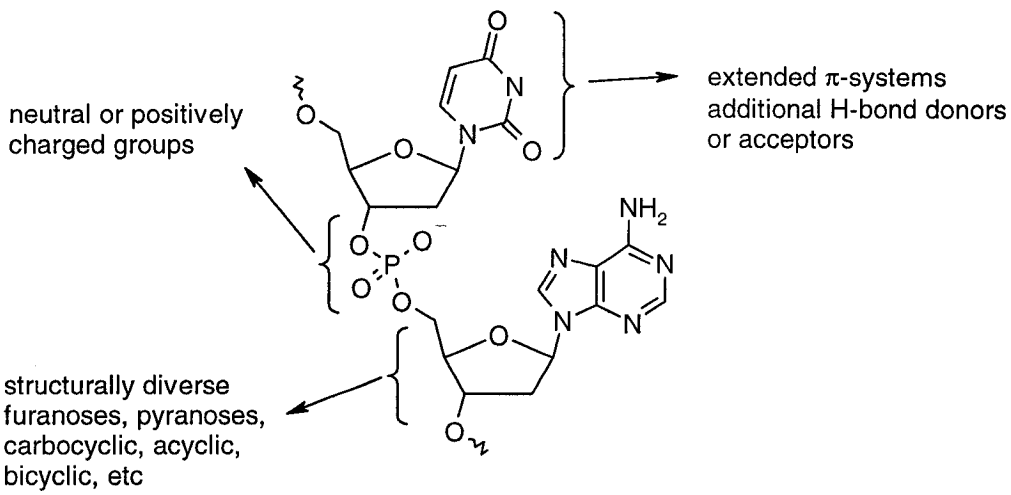
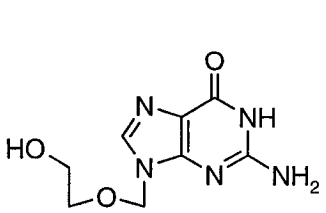
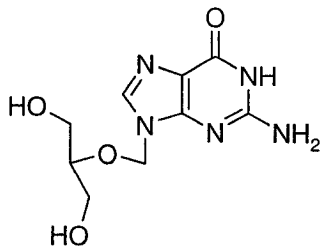


Figure 1.6: Strategies for Structural Modification of DNA.

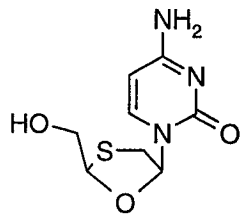
While most modified nucleosides serve as building blocks for preparation of the corresponding modified oligonucleotides, a large range of applications center on modified nucleosides on their own as antiviral (or anticancer) agents. For the latter application, sugar modifications are used almost exclusively (Figure 1.7). Modification of the sugar moiety has come in many forms ranging from simple inversion of stereocenters on the ribose ring to complete replacement of the furanose skeleton with other structures (Figure 1.8).[27] When modified nucleosides are used as building blocks for oligonucleotide synthesis, the goals can vary from development of an antisense candidate to other interesting studies, such as, investigation of the recognition elements between enzyme and nucleic acids,[28] to the chemical etiology of RNA and DNA as nature's bearer of genetic information.[29]



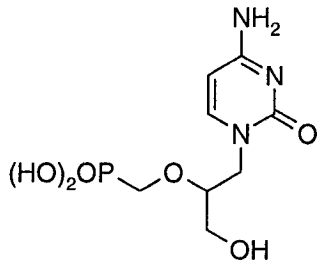
acyclovir (**8**)



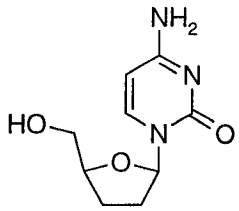
gangciclovir (**9**)



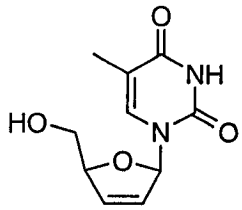
lamivudine (**10**)



cidofovir (**11**)

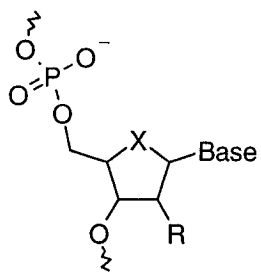


zalcitabine (**12**)



stavudine (**13**)

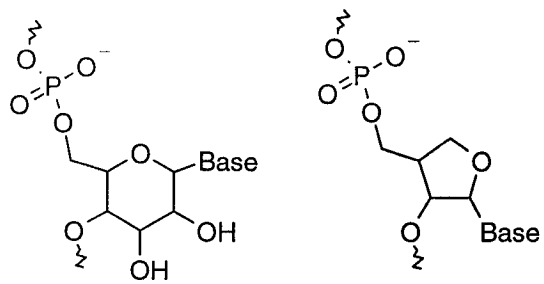
Figure 1.7: Antiviral and Anticancer Nucleoside Analogues.[18]



14

atom substitutions

R = F, OMe, OMEM, $OCH_2CH_2CH_2NH_2$
X = O, S, CH_2

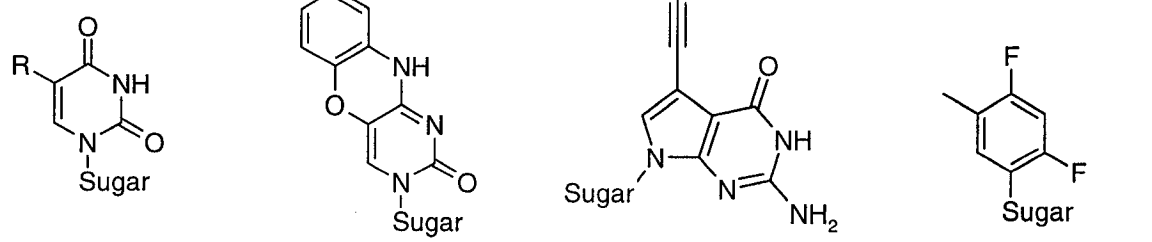


15

other carbohydrates

Figure 1.8: Other Sugar Modified Nucleoside Analogues.[28,29]

Modification of the purine or pyrimidine base usually involves extending the π-system or the hydrophobic surface of a base in order to increase the stacking potential between base pairs, leading to an increase in the thermodynamic stability of the duplex (Figure 1.9). The bases have also been modified to include additional hydrogen bond donors/acceptors, which also aids complex stability and sequence recognition.[30]



16 **17** **18** **19**

C5 substituted uracil
R = C≡C-R'
CH_2OH
CN
alkyl
CO_2H
CF_3

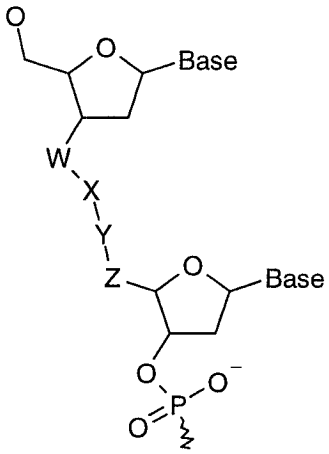
extended π-systems

nucleobase isosteres

Figure 1.9: Selected Base Modifications.[30]

To date, most of the research that has been reported deals with modification of the phosphodiester linkage (Figure 1.10).[31] These types of modifications can potentially solve both the problems of *in vivo* stability and cell membrane permeability, since both of these properties are linked to the phosphodiester backbone. Many alternative linkages have been reported;[31] the phosphorothioate linkage (**20**), where one of the non-bridging oxygen atoms has been replaced with a sulfur atom, is the most successful analogue

reported in terms of its application in antisense technology. This oligonucleotide analogue exhibits good DNA and RNA recognition properties[32] and also enhanced *in vivo* stability.[33] A hallmark of the phosphorothioate linked DNA is that it forms duplex structures with RNA that are recognized by RNase H leading to degradation of the target RNA by the enzyme.[34] However, one drawback of this modification is that the oligomers exhibit non-specific interactions with proteins.[35] Nonetheless, phosphorothioate oligonucleotides were the first antisense oligomers to receive FDA approval for the treatment of retinitis caused by the human cytomegalovirus and are currently in various phases of clinical trials for the treatment of a number of other diseases.[36]



	W =	X =	Y =	Z =
20	O	O=P-S⁻	O	CH_2
21	O	S=P-S⁻	O	CH_2
22	O	O=P-BH_3⁻	O	CH_2
23	O	O=P-CH_3	O	CH_2
24	CH_2	NH	C=O	CH_2
25	NH	O=P-O⁻	O	CH_2
26	CH_2	NCH_3	O	CH_2
27	O	O=P=O⁻	CH=	=CH
28	O	CH_2	S	CH_2

Figure 1.10: Examples of Linker Modifications.[31]

1.6 - Peptide Nucleic Acids (PNAs).

First introduced by Nielsen and coworkers in 1991,[37] peptide nucleic acids (PNAs) are DNA analogues having the phosphodiester/deoxyribose backbone completely replaced by N-(2-aminoethyl)-glycine units to which the nucleobases are attached via acetate linkers (Figure 1.11). The design of PNAs (as DNA mimics) was based on molecular modeling calculations where a conformationally rigidified single strand of a DNA duplex was structurally modified from the native DNA skeleton to a polyamide backbone. In this way, the distance between adjacent nucleobases was kept constant at 6 σ-bonds and the distance from the base to the backbone remained as 3 σ-bonds. The original hypothesis was that the bases would remain optimally spaced for Watson-Crick base pairing and the polyamide structures would retain the recognition properties of DNA.

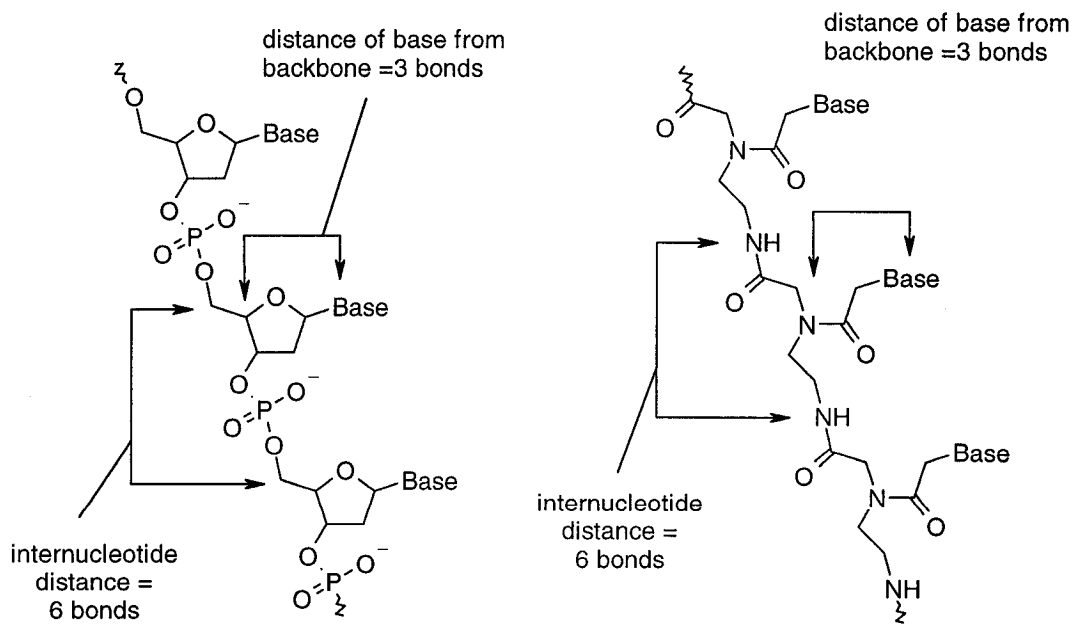
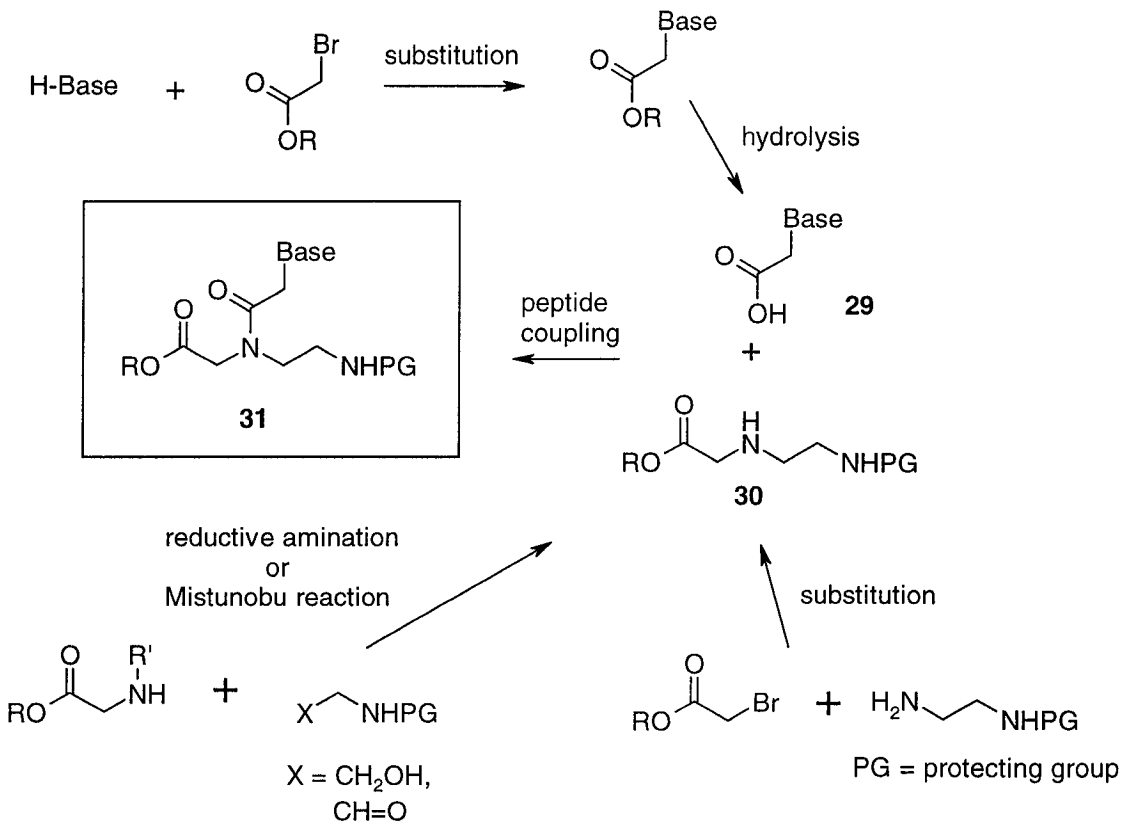


Figure 1.11 Comparison of the Structures of DNA and PNA.

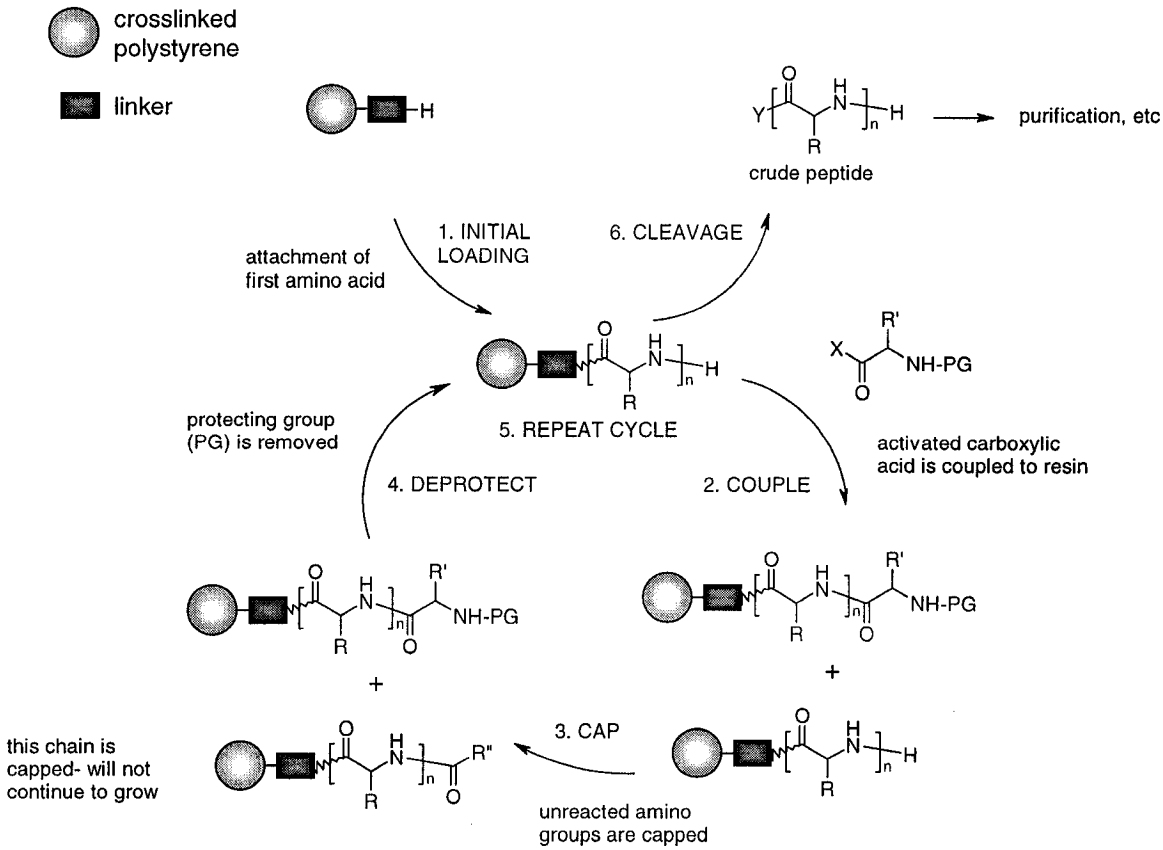
Since the repeating units of PNA monomers are acyclic and achiral, the synthesis of these molecules is relatively straightforward compared to other modified oligonucleotide analogues (Scheme 1.1). In most approaches, the PNA backbone (**30**) is assembled first and then the nucleobase acetic acid derivative (**29**) is attached to complete the monomer framework (**31**).[38] In essence, the fully assembled monomer is an unnatural amino acid. This property allows for the PNA oligomers to be synthesized using the robust methods of modern peptide chemistry, rather than the more sensitive methods employed in the synthesis of oligonucleotides.

Scheme 1.1 Basic Representation of PNA Monomer Construction.



Peptide nucleic acid oligomers are usually assembled using modern solid-phase peptide chemistry based on Merrifield's approach,[39,40] although more recently innovative sub-monomer[41] and solution phase strategies[42] have been successfully employed. The Merrifield approach involves growing the peptide chain on an insoluble polymer, usually made of polystyrene crosslinked with divinylbenzene (Scheme 1.2). The peptide is attached to the polymer by a linker group that will allow for cleavage of the final product from the resin under conditions orthogonal to those used during the chain elongation. The peptide is assembled one monomer at a time usually by coupling carboxyl-activated, N-protected amino acids to resin-bound free amino groups. Since the product of the coupling reaction is bound on the solid support, large excess of the starting materials can be used to drive the reaction to completion. When the reaction is complete, excess reagents are removed by filtration and washing of the resin. Unreacted amino groups are usually "capped" in the next step of the coupling cycle in order to prevent accumulation of failure sequences on the solid support. The final step of the coupling cycle is the deprotection of the (n+1)-elongated peptide to liberate a free amino group for the next round of coupling. After the desired sequence has been synthesized, a cleavage step releases the peptide from the solid support where it is further purified to homogeneity.

Scheme 1.2: Solid-Phase Peptide Synthesis.



Since the first reports on the synthesis of PNAs in the early 1990's, considerable effort has been made to refine the solid-phase methodology used for making these oligomers. Although some potentially useful advances have been made, the conditions described originally by Nielsen and coworkers are suitable for the synthesis of most PNA analogues.[40] After synthesis and cleavage from the solid support, the oligomers are generally purified by preparative reverse-phase HPLC using slow gradients from water to acetonitrile. The HPLC column is usually heated to 50°C, in order to separate failure sequences from the desired compounds.[43]

1.7 - Structure and Physicochemical Properties of PNA Hybrids with DNA, RNA and PNA

Although PNA was originally designed as a DNA mimic, this molecule exhibits several unique and unexpected molecular recognition properties.[44] In the earliest hybridization studies of PNAs with natural oligonucleotides, it was discovered that PNAs bind to DNA and RNA with remarkably higher thermal stability relative to all other synthetic or natural oligonucleotides. This discovery was accompanied by the observation that homopyrimidine PNA oligomers bind to their target homopurine strands to form exclusively triplex structures (Figure 1.12a). During the hybridization experiments, no intermediate duplex could be observed, suggesting that PNA_2:DNA and PNA_2:RNA triplexes are far more stable than the corresponding PNA:DNA and PNA:RNA duplexes. Another very unique binding characteristic of PNAs is the strand invasion mechanism by which PNAs hybridize to double stranded DNA (Figure 1.12b). In this scenario, a short antisense PNA molecule displaces the "antisense" strand of a longer DNA duplex to form a loop of the displaced DNA oligomer and a PNA:DNA duplex with the "sense" strand. Furthermore, when the PNA strand is rich in pyrimidines, the analogous invasion triplex is formed. Another remarkable feature of PNAs is that they are capable of forming complexes with complementary PNA strands in both the duplex and triplex binding modes. While this discovery has little therapeutic value, it is intriguing that oligomers composed of flexible, acyclic monomers are capable of forming complexes with such remarkable properties. PNAs also exhibit remarkable

sequence selectivity and the order of increasing complex stability has been shown to be: PNA:PNA > PNA:RNA > PNA:DNA.[45]

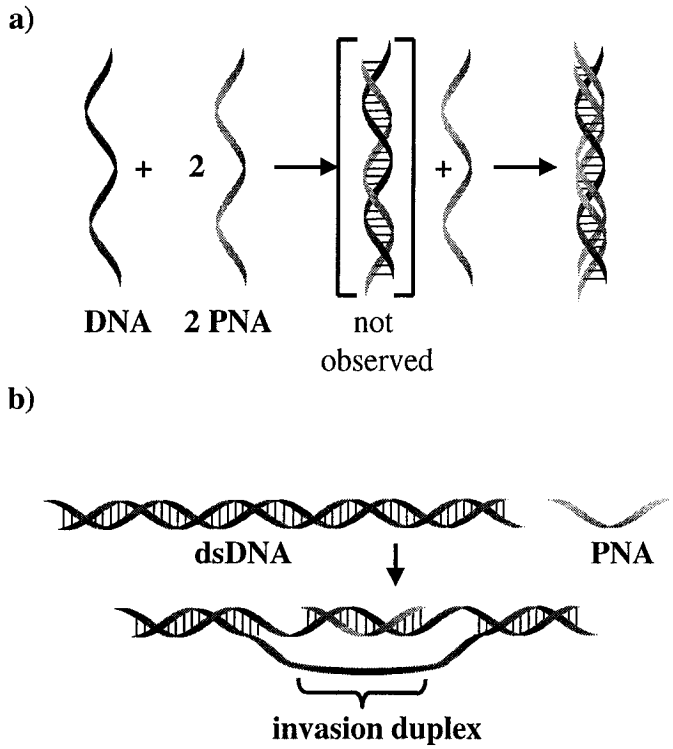


Figure 1.12 a) Representation of the Binding of Homopyrimidine PNA to DNA; b) Strand Invasion Mechanism of PNA.

The discovery of the unusual properties of PNAs initiated a number of studies aimed at understanding why PNAs are capable of forming such stable complexes with DNA and RNA. A number of NMR and molecular modeling studies have been used to explore the three dimensional structure of PNA hybrids.[46,47,48,49] Consequently, it is now known that duplexes involving PNAs are neither B-form nor A-form. Both NMR and circular dichroism studies have revealed that the bases of PNA hybrids can stack as either in the A-form helices preferred by RNA or as in B-form helices preferred by DNA.

However, PNAs are unique in that they prefer to form double helices with wider and larger helical rise than native duplexes (about 18 base pairs per turn). This finding suggests that PNAs are not perfect DNA mimics and that there is room for improvement in designing peptidic DNA analogues.

While PNA:DNA and PNA:RNA complexes are all right-handed helices (presumably due to the stereochemistry of the natural strand), complexes composed only of PNAs exist as mixtures of right- and left-handed helices.[50] Nielsen's group has shown that the sense of a PNA:PNA:PNA triplex could be controlled by incorporation of a single insert of a chiral PNA monomer in the middle of the homopurine strand.[50] To date, it is still not fully understood why PNA forms such stable complexes with DNA, RNA and itself. However, due to the potential applications of PNAs in antisense chemotherapy and/or in diagnostics tools, studies which address this question will continue to be of importance.

1.8 - Biological Properties and Applications of PNA.

A key feature of PNAs that makes them attractive candidates for drug development is that they are stable toward degradation by protease and nuclease enzymes, a critical requirement for oligomers used in antisense/antigene drug development.[51] Although the microinjection of PNA oligomers into cells has been shown to induce transcription and translation arrest,[52] their low water solubility and poor cell membrane permeability seriously compromise their therapeutic utility. In an effort to

overcome poor membrane permeability, PNA conjugates of nuclear internalizing peptides were recently explored. For example, biotinylated PNA conjugates linked to the monoclonal antibody OX26 were shown to undergo *in vivo* receptor-mediated transcytosis through the blood-brain barrier,[53] thus effectively transporting the PNA oligomers into brain cells. PNAs have also been transported into cells as unmodified PNA-DNA duplexes encapsulated into cationic lipid micelles. This latter approach has been used to effectively inhibit telomerase activity in DU145 prostate-derived tumor cells.[54] More recently, PNA has been found to permeate certain types of cell membranes including the cell membranes of nerve cells in culture[55] and *in vivo*[56] and the cell membranes of various strains of *E. coli*.[57] Although in both of these instances, down regulation of certain genes was reported, the latter findings have great implications for the use of PNA as a general antisense antibiotic. It is noteworthy that although PNA:RNA duplexes do not recruit RNase H, the potential applications of PNAs in antisense chemotherapy is still of significant interest.[58] In addition to antisense applications, PNAs are also valuable tools in clinical research. Examples include *in situ* hybridization techniques,[59] PCR modulation,[60] array hybridization[61] and hybridization detection by mass spectroscopy.[62]

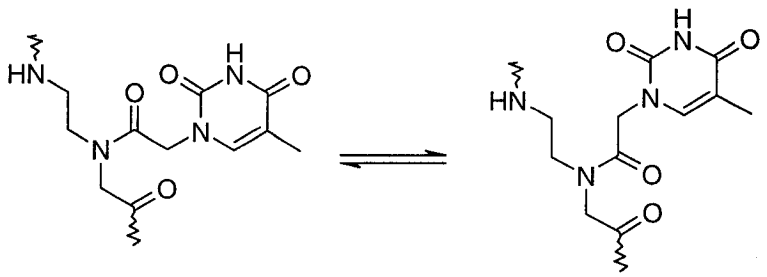
1.9 - Structural Analogues of PNA.

A great deal of effort has been devoted to the design of PNA analogues that retain the unique properties of PNA while at the same time attempting to address the problems

associated with these molecules. Much of the effort has centered on improving the water solubility of PNA oligomers. However, a considerable amount of work has also focused on understanding the importance of certain structural elements of PNAs in modulating its hybridization affinity towards DNA and RNA.

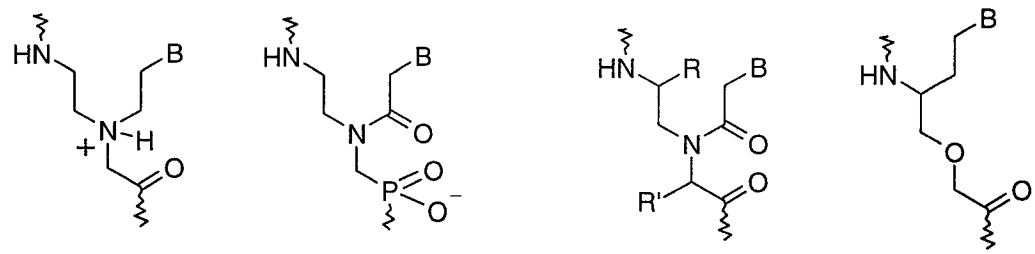
Since PNAs are acyclic and therefore conformationally quite flexible, it has been speculated that structural changes that would rigidify the PNA backbone may help to preorganize the oligomers in favor of duplex or triplex formation.[63] The tertiary amide moiety of the PNA monomer that tethers the nucleic acid base to the backbone has received much attention as a target for modification. NMR and X-ray crystallographic studies have independently shown that this amide bond is oriented with the carbonyl oxygen toward the C-terminal of the peptide [**Z(O)-32**] in PNA:DNA,[46] PNA:RNA[47] and PNA:PNA[48] complexes. However, the two amide rotamers are in rapid equilibrium in the single stranded PNA oligomers. Early molecular modeling studies suggested that the carbonyl oxygen atom of the tertiary amide might participate in an intramolecular hydrogen bond, helping to preorganize PNA oligomers for duplex formation.[49] Although these studies are not conclusive, Nielsen[64] and Leumann[65] have independently confirmed the importance of this carbonyl moiety. For example, Nielsen and coworkers first demonstrated that the Eth-PNA analogues (**33**, Figure 1.13),[64] although being highly water-soluble and polycationic, displayed no hybridization affinity for complementary DNA or RNA. Leumann later provided evidence that the conformation and structural rigidity of the backbone amide are not the sole contributing factor for hybridization stability through experiments involving E- and Z-Olefinic Peptide Nucleic Acids (OPAs, **37**, **38**).[65] While the E-OPA modification (**37**), which is analogous to the Z(O)

conformation of PNA, was found to have better hybridization properties than the Z-OPA modification (**38**), consistent with NMR and X-ray data, both analogues gave less stable complexes with DNA and RNA than the corresponding complexes with unmodified PNAs. Neilsen subsequently showed that entropic factors remain critical for good binding properties since the conformationally rigid pyrollidinone- (**39**) and pyrollidine-PNAs (**40**) form complexes with complementary DNA, RNA, and also PNA of similar stability to that of unmodified PNAs.[66] Although more experiments are still required to fully elucidate the role of the branching amide moiety, it is generally believed that it is both the amide geometry, as well as intramolecular electrostatic interaction with adjacent residues that contribute to the molecular recognition of PNAs for natural oligonucleotides.



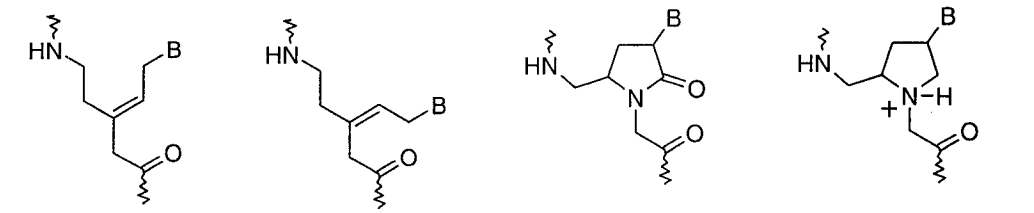
E(O)-32 **Z(O)-32**

a) PNA analogues having increased water solublity.[64,67,68,69]



Eth-PNA (**33**) | phosphono-peptide nucleic acids PHONA (**34**) | chiral PNA (**35**), R or R' = charged amino acid side chain | oxy- peptide nucleic acids OPNA (**36**)

b) Structurally rigid PNA analogues.[65,66]



olefinic peptide nucleic acids

E-OPA (**37**) | Z-OPA (**38**) | pyrollidinone-PNA (**39**) | pyrollidine-PNA (**40**)

Figure 1.13: Selected Examples of Modified PNA.

1.10 - Aromatic Peptide Nucleic Acids (APNA).

To this date, no general solution to the cell membrane permeability problem of PNAs has been reported. Although it is known that PNAs are capable of exerting an antigene or antisense effect in cell culture and *in vivo* in special cases, these molecules are unable to permeate the membrane of most eukaryotic cells. Thus, at the time that the work described herein was initiated, it was speculated that if PNA analogues were

designed to be more hydrophobic, their cell membrane penetrating properties might be improved. A class of PNA analogues, which incorporated an aromatic ring as an integral part of the backbone, was developed (**41**, Figure 1.14).[70] In addition to being more lipophilic, these aromatic peptide nucleic acids (APNAs) were also designed in order to investigate possible π-stacking or dipole-quadrupole interactions along the backbone of the APNA oligomers and the possible stabilizing effects on a duplex or triplex structure. Although it is widely believed that the major stabilizing forces of double and triple helices are the π-stacking interactions between adjacent base pairs, the study of such interactions within the backbone of oligonucleotide analogues is completely unprecedented and, therefore, may be of considerable interest. More recently, another analogue of APNAs was designed with the following requirements in mind: a) the basic scaffold would be modular so that simple and readily available components could be brought together in a convergent manner, making the process amenable to combinatorial chemistry and b) the synthetic scheme should be general enough to permit the incorporation of all five natural nucleic acid bases into the oligomers. The general structure **42** seemed attractive as an initial target since it fulfilled these requirements. It was thought that the C-terminal fragment could be derived from any of the 21 naturally occurring amino acids, as well as any amino acid moiety available through the plethora of synthetic protocols developed for the synthesis of non-natural amino acids.[71] It was reasoned that this approach allows for the incorporation of a wide range of functionality and chirality into the backbone of the APNA oligomers. In addition, the N-terminal fragment was designed to allow for the incorporation of substituted or unsubstituted aromatic and heteroaromatic ring systems. Substituents (Y) at the A, B, C or D positions

of the ring could be polar or charged groups to aid water solubility and induce a dipole within the ring system, enhancing π-stacking interactions. Substituents could also be used to induce an extended π-system for better face-to-face overlap (π-stacking) of adjacent rings. Finally, the attachment of the nucleobase via the acetate linker can be carried out in the same way as with PNA. This also allows the incorporation of non-natural bases such as those described in Section 1.5, as well as new bases designed to expand the genomic alphabet.[72]

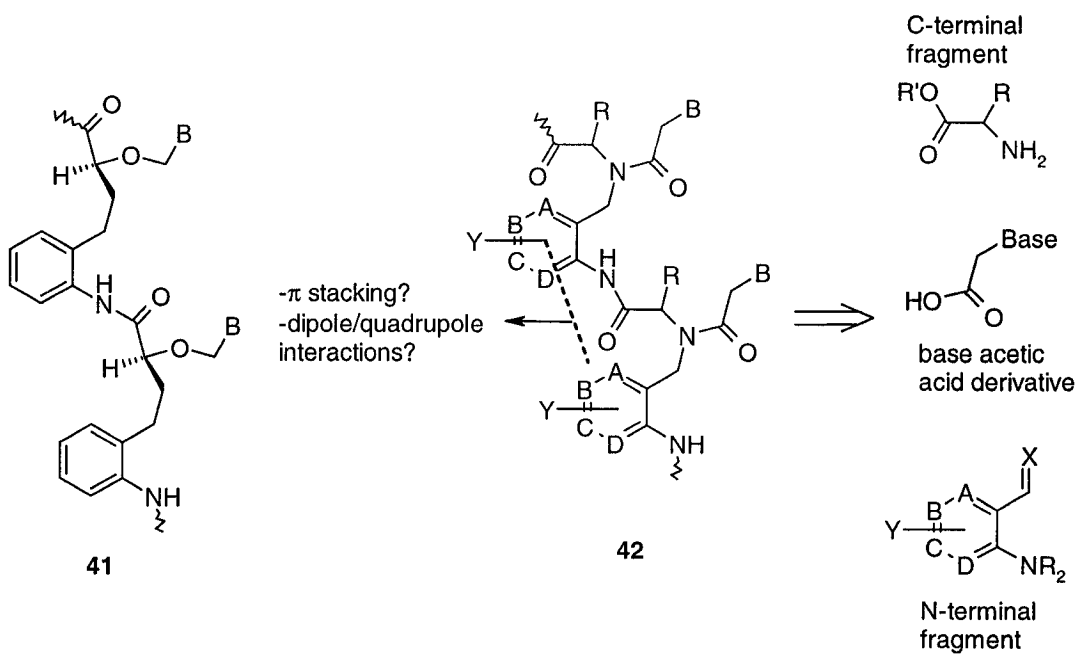


Figure 1.14: Design of Aromatic Peptide Nucleic Acids.

1.11 - Study of Modified Oligonucleotides.

A number of different experimental techniques have been used to explore the biological/biophysical properties of oligonucleotides. Usually, when a novel antisense agent is in development, it is important to determine if oligomers comprised of the new synthetic monomers can form stable complexes with RNA and to evaluate if they do so with high sequence discrimination. By far, the most exploited method for addressing these questions is the thermal melt (T_m) technique. In this method, some observable property of the complex is measured as a function of temperature (Figure 1.15). Commonly, UV absorbance is used since a hypochromic effect causes the molar extinction coefficient of a duplex or triplex to be less than the sum of the extinction coefficients of the constituent oligomers. Therefore, as the temperature rises and the complex dissociates the absorbance of the solution increases. The inflection point of the resulting sigmoidal curve is known as the "T_m value" and corresponds to the temperature at which half of the complexes in solution are dissociated. Comparison of T_m values between modified and unmodified complexes of the same sequence provides a means to assess the relative stability of a modified oligonucleotide.

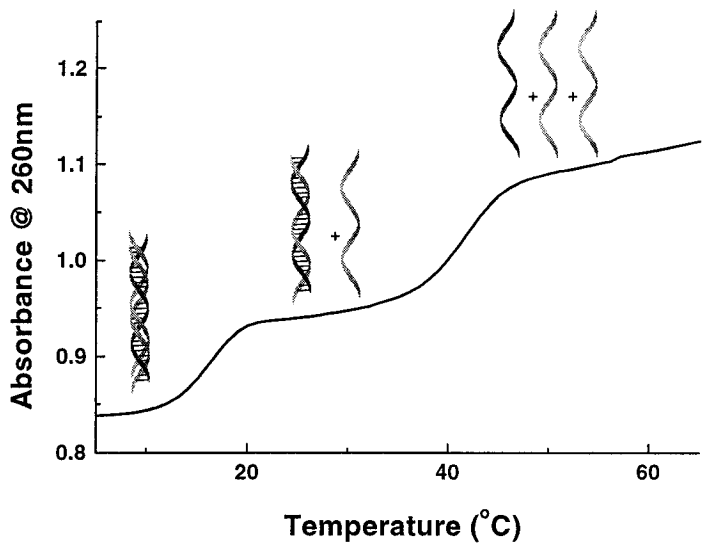


Figure1.15: Example of the thermal denaturation (T_m) curve of a triplex structure.

The chiral nature of duplexes and triplexes and the way in which the chromophoric bases are stacked on top of each other allows for circular dichroism (CD) to also be used as a tool for studying structural aspects of oligonucleotide complexes. For example, A-form and B-form helices give rise to diagnostic CD profiles allowing for differentiation of one type of helix from the other.[68] However, the CD signature of duplexes and triplexes is sequence dependent and so valuable comparisons can only be made between complexes with the same base sequence.[73] CD signatures can also be monitored as a function of temperature to create T_m curves since the CD profile of double or triple stranded complexes is typically very different from the sum of the CD signatures of the single strands.

Nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography are powerful techniques used to acquire high-resolution structural data of duplexes and triplexes involving natural and modified oligonucleotides. In particular, the information

gained through NMR experiments has become the modern standard for studying oligonucleotides that show great potential for clinical applications. These studies are particularly insightful because the three-dimensional structures that are solved are considered to be the most accurate picture of how the molecules behave in solution.[74]

1.12 – References.

[1] Karp, G. *Cell and Molecular Biology: Concepts and Experiments, 2nd* Ed., Karp, G., Ed., John Wiley and Sons, Inc., New York, NY, **1999**.

[2] Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D.; *Molecular Biology of the Cell*, 3rd Ed., Garland Publishing, Inc., New York, NY, **1994**.

[3] Stryer, L.; *Biochemistry*, 3rd Ed., W.H. Freeman and Company, New York, NY, **1988**.

[4] McPherson, J. D.; Marra, M.; Hillier, L.; Waterston, R. H.; Chinwalla, A.; Wallis, J.; Sekhon, M.; Wylie, K.; Mardis, E. R.; Wilson, R. K.; Fulton, R.; Kucaba, T. A.; Wagner-McPherson, C.; Barbazuk, W. B.; Humphray, S. J.; French, L.; Evans, R. S.; Bethel, G.; Whittaker, A.; Holden, J. L.; McCann, O. T.; Dunham, A.; Soderlund, C.; Scott, C. E.; Bentley, D. R.; Chen, H.-C.; Jang, W.; Green, E. D.; Idol, J. R.; Maduro, V. V. Braden; M., Kate T.; Lee, E.; Miller, A.; Emerling, S.; Kucherlapati, R.; Gibbs, R.; Scherer, S.; Gorrell, J. H.; Sodergren, E.; Clerc-Blankenbrg, K.; Tabor, P.; Naylor, S.; Garcia, D.; de Jong, P. J.; Catanese, J. J.; Nowak, N.; Osoegawa, K.; Qin, S.; Rowne, L.; Madan, A.; Dors, M.; Hood, L.; Trask, B.; Friedman, C.; Massa, H.; Cheung, V. G.; Kirsch, I. R.; Reid, T.; Yonescu, R.; Weissenbach, J.; Bruls, T.; Heilig, R.; Branscomb, E.; Olsen, A.; Doggett, N.; Cheng, J.-F.; Hawkins, T.; Myers, R. M.; Shang, J.; Ramirez, L.; Schmutz, J.; Velasquez, O.; Dixon, K.; Stone, N. E.; Cox, . R.; Haussler, D.; Kent, W. J.; Furey, T.; Rogic, S.; Kennedy, S.; Jones, S.; Rosenthal, A.; Wen, G.; Schihabel, M.; Gloeckner, G.; Nyakatura, G.; Siebert, R.; Schlegelberger, B.; Korenberg, J.; Chen, X.-N.; Fujiyama, A.; Hattori, M.; Toyoda, A.; Yada, T.; Park, H.-S.; Sakaki, Y.; Shimizu, N.; Asakawa, S.; Kawasaki, K.; Sasaki, T.; Shintani, A.; Shimizu, A.; Shibuya, K.; Kudoh, J.; Minoshima, S.; Ramser, J.; Seranski, P.; Hoff, C.; Poustka, A.; Reinhardt, R.; Lehrach, H. *Nature* **2001**, *409*, 934.

[5] a) Boguski, M. S. *Science* **1999**, *286*, 453; b) Service R. F. *Science* **2001**, *294*, 2079.

[6] Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.

[7] Created using the templates provided with the Hyperchem (v5.1) Molecular Modeling package.

[8] a) Hurley, L. H. *Nature Rev. Cancer* **2002**, *2*, 188.; b) Hurley, L. H. *Biochem. Soc. Trans.* **2001**, *29*, 692.; c) Mergny, J.-L.; Helene, C. *Nature Med.* **1998**, *4*, 1366.

[9] a) Xavier, K. A.; Eder, P. S.; Giordano, T. *Trends Biotechnol.* **2000**, *18*, 349.; b) Hermann, T.; Westhof, E. *Curr. Opin. Biotechnol.* **1998**, *9*, 66.; c) Pearson, N. D.; Prescott, C. D. *Chem. Biol.* **1997**, *4*, 409.

[10] a) Reid, D. G.; Salisbury, S. A.; Williams, D. H. *Biochem.* **1983**, *22*, 1377.; b) Brown, S. C.; Mullis, K.; Levenson, C.; Shafer, R. H. *Biochem.* **1984**, *23*, 403.

[11] a) Rinehart, K. L.; Holt, T. G.; Fregeau, N. L.; Stroh, J. G.; Keifer, P. A.; Sun, F.; Li, L. H.; Martin, D. G. *J. Org. Chem.* **1991**, *56*, 1676.; b) Pommier, Y.; Kohlhagen, G.; Bailly, C.; Waring, M.; Mazumder, A.; Kohn, K. W. *Biochem.* **1996**, *35*, 13303.; c) Moore, B. M., II; Seaman, F. C.; Hurley, L. H. *J. Am. Chem. Soc.* **1997**, *119*, 5475.; d) Villalona-Calero, M. A.; Eckhardt, S. G.; Weiss, G.; Hidalgo, M.; Beijnen, J. H.; Van Kesteren, C.; Rosing, H.; Campbell, E.; Kraynak, M.; Lopez-Lazaro, L.; Guzman, C.; Von Hoff, D. D.; Jimeno, J.; Rowinsky, E. K. *Clin. Cancer Res.* **2002**, *8*, 75.

[12] Bailly, C.; Chaires, J. B. *Bioconjugate Chem.* **1998**, *9*, 513.

[13] Goldberg, I. H. *Acc. Chem. Res.* **1991**, *24*, 191.

[14] Gewirtz, D. A. *Biochem. Pharmacol.* **1999**, *57*, 727.

[15] a) Dunn, C. J.; Goa, K. L. *Drugs Aging* **1996**, *9*, 122.; b) Faulds, D.; Balfour, J. A.; Chrisp, P.; Langtry, H. D. *Drugs* **1991**, *41*, 400.

[16] Li, L. H.; Krueger W. C. *Pharmacol. Ther.* **1991**, *51*, 239.

[17] Greider, C. W.; Blackburn, E. H. *Cell* **1987**, *51*, 887.

[18] a) Elion, G. B. *Adv. Enzyme Regul.* **1985**, *24*, 323.; b) Rando, R. F.; Nguyen-Ba, N. *Drug Discovery Today* **2000**, *5*, 465.

[19] Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543.

[20] For examples see: a) Bigey, P.; Knox, J.D.; Croteau, S.; Bhattacharya, S. K.; Theberge, J.; Szyf, M. *J. Biol. Chem.* **1999**, *274*, 4594.; b) Bridonneau, P.; Chang, Y.-F.; O'Connell, D.; Gill, S. C.; Snyder, D. W.; Johnson, L.; Goodson Jr., T.; Herron, D. K.; Parma, D. H. *J. Med. Chem.* **1998**, *41*, 778.; c) Jing, N.; De Clercq, E.; Rando, R. F.; Pallansch, L.; Lackman-Smith, C.; Lee, S.; Hogan, M. E. *J. Biol. Chem.* **2000**, *275*, 3421.

[21] For example as tools for *in situ* hybridization: Werner, M.; Wilkens, L.; Aubele, M.; Nolte, M.; Zitzelsberger, H., Komminoth, P. *Histochem. Cell Biol.* **1997**, *108*, 381.; or as biosensors: Pividori, M. I.; Merkoci, A.; Alegret, S. *Biosens. Bioelectron.* **2000**, *15*,.

[22] Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 280.

[23] Glukhov, A. I.; Zimnik, O. V.; Gordeev, S. A.; Severin, S. E. *Biochem. Biophys. Res. Commun.* **1998**, *248*, 368.

[24] Helene, C. *Anti-Cancer Drug Des.* **1991**, *6*, 569.

[25] Egholm, M.; Christensen, L.; Dueholm, K. L.; Buchardt, O.; Coull, J.; Nielsen, P. E. *Nucleic Acids Res.* **1995**, *23*, 217.

[26] a) Coman, D.; Russu, I. M. *Biochem.* **2002**, *41*, 4407-4414; b) Lin, K.-Y.; Matteucci, M. D. *J. Am. Chem. Soc.* **1998**, *120*, 8531.

[27] For a recent reviews see: a) Verma, S.; Eckstein, F. *Ann. Rev. Biochem.* 1998, *67*, 99; b) Mangos, M. M.; Damha, M. J. *Curr. Top. Med. Chem.* **2002**, *2*, 1147.

[28] For recent examples see: Vastmans, K.; Rozenski, J.; Van, A.; Herdewijn, P. *Biochim. Biophys. Acta* **2002**, *1597*, 115.; b) Trempe, J.-F.; Wilds, C. J.; Denisov, A. Y.; Pon, R. T.; Damha, M. J.; Gehring, K. *J. Am. Chem. Soc.* **2001**, *123*, 4896.

[29] a) Wu, X.; Guntha, S.; Ferencic, M.; Krishnamurthy, R.; Eschenmoser, A. *Org. Lett.* **2002**, *4*, 1279; Eschenmoser, A. *Science* **1999**, *284*, 2118.

[30] For a recent review see: Luyten, I.; Herdewijn, P. *Eur. J. Med. Chem.* **1998**, *33*, 515.

[31] For a recent review see: Micklefield, J. *Curr. Med. Chem.* **2001**, *8*, 1157.

[32] De Clercq, E.; Eckstein, F.; Merigan, T. C. *Science* **1969**, *165*,1137.

[33] For reviews see: a) Agrawal, S.; Iyer, R. P. *Curr. Opin. Biotechnol.* **1995**, *6*,12.; b) see also refs 19 and 25.

[34] Kibler-Herzog, L.; Zon, G.; Uznanski, B.; Whittier, G.; Wilson, W. *Nucleic Acids Res.* **1991**, *19*, 2979.

[35] Stein C. *Nature Med.* **1995**, *1*, 1119.

[36] Dove, A. *Nature Biotechnol.* **2002**, *20*, 121.

[37] Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.

[38] Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767.

[39] Merrifield, R.B. *J. Am. Chem. Soc.* **1963**, *85*, 2149.

[40] a) Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P.E.; Coull, J.; Berg, R. H. *J. Pept. Sci.* **1995**, *3*, 175.; b) Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielson, P. E. *J. Pept. Sci.* **1995**, *1*, 185.; c) Koch, T.;

Hansen, H. F.; Andersen, P.; Larsen, T.; Batz, H. G.; Otteson, K.; Oerum, H. *J. Pept. Res.* **1997**, *49*, 80.

[41] a) Viirre, R. D.; Hudson, R. H. E. *Org. Lett.* **2001**, *3*, 3931.; b) Richter, L. S.; Zuckermann, R. N. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1159.; c) Seitz, O.; Bergmann, F.; Heindl, D. *Angew. Chem. Int. Ed.* **1999**, *38*, 2203.

[42] Challa, H.; Woski, S. A. *Tet. Lett.* **1999**, *40*, 419.

[43] Wei, Y.; Marino, M.; Thompson, B.; Girard, J. E. *J. Chromatogr. A* **1999**, 49.

[44] a) Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624.; b) Nielsen, P. E. *Curr. Opin. Struc. Biol.* **1999**, *9*, 353.

[45] Hyrup, B.; Egholm, M.; Nielson, P. E.; Wittung, P.; Nordén, B; Buchardt, O. *J. Am. Chem. Soc.* **1994**, *116*, 7964.

[46] a) Leijon, A.; Graslund, A.; Nielsen, P. E.; Buchardt, O.; Norden, B.; Kristensen, M.; Eriksson, M. *Biochem.* **1994**, *33*, 9820.; b) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S.R. *Science* **1995**, *270*, 1838.

[47] Brown, S. C.; Thomson, S. A.; Veal, J. M.; Davis, D. G. *Science* **1994**, *265*, 777.

[48] Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E. *Nat. Struct. Biol.* **1997**, *4*, 98.

[49] a) Almarsson, O.; Bruice, T. C.; Kerr, J.; Zuckermann, R. N. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 7518.; b) Almarsson, O.; Bruice, T. C. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9542.

[50] a) Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org. Chem.* **1999**, 197.; b) Wittung, P., Nielsen, P.; Norden, B. *J. Am. Chem. Soc.* **1997**, *119*, 3189.; c)

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

[51] Demidov, V.; Potaman, V. N.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchard, O.; Sonnichsen, S. H.; Nielsen, P. E. *Biochem. Pharmacol.* **1994**, *48*, 1310.

[52] Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadila, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Bruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481.

[53] a) Bickel, U.; Yoshikawa, T.; Landaw, E. M.; Faull, K. F.; Pardrige, W. M. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 2618.; b) Pardridge, W. M.; Boado, R. J.; Kang, Y.-S. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5592.

[54] Hamilton, S. E.; Simmons, C. G.; Kathiriya, I. S.; Corey, D. R. *Chem. Biol.* **1999**, *6*, 343.

[55] Aldrian-Herrada, G; Desarmenien M. G.; Orcel, H.; Boissin-Agasse, L; Mery, J.; Brugidou, J.; Rabie, A. *Nucleic Acids Res.* **1998**, *26*, 4910.

[56] Tyler, B. M.; McCormick, D. J.; Hoshall, C. V.; Douglas, C. L.; Jansen, K.; Lacy, B. W.; Cusack, B; Richelson, E. *FEBS Lett.* **1998**, *421*, 280.

[57] a) Good, L; Nielsen, P. E.; *Nature Biotechnol.* **1998**, *16*, 355.; b) Good, L.; Nielsen, P. E. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2073.

[58] Knudsen, H.; Nielsen, P. E. *Nucleic Acids Res.* **1996**, *24*, 494.

[59] Lansdorp, P. M.; Verwoerd, N. P.; van de Rijke, F. M.; Dragowska, V.; Little, M.-T.; Dirks, R. W.; Raap, A. K.; Tanke, H. J. *Hum. Mol. Genet.* **1996**, *5*, 685.

[60] Orum, H.; Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O.; Stanely, C. *Nucleic Acids Res.* **1993**, *21*, 5332.

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

[62] Griffin, T.; Tang, W.; Smith, L. M. *Nature Biotechnol.* **1997**, *15*, 1368.

[63] a) D'Costa, M.; Kumar, V.; Ganesh, K. N. *Org. Lett.* **2001**, *3*, 1269.; b) Vilaivan, T.; Suparpprom, C.; Harnyuttanakorn, P.; Lowe, G. *Tetrahedron. Lett.* **2001**, 42, 5533.; c) Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org. Chem.* **1999**, 197.

[64] Hyrup, B.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1083.

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

[66] a) Puschl, A.; Boesen, T.; Zuccarello, G.; Dahl, O.; Pitsch, S.; Nielsen, P. E. *J. Org. Chem.* **2001**, *66*, 707.; b) Pueschl, A.; Tedeschi, T.; Nielsen, P. E. *Org. Lett.* **2000**, *2*, 4161.

[67] a) Peyman, A.; Uhlmann, E.; Wagner, K.; Augustin, S.; Weiser, C.; Will, D. W.; Breipohl, G. *Angew. Chem. Int. Ed.* **1998**, *36*, 2809; b) Efimov, V. A.; Choob, M. V.; Buryakova, A. A.; Kalinkina, A. L.; Chakhmakhcheva, O. G. *Nucleic Acids Res.* **1998**, *26*, 566.

[68] a) Sforza, S.; Haaima, G.; Marchelli, R.; Nielsen, P. E. *Eur. J. Org. Chem.* **1999**, *1*, 97.; b) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed.* **1996**, *35*, 1939.; c) Dueholm, K. L.; Petersen, K. H.; Jensen, D. K.; Egholm, M.; Nielsen, P. E.; Buchardt, O. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1077.

[69] a) Kuwahara, M.; Arimitsu, M.; Shigeyasu, M.; Saeki, N.; Sisido, M. *J. Am. Chem. Soc.* **2001**, *123*, 4653.; b) Kuwahara, M.; Arimitsu, M.; Sisido, M. *J. Am. Chem. Soc.* **1999**, *121*, 256.

[70] Tsantrizos, Y. S.; Lunetta, J. F.; Boyd, M.; Fader, L. D.; Wilson, M.-C. *J. Org. Chem.* **1997**, *62*, 5451.

[71] a) O'Donnell, M. J. *Aldrichimica Acta* **2001**, *34*, 3.; b) Cativiela, C.; Diaz-De-Villegas, M. D. *Tetrahedron: Asymm.* **2000**, *11*, 645.; c) Calmes, M.; Daunis, J. *Amino Acids* **1999**, *16*, 215.; d) Burk, M. J. *Acc. Chem. Res.* **2000**, *33*, 363.; e) Fehr, C. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2567.; f) Ohfune, Y. *Acc. Chem. Res.* **1992**, *25*, 360.

[72] For examples see: a) Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. *Nature Biotechnol.* **2002**, *20*, 177.; b) Tae, E. L.; Wu, Y.; Xia, G.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2001**, *123*, 7439.

[73] a) Gray, D. M.; Ratliff, R. L. *Bioploymers* **1975**, *14*, 487.; b) Hung, S.-H.; Yu, Q.; Gray, D. M.; Ratliff, R. L. *Nucleic Acids Res.* **1994**, *22*, 4326.; c) Ratmeyer, L.; Vinayak, R.; Zhong, Y. Y.; Zon, G.; Wilson, W. D. *Biochem.* **1994**, *33*, 5298.; d) Lesnik, E. A.; Freier, S. M. *Biochem.* **1995**, *34*, 10807.; e) Gyi, J. I.; Conn, G. L.; Lane, A. N.; Brown, T. *Biochem.* **1996**, *35*, 12538.; f) Roberts, R. W.; Crothers, D. M. *Science* **1992**, *258*, 1463.

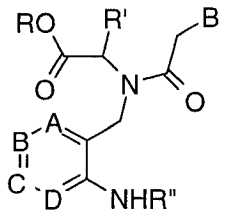
[74] Matthews, S. J. *Nucl. Magn. Reson.* **2002**, *31*, 312.

CHAPTER 2

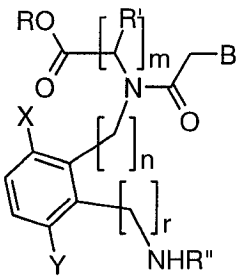
RESULTS AND DISCUSSION: DEVELOPMENT OF A NOVEL CLASS OF OLIGONUCLEOTIDE ANALOGUES; AROMATIC PEPTIDE NUCLEIC ACIDS (APNAs).

2.1 - Objectives.

Studies on APNA-type oligomers were initiated with analogues represented by the general structure **42** (A=B=C=D=CH). The effects of benzene ring moieties along the backbone of a PNA-like oligomer were evaluated by measuring the thermal stability of complexes formed between PNA-APNA hybrids and DNA or RNA. For these studies, an efficient synthetic route to APNA monomers that allowed for some structural diversity was developed. Monomers **43-46** were designed to modulate the relative position of the backbone aromatic ring, as well as the nucleobase, in order to identify a promising lead structure for further optimization. Substituted analogues of **43** were also synthesized in order to examine the effect of chirality within the backbone and to introduce functionality to the basic APNA scaffold (**47-50**).



42

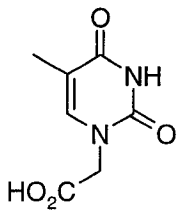


	m	n	r	X	Y	R'
43	1	1	0	H	H	H
44	2	1	0	H	H	H
45	1	0	1	H	H	H
46	2	0	0	H	H	H
47	1	1	0	H	H	Me
48	1	1	0	CH_2OH	H	H
49	1	1	0	CH_2NMe_2	H	H
50	1	1	0	H	CO_2CH_3	H

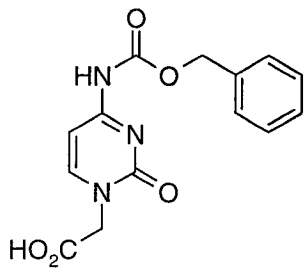
2.2 - Synthesis of Thymine, Cytosine, Adenine and Guanine Acetic Acid Derivatives.

As alluded to in Chapter 1 (Fig. 1.14), the design of the APNA monomers allowed for the nucleobases to be attached to the APNA backbone using the known acetic acid derivatives **51-54**.[75,76,77] Compounds **51-54** were prepared following slightly

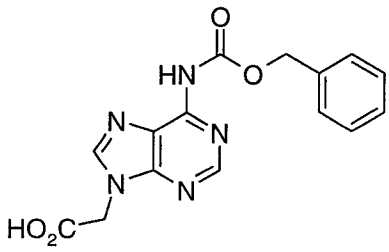
modified literature procedures; a number of previously undocumented observations were made which will be discussed below.



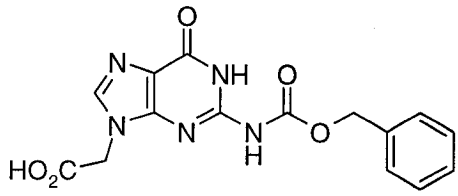
51



52



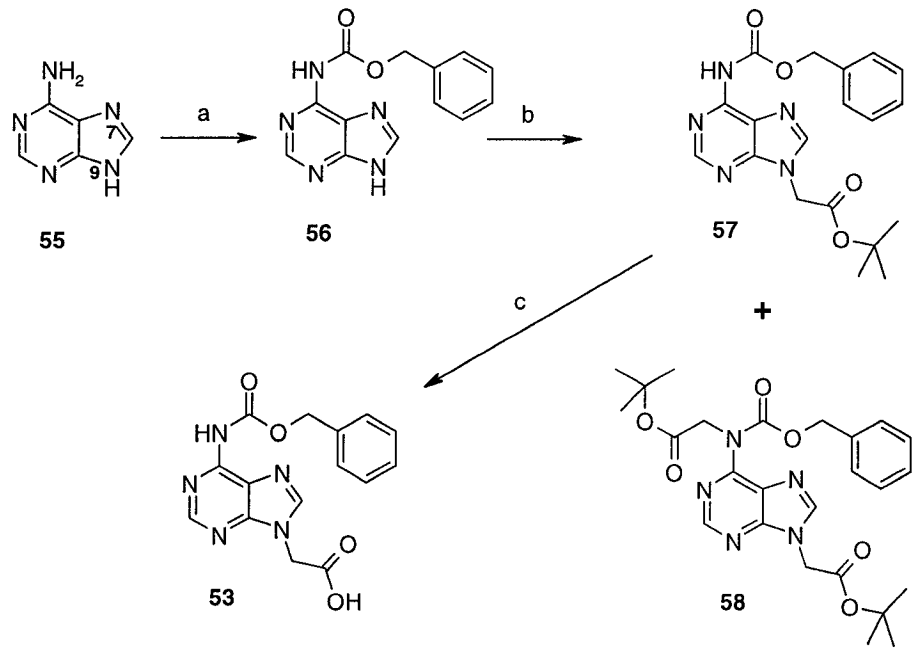
53



54

Preparation of derivative **53** began with Cbz-protection of the exocyclic amino group of adenine (**55**) to give carbamate **56** in high yields (Scheme 2.1).[78] Although alkylation at the N-9 position of adenine with *tert*-butyl bromoacetate under basic conditions was usually a reliable method for obtaining ester **57**, the dialkylated derivative **58** was often isolated as a minor product. After removal of this side product by column chromatography, TFA mediated conversion of ester **57** to adenine derivative **53** proceeded in high yield with no detectable loss of the Cbz protecting group.[79]

Scheme 2.1: Synthesis of Adenine Derivative **53**.

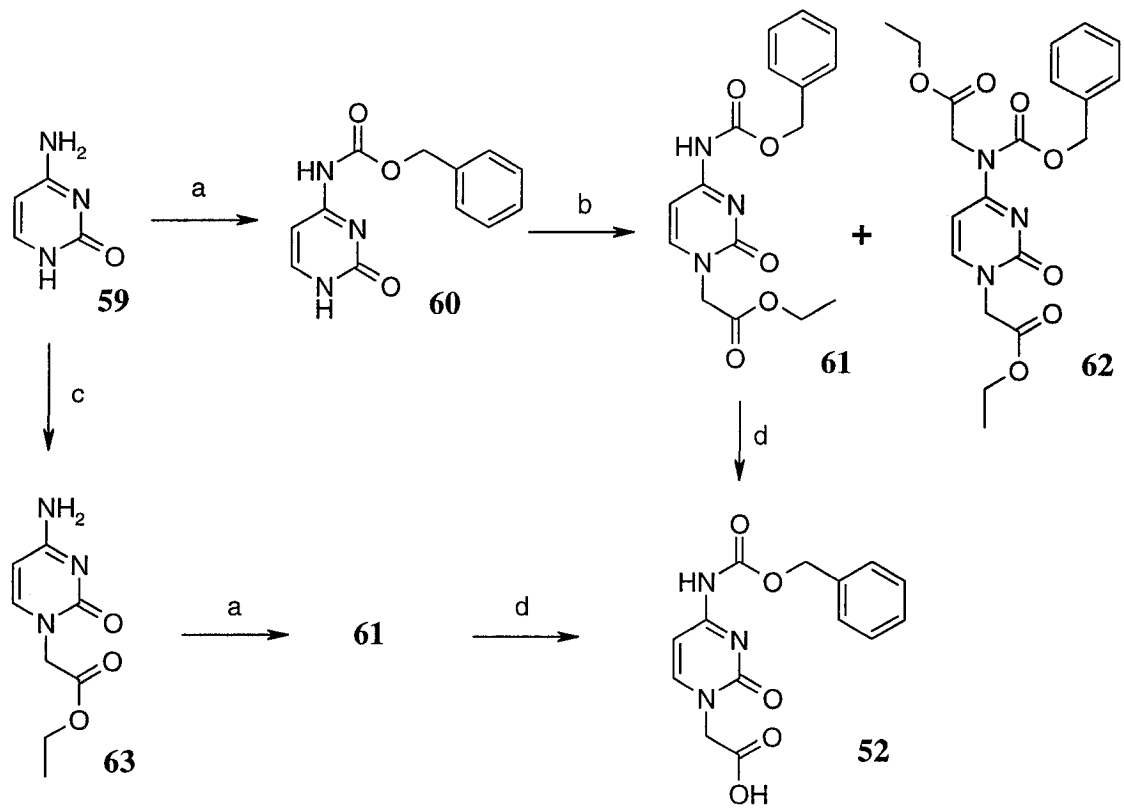


Conditions: a) NaH, Cbz-Cl, THF; b) $BrCH_2CO_2tBu$, K_2CO_3, Cs_2CO_3, DMF; c) TFA/DCM (2:1 v/v).

Two previously reported methods for the preparation of the cytosine derivative **52** were explored and compared (Scheme 2.2).[76,80] In the first route, the exocyclic amino group of cytosine (**59**) was Cbz-protected under standard conditions giving carbamate **60** in modest yields. While alkylation at the N-1 position of cytosine derivative **60** has been reported to proceed in modest yield, this reaction was somewhat unreliable with dialkylated derivative **62** being isolated in as much as a 70% yield. In an alternative approach, the sequence was reversed with the N-1 position being selectively alkylated first to give ester **63** in high yield and excellent purity.[81] Ester **63** was then protected at the exocyclic amino position to give fully protected cytosine derivative **61** in high yield and purity. Overall, it was observed that the sequence of **59** -> **63** -> **61** leads to superior

yield and purity of the ethyl ester **61**. Finally, the ethyl ester was saponified to provide the free acid **51** in nearly quantitative yield.

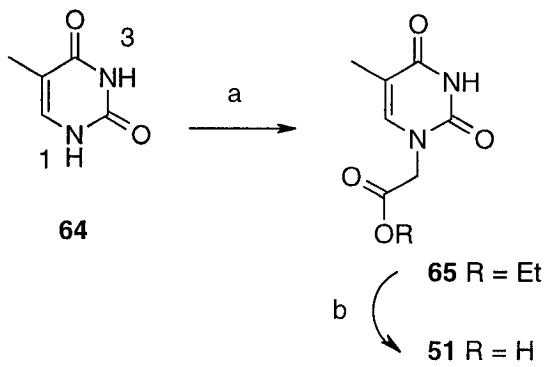
Scheme 2.2: Synthesis of Cytosine Derivative **52**.



Conditions: a) Cbz-Cl, pyr; b) $BrCH_2CO_2Et$, K_2CO_2, DMF; c) $BrCH_2CO_2Et$, KOtBu, DMF; d) $LiOH/H_2O$ then HCl.

The thymine and guanine acetic acid derivatives **51** (Schemes 2.3)[78] and **54** (Schemes 2.4)[82] were prepared following literature procedures without modification. Among the four bases, preparation of the thymine acetic acid derivative was the simplest, involving only alkylation at N1 with ethylbromoacetate, followed by saponification of the ethyl ester (Scheme 2.3).

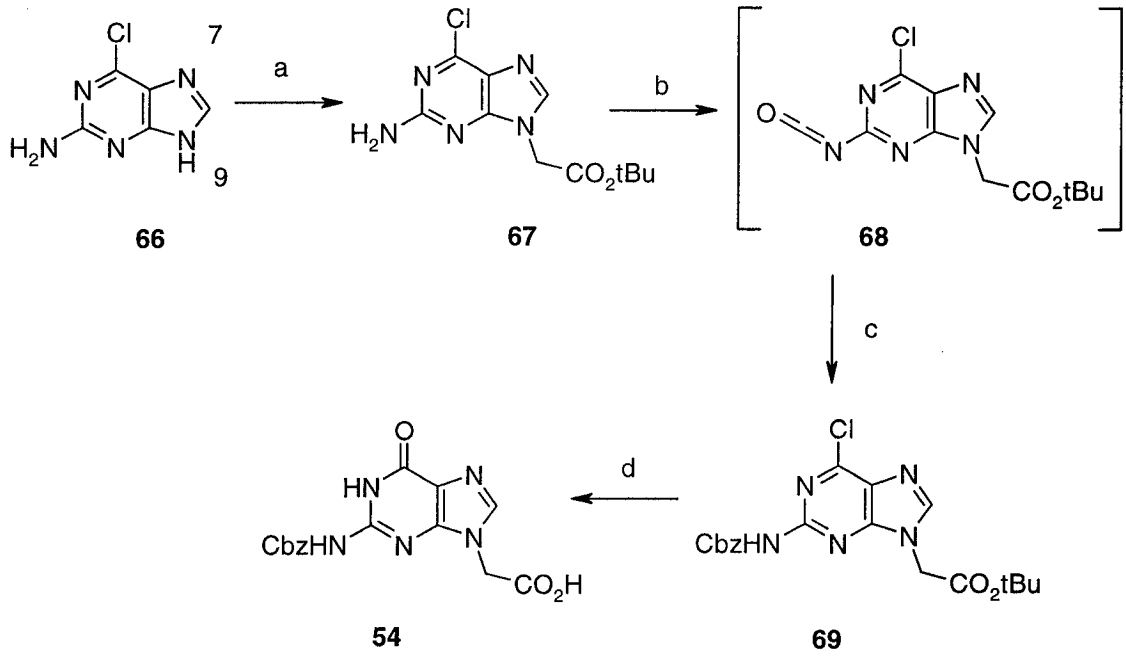
Scheme 2.3: Synthesis of Thymine Derivative **51**.



Conditions: a) $BrCH_2CO_2Et$, K_2CO_3, DMF; b) NaOH, then HCl.

In contrast to the thymine derivative **51,** the guanine acetic acid analogue is usually the most difficult to prepare. Several synthetic routes have already been published, each providing some advantages and some disadvantages.[78,82] In our hands, the method developed by Coull and Hodge which involves alkylation of 6-chloro-2-aminopurine (**66**) at N9 with *tert*-butyl bromoacetate followed by introduction of the Cbz group on the exocyclic amine, presumably through isocyanate **68** (Scheme 2.4) proved to be the most efficient in providing the key intermediate compound **69**.[82] The *tert*-butyl group of **69** can be easily removed and simultaneously the heterocycle converted to the guanine nucleus of compound **54** (Scheme 2.4).

Scheme 2.4: Synthesis of Guanine Derivative **54**.



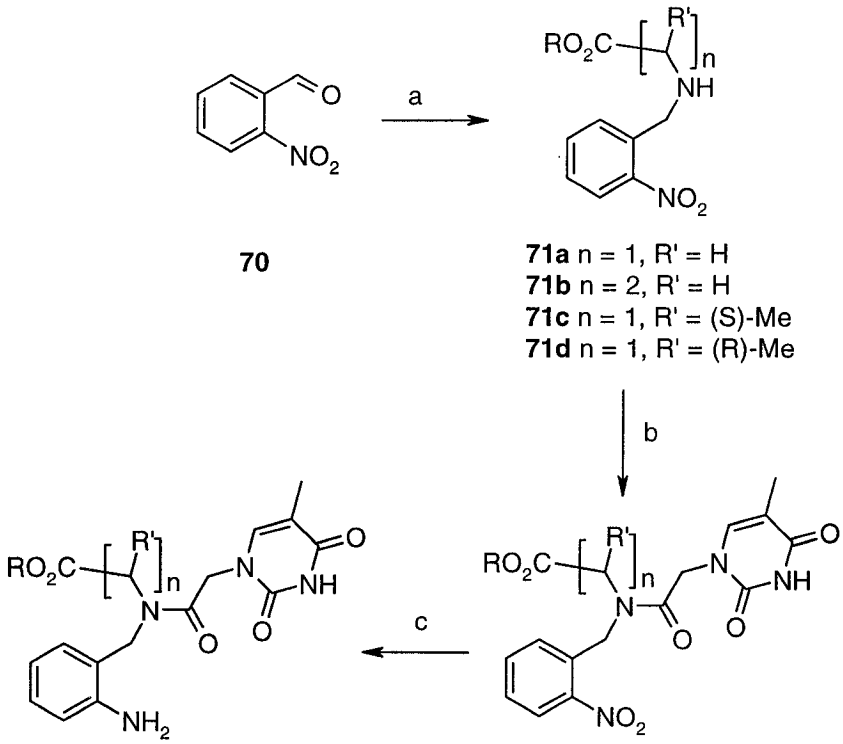
Conditions: a) $BrCH_2CO_2tBu$, K_2CO_3, DMF; b) $(CCl_3O)_2CO$, DIPEA, THF, c) BnOH/THF; d) $NCCH_2CH_2OH$, NaH, THF, then HCl/H_2O.

2.3 - Synthesis of APNA Monomers 43, 44, (S)-47 and (R)-47.

The synthesis of APNA monomers **43, 44, (S)-47** and **(R)-47** is outlined in Scheme 2.5. Reductive alkylation of each corresponding amino ester with commercially available aldehyde **70** in the presence of sodium triacetoxyborohydride gave the required secondary amine **71a-d** (Scheme 2.5).[83] Following work-up, each of the amines **71a-d** was precipitated from the organic solution (Et_2O/EtOAc) of the crude material as the corresponding HCl salt in good yield and high purity; purity of sample confirmed by ^1H NMR to be >97% of the desired product. Amines **71a-d** were then coupled to the thymine derivative **51**, using O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium

hexafluoro-phosphate (HATU)[84] as the coupling reagent, to give the desired tertiary amides **72a-d** in good to excellent yields. Following workup, the pure thymine derivatives **72a-d** were isolated after trituration with ethyl acetate/hexane. The subsequent catalytic hydrogenation of the nitro moiety was carried out using palladium on charcoal, either under a hydrogen atmosphere (1-5 atm) or with triethylammonium formate as the hydrogen source, leading to the isolation of the aniline intermediates **43**, **44** and **47**[85] in nearly quantitative yields and excellent purity.

Scheme 2.5: Synthesis of APNA Monomers **43, 44 (S)-47** and **(R)-47**.



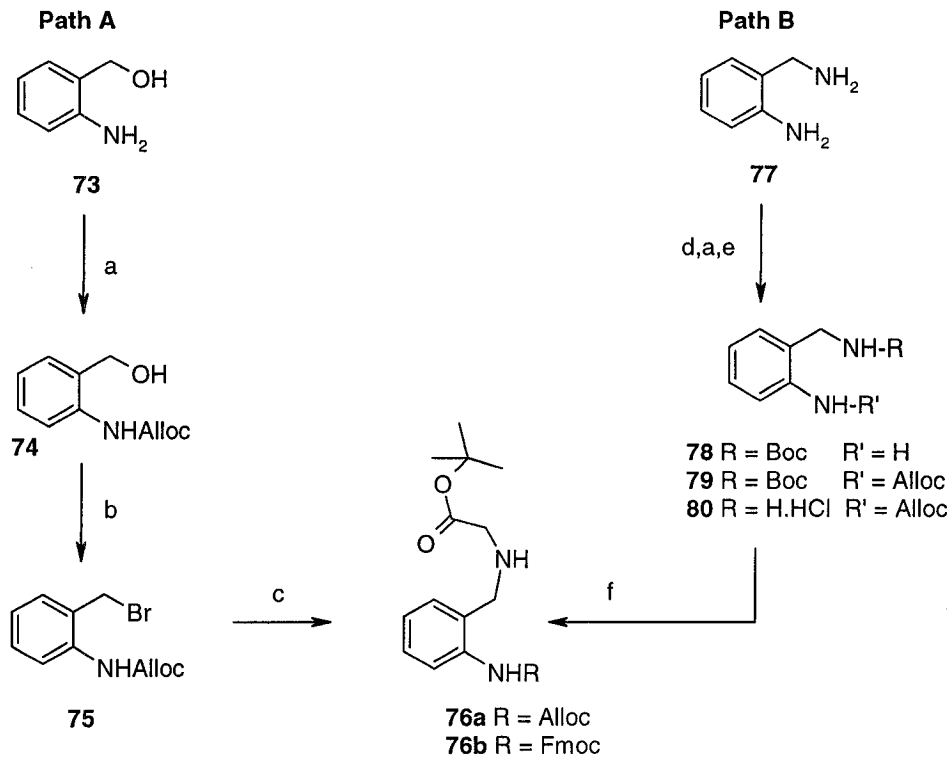
	n	R	R'	R"
43a	1	Me	H	H
44a	2	Et	H	H
(R)-47a	1	Me	(R)-Me	H
(S)-47a	1	Me	(S)-Me	H

Conditions: a) HCl.H-X-Ome, X = Glycine, (R)- or (S)-Alanine or β-Alanine, $NaBH(OAc)_3$, Et_3N, AcOH, DCM, 65-80%; b) Th-CH_2-CO_2H (**51**), HATU, DIPEA, DMF, 80-85%; c) HCO_2H, Et_3N, Pd-C, DMF, >95%.

2.4 - Synthesis of N-Fmoc Protected APNA Derivatives of All Four DNA Bases.

Preparation of mixed base sequences of APNA oligomers often required the use of Fmoc-protected *tert*-butyl esters of the APNA monomers containing each of the four DNA bases. Thus the approach of Scheme 2.5 was modified in order to accommodate Cbz-protected bases and *tert*-butyl esters. In one approach, aminobenzyl alcohol **73** was first converted to the allyl carbamate derivative **74**,[86] and the hydroxy moiety was subsequently converted to the bromide to obtain compound **75** (Scheme 2.6, Path A). Bromide **75** was converted to secondary amine **76a** in the presence of excess *tert*-butyl ester of glycine hydrochloride in order to avoid dialkylation of the amine. Although this synthetic route is very efficient, it requires the use of the relatively expensive glycine hydrochloride *tert*-butyl ester (5g, 178.90 CDN, Aldrich Chemical Company, 2000-2001). Therefore, an alternative synthetic route was optimized utilizing only commodity chemicals (Scheme 2.6, Path B). Commercially available 2-aminobenzylamine **77** was selectively protected by sequential treatment with Boc_2O giving aniline **78**, Alloc-Cl giving dicarbamate **79** and finally HCl (~4M dioxane) to provide derivative **80**. The free benzylamine hydrochloride salt **80** was alkylated with *tert*-butyl 2-bromoacetate in the presence of DIPEA to reproducibly give the fully protected backbone **76a** in ~70% yield with less than 5% of the dialkylated side product. Using this protocol, the cost of producing **76a** was calculated to be roughly $6.00 CDN/g, compared to $64.00 CDN/g using Path A.

Scheme 2.6: Synthesis of Secondary Amine **82**.

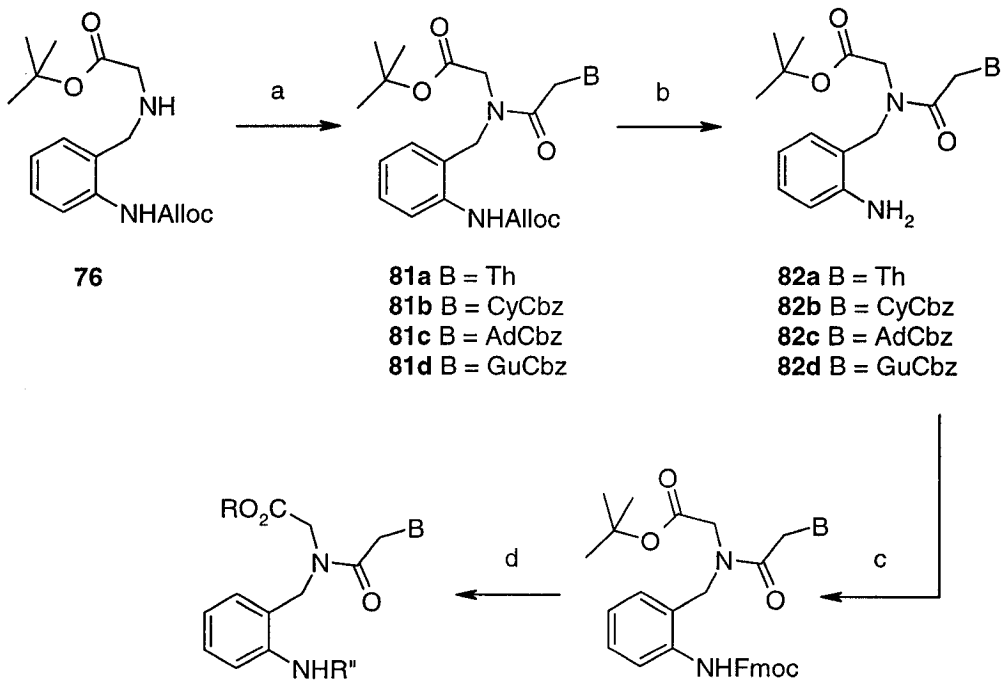


Conditions: a) AllocCl, pyridine, DCM; b) CBr_4, PPh_3, THF; c) $HCl{\cdot}H_2NCH_2CO_2R$, DIPEA, DMF; d) Boc_2O, THF; e) HCl, 4M in 1,4-dioxane, 88%, three steps; f) $BrCH_2CO_2R$, DIPEA, DMF, 70%.

Subsequently, the thymine, and Cbz-protected adenine, cytosine and guanine acetic acid derivatives were attached to the backbone in the presence of HATU/DIPEA to give the fully protected monomers **81a-d** in excellent yields (Scheme 2.7). The Alloc protecting groups were removed by treatment with $Pd(Ph_3P)_4$ in DCM/MeOH using benzenesulfinic acid as the allyl scavenger to give anilines **82a-d**.[87] Although the use of stoichiometric amounts of tributyl tin hydride as the allyl scavenger has been widely reported in the literature,[88] it was discovered that this method leads to side products resulting from N-allylation of the nucleobases (i.e. compound **84**), which were formed in as much as a 40% yield. This side reaction could only be suppressed if excess Bu_3SnH was employed, however, due to the highly toxic nature of this reagent, benzene sulfinic

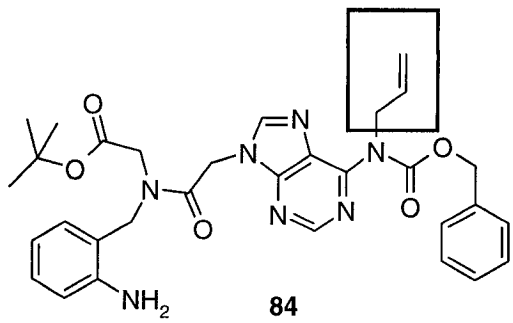
acid is a better alternative. Anilines **82a-d** were reprotected with Fmoc-Cl in the presence of aqueous Na_2CO_3 giving derivatives **83a-d**, which were treated with TFA/DCM to give the free acids **43a-d** (Scheme 2.7). It should be noted that the deprotection, reprotection strategy used to install the Fmoc group of compound **83a-d** was necessary because backbone **76b** was unstable in solution as its free amine due to rapid loss of the Fmoc group.

Scheme 2.7: Synthesis of Fmoc Protected Free Acid APNA Monomers **43a-d**.



	B	R	R"
43a	Th	H	Fmoc
43b	Cy^{Cbz}	H	Fmoc
43c	Ad^{Cbz}	H	Fmoc
43d	Gu^{Cbz}	H	Fmoc

Conditions: a) B-CH_2CO_2H, HATU, DIPEA, DMF: B = Th, 82%, B = Cy^{Cbz}, 97%, B = Ad^{Cbz}, 92%, B = Gu^{Cbz}, 76%; b) $Pd[PPh_3]_4$, DCM, $C_6H_6SO_2H$, MeOH: B = Th, 85%, B = Cy^{Cbz}, 85%, B = Ad^{Cbz}, 84%, B = Gu^{Cbz}, 86%; c) Fmoc-Cl, Na_2CO_3, 1,4-dioxane, H_2O: B = Th, 89%, B = Cy^{Cbz}, 83%, B = Ad^{Cbz}, 87%, B = Gu^{Cbz}, 81%; d) 1:1 TFA/DCM, B = Th, Cy^{Cbz}, Ad^{Cbz}, Gu^{Cbz}, >97%.



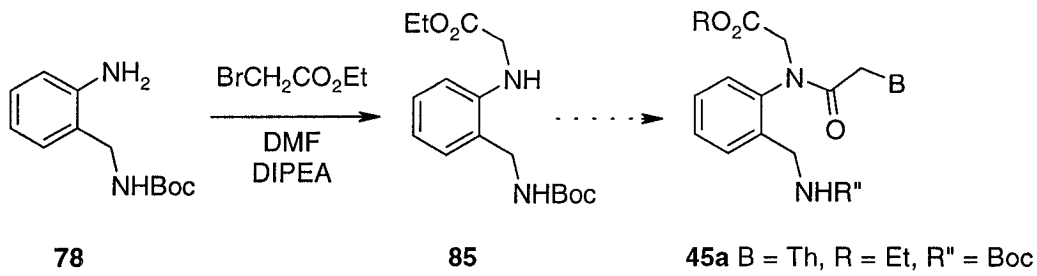
84

2.5 - Synthesis of APNA Monomers 45 and 46.

The synthesis of the APNA monomers **45** was first attempted following the route outlined in Scheme 2.8. Alkylation of the known mono-Boc protected aromatic diamine **78** with ethyl bromoacetate under basic conditions gave the secondary aniline **85** in good yield (Scheme 2.8). However, all subsequent attempts to acylate aniline **85** with thymine derivative **51** failed under a wide variety of peptide coupling conditions. Peptide coupling reactions involving secondary anilines are notoriously sluggish and in the case of compound **85** did not proceed at all under standard conditions. This result could possibly be due to an inductive effect of the adjoined ester, which would render the aniline far less nucleophilic. In an alternative approach, compound **78** was allylated with allyl bromide giving the masked glycine derivative **86** in good to excellent yield (Scheme 2.9). In contrast to the case of aniline **85**, the nucleic acid base derivatives **51** and **53** were easily attached to the backbone via EDC mediated amide formation to give amides **87a** and **87b**. Finally, the terminal olefins of **87a,b** were dihydroxylated to give diols **88a,b**, which were then oxidatively cleaved to the corresponding aldehydes **89a,b**.

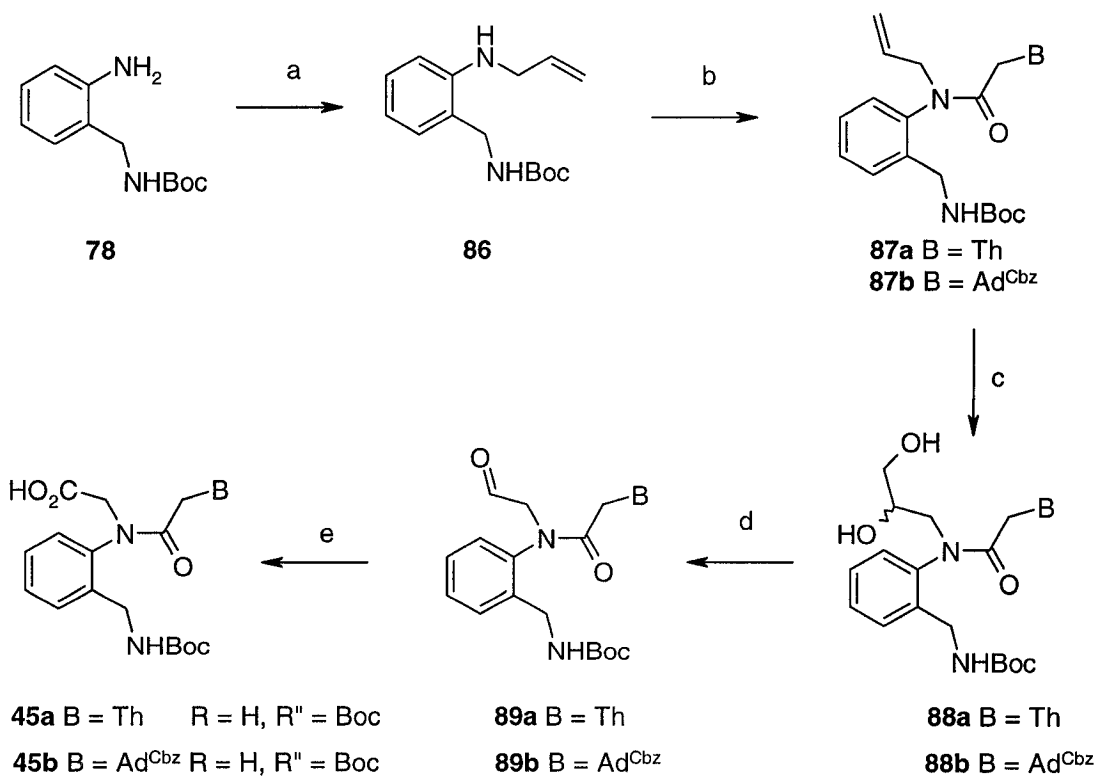
Aldehydes **89a,b** were in turn oxidized to their corresponding carboxylic acids **45a,b** (Scheme 2.9).

Scheme 2.8: Proposed Synthesis of Monomer **45a**.



78 **85** **45a** B = Th, R = Et, R" = Boc

Scheme 2.9: Synthesis of APNA Monomers **45a,b**.



78 **86** **87a** B = Th **87b** B = Ad^{Cbz}

88a B = Th **88b** B = Ad^{Cbz}

89a B = Th **89b** B = Ad^{Cbz}

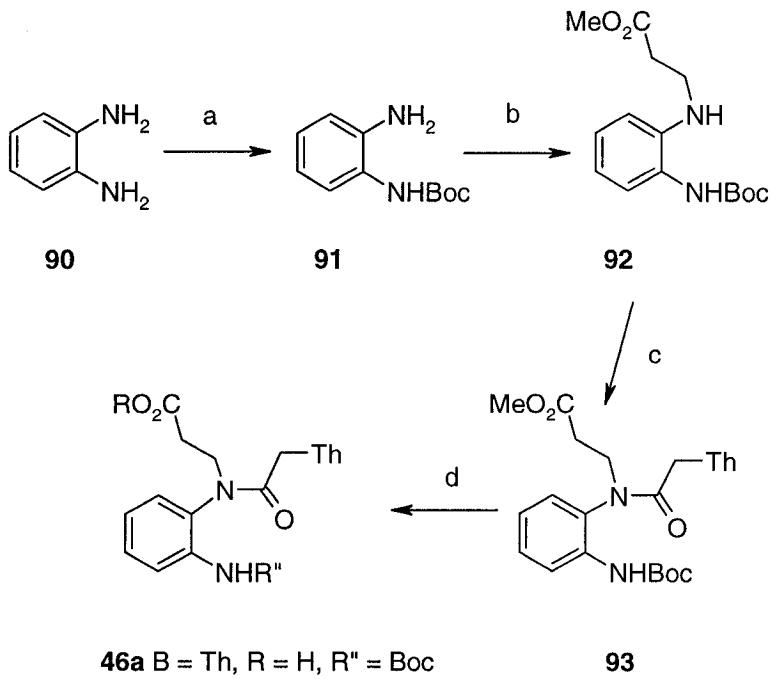
45a B = Th R = H, R" = Boc

45b B = Ad^{Cbz} R = H, R" = Boc

Conditions: a) $BrCH_2CHCH_2$, DIPEA, DMF 59%; b) Compound **51** or **53**, EDC, DMF 89%; c) OsO_4, NMO, THF/tBuOH/H_2O, 86%; d) $NaIO_4$, THF/H_2O, 98%; e) $NaClO_2$, 2-methyl-2-butene, NaH_2PO_4, 75-90%.

The synthesis of the monomers represented by structure **46** was carried out starting with diamine **90**, which was selectively mono-Boc protected to give aniline **91** in excellent yield (Scheme 2.10). Aniline **91** was alkylated with methyl 3-bromoproprionate under neutral conditions at high temperature giving secondary aniline **92**, in low yields. Compound **51** was then coupled to aniline **92** to give fully protected monomer **93** in modest overall yields. Interestingly, in the case of aniline **93** (the one carbon extended homologue of aniline **85**) the coupling reaction proceeded reasonably well, which is consistent with the apparent inductive effect observed with aniline **85**. Finally, the methyl ester of monomer **93** was saponified to yield the free acid monomer **46** in quantitative yield.

Scheme 2.10: Synthesis of APNA Monomer **46a**.

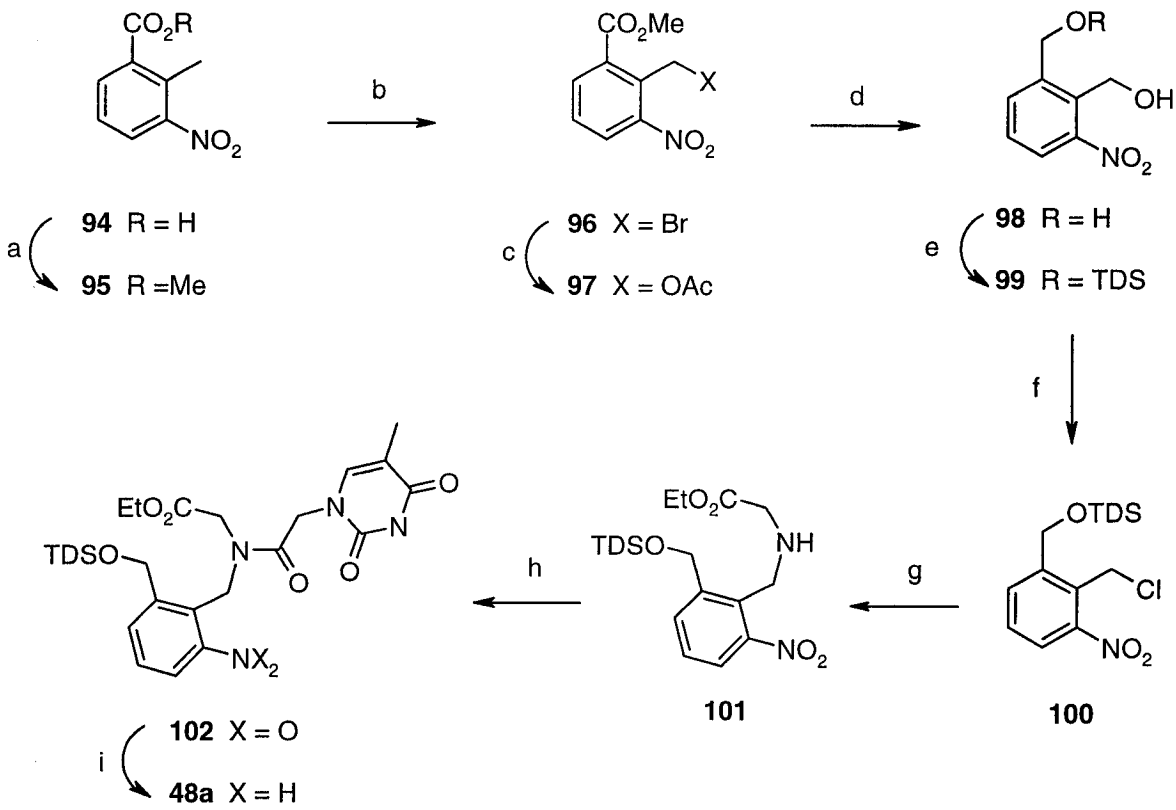


Conditions: a) Boc_2O, THF 65%; b) $BrCH_2CH_2CO_2Me$, DIPEA, DMF 36%; c) Compound **51**, EDC, DMF 49%; d) LiOH, THF/H_2O, 95%.

2.6 - Synthesis of APNA Monomers 48, 49 and 50.

The synthesis of monomer **48** began with a Fischer esterification of 2-methyl-3-nitrobenzoic acid (**94**) followed by free radical bromination of ester **95** to give bromide **96** (Scheme 2.11). Subsequent displacement of the bromide with sodium acetate, followed by $LiBH_4$ reduction of the diester **97** gave the diol **98** in good overall yield. Selective mono silylation of the diol was achieved using the procedure of Yu *et al* in moderate yield.[89] A series of nuclear Overhauser enhancement (nOe) NMR experiments unambiguously confirmed that the less hindered alcohol was protected during the silylation reaction (Scheme 2.11). In a somewhat unusual way, the benzyl alcohol **99** was then converted to the benzyl chloride **100** by allowing the alcohol **99** to react with mesyl chloride (MsCl) under basic conditions. It should be noted that the intermediate mesylate was observed and could be isolated if the reaction was stopped early, indicating that the overall process involves an *in situ* Finklestein reaction; this type of reaction is commonly observation with allylic and benzylic alcohols.[90] The chloride moiety of compound **100** was then displaced with glycine ethyl ester, under basic conditions, to give the backbone fragment **101**. Compound **101** was then coupled with the thymine derivative **51** to give the tertiary amide **102** in excellent yield. Finally, the nitro group was reduced to give the APNA monomer **48a** in an appropriately protected form for oligomer synthesis.

Scheme 2.11: Synthesis of APNA Monomers **48a**.

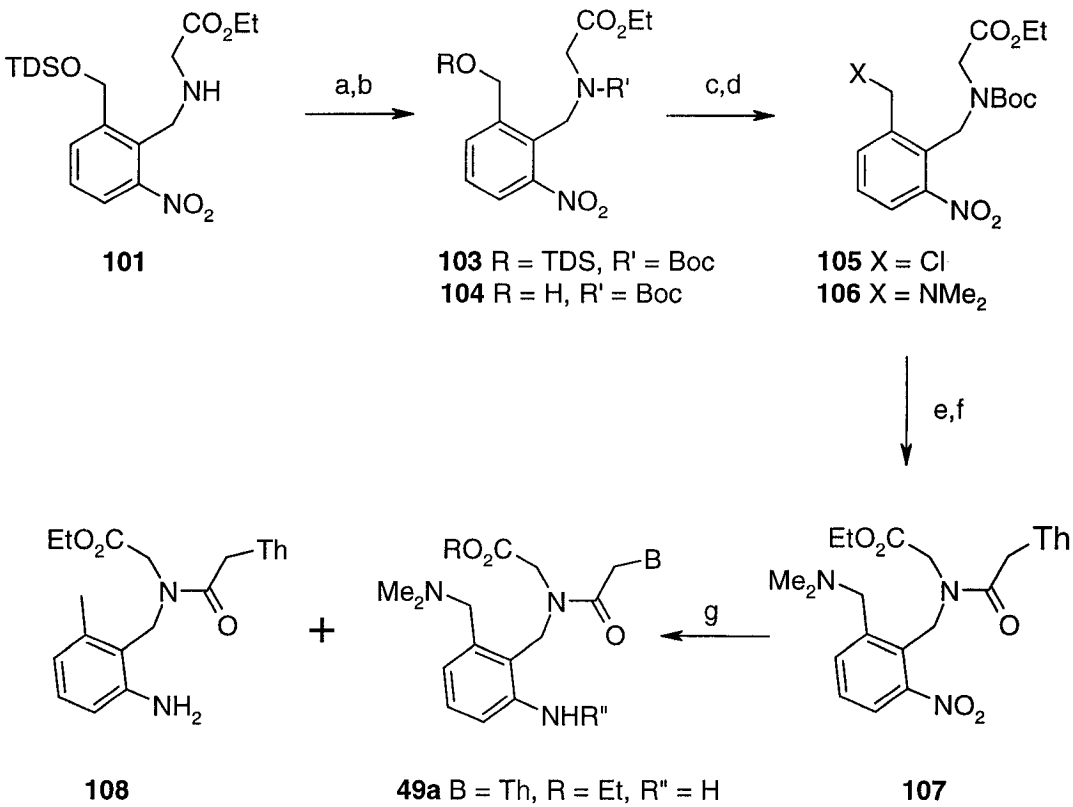


Conditions: a) MeOH, H_2SO_4, Δ, 89-100%; b) NBS, Bz_2O_2, CCl_4, Δ, 74-80%; c) NaOAc, DMF, >99%; d) $LiBH_4$, THF, 89-95%%; e) TDS-Cl, DIPEA/DMF 50-59%; f) MsCl, DIPEA, DCM; g) $ClH_3NCH_2CO_2Et$, DIPEA, DMF, 69-81% over two steps; h) Compound **51**, EDC, DMF, 94-99%; i) H_2, Pd-C, DMF, 63%.

Intermediate **101** was also used in the synthesis of monomer **49** (Scheme 2.12). Its secondary amine was Boc-protected giving carbamate **103** in excellent yield. Removal of the *tert*-butyldiphenylsilyl (TDS) group followed by treatment of the resulting alcohol **104** with MsCl and DIPEA gave chloride **105** in good yield. Displacement of the chloride with dimethylamine under basic conditions gave benzyl amine **106** in quantitative yield. The Boc protecting group was removed and the thymine derivative **51** was coupled to the resulting secondary amine to give intermediate **107** in excellent yield. The nitro group was then reduced to provide APNA monomer **49**.

However, during the reduction step, extreme care had to be taken in order to avoid hydrogenation of the benzylamine to the side product **108**. For example, the use of formic acid/triethyl amine as a hydrogen donor gave only toluene derivative **108**, whereas, using a low catalyst loading, under 1atm H_2, resulted in selective reduction of the nitro group.

Scheme 2.12: Synthesis of APNA Monomer **49**.

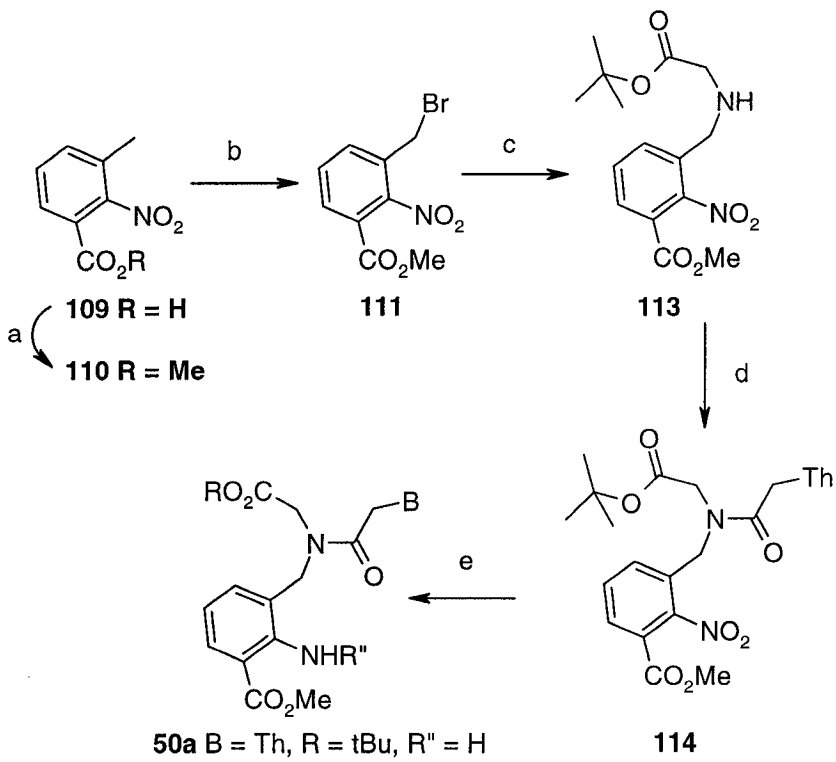


Conditions: a) Boc_2O, THF, 94%; b) TBAF, THF, 93%; c) MsCl, DIPEA, DCM, 88%; d) $Me_2NH.HCl$, K_2CO_3, DMF, 100%; e) HCl, dioxane; f) Compound **51**, EDC, DMF, 71% over two steps; g) H_2, Pd-C, DMF, 81%.

The synthesis of monomer **50** was achieved following a scheme similar to that described for compound **43** (Scheme 2.13). 3-Methyl-2-nitrobenzoic acid was converted to ester **110** under acidic conditions in methanol. The ester was then brominated to give

compound **111** in good yield. Bromide derivative **111** was used to alkylate amine **112** under basic conditions to give the backbone fragment **113** in quantitative yield. The thymine dervative **51** was then coupled to the backbone moiety and the nitro group of the resulting tertiary amide **114** was reduced to give the protected monomer **50a** in excellent overall yield.

Scheme 2.13: Synthesis of APNA Monomer **50**.



Conditions:) MeOH, H_2SO_4, Δ, 89-100%; b) NBS, Bz_2O_2, CCl_4, Δ, 74-80%; c) $H_2NCH_2CO_2t$-Bu (**112**), DIPEA, DMF, >98%; d) Compound **51**, EDC, DMF, 93%; e) HCO_2H, DIPEA, Pd-C, DMF, 89%.

2.7 - NMR Characterization of APNA Monomers.

Amide nitrogen atoms typically favor the sp^2 geometry and are coplanar with the carbonyl moiety so that the lone pair on the nitrogen atom is conjugated with C=O π-system. Furthermore, since the two substituents on the nitrogen atom of the tertiary amide intermediates in Schemes 2.5, 2.7 and 2.11-2.13 are sterically and electronically similar, the two amide rotamers [i.e. E(O)-**72a** and Z(O)-**72a**] are of similar energy. The free energy barriers for interconversion between amide rotamers of this type are typically in the range of 18-21 kcal/mol, which renders the rate of interconversion slow enough for individual rotamers to be observed on the NMR time scale. Thus, the ^1H and ^{13}C NMR spectra of intermediates such as **72a** displayed separate resonances for the E(O) and Z(O) isomers (Figure 2.1).

Variable temperature ^1H NMR was used to confirm that the two sets of signals arising from the tertiary amides in Schemes 2.5, 2.7 and 2.11-2.13 corresponded to the expected amide rotamers and *not* to two separate compounds. Upon heating a pure sample of compound **72a**, the pairs of ^1H NMR resonances observed at room temperature began to coalesce at temperatures over 90-100°C, giving rise to a single set of signals at approximately 140°C. Theoretically, the temperature at which two sets of resonances will coalesce is dependent on the difference in the proton chemical shifts of the rotamers and it is also related to the free energy barrier of rotation about the C-N bond; this energy can be calculated by equation 2.1:

$$\Delta G^{\neq} = (1.987)(T_{coal})(23.76 + \ln[T_{coal}/k_{coal}]), \qquad (2.1)$$

$$k_{coal} = (0.5)(\sqrt{2})\pi\Delta$$
$$\Delta = \delta_{rotamer\ a} - \delta_{rotamer\ b}$$

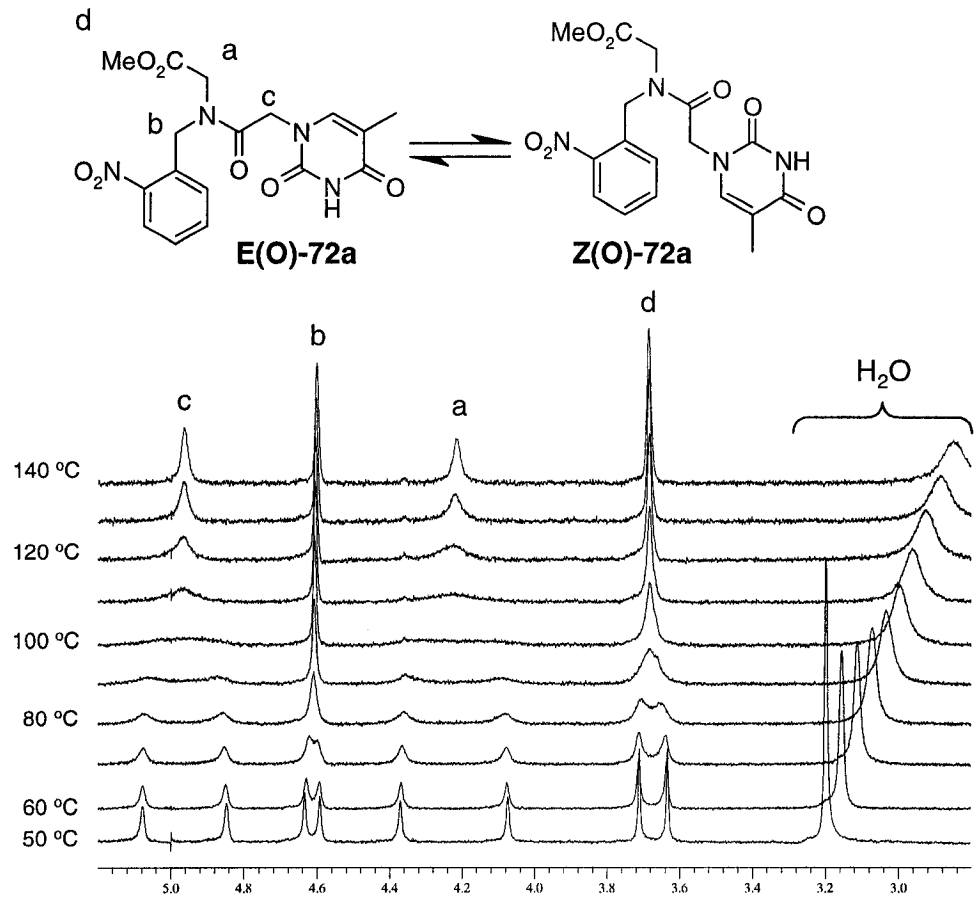


Figure 2.1: Temperature Dependence of the ^{1}H Methylene Resonances of Compound **72a**.

For example, based on the ^{1}H NMR data of compound **72a**, the barrier to interconversion between E(O)-**72a** and Z(O)-**72a** was calculated to be 18.9 kcal/mol, which is consistent with literature data.[91] Although it was also possible to record the ^{13}C spectra at 140°C for most compounds, the spectra were usually recorded at ambient temperature.[92] Similar results were obtained with all N,N-dialkyl tertiary amide intermediates presented in Schemes 2.5, 2.7 and 2.11-2.13.

2.8 - Conformational Analysis of the N-phenyl APNA Monomers

Monomers **45** and **46** are particularly interesting as PNA analogues since it is well known that *N*-phenyl-*N*-alkyl tertiary amides adopt the E(O) geometry exclusively, with the plane of the phenyl moiety oriented perpendicularly to the plane of the amide bond (Figure 2.2a).[93] Thus, if the R group of such an amide was to extend in the C-terminal direction of an APNA monomer and the R' group to a nucleobase [i.e. E(O)-**45**], then this amide conformation would be analogous to the more favourable Z(O) rotamer of the PNA monomers for binding to natural oligonucleotides (see Section 1.8, page 24). Unlike the previously discussed molecules (e.g. compound **72a**, Fig. 2.1), for compounds **45a,b** and **46a** the rotamer equilibrium is heavily shifted towards one isomer, since only one set of signals were observed for these compounds in both their ^{1}H and ^{13}C NMR spectra. Therefore, nOe experiments were conducted to confirm that these monomers exist predominately in the expected E(O) conformation. An unambiguous assignment of the preferred rotamer was obtained by studying the adenine derivative **87b**.[94] The key nOe correlations are shown in Figure 2.3. Compound **87b**, as well as all of the N-phenyl derivatives described in this chapter, are found in solution as a 1:1 mixture of atropisomers, thus H_a and H_b are diastereotopic and appear at very different chemical shifts in the ^{1}H NMR spectrum. Chemical shift assignments for both H_a and H_b were confirmed by the nOe correlations observed, including a strong nOe between both H_a, H_b and H_8 of the adenine base. Strong nOe was also observed between H_a and aromatic H_c proton but not between H_d and H_e, consistent with the desired E(O) rotamer conformation.

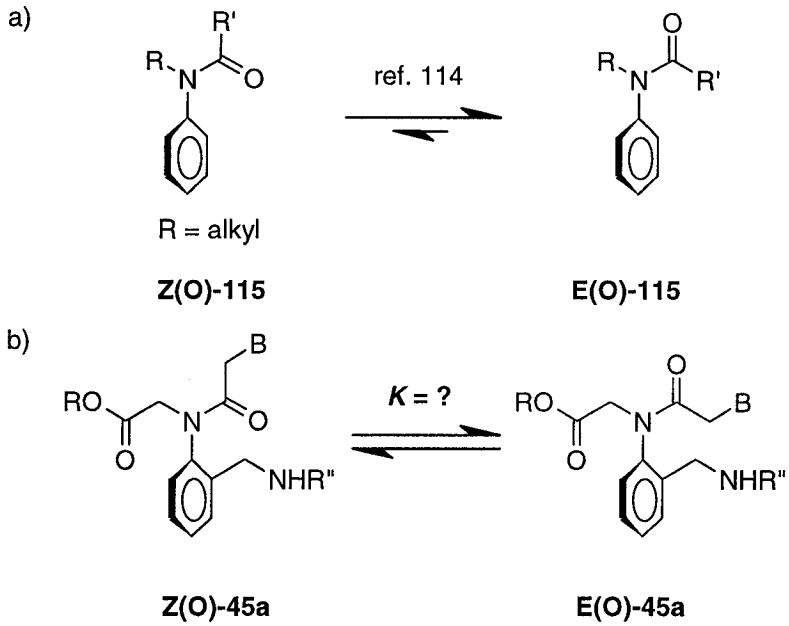


Figure 2.2: Amide Rotamers of Conformationally Preorganized APNA Monomers.

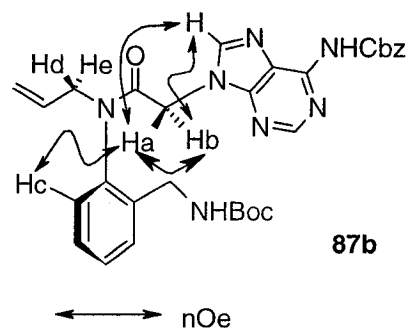


Figure 2.3: Key nOe Correlations for Compound **87b**.

2.9 - Estimation of the Barrier of Rotation for Atropisomers of APNA Monomers.

The acylation reaction of aniline **86**, to give the masked monomer precursors **87a,b** (Scheme 2.9), was done without control of the stereochemistry. Consequently, an obvious concern that needs to be addressed is what effect the incorporation of a racemic mixture of monomers **45** (**a** or **b**) into PNA oligomers would have on the hybridization properties of these oligomers with DNA and RNA. Furthermore, it is likely that one enantiomer may possess better DNA or RNA recognition properties than the other. In such a case, it is questionable whether the unfavored enantiomer would interconvert into the favoured enantiomer, at ambient temperatures, during hybridization by bond rotation about the C_{Ar}-N bond. In order to obtain some insight into these questions, molecular mechanics calculations and NMR experiments were used to estimate the barrier of rotation for interconversion of the two atropisomers.

To simplify the molecular modeling calculations, the compound **116** was chosen as a model structure; the groups omitted from compound **87b** were expected to have little impact on the overall outcome of the calculations. In this experiment, the C_{Ar}-N bond was rotated as shown in Figure 2.4 in increments of 10°, while holding the dihedral angle of interest constant. The structure was minimized using the MM+ force field provided in the HyperChem Molecular Modelling software (v5.1). Finally, the structure was then further refined by a semi-empirical calculation at the PM3 level. The bond was rotated so that the $-CH_2NHCHO$ moiety passed in proximity to the allyl fragment of the molecule, which is the path calculated to be subject of the least steric interaction. The energies obtained by this method were then plotted as a function of torsion angle to generate the free energy diagram shown in Figure 2.4. Based on these calculations, the barrier to

interconversion was estimated at approximately 23.5 kcal/mol, which should allow for resolution of the enantiomers at room temperature.

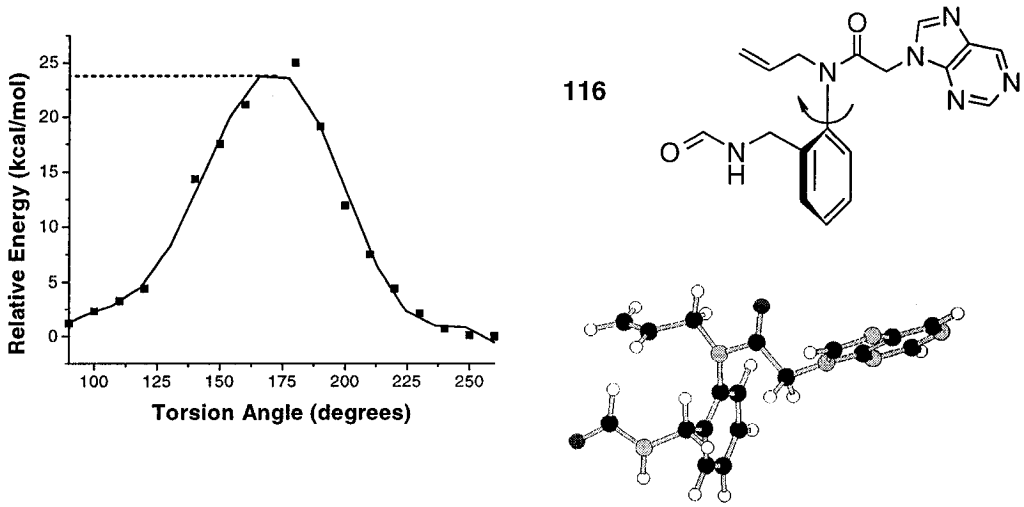


Figure 2.4: Free Energy Diagram for Interconversion of Atropisomers of Compound **116.**

A more direct method for determining the barrier to rotation is by variable temperature NMR spectroscopy; provided that the rate of interconversion of rotational isomers is slow enough at room temperature to be observed on the NMR timescale. The temperature dependence of all the methylene ^1H chemical shift of compound **87a** is shown in Figure 2.5. It was expected that at the point of free rotation about the C_{Ar}-N bond, each signal for the diasteriotopic geminal protons of compound **87a** would coalesce into a single resonance. However, this phenomenon was not observed for resonances that were completely resolved and so an accurate determination of the barrier of rotation could not be made in this way. Due to the temperature limitation imposed by the instruments available, the only signals that were appropriately spaced for these experiments were those of geminal protons H_c and H_d. However, these signals were

obscured by the resonance for proton H_b as this signal coalesced with H_a. Furthermore, while it is not clear exactly when the resonances for H_a and H_b coalesce, it seems as though the coalescence may occur at approximately 140-150°C. This range would correspond to a barrier to rotation of approximately 22 kcal/mol, within reasonable agreement with that calculated by molecular modelling above.

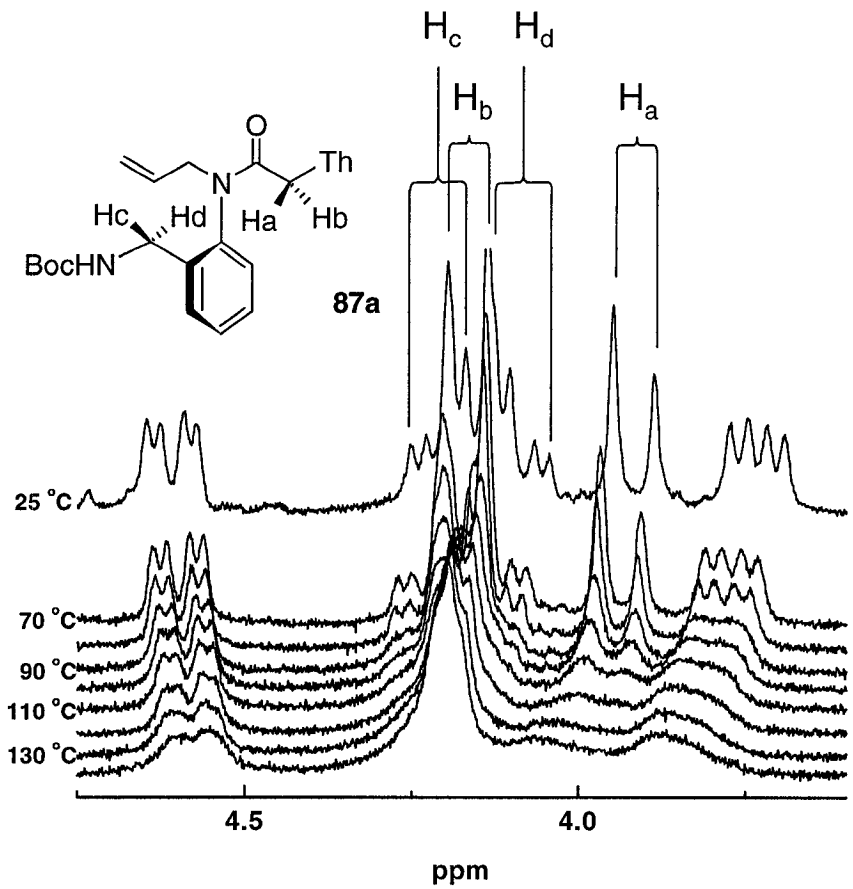


Figure 2.5: Temperature Dependence of Chemical Shift of the Resonances for the Methylene Region of Compound **87a**.

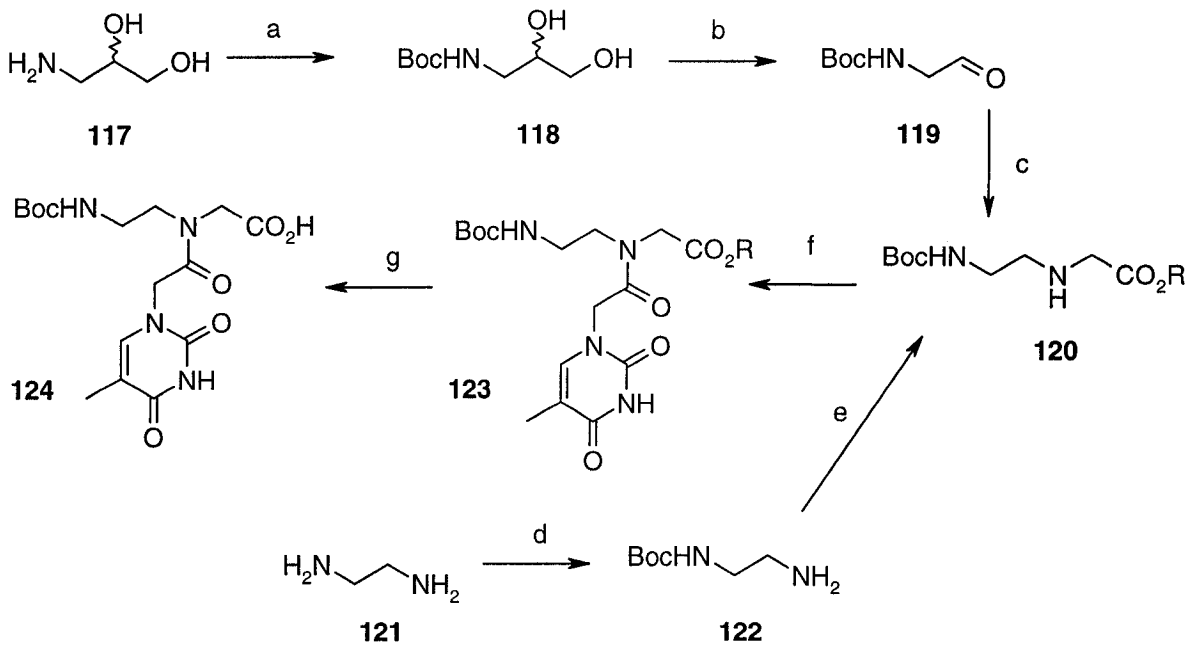
Atropisomerism has been arbitrarily defined by Oki as being present when one atropisomer can be isolated from the other and have a half life of greater than 1000 seconds.[95] This definition does not specify a barrier of rotation, but at 300K the critical barrier is 22.3 kcal/mol. Since each atropisomer of the APNA monomers described above may have different DNA or RNA recognition properties, a method for separation of each enantiomer would be of significant value. The dihydroxylation reaction of intermediates **87a,b** gave a 1:1 mixture of diasteromers of diols **88a,b** (Scheme 2.9).[96] Although these diasteromers could not be separated by silica gel chromatography, partial separation of the diasteriomers was achieved by chiral HPLC. However, since the barrier of interconversion is very similar to the minimum free energy barrier that would allow for separation of the diastereomers at room temperature, a pure enantiomer of **45a,b** once incorporated into an oligomer would lose all or most of its optical purity under the conditions used for annealing the complementary strands in hybridization experiments described in Chapter 4. Therefore, no further attempt was made to resolve the enantiomers of monomer **45a,b**.

2.10 - Synthesis of PNA Monomers.

The oligomers prepared as part of this thesis required the use of Boc/Cbz- and Fmoc/Cbz-protected PNA monomers, which were prepared following literature procedures. The Boc-protected thymine PNA monomers were prepared using one of two methods (Scheme 2.14). In the first approach, 3-amino-1,2-propanediol (**117**) was N-Boc

protected by reaction with Boc_2O under basic conditions giving diol **118**. Oxidative cleavage of the diol, followed by reductive amination of the corresponding aldehyde **119** with glycine methyl ester gave the desired secondary amine **120** in low to modest yield (Scheme 2.14). Alternatively, the PNA backbone **120** could be prepared by mono-Boc protected of ethylenediamine by reaction with 0.13 equivalents of Boc_2O to give pure compound **122** in >98% yield. Compound **122** was further reacted with 0.13 equivalents of ethyl bromoacetate in the presence of K_2CO_3, giving compound **120** in 96% yield. In previously reported procedures, the preparation of compound **120** was carried out using a 1:1 molar ratio of Boc-protected ethylenediamine to ethyl bromoacetate.[97] In our hands, this method led to uncontrollable dialkylation as a major side reaction. Excess amine **122** was required in this reaction in order to avoid dialkylation, however, the unreacted starting material was usually recovered quantitatively. Each of these methods of PNA backbone construction has advantages and disadvantages. Although the first route involves more steps and the yield of the reductive amination is relatively low, with poor mass recovery, this method was practical for production PNA monomer in a larger than 10g scale. In contrast, the second route was far more efficient and allowed for near quantitative mass recovery of the amine **122**. However, the large excess of amine **122** required and the high dilution conditions made large-scale preparation impractical. In order to complete the synthesis of the thymine monomer **124**, thymine derivative **51** was coupled with secondary amine **120** giving fully protected thymine monomer **123**, which was then hydrolyzed to the free carboxylic acid **124** (Scheme 2.14).

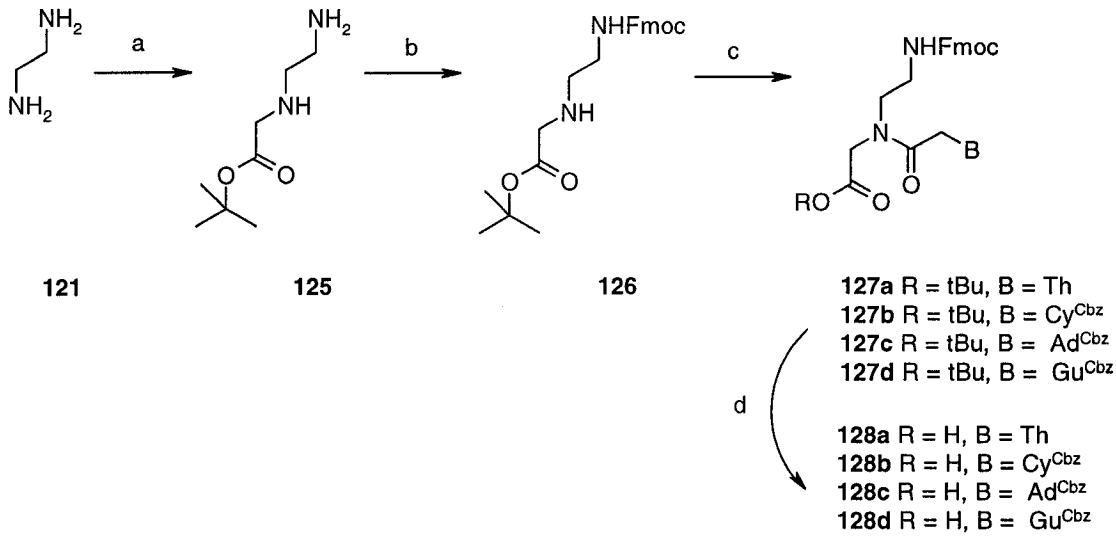
Scheme 2.14: Synthesis of Boc-Protected, Thymine PNA Monomer.



Conditions: a) Boc_2O, NaOH, H_2O, >99%; b) $NaIO_4$, THF/H_2O, >98%; c) $HCl.H_2NCH_2CO_2R$,$NaBH_3CN$, EtOH, 25-45%; d) 0.13 equivalent Boc_2O, THF, >99%; e) 0.13 equivalent $BrCH_2CO_2R$, K_2CO_3, $CHCl_3/CH_2CN$, 90-95%; f) **51**, EDC, DMF, 75-90%; g) LiOH, THF/H_2O, then HCl, >90%.

Fmoc-protected PNA monomers were also prepared following a literature method without modification (Scheme 2.15). Ethylenediamine was mono-alkylated with 0.1 equivalents of *tert*-butyl bromoacetate giving diamine **125**. The primary amino group of compound **125** was selectively protected with N-(9-Fluorenylmethoxycarbonyloxy) succinimide (Fmoc-ONSu) to give the PNA backbone fragment **126**. The nucleobase derivatives were coupled to secondary amine **126** under standard conditions and the *tert*-butyl esters converted to the free acids **128a-d** after treatment with TFA in DCM.

Scheme 2.15: Synthesis of Fmoc/Cbz-Protected PNA Monomers.



Conditions: a) $BrCH_2CO_2tBu$; b)Fmoc-OSucc, DIPEA; c) B-CH_2CO_2H, EDC, DMF; d) TFA/DCM.

2.11 - Conclusions.

Synthetic methods were developed for the preparation of several structurally diverse APNA monomers. These analogues vary in the position of the aromatic ring and nucleobase. Furthermore, APNA monomers containing a variety of substituents and functional groups on the backbone and aromatic ring were prepared. Monomers containing N-phenyl-N-alkyl amide moieties are acyclic PNA analogues whose tertiary amide rotamer equilibrium heavily favours the conformation preferred by unmodified PNAs when hybridized to DNA and RNA. This set of monomer building blocks is useful for the preparation of oligomers designed to evaluate the hybridization properties of novel PNA analogues, the APNAs, that incorporate aromatic rings into their backbone.

2.12 – References.

[75] Wang, S. Y.; Nnadi, J. C.; Greenfeld, D. *Chem. Commun.* **1968**, *19*, 1162.

[76] Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767.

[77] Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677.

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

[79] This is contrary to that observed in the literature for the removal of the t-butyl group of compound **57**. See ref 78.

[80] Cook, P. D.; Sprankle, K. WO 95/23163, **1995**.

[81] For the initial alkylation of cytosine, the method of Cook and Sprankle was used. See: ref. 76. The Cbz- protection of ester **61** was done using an independent method.

[82] Coull, J. M.; Hodge, R. P. WO 95/17403, **1995**.

[83] a) Abdel-Magid, A. F.; Carson, K. G.; Harris, B. D.; Maryanoff, C. A.; Shah, R. D. *J. Org. Chem.* **1996**, *61*, 3849.; b) Abdel-Magid, A. F.; Maryanoff, C. A.; Carson, K. G. *Tetrahedron Lett.* **1990**, *31*, 5595.

[84] a) Carpino, L. A.; El-Faham, A. *J. Org. Chem.* **1994**, *59*, 695; b) Carpino, L.; El-Faham, A.; Minor, C. A.; Albericio, F. *J. Chem. Soc. Chem. Commun.* **1994**, *2*, 201.

[85] For synthetic details and characterization of monomer **47**, see: Boyd, M. J. *M.Sc. Thesis*, **1999**, Concordia University, Montreal, Canada.

[86] Corey, E. J.; Suggs, J. W. *J. Org. Chem.* **1973**, *38*, 3223.

[87] Honda, M.; Morita, H.; Nagakora, I. *J. Org. Chem.* **1997**, *62*, 8932.

[88] Dangles, O.; Guibe, F.; Balavoine, G.; Lavielle, S; Marquet, A. *J. Org. Chem.* **1987**, *52*, 4984.

[89] Yu, C.; Liu, B.; Hu, L. *Tetrahedron. Lett.* **2000**, *41*, 4281.

[90] For recent examples of benzylic chlorides see: a) Hamashima, Y.; Sawada, D.; Nogami, H.; Kanai, M.; Shibasaki, M. *Tetrahedron* **2001**, *57*, 805.; b) Maezaki, N.; Izumi, M.; Yuyama, S.; Sawamoto, H.; Iwata, C.; Tanaka, T. *Tetrahedron* **2000**, *56*, 7927.; c) Brondsted Nielsen, M.; Hansen, J. G.; Becher, J. *Eur. J. Org. Chem.* **1999**, *11*, 2807.; d) Abe, Y.; Kayakiri, H.; Satoh, S.; Inoue, T.; Sawada, Y.; Inamura, N.; Asano, M.; Aramori, I.; Hatori, C.; Sawai, H.; Oku, T.; Tanaka, H. *J. Med. Chem.* **1998**, *41*, 4062.; e) Davies, P. J.; Veldman, N.; Grove, D. M.; Spek, A. L.; Lutz, B. T. G.; van Koten, G. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1959.; f) Douty, B. D.; Salvino, J. M.; Seoane, P. R.; Dolle, R. E. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 363.; g) Ciufolini, M. A.; Qi, H. B.; Browne, M. E. *J. Org. Chem.* **1988**, *53*, 4149.; h) Ciufolini, M. A.; Browne, M. E. *Tetrahedron. Lett.* **1987**, *28*, 171. For recent examples of allylic chlorides see: a) Myers, A. G.; Glatthar, R.; Hammond, M.; Harrington, P. M.; Kuo, E. Y.; Liang, J.; Schaus, S. E.; Wu, Y.; Xiang, J. -N. *J. Am. Chem. Soc.* **2002**, *124*, 5380.; b) Deiters, A.; Hoppe, D. *J. Org. Chem.* **2001**, *66*, 2842.; c) Deiters, A.; Mueck-Lichtenfeld, C.; Froehlich, R.; Hoppe, D. *Org. Lett.* **2000**, *2*, 2415.; d) Amano, S.; Ogawa, N.; Ohtsuka, M.; Chida, N. *Tetrahedron* **1999**, *55*, 2205.; e) Heathcock, C. H.; Brown, R. C. D.; Norman, T. C. *J. Org. Chem.* **1998**, *63*, 5013.; f) Amano, S.; Ogawa, N.; Ohtsuka, M.;

Chida, N. *Chem. Comm.* **1998**, *12*, 1263.; g) Hall, D. G.; Deslongchamps, P. *J. Org. Chem.* **1995**, *60*, 7796.

[91] Eliel, E. L.; Wilen, S. H., *Stereochemistry of Organic Compounds*, John Wiley and Sons Inc., New York, NY, **1994**, p 554.

[92] In cases where a Boc group was present in the molecule, some decomposition was observed during the prolonged acquisition time required for ^{13}C NMR. Thermal removal of Boc protecting groups has been documented; see: a) Rawal, V. H.; Jones, R. J.; Cava, M. P. *J. Org. Chem.* **1987**, *52*, 19.; b) Wasserman, H. H.; Berger, G. D.; Cho, K. R. *Tetrahedron. Lett.* **1985**, *23*, 465.

[93] Manea, V. P.; Wilson, K. J.; Cable, J. R. *J. Am. Chem. Soc.* **1997**, *119*, 2033 and references cited therein.

[94] H_a and H_b are most easily distinguished from other isolated methylene protons by nOe correlation to a nucleic acid base proton (H6 for thymine, H8 for adenine). In all cases studied, the resonance of H6 of thymine overlapped with aromatic protons of the APNA backbone lending to some ambiguity. This was not the case for H8 of adenine.

[95] Oki, M. *Top. Stereochem.* **1983**, *14*, 1.

[96] Estimated by integration of the ^{13}C resonances for each diasteromer in the crude mixture.

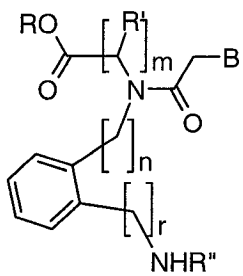
[97] Aldrian-Herrada, G.; Rabié, A.; Wintersteiger, R.; Brugidou, J. *J. Pept. Sci.* **1998**, *4*, 266.

CHAPTER 3

RESULTS AND DISCUSSION: SOLID-PHASE SYNTHESIS OF APNA HOMOPOLYMERS AND APNA-PNA CHIMERAS.

3.1 - Objectives.

In order to evaluate the biophysical properties of the novel APNA monomers described in Chapter 2, and to identify a lead structure for further optimization, the monomer units **43-47** were incorporated into the middle of all thymine PNA hexamers, which were used for hybridization studies with DNA and RNA. Subsequently, the most promising lead structure was used to prepare APNA homopolymers and APNA-PNA chimeras for further evaluation of the biophysical properties associated with APNA containing oligomers.



	m	n	r	R'
43	1	1	0	H
44	2	1	0	H
45	1	0	1	H
46	2	0	0	H
47	1	1	0	Me

3.2 - Synthesis of PNA Containing Oligomers.

The PNA segments of all the oligomers used in the hybridization experiments described herein were prepared by solid-phase peptide synthesis. Oligomerization was carried out on a methylbenzhydrylamine (MBHA) resin following previously established procedures.[40] 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) was used as the coupling reagent in the presence of diethylcyclohexylamine (DECA) and in a solvent mixture of 1:1 pyridine/DMF. After each step, unreacted free amino residues were capped by treatment of the resin with Ac_2O in the presence of pyridine. Capping of the *N*-terminus of the PNA oligomers was done in order to avoid intramolecular *N*-acyl transfer of the last acetylnucleobase moiety (Scheme 3.1). The N-Boc protecting groups were removed by treating the peptide with 25% TFA in dichloromethane and the progress of both the coupling and deprotection steps was monitored using the Kaiser test.[98] In cases where the N-Fmoc protecting group was used, the protected amines were liberated from the resin with 20% piperidine in DMF. It should be noted that a lysine residue was included at the *C*-terminus of each peptide in order to prevent aggregation of the polymer-bound oligomer during its synthesis on the solid support.[99] However, based on the literature,[99] the presence of this amino acid at the *C*-terminal of an oligomer does not have any significant impact on the hybridization properties of these molecules. Cleavage from the resin was achieved by treatment of the polymer-bound peptide with trifluromethanesulfonic acid (TFMSA) in TFA. The crude peptides were first precipitated from the TFA solution, by diluting the mixture with anhydrous diethyl ether (20-fold dilution) and then purified by preparative C18 reversed-

phase HPLC. Following this methodology, the PNA oligomers **129-131** listed in Table 3.1 were prepared.

Scheme 3.1: Acyl Transfer Reaction of Resin Bound PNA Oligomers.

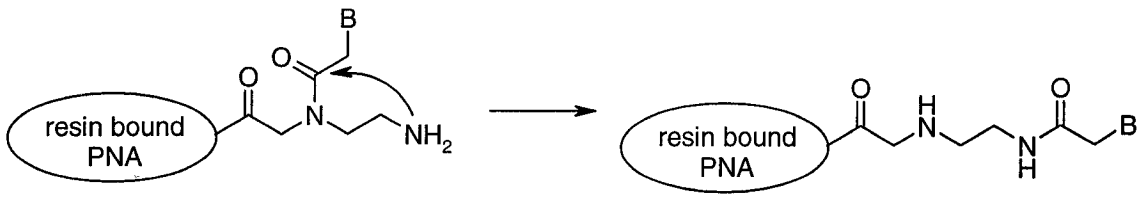


Table 3.1: PNA Control Sequences Used in Hybridization Studies.

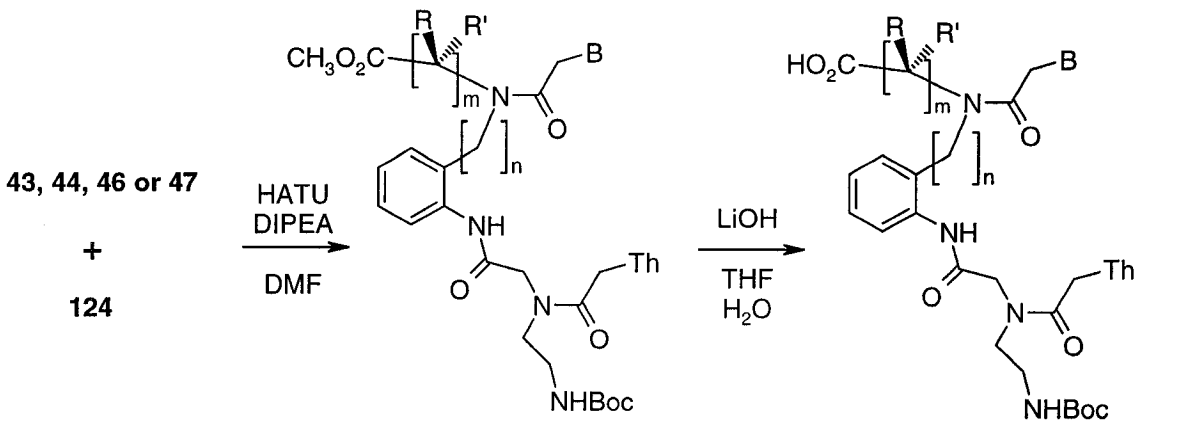
PNA		Sequence	
129	N	Ac-TTTTTT-Lys-NH_2	C
130	N	Ac-ATCATTCTCT-Lys-NH_2	C
131	N	H-GTAGATCAACT-Lys-NH_2	C

3.3 - Synthesis of APNA-PNA Homo-Thymine Hexamers.

The *ortho*-substituted aniline moieties of monomers **43**, **44**, **46** and **47** were found to react sluggishly during peptide synthesis, even when an excess of the free acids and coupling reagents were used. Furthermore, the free aniline moieties cannot be detected

using the Kaiser test and so a different method for monitoring the coupling cycles of each APNA unit would have to be explored. Therefore, all of the APNA monomers were incorporated into the oligomers using the APNA-PNA dimers **138-143**. These were easily synthesized in solution by the coupling of compounds **43**, **44**, **46** and **47** with the PNA monomer **124** using HATU/DIPEA (Scheme 3.2).[100] The ester moieties of intermediates **132-137** were hydrolyzed to give the free acids **138-143** which were used directly for the solid-phase synthesis of the Lys-PNA_2-APNA-PNA_3-Ac hexamers (**144-150**) following the same protocol as for the synthesis of the PNA control (Scheme 3.3). After cleavage of these oligomers from the solid support, HPLC analysis of the crude material indicated the presence of the desired PNA-APNA chimeras in >65% of the total sample. However, oligomers **144-150** were isolated in approximately 30-50% yield after HPLC purification.

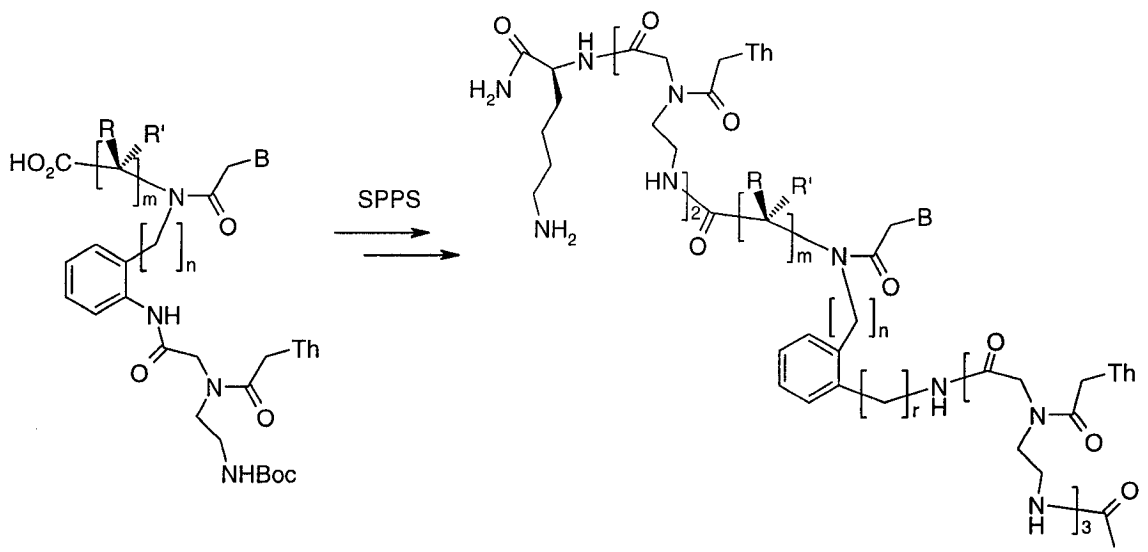
Scheme 3.2: Synthesis of APNA-PNA Dimers **138-143**.



	B	m	n	R	R'
132	Th	1	1	H	H
133	Th	2	1	H	H
134	Th	2	0	H	H
135	Th	1	1	Me	H
136	Th	1	1	H	Me
137	Cy^{Cbz}	1	1	H	H

	B	m	n	R	R'
138	Th	1	1	H	H
139	Th	2	1	H	H
140	Th	2	0	H	H
141	Th	1	1	Me	H
142	Th	1	1	H	Me
143	Cy^{Cbz}	1	1	H	H

Scheme 3.3: Synthesis of APNA-PNA Chimeras **144-150**.

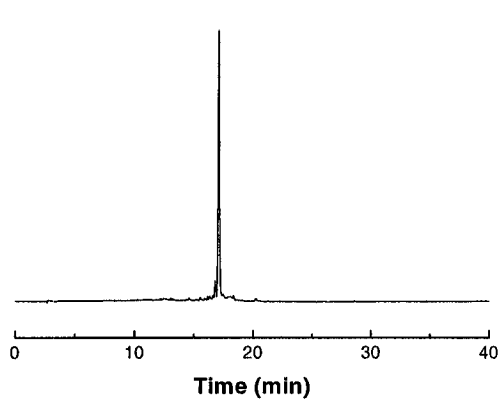


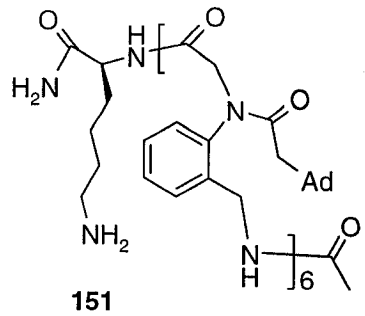
138-143

	m	n	r	B	R	R'
144	1	1	0	Th	H	H
145	1	0	1	Th	H	H
146	2	1	0	Th	H	H
147	2	0	0	Th	H	H
148	1	1	0	Th	Me	H
149	1	1	0	Th	H	Me
150	1	1	0	Cy	H	H

An appealing structural feature of monomer **45** is that the N- terminal of this analogue is a benzyl amine, in contrast to the aniline-based APNA monomers **43**, **44**, **46** and **47**. The superior nucleophilic character of this amine led to more efficient solid phase coupling reactions during chain elongation. In all cases, coupling reactions with monomers **45a,b** were done under the conditions developed for PNA chemistry which were discussed in Section 3.2. Generally, these monomers were completed within 30 to 40 minutes and the progress of the reaction could be easily monitored using the Kaiser test.[101] Following these procedures, oligomer **145** was prepared using APNA monomer

45a and the APNA homopolymer **151** was prepared from monomer **45b** in excellent yield and purity (Figure 3.1). In all cases, the purity of each oligomer was confirmed to be greater than 90% by analytical C18 HPLC (run at 60 °C with a slow solvent gradient from H_2O to CH_3CN) and characterized by electrospray mass spectroscopy.



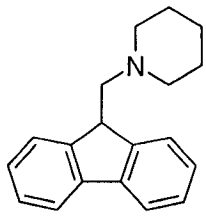


151

Figure 3.1 HPLC Analysis of Crude Hexamer **151**.

3.4 - Optimization of Solid-Phase Peptide Chemistry for Aniline-Based APNA Modifications.

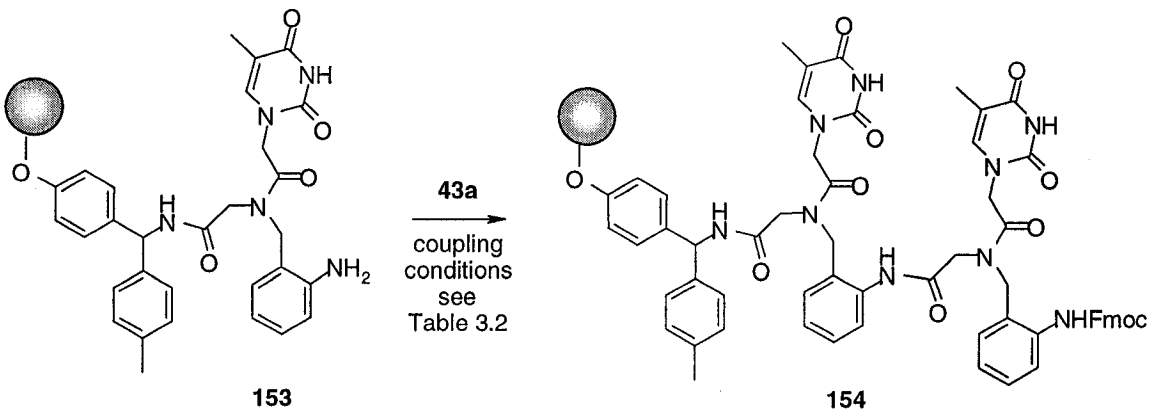
The Fmoc-based coupling strategy is an approach which is often used in solid-phase synthesis so that the coupling yields at each cycle can be quantified by spectroscopic analysis of the dibenzofullvene adduct **152**.[102] This approach was chosen for the solid-phase synthesis of oligomers containing more than one APNA insert. PNA-APNA chimeras were synthesized using Fmoc-protected PNA monomers **128a-d**, following the coupling protocol described in Section 3.2.



152

A number of reagents were explored in order to find the most efficient protocol for the coupling of anilines. Preliminary solution phase studies indicated that the coupling reactions involving APNA free aniline monomers were most efficient when the uronium based reagent HATU was used to activate the carboxylic acid. In general, other reagents based on the formation of HOBt esters, such as HBTU and TBTU, as well as the phosphonium reagent BOP-Cl, gave lower coupling yields, whereas, the carbodiimide reagents DCC and EDC were ineffective in providing the desired amides in any acceptable quantities. Therefore, these reagents were not tested during optimization of the solid-phase synthesis protocol. Instead, HOAt derived active esters, as well as 2-chloro-1-methylpyridinium iodide (MCPI, Mukayama's Reagent) were explored.[103] In order to optimize the coupling conditions for APNA anilines, compound **154** was synthesized on MBHA resin using monomer **43a** (B = Th, R = H, R" = Fmoc). The yields of the coupling reaction under different conditions are summarized in Table 3.2. Using HATU as the coupling reagent, DIPEA as the base and 1:1 DMF/pyr as the solvent (optimal for PNA synthesis), a coupling yield of 33% was obtained (entry 1). In contrast, when MCPI was used with the same solvent the coupling yield increased to 70% (entry 2). Changing the solvent to DMF when HATU was used (entry 3) improved the coupling yield slightly to 55%. Coupling reactions in DCM, even with MCPI as the coupling reagent (entry 4), gave poor yield (29%); however, in this solvent the APNA monomers

were poorly soluble. When HATU was used in DMF as the coupling reagent and the base was changed to 2,4,6-collidine the coupling yield also improved very slightly to 57% (compare entries 3 and 5). Finally, if 1 equivalent of HOAt and 3 equivalents of collidine were used, the coupling yield improved to 61.9% (entry 6). On the basis of these experiments, MCPI with DIPEA in DMF/pyr or HATU/HOAt with collidine in DMF appeared to be the best combination of reagents and solvent for coupling APNA monomers to resin bound free anilines.



153 **154**

Table 3.2: Optimization of Coupling Yield for Solid-Phase Synthesis of APNA Oligomers and APNA-PNA Chimeras.[a]

entry	Coupling Reagent Used	Solvent	Yield of Coupling[c]
1	HATU / DIPEA (1 equ / 2 equ)[b]	1:1 DMF/pyr	33.4%
2	MCPI / DIPEA (4 equ / 6 equ)	1:1 DMF/pyr	70.4%
3	HATU / DIPEA (1 equ / 2 equ)	DMF	54.8%
4	MCPI / DIPEA (4 equ / 6 equ)	DCM	29.1%
5	HATU / collidine (1 equ / 2 equ)	DMF	57.3%
6	HATU / HOAt / collidine (1 equ / 1 equ / 3 equ)	DMF	61.9%

[a]- 3 equivalents of monomer **43a** relative to the loading of the resin were used in all cases. [b]- relative to free acid **43a**; [c]- reaction time was 4.5h in all cases.

Finally, for the synthesis of oligomers, the conditions given in entry (6) (Table 3.2), were altered slightly by a) increasing the number of equivalents of free acid, HATU, HOAt and collidine to a ratio of 5:5:5:15 relative to the initial loading of the resin bound aniline, and b) by increasing the reaction time to 10 h. These changes resulted in average coupling efficiency of approximately 95%, based on quantification of the Fmoc chromophore. However, for oligonucleotide synthesis an average coupling yield of 95% is not efficient enough for long oligomers. For example, with a 95% yield at each coupling step, the overall yield for the synthesis of an 18-mer would be a maximum of 39% (assuming all other steps, such as deprotections and capping, are quantitative). To compensate for the less than ideal coupling efficiency of the aniline APNA monomers, and also aid oligomer purification, Fmoc protected APNA dimers were employed as building blocks for oligomer synthesis. The required dimers were synthesized in solution

by coupling the appropriate aniline **82a-d** with the appropriate free acid **43a-d**, followed by a deprotection step to give the free acid dimers indicated in Scheme 3.4.

Scheme 3.4: Synthesis of APNA Dimers Used in Solid-Phase Synthesis.

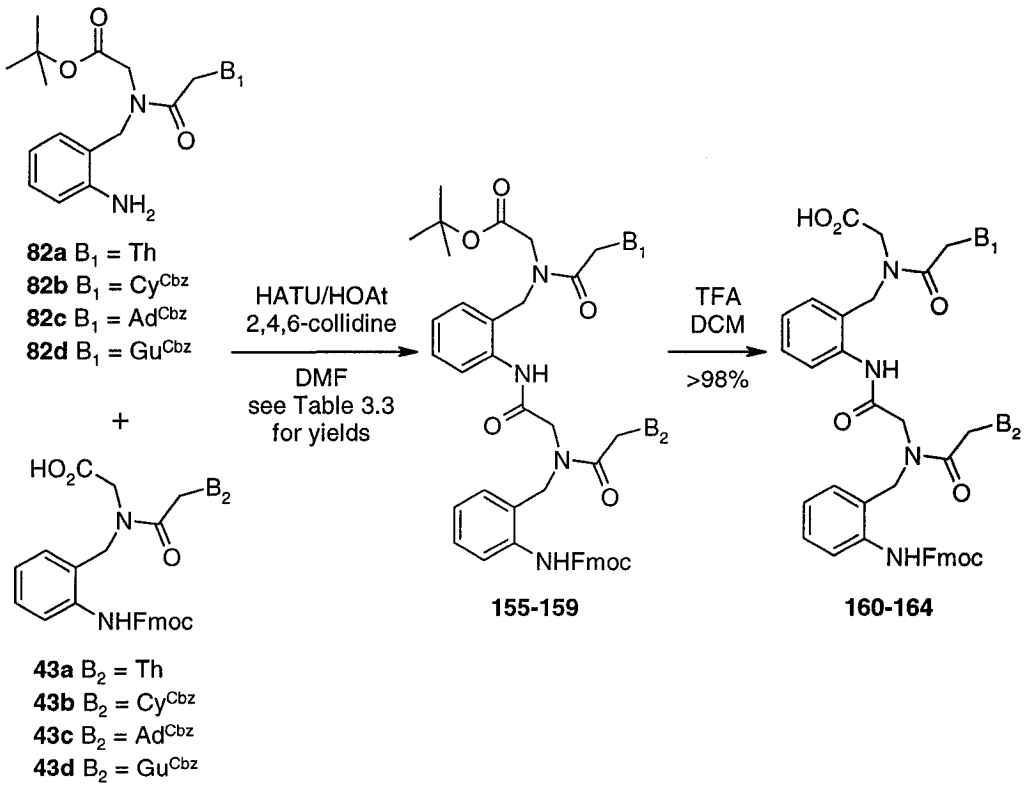


Table 3.3: Summary of Yields for APNA Dimer Synthesis.

Compound	B_1	B_2	% yield of dimer	Corresponding Free Acid
155	T	T	58	**160**
156	A^{Cbz}	C^{Cbz}	70	**161**
157	T	A^{Cbz}	74	**162**
158	G^{Cbz}	A^{Cbz}	40	**163**
159	T	G^{Cbz}	49	**164**

The Fmoc protecting group on the last residue was retained on the oligomer in order to render the final product more lypophilic than the *n*-2 failure sequences and consequently was easier to purify by reversed-phase HPLC. Once oligomers had been synthesized on the solid support, they were cleaved from the resin in one step by treatment with TFMSA in TFA with thioanisole as the scavenger. The reaction mixture was then diluted with ether (20-fold dilution) and the fluffy, white precipitate was centrifuged to a pellet. Upon analysis of the crude mixture by HPLC, often numerous peaks were detected whose mass spectra corresponded to full-length decamers containing from one to five Cbz protecting groups. In those cases, the isolated solid was re-treated with the cleavage cocktail in order to remove the remaining protecting groups (Figure 3.2). Generally, the purity of the final product was superior if the crude solid obtained after cleavage of the peptide from the resin and ether precipitation was re-subjected to the cleavage conditions rather than prolonging the original cleavage reaction. Using this protocol, compounds **165-173** were prepared (Table 3.4); the reasons for this are not obvious.

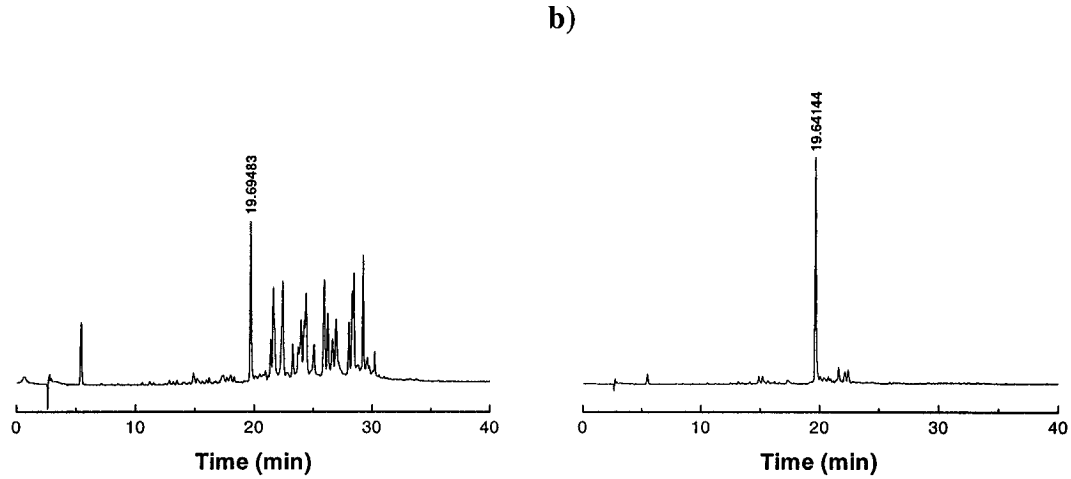


Figure 3.2: HPLC Analysis of: a) Crude Sample of Decamer Fmoc-**168** (Table 3.4); and b) the Same Sample after Subsequent Treatment with TFMSA/TFA/thioanisole.

Table 3.4: APNA Containing Sequences Used in Binding Experiments.[a]

Compound	Sequence
165	Ac-ATCAT-$\mathbf{T}_{\alpha}$-CTCT-Lys-NH_2
166	Ac-ATCA-$\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}$-CTCT-Lys-$NH_2$
167	H- $\mathbf{A}_{\alpha}\mathbf{T}_{\alpha}\mathbf{C}_{\alpha}\mathbf{A}_{\alpha}$-TTCTCT-Lys-$NH_2$
168	H- $\mathbf{A}_{\alpha}\mathbf{T}_{\alpha}\mathbf{C}_{\alpha}\mathbf{A}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{C}_{\alpha}\mathbf{T}_{\alpha}\mathbf{C}_{\alpha}\mathbf{T}_{\alpha}$ -Lys-NH_2
169	H-GTAGA-$\mathbf{T}_{\alpha}$-CACT-Lys-NH_2
170	H-GTAG-$\mathbf{A}_{\alpha}\mathbf{T}_{\alpha}$-CACT-Lys-$NH_2$
171	H-$\mathbf{G}_{\alpha}\mathbf{T}_{\alpha}\mathbf{A}_{\alpha}\mathbf{G}_{\alpha}$-ATCACT-Lys-$NH_2$
172	Ac-$\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}$ -Lys-NH_2
173	H-Asp-$\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}\mathbf{T}_{\alpha}$ -Lys-NH_2

Notes: [a] APNA Residues with subscript "α"and bold text, PNA residues in normal text.

3.5 – References.

[98] Sarin, V. K.; Kent, S. B.; Tam, J. P.; Merrifield, R. B. *Anal. Biochem.* **1981**, *117*, 147.

[99] Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895.

[100] For synthetic details and characterization of dimers **135**, **136**, **141** and **142**, see ref. 81.

[101] In these cases, the polymer beads turned the same characteristic blue color observed for unprotected PNA amino groups.

[102] a) Chang, C. -D.; Felix, A. M.; Jimenez, M. H.; Meienhofer, J. *Int. J. Pept. Protein Res.* **1980**, *15*, 485.; b) Peptide and Peptidomimetic Synthesis. Reagents for Drug Discovery. Fluka Chemie GmbH, Buchs, **2000**, p123.

[103] Mukaiyama, T.; Usui, M.; Shimada, E.; Saigo, K. *Chem. Lett.* **1975**, *10*, 1045.; b) Bald, E.; Saigo, K.; Mukaiyama, T. *Chem. Lett.* **1975**, *11*, 1163.

CHAPTER 4

RESULTS AND DISCUSSION: PHYSICOCHEMICAL PROPERTIES OF APNA-PNA CHIMERAS CONTAINING THYMINE, CYTOSINE, ADENINE AND GUANINE.

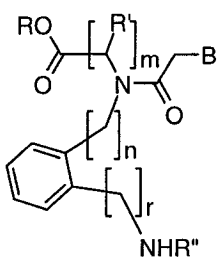
4.1 - Objectives.

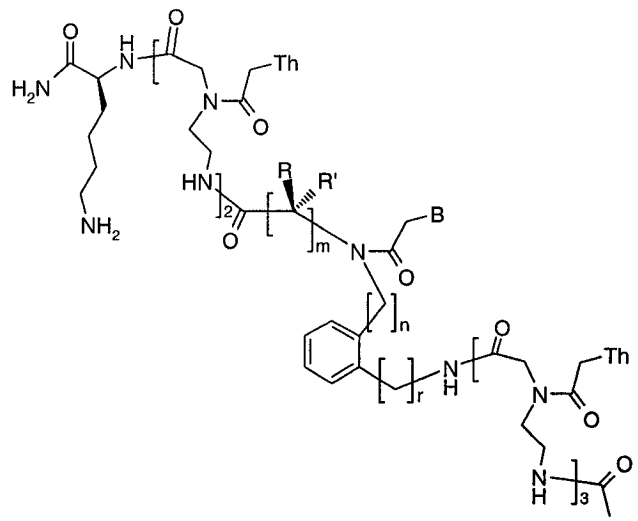
It was assumed that once an APNA monomer was identified that exhibited good conformational rigidity and reasonable molecular recognition for natural nucleic acids, its backbone moiety could be further modified in order to optimize the biophysical properties of oligomers composed of such monomers. Thermal denaturation experiments involving hexamers **129**, **144-150** were first used to identify APNA monomers with promising biophysical properties. Oligomers containing multiple inserts of the APNA monomers (PNA-APNA chimeras) were examined by thermal denaturation and cyanine dye binding experiments.

4.2 - Physicochemical Properties of APNA-PNA Hexamers.

In order to gain some insight into the physicochemical properties of each APNA monomer, the hybridization properties of hexamers **144-150** with polyriboadenylate (poly

rA) and polydeoxyriboadenylate (poly dA) were compared to those of the control PNA-T_6 (**129**, Table 4.1). In all cases, the stability of triplex formation was evaluated by UV-thermal melting (T_m) experiments at nearly physiological conditions (pH = 7.1, 150 mM NaCl, 10 mM NaH_2PO_4, 1 mM EDTA). Hexamers were mixed in a 2:1 ratio (T_2:A) based on the expectation that each PNA strand (or APNA-PNA chimera) would bind to the DNA or RNA strand to form PNA_2:DNA or PNA_2:RNA triplexes.[104] Complex formation in a 2:1 ratio was also confirmed by Job plot titration. [105]





	m	n	r	R'
43	1	1	0	H
44	2	1	0	H
45	1	0	1	H
46	2	0	0	H
47	1	1	0	Me

	m	n	r	B	R	R'
144	1	1	0	Th	H	H
145	1	0	1	Th	H	H
146	2	1	0	Th	H	H
147	2	0	0	Th	H	H
148	1	1	0	Th	Me	H
149	1	1	0	Th	H	Me
150	1	1	0	Cy	H	H

The circular dichroism (CD) spectrum of the complex formed between the Lys-PNA_2-APNA-PNA_3-Ac chimera **144** and poly rA was very similar to that observed for

the triplex of the control PNA-T_6 (**129**) with poly rA (Figure 4.1). Similar CD spectra were also observed for the complexes formed between the PNA-T_6 and poly dA, as well as chimeras **144** and poly dA, strongly suggesting that the incorporation of a single monomer of structure **43** into the PNA homopolymer did not alter the known hybridization characteristics of these oligomers.

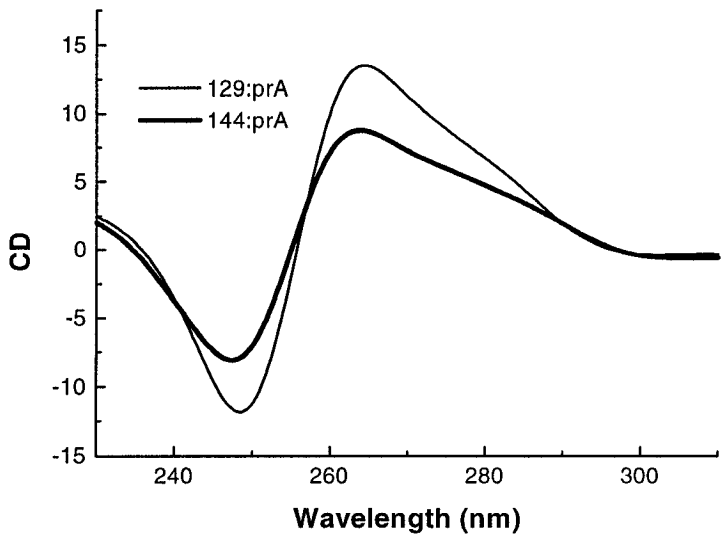


Figure 4.1: CD Spectra of the Complexes **129**:prA and **144**:prA.

The melting temperatures (T_m) of the complexes between oligomers **129**, **144-150** and poly rA and poly dA are summarized in Table 4.1. The PNA-T_6 control strand (**129**) gave T_m values of 65°C and 56°C when hybridized to poly rA and poly dA, respectively (Table 4.1), consistent with those reported previously for triplexes of PNAs with natural oligonucleotides.[106] However, destabilization was observed upon insertion of one APNA monomer of analogue **43** into the middle of the hexamer (chimera **144**), resulting in a significant decrease in the T_m values observed for the complexes formed between the

Lys-PNA_2-APNA-PNA_3-Ac chimeras (**144**) and poly rA or poly dA as compared to those of the control. Incorporation of monomer **45** into a hexamer (i.e. chimera **145**) resulted in only a slight further decrease of the T_m value for the triplex with RNA (ΔT_m = -3°C) as compared to hexamer **144**, whereas a more significant difference was observed with DNA (ΔT_m = -6°C). Increasing the backbone distance between adjacent nucleobases, as in the case of the APNA β-alanine analogue (monomer **44**, chimera **146**), also had a detrimental effect on the stability of the complex formed with poly rA, whereas there was no hybridization observed between **146** and poly dA. Finally, when monomer **46** was incorporated into the PNA hexamer (chimera **147**), there was no hybridization with either poly rA or poly dA. It should be noted that in both monomers **44** and **46** the space between adjacent nucleic acid bases is significantly altered, in particular at the junction points between the APNA and PNA units, which most likely results in a dramatic destabilization of any complex that may form with DNA or RNA. Thus, it is conceivable that homopolymers of monomers **44** or **46** could potentially exhibit much better hybridization properties than those observed with the PNA-APNA chimeras **146** and **147,** respectively.

Substitution at C_α of the amino acid moiety in the APNA backbone also had a negative effect on the binding properties of the APNA-PNA chimeras, although methyl substitution of the *R* configuration (chimera **148**) was slightly better than the *S* (chimera **149**). The destabilizing effects observed due to backbone substitution of the APNA analogues are far more significant than those previously reported for the C_α-substituted analogues of PNAs.[107]

Table 4.1: Results of Thermal Denaturation Experiments[a] for Complexes of Hexamers **126, 141-147** with poly rA and poly dA.

compound	complex with poly rA (T_m, °C)	complex with poly dA (T_m, °C)
129	65	56
144	50	45
145	47	39
146	26	*<5*
147	<5	<5
148	39	30
149	30	26
150	17	<5

[a] Melting temperatures of duplexes were determined on a Cary 3 UV-vis spectrophotometer. All T_m solutions were 150mM in NaCl, 10mM in NaH_2PO_4, 1mM in EDTA, the pH was adjusted to 7.1. The solutions were heated to 90°C for 10 min then cooled slowly to 4°C and stored at that temperature for at least 1h. The melting curves were recorded by heating the solutions from 5°C to 90°C in steps of 0.5°C/min.

In all cases, the known selectivity of PNAs for recognition of RNA over DNA was maintained in the hybridization properties of the PNA-APNA chimeras (Table 4.1).[27] This discrimination between RNA and DNA is particularly important for DNA-PNA chimeras and their applications in antisense technology, since a key approach to halting the flow of genetic information is through activation of RNase H and degradation of the target RNA molecules.[108]

Once the APNA monomer derived from glycine (**43**) was identified as the most promising lead structure in this series, it was important to confirm that the decrease in the melting temperatures observed with chimera **144** (ΔT_m = -15°C and -11°C for hybridization with poly rA and poly dA, respectively), as compared to the PNA control,

was not due to lack of base pairing between the APNA thymine residue and its complementary adenosine residue.[109] Examination of the stability of the triplexes formed between a hexamer containing an APNA monomer with the wrong base for Watson-Crick hydrogen bonding (e.g. hexamer **150**) with poly rA and poly dA revealed a much lower hybridization affinity than that observed with the fully complementary chimera **144**. Therefore, the effects of a cytosine-adenosine mismatch, resulting from the incorporation of a non-complemetary APNA unit (Table 4.1), strongly suggests that the APNA monomer of hexamer **144** is indeed participating in Watson-Crick and Höogsteen base pairing with the complementary nucleobase of poly rA and poly dA.

4.3 - Hybridization of (APNA-PNA)$_2$:DNA or (APNA-PNA)$_2$:RNA Triplexes.

The hybridization affinity for the formation of triplex structures between a control PNA decamer (**130**) and complementary DNA or RNA was examined by T_m measurements and compared to those of APNA-PNA chimeras containing 1, 2 or 4 APNA modifications (Table 4.2, decamers **165-167**). The hybridization affinities of the PNA control, as well as the APNA-PNA chimeras with the complementary antiparallel RNA strand, were always greater than those observed with the complementary antiparallel DNA strands, consistent with the known properties of PNAs (Table 4.3). It should be noted that although the APNA-PNA chimeras **165-167** gave complexes with both DNA and RNA that were thermally less stable than those of the control PNA decamer **130**, they were significantly more stable than the corresponding DNA:DNA (T_m

= <5°C) and DNA:RNA (T_m = 11°C) complexes under the same conditions. The destabilization observed with complexes involving chimera **165** was well within agreement with previous studies described in Section 4.2 on the binding of this class of APNA analogues in a triplex binding mode. Moreover, since APNA-PNA decamer **167** was found to hybridize equally or more favorably than the APNA-PNA decamers **165** and **166**, it seemed that multiple insertions of APNA monomers into PNA oligomers may be well tolerated. Job plots confirmed that the stoichiometry of binding of the PNA strands and APNA-PNA chimeras with DNA or RNA was 2:1 (Figure 4.2) and the T_m values showed a dependence on the pH of the buffer solution (Figure 4.3), indicating that the complexes formed were most likely triplexes involving Watson-Crick and Höogsteen base pairing.

Table 4.2: Sequences Used in Triplex Binding Experiments.[a]

Compound	**Sequence**
130	Ac-ATCATTCTCT-Lys-NH_2
165	Ac-ATCAT-$\mathbf{T}_\alpha$-CTCT-Lys-NH_2
166	Ac-ATCA-$\mathbf{T}_\alpha\mathbf{T}_\alpha$-CTCT-Lys-$NH_2$
167	H- $\mathbf{A}_\alpha\mathbf{T}_\alpha\mathbf{C}_\alpha\mathbf{A}_\alpha$-TTCTCT-Lys-$NH_2$
174	5'-dAGAGAATGAT-3'
175	5'-rAGAGAAUGAU-3'

Notes: [a] APNA Residues with Subscript "α"and bold text, PNA residues in normal text.

Table 4.3: T_m Data for Mixed Sequence Triplexes.

	T_m[a] (ΔT_m)			
	Antiparallel PNA:DNA **174**		Antiparallel PNA:RNA **175**	
130	34		51	
165	27	(-7)	40	(-11)
166	26	(-8)	35	(-16)
167	29	(-5)	38	(-13)

Notes: [a] PNA:DNA or PNA:RNA molar ratio was 1:1. Solutions were 4-5μM in both PNA and DNA or RNA. Samples were heated from 5 to 95 °C and/or cooled from 90 to 5 °C at a rate of 0.5°C/min and the absorbance at 260nm was monitored as a function of temperature. T_m values are the maxima of the first derivative plots of the absorbance versus temperature data. Buffer conditions: 150mM NaCl, 10mMNaH_2PO_4, 0.1mM EDTA, pH = 7.0.

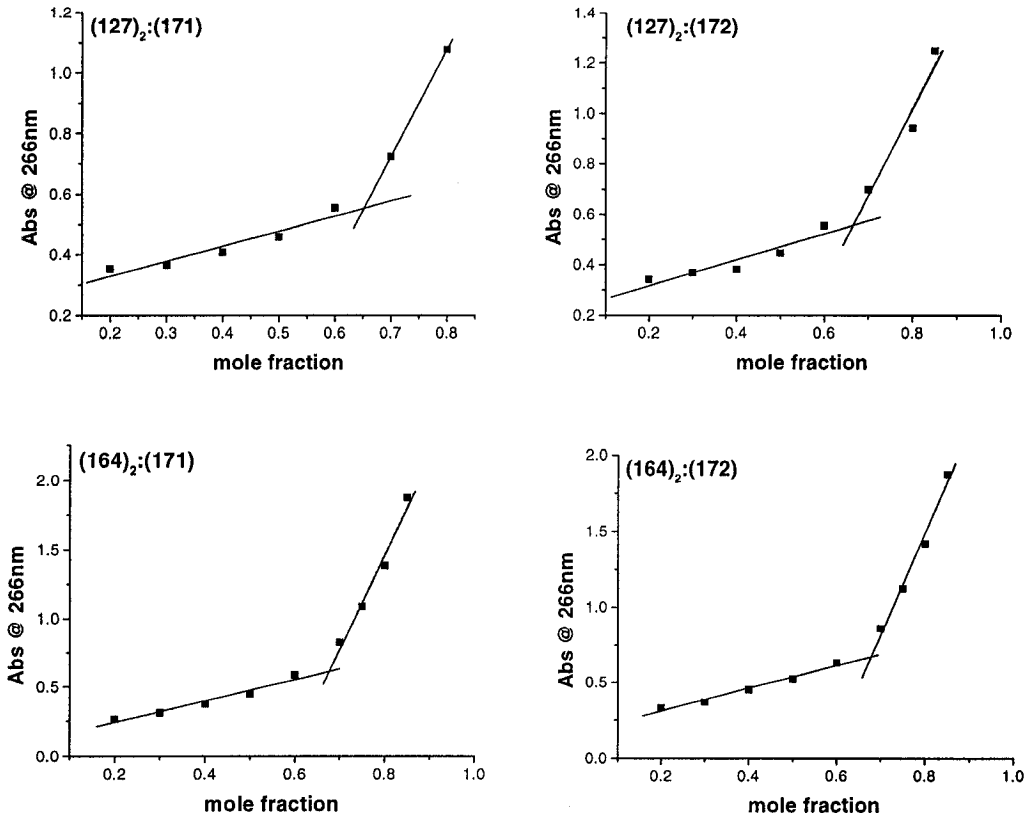


Figure 4.2: Job Plots for Complexes **130** and **167** with DNA and RNA.

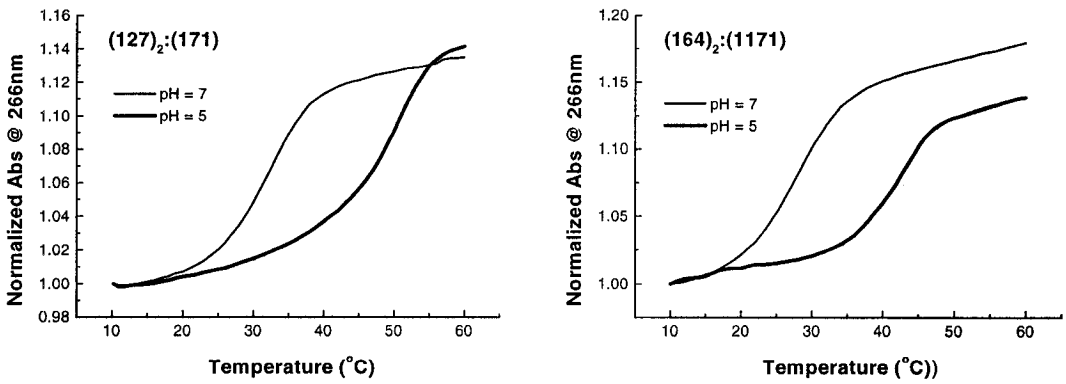


Figure 4.3: T_m curves for complexes of compounds **130** and **167** with DNA **174** at pH = 5 and 7.

4.4 - Hybridization of APNA-PNA:DNA or APNA-PNA:RNA Duplexes.

Thermal denaturation experiments were conducted with APNA-PNA chimeras designed to form duplex structures with DNA and RNA. As indicated previously (Chapter 1), pyrimidine rich PNA sequences form triplex structures with complementary DNA and RNA and the intermediate duplex is not observable. Therefore, a previously described base sequence composed of 1:1 pyrimidine:purine content was chosen in order to favor duplex formation and obtain some hybridization data that could be compared to published results (i.e. oligomer **131**, Table 4.4).[45,50,110] The PNA control decamer **131** gave T_m values of 54 °C (antiparallel, **131:177**) and 39 °C (parallel, **131:178**) with DNA (Table 4.5).[111] As expected, the corresponding RNA complexes gave higher T_m values of 58 °C (antiparallel, **131:179**) and 42 °C (parallel, **131:180**). Surprisingly, the degree of destabilization observed upon introduction of a single APNA unit in the middle of an antiparallel PNA:DNA duplex (Table 4.5, **169:177**) was far greater than that previously observed for a PNA_2:DNA triplex of the same length (Table 4.3, **165_2:174**). Subsequently, the effects of longer APNA inserts into the APNA-PNA chimeras were examined. Incorporation of two APNA units (decamer **170**) lead to further destabilization of the complexes formed with DNA (**170:177** and **170:178**). A similar effect on the stability of the duplexes formed was also observed with RNA (**170:179** and **170:180**). However, chimera **171**, composed of four APNA units at the N-terminal, respectively, did not exhibit the same degree of destabilization as chimera **170** relative to the control PNA decamer **131** (Table 4.5). These results suggest (albeit do not prove) that the destabilization observed upon insertion of a APNA monomer into PNA

oligomers may be due to an incompatibility between the two types of backbones that are being linked together. Furthermore, oligomers composed of multiple consecutive APNA units may benefit from some cooperative intramolecular interactions between the APNA monomers, which may favor hybridization. It should be noted that well-defined melting curves were obtained for all decamer duplexes reported in Table 4.5. These melting curves were distinctly different from those observed by measuring the UV of each corresponding single strand as a function of temperature. In cases where the T_m of the transition that was assigned to the melting of the PNA:DNA or PNA:RNA duplex was very similar to an apparent transition of the single strand alone, the genuine transitions corresponding to melting of the duplexes was confirmed by recording the T_m experiments as a function of both increasing and decreasing temperature (Figure 4.4). The single strands characteristically show pronounced hysteresis, whereas the duplexes did not, suggesting that the sigmoidal curves obtained with the PNA:DNA or PNA:RNA mixtures are due to the melting of the expected duplexes. In addition, the melting curves observed for duplexes of a PNA strand corresponding to the PNA portion of the APNA-PNA chimera **171** (i.e. hexamer **176**) with the same complementary DNA or RNA decamers were all clearly different and gave significantly lower T_m values than decamer **171**.

Table 4.4: Sequences Used in Duplex Binding Experiments.[a]

Compound	Sequence
131	H-GTAGATCACT-Lys-NH_2
169	H-GTAGA-$\mathbf{T}_\alpha$-CACT-Lys-NH_2
170	H-GTAG-$\mathbf{A}_\alpha\mathbf{T}_\alpha$-CACT-Lys-$NH_2$
171	H-$\mathbf{G}_\alpha\mathbf{T}_\alpha\mathbf{A}_\alpha\mathbf{G}_\alpha$-ATCACT-Lys-$NH_2$
176	H-ATCACT-Lys-NH_2
177	5'-dAGTGATCTAC-3'
178	5'-dCATCTAGTGA-3'
179	5'-rAGUGAUCUAC-3'
180	5'-rCAUCUAGUGA-3'
181	5'-dAGTGA-**A**-CTAC-3'
182	5'-dAGTG-**T**-TCTAC-3'
183	5'-dAGTGATC-**A**-AC-3'

[a] APNA Residues with Subscript "α"and bold text, PNA residues in normal text. DNA mismatched residue in bold text.

Table 4.5: T_m Data for Mixed Sequence Duplexes.

	T_m[a] (ΔT_m)							
	PNA:DNA				PNA:RNA			
	Antiparallel		Parallel		Antiparallel		Parallel	
	177		**178**		**179**		**180**	
131	54		39		58		42	
169	40	(-14)	34	(-5)	48	(-10)	31	(-11)
170	30	(-24)	26	(-13)	38	(-20)	27	(-15)
171	43	(-11)	28	(-7)	45	(-13)	28	(-14)
176	24		<10		26		20	

Notes: [a] PNA:DNA or PNA:RNA molar ratio was 1:1. Solutions were 4-5μM in both PNA and DNA or RNA. Samples were heated from 5 to 95 °C and/or cooled from 90 to 5 °C at a rate of 0.5°C/min and the absorbance at 260nm was monitored as a function of temperature. T_m values are the maxima of the first derivative plots of the absorbance versus temperature data. Buffer conditions: 100mM NaCl, 10mM NaH_2PO_4, 0.1mM EDTA, pH = 7.0.

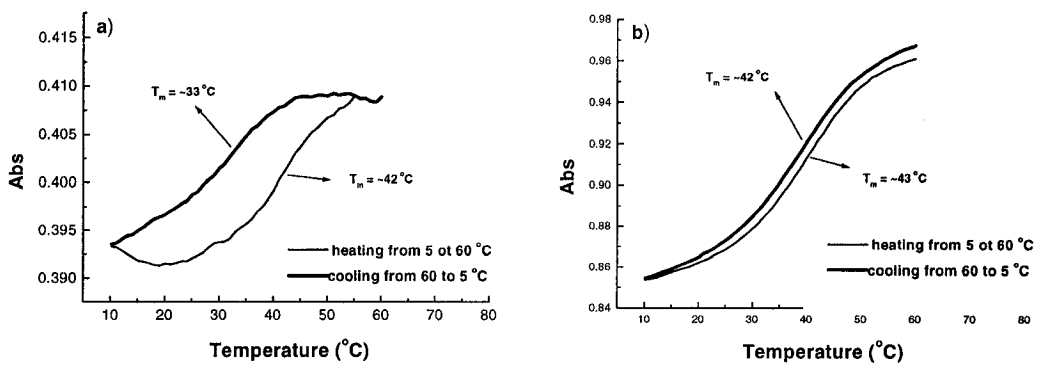


Figure 4.4: T_m Curves as a Function of Increasing and Decreasing Temperature for a) APNA-PNA Chimera **171** Alone, and; b) Duplex **171:177**.

The destabilization observed upon incorporation of APNA units into the PNA decamers is unlikely due to the inability of the APNA monomers to recognize their complementary bases and more likely due to the yet unoptimized conformation and/or distances between units at the PNA---APNA---PNA junctions. In support of this assumption, the introduction of single point mismatches in the complementary DNA strands led to a far greater decrease in the T_m values of the corresponding duplexes than those observed upon insertion of an APNA unit at the same location as the mismatch (Table 4.6). The difference in thermal stability observed for duplex **131:177** compared to duplex **131:181**, **131:182** or **131:183** (ΔT_m = -10 to -20°C) is consistent with the drop in T_m reported previously for a single mismatch in a PNA:DNA duplex. Furthermore, duplexes involving oligomers **169-171** that contained mismatches opposite an APNA residue were of lower thermal stability relative to their fully complementary complexes. These results strongly support the conclusion that APNA monomers participate in Watson-Crick interactions and contribute to the overall sequence specific recognition and binding of the APNA-PNA decamers **169-171** with complementary natural oligonucleotides.

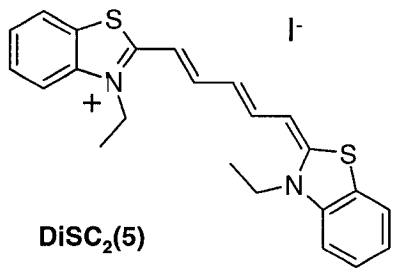
Table 4.6: T_m Data for Mixed Sequence Duplexes Containing Single Mismatches.

	T_m[a] (ΔT_m)[b]					
	181		**182**		**183**	
131	34	(-20)	36	(-18)	44	(-10)
169	27	(-13)	*n.d*[c]		*n.d*	
170	*n.d*		17	(-13)	*n.d*	
171	*n.d*		*n.d*		38	(-5)

Notes: [a] PNA:DNA or PNA:RNA molar ratio was 1:1. Solutions were 4-5μM in both PNA and DNA or RNA. Samples were heated from 5 to 95 °C and/or cooled from 90 to 5 °C at a rate of 0.5°C/min and the absorbance at 260nm was monitored as a function of temperature. T_m values are the maxima of the first derivative plots of the absorbance versus temperature data. Buffer conditions: 100mM NaCl, 10mM NaH_2PO_4, 0.1mM EDTA, pH = 7.0. [b] ΔT_m is relative to the corresponding fully matched duplex. [c] *n.d.*- not determined.

Complex formation between compounds **131**, **169-171** and oligonucleotides **177** and **179** to form duplex structures was also confirmed using the molecular recognition methodology developed by Armitage and coworkers.[112] Upon addition of the cyanine dye 3-ethyl-2-[5-(3-ethyl-3H-benzothiazol-2-ylidene)-penta-1,3-dienyl]-benzothiazol-3-ium iodide (**$DiSC_2(5)$**, **184**) to a duplex or triplex composed in whole or in part of PNA oligomers, the dye binds in a spontaneous and cooperative fashion to the minor groove of the complex as a highly ordered aggregate. The binding is accompanied by a hypsochromic shift of the main visible absorption band of the dye allowing for convenient detection of PNA containing duplexes or triplexes. The absorption spectra of antiparallel[113] duplexes **131:177** and **131:179** are given in Figures 4.5a and 4.5b, respectively. In the case of the PNA:DNA complex **131:177**, the expected blue shift of ~115 nm was observed upon cooling a solution containing **131:179** and **$DiSC_2(5)$** from 40°C to 7°C.[114] Furthermore, similar spectral changes were observed with the PNA:RNA

complex **131:179**. This is the first documented example of complex formation between **DiSC$_2$(5)** and a PNA:RNA duplex, which expands the scope of this methodology. The absorption spectra of antiparallel duplexes **170:177** and **170:177** are given in Figures 4.5c and 4.5d, respectively. The data shown in Figures 4.5c and 4.5d support the conclusion that introduction of two APNA units into the middle of a PNA decamer does not seriously perturb the three dimensional topography of the complex, since the binding of **DiSC$_2$(5)** to PNA:DNA duplexes has been shown to be sensitive to conformational distortions of the duplex.[112] Similar results were obtained with all the complexes formed between compounds **131**, **169-171** and oligonucleotides **177** and **179**.[115]



DiSC$_2$(5)

184

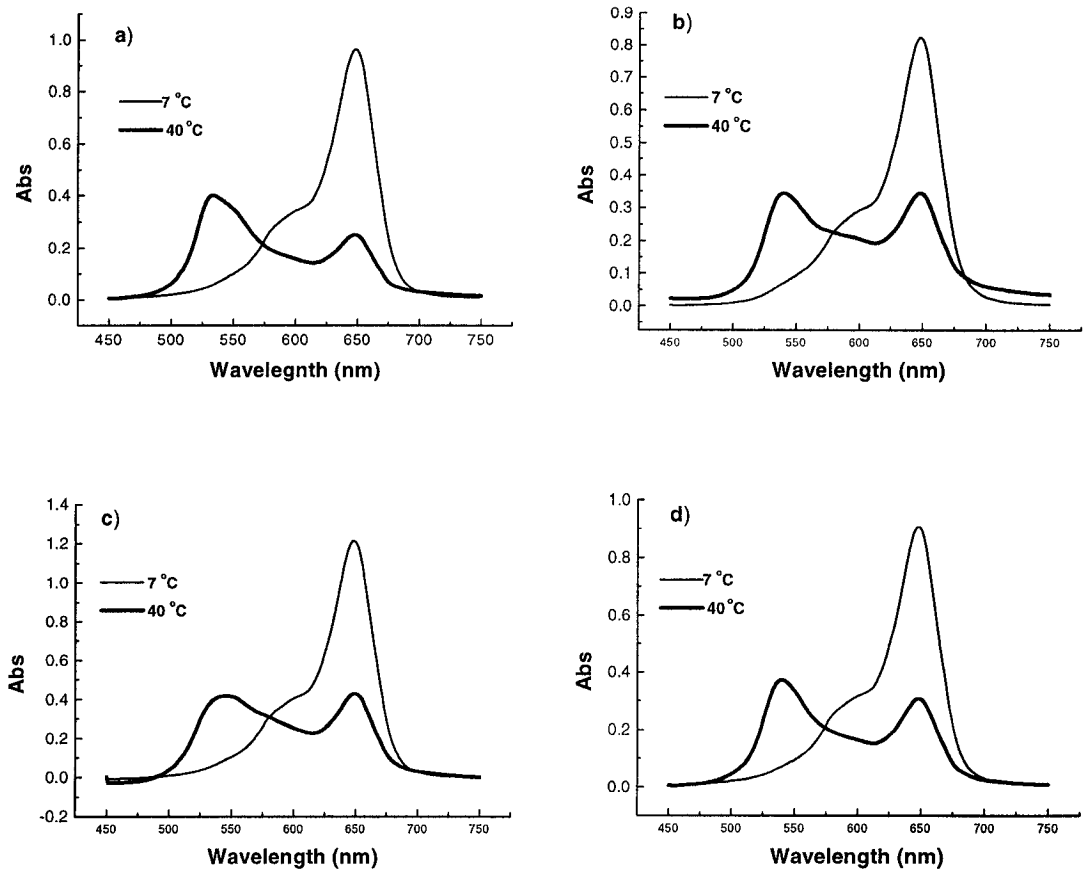


Figure 4.5: a) UV-vis Spectra of **DiSC$_2$(5)** Added to Duplex **131:177**; b) UV-vis Spectra of **DiSC$_2$(5)** Added to Duplex **131:179**; c) UV-vis Spectra of **DiSC$_2$(5)** Added to Duplex **170:177**; d) UV-vis Spectra of **DiSC$_2$(5)** Added to Duplex **170:179**.

4.5 – Conclusions.

Thermal denaturation experiments involving hexamers **129**, **144-150** revealed that APNA monomers represented by structure **43** and **45** exhibited good binding properties with DNA and RNA. Further studies with the most promising monomer **43** showed that this APNA analogue displayed good sequence recognition for DNA and RNA, most likely through Watson-Crick and Höogsteen base pairing. In general, it appears that continuous stretches of APNA monomers are well tolerated in both duplex and triplex binding modes. The binding of APNA modified oligomers was verified using the cyanine dye **184**. The fact that the dye binds to the APNA-PNA:DNA and APNA-PNA:RNA complexes suggests that the minor groove of duplexes involving APNA is of similar topology to the unmodified PNA complexes. The current results are very promising and have provided a new lead structure of peptide nucleic acid analogues having unique physicochemical properties.

4.6 – References.

[104] Kim, S. K.; Nielson, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477.

[105] Job, P. *Ann. Chim. (Paris)* **1928**, *9*, 113.

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

[107] Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1939.

[108] Seitz, O. *Angew. Chem. Int. Ed.* **1999**, *38*, 3466.

[109] Hyrup, B.; Egholm, M.; Nielson, P. E.; Wittung, P.; Nordén, B; Buchardt, O. *J. Am. Chem. Soc.* **1994**, *116*, 7964.

[110] In some cases, the N-terminal of the oligomers are capped with an acetate unit in order to prevent acyl transfer of the last nucleobase-carboxymethyl moiety. However, this measure was not taken with chimeras **131**, **169-171** and **176** since the sequences studied in ref. 106 were not capped in this way. For comments on the influence a positively charged N-terminal has on the stability of PNA:DNA complexes, see: Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895.

[111] The T_m value given for this PNA sequence in reference 106 was 50 °C in the case of the antiparallel complex.

[112] Smith, J. O.; Olson, D. A.; Armitage, B. A. *J. Am. Chem. Soc.* **1999**, *121*, 2686.

[113] Only the antiparallel complexes were examined for binding to **DiSC$_2$(5)**. The authors of ref. 112 have demonstrated that the dye binds with very low affinity to parallel PNA complexes.

[114] The dye aggregates bound to PNA:DNA duplexes give T_m values of ~29 °C. See ref. 112 for complete experimental details.

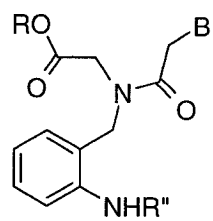
[115] In a series of control experiments, the UV-vis spectra of compounds **131, 169-171** in the presence of **DiSC2(5)** were recorded at 40°C and then 7°C, in the absence of complimentary oligonucleotide **177** or **179**. No significant change in the spectra was observed, indicating that a $(\mathbf{DiSC_2(5)})_n$:ssAPNA-PNA aggregate was not responsible for the observed spectral changes.

CHAPTER 5

RESULTS AND DISCUSSION: PHYSICOCHEMICAL PROPERTIES OF APNA HOMOPOLYMERS.

5.1 – Objectives.

Preliminary binding results showed that oligomers containing APNA monomer **43** were capable of recognizing complementary DNA and RNA residues. Therefore, it was of significant interest to evaluate the binding properties of APNA homopolymers. Decamer **168** and hexamer **172** (Table 5.1) were synthesized in order to compare the binding properties of APNA homopolymers to those of PNA oligomers and APNA-PNA chimeras, such as oligomers **129** and **130** and **144**, **165-167**, respectively (Tables 3.1 and 3.4). The modified hexamer **173** was also synthesized and used in these studies; the N-terminal aspartic acid residue was added to this oligomer in order to improve aqueous solubility.



43

Table 5.1: APNA Homopolymers Prepared for Binding Experiments.[a]

Compound	Sequence
168	H- $\mathbf{A_\alpha T_\alpha C_\alpha A_\alpha T_\alpha T_\alpha C_\alpha T_\alpha C_\alpha T_\alpha}$ -Lys-NH_2
172	Ac-$\mathbf{T_\alpha T_\alpha T_\alpha T_\alpha T_\alpha T_\alpha}$ -Lys-NH_2
173	H-Asp-$\mathbf{T_\alpha T_\alpha T_\alpha T_\alpha T_\alpha T_\alpha}$ -Lys-NH_2

[a] APNA Residues are indicated in bold with the subscript "α"; PNA residues in normal text.

5.2 – Aqueous Solubility of APNA homopolymers.

In order to evaluate the binding properties of fully modified APNA oligomers, compounds **168** and **172** were prepared as described in Chapter 3. Unfortunately, saturated solutions of these oligomers in various aqueous buffers gave near zero absorbance at 260 nm, indicating they were essentially insoluble under the conditions used for the hybridization studies. PNA oligomers having an aspartate residue at N-terminal were previously synthesized in order to enhance aqueous solubility without altering the hybridization properties of these compounds.[46] Similarly, the APNA hexamer **173** was found to be much more soluble in water than the equivalent hexamer **172** (Table 5.1), allowing us to evaluate the hybridization properties of APNA homopolymers with DNA and RNA.

5.3 – Hybridization of Hexamer 173 to poly dA and poly rA.

Thermal denaturation experiments were conducted following the usual protocol of first mixing the APNA oligomer **173** with poly dA or poly rA at high temperatures and then allowing the sample to cool so that the complex could form between the synthetic oligomer and the complementary DNA or RNA strand. Surprisingly, the UV absorbance *vs* temperature plots did not exhibit clear sigmoidal melting profiles. This result was somewhat unexpected given that APNA monomer **43** was previously shown to recognize DNA and RNA in a sequence selective way in both the triplex and duplex binding modes (Chapter 4). Ambiguous T_m curves are not uncommon with complexes involving PNAs.[45] In those cases, circular dichroism spectropolarimetry (CD) has proven to be a valuable tool in assessing the stability of PNA:DNA hybrids.[45]

Complex formation between oligomer **173** and poly dA or poly rA was investigated by CD. Upon addition of hexamer **173** to a solution of poly dA, significant changes in the CD spectra were observed (Figure 5.1). Maxima at 266 nm and 230 nm and minima at 248 nm and 205 nm were consistent with previously reported triplexes involving PNAs.[98] The CD spectrum of each complex was clearly different from that obtained from the addition of the CD spectrum of oligomer **173** alone and poly dA alone. Saturation of the maxima at 266 nm and minima at 300 nm occurred at approximately a 2:1 ratio, indicating that a triplex was formed between hexamer **173** and poly dA. The temperature dependence of the CD spectrum was also examined in order to evaluate the stability of the complex formed between hexamer **173** and poly dA (Figure 5.2). As the temperature of the solution was increased, clear changes were observed in the CD

spectrum of **173**:poly dA (Figure 5.2a) and a T_m value of 23°C could be assigned to the complex (Figure 5.2b). This value is much lower than that of the corresponding PNA_2:DNA triplex (**129_2**:poly dA, T_m = 56°C, Table 4.1). However, it should be noted that a T_m value could not be determined for the corresponding DNA T_6 hexamer under the same buffer conditions and only at high salt concentrations (1M NaCl) a T_m value of 17°C could be measured.

Unfortunately, only a minor change of the CD spectrum was observed upon addition of hexamer **173** to a solution of poly rA. While the reasons for this result remain unclear, a plausible explanation is that the backbone of poly rA is much more rigid than that of DNA, thus more resistant to the conformational changes required to allow triplex formation with the APNA homopolymer at 7°C. Therefore, higher temperatures may be required in order for the proper conformational alignment to occur for annealing of the two oligomers.

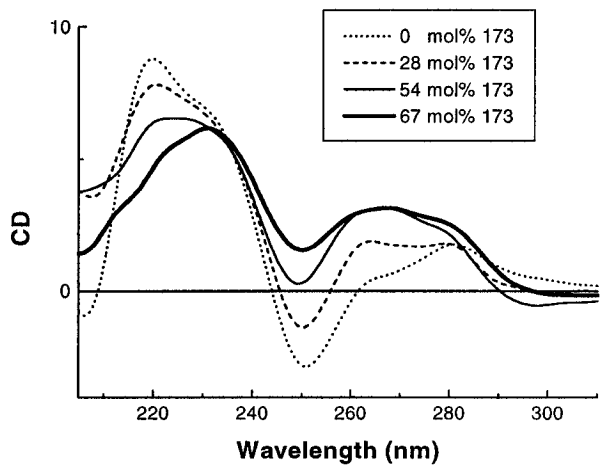


Figure 5.1: CD Spectra of the complex between APNA Homopolymer **173** and poly dA. Buffer conditions: 10 mM $NaHPO_4$, 1 mM EDTA, pH = 7.

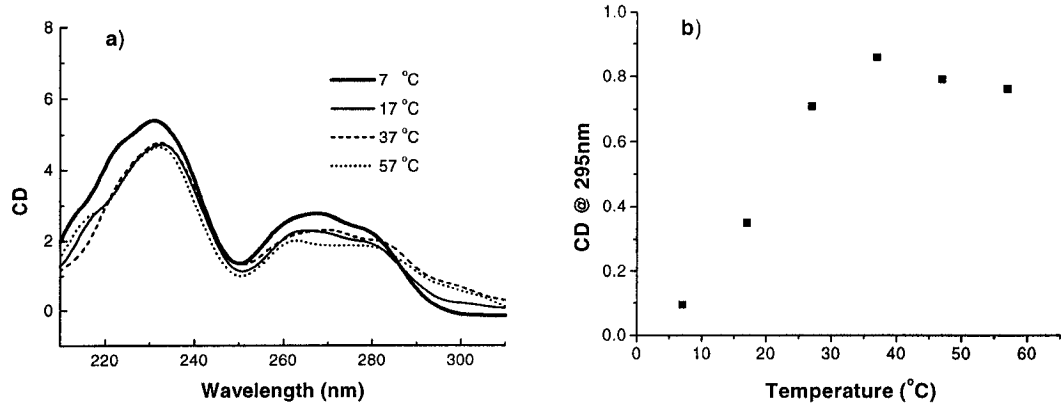


Figure 5.2: a) Temperature Dependance of the CD Spectra of **173**:poly dA. b) CD at 295 nm *vs* Temperature Plot. The same buffer conditions were used as in Figure 5.1.

The binding of the APNA hexamer **173** was also investigated using binding experiments with the dye **DiSC$_2$(5)**, similar to those described in Chapter 4. When the dye was added to the pre-cooled solutions containing hexamer **173** and poly dA or poly rA, a blue shift of the main visible absorption band for the dye was observed in the UV-visible spectrum with a new maximum centered at 550 nm (Figure 5.3). Surprisingly, control experiments revealed that the dye also bound to the poly dA and poly rA, giving blue shifts of the λ_{max} from 625 nm to 525 nm (Figure 5.3). These results are in contrast to previously reported observations with short oligomers composed of less than twenty bases and polymeric systems similar to those used in the present study.[112] To the best of our knowledge, this is the first time that binding of the dye has been observed with a single-stranded DNA or RNA (also refer to Chapter 4). It is noteworthy that the λ_{max} of this complex occurs at 525 nm, a blue shift of 120 nm from the unbound dye. This suggests that the aggregate formed by the dye on the single-stranded template is a higher

order aggregate similar to those characteristic of PNA:DNA duplexes (λ_{max} = 539 nm),[112] rather than that of the dimeric species that assembles on a DNA:DNA template (λ_{max} = 590 nm).[116]

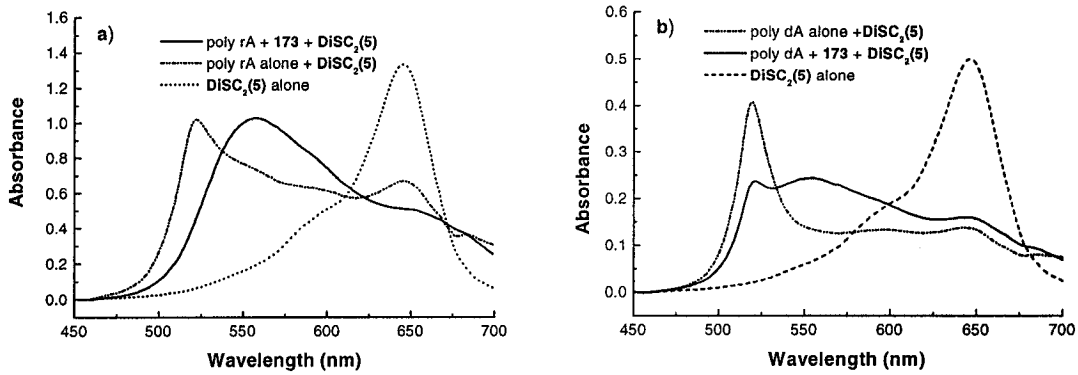


Figure 5.3: UV-vis Spectra of a) **$DiSC_2(5)$** alone, **$DiSC_2(5)$** poly rA and **$DiSC_2(5)$ + 173** + poly rA; and b) **$DiSC_2(5)$** alone, **$DiSC_2(5)$** poly dA and **$DiSC_2(5)$ + 173** + poly dA.

The effect of temperature on the UV-visible spectrum for the **$DiSC_2(5)$**:poly dA and the **$DiSC_2(5)$**:poly rA was examined in order to further characterize these unusual complexes. As the temperature of the solution of each complex was increased, the band at 525 nm gradually decreased in magnitude, while that of the monomeric form of the unbound dye at 645 nm increased in magnitude (Figures 5.4a and 5.4b). In each case, the plot of UV absorbance *vs* temperature gave a clear sigmoidal curve with T_m values of ~52°C and ~53°C for the DNA and RNA complexes respectively (Figure 5.5b and 5.5b). These T_m values are for the dissociation of **$DiSC_2(5)$** from the single-stranded poly dA and poly rA, respectively. Interestingly, these values are 20-25°C higher than those

reported for the complexes of the dye with PNA:DNA and DNA:DNA complexes.[112,116,117]

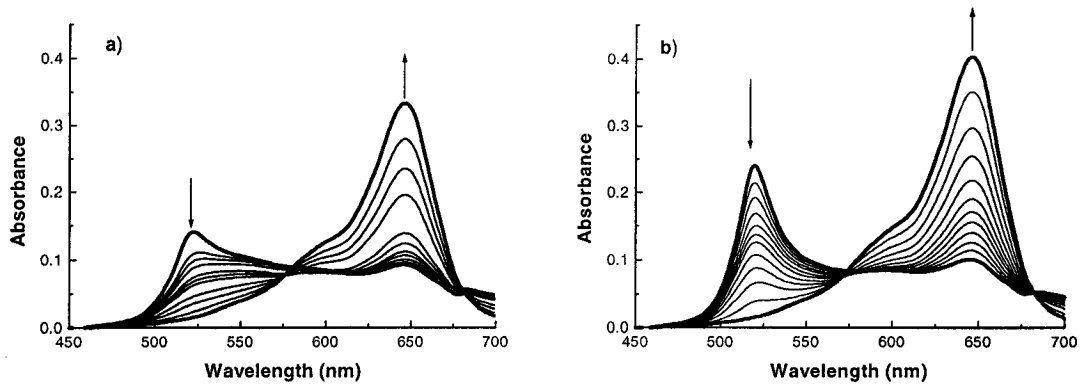


Figure 5.4: a) Temperature dependence of the UV-vis Spectra of poly rA + **DiSC$_2$(5)**; Temperature dependence of the UV-vis Spectra of poly dA + **DiSC$_2$(5)**.

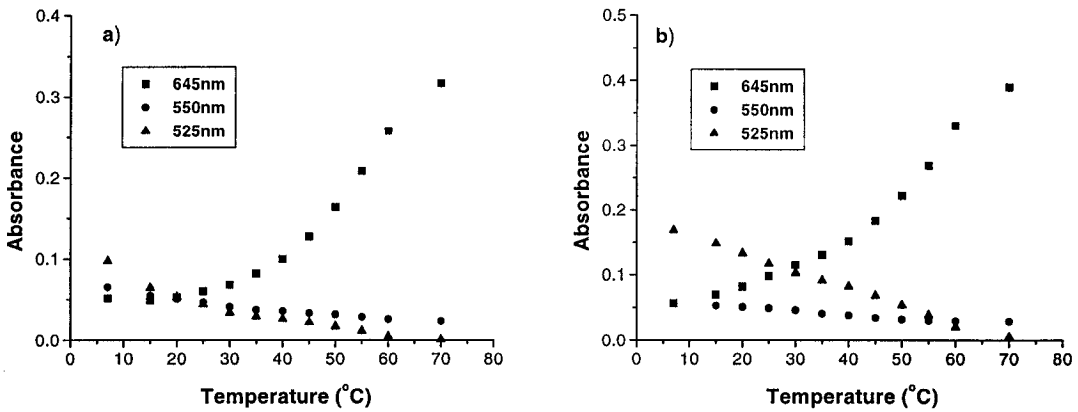


Figure 5.5: Absorbance *vs* Temperature Plots at 645, 550 and 525 nm for a) poly rA + **DiSC$_2$(5)**; and b) poly dA and **DiSC$_2$(5)**.

The effect of temperature on the UV-visible spectrum of the **DiSC$_2$(5)**:APNA:DNA and **DiSC$_2$(5)**:APNA:RNA complexes was also examined (Figure 5.6 and 5.7). As the temperature of the solution containing the APNA:DNA complex was increased, the absorption maximum at 550 nm gradually decreased in magnitude, while the band for the **DiSC$_2$(5)** complex with DNA or RNA alone at 525 nm first gradually increased and then decreased at higher temperatures near the T_m values (Figure 5.6 and 5.7, curved arrow indicates this observation). As the band at 525 nm began to decrease, an increase in the monomeric band at 645 nm was also observed. These observations indicate that there may be an equilibrium between three different states of the dye in solution: the unbound monomeric form of the dye (λ_{max} = 645 nm), the dye bound to the APNA:DNA or APNA:RNA complex as a higher order aggregate (λ_{max} = 550 nm), or the dye bound to the single-stranded poly dA or poly rA as a higher order aggregate (λ_{max} = 525 nm).

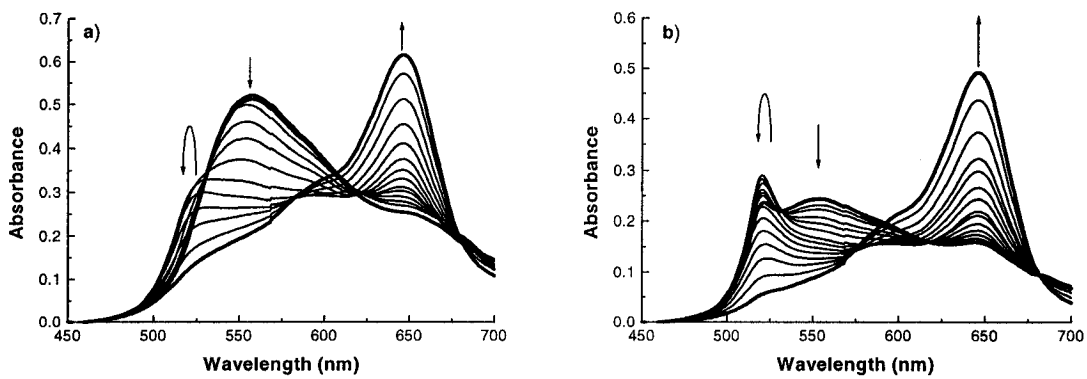


Figure 5.6: a) Temperature dependence of the UV-vis Spectra of Compound **173** + poly rA + **DiSC$_2$(5)**; b) Temperature dependence of the UV-vis Spectra of Compound **173** + poly dA + **DiSC$_2$(5)**.

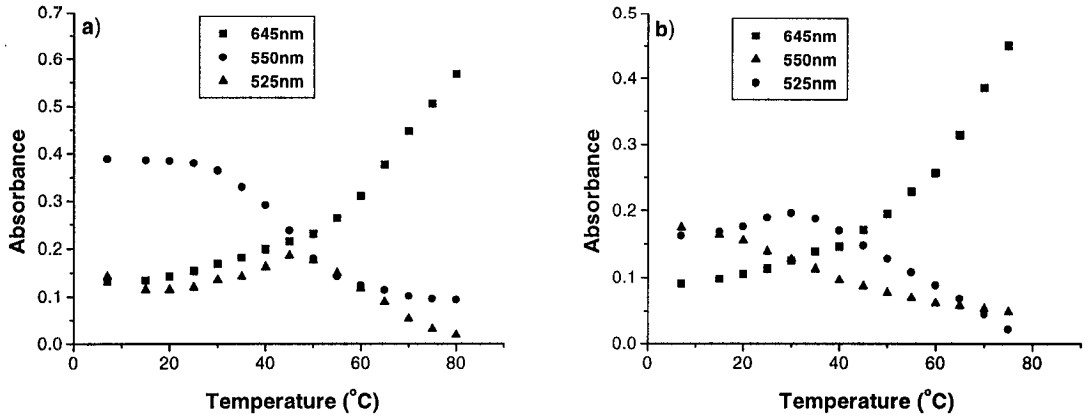


Figure 5.7: Absorbance *vs* Temperature Plots at 645, 550 and 525nm for a) Compound **173** + poly rA + **DiSC$_2$(5)**; and b) Compound **173** + poly dA and **DiSC$_2$(5)**.

In an attempt to determine if the recognition event between hexamer **173** and poly dA or poly rA was mediated by Watson-Crick and/or Höogsteen base pairing, similar experiments were conducted with the fully mismatched RNA molecule poly rU. When **DiSC$_2$(5)** was added to a solution containing compound **173** and poly rU, a blue shift was observed that was identical to that observed when the dye was added to poly rU alone. Furthermore, the absorbance *vs* temperature plots for this mixture indicated the presence of only the **DiSC$_2$(5)**:poly rU and monomeric **DiSC$_2$(5)**. These observations strongly suggest that hexamer **173** binds to poly rA (and most likely poly dA) in a sequence selective manner, most likely through Watson-Crick and Höogsteen base pairing.

A number of reports have already described the use of CD spectropolarimetry to study the complexes formed between **DiSC$_2$(5)** and DNA:DNA[116,117] or PNA:DNA[112] duplexes. In the case of PNA:DNA duplexes, assembly of the dye in the minor groove of the complex results in a bisignant CD signature with a positive Cotton effect at 558 nm, a

negative Cotton effect at 529 nm and isoelliptic point at 544 nm.[112] Addition of **DiSC$_2$(5)** to a solution containing either **173**:poly dA or **173**:poly rA had a similar effect on the CD spectrum in the region of 450-700 nm (Figure 5.8). However, when the dye was added to a solution of poly dA or poly rA alone, the results obtained were distinctly different. For example, addition of the dye to poly dA gave positive Cotton effects at 513 nm and 544 nm, while the **DiSC$_2$(5)**:**173**:poly dA complex gave a weak positive Cotton effect at around 616 nm, a strong Cotton effect around 547 nm with an isoelliptic point at 575 nm. These results indicate that, in each case, there are discretely different complexes present and that the dye molecules of the **DiSC$_2$(5)**:**173**:poly dA complex are likely assembled in a right-handed helical manner similar to that of the dye bound to PNA:DNA duplexes.

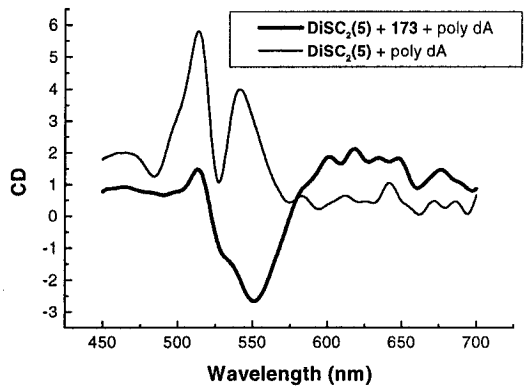


Figure 5.8: CD Spectra of **DiSC$_2$(5)**:**173**:poly dA and **DiSC$_2$(5)**:poly dA.

5.4 – Conclusions.

Complex formation between fully modified APNA homopolymer **173** was confirmed by CD and **DiSC$_2$(5)** binding experiments. These studies strongly suggest that the APNA homopolymer binds to DNA and RNA, apparently through Watson-Crick and/or Höogsteen base pairing, and form a complex with DNA that is more stable than that formed with the corresponding homothymine DNA hexamer with complementary DNA under similar conditions.

5.5 – References.

[116] Garoff, R. A.; Litzinger, E. A.; Connor, R. E.; Fishman, I.; Armitage B. A. *Langmuir* **2002**, *18*, 6330.

[117] Seifert, J. L.; Connor, R. E.; Kushon, S. A.; Wang, M; Armitage, B. A. *J. Am. Chem. Soc.* **1999**, *121*, 2987.

CHAPTER 6

CONTRIBUTIONS TO KNOWLEDGE

6.1 – Summary of Thesis.

Synthetic methods were developed for the preparation of several structurally diverse APNA monomers. These analogues vary in the position of the aromatic ring and nucleobase and in the substitution pattern along the backbone. Monomers containing N-phenyl-N-alkyl amide moieties are acyclic PNA analogues whose tertiary amide rotamer equilibrium heavily favours the conformation preferred by unmodified PNAs when hybridized to DNA and RNA. This set of monomer building blocks is useful for the preparation of oligomers designed to evaluate the hybridization properties of a novel class of peptide based DNA mimics.

Thermal denaturation experiments involving hexamers **129**, **144-150** revealed that APNA monomers represented by structure **43** and **45** exhibited good binding properties with DNA and RNA. Further studies with the most promising monomer **43** showed that this APNA analogue displayed good sequence recognition for DNA and RNA, most likely through Watson-Crick and Höogsteen base pairing. The binding of APNA modified oligomers was verified using the cyanine dye **184**. These studies suggested that the local environment of the minor groove of duplexes involving APNA is similar to the unmodified PNA complexes. The current results are very promising and have provided a

new lead structure of peptide nucleic acid analogues having unique physicochemical properties.

Finally, fully modified APNA homopolymers were prepared in order to investigate their physicochemical properties. The short homothymine hexamer **173** was synthesized which was modified at the C- and N-terminals so that it can be dissolved in water in sufficient quantities to allow for evaluation of its recognition and binding to DNA and RNA. Complex formation was confirmed by CD and **DiSC$_2$(5)** binding experiments. These studies indicated that the APNA homopolymer bound to DNA and RNA, apparently through Watson-Crick base pairing, and formed a complexe that was more stable than those formed by the corresponding homothymine DNA oligomer with complementary DNA under similar conditions.

6.2 - Publications Resulting from Thesis Work.

As a direct result of the work described in this thesis, the following publications have appeared:

1. L. D. Fader and Y. S. Tsantrizos. Hybridization Properties of Aromatic Peptide Nucleic Acids: A Novel Class of Oligonucleotide Analogues. *Organic Letters* **2002**, *4*, 63-66.

2. L. D. Fader, M. Boyd, Y. S. Tsantrizos. Backbone Modifications of Aromatic Peptide Nucleic Acid (APNA) Monomers and their Hybridization Properties with DNA and RNA. *Journal of Organic Chemistry* **2001**, *66,* 3372-3379.

*Note: M. Boyd's contribution to this work was the preparation of monomer **47**, synthesis of dimers **140** an **141** and their incorporation into hexamers for binding studies, as indicated in references 83 and 94.*

6.3 – Conference Proceedings Resulting from Thesis Work.

The work described in this thesis has been presented in the following conference proceedings:

1. L.D. Fader and Y. S. Tsantrizos (June 2001). Synthesis of Novel Peptidic DNA Mimics. Bioorganic Gordon Research Conference, Andover, New Hampshire.

2. L.D. Fader, Y. S. Tsantrizos, (2001). Synthesis of Novel Peptidic DNA Mimics. 84th Canadian Society for Chemistry Conference and Exhibition, Montreal, QC.

3. Y. S. Tsantrizos, M. Boyd, L.D. Fader, J. Mancuso, E. Myers, J.F. Lunetta (2001). Aromatic Peptide Nucleic Acids: A Paradigm of Nucleic Acids. 84th Canadian Society for Chemistry Conference and Exhibition, Montreal, QC.

4. Y.S. Tsantrizos, L.D. Fader (2000). Synthesis and Physicochemical Properties of a Novel Class of Peptide Nucleic Acids. Pacifichem 2000 Meeting, Honolulu, Hawaii.

5. L.D. Fader, Y. S. Tsantrizos, M. Boyd (2000). Peptide Modulators of Gene Expression. 3rd Annual Chemistry Biochemistry Graduate Research Conference, Concordia University, Montreal, QC.

6. L.D. Fader, Y. S. Tsantrizos, M. Boyd (1999). Peptide Modulators of Gene Expression. 10th Annual Quebec-Ontario Minisymposium in Synthetic and Bioorganic Chemistry, Saint Saveur, QC.

7. L.D. Fader, Y. S. Tsantrizos, M. Boyd (1999). Peptide Modulators of Gene Expression. 82nd Canadian Society for Chemistry Conference and Exhibition, Toronto, Ont.

8. Y.S. Tsantrizos, J.F. Lunetta, L.D. Fader, M. Boyd and J. Mancuso (1998) Aromatic Peptide Nucleic Acids: Novel Oligonucleotide Analogs. 81st Canadian Society for Chemistry Conference and Exhibition, Wistler, B.C.

CHAPTER 7

EXPERIMENTAL SECTION.

7.1-General Methods.

Solvents were purchased from Fischer Scientific and purified as follows: THF was distilled from sodium/benzophenone ketyl; CH_2Cl_2 distilled from P_2O_5 or CaH_2; DMF treated with KOH overnight at RT, then vacuum distilled from CaO or BaO and stored over activated 4Å molecular sieves; MeCN distilled from CaH_2; pyridine distilled from CaH_2. HPLC solvents were HPLC grade and were filtered through 0.45μm filters (Supelco, Bellefonte, PA) prior to use. HATU was purchased from PerSeptive Biosystems Ltd. MBHA resin was purchased from Nova Biochem Ltd. All other starting materials and reagents were purchased from Sigma/Aldrich Canada and were used without further purification, except for DIPEA and Et_3N, which were refluxed over CaH_2 and then distilled and stored over activated 4Å molecular sieves. Thin-layer chromatography was carried out on aluminum-backed silica gel 60 F_{254} plates (EM Science, Germany) using the solvent systems indicated. Reversed-phase thin layer chromatography was carried out on glass-backed RP-18 F_{254}s plates (EM Science, Germany) using the solvent systems indicated. Oligonucleotides **139**, **146, 147** and **150-152** were purchased from UCDNA Lab (University of Calgary, Calgary, AB) and were used as received. Oligonucleotides **140**, **148** and **149** were purchased from Dharmacon

Research, Inc. (Lafayette, CO), deprotected following the procedure recommended by the manufacturer and used without further purification. Polydeoxyriboadenylate and polyriboadenylate were purchased from Sigma.

7.2-NMR Spectroscopy.

Deuterated NMR solvents were purchased from Isotec Inc. (Miamisburg, OH). NMR spectra were obtained at ambient temperature unless otherwise indicated. ^{1}H and ^{13}C NMR chemical shifts are quoted in ppm and are referenced to the internal deuterated solvent. A mixture of rotamers was often observed by NMR; in those cases the ratio is indicated and the signals are denoted as major (ma) and minor (mi). Some NMR spectra were also recorded at the elevated temperature indicated in parenthesis, in order to confirm the presence of rotamers. Variable temperature experiments were performed using the VT accessory on a JEOL 270 MHz spectrometer by heating the sample slowly over ca. 5 min to the desired temperature followed by a minimum equilibration time of 5 min. All ^{1}H NMR spectra were recorded on a Varian Mercury (300 or 400 MHz), a Unity Inova (300 or 500 MHz) or a JEOL spectrometer (270 MHz) Spectrometer. ^{13}C NMR spectra were recorded on a JEOL spectrometer (67.5 MHz) or Varian Mercury spectrometers (75 or 100 MHz).

7.3-UV Spectroscopy and Thermal Denaturation Experiments.

Melting temperatures of duplexes were determined on a Cary 3 UV-vis spectrophotometer. Solutions were 4-5μM in $(PNA)_n$:DNA or $(PNA)_n$:RNA. All T_m solutions were 10 mM in NaH_2PO_4, 1 mM in EDTA and the pH was adjusted to 7.1 or 5.0 as indicated. In some cases, as indicated in the Results and Discussion, T_m buffers containing 150 or 100 mM NaCl were used. The solutions were heated to 90°C for 10 min then cooled slowly to 4°C and stored at that temperature for at least 1 h but usually overnight. Samples were heated from 5 to 95°C and/or cooled from 90 to 5°C at a rate of 0.5°C/min and the absorbance at 260 nm was monitored as a function of temperature. In all cases, T_m experiments and UV spectra were recorded in quartz cuvettes purchased from HELMA Canada (matching 281, 282, or 284). T_m values are the maxima of the first derivative plots of the absorbance *versus* temperature data.

7.4-Circular Dichroism Spectropolarimetry.

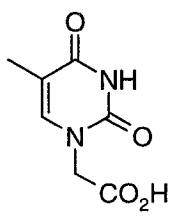
Circular dichroism spectra were recorded on a Jasco J-710 spectropolarimeter equiped with a NESLAB RTE-111 cooling unit and jacketed cell holder. Duplex or triplex solutions were preannealed as described in Section 7.3 prior to measurements. For T_m determination by CD, the sample was heated in steps of 5°C and equilibrated at the desired temperature for at least 5 min. Spectra are the average of at least 3 acquisitions that were recorded at a scan rate of 50 nm/min and a step resolution of 0.2

nm for the range of 200-320 nm and 100 nm/min at 0.5 nm resolution for the range of 450-700 nm.

7.5-Synthesis of PNA Monomers.

All PNA monomers used in this thesis were synthesized using literature procedures. In cases of exceptions to this, alternate procedures are given below.

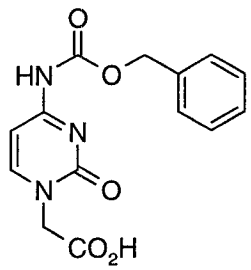
Synthesis of Thymin-1-yl acetic acid (51):



51

Thymine (18.9g, 150mmol) and K_2CO_3 (20.7g, 150mmol) were suspended in anhydrous DMF (300mL) under N_2 with vigorous stirring. To the mixture was added ethyl bromoacetate (25.0g, 150mmol) and the reaction was allowed to stir overnight. The reaction mixture was then filtered, evaporated to dryness and taken up in H_2O (200mL) and acidified by addition of 4M HCl (60mL). After stirring for 30 min, the solid was filtered off and washed with water (2 x 200mL). The filter cake was then suspended in H_2O (200mL) and 2M NaOH (200mL) was added. The solution was boiled for 20 minutes and then allowed to cool to approximately 45°C. The pH was then adjusted to 2 and cooling was continued slowly to 0°C. The resulting white crystals were collected by filtration and dried overnight, *in vacuo*, over P_2O_5 giving 14.3g (52%) of pure product.

Synthesis of N^4-benzyloxycarbonylcytisin-1-yl acetic acid (53):

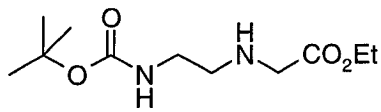


52

A suspension of cytosin-1-yl acetic acid ethyl ester **63** (5.14g, 26.1mmol) and DMAP (1.87g, 15.3) in dry pyridine (90mL) was placed under N_2 with stirring and cooled to 0°C. Benzylchloroformate (12.0g, 100.0mL, 69.9mmol) was added dropwise to the mixture over 15 h and the reaction was allowed to warm to RT and stir overnight. The solvent was removed under reduced pressure and 1M HCl (aqu.) (80mL) was added. After stirring for 10 min, the white solid was collected by filtration, washed with cool 1M HCl (2 x 20mL) and the wet filtercake used directly in the next step without further purification.

N^4-benzyloxycarbonylcytosine acetic acid ethyl ester **61** from the above step was suspended in H_2O (50mL). $LiOH.H_2O$ (3.6g, 78mmol) in H_2O (30mL) was then added to the suspension and the reaction was allowed to stir for 1 h (Note the pH of the reaction should be checked at the beginning to ensure the hydrolysis will proceed). The pH was adjusted to 2 and the resultant precipitate collected by filtration and dried over P_2O_5, *in vacuo*, to give 4.40g (57% over 2 steps) of the desired product. ^{1}H NMR (DMSO-d_6) δ: 4.49 (s, 2H), 5.15 (s, 2H), 6.98 (d, 1H), 7.25-7.46 (m, 5H), 7.99 (d, 1H), 10.79 (br s, 1H).

Synthesis of Ethyl N-[2-(tert-butoxycarbonyl amino)ethyl]glycinate (120):

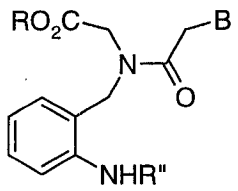


120

To a solution of compound **122** (18.5g, 115mmol) in 1:1 $CHCl_3/CH_3CN$ (500mL) was added K_2CO_3 (2.18g, 15.8mmol). This mixture was placed under N_2 with stirring and ethyl bromoacetate (2.401g, 14.4mmol) dissolved in $CHCl_3$ (100mL) was added dropwise over 7 h. The reaction was allowed to proceed for an additional 2 h and then filtered and concentrated to dryness. The residue was taken up in ~100mL EtOAc and filtered through silica gel (4cm x 6cm diameter). The silica gel was washed twice with EtOAc (2 x 200mL) and the combined fractions evaporated to dryness to give a crude mixture that was primarily the desired amine **120**. The silica gel was then washed with 1:1 MeOH/$CHCl_3$ (1.5L) to recover the amine starting material **122,** which was pure by ^{1}H NMR (16.4g, 102mmol). The product was purified by flash column chromatography (100% EtOAc, 4 x 21cm silica) to give 2.47g (96%) of the desired secondary amine as a yellowish oil. ^{1}H NMR ($CDCl_3$) δ: 1.26 (t, 3H, ^3J = 7.2 Hz), 1.42 (s, 9H), 2.42 (br, 2H), 2.76 (t, 2H, ^3J = 5.7 Hz), 3.22 (q, 2H, ^3J = 5.7), 3.41, (s, 2H), 4.16 (q, 2H, ^3J = 7), 4.07 (br, 1H).

7.6 - Synthesis of APNA Monomers.

Synthesis of N-(2-aminobenzyl)-N-(thymin-1-ylacetyl)glycine methyl ester (43a, R = Me, R" = H, B = Th):



43a-d

Compound **72a** (434mg, 1.1mmol) was dissolved in anhydrous DMF (3mL) and Pd-C (65mg) was added. Triethylamine (0.6mL, 4.3mmol) and formic acid (0.13mL, 3.3mmol) were added and the reaction mixture was stirred at RT for 4 h. EtOH (10mL) was added and the mixture filtered through celite and concentrated to dryness. The resulting white solid was dried *in vacuo*, giving 388mg (97% yield) of compound **43a** which was determined to be pure by ^{1}H NMR.

TLC R_f (EtOAc) = 0.213.

^{1}H NMR (400MHz, DMSO-d_6, mixture of rotamers in 1:1.5 ratio) δ: 1.75 (*mi*) & 1.76 (*ma*) (d, 3H, J = 1.2 Hz), 3.59 (*mi*) & 3.62 (*ma*) (s, 3H), 3.97 (*mi*) & 4.20 (*ma*) (s, 2H), 4.37 (*ma*) & 4.50 (*mi*) (s, 2H), 4.56 (*ma*) & 4.61 (*mi*) (s, 2H), 5.01 (*mi*) & 5.06 (*ma*) (bs, 2H), 6.47 (dt, 0.6H, J = 6.3 & 1.2 Hz), 6.60 (m, 0.9H, *J* = 7.0 & 1.2 Hz), 6.69 (dd, 0.5H, *J* = 7.6 & 1.2 Hz), 6.96-7.01 (m, 2H), 7.35 (*ma*) & 7.40 (*mi*) (d, 1H, *J* = 1.2 Hz), 11.29 (*mi*) & 11.32 (*ma*) (bs, 1H).

^{13}C NMR (67.5MHz, DMSO-d_6, 130°C) δ: 12.0, 48.0, 48.6, 52.2, 109.0, 116.3, 117.2, 119.5, 129.0, 130.0, 142.3, 147.1, 151.5, 164.6, 168.7, 169.6.

FAB$^+$ HRMS (glycerol) *m/z* found 361.1512 $(M+H)^+$; calculated for $(C_{17}H_{21}N_4O_5)^+$ = 361.1511.

Synthesis of N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(thymin-1-ylacetyl)glycine (43a, R = H, R" = Fmoc, B = Th):

Compounds **43a-d** were synthesized using the following procedure outlined for carbamate **43a**: *t*-Butyl ester **83a** (2.24g, 3.59mmol) was dissolved in 2:1 TFA/CH_2Cl_2 (35mL) and stirred at RT for 2 h. The solution was then rotovapped to dryness, diluted with PhMe (50mL) and again concentrated to dryness. This was repeated two more times giving pure product as a white powder (>98% yield). ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.68 (*ma.*) and 1.72 (*mi.*) (s, 3H), 3.87 (*mi.*) and 4.16 (*ma.*) (s, 2H), 4.25 (*ma.*) and 4.26 (*mi.*) (t, 1H, *J* = 7.2Hz), 4.34 (*ma.*) and 4.44 (*mi.*) (d, 2H, J = 7.2Hz), 4.52, (s, 2H), 4.59 (*ma.*) and 4.61 (*mi.*) (s, 2H), 6.83-7.90 (m, 14H), 9.17 (*mi.*) and 9.27 (*ma.*) (br 1H), 11.28 (*mi.*) and 11.30 (*ma.*) (br, 1H), 12.70 (*mi.*) and 13.09 (*ma.*) (br, 1H). ^{13}C NMR (75MHz, DMSO-d_6) δ: 12.54, 47.25, 47.37, 47.76, 48.35, 48.64, 48.90, 66.51, 66.69, 108.68, 108.91, 120.76, 120.82, 125.19, 125.82, 126.00, 126.59, 127.24, 127.77, 128.34, 128.60, 129.43, 131.36, 136.12, 136.89, 141.40, 141.46, 142.74, 142.93, 144.38, 151.65, 151.71, 154.73, 155.13, 165.05, 165.12, 168.43, 169.39, 170.76, 170.95. FAB+ HRMS (glycerol) *m/z* calculated for $C_{31}H_{28}N_4O_7 + H^+$: 569.2036, found: 569.2035.

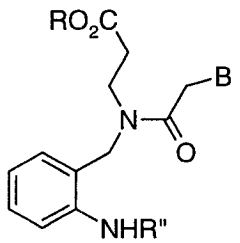
Analytical Data for Free Acid **43b** (R = H, R" = Fmoc, B = Cy^{Cbz}): Yield: >98% ^{1}H NMR (400MHz, DMSO-d_6) δ: 3.88 (*mi.*) and 4.22 (*ma.*) (s, 2H), 4.18-4.37 (m, 1H), 4.33 (*ma.*) and 4.44 (*mi.*) (d, 2H, J = 6.4Hz), 4.54 (*ma.*) and 4.65 (*mi.*) (s, 2H), 4.70 (*mi.*) and 4.75 (*ma.*) (s, 2H), 5.15 (*ma.*) and 5.18 (*mi.*) (s, 2H), 6.98-7.43 (m, 17H), 7.49 (*mi.*) and

7.70 (*ma.*) (br, 1H), 7.87 (*ma.*) and 7.90 (*mi.*) (d, 1H, J = 7.6Hz), 7.96 (*mi.*) and 8.00 (*ma.*) (d, 1H, J = 7.6Hz), 9.25 (br, 1H).

Analytical Data for Free Acid **43c** (R = H, R" = Fmoc, B = Ad^{Cbz}):: Yield: >98%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 3.91 (*mi.*) and 4.32 (*ma.*) (s, 2H), 4.16-4.22 (m, 1H), 4.31 (*ma.*) and 4.46 (*mi.*) (d, 2H, J = 6.4Hz), 4.55 (*ma.*) and 4.77 (*mi.*) (s, 2H), 5.19 (*ma.*) and 5.22 (*mi.*) (s, 2H), 5.21 (*ma.*) and 5.31 (*mi.*) (s, 2H), 7.12-7.70 (m, 17H), 7.86 (*ma.*) and 7.89 (*mi.*) (d, 1H, J = 8.0Hz), 8.38-8.43 (m, 1H), 8.57-8.61 (m, 1H), 9.22 (br, 1H), 10.71 (br, 1H).

Analytical Data for Free Acid **43d** (R = H, R" = Fmoc, B = Gu^{Cbz}):: Yield: 100%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 3.89-4.71 (m, 7H), 5.01-5.29 (m, 4H), 7.10-7.94 (m, 18H), 9.23 (br, 1H), 11.30-11.60 (m, 2H).

Synthesis of N-(2-aminobenzyl)-N-(thymin-1-ylacetyl)-β-alanine methyl ester (44a, R = Et, R' = H, B = Th):



44a

Compound **44a** was synthesized using the same procedure as that described for **43a** (R = Me, R" = H):

Yield: 80 %. TLC R_f (EtOAc) = 0.130.

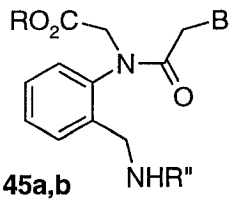
^{1}H NMR (300MHz, DMSO-d_6, mixture of rotamers in 1:2.3 ratio) δ: 1.14 (*mi*) & 1.18 (*ma*) (t, 3H, *J* = 7.1 Hz), 1.74 (*mi*) & 1.75 (*ma*) (bs, 3H), 2.49 (*mi*) & 2.69 (*ma*) (t, 2H, *J*

= 7.5 Hz), 3.42 (*mi*) & 3.48 (*ma*) (t, 2H, *J* = 7.5 Hz), 4.00 (*mi*) & 4.07 (*ma*) (q, 2H, *J* = 7.1 Hz), 4.37 (*ma*) & 4.42 (*mi*) (s, 2H), 4.49 (*mi*) & 4.67 (*ma*) (s, 2H), 4.99 (*mi*) & 5.05 (*ma*) (s, 2H), 6.49 (t, 0.6H, *J* = 6.9 Hz), 6.57-6.61 (m, 1H), 6.67 (d, 0.4 H, *J* = 7.8 Hz), 6.90-7.03 (m, 2H), 7.39 (*mi*) & 7.41 (*ma*) (bs, 1H), 11.28 (bs, 1H).

^{13}C NMR (100MHz, DMSO-d_6) δ: 11.9, 14.0, 32.1, 32.20, 40.91, 42.3, 44.7, 46.5, 48.1, 48.2, 60.0, 60.2, 107.9, 108.0, 114.7, 115.3, 115.7, 116.6, 119.0, 119.3, 126.2, 127.9, 128.4, 129.8, 142.3, 142.4, 146.0, 146.6, 151.1, 164.5, 167.3, 171.0, 171.1.

FAB$^+$ HRMS (glycerol) *m/z* found 389.1825 $(M+H)^+$; calculated for $(C_{19}H_{25}N_4O_5)^+$ = 389.1826.

Synthesis of N-(2-N'-tert-Butoxycarbonylaminomethylphenyl)-N-(thymin-1-ylacetyl)glycine (45a, R = H, R" = Boc, B = Th) and N-(2-N'-tert-Butoxycarbonylaminomethylphenyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl)glycine (45b, R = H, R" = Boc, B = Ad^{Cbz}):



45a,b

Compounds **45a,b** were prepared using the following general procedure: A solution of $NaClO_2$ (8 equivalents) and $NaHPO_4$ (5 equivalents) in H_2O (3.5M in $NaClO_2$) was added to a solution of aldehyde **88a,b** dissolved in a 1:8:4 mixture of 2-methyl-2-butene/tBuOH/THF (0.08M). After complete consumption of the aldehyde as determined by TLC, the reaction was partitioned between aqueous NaOH (0.5M) and EtOAc. The aqueous phase was then acidified to pH = 3 by addition of HCl (3M) and then extracted

with EtOAc. The organic layer was then dried over anhydrous $MgSO_4$ and concentrated to a white foam. Residual H_2O and acetic acid were removed by suspending the product in toluene and concentrating to dryness.

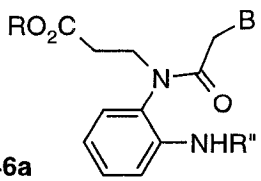
Analytical Data for Acid **45a**: ^{1}H NMR (300MHz, DMSO-d_6) δ: 1.37 (s, 9H), 1.71 (m, 3H), 3.81-4.63 (m, 6H), 7.13-7.55 (m, 5H), 11.30 (br, 1H). ^{13}C NMR (75MHz, DMSO-d_6) δ: 11.90, 28.16, 48.28, 50.41, 78.06, 107.78, 125.09, 127.98, 128.67, 128.79, 129.08, 129.16, 137.82, 138.02, 141.70, 150.60, 155.69, 164.07, 166.59.

ES^+ MS *m/z*: 469.2 $(M+H)^+$; ES^- MS *m/z*: 445.2 $(M-H)^-$.

Analytical Data for Acid **45b**: ^{1}H NMR (300MHz, DMSO-d_6) δ: 1.38 (s, 9H), 3.90 (d, 1H, J = 17.1Hz), 4.19 (d, 1H, J = 15.9Hz), 4.35 (d, 1H, J = 15.9), 4.62 (d, 1H, J = 16.8Hz), 4.69 (d, 1H, J = 17.4Hz), 4.90 (d, 1H, 17.1Hz), 5.20 (s, 2H), 7.32-7.62 (m, 9H), 8.30 (s, 1H), 8.57 (s, 1H), 10.65 (s, 1H). ^{13}C NMR (75MHz, $CDCl_3$) δ: 28.26, 39.43, 44.81, 51.11, 67.37, 79.71, 120.20, 128.12, 128.38, 129.10, 130.22, 130.64, 135.45, 137.37, 138.10, 144.52, 149.36, 150.91, 151.31, 152.52, 156.47, 166.21, 171.29.

ES^+ MS *m/z*: 590.3 $(M+H)^+$; ES^- MS *m/z*: 588.3 $(M-H)^-$.

Synthesis of N-(2-N'-*tert*-Butoxycarbonylaminophenyl)-N-(thymin-1-ylacetyl)-β-alanine methyl ester (46a, R = Me, R" = Boc, B = Th):

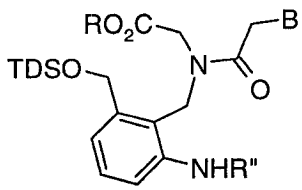


46a

Secondary aniline **92** (40mg, 0.136mmol) and thymine derivative **51** (50mg, 0.27mmol) were dissolved in anhydrous DMF (500μL). To the solution was added EDC (52mg,

0.27mmol) and the reaction was allowed to stir overnight under N_2. The reaction mixture was diluted with EtOAc (5mL) and extracted with saturated $NaHCO_3$ (2 x 5mL) and H_2O (1 x 5mL). The organic layer was dried over $MgSO_4$ and then concentrated to dryness. The product was isolated as a white foam after purification by silica gel chromatography (EtOAc as eluant) in 49% yield. ^{1}H NMR (300MHz, DMSO-d_6) δ: 1.53 (s, 9H), 1.91 (m, 3H), 2.48-2.56 (m, 1H), 2.60-2.67 (m, 1H), 3.66 (sw, 3H), 3.83-3.90 (m, 1H), 4.06-4.14 (m, 3H), 6.94 (m, 1H), 7.15-7.22 (m, 2H), 7.41-7.45 (m, 1H), 8.00 (br, 1H), 8.05 (d, 1H, J = 7.2Hz). ^{13}C NMR (67.7MHz, DMSO-d_6) δ: 12.22, 28.23, 32.09, 44.56, 49.46, 51.98, 81.18, 110.40, 123.41, 124.69, 129.38, 130.00, 130.05, 135.65, 141.32, 151.32, 153.31, 164.76, 167.41, 172.01.

Synthesis of {[2-Amino-6-(tert-butyl-diphenyl-silanyloxymethyl)-benzyl]-[2-(5-methyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-acetyl]-amino}-acetic acid ethyl ester (48a, R = Et, R" = H, B = Th).



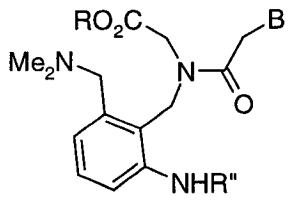
48a

To solution of nitro derivative **102** (25mg, 0.04mmol) in 1:1 DMF/EtOH (1mL) was added 5mg of Pd-C (10% w/w). The flask was then flushed with N_2 followed by H_2. The reaction was allowed to stir under a H_2 atmosphere for 2.5 h and then diluted with EtOH (5mL) and filtered through celite. The celite was washed with EtOH (1mL) and the combined filtrates evaporated to dryness. The crude residue was then dissolved in EtOAc

(5mL) and washed with saturated $NaHCO_3$ (aqueuos, 3 x 5mL). The organic layer was then dried with Na_2SO_4 and concentrated to dryness. The product was obtained as a glassy solid after flash column chromatography (1-2% MeOH/DCM) in 63% yield. ^1H NMR (400MHz, $(CD_3)_2SO$) δ: 0.98-1.05 (m, 12H), 1.74 (s, 3H), 3.60 (*mi.*) and 3.96 (*ma.*) (s, 2H), 3.81 (*mi.*) and 3.89 (*ma.*) (q, 2H, J = 7.2Hz), 4.50 (*ma.*) and 4.56 (*mi.*) (s, 2H), 4.52 (*ma.*) and 4.74 (*mi.*) (s, 2H), 4.65 (*ma.*) and 4.74 (*mi.*) (s, 2H), 5.05 (br, 2H), 6.60-7.62 (m, 14H), 11.31 (*ma.*) and 11.32 (*mi.*) (br, 1H). ^{13}C NMR (75MHz, $(CD_3)_2SO$) δ: 12.56, 14.36, 19.44, 27.18, 46.54, 48.57, 61.47, 64.35, 108.61, 115.00, 115.15, 115.82, 128.26, 129.07, 130.27, 133.19, 133.30, 135.32, 135.38, 140.87, 140.96, 142.29, 148.20, 151.28, 151.84, 164.73, 168.32, 168.41, 169.04.

ES^+ MS *m/z*: 665.3 $(M+Na)^+$; ES^- MS *m/z*: 641.2 $(M-H)^-$.

Synthesis of {(2-Amino-6-dimethylaminomethyl-benzyl)-[2-(thymin-1-yl)-acetyl]-amino}-acetic acid ethyl ester (49a, R = Et, R" = H, B = Th):



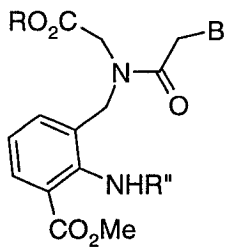
49a

To a solution of compound **107** (363mg, 0.79mmol) dissolved in DMF (4mL) was added Pd-C (40mg). The flask was then sealed and flushed with H_2. The reaction was allowed to stir under a H_2 atmosphere until the precise moment the starting material had been completely consumed (~3 h). The reaction was then diluted with EtOH and filtered through celite. The residue was then purified by flash column chromatography (5-15%

MeOH/$CHCl_3$) to give aniline **49a** in 81% yield. ^{1}H NMR (300MHz, $(CD_3)_2SO$) δ: 1.09 (mi) and 1.19 (*ma.*) (t, 3H, J = 6.9Hz), 1.77 (s, 3H), 2.06 (br, 6H), 3.17 (*ma.*) and 3.26 (*mi.*) (s, 2H), 3.69 (*mi.*) and 4.12 (*ma.*) (s, 2H), 3.95 (*mi.*) and 4.09 (*ma.*) (q, 2H, J = 6.6Hz), 4.50 (*ma.*) and 4.68 (*mi.*) (s, 2H), 4.57 (*ma.*) and 4.82 (*mi.*) (s, 2H), 4.95 (br, 2H), 6.36 (*ma.*) and 6.46 (*mi.*) (d, 1H, J = 6.9Hz), 6.56 (*ma.*) and 6.68 (*mi.*) (d, 1H, J = 8.1Hz), 6.89-6.99 (m, 1H), 7.30 (*ma.*) and 7.42 (*mi.*) (s, 1H), 11.29 (*ma.*) and 11.32 (*mi.*) (br, 1H). ^{13}C NMR (75MHz, DMSO-d_6) δ: 12.54, 14.57, 42.01, 45.17, 45.46, 46.36, 48.58, 61.43, 62.64, 108.41, 108.65, 114.97, 117.04, 119.32, 120.28, 128.44, 139.13, 139.38, 142.33, 142.70, 148.38, 148.83, 151.29, 164.74, 168.94, 169.12, 169.26.

ES^+ MS *m/z*: 453.4 $(M+Na)^+$; ES^- MS *m/z*: 430.3 $(M-H)^-$.

Synthesis of 2-Amino-3-({tert-butoxycarbonylmethyl-[2-(thymin-1-yl)-acetyl]-amino}-methyl)-benzoic acid methyl ester (50a, R = tBu, R" = H, B = Th):



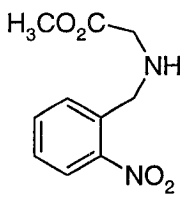
50a

To a solution of compound **114** (578mg, 1.16mmol) dissolved in DMF (5mL) was added Pd-C (75mg) followed by DIPEA (670μL, 3.85mmol) and HCO_2H (120μL, 3.50mmol). After stirring for 2.5 h, the reaction was diluted with EtOH (10x) and filtered through celite. The filtrate was then evaporated to dryness and the residue purified by flash column chromatography (2-5%MeOH/DCM) to give 480mg of aniline **50a** (89%). ^{1}H NMR (400MHz, $(CD_3)_2SO$) δ: 1.25 (*ma.*) and 1.37 (*mi.*) (s, 9H), 1.74 (*mi.*) and (1.76

(*ma.*) (s, 3H), 3.77 (*ma.*) and 3.80 (*mi.*) (s, 3H), 3.88 (*mi.*) and 4.19 (*ma.*) (s, 2H), 4.44 (*ma.*) and 4.51 (*mi.*) (s, 2H), 4.56 (br, 2H), 6.50-6.70 (m, 3H), 7.22-7.40 (m, 2H), 7.69-7.74 (m, 1H), 11.28 (*mi.*) and 11.31 (*ma.*) (br, 1H). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 12.40, 27.75, 28.14, 47.78, 48.01, 48.50, 48.85, 49.29, 52.05, 81.44, 82.20, 108.47, 108.74, 109.74, 114.83, 115.59, 121.43, 130.59, 131.32, 132.52, 136.20, 142.53, 142.75, 149.22, 150.03, 151.47, 164.88, 167.94, 168.52, 169.21.

ES^+ MS *m/z*: 483.1 $(M+Na)^+$; ES^- MS *m/z*: 459.0 $(M-H)^-$.

Synthesis of N-(2-nitrobenzyl)glycine methyl ester (71a):



71a

2-Nitrobenzaldehyde (5.70g, 38mmol), methyl glycinate hydrochloride (4.30g, 34mmol) and triethylamine (5.0mL, 38mmol) were dissolved in dry CH_2Cl_2 (160mL). This mixture was stirred under N_2 for 30 min and then $NaBH(OAc)_3$ was added. The reaction was stirred for 5 h, then quenched by the addition of H_2O (30mL) and the aqueous layer was discarded. The crude product was extracted into 0.1M HCl (100mL) and the water layer was back-extracted with CH_2Cl_2 (2 x 300mL). The pH was adjusted to 4 by drop-wise addition of NaOH (1M) and the product extracted into CH_2Cl_2 (4 x 500mL). The combined organic layers were dried over anhydrous $MgSO_4$ and concentrated to dryness giving the title compound as an orange oil in 68% yield and in high purity. For more convenient storage and handling of the material, compound **71a** was converted

quantitatively to its hydrochloride salt by dissolution of the oil in 1:1 Et_2O/EtOAc (60mL) followed by addition of HCl (1.5 eq., 2M HCl in Et_2O). The white precipitate was collected by filtration and dried *in vacuo*.

TLC R_f (60% hex in EtOAc) = 0.246.

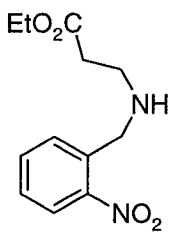
^{1}H NMR (300 MHz, $CDCl_3$) δ: 2.41 (bs, 1H), 3.43 (s, 2H), 3.69 (s, 3H), 4.07 (s, 2H), 7.40 (dt, 1H, J = 8.0 & 1.4 Hz), 7.53-7.62 (m, 2H), 7.93 (d, 1H, J = 8.0Hz).

^{13}C NMR (75MHz, $CDCl_3$) δ: 50.27, 50.36, 51.94, 124.91, 128.23, 131.06, 133.31, 135.12, 149.17, 172.70.

HCl salt of **71a**: ^{1}H NMR (300MHz , CD_3OD) δ: 3.86 (s, 3H), 4.14 (s, 2H), 4.58 (s, 2H), 7.74-7.89 (m, 3H), 8.29 (dd, 1H, J = 8.4 & 1.0 Hz).

FAB$^+$ HRMS (glycerol/KCl) *m/z* found 225.0875 $(M+H)^+$; calculated for $(C_{10}H_{13}N_2O_4)^+$ = 225.0876.

Synthesis of N-(2-nitrobenzyl)β-alanine ethyl ester (71b):



71b

Amine 4b was synthesized using the same procedure as that described for **71b**:

Yield: 70%. TLC R_f (free amine, 1:1 hex / EtOAc) = 0.182.

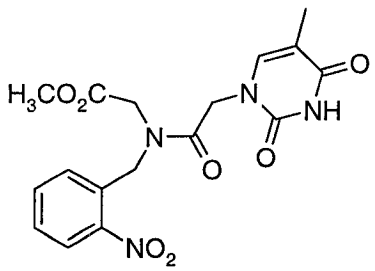
^{1}H NMR (300MHz, $CDCl_3$) δ: 1.25, (t, 3H, J = 6.9 Hz), 3.02 (t, 2H, J = 5.9 Hz), 3.36, (t, 2H, J = 6.4 Hz), 4.17, (q, 2H, J = 7.0 Hz), 4.37, (s, 1H), 7.63, (dt, 1H, J = 7.6 & 1.2 Hz),

7.76 (dt, 1H, J = 7.6 & 1.8 Hz), 8.01 (dd, 1H, J = 6.4 & 1.2 Hz), 8.16 (dd, 1H, J = 7.0 & 1.2 Hz).

^{13}C NMR (100MHz, DMSO-d_6) δ: 14.7, 30.8, 43.3, 47.6, 61.2, 125.9, 127.9, 131.3, 134.0, 134.9, 149.0, 170.6.

FAB$^+$ HRMS (glycerol): *m/z* found 253.1188 $(M+H)^+$; calculated of $(C_{12}H_{17}N_2O_4)^+$ = 253.1188.

Synthesis of N-(2-nitrobenzyl)-N-(thymin-1-ylacetyl)glycine methyl ester (72a):



72a

Thymidyl acetic acid **51** (1.50g, 8.1mmol) and HATU (3.09g, 8.1mmol) were dissolved in anhydrous DMF (20mL) under N_2 and cooled to 0°C. To this solution, DIPEA (1.5mL, 8.6mmol) was added and the reaction mixture was stirred for 10 min at 0°C. Compound **71a** (1.69g, 7.5mmol) was than added as a solution in anhydrous DMF (10mL) and the reaction was allowed to warm-up to RT and stir overnight. The reaction mixture was then concentrated to ~5mL and EtOAc (50mL) was added. This solution was washed with water (3 x 15mL) and the organic layer concentrated to about 10mL and cooled to 0°C. The precipitate formed was collected by filtration and washed once with 10% hexanes in EtOAc giving 2.50g of compound **72a** (85% yield) in high purity. For large-scale synthesis of **72a** (i.e. >20g), it was far more cost effective to use a

DCC/pentafluorophenol pre-activation protocol, however, this procedure gave a lower yield of **72a** (~70%).

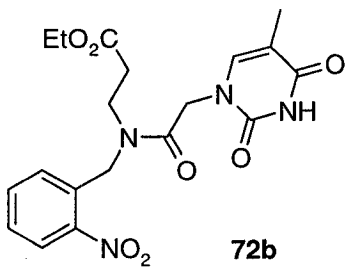
Compound **72a**: TLC R_f (EtOAc) = 0.288.

^{1}H NMR (300MHz, DMSO-d_6, mixture of rotamers in 1:1 ratio) δ: 1.74 & 1.75 (s, 3H), 3.62 & 3.69 (s, 3H), 4.05 & 4.36 (s, 2H), 4.58 & 4.83 (s, 2H), 4.64 & 5.06 (s, 2H), 7.39-7.44 (m, 1.5H), 7.53 (dt, 0.5H, J = 7.0 & 1.2 Hz), 7.58-7.62 (m, 1H), 7.70 & 7.80 (dt, 1H, J = 7.0 & 1.2 Hz), 8.05 & 8.16 (dd, 1H, J = 7.0 & 1.2 Hz), 11.28 & 11.31 (s, 1H). Coalescence of each pair of rotamer resonances was achieved at ~90°C.

^{13}C NMR (67.5MHz, DMSO-d_6, 120°C) δ: 12.0, 48.6, 49.4, 52.4, 109.0, 125.2, 129.1, 129.6, 132.3, 134.2, 142.2, 149.1, 151.5, 164.6, 169.0, 169.6.

FAB$^+$ MS (glycerol): *m/z* 391 $(M+H)^+$.

Synthesis of N-(2-nitrobenzyl)-N-(thymin-1-ylacetyl)-β-alanine methyl ester (72b):



72b

Compound **72b** was synthesized using the same procedure as that described for **72a**:

Yield: 93%. TLC R_f (EtOAc) = 0.226.

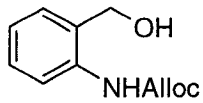
^{1}H NMR (400MHz, DMSO-d_6, mixture of rotomers in 1:1.7 ratio) δ: 1.14 (*mi*) & 1.15 (*ma*) (t, 3H, J = 7.0 Hz), 1.73 (bs, 3H), 2.55 (*mi*) & 2.71 (*ma*) (t, 2H, J = 8.1 Hz), 3.45 (*mi*) & 3.62 (*ma*) (t, 2H, J = 8.1 Hz), 3.99 (*mi*) & 4.01 (*ma*) (q, 2H, J = 7.8 Hz), 4.45 (*mi*) & 4.75 (*ma*) (s, 2H), 4.81 (*ma*) & 5.01 (*mi*) (s, 2H), 7.31-7.37 (m, 1.6H), 7.44 (d, 0.4H, J

= 7.6 Hz), 7.53 (*ma*) & 7.60 (*mi*) (t, 1H, *J* = 7.2 Hz), 7.69 (*ma*) & 7.81 (*mi*) (t, 1H, *J* = 7.6 Hz), 8.06 (*ma*) & 8.17 (*mi*) (d, 1H, *J* = 8.0 Hz), 11.25 (bs, 1H).

^{13}C NMR (75MHz, DMSO-d_6) δ: 12.4, 12.5, 14.5, 32.5, 33.2, 34.2, 35.4, 43.2, 43.4, 46.7, 48.6, 48.7, 49.0, 49.7, 60.6, 60.7, 108.5, 108.6, 125.4, 125.9, 128.0, 128.3, 128.7, 129.1, 133.2, 133.3, 134.3, 135.0, 142.7, 148.2, 148.6, 151.7, 151.9, 165.3, 168.4, 171.6, 171.8.

FAB$^+$ MS (glycerol): *m/z* 419 $(M+H)^+$.

Synthesis of N-(allyloxycarbonyl)-2-aminobenzylalcohol (74):

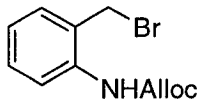


74

2-Aminobenzyl alcohol (**73**, 1.04g, 8.5mmol) was dissolved in 1:1 DCM/pyridine (20mL) and Alloc-Cl (1.12g, 9.3mmol) was added drop-wise *via* syringe over a period of 15 min. After 1.5 h the reaction mixture was diluted with DCM (200mL) and extracted with H_2O (100mL). The organic layer was then dried over anhydrous $MgSO_4$ and concentrated to dryness. Flash column chromatography, using 12% EtOAc in hexanes as the eluent, provided 1.36g pure carbamate **74** (77% yield) as a colorless oil.

^{1}H NMR (300MHz, $CDCl_3$) δ: 2.34 (bd, 1H, *J* = 11.7 Hz), 4.65 (ddd, 2H, *J* = 5.1 & 1.8 Hz), 4.67 (s, 2H), 5.26 (dm, 1H, *J* =10.5 Hz), 5.36 (dm, 1H, *J* = 17.1 Hz), 5.91-6.04 (m, 1H), 7.03 (ddd, 1H, *J* = 7.0 & 1.2 Hz), 7.15 (dd, 1H, *J* = 7.0 & 1.2 Hz), 7.31 (ddd, 1H, *J* = 8.2 & 1.8 Hz), 7.90 (bd, 1H, *J* = 8.1 Hz), 7.97 (bs, 1H).

Synthesis of N-(allyloxycarbonyl)-2-aminobenzyl bromide 75:

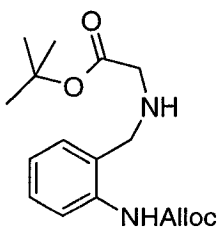


75

Carbamate **74** (1.35g, 3.7mmol) was dissolved in anhydrous THF (50mL). CBr_4 (2.2g, 6.5mmol) followed by PPh_3 (1.70g, 6.5mmol) were added and the reaction mixture was stirred at RT for 2 h. Upon completion of the reaction, the solvent was evaporated to dryness and crude product was re-dissolved in EtOAc (20mL). The precipitate ($OPPh_3$) was removed by filtration and the filtrate was concentrated to dryness. The residue was then purified by flash column chromatography, using 5% EtOAc in hexanes as the eluent, to give bromide **75** as a fluffy white solid (920mg, 58% yield).

^{1}H NMR (400MHz, $CDCl_3$) δ: 4.50 (s, 2H), 4.68 (dt, 2H, J = 5.6 & 1.6Hz), 5.27 (dm, 1H, J = 10.8 Hz), 5.39 (dm, 1H, J = 17.2 Hz), 5.95-6.05 (m, 1H), 6.88 (bs, 1H), 7.11 (ddd, 1H, J = 7.1 & 1.2 Hz), 7.28 (dd, 1H, J = 8.0 & 1.6 Hz), 7.34 (ddd, 1H, J = 7.8 & 1.6 Hz), 7.83 (bd, 1H, J = 7.2 Hz).

Synthesis of N-[N-(allyloxycarbonyl)-2-aminobenzyl]glycine *tert*-butyl ester 76a:



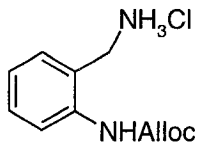
76a

Route A: To a solution of bromide **75** (268mg, 1mmol) in anhydrous DMF (3mL), HCl·GlyOtBu (500mg, 3mmol) and DIPEA (1.0mL, 6.0mmol) were added. The reaction

mixture was stirred overnight at RT under N_2. The crude product was then partitioned between H_2O (12mL) and EtOAc (25mL). The organic layer was washed once with H_2O (12mL) dried over anhydrous $MgSO_4$ and evaporated to dryness. The residue was purified by flash column chromatography, using 10% EtOAc in hexanes as the eluent, to give 293mg of pure amine **76a** (92% yield) as a colorless oil.

Route B: To a solution of ammonium salt **80** (31.3g, 129mmol) and DIPEA (44.9mL, 258mmol) in DMF (900mL) was added *t*-butyl bromoactetate (25.2g, 129mmol) dropwise over 2 h. The reaction was stirred overnight at RT under N_2. The product mixture was then concentrated to 1/3 of the original volume and partitioned between H_2O (900mL) and EtOAc (1000mL). The organic layer was then washed twice with H_2O (900mL) dried over anhydrous Na_2SO_4 and evaporated to dryness. The residue was then purified by flash column chromatography (10% EtOAc/hex, 2 batches) to give pure amine **76a** (64%) as a colorless oil. ^{1}H NMR (400MHz, $CDCl_3$) δ: 1.48 (s, 9H), 3.43 (s, 2H), 3.92 (s, 2H), 4.67 (dt, 2H, J = 5.6Hz, 1.6Hz), 5.24 (dq, 1H, J = 1.2Hz, 10.4Hz), 5.38 (dq, 1H, J = 1.2Hz, 17.2Hz), 5.94-6.04 (ddt, 1H, J = 5.6, 17.2Hz, 10.4Hz), 7.03 (dt, 1H, J = 7.6Hz, 1.2Hz), 7.18 (dd, 1H, J = 13.6Hz, 1.2Hz), 7.31 (dt, 1H, J = 15.6Hz, 1.6Hz), 7.83 (bd, 1H), 7.90 (bs, 1H), 9.31 (bs, 1H). ^{13}C NMR (67.5MHz, $CDCl_3$) δ: 28.2, 50.1, 52.2, 65.5, 81.5, 117.6, 120.2, 122.6, 126.3, 128.6, 129.7, 133.0, 138.9, 153.8, 171.1. FAB+ HRMS (glycerol/KCl) m/z calc for $C_{17}H_{24}N_2O_4 + H^+$: 321.1814, found: 321.1813.

Synthesis of Alkyl Ammonium Chloride Salt 80:

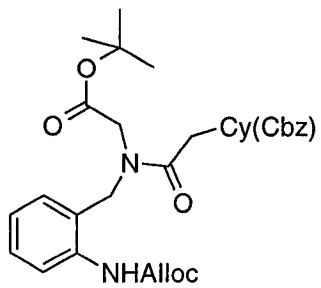


80

2-Aminobenzyl amine (20.6g, 171mmol) was dissolved in dry THF (425mL). To the solution was added Boc_2O (40.9g, 188mmol) as a solid in four equal portions at 10 min intervals. After 35 min the reaction was complete by TLC. The reaction was then concentrated to an orange oil and redissolved in dry pyridine (200mL) and dry CH_2Cl_2 (200mL). The solution was then cooled to 0°C under N_2 and AllocCl (36mL, 340mmol) was added dropwise. After 3 h, the reaction was shown to be complete by TLC. The reaction was diluted with CH_2Cl_2 (1L) and extracted with 1M HCl (3 x 500mL). The organic layer was then dried over anhydrous $MgSO_4$ and concentrated providing 49.4g (95%) of the dicarbamate precursor **79**. ^{1}H NMR (300MHz, $CDCl_3$) δ: 1.42 (s, 9H), 4.25 (d, 2H J = 6.8Hz), 4.65 (dt, 2H, J = 5.6, 1.6Hz), 4.96 (br, 1H), 5.22 (m, 1H, J = 10.4, 1.6 Hz), 5.37 (m, 1H, J = 17.2, 1.6Hz), 5.97 (m, 1H), 7.04 (dt, 1H, J = 7.6, 1.2Hz), 7.12 (dd, 1H, J = 7.2, 1.6Hz), 7.28 (dt, 1H, J = 7.6, 1.6Hz), 7.99 (br d, 1H), 8.59 (br, 1H). ^{13}C NMR (75MHz, $CDCl_3$) δ: 28.50, 41.92, 65.51, 80.69, 117.79, 121.57, 123.67, 129.08, 130.27, 130.40, 132.91, 137.05, 154.35, 157.01. Hydrogen chloride gas was bubbled through a solution of the dicarbamate (49.4g, 161mmol) dissolved in 1,4-dioxane (750mL) until CO_2 evolution began. After 20 min the reaction was shown to be complete by TLC. The reaction was then concentrated to dryness and the resulting solid triturated with EtOAc to give 36.0g (3 crops, 93%) of pure salt **80** as a pale orange solid. ^{1}H NMR (300MHz, CD_3OD) δ: 4.09 (s, 2H), 4.67 (dt, 2H, *J* = 6.0, 1.4Hz), 5.25 (m, 1H, J = 10.5,

1.2Hz), 5.39 (m, 1H, J = 16.8, 1.2Hz), 5.95-6.08 (m, 1H), 7.30-7.51 (m, 4H). ^{13}C NMR (75MHz, CD_3OD) δ: 40.50, 67.26, 118.23, 127.24, 127.84, 129.33, 131.14, 131.82, 133.99, 137.75, 157.79.

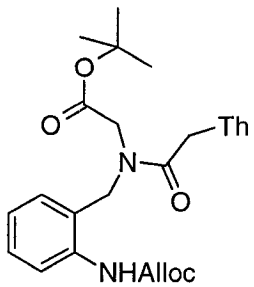
Synthesis of N-(N'-(allyloxycarbonyl)-2-aminobenzyl)-N-(N^4-benzyloxycarbonylcytisin-1-ylacetyl) glycine *tert*-Butyl Ester (81b):



81b

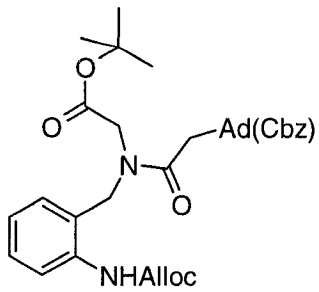
Compounds **81a-d** were synthesized using the following procedure outlined for carbamate **81b**: Amine **76** (170mg, 0.531mmol), $Cy^{Cbz}CH_2CO_2H$ **53** (241mg, 0.797mmol) and HATU (333mg, 0.877mmol) were suspended in DMF (5mL) under N_2 at 0°C. To the stirred suspension was added DIPEA (0.3mL, 1.6mmol). After 10 min, the reaction was allowed to come to RT and was then stirred for 8 h. The product mixture was diluted with EtOAc (100mL) and washed with H_2O (1 x 75mL), 5% $NaHCO_3$ (1 x 75mL) and H_2O (1 x 75mL). The organic layer was then dried over anhydrous $MgSO_4$ and concentrated to dryness. Pure monomer **81b** was obtained after flash column chromatography as a white foam. Yield: 97%. Rotational isomerism was present in the NMR spectra of this compound. Rotational isomers were present in approximately a 1:1 ratio. ^{1}H NMR (400MHz, $CDCl_3$) δ: 1.25 (s, 9H), 3.74 and 3.76 (s, 2H), 4.05 (bs, 2H), 4.50, 4.59, 4.61, 4.65 (s, 4H), 5.15 (overlapping m, 3H), 5.25 and

5.33 (bs, 1H), 5.85-5.96 (m, 1H), 6.90-6.98 (overlapping m, 1H), 7.05-7.10 (m, 1H), 7.13-7.38 (m, 6H), 7.71-7.77 (br, 1H), 7.95-8.05 (br, 1H). ^{13}C NMR (67.5MHz, DMSO-d_6) δ: 28.0, 28.2, 47.3, 47.7, 49.0, 49.8, 50.2, 50.4, 94.4, 94.5, 117.7, 128.5, 128.5, 128.7, 129.0, 133.8, 133.8, 136.1, 136.5, 137.0, 154.4, 154.8, 163.7, 168.2, 168.2, 168.3, 168.3, 169.1. FAB+ HRMS (glycerol/KCl) *m/z* calculated for $C_{31}H_{35}N_5O_8$ + K^+: 644.2123, found: 644.2125.



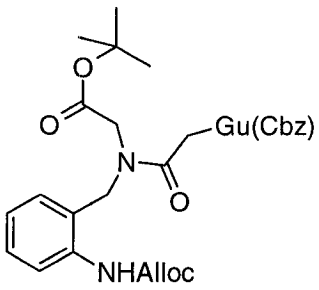
81a

Analytical Data for Carbamate **81a**: Yield: 82%. ^{1}H NMR (300MHz, DMSO-d_6) δ: 1.32 (*ma.*) and 1.37 (*mi.*) (s, 9H), 1.74 (*mi.*) and 1.75 (*ma.*) (s, 3H), 3.83 (*mi.*) and 4.17 (*ma.*) (s, 2H), 4.50-4.62 (m, 6H), 5.15-5.37 (m, 2H), 5.86-6.01 (m, 1H), 7.05-7.41 (m, 5H), 7.58 (br, 1H), 9.17 (br, 1H), 11.31 (br, 1H). ^{13}C NMR (67.5MHz, $CDCl_3$) δ: 12.46, 14.30, 21.15, 27.91, 47.83, 48.39, 48.95, 60.50, 65.58, 77.43, 83.63, 110.92, 117.32, 121.15, 123.15, 123.72, 129.85, 131.28, 132.80, 137.78, 141.05, 151.34, 154.01, 164.53, 167.49, 168.75.



81c

Analytical Data for Carbamate **81c**: Yield: 92%. ^{1}H NMR (300MHz, DMSO-d_6) δ: 1.36 (*ma.*) and 1.37 (*mi.*) (s, 9H), 3.71 (*mi.*) and 3.87 (*ma.*) (s, 2H), 4.31 (*mi.*) and 4.52 (*ma.*) (s, 2H), 4.50 (*ma.*) and 4.60 (*mi.*) (br d, 2H). ^{13}C NMR (75MHz, $CDCl_3$) δ: 27.82, 43.74, 48.91, 49.50, 65.48, 67.66, 83.81, 117.48, 121.08, 121.20, 123.12, 123.44, 128.45, 128.61, 128.65, 129.86, 131.20, 132.45, 135.55, 137.64, 144.24, 149.60, 151.39, 151.59, 153.84, 167.01, 167.68. FAB+ HRMS (glycerol) m/z calc. for $C_{32}H_{35}N_7O_7$ + H^+: 630.2676, found: 630.2676.

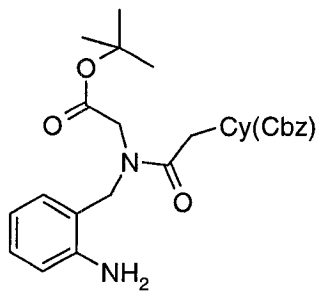


81d

Analytical Data for Carbamate **81d**: Yield: 76%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.35 (s, 9H), 3.85 (*mi.*) and 4.28 (*ma.*) (s, 2H), 4.49 (*ma.*) and 4.70 (*ma.*) (s, 2H), 4.52 (*ma.*) and 4.57 (*mi.*) (m, 2H), 5.00 (*ma.*) and 5.04 (*mi.*) (s, 2H), 5.12-5.35 (m, 4H), 5.84-6.00 (m, 1H), 7.06-7.56 (m, 9H), 7.86 (mi) and 7.87 (*ma.*) (s, 1H), 9.07 (*ma.*) and 9.19 (*mi.*) (br, 1H). ^{13}C NMR (75MHz, $CDCl_3$) δ: 27.95, 28.06, 44.20, 44.35, 48.85, 49.34, 49.57,

65.70, 66.33, 68.42, 68.59, 82.36, 83.44, 117.75, 118.43, 119.35, 119.67, 121.32, 123.31, 124.17, 126.49, 128.56, 128.77, 128.86, 129.00, 129.76, 129.98, 131.44, 132.48, 134.74, 134.87, 137.64, 140.14, 140.31, 147.19, 149.21, 149.47, 154.11, 155.24, 155.80, 155.88, 167.41, 167.61, 168.12.

Synthesis of N-(2-aminobenzyl)-N-(N^4-benzyloxycarbonylcytisin-1-ylacetyl) glycine *tert*-Butyl Ester 82b:



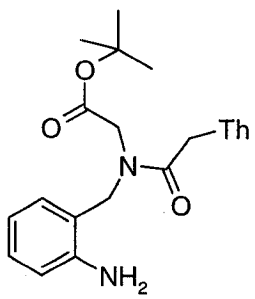
82b

Anilines **82a-d** were prepared using one of the two deprotection procedures described for aniline **82b**.

Procedure A: To a solution of allyl carbamate **81b** (1.65g, 2.73mmol) dissolved in DCM (20mL) was added Bu_3SnH (793mg, 2.73mmol), followed by $Pd(PPh_3)_4$ (157mg, 0.14mmol) and then H_2O (150uL). After 30 min the reaction was shown to be complete by TLC. The reaction mixture was filtered through silica gel and the silica gel was washed with EtOAc (100mL). The combined filtrates were concentrated to dryness and the resulting solid triturated with 1:1 EtOAc/hexanes. Pure aniline **82b** was isolated as a white solid (1.19g, 84%).

Procedure B: To a solution of allyl carbamate **81b** (1.65g, 2.73mmol) dissolved in DCM (20mL) and MeOH (10mL) was added benzenesulfinic acid (793mg, 2.73mmol)

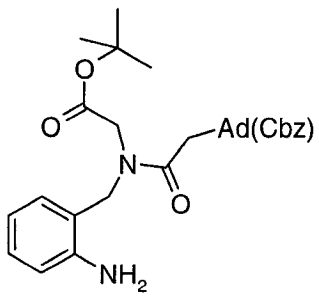
followed by $Pd(PPh_3)_4$ (157mg, 0.14mmol). After 20-40 min, the reaction was shown to be complete by TLC. The reaction mixture was then concentrated to dryness and the residue purified by flash column chromatography (0-8% MeOH/DCM) giving aniline **81b** as a glassy solid (86%). ^{1}H NMR (DMSO-d_6, 300MHz) δ: 1.30 (*ma.*) and 1.35 (*mi.*) (s, 9H), 1.75 (bs, 3H), 3.81 (*mi.*) and 4.10 (*ma.*) (s, 2H), 4.35 (*ma.*) and 4.44 (*mi.*) (s, 2H), 4.54 (*ma.*) and 4.61 (*mi.*) (s, 2H), 4.99 (*mi.*) and 5.06 (*ma.*) (s, 2H), 6.46 (t, 0.8H), 6.58-6.71 + 6.90-7.3 (m, 3.2H), 7.33 (*ma.*) and 7.39 (*mi.*) (d, 1H ^4J = 1.3Hz), 11.29 (br, 1H). ^{13}C NMR (75MHz, DMSO-d_3) δ: 27.95, 28.16, 47.66, 48.08, 48.15, 49.14, 50.33, 50.41, 67.01, 81.41, 82.20, 94.36, 94.48, 115.17, 115.84, 116.13, 117.05, 118.90, 119.06, 128.34, 128.42, 128.46, 128.66, 128.78, 128.99, 129.20, 130.96, 136.48, 146.94, 147.51, 151.41, 151.51, 153.72, 155.56, 163.70, 164.06, 168.12, 168.23, 168.26, 168.50, 168.64. FAB+ HRMS (glycerol/KCl) *m/z* calculated for $C_{20}H_{26}N_4O_5$ + H^+: 403.1981, found: 403.1981.



82a

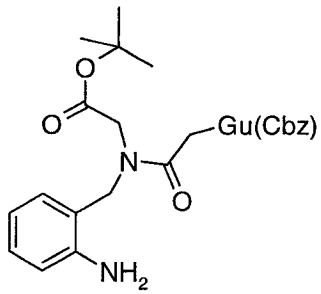
Analytical Data for Aniline **82a**: Aniline **82a** precipitated from the reaction mixture (procedure A) and was collected by filtration and washed with hexanes. Yield: 85%.^{1}H NMR (300MHz, DMSO-d_3) δ: 1.30 (*ma.*) and 1.35 (*mi.*) (s, 9H), 1.75 (br, 3H), 3.81 (*mi.*) and 4.11 (*ma.*) (s, 2H), 4.36 (*ma.*) and 4.45 (*mi.*) (s, 2H), 4.54 (*ma.*) and 4.61 (*mi.*) (s, 2H), 5.00 (*mi.*) and 5.07 (*ma.*) (s, 2H), 6.47 (*ma.*) and 6.57 (*mi.*) (t, 1H, J = 7.2Hz), 6.58

(*ma.*) and 6.68 (*mi.*) (d, 1H, J = 8.4Hz), 6.95-7.05 (m, 2H), 7.34 (*ma.*) and 7.39 (*mi.*) (s, 1H), 11.29-11.30 (br, 1H). ^{13}C NMR (75MHz, DMSO-d_3) δ: 12.44, 12.48, 27.90, 28.16, 47.66, 47.96, 48.12, 48.54, 48.99, 81.41, 82.22, 108.51, 108.73, 115.13, 115.83, 116.11, 117.06, 118.86, 119.05, 128.18, 128.80. 129.24, 131.06, 142.58, 143.81, 146.89, 147.53, 151.61, 165.09, 168.14, 168.27, 168.35, 168.73.



82c

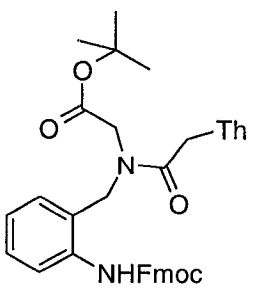
Analytical Data for Aniline **82c**: Aniline **82c** was prepared using procedure B and purified by flash column chromatography (using EtOAc as eluant). Yield: 85%. ^{1}H NMR (300MHz, $CDCl_3$) δ: 1.36, (s, 9H), 3.96 (*mi.*) and 4.12 (*ma.*) (s, 2H), 4.57-4.77 (m, 4H), 4.99-5.31 (m, 6H), 5.83-5.94 (m, 1H), 6.99-7.46 (m, 7H), 8.08-8.78 (m, 3H). ^{13}C NMR (75MHz, DMSO-d_6) δ: 27.94, 28.10, 44.66, 47.76, 48.16, 48.29, 49.09, 55.37, 66.72, 81.40, 83.32, 115.15, 116.12, 118.69, 119.00, 123.42, 128.30, 128.42, 128.85, 129.11, 129.24, 129.28, 131.07, 131.87, 132.02, 132.45, 132.49, 133.96, 136.84, 145.73, 145.77, 146.94, 147.47, 149.82, 149.87, 151.96, 152.64, 152.89, 167.67, 168.04, 168.12, 168.15.



82d

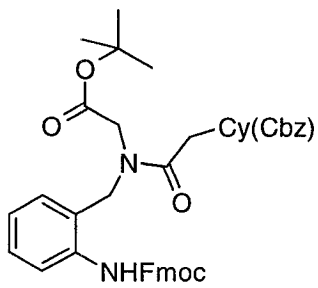
Analytical Data for Aniline **82d**: Aniline **82d** was prepared using procedure B and was purified by triturated with 1:5 EtOAc/hexanes and isolated as a pale yellow solid in 86% yield. ^{1}H NMR (300MHz, DMSO-d_6) δ: 1.33 (s, 9H), 3.84 (*mi.*) and 4.23 (*ma.*) (s, 2H), 4.36 (*ma.*) and 4.54 (*mi.*) (s, 2H), 4.99 (*ma.*) and 5.10 (*mi.*) (s, 2H), 5.04 (br, 2H), 5.24 (s, 2H), 6.47 –6.72 (m, 2H), 6.95-7.06 (m, 2H), 7.33-7.45 (m, 5H), 7.84 (s, 1H), 11.40-11.48 (br, 1H). ^{13}C NMR (75MHz, DMSO-d_3) δ: 28.04, 28.15, 44.57, 44.71, 47.74, 48.06, 49.12, 67.65, 81.47, 82.33, 115.17, 115.97, 116.16, 117.29, 118.67, 119.00, 119.71, 119.78, 128.59, 128.81, 129.30, 131.13, 136.04, 140.94, 141.09, 146.98, 147.47, 147.75, 147.82, 150.02, 150.13, 155.21, 155.72, 167.66, 168.19, 168.27, 168.37.

Synthesis of N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(thymin-1-ylacetyl) glycine tert-Butyl Ester 83a:



83a

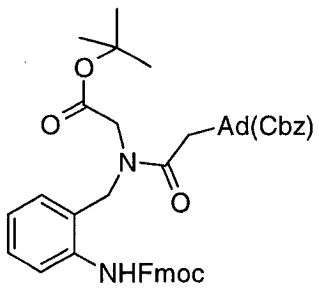
Compounds **83a-d** were synthesized using the following procedure outlined for carbamate **83a**: To a vigorously stirred suspension of aniline **82a** (2.30g, 5.74mmol) suspended in 1,4-dioxane (5mL) was added an aqueous solution of Na_2CO_3 (620mg in 5mL H_2O) followed by solid Fmoc-Cl (1.48g, 5.74mmol). After 45 min, the reaction was diluted with H_2O (30mL) and extracted once with EtOAc (75mL). The organic layer was dried over anhydrous $MgSO_4$ and concentrated to dryness. The white solid was purified by dissolution in a minimal amount of MeOH and precipitation with Et_2O giving 3.25g (89%) of pure product. ^{1}H NMR (300MHz, $CDCl_3$) δ: 1.37 (s, 9H), 1.71 (s, 3H), 4.07 (s, 2H), 4.25 (t, 1H), 4.39 (d, 2H), 4.49 (s, 2H), 4.64 (s, 2H), 7.02-7.42 (m, 10H), 7.76 (m, 4H), 8.11-8.69 (m, 2H). ^{13}C NMR (67.5MHz, $CDCl_3$) δ: 12.35, 27.97, 47.19, 48.00, 48.93, 49.27, 60.56, 77.43, 83.67, 111.06, 119.96, 120.84, 123.10, 123.51, 125.63, 127.17, 127.81, 129.97, 131.44, 137.86, 140.73, 141.31, 144.05, 151.31, 154.10, 164.42, 167.59, 168.88.



83b

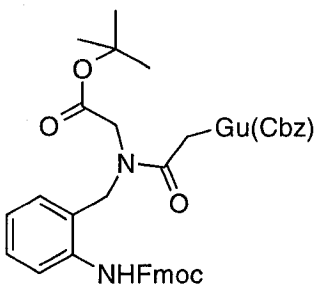
Analytical Data for Fmoc Derivative **83b**: Product was purified by flash column chromatography (EtOAc as eluant). Yield: 83%. ^{1}H NMR (400MHz, $CDCl_3$) δ: 1.35 (s, 9H), 4.16-4.25 (overlapping s and m, 3.7H), 4.35-4.37 (bs, 2.3H), 4.62-4.65 (m, 5.6H), 5.08 (br, 2.4H), 5.85-5.96 (m, 1H), 6.90-6.98 (overlapping m, 1H), 7.05-7.10 (m, 1H), 7.24-7.35 (m, 21.8H) 7.67-7.79 (m, 7.2H), 8.10-8.20 (br, 1H), 8.71-8.73 (br, 1H). ^{13}C

NMR (75MHz, $CDCl_3$) δ: 27.77, 27.89, 47.08, 48.91, 49.51, 49.76, 66.95, 67.62, 77.43, 82.02, 83.37, 95.22, 119.70, 119.94, 120.89, 123.00, 123.72, 125.12, 125.60, 127.13, 127.55, 128.25, 128.53, 128.58, 129.71, 131.24, 131.99, 132.12, 135.08, 137.72, 141.12, 143.70, 143.99, 149.63, 152.04, 154.05, 155.73, 162.86, 167.65, 168.02, 168.92.



83c

Analytical Data for Fmoc Derivative **83c**: Product was purified by flash column chromatography (EtOAc as eluant). ^1H NMR (400MHz, $CDCl_3$) δ: 1.35 (s, 9H), 4.16-4.25 (overlapping s and m, 3.7H), 4.35-4.37 (bs, 2.3H), 4.62-4.65 (m, 5.6H), 5.08 (br, 2.4H), 5.85-5.96 (m, 1H), 6.90-6.98 (overlapping m, 1H), 7.05-7.10 (m, 1H), 7.24-7.35 (m, 21.8H) 7.67-7.79 (m, 7.2H), 8.10-8.20 (br, 1H), 8.71-8.73 (br, 1H).

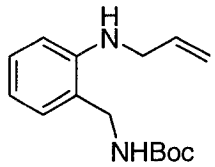


83d

Analytical Data for Fmoc Derivative **83d**: Product was purified by flash column chromatography (10%MeOH/EtOAc as eluant). Yield: 81%. ^1H NMR (400MHz, $CDCl_3$) δ: 1.33 (s, 9H), 3.84 (*mi.*) and 4.28 (*ma.*) (s, 2H), 4.19-4.68 (m, 5H), 5.00 (*mi.*) and 5.03

(*ma.*) (s, 2H), 5.18 (*ma.*) and 5.23 (*mi.*) (s, 2H), 7.07-7.90 (m, 18H), 9.20 (br, 1H), 11.28 (*ma.*) and 11.34 (*mi.*) (br, 1H), 11.38 (*mi.*) and 11.52 (*ma.*) (br, 1H). ^{13}C NMR (75MHz, DMSO-d_6) δ: 20.04, 28.15, 44.42, 47.02, 47.17, 47.60, 47.73, 48.80, 49.83, 66.42, 66.59, 67.66, 67.71, 81.59, 82.51, 119.70, 119.87, 120.60, 120.67, 123.62, 124.81, 125.60, 126.77, 127.52, 128.13, 128.19, 128.45, 128.61, 128.67, 128.72, 128.83, 129.00, 130.01, 131.15, 135.95, 136.05, 136.91, 140.89, 141.08, 141.20, 141.29, 144.12, 144.78, 147.60, 147.68, 149.91, 150.05, 154.41, 154.97, 155.02, 155.61, 155.67, 167.55, 168.17, 168.54, 168.82.

Synthesis of N-(2-N'-tert-Butoxycarbonylaminomethylphenyl) allylamine (86):



86

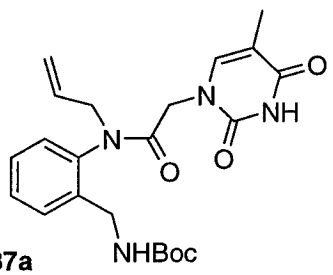
To a solution of aniline **85** (9.6g, 43mmol) in anhydrous DMF (420mL) was added DIPEA (3.7mL, 21mmol) and allyl bromide (25.8g, 21mmol). The reaction was allowed to stir overnight under N_2 and then concentrated to an oil. The residue was then taken up in EtOAc (400mL) and washed with H_2O (3 x 300mL). The organic layer was dried over Na_2SO_4 (anhydrous) and the solvent removed under reduced pressure. The product was obtained as an off-white solid after silica gel chromatography (10% EtOAc/hexanes) in 59% yield.

^{1}H NMR (300MHz, $CDCl_3$). δ: 1.44 (s, 9H), 3.83 (m, 2H), 4.27 (d, 2H, J = 6.6Hz), 4.82 (br, 1H), 5.18 (m, 1H), 5.29 (m, 1H), 5.90-6.03 (m, 1H), 6.67-6.72 (m, 2H), 7.06 (d, 1H, J

= 6.3), 7.21 (t, 1H, J = 7.7Hz). ^{13}C NMR (67.7MHz, DMSO-d_6) δ: 14.39, 22.52, 28.65, 31.42, 45.77, 78.48, 110.41, 115.69, 115.90, 123.85, 128.37, 136.41, 146.16, 156.67. ES$^+$ MS *m/z*: 263.1 $(M+H)^+$.

N-(2-N'-tert-Butoxycarbonylaminomethylphenyl)-N-(thymin-1-ylacetyl) allylamine (87a) and N-(2-N'-tert-Butoxycarbonylaminomethylphenyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl) allylamine (87b):

Compounds **87a,b** were prepared using the following general procedure: Secondary aniline **87a** (1 equivalent) was dissolved in anhydrous DMF (1M). To the solution was added carboxylic acid **51** or **53** (2 equivalents) followed by EDC (2 equivalents). The reaction was allowed to stir for 5-7 h under N_2 and then concentrated to a yellow oil. The residue was partitioned between EtOAc and saturated $NaHCO_3$. The organic layer was further extracted with saturated $NaHCO_3$ (1x) and H_2O (1x). The organic layer was concentrated to give a yellowish solid and the products purified as indicated.

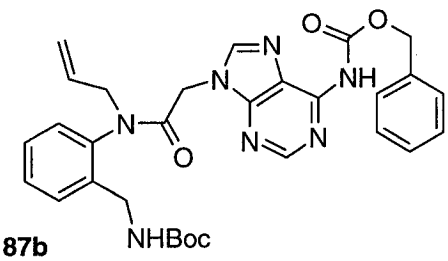


87a

Analytical Data for Thymine Derivative **87a**: Product was purified by dissolution in a minimal amount of EtOAc, followed by precipitation by slow addition of hexanes. Yield: 89% (white solid). ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.38 (s, 9H), 1.72 (s, 3H), 3.72 (dd, 1H, J = 14.4, 7.2), 3.91 (d, 1H, J = 16.8Hz), 4.07 (dd, 1H, J = 16.0, 6.0Hz), 4.14-4.22 (m, 2H), 4.60 (dd, 1H, J = 14.4, 5.6), 5.06-5.11 (m, 4H), 5.74-5.85 (m, 1H), 7.20-7.59 (m,

5H), 11.29 (br, 1H).[13]C NMR (67.7MHz, DMSO-d_6) δ: 12.34, 28.67, 49.20, 51.67, 78.69, 108.39, 118.95, 128.55, 129.25, 129.66, 130.01, 133.15, 137.93, 138.65, 142.50, 151.42, 156.50, 164.84, 166.63.

ES^+ MS *m/z*: 429.3 $(M+H)^+$; ES^- MS *m/z*: 427.2 $(M-H)^-$.



87b

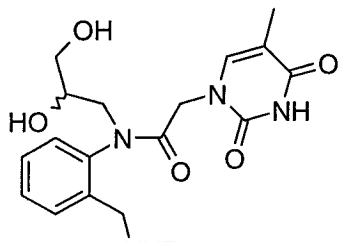
Analytical Data for Adenine Derivative **87b**: Product was purified by silica gel chromatography using EtOAc as eluant. Yield: 85% (white foam). ^{1}H NMR (400MHz, $CDCl_3$) δ: 1.44 (s, 9H), 3.92 (dd, 1H, J = 14.4, 7.2Hz), 4.25-4.62 (m, 4H), 4.91 (d, 1H, 16.8Hz), 5.12 (m, 1H), 5.18 (m, 1H), 5.30 (s, 2H), 5.77-5.93 (m, 1H), 6.64 (br, 1H), 7.22-7.53 (m, 4H), 7.71 (d, 1H, J = 7.2Hz), 8.29 (br, 1H), 8.79 (s, 1H). ^{13}C NMR (75MHz, DMSO-d_6) δ: 12.45, 14.59, 21.26, 28.59, 49.13, 50.65, 60.28, 80.10, 108.41, 118.65, 125.78, 126.21, 129.64, 130.81, 133.15, 133.24, 135.87, 142.49, 151.59, 153.96, 164.87, 167.21, 170.83.

ES^+ MS *m/z*: 572.3 $(M+H)^+$; ES^- MS *m/z*: 570.3 $(M-H)^-$.

N-(2-N'-tert-Butoxycarbonylaminomethylphenyl)-N-(thymin-1-ylacetyl)-3-amino-1,2-propanediol (88a) and N-(2-N'-tert-Butoxycarbonylaminomethylphenyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl)-3-amino-1,2-propanediol (88b):

Compounds **88a,b** were prepared using the following general procedure: Compound **87a** or **87b** was dissolved in 1:1:1 H_2O/THF/tBuOH (0.34M in **87a,b**). To the solution was

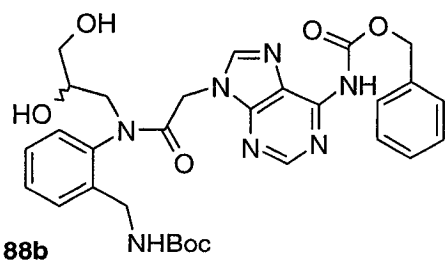
added OsO_4 (0.05 eq., 0.052M in benzene) followed by NMO (1.1 eq.) and the reaction was allowed to stir for 5-10 h. The reaction was then quenched by addition of a 1:1 mixture of saturated Na_2SO_3/saturated $NaHCO_3$ (2mL for every 1mL of reaction mixture). This mixture was stirred for 30-60 min and then extracted with EtOAc (1x). The products were then purified by silica gel chromatography using a gradient from 0-10% MeOH/DCM.



88a

Analytical Data for Diol **88a:** Yield: 86%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.39 (s, 9H), 1.72 (s, 3H), 3.27-4.76 (m, 6H), 7.16-7.58 (m, 5H), 11.26 (s, 1H). ^{13}C NMR (67.6MHz, $CDCl_3$) δ: 11.93, 28.45, (40.17, 40.48), 50.41, (52.69, 53.57), (64.92, 65.06), (70.43, 70.45), 80.00, (110.34, 110.38), 129.17, (129.68, 129.81), (130.21, 130.32), 130.40, (138.65, 138.74), 139.68, (143.29, 143.34), 152.61, 157.97, 166.41, (169.04, 169.39).

ES$^+$ MS *m/z*: 463.3 $(M+H)^+$; ES$^-$ MS *m/z*: 461.2 $(M-H)^-$.

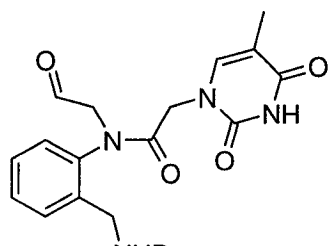


88b

Analytical Data for Diol **88b**:Yield: 89%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.39 (s,9H), 3.24-4.84 (m, 6H), 5.21 (s, 2H), 7.22-7.71 (m, 9H), 8.31 and 8.32 (s, 1H), 8.58 and 8.59 (s, 1H), 10.69 (br, 1H).

N-(2-N'-tert-Butoxycarbonylaminomethylphenyl)-N-(thymin-1-ylacetyl)glycinal (89a) and N-(2-N'-tert-Butoxycarbonylaminomethylphenyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl)glycinal (89b):

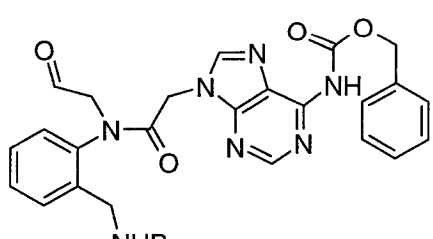
Compounds **89a,b** were prepared using the following general procedure: An aqueous solution of $NaIO_4$ (1.05 eq., 0.25M) was added to a solution of diol **88a** or **88b** dissolved in THF (0.5M) and the reaction was carefully monitored by TLC. Upon complete consumption of the starting material, the reaction was filtered and the product was extracted from the reaction mixture with EtOAc (3x). The combined organic layers were dried over anhydrous $MgSO_4$, concentrated to dryness and the resulting white foam was used immediately in the next step.



89a

Analytical Data for Aldehyde **89a**: Yield (crude): >98%. TLC (EtOAc): ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.38 (s, 9H), 1.72 (s, 3H), 4.04-4.34 (m, 4H), 4.68 (d, 1H, J = 18.0Hz), 5.16 (d, 1H, J = 6.8Hz), 7.34-7.56 (m, 5H), 9.51 (s, 1H).

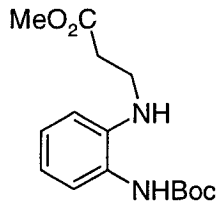
ES^+ MS *m/z*: 431.2 $(M+H)^+$; ES^- MS *m/z*: 429.2 $(M-H)^-$.



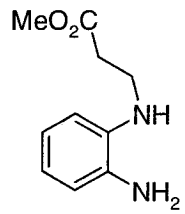
89b

Analytical Data for Aldehyde **89b**: Yield (crude): >98%. TLC (EtOAc): ^{1}H NMR (300MHz, DMSO-d_6) δ: 1.39 (s, 9H), 4.04 (d, 1H, J = 18.0Hz), 4.40 (d, 2H, J = 6.0Hz), 4.98 (d, 1H, J = 18.0Hz), 5.31 (s, 2H), 7.29-7.67 (m, 5H), 8.48 (br, 1H), 8.79 (s, 1H), 9.64 (s, 1H). ^{13}C NMR (75MHz, $CDCl_3$) δ: 14.24, 21.11, 28.48, 39.66, 44.70, 59.47, 60.46, 67.64, 79.81, 121.25, 128.45, 128.59, 128.66, 129.20, 129.32, 130.49, 131.16, 135.47, 137.69, 138.12, 144.53, 149.67, 151.19, 151.28, 152.68, 156.55, 166.30, 171.27, 195.04.

Synthesis of N-(2-N'-*tert*-Butoxycarbonylaminophenyl)–β-alanine methyl ester (92):



92

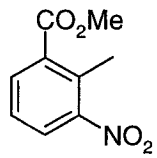


188

A flask containing mono-Boc protected diamine **91** (711mg, 3.4mmol) was charged with anhydrous DMF (750μL) and methyl 3-bromoproprionate (230mg, 1.4mmol). The mixture was rapidly (over ~5 min) heated to ~110-130°C and then allowed to cool to RT with stirring over 1 h. The reaction was then evaporated to dryness and the mixture purified by silica gel chromotography (gradient of 25-50% EtOAc/hexanes as eluant) to give 36% of secondary aniline **92**, 34% of deprotected backbone **188** and 23% of starting aniline **91**. ^{1}H NMR (300MHz, $CDCl_3$) δ: 1.52 (s, 9H), 2.66 (t, 2H, J = 6.6Hz), 3.44 (t, 2H, J = 6.3Hz), 3.72 (s, 3H), 6.76-6.83 (m, 2H), 7.08 (dt, 1H, J = 7.5, 1.8Hz), 7.37 (d, 1H, J = 7.8Hz). ^{13}C NMR (75MHz, DMSO-d_6) δ: 28.34, 33.83, 39.90, 51.76, 80.36, 113.10, 118.65, 124.89, 125.19, 126.23, 141.22, 154.13, 172.76.

ES^+ MS *m/z*: 295.2 $(M+H)^+$.

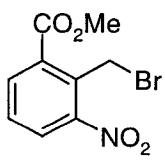
Synthesis of 2-Methyl-3-nitrobenzoic acid methyl ester (95).



95

2-Methyl-3-nitrobenzoic acid (52.0g, 287mmol) was dissolved in 500mL reagent grade MeOH and concentrated H_2SO_4 (3mL) was added to the solution. This mixture was heated to reflux for 24 h, cooled to RT and concentrated to 1/10 of the original volume on a rotary evaporator. The mixture was diluted with EtOAc (500mL) and washed with saturated $NaHCO_3$ (2 x 500mL) and H_2O (1 x 500mL). The organic layer was dried over Na_2SO_4 and concentrated to dryness to give the title compound in 89% yield as a pale orange solid. ^1H NMR (270MHz, $(CDCl_3)$ δ: 2.61 (s, 3H), 3.92 (s, 3H), 7.37 (t, 1H, J = 8.4Hz), 7.83 (d, 1H, J = 9.2Hz), 7.98 (d, 1H, J = 6.7Hz). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 15.99, 52.88, 127.35, 127.83, 132.74, 134.19, 134.46, 153.06, 167.35.

Synthesis of 2-Bromomethyl-3-nitro-benzoic acid methyl ester (96).

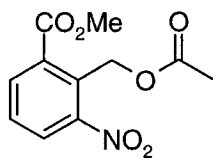


96

2-Methyl-3-nitrobenzoic acid methyl ester (5.78g, 29.6mmol) and N-bromosuccinimde (5.8g, 33mmol) were suspended in anhydrous CCl_4 (50mL) and the mixture heated to reflux. The reflux condenser was then removed, benzoyl peroxide (1.8g, 7.4mmol) quickly added and the condenser reattached. The reaction was maintained at reflux until decomposition of the product became competitive with product formation, at which time the reaction was cooled to RT. The solution was concentrated to near dryness and the resulting residue taken up in EtOAc (300mL) and washed with a 1:1 mixture of saturated $NaHCO_3$ and saturated $Na_2S_2O_3$ (300mL) followed by H_2O (300mL). The organic layer was dried over Na_2SO_4 and concentrated to an orange oil. The residue was purified by silica gel chromatography (elution with 10% EtOAc/hexanes until the starting material had been collected, and then elution with 20% EtOAc/hexanes) to give 6.0g of the title compound (74%) and 1.4g of starting material (25%).

Compound **96**: ^{1}H NMR (400MHz, $CDCl_3$) δ: 3.99 (s, 3H), 5.14 (s, 2H), 7.52 (t, 1H, J = 8.0Hz); 7.94 (d, 1H, J = 8.4Hz), 8.09 (d, 1H, J = 8.0Hz). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 23.99, 53.37, 128.71, 130.72, 132.64, 133.44, 135.44, 151.62, 166.65.

Synthesis of 2-Acetoxymethyl-3-nitro-benzoic acid methyl ester (97).

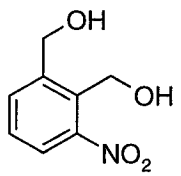


97

2-Bromomethyl-3-nitro-benzoic acid methyl ester (5.36g, 19.6mmol) and NaOAc (16.1g, 196mmol) were dissolved in anhydrous DMF (125mL). After stirring for 15 h, the

reaction mixture was diluted with EtOAc (400mL) and H_2O (400mL). The organic layer was washed with H_2O (2 x 400mL), dried over Na_2SO_4 and concentrated onto 5g of silica gel. The silica gel mixture was added to a one inch plug of silica gel and the product was eluted from the column with 20% EtOAc/hexanes. The fractions containing the title compound were concentrated to dryness giving 2-acetoxymethyl-3-nitro-benzoic acid methyl ester as a colorless oil (99%). ^{1}H NMR (400MHz, $(CD_3)_2SO$) δ: 2.03 (s, 3H), 3.83 (s, 3H), 5.12 (s, 2H), 7.75 (t, 1H), 7.88 (d, 1H), 7.95 d, 1H). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 20.18, 53.26, 59.89, 127.83, 130.54, 130.93, 134.43, 134.65, 152.13, 167.13, 170.36.

Synthesis of 2-Hydroxymethyl-6-nitro-phenyl)-methanol (98).

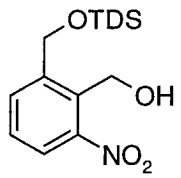


98

2-Acetoxymethyl-3-nitro-benzoic acid methyl ester (4.4g, 19mmol) was dissolved in dry THF (165mL). To the solution was added $LiBH_4$ (2.5g, 110mmol) in six equal portions at 5 minute intervals. The reaction was allowed to stir overnight and then quenched by the addition of H_2O. After stirring for 30 min, the solution was acidified to pH = 1 and then extracted with EtOAc (2 x 300mL). The combined organic layers were then washed with saturated $NaHCO_3$ (1 x 300mL) and H_2O (1 x 300mL), dried over Na_2SO_4 and concentrated to give the title compound in 89% yield as an orange oil which solidified on standing. ^{1}H NMR (400MHz, $(CD_3)_2CO$) δ: 4.34 (t, 1H, J = 5.6Hz), 4.56 (t, 1H, J = 5.6),

4.83 (d, 2H, J = 6.0Hz), 4.88 (d, 2H, J = 5.6Hz), 7.51 (t, 1H, J = 8.0Hz), 7.69 (d, 1H, J = 8Hz), 7.79 (d, 1H, J = 7.6Hz). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 57.13, 123.35, 129.22, 132.53, 133.32, 144.31, 151.84.

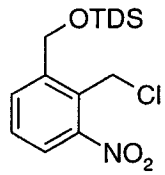
Synthesis of [2-(tert-Butyl-diphenyl-silanyloxymethyl)-6-nitro-phenyl]-methanol (99).



99

To a solution of diol **98** (3.0g, 16.5mmol) in anhydrous DMF (50mL) and distilled DIPEA (30mL), was added dropwise TDS-Cl (9.1g, 33mmol) under an N_2 atmosphere. After stirring overnight, the reaction was quenched with H_2O (300mL) and extracted with EtOAc (300mL). The organic layer was then washed with dilute HCl (2 x 300mL), saturated $NaHCO_3$ (2 x 300mL) and H_2O (300mL). The organic layer was then dried over Na_2SO_4 and concentrated to dryness. The residue was then purified by silica gel chromatography (gradient from 5% to 20% EtOAc/hexanes) to give the title compound as the only silylated product in 50% yield. ^{1}H NMR (400MHz, $(CD_3)_2SO$) δ: 1.04 (s, 9H), 4.52 (d, 2H, J = 5.2Hz), 4.95 (s, 2H), 5.19 (t, 1H, J = 5.4Hz), 7.35-7.48 (m, 6H), 7.53 (t, 1H, J = 7.6Hz), 7.61 (d, 4H, J = 6.8Hz), 7.69 (d, 1H, J = 8Hz), 7.78 (d, 1H, J = 7.2Hz). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ:19.84, 27.23, 57.07, 63.95, 123.36, 128.78, 129.25, 130.86, 131.61, 132.92, 133.86, 136.27, 143.02, 151.80.

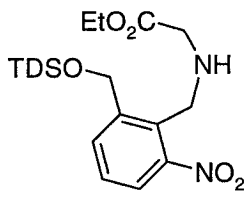
Synthesis of tert-Butyl-(2-chloromethyl-3-nitro-benzyloxy)-diphenyl-silane (100).



100

To a solution of alcohol **99** (3.2g, 7.6mmol) in dry DCM (20mL) was added distilled DIPEA (1.6mL, 9.2mmol) followed by dropwise addition of MsCl (660μL, 8.4mmol). After ca. 5 h of stirring, the reaction mixture was evaporated to dryness and used directly in the next step. A small portion of the reaction mixture was further purified by silica gel chromatography (20%EtOAc/hexanes) in order to confirm the identity of the product. ^{1}H NMR (270MHz, $(CD_3)_2CO$) δ: 1.12 (s, 9H), 4.93 (s, 2H), 5.13 (s, 2H), 7.42-7.53 (m, 6H), 7.66-7.78 (m, 5H), 7.90-7.96 (m, 2H). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 19.80, 27.20, 37.71, 63.58, 124.48, 128.79, 129.34, 130.70, 130.92, 132.78, 133.58, 136.27, 143.02, 151.00.

Synthesis of [2-(tert-Butyl-diphenyl-silanyloxymethyl)-6-nitro-benzylamino]-acetic acid ethyl ester (101).



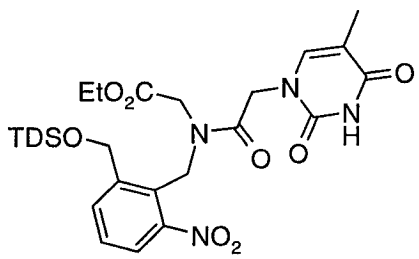
101

The crude mixture from the previous reaction was dissolved in anhydrous DMF (15mL) and added *via* syringe to a solution of glycine ethyl ester hydrochloride (10.6g, 76mmol) and DIPEA (13.2mL, 76mmol) in DMF (200mL). After stirring for 15 h, the reaction

was diluted with H_2O (600mL) and extracted with EtOAc (600mL). The organic layer was further extracted with saturated $NaHCO_3$ (2 x 600mL) and dried over Na_2SO_4. The product was isolated after silica gel chromatography (10%EtOAc/hexanes) as a pale yellow gum in 69% yield over the two steps. ^1H NMR (400MHz, $(CD_3)_2CO$) δ: 1.12 (s, 9H), 1.20 (t, 3H, J = 7.2Hz), 2.84 (br, 1H), 3.24 (s, 2H), 3.85 (s, 2H), 4.08 (q, 2H, 7.2Hz), 5.11 (s, 2H), 7.41-7.50 (m, 6H), 7.55 (t, 1H, J = 7.6Hz), 7.69 (d, 1H, J = 8Hz), 7.73 (d, 4H, 7.6Hz), 7.90 (d, 1H, J = 8.0Hz). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 14.54, 19.85, 27.26, 45.28, 51.15, 60.85, 63.84, 123.27, 128.74, 129.00, 130.82, 131.48, 131.55, 133.95, 136.31, 143.49, 152.44, 172.38.

ES^+ MS *m/z*: 507.1 $(M+H)^+$.

Synthesis of {[2-(tert-Butyl-diphenyl-silanyloxymethyl)-6-nitro-benzyl]-[2-(5-methyl-2,4-dioxo-3,4-dihydro-2H-pyrimidin-1-yl)-acetyl]-amino}-acetic acid ethyl ester (102).



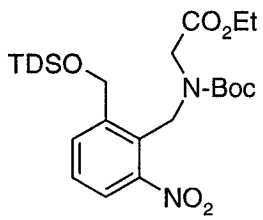
102

Secondary amine **101** (201mg, 0.397mmol) and thymine derivative **51** (146mg, 0.794mmol) were dissolved in anhydrous DMF (1200μL). To the solution was added EDC (150mg, 0.785mmol) and the reaction was allowed to stir overnight under N_2. The reaction mixture was diluted with EtOAc (5mL) and extracted with saturated $NaHCO_3$ (2

x 5mL) and H_2O (1 x 5mL). The organic layer was dried over $MgSO_4$ and then concentrated to dryness. The product was isolated as a white powder after purification by silica gel chromatography (1% MeOH/DCM) in 96% yield. ^1H NMR (400MHz, $(CD_3)_2SO$) δ: 1.03 (s, 9H), 1.13-1.17 (m, 3H), 1.70 (*ma.*) and 1.73 (*mi.*) (br, 3H), 3.52 (*mi.*) and 4.00 (*ma.*) (s, 2H), 3.88 (*mi.*) and 4.06 (*ma.*) (q, 2H, J = 6.8), 4.39 (*ma.*) and 4.63 (*mi.*) (s, 2H), 4.55 (*ma.*) and 4.81 (*mi.*) (s, 2H), 4.85 (*ma.*) and 4.95 (*mi.*) (s, 2H), 6.98 (*ma.*) and 7.05 (*mi.*) (s, 1H), 7.36-7.44 (m, 6H), 7.55-7.61 (m, 5H), 7.75 (d, 1H, J = 8.4Hz), 7.84 (d, 1H, J = 7.2Hz), 11.29 (br, 1H). ^{13}C NMR (67.7MHz, $(CD_3)_2SO$) δ: 12.55, 14.40, 14.48, 27.24, 42.45, 48.19, 48.27, 61.05, 61.68, 62.98, 108.62, 108.70, 123.12, 126.02, 128.33, 128.39, 129.58, 130.37, 130.53, 132.66, 132.96, 135.32, 135.42, 141.96, 142.59, 143.09, 150.95, 151.11, 151.60, 164.64, 168.41, 168.52, 168.95.

ES^+ MS *m/z*: 695.3 $(M+Na)^+$; ES^- MS *m/z*: 671.2 $(M-H)^-$.

Synthesis of {tert-Butoxycarbonyl-[2-(tert-butyl-diphenyl-silanyloxymethyl)-6-nitro-benzyl]-amino}-acetic acid ethyl ester (103).



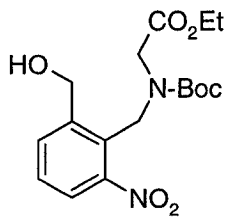
103

Di-*tert*-butyl dicarbonate (496mg, 2.28mmol) was added in one portion to a solution of secondary amine **101** dissolved in dry THF (50mL). The reaction was heated to 50°C and stirred until all of the starting material had been consumed (TLC monitoring). The reaction was then concentrated and purified by flash column chromatography (15%

EtOAc/hexanes) to give carbamate **103** in 94% yield. ^{1}H NMR (400MHz, $CDCl_3$) δ: 1.08 (*mi.*) and 1.11 (*ma.*) (s, 9H), 1.13 (*mi.*) and 1.27 (*ma.*) (s, 9H), 1.19-1.27 (m, 3H), 3.58 (*ma.*) and 3.68 (*mi.*) (s, 2H), 4.09-4.15 (m, 2H), 4.50 (*mi.*) and 4.55 (*ma.*) (s, 2H), 4.83 (*mi.*) and 4.92 (*ma.*) (s, 2H), 7.33-7.63 (m, 12H), 7.84 (mi) and 7.90 (*ma.*) (d, 1H, J = 8.0Hz). ^{13}C NMR (67.7MHz, $CDCl_3$) δ: 14.42, 19.55, 27.02, 27.15, 28.05, 28.19, 43.01, 44.23, 47.89, 48.28, 61.18, 62.65, 62.84, 80.89, 81.28, 122.48, 122.69, 126.78, 127.99, 128.09, 128.52, 128.81, 130.05, 130.19, 130.53, 132.87, 133.22, 135.66, 143.96, 143.98, 151.24, 151.76, 155.17, 155.46, 169.88.

ES$^+$ MS *m/z*: 629.1 $(M+Na)^+$.

Synthesis of [tert-Butoxycarbonyl-(2-hydroxymethyl-6-nitro-benzyl)-amino]-acetic acid ethyl ester (104).



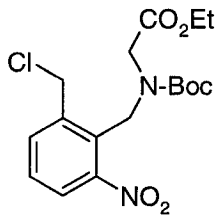
104

To a solution of carbamate **103** (1.64g, 2.72mmol) dissolved in dry THF (10mL) was added TBAF (3.1mL, 4M in THF). After 1h of stirring at RT, the reaction was concentrated onto 4g of silica gel and then applied to a column of silica gel. The product was eluted the column with 1:1 EtOAc/hexanes to give pure alcohol **104** in 93% yield. ^{1}H NMR (300MHz, $(CD_3)_2SO$) δ: 1.15-1.21 (m, 3H), 1.22 and 1.33 (s, 9H), 3.65 and 3.72 (s, 2H), 4.02-4.12 (m, 2H), 4.57-4.61 (m, 4H), 5.40-5.50 (m, 1H), 7.45-7.55 (m, 1H), 7.65-7.78 (m, 2H). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 14.34, 14.27, 28.16, 44.03, 44.82,

48.98, 60.43, 61.30, 61.60, 62.04, 80.75, 80.89, 122.91, 123.05, 128.11, 128.99, 129.26, 129.49, 129.85, 135.35, 144.31, 145.14, 151.87, 152.35, 170.17.

ES^+ MS *m/z*: 390.9 $(M+Na)^+$.

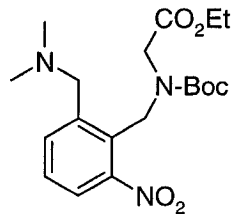
Synthesis of [tert-Butoxycarbonyl-(2-chloromethyl-6-nitro-benzyl)-amino]-acetic acid ethyl ester (105).



105

To a solution of alcohol **104** (1.38g, 3.80mmol) and DIPEA (1.4mL, 8.0mmol) in dry DCM (15mL) was added dropwise MsCl (330μL, 4.2mmol). The reaction was allowed to stir for 20 h at RT and then concentrated and applied directly to a column of silica gel. Elution with 20% EtOAc/hexanes afforded the chloride **105** in 88% yield as a pale yellow oil. ^{1}H NMR (400MHz, ($CDCl_3$) δ:1.21-1.29 (m, 3H), 1.43 (s, 9H), 3.78 (*ma.*) and 3.84 (*mi.*) (s, 2H), 4.13-4.22 (m, 2H), 4.74-4.78 (m, 4H), 7.42-7.46 (m, 1H), 7.60-7.64 (m, 2H). ^{13}C NMR (67.7MHz, ($CDCl_3$) δ: 14.48, 28.42, 42.96, 43.55, 44.12, 61.36, 81.49, 81.9, 124.23, 124.37, 129.38, 129.52, 129.81, 134.44, 134.74, 140.00, 140.68, 152.14, 155.46, 155.65, 170.03.

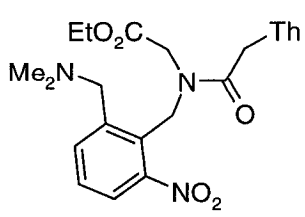
Synthesis of [tert-Butoxycarbonyl-(2-dimethylaminomethyl-6-nitro-benzyl)-amino]-acetic acid ethyl ester (106).



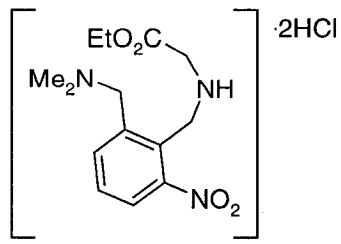
106

To a solution of chloride **105** (1.20g, 3.16mmol) in dry DMF (75mL) was added K_2CO_3 (4.7g, 34mmol) followed by $(CH_3)_2NH_2Cl$ (2.6g, 31mmol) and the reaction was stirred overnight at RT. The reaction was then diluted with H_2O (250mL) and EtOAc (250mL). The organic layer was washed with saturated $NaHCO_3$ (3 x 250mL), dried over Na_2SO_4 and concentrated to give the desired tertiary amine in >98% yield in analytically pure form as determined by ^1H NMR. ^1H NMR (400MHz, ($CDCl_3$) δ: 1.23-1.27 (m, 3H), 1.40 (*mi.*) and 1.43 (*ma.*) (s, 9H), 2.19 (s, 6H), 3.46 (*mi.*) and 3.49 (*ma.*) (s, 2H), 3.69 (*ma.*) and 3.75 (*mi.*) (s, 2H), 4.13-4.18 (m, 2H), 4.77 (*mi.*) and 4.80 (*ma.*) (s, 2H), 7.33-3.37 (m, 1H), 7.49-7.56 (m, 2H).

Synthesis of {(2-Dimethylaminomethyl-6-nitro-benzyl)-[2-(thymin-1-yl)-acetyl]-amino}-acetic acid ethyl ester (107).



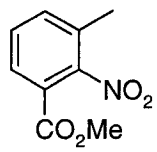
107



189

Carbamate **106** (605mg, 1.56mmol) was treated with HCl (7.8mL, 4M/dioxane) for 2 h at RT. At this time, the reaction was shown to be complete by HPLC. The solution was then concentrated to dryness to give the dihydrochloride salt **189·2HCl** as a white powder, which was used directly in the next step without further characterization. The free amine was then isolated by addition of EtOAc (25mL) and 5% K_2CO_3 (25mL), drying the organic layer over Na_2SO_4 and concentrating to dryness. To a solution of compound **189**, carboxylic acid **51** (574mg, 3.12mmol) and HATU (596mg, 3.12mmol) in dry DMF (9.3mL) was added collidine dropwise *via* syringe. After stirring overnight, the reaction was judged to be complete by TLC analysis. The reaction was diluted with H_2O (50mL) and EtOAc (50mL) and the organic layer washed with saturated $NaHCO_3$ (3 x 50mL). The organic layer was then dried over Na_2SO_4 and concentrated to dryness. The residue was then triturated with EtOAc/hexanes (3:1) and the product was isolated by filtration. Analytically pure product was obtained after silica gel chromatography (0-5% MeOH/DCM) in 71% yield over two steps. ^{1}H NMR (300MHz, $(CD_3)_2SO$) δ: 1.10 (*mi.*) and 1.18 (*ma.*) (t, 3H, J = 6.9Hz), 1.75 (br, 3H), 2.09 (*ma.*) and 2.13 (*mi.*) (s, 6H), 3.38 (*ma.*) and 3.52 (*mi.*) (s, 2H), 3.62 (*mi.*) and 4.00 (*ma.*) (s, 2H), 3.98 (*mi.*) and 4.10 (*ma.*) (q, 2H, J = 7.1Hz), 4.46 (*ma.*) and 4.76 (*mi.*) (s, 2H), 4.79 (*ma.*) and 4.60 (*mi.*) (s, 2H), 7.21 (*ma.*) and 7.26 (*mi.*) (br, 1H), 7.48-7.85 (m, 3H), 11.29 (br, 1H). ^{13}C NMR (67.7MHz, DMSO-d_6) δ: 12.38, 14.42, 31.15, 42.12, 43.61, 45.37, 46.43, 47.48, 48.33, 60.79, 60.99, 61.28, 61.53, 61.62, 108.77, 123.37, 127.65, 128.15, 129.38, 129.95, 134.70, 141.40, 141.82, 142.28, 151.37, 151.40, 152.12, 164.84, 168.81, 168.93, 169.23. ES^+ MS *m/z*: 483.2 $(M+Na)^+$; ES^- MS *m/z*: 459.0 $(M-H)^-$.

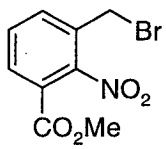
Synthesis of 3-Methyl-2-nitro-benzoic acid methyl ester (110).



110

3-Methyl-2-nitrobenzoic acid (52.0g, 287mmol) was dissolved in 500mL reagent grade methanol and concentrated H_2SO_4 (3mL) was added to the solution. This mixture was heated to reflux for 24h, cooled to RT and concentrated to 1/10 the original volume on a rotary evaporator. The mixture was diluted with EtOAc (500mL) and washed with saturated $NaHCO_3$ (2 x 500mL) and H_2O (1 x 500mL). The organic layer was dried over Na_2SO_4 and concentrated to dryness to give the title compound in 89% yield as a pale orange solid. ^1H NMR (270MHz, ($CDCl_3$) δ: 2.34 (s, 3H), 3.89 (s, 3H), 7.42-7.49 (m, 2H), 7.84 (d, 1H, J = 6.0Hz), 7.98. ^{13}C NMR (67.7MHz, $CDCl_3$) δ: 17.25, 53.14, 123.26, 128.90, 130.13, 130.66, 135.83, 150.84, 164.14.

Synthesis of 3-Bromomethyl-2-nitro-benzoic acid methyl ester (96).

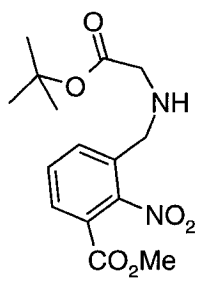


96

3-Methyl-2-nitrobenzoic acid methyl ester (1.00g, 5.13mmol) and N-bromosuccinimde (1.00g, 5.64mmol) were suspended in anhydrous CCl_4 (50mL) and the mixture heated to reflux. The reflux condenser was then removed, benzoyl peroxide (319mg, 1.28mmol) quickly added and the condenser reattached. The reaction was maintained at reflux until

decomposition of the product became competitive with product formation, at which time the reaction was cooled to RT. The solution was concentrated to near dryness and the resulting residue taken up in EtOAc (50mL) and washed with a 1:1 mixture of saturated $NaHCO_3$ and saturated $Na_2S_2O_3$ (50mL) followed by H_2O (50mL). The organic layer was dried over Na_2SO_4 and concentrated to an orange oil. The residue was purified by silica gel chromatography (elution with 5% EtOAc/hexanes until the starting material had been collected, and then elution with 20% EtOAc/hexanes) to give 689mg of the title compound (49%). ^1H NMR (400MHz, $(CD_3)_2CO$) δ: 3.89 (s, 3H), 4.43 (s, 2H), 7.56 (t, 1H, J = 6.7Hz), 7.72 (d, 1H, J = 7.1Hz), 7.94 (d, 1H, J = 7.9Hz). ^{13}C NMR (67.7MHz, $CDCl_3$) δ: 25.82, 53.52, 124.48, 130.54, 130.95, 131.38, 135.49, 149.89, 163.62.

Synthesis of 3-[(tert-Butoxycarbonylmethyl-amino)-methyl]-2-nitro-benzoic acid methyl ester (113):



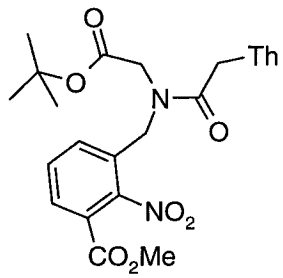
113

To a solution of glycine, *tert*-butyl ester (1.02g, 7.54mmol) and DIPEA (0.44mL, 2.5mmol) in anhydrous DMF (20mL) was added bromide **111** (689mg, 2.51mmol). The reaction was allowed to stir for 3 h, after which the reaction had gone to completion. The solution was diluted with EtOAc (100mL) and extracted with H_2O (60ml) and saturated $NaHCO_3$ (aqueous, 3 x 60mL). The organic layer was dried over Na_2SO_4 and

concentrated to a pale yellow oil, which was pure by ^{1}H NMR. Yield >98%. ^{1}H NMR (300MHz, $(CD_3)_2CO$) δ: 1.45 (s, 9H), 3.25 (s, 2H), 3.86 (s, 2H), 3.87 (s, 3H), 7.71 (t, 1H, J = 7.6Hz), 7.93 (dd, 1H J = 8.0, 1.6Hz), 7.97 (dd, 1H, J = 7.6, 1.6Hz). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ: 28.00, 48.35, 51.36, 53.20, 81.11, 124.38, 130.30, 131.36, 134.23, 135.36, 150.64, 164.86, 171.88.

ES^+ MS *m/z*: 324.9 $(M+H)^+$; ES^- MS *m/z*: 305.0 $(M-CH_3)^-$.

Synthesis of 3-({tert-Butoxycarbonylmethyl-[2-(thymin-1-yl)-acetyl]-amino}-methyl)-2-nitro-benzoic acid methyl ester (114):



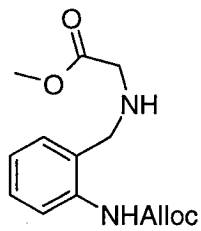
114

Amine **113** (501mg, 1.54mmol), thymine derivative **51** (311mg, 1.69mmol) and EDC (311mg, 1.63mmol) were dissolved in anhydrous DMF (3mL) and the reaction was stirred overnight. The mixture was then diluted with EtOAc (15mL) and extracted with H_2O (15mL) and saturated $NaHCO_3$ (3 x 15mL). The organic layer was then dried over Na_2SO_4 and concentrated to dryness to give 705mg (93%) of the desired amide in an analytically pure form. ^{1}H NMR (400MHz, $(CD_3)_2SO$) δ: 1.37 (*mi.*) and 1.40 (*ma.*) (s, 9H), 1.74 (s, 3H), 3.83 (*ma.*) and 3.84 (*mi.*) (s, 3H), 3.87 (*mi.*) and 4.23 (*ma.*) (s, 2H), 4.50 (*ma.*) and 4.61 (*mi.*) (s, 2H), 4.57 (*ma.*) and 4.71 (*mi.*) (s, 2H), 7.34 (*ma.*) and 7.40 (*mi.*) (s, 1H), 7.70-7.96 (m, 3H), 11.30 (br, 1H). ^{13}C NMR (67.7MHz, $(CD_3)_2CO$) δ:

12.25, 28.05, 28.13, 46.77, 47.38, 48.84, 48.92, 49.22, 50.40, 53.33, 53.43, 82.10, 83.10, 109.81, 109.89, 124.55, 125.13, 130.37, 130.83, 131.02, 131.19, 131.82, 132.41, 133.74, 134.61, 142.46, 142.64, 150.18, 151.95, 164.51, 164.59, 164.86, 168.39, 168.72, 168.81, 169.47.

ES^+ MS *m/z*: 513.0 $(M+Na)^+$; ES^- MS *m/z*: 489.1 $(M-H)^-$.

Synthesis of methyl ester backbone 185:



185

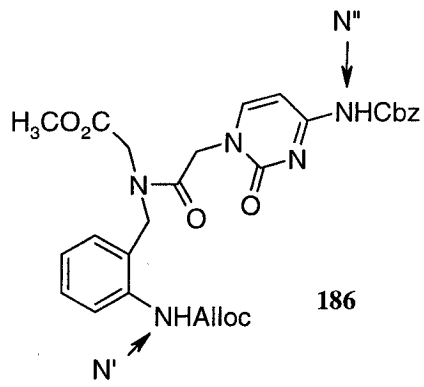
Using the conditions outlined for amine **76a**, compound **185** was also prepared.

^1H NMR (400MHz, $CDCl_3$) δ: 3.46 (s, 2H), 3.75 (s, 3H), 3.90 (s, 2H), 4.67 (ddd, 2H, *J* = 6.0 & 0.8 Hz), 5.24 (dm, 1H, *J* = 10.4 Hz), 5.38 (dm, 1H, *J* = 17.2 Hz), 5.94-6.04 (m, 1H), 7.03 (ddd, 1H, *J* = 7.2 & 0.8 Hz), 7.18 (d, 1H, *J* = 7.2 Hz), 7.31 (ddd, 1H, *J* = 7.8 & 1.6 Hz), 7.94 (bd, 1H, *J* = 8.0 Hz), 9.31 (bs, 1H).

^{13}C NMR (67.5MHz, $CDCl_3$) δ: 28.2, 50.1, 52.2, 65.5, 81.5, 117.6, 120.2, 122.6, 126.3, 128.6, 129.7, 133.0, 138.9, 153.8, 171.1.

Synthesis of protected N-[N'-(allyloxycarbonyl)-2-aminobenzyl]-N-[N"-benzyloxycarbonylcytisin-1-ylacetyl] glycine methyl ester (186):

Cytosine monomer **186** was also prepared using the procedure described for compounds **81a-d**:



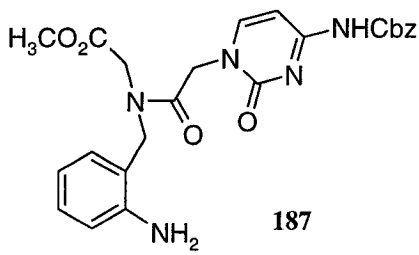
186

Amine **185** (328mg, 1.2mmol), $Cy^{Cbz}CH_2CO_2H$ **52** (466mg, 1.5mmol) and HATU (583mg, 1.5mmol) were suspended in anhydrous DMF (7mL) under N_2 at 0°C. To the stirred suspension, DIPEA (0.54mL, 3.1mmol) was added and, after stirring at 0°C for 10 min, the reaction mixture was allowed to warm-up to RT and stir for an additional 8 h. The crude product was diluted with EtOAc (100mL) and washed with H_2O (1 x 75mL), 5% $NaHCO_3$ (1 x 75mL) and H_2O (1 x 75mL). The organic layer was then dried over anhydrous $MgSO_4$ and concentrated to dryness. Pure monomer **186** was obtained after flash column chromatography, using pure EtOAc as the eluent, as a white foam (642mg, 72% yield).

^1H NMR (300MHz, $CDCl_3$, mixture of rotamers in 1:3.5 ratio) δ: 3.69 (bs, 3H), 4.00 (*mi*) & 4.26 (*ma*) (bs, 2H), 4.59-4.85 (m, 6H), 5.19-5.37 (m, 4H), 5.85-6.02 (m, 1H), 6.98-7.03 (m, 1H), 7.10-7.12 (m, 1H), 7.22-7.39 (m, 7H), 7.53 (bs, 1H), 7.70 (bd, 1H, J = 6.9 Hz), 8.04 (bd, 1H, J = 6.9 Hz), 8.34 (bs, 1H).

ES^+ MS: *m/z* 564 $(M+H)^+$; ES^- MS: *m/z* 562 $(M-H)^-$.

Synthesis of N-(2-aminobenzyl)-N-(N^4-benzyloxycarbonylcytisin-1-ylacetyl) glycine methyl ester 187:



187

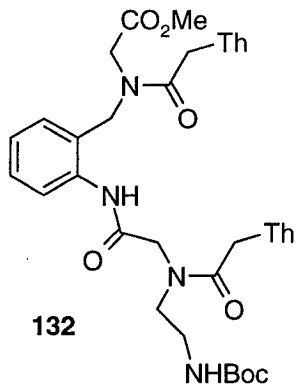
Aniline **187** was prepared using procedure A used for anilines **83a-d**: Allyl carbamate **186** (164mg, 0.3mmol) was dissolved in DCM (6mL) and H_2O (100μL), and then Bu_3SnH (81μL, 0.3mmol) followed by $Pd(PPh_3)_4$ (16mg, 5mol%) were added. After 15 min, the reaction mixture was concentrated to an oil, which was then purified by flash column chromatography, using a solvent gradient of 0-5% MeOH in $CHCl_3$ as the eluant. Pure aniline **187** was isolated as a glassy solid (140mg, 91%).

^{1}H NMR (300MHz, $CDCl_3$, mixture of rotamers in 1:1.3 ratio) δ: 3.58 (*mi*) & 3.62 (*ma*) (s, 3H), 3.96 (*mi*) & 4.23 (*ma*) (s, 2H), 4.37 (*ma*) & 4.51 (*mi*) (s, 2H), 4.70 (*ma*) & 4.77 (*mi*) (s, 2H), 5.01 (*mi*) & 5.02 (*ma*) (s, 2H), 5.17 (*mi*) & 5.18 (*ma*) (s, 2H), 6.47 (ddd, 0.6H, J = 7.2 & 1.2 Hz), 6.56-6.60 (m, 0.9H), 6.69 (d, 0.5H, J = 7.6 Hz), 6.95-7.03 (m, 3H), 7.33-7.40 (m, 5H), 7.94-7.99 (m, 1H), 10.74 (bs, 1H).

ES^+ MS: *m/z* 480 $(M+H)^+$; ES^- MS: *m/z* 478 $(M-H)^-$.

7.7-Synthesis of APNA-PNA Dimers.

Synthesis of {N-[2-(tert-butoxycarbonyl amino)ethyl]-N-[thymin-1-ylactetyl] glycyl}-{N-(2-aminobenzyl)-N-(thymin-1-ylacetyl)glycine} methyl ester (132):



132

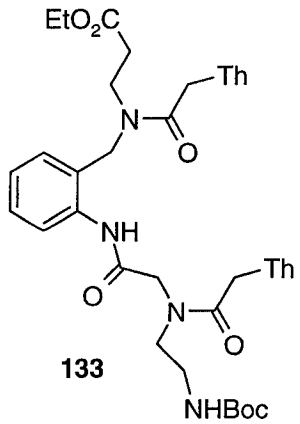
Free acid **124** (830mg, 2.2mmol), aniline **43a** (773mg, 2.2mmol) and HATU (828mg, 2.2mmol) were placed in a flask under N_2 and anhydrous DMF (5mL) was added. The mixture was cooled to 0°C and a solution of DIPEA (0.11mL, 10% solution in anhydrous DMF, 4.3mmol) was added drop-wise. The reaction mixture was allowed to warm-up to RT and stir for 6 h before quenching with H_2O (15mL). A mixture of 10% MeOH in $CHCl_3$ (50mL) was added and the organic layer was washed with 5% $NaHCO_3$ (2 x 15mL) and brine (1 x 15mL), dried over anhydrous $MgSO_4$ and concentrated to dryness. The crude product was purified by flash column chromatography, using a solvent gradient from 0-8% MeOH in $CHCl_3$ as the eluent to give 1.56g (79% yield) of pure dimer **132**.

TLC R_f(8% MeOH in $CHCl_3$) = 0.211.

^{1}H NMR (300MHz, DMSO-d_6) δ: 1.36-1.38 (bs, 9H), 1.74 (bs, 6H), 3.06-3.44 (m, 4H), 3.59-3.62 (bs, 3H), 3.97-4.68 (m, 8H), 6.74-7.68 (m, 6H), 9.55-9.86 (m, 1H), 11.29-11.33 (m, 2H).

FAB$^+$ HRMS (glycerol/KCl) *m/z* found 765.260910 $(M+K)^+$; calculated for $(C_{33}H_{42}N_8O_{11}K)^+$ = 765.261013.

Synthesis of Dimers 133 and 137:



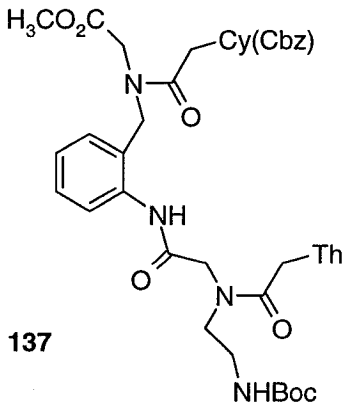
133

Dimers **133** and **137** were synthesized using the same procedure as for dimer **132**:

Dimer **133**: Yield: 71%. TLC R_f (12% MeOH in EtOAc) = 0.121.

^{1}H NMR (300MHz, DMSO-d_6) δ: 1.09-1.19 (m, 3H), 1.35-1.37 (m, 9H), 1.74 (bs, 6H), 2.41-2.71 (m, 2H), 3.00-3.52 (m, 6H), 3.93-4.73 (m, 10H), 6.74-7.66 (m, 7H), 9.57-9.90 (m, 1H), 11.27-11.30, (m, 2H).

ES$^+$ MS: *m/z* 755 $(M+H)^+$; ES$^-$ MS: *m/z* 753 $(M-H)^-$.



137

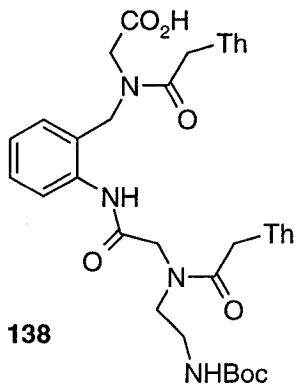
Dimer **137**:

Yield: 69%. TLC R_f (5%MeOH in $CHCl_3$) = 0.19.

^{1}H NMR (300MHz, $CDCl_3$) δ: 1.35-1.42 (m, 9H), 1.83-1.88 (m, 3H), 3.26-3.75 (m, 7H), 4.02-4.79 (m, 10H), 5.19-5.43 (m, 2H), 6.97-8.47 (m, 13H), 8.81-9.03 (m, 1H), 9.50-9.85 (m, 1H), 10.68 (bs, 1H).

ES^+ MS: *m/z* 832 $(M+H)^+$; ES^- MS: *m/z* 830 $(M-H)^-$.

Synthesis of {N-[2-(tert-butoxycarbonyl amino)ethyl]-N-[thymin-1-ylactetyl] glycyl}-{N-(2-aminobenzyl)-N-(thymin-1-ylacetyl)glycine} (138):



138

To a solution of ester **132** (1.97g, 2.7mmol) in THF (20mL), aqueous LiOH (10mL, 0.8M, 8.0mmol) was added and the reaction was stirred at RT for 45 min. The mixture was then diluted with H_2O (30mL), cooled to 0°C and acidified to pH 3 by the drop-wise addition of aqueous HCl (0.1M). The solution was then extracted with 10% MeOH in $CHCl_3$ (3 x 100mL). The combined organic layers were dried over anhydrous $MgSO_4$ and concentrated to give dimer **138** as a white foam in 88% yield (1.70g) in high purity.

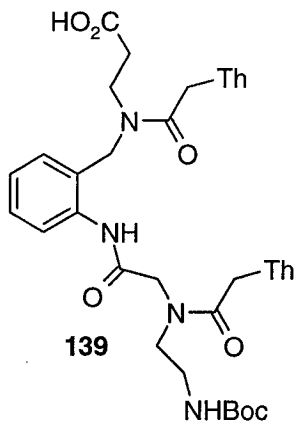
RP-TLC R_f (80% MeOH in H_2O) = 0.1.

^{1}H NMR (300MHz, DMSO-d_6, mixture of four rotamers) δ: 1.34-1.36 (m, 9H), 1.71-1.75 (m, 6H), 3.02-3.48 (m, 4H), 3.82-4.69 (m, 10H), 5.19 (s, 2H), 5.31 (*ma*) & 5.35 (*mi*) (s, 2H), 6.70-7.70 (m, 6H), 9.57-9.68 (m, 1H), 10.66 (bs, 1H), 11.26-11.31 (m, 2H).

FAB$^+$ HRMS (glycerol/NaCl) *m/z* found 735.2714 $(M+Na)^+$; calculated for $(C_{32}H_{40}N_8O_{11}+Na)^+ = 735.2712$.

Synthesis of dimers 139 and 143:

The free acids of dimers **139** and **143** were synthesized using the same procedure as for **138**:

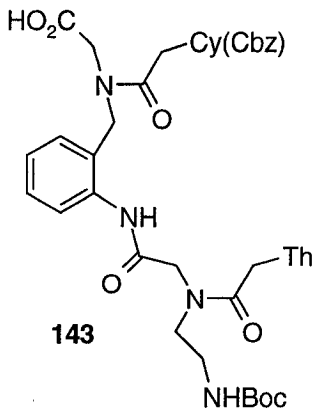


139

Free acid **139**: Yield: 90%.

^{1}H NMR (400MHz, DMSO-d_6) δ: 1.35 and 1.37 (m, 9H), 1.73 and 1.75 (m, 6H), 2.36-2.65 (m, 2H), 3.03-3.61 (m, 6H), 4.04-4.73 (m, 10H), 6.74-7.71 (m, 6H), 9.64-9.97 (m, 1H), 11.29-11.32, (m, 2H), 12.26 (br, 1H).

ES$^+$ MS: *m/z* 727 $(M+H)^+$; ES$^-$ MS: *m/z* 725 $(M-H)^-$.



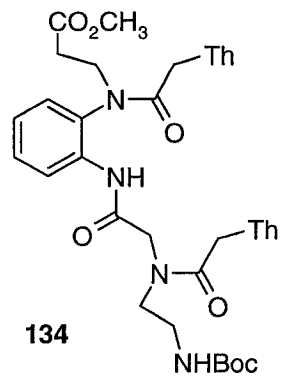
143

Free acid **143**: Yield: 88%.

^{1}H NMR (300 MHz, DMSO-d_6) δ: 1.34-1.36 (m, 9H), 1.71-1.75 (m, 3H), 3.02-3.48 (m, 4H), 3.86-4.83 (m, 10H), 5.18 (s, 2H), 6.70-7.92 (m, 13H), 9.54-9.85 (m, 1H), 10.79 (bs, 1H), 11.30 (m, 1H).

ES$^+$ MS: *m/z* 846 $(M+H)^+$; ES$^-$ MS: *m/z* 844 $(M-H)^-$.

Synthesis of {N-[2-(tert-butoxycarbonyl amino)ethyl]-N-[thymin-1-ylactetyl] glycyl}-{N-(2-N'-*tert*-Butoxycarbonylaminophenyl)-N-(thymin-1-ylacetyl)-β-alanine} methyl ester (134):

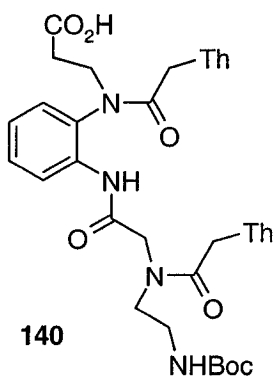


134

Thymine derivative **93** (302mg, 0.66mmol) was treated with HCl in dioxane (20 eq.) at RT. After complete consumption of the starting material, the solution was evaporated to dryness to give an off-white powder. To the flask was added PNA thymine monomer

124 (300mg, 0.79mmol), HATU (298mg, 0.79mmol) and HOAt (106mg, 0.79mmol). The flask was cooled to 0°C and charged with a solution of collidine (490μL, 3.7mmol) in anhydrous DMF (4mL). The reaction was allowed to warm to RT and stir overnight under N_2. The reaction was diluted with EtOAc (12mL) and extracted with saturated $NaHCO_3$ (2 x 12mL) and H_2O (1 x 12mL). The organic layer was dried over $MgSO_4$ and concentrated to an orange residue. The product was purified by silica gel chromatography giving dimer **134** in 47% yield. ^{1}H NMR (400MHz, $CDCl_3$) δ: 1.42-1.44 (m, 9H), 1.85-1.91 (m, 6H), 2.41-2.69 (m, 2H), 3.27-4.85 (m, 12H), 5.86 (br, 1H), 6.89-7.52 (m, 6H), 8.22-8.24 (*ma.*) and 8.37-8.39 (*mi.*) (m, 1H), 9.27 (*ma.*) and 9.57 (*mi.*) (s, 1H), 9.39 (*ma.*) and 9.86 (*mi.*) (br, 2H).

Synthesis of {N-[2-(tert-butoxycarbonyl amino)ethyl]-N-[thymin-1-ylactetyl] glycyl}-{N-(2-N'-*tert*-Butoxycarbonylaminophenyl)-N-(thymin-1-ylacetyl)-β-alanine} (140):



140

A solution of aqueous LiOH (4 eq., 0.75M) was added to a solution of dimer **134** (142mg, 0.19mmol), which was dissolved in THF (1mL). After 20 min of stirring, the reaction was diluted with H_2O (5mL) and acidified to pH = 3 by dropwise addition of HCl (3M). The product was then extracted into 5% MeOH/DCM (5 x 10mL). The organic layer was dried over $MgSO_4$ and concentrated to give dimer **140** as a glassy

white solid (97%). ^{1}H NMR (400Hz, DMSO-d_6) δ: 1.34 (*mi.*) and 1.37 (*ma.*) (s, 9H), 1.67-1.69 (m, 6H), 2.36-2.53 (m, 2H), 3.02-3.48 (m, 6H), 3.85-4.74 (m, 6H), 6.76 (*mi.*) and 6.98 (*ma.*) (m, 1H), 7.17-7.49 (m, 6H) 7.83-7.85 (m, 1H), 9.45 (*ma.*) and 9.88 (*mi.*) (s, 1H), 11.28-11.31 (m, 2H). ES^+ MS *m/z*: 713.3 $(M+H)^+$, 735.3 $(M+Na)^+$; ES^- MS *m/z*: 711.3 $(M-H)^-$.

7.8-Synthesis of APNA Dimers.

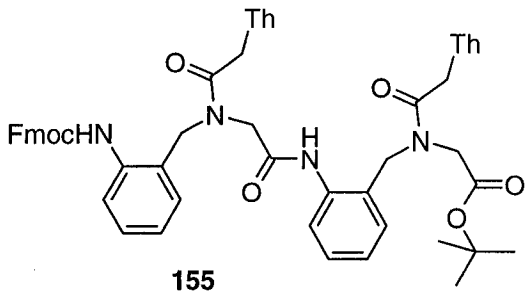
Dimers **155-159** were synthesized using one of the two following procedures:

Procedure A: Free acid **43a-d** (1.2 eq.), aniline **82a-d** (1.0 eq.) and HATU (1.2 eq.) were dissolved in anhydrous DMF (0.5M solution) and cooled to 0°C under N_2. The reaction was initiated by the dropwise addition of 2,4,6-collidine (2.2 eq.) and the reaction was allowed to come to RT and monitored by TLC. When TLC detected no further progress, the reaction was diluted with EtOAc (3 volumes) and washed with H_2O (1 x 3 volumes), saturated $NaHCO_3$ (2 x 3 volumes), dilute HCl (1 x 3 volumes) and H_2O (1 x 3 volumes). The organic layer was then dried over anhydrous Na_2SO_4 and concentrated to dryness. The products were then purified by flash column chromatography using the conditions indicated.

Procedure B: Free acid **43a-d** (1.2 eq.), aniline **82a-d** (1.0 eq.), HOAt (1.2eq.) and HATU (1.2 eq.) were dissolved in anhydrous DMF (0.5M solution) and cooled to 0°C under N_2. The reaction was initiated by the dropwise addition of 2,4,6-collidine (3.4 eq.). The reaction was allowed to come to RT and monitored by TLC. When TLC detected no

further progress, the reactions were worked up and the products purified as described in procedure A.

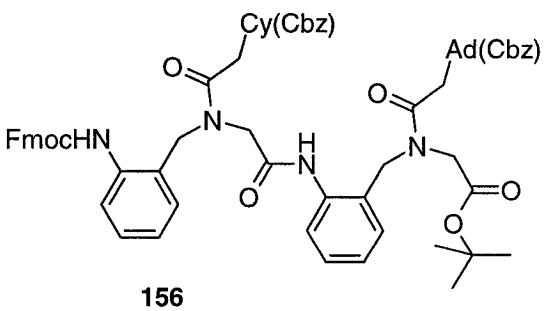
{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(thymin-1-ylacetyl)glycyl}-{N-(2-aminobenzyl)-N-(thymin-1-ylacetyl) glycine} ***tert*****-Butyl Ester** (Dimer **155**):



155

Synthesized using procedure A. Column conditions: 0-8% MeOH/CH_2Cl_2. Yield: 549mg pure (58%). ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.23-1.35 (9H), 1.66 (6H), 3.80-5.08 (15H), 6.47-7.89 (19H), 9.30-9.76 (1H), 11.30-11.35 (2H).

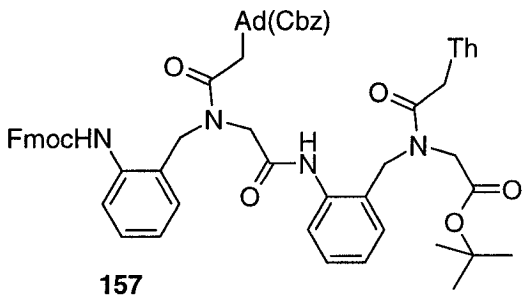
{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-[(N"-benzyloxycarbonyl)cytosin-1-ylacetyl]glycyl}-{N-(2-aminobenzyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl) glycine} ***tert*****-Butyl Ester** (Dimer **156**):



156

Synthesized using procedure A. Column conditions: 0-10% MeOH/EtOAc. Yield: 736mg (70%). ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.29-1.37 (9H), 3.85-5.28 (19H), 6.42-8.03 (29H), 8.31-8.38 (1H), 8.52-8.60 (1H), 9.16-9.85 (2H), 10.67-10.78 (1H).

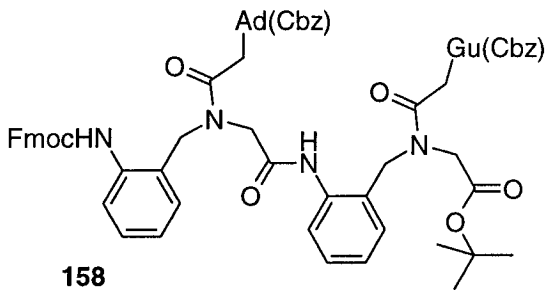
{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl)glycyl}-{N-(2-aminobenzyl)-N-(thymin-1-ylacetyl) glycine} *tert*-Butyl Ester (Dimer **157**):



157

Synthesized using procedure B. Column conditions: 3-5% MeOH/CH_2Cl_2. Yield: 321mg (74%). ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.18-1.35 (9H), 1.63-1.76 (3H), 3.79-5.37 (17H), 6.45-7.97 (22H), 8.28-8.62 (2H), 9.23 (1H), 9.50-9.82 (1H), 10.65-10.66 (1H), 11.25-11.50 (1H).

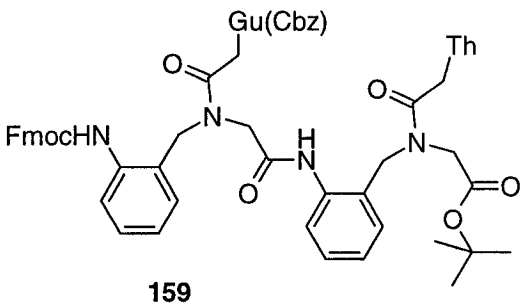
{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl)glycyl}-{N-(2-aminobenzyl)-N-(N"-benzyloxycarbonyl)guanin-1-ylacetyl) glycine} *tert*-Butyl Ester (Dimer **158**):



158

Synthesized using procedure B. Column conditions: 3-5% MeOH/CH_2Cl_2. Yield: 140mg (40%). ^{1}H NMR (400MHz, $CDCl_3/CD_3OD$, 20/1) δ: 1.21-1.36 (9H), 3.82-5.26 (19H), 6.96-7.88 (29H), 8.19 (m, 1H), 8.45 (m, 1H), 9.59 (m, 1H).

{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(N"-benzyloxycarbonyl)guanin-1-ylacetyl)glycyl}-{N-(2-aminobenzyl)-N-(thymin-1-ylacetyl) glycine} *tert*-Butyl Ester (Dimer **159**):



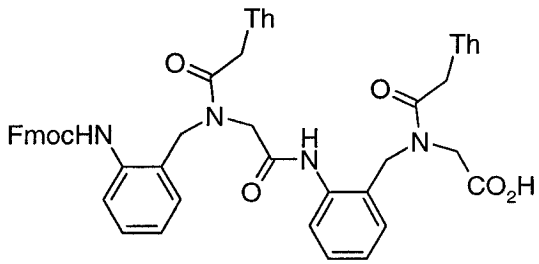
159

Synthesized using procedure B. Column conditions: 3-5% MeOH/CH_2Cl_2. Yield: 437mg (49%). ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.21-1.34 (9H), 1.65-1.72 (3H), 3.77-5.22 (17H), 7.08-7.98 (24H), 9.24 (1H), 9.47-9.85 (1H), 11.22-11.60 (2H).

Synthesis of Free Acid Fmoc Protected Dimers 160-164:

Dimers **160-164** were synthesized using the procedure described for free acids **43a-d**.

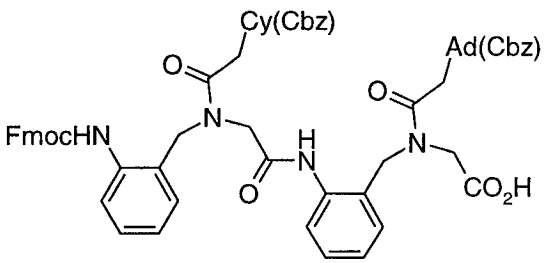
{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(thymin-1-ylacetyl)glycyl}-{N-(2-aminobenzyl)-N-(thymin-1-ylacetyl) glycine} (Dimer **160**):



160

Yield: 100%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.23-1.35 (9H), 1.66 (6H), 3.80-5.08 (15H), 6.47-7.89 (19H), 9.30-9.76 (1H), 11.30-11.35 (2H).

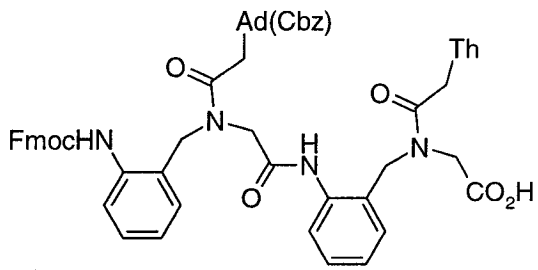
{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-[(N"-benzyloxycarbonyl)cytosin-1-ylacetyl]glycyl}-{N-(2-aminobenzyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl) glycine} (Dimer **161**):



161

Yield: 99%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 3.91-5.30 (m, 19H), 6.96-8.01 (m, 29H), 8.37-8.60 (m, 2H), 9.12-9.84 (m, 2H), 10.79 (br, 1H).

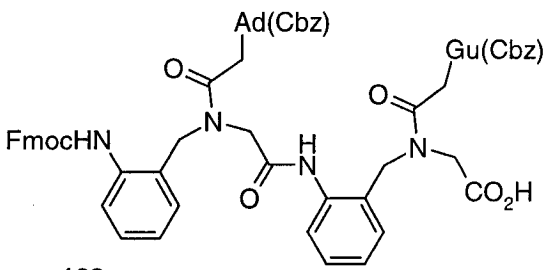
{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl)glycyl}-{N-(2-aminobenzyl)-N-(thymin-1-ylacetyl) glycine} (Dimer **162**):



162

Yield: 100%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 1.64-1.74 (m, 3H), 3.86-5.40 (m, 17H), 7.15-7.90 (24H), 8.40-8.63 (1H), 9.23-9.84 (m, 2H), 11.23-11.41 (m, 1H).

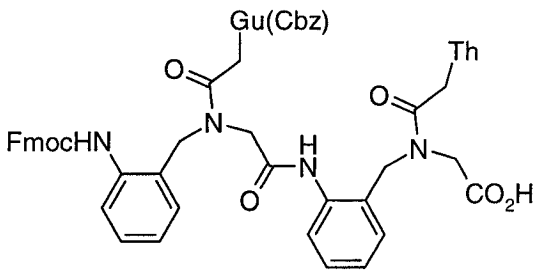
{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(N"-benzyloxycarbonyl)adenin-1-ylacetyl)glycyl}-{N-(2-aminobenzyl)-N-(N"-benzyloxycarbonyl)guanin-1-ylacetyl) glycine} (Dimer **163**):



163

Yield: 97%. ^{1}H NMR (400MHz, DMSO-d_6) δ: 3.86-5.37 (m, 19H), 7.10-7.86 (m, 30H), 8.30-8.57 (m, 1H), 9.20 (br, 1H), 9.52-9.87 (m, 1H), 10.61 (br m, 1H), 11.26-11.51 (m, 2H).

{N-(N'-(9-fluorenylmethoxycarbonyl)-2-aminobenzyl)-N-(N"-benzyloxycarbonyl)guanin-1-ylacetyl)glycyl}-{N-(2-aminobenzyl)-N-(thymin-1-ylacetyl) glycine} (Dimer **164**):



164

Yield: >98%. ^{1}H NMR (400MHz, DMSO-d_6, 40°C) δ: 1.66-1.71 (3H), 3.84-5.22 (17H), 7.10-7.93 (24H), 9.23 (1H), 9.49-9.82 (1H), 11.23-11.59 (2H).

7.9-Solid Phase Synthesis of APNA-PNA Chimeras and APNA Homopolymers.

Oligomers **126-128**, **144-150**, **165-173** (0.02- 0.03mmol) were synthesized on MBHA resin with a loading capacity of 0.19-0.40mmol/g, using the following protocol:

1) Derivatization of MBHA Resin (per 100mg of dry resin): The dry MBHA.HCl resin was suspended in DCM (5mL) for 24 h. The resin was washed with 5% DIPEA in DMF (2 x 3mL) for 10 min each time. Boc-Lys(2-Cl-Cbz)-OH (5 eq.) was cross-linked to the resin using HATU (4.8 eq.) as the coupling agent in the presence of DIPEA (10 eq.) until a negative Kaiser test was observed (~3 h). The resin was treated with a mixture of Ac_2O/pyridine/DCM (1:25:25 ratio, 3mL) for 10 min and then washed with DCM (2 x 30 sec, 3mL), DMF (2 x 30 sec, 3mL) and DCM (2 x 30 sec, 3mL).

2) Coupling of Monomer or Dimer Fragments to Fmoc- protected Amino Groups: Deprotection: The resin was treated with 5% piperidine/DMF (1 x 5 min) and then washed with DMF (2 x 30 sec) and the washings collected. This process was repeated once. The resin was washed with DCM (3 x 30 sec) and DMF (3 x 30 sec). Deprotection was monitored by the Kaiser test. Coupling: The resin was treated with a solution of free acid/HATU/2,4,6-collidine (3eq/2.85eq/6eq) in 1:1 DMF/pyridine (total free acid concentration 0.07M) which had been pre mixed for 1 min. The coupling was allowed to proceed until a negative Kaiser test was observed (~30-240 min). The resin was washed with DMF (3 x 30 sec), DCM (3 x 30 sec), DMF (3 x 30 sec) and DCM (3 x 30 sec). Capping: The resin was the treated with Ac_2O/pyridine/DCM (1:25:25) for 5 min and washed with DCM (4 x 30 sec), DMF (2 x 30 sec) and DCM (2 x 30 sec). A small portion of the resin was removed and the coupling sequence quantitated as described in reference 9.

3) Coupling of Monomer or Dimer Fragments to Fmoc- protected Anilines: Deprotection: The resin was treated with 5% piperidine/DMF (1 x 5 min) and then washed with DMF (2 x 30 sec) and the washings collected. This process was repeated once. The resin was washed with DCM (3 x 30 sec) and DMF (3 x 30 sec). Coupling: The resin was treated with a solution of free acid/HATU/2,4,6-collidine (5eq./5eq./10eq.) in DMF (total free acid concentration 0.07M) which had been pre mixed for 1 min. The coupling was allowed to proceed for 6-12 h. The resin was washed with DMF (3 x 30 sec), DCM (3 x 30 sec), DMF (3 x 30 sec) and DCM (3 x 30 sec). Capping: The resin was treated with Ac_2O/pyridine/DCM (1:25:25) for 5 min and washed with DCM (4 x 30 sec), DMF (2 x

30 sec) and DCM (2 x 30 sec). A small portion of the resin was removed and the coupling sequence quantitated as described in reference 9.

4) Cleavage of Oligomers from the Resin: The resin was dried *in vacuo* for at least 12h. The dry resin was then washed with TFA (2 x 1 min) and then subjected to a precooled (0°C) solution of TFMSA/thioanisole/TFA (1/1/8) and mixed for 1 h while coming to RT. The resin was then removed by filtration and the filtrate diluted with anhydrous Et_2O (20x). The resulting mixture was then centrifuged to give a white pellet. The ethereal supernatant was decanted off, fresh Et_2O was added and the pellet was broken up and resuspended. After centrifugation the supernatant was decanted off and the pellet dried *in vacuo*. The oligomers were then purified by reversed-phase HPLC and analyzed by mass spectrometry.

HPLC conditions A: C_{18} Vydac reversed phase column (4.6 x 125 mm, 5μm), flow rate 1.5mL/min, linear gradient from 5% aqueous CH_3CN to 100% CH_3CN (all solvents contained 0.06% TFA) in 35 min, UV monitored at λ=260 nm, RT.

HPLC conditions B: HP Zorbax RX-C_{18} reversed phase column (9.4 x 250 mm, 5μm), flow rate 4.2mL/min, linear gradient from 100% H_2O to 30% aqueous CH_3CN (all solvents contained 0.06% TFA) in 55 min, UV monitored at λ=260 nm, at 55°C.

Hexamer **129**: HPLC conditions A: retention time = 8.1 min (peak area 100%).

ES^+ MS: *m/z* 1807 $(M+H)^+$, 1830 $(M+Na)^+$; ES^- MS: *m/z* 1805 $(M-H)^-$, 1828 $(M+Na-H)^-$.

Decamer **130**: HPLC conditions B: retention time = 19.8 min (peak area 93%).

ESI^+ MS: *m/z* 1412.1 $(M+2H)^{2+}$, 941.9 $(M+3H)^{3+}$, 706.9 $(M+4H)^{4+}$.

Decamer **131**: HPLC conditions A: retention time = 16.9min (peak area 100%).

MALDI ToF MS: *m/z* 2862.

Hexamer **144**: HPLC conditions A: retention time = 9.6 min (peak area 100%).

ES^+ MS: *m/z* 1846.6 $(M+H)^+$; ES^- MS: *m/z* 1844.6 $(M-H)^-$.

Hexamer **146**: HPLC conditions B: retention time = 24.8 min (peak area 100%).

ES^+ MS: *m/z* 1861 $(M+H)^+$; ES^- MS: *m/z* 1859 $(M-H)^-$.

Hexamer **150**: HPLC conditions B: retention time: 24.3 min (peak area 100%).

ES^+ MS: *m/z* 1856 $(M+Na)^+$; ES^- MS: *m/z* 1854 $(M+Na-H)^-$.

Hexamer **145**: HPLC conditions B: retention time: 22.1 min (peak area 99%).

FAB^+ MS: *m/z* 1848 $(M+H)^+$.

Hexamer **147**: HPLC conditions B: retention time: 24.5 min (peak area 92%).

FAB^+ MS: *m/z* 1848 $(M+H)^+$.

Hexamer **151**: HPLC conditions B: retention time: 49.7 min (peak area 99%).

ES^+ MS: *m/z* 1106.7 $(M+2H)^{2+}$.

Decamer **165**: HPLC conditions A: retention time = 32.1 min (peak area 100%).

ESI^+ MS: *m/z* 1443.1 $(M+2H)^{2+}$, 962.6 $(M+3H)^{3+}$, 722.3 $(M+4H)^{4+}$.

Decamer **166**: HPLC conditions A: retention time = 36.9 min (peak area 97%).

ESI^+ MS: *m/z* 1474.1 $(M+2H)^{2+}$, 983.3 $(M+3H)^{3+}$, 737.9 $(M+4H)^{4+}$.

Decamer **167**: HPLC conditions A: retention time = 44.3 min (peak area 94%).

ESI^+ MS: *m/z* 1515.5 $(M+2H)^{2+}$, 1010.9 $(M+3H)^{3+}$, 758.5 $(M+4H)^{4+}$.

Decamer **168**: HPLC conditions B: retention time = 19.6 min (peak area 99%).

MALDI ToF MS: *m/z* 3402.

Decamer **169**: HPLC conditions A: retention time: 28.7 min (peak area 98%).

MALDI ToF MS: *m/z* 2921.

Decamer **170**: HPLC conditions A: retention time = 34.0 min (peak area 89%).

ESI$^+$ MS: *m/z* 1490.1 $(M+2H)^{2+}$, 994.1 $(M+3H)^{3+}$, 745.9 $(M+4H)^{4+}$.

Decamer **171**: HPLC conditions B: retention time = 23.9 min (peak area 92%).

ESI$^+$ MS: *m/z* 1552.1 $(M+2H)^{2+}$, 1035.2 $(M+3H)^{3+}$, 776.8 $(M+4H)^{4+}$.

Hexamer **172**: HPLC conditions A: retention time: 52.1 min (peak area 98%).

MALDI ToF MS: *m/z* 2184.2.

Hexamer **173**: HPLC conditions A: retention time: 46.6 min (peak area 99%).

ESI$^+$ MS: *m/z* 1116.1 $(M+2H)^{2+}$.

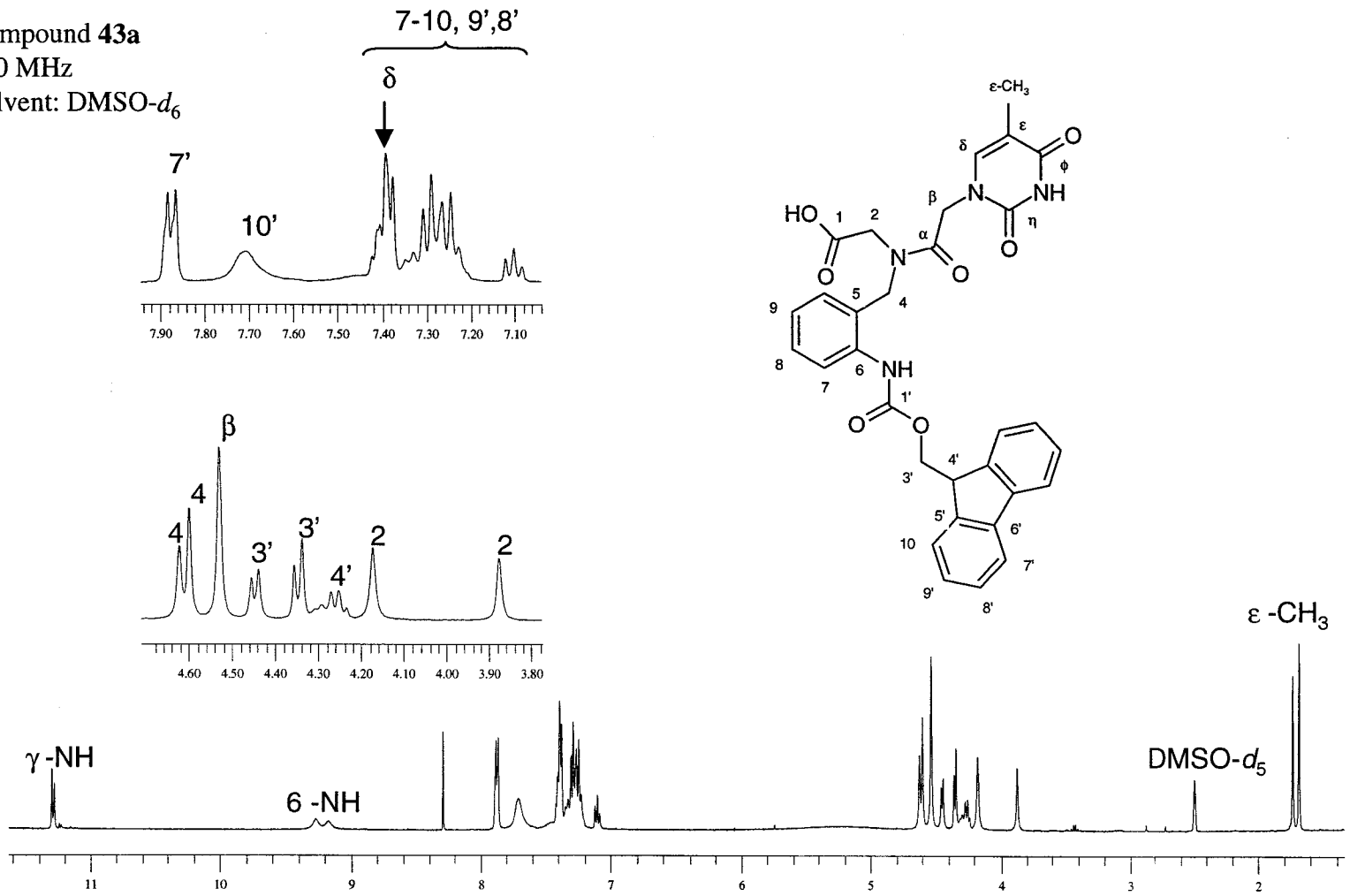
Hexamer **176**: HPLC conditions B: retention time: 12.8 min (peak area 100%).

ESI$^+$ MS: *m/z* 1752.7 $(M+H)^{+}$, 888.2 $(M+2H)^{2+}$.

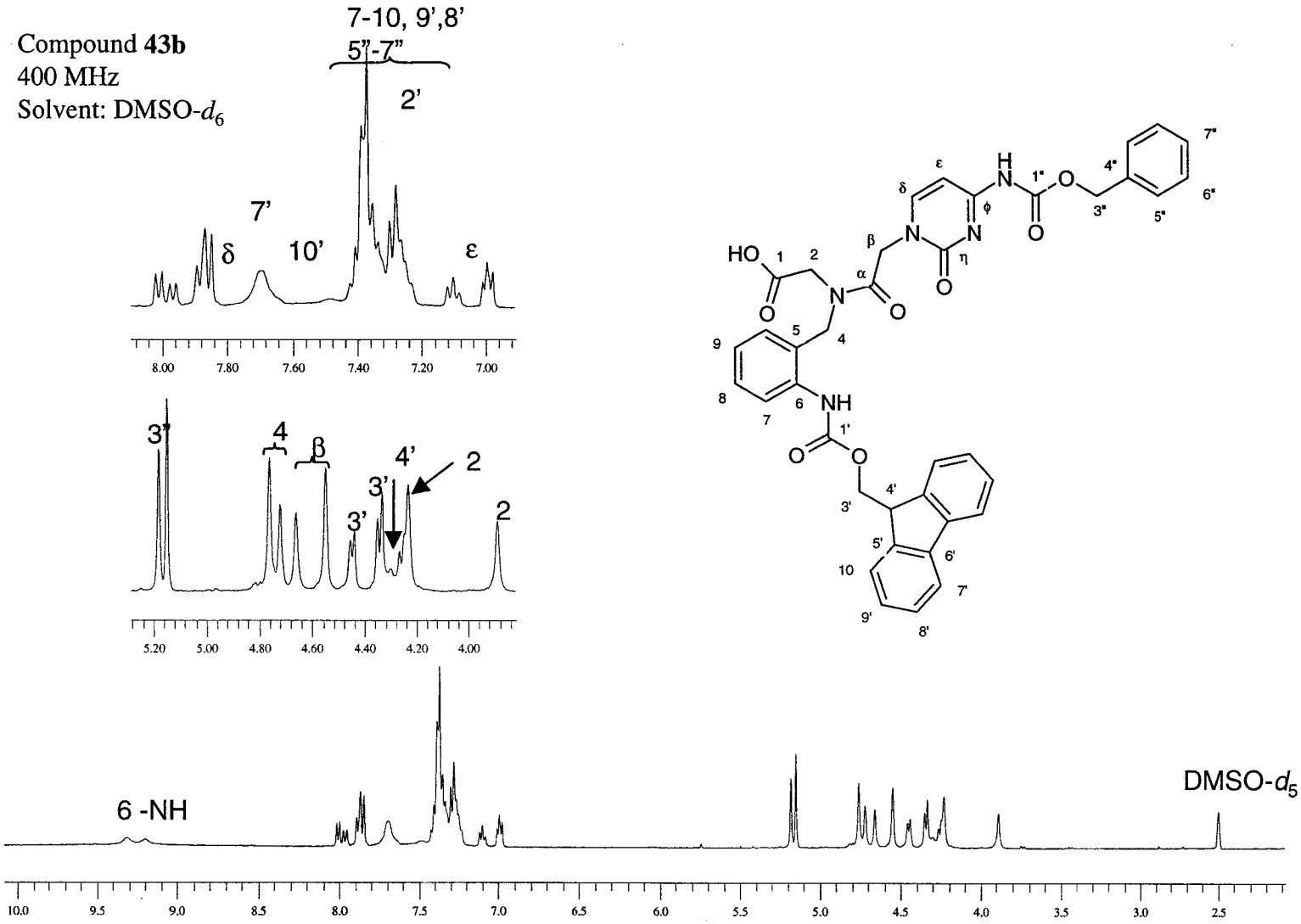
APPENDIX A

^{1}H NMR SPECTRA OF APNA MONOMERS 43a-d.

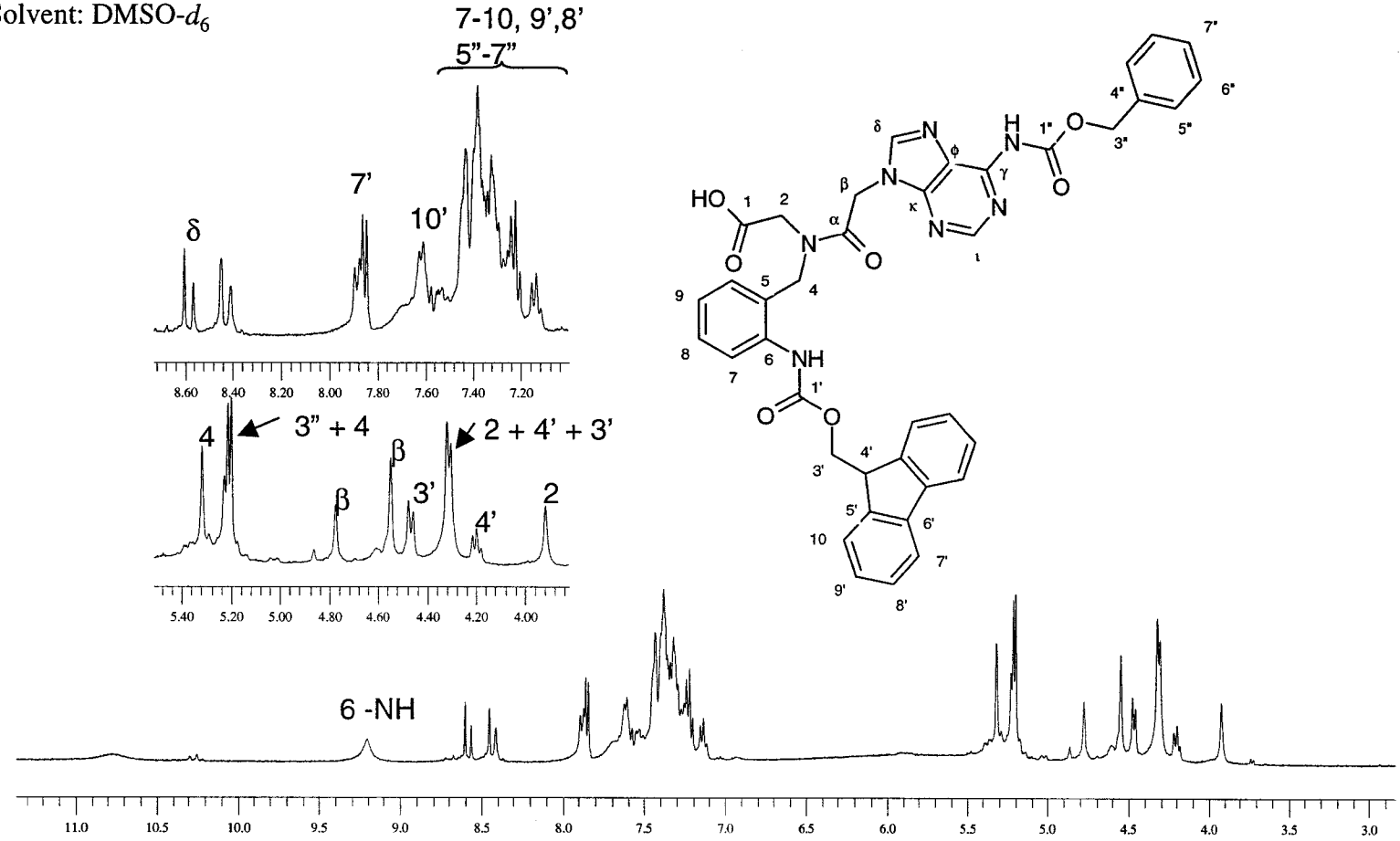
Compound **43a**
400 MHz
Solvent: DMSO-d_6



Compound **43b**
400 MHz
Solvent: DMSO-d_6



Compound **43c**
400 MHz
Solvent: DMSO-d_6



Compound **43d**
400 MHz
Solvent: DMSO-d_6